Solutions to Quantum Computation and Quantum Information

Ryohei (Rio) Weil & Arnab Adhikary

This document was typeset on June 26, 2026

Introduction:

This document is a collection of comprehensive solutions to the exercises and problems in Nielsen and Chuang's "Quantum Computation and Quantum Information". To our knowledge, no official solution manual to the text exists, and while many valiant attempts can be found, to date there does not exist a complete set (understandably so, as we count 529 exercises and 63 end-of-chapter problems). This manual thus exists to fill this gap, and exist as a companion reference for students using the text. Though quantum computation/information as a research area has evolved significantly since its publication, "Mike and Ike" still remains a standard for those entering the field, and we thus hope this manual will be of pedagogical use. Additionally, results of exercises are still frequently quoted in the quantum information literature, further motivating a collection of solutions.

Each solution has the involved concepts (and hence rough pre-requisite knowledge) necessary for the problem in addition to the solution. We have also noted where we have felt that exercises may be incorrect as stated, and in most cases presented a solution for a corrected version. For some exercises, we have added additional remarks about implications of the result and/or alternative perspectives. We do not present solutions to the problems labelled as "Research", but instead overview the current status of the problem, providing references to the relevant literature if a solution has been found or progress has been made.

Additionally, while Nielsen and Chuang remark in the preface that "with few exceptions the exercises can be worked out in a few minutes", having completed all of them it appears to us as this is either a severe overestimate of the average reader, or a severe underestimate of the difficulty of the exercises - many of the exercises took us much longer than a few minutes! We have thus come up with a finer-resolution grading system to give students a better gauge of the difficulty level of problems. In particular, we have starred exercises which we have considered difficult/non-trivial, requiring some insight/thought/time (in comparison, no stars indicates that the exercise indeed should probably take ≤ 15 minutes, and generally be straightforwards or follow directly from the material of the text). Two stars indicates considerable difficulty and three stars indicates formidable difficulty (with the latter being reserved for research problems and problems that we felt were of the highest difficulty/effort amongst those in the text).

Currently, Chapters 2/4/7/8/9/10/11 and all appendices are fully complete. Most exercises/problems from the remaining chapters have been completed. To be completed are (E=Exercises, P=Problems) P1.1, P1.2, P3.1, P3.7, P3.9, P3.10, E5.17, E5.26, E5.27, E5.28, E5.29, P5.1, P5.5, E6.14, P6.1, P6.3, P6.4, E12.6, E12.7, E12.12, E12.15, P12.3.

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

Contents

1 Introduction and overview

Exercise 1.1: Probabilistic Classical Algorithm

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

Solution

Concepts Involved: Deutsch's Problem, Probability

Recall that a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is said to be *balanced* if $f(x) = 1$ for exactly half of all possible 2^n values of x.

A single evaluation tells us no information about whether f is constant or balanced, so our success rate/error rate after a single evaluation is $\epsilon = \frac{1}{2}$ (random guessing!). Therefore, consider the case where we do two evaluations. If we obtain two different results, then we immediately conclude that f is balanced. Suppose instead that we obtain two results that are the same. If f is balanced, then the probably that the first evaluation returns the given result is $\frac{1}{2}$, and the probability that the second evaluation returns the same result is $\frac{2^n/2-1}{2^n-1}$ (as there are $2^n/2$ of each result of 0 and 1, 2^n total results, $2^n/2 - 1$ of the given result left after the first evaluation, and $2^n - 1$ total uninvestigated cases after the first evaluation). Therefore, if f is balanced, this occurs with probability $\frac{1}{2} \cdot \frac{2^n/2-1}{2^n-1}$, which we can see is less than $\frac{1}{2}$ as:

$$2^n < 2^{n+1} \implies 2^n - 2 < 2^{n+1} - 2 \implies \frac{2^n/2 - 1}{2^n - 1} < 1 \implies \frac{1}{2}\frac{2^n/2 - 1}{2^n - 1} < \frac{1}{2}$$

Hence, if we get the same result in two evaluations, we can conclude that f is constant with error $\epsilon < \frac{1}{2}$. We conclude that only 2 evaluations are required for this algorithm. □

Exercise 1.2

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

Solution

Concepts Involved: Quantum Distinguishability, Quantum Measurement

---

Given access to a device which can distinguish non-orthogonal quantum states $|\psi\rangle, |\varphi\rangle$ (without measurement), we can then either run a circuit U_ψ with action $U_\psi |0\rangle \to |\psi\rangle$ or U_φ with action $U_\varphi |0\rangle \to |\varphi\rangle$ (for $|0\rangle$ some reference/blank state as the cloning substrate), which allows us to clone $|\psi\rangle, |\varphi\rangle$ as we desired. Conversely, given a cloning device, we could clone $|\psi\rangle$ and $|\varphi\rangle$ an arbitrary number of times. Then, we can measure half of the cloned states in measurement basis M_1, and the other half in basis M_2, the first basis containing $|\psi\rangle$ and the second containing $|\varphi\rangle$. Given enough measurements, we would either find that one of the M_1 basis measurements yielded a non-$|\psi\rangle$ outcome (signifying the state was $|\varphi\rangle$), or one of the M_2 basis measurements yielded a $|\varphi\rangle$ outcome (signifying the state was $|\psi\rangle$). Thus, we would be

able to distinguish $|\psi\rangle, |\varphi\rangle$. □

Problem 1.1: (Feynman-Gates conversation)

($\star\star\star$) Construct a friendly imaginary discussion of about 2000 words between Bill Gates and Richard Feynman, set in the present, on the future of computation (Comment: You might like to try waiting until you've heard the rest of the book before attempting this question. See 'History and further reading' below for pointers to one possible answer for this question).

Problem 1.2

($\star\star\star$) What is the most significant discovery yet made in quantum computation and quantum information? Write an essay of about 2000 words to an educated lay audience about the discovery (Comment: As for the previous problem, you might like to try waiting until you've read the rest of the book before attempting this question.)

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. □

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 represen-

tation:

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

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:

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

□

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. □

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

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

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

From which we obtain that A has the matrix representation:

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

□

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:

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

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:

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

(2) We have:

$$\begin{aligned}((y_1,\ldots,y_n),(z_1,\ldots,z_n)) &= \sum_i y_i^* z_i \\ &= \sum_i (y_i z_i^*)^* \\ &= \left(\sum_i z_i^* y_i\right)^* \\ &= \big(((z_1,\ldots,z_n),(y_1,\ldots,y_n))\big)^*\end{aligned}$$

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

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

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

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

which proves the claim.

□

Exercise 2.6

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

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

Solution

Concepts Involved: Inner Products

Applying properties (2) (conjugate symmetry), (1) (linearity in second argument), and (2) (again) in

succession, we have:

$$\begin{aligned}\left(\sum_i \lambda_i \left|w_i\right\rangle, \left|v\right\rangle\right) &= \left(\left|v\right\rangle, \sum_i \lambda_i \left|w_i\right\rangle\right)^* \\ &= \left(\sum_i \lambda_i (\left|v\right\rangle, \left|w_i\right\rangle)\right)^* \\ &= \sum_i \lambda_i^* (\left|v\right\rangle, \left|w_i\right\rangle)^* \\ &= \sum_i \lambda_i^* (\left|w_i\right\rangle, \left|v\right\rangle)\end{aligned}$$

□

Exercise 2.7

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

Solution

Concepts Involved: Inner Products, Orthogonality, Normalization
Recall that two vectors $\left|v\right\rangle, \left|w\right\rangle$ are orthogonal if $\langle v|w\rangle = 0$, and the norm of $\left|v\right\rangle$ is given by $\left\|\left|v\right\rangle\right\| = \sqrt{\langle v|v\rangle}$.

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:

$$\begin{aligned}\left\|\left|w\right\rangle\right\| &= \sqrt{\langle w|w\rangle} = \sqrt{1^2 + 1^2} = \sqrt{2}, \\ \left\|\left|c\right\rangle\right\| &= \sqrt{\langle v|v\rangle} = \sqrt{1^2 + (-1)^2} = \sqrt{2}\end{aligned}$$

So the normalized forms of the vectors are:

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

□

Exercise 2.8

($\star$) 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 $|w_1\rangle, \ldots, |w_d\rangle$ as a basis set for a vector space V, the Gram-Schmidt procedure constructs a basis set $|v_1\rangle, \ldots, |v_d\rangle$ by defining $|v_1\rangle \equiv |w_1\rangle / \big\||w_1\rangle\big\|$ and then defining $|v_{k+1}\rangle$ inductively for $1 \leq k \leq d-1$ as:

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

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 $|v_1\rangle, \ldots, |v_d\rangle$ 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:

$$|v_2\rangle = \frac{|w_2\rangle - \langle v_1|w_2\rangle\, |v_1\rangle}{\big\||w_2\rangle + \langle v_1|w_2\rangle\, |v_1\rangle\big\|}$$

Therefore:

$$\langle v_1|v_2\rangle = \frac{\langle v_1|w_2\rangle - \langle v_1|w_2\rangle\, \langle v_1|v_1\rangle}{\big\||w_2\rangle + \langle v_1|w_2\rangle\, |v_1\rangle\big\|} = \frac{\langle v_1|w_2\rangle - \langle v_1|w_2\rangle}{\big\||w_2\rangle + \langle v_1|w_2\rangle\, |v_1\rangle\big\|} = 0$$

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 $|v_1\rangle, \ldots, |v_k\rangle$ are orthogonal and linearly independent. We then have:

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

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

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

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

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

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 $\big\||v_1\rangle\big\| = \big\||v_2\rangle\big\|$, 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. □

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

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

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

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 $|1\rangle, \ldots, |k\rangle$ 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

$\boxed{\Longrightarrow}$ 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 $U^\dagger U = UU^\dagger = I$. 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:

$$\left(|v\rangle, A^\dagger A|v\rangle\right) = \left((A^\dagger)^\dagger|v\rangle, A|v\rangle\right) = \left(A|v\rangle, A|v\rangle\right).$$

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:

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

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

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

□

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:

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

□

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:

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

□

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:

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

□

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 $|v\rangle \otimes |w\rangle$:

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

□

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. □

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

($\star$) 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:

$$\operatorname{tr}(X)=\operatorname{tr}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}=0+0=0$$

$$\operatorname{tr}(Y)=\operatorname{tr}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}=0+0=0$$

$$\operatorname{tr}(Z)=\operatorname{tr}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}=1-1=0$$

□

Exercise 2.37: Cyclic property of the trace

If A and B are two linear operators show that

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

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:

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

□

Exercise 2.38: Linearity of the trace

If A and B are two linear operators, show that

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

and if z is an arbitrary complex number show that

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

Solution

Concepts Involved: Trace

From the definition of trace, we have:

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

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

□

Exercise 2.39: The Hilbert-Schmidt inner product on operators

($\star$) 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:

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

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 $\big(|v_i\rangle\langle v_j|\big)^\dagger = |v_j\rangle\langle v_i| \neq |v_i\rangle\langle v_j|$. 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:

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

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 $\{\sigma_i, \sigma_j\} = 2\delta_{ij} I$

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:

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

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

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:

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

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.

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:

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

□

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:

$$H = \sum_E E \left|E\right\rangle\!\left\langle E\right|$$

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

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

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

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

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:

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

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

($\star\star$) 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 $|\psi\rangle \in \mathcal{H}$:

$$\langle\psi|E_{i+1}|\psi\rangle = \langle\psi|I|\psi\rangle - \sum_{i=1}^m \langle\psi|E_i|\psi\rangle = 1 - \sum_{i=1}^m \langle\psi|E_i|\psi\rangle = 1 - \sum_{i=1}^m \frac{\left|\langle\psi|p_i\rangle\right|^2}{m} \geq 1 - \sum_{i=1}^m \frac{1}{m} = 0$$

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:

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

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

Remark: Our construction is not optimal (in the sense of minimizing $p_{m+1} = \langle\psi| E_{i+1} |\psi\rangle$). We thank Thibaud Ardant for discussions about optimality of POVMs, and in particular for pointing out that a more optimal POVM can be chosen by taking the denominators of E_i to be α_i instead of m (i.e. $E_i = \frac{|p_i\rangle\langle p_i|}{\alpha_i}$), with $\alpha_i = \max_{|\varphi\rangle} \sum_{i=1}^{m} \left|\langle p_i|\varphi\rangle\right|^2$. More generally, one wants to minimize α_i subject to the positivity constraint $\sum_{i=1}^{m} \frac{1}{\alpha_i}\left|\langle p_i|\varphi\rangle\right|^2 \leq 1$, which is a problem which cannot be generally determined without additional data.

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 $X_1 Z_2$ 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 $X_1 Z_2$, we get:

$$\begin{aligned}
\langle X_1 Z_2 \rangle &= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}} \right) X_1 Z_2 \left(\frac{|00\rangle + |11\rangle}{\sqrt{2}} \right) \\
&= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}} \right) \left(\frac{X_1 Z_2 |00\rangle + X_1 Z_2 |11\rangle}{\sqrt{2}} \right) \\
&= \left(\frac{\langle 00| + \langle 11|}{\sqrt{2}} \right) \left(\frac{|10\rangle - |01\rangle}{\sqrt{2}} \right) \\
&= \frac{1}{2} \left(\langle 00|10\rangle - \langle 00|01\rangle + \langle 11|10\rangle - \langle 11|01\rangle \right) \\
&= \frac{1}{2} \left(0 + 0 + 0 + 0 \right) \\
&= 0.
\end{aligned}$$

□

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

Let us decompose $V = W \cup (V \perp W)$, and suppose $\dim(V) = d$ and $\dim(W) = n \leq d$. Let $\left\{ |w_i\rangle \right\}_{i=1}^{n}$ be an orthonormal basis of W and let $\left\{ |w_i\rangle \right\}_{i=n+1}^{d}$ be an orthonormal basis of $V \perp W$ such that $\left\{ |w_i\rangle \right\}_{i=1}^{d}$ forms an orthonormal basis over the whole space. Then, let us define the unitary extension U' to be:

$$U' = U + I_{V \perp W} = \sum_{i=1}^{d} \sum_{j=1}^{n} u_{ij} |i\rangle\langle j| + \sum_{i=n+1}^{d} |i\rangle\langle i|$$

evidently U' is an operator defined on the entire space V. Let us check that it has the desired properties. Suppose $|w\rangle \in W$. Then, $|w\rangle = \sum_{k=1}^{n} w_k |w_k\rangle$, so:

$$U' |w\rangle = U |w\rangle + \underbrace{I_{V \perp W} |w\rangle}_{0} = U |w\rangle$$

Then, let us check that U' is unitary. Take any $|v\rangle \in V$, then we can write:

$$|v\rangle = \sum_{k=1}^{n} v_k |v_k\rangle + \sum_{k=n+1}^{d} v_k |v_k\rangle = |w\rangle + |v \perp w\rangle$$

Then, let us compute:

$$\begin{aligned}
\langle v| (U')^\dagger U' |v\rangle &= (\langle w| + \langle v \perp w|)(U^\dagger + I_{V\perp W})(U + I_{V\perp W})(|w\rangle + |v \perp w\rangle) \\
&= (\langle w| U^\dagger + \langle v \perp w|)(U|w\rangle + \langle v \perp w|) \\
&= \langle w| U^\dagger U |w\rangle + \langle v \perp w| v \perp w\rangle \\
&= \langle w|w\rangle + \langle v \perp w| v \perp w\rangle \\
&= \langle v|v\rangle .
\end{aligned}$$

thus, U' is a unitary on V. □

Remark: We thank Thibaud Ardant for pointing out an error in the first version of this solution.

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:

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

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:

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

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

Exercise 2.70

($\star$) 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:

$$E=a_1I+a_2X+a_3Y+a_4Z$$

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:

$$\begin{aligned}\langle B_{00}|E\otimes I|B_{00}\rangle&=\left(\frac{\langle 00|+\langle 11|}{\sqrt{2}}\right)E\otimes I\left(\frac{|00\rangle+|11\rangle}{\sqrt{2}}\right)\\ &=\left(\frac{\langle 00|+\langle 11|}{\sqrt{2}}\right)(a_1I+a_2X+a_3Y+a_4Z)\otimes I\left(\frac{|00\rangle+|11\rangle}{\sqrt{2}}\right)\\ &=\left(\frac{\langle 00|+\langle 11|}{\sqrt{2}}\right)\left(a_1\frac{|00\rangle+|11\rangle}{\sqrt{2}}+a_2\frac{|10\rangle+|01\rangle}{\sqrt{2}}+a_3\frac{i|10\rangle-i|01\rangle}{\sqrt{2}}+a_4\frac{|00\rangle-|11\rangle}{\sqrt{2}}\right)\\ &=\frac{1}{2}(a_1+a_1+a_4-a_4)\\ &=a_1\end{aligned}$$

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 $\operatorname{tr}(\rho) = 1$, hence:

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

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:

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

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 $\operatorname{tr}(\rho) = 1$ 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:

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

□

Exercise 2.72: Bloch sphere for mixed states

($\star$) 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 $\operatorname{tr}\left(\rho^2\right) \leq 1$, 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 ($\operatorname{tr}\left(\rho^2\right)$ is minimized, with $\operatorname{tr}\left(\rho^2\right) = \frac{1}{2}$).

(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:

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

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

($\star\star$) 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. □

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:

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

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

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

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:

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

□

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

($\star\star$) 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. □

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:

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

Identically:

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

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:

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

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

($\star$) 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:

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

(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

($\star$) 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

($\star\star$) 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

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

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. $N < \mathrm{Sch}(\psi)$.

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 $\mathrm{Sch}(\psi)$, 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

($\star$) 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:

$$\begin{aligned} N^2 &= (\sum_{i=1}^{3} n_i \sigma_i)^2 \\ &= n_1^2\sigma_1^2 + n_2^2\sigma_2^2 + n_3^2\sigma_3^2 + n_1 n_2(\sigma_1\sigma_2 + \sigma_2\sigma_1) + n_1 n3(\sigma_1\sigma_3 + \sigma_3\sigma_1) + n_2 n_3(\sigma_2\sigma_3 + \sigma_3\sigma_2) \end{aligned}$$

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:

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

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 $\langle [Q, R] \otimes [S, T] \rangle \leq 4$ 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 A1.4) that:

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

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. □

3 Introduction to computer science

Exercise 3.1: Non-computable processes in Nature

($\star$) How might we recognize that a process in Nature computes a function not computable by a Turing machine?

Solution

Concepts Involved: Turing Machines, Computability

One criteria is natural phenomena that appear to be truly random; Turing machines as defined in the text are deterministic (though there are probabilistic variations that would solve this issue) and hence would not be able to compute a random function. From a more direct point, if a process in Nature was to be found to compute a known non-computable problem (e.g. solve the Halting problem or the Tiling problem) then we may conclude (trivially) that the process would not be computable. However since the domain of inputs that we could provide top such a natural process would have to be finite, there would be no concrete method in which one could actually test if such a process was truly computing a non-Turing computable function (as a Turing machine that works on a finite subset of inputs for an uncomputable problem could be devised). □

Exercise 3.2: Turing numbers

($\star\star$) Show that single-tape Turing machines can each be given a number from the list $1, 2, 3, \ldots$ in such a way that the number uniquely specifies the corresponding machine. We call this number the *Turing number* of the corresponding Turing machine. (Hint: Every positive integer has a unique prime factorization $p_1^{a_1} p_2^{a_2} \ldots p_k^{a_k}$, where p_i are distinct prime numbers, and $a_1, \ldots a_k$ are non-negative integers.)

Solution

Concepts Involved: Turing Machines, Cardinality

Lemma 1. If $\{A_n\}_{n=1}^{\infty}$ is a sequence of sets that are countably infinite (that is, they can be put in bijection with the natural numbers $\mathbb{N}$) then their union $A = \bigcup_{n=1}^{\infty} A_n$ is also countably infinite.
Proof. Write $A_n = \{x_{n1}, x_{n2}, x_{n3}, \ldots\}$ (which we can do as each of the A_ns are countably infinite). Then, we form an array:

$$\begin{aligned} A_1 &= x_{11} \nearrow \quad x_{12} \nearrow \quad x_{13} \nearrow \quad \cdots \\ A_2 &= x_{21} \quad x_{22} \nearrow \quad x_{23} \nearrow \quad \cdots \\ A_3 &= x_{31} \quad x_{32} \quad x_{33} \nearrow \quad \cdots \\ &\ldots \end{aligned}$$

Then, we can re-number the elements along the diagonal lines (i.e. $x_{11}, x_{21}, x_{12}, x_{31}, x_{22}, x_{13}, \ldots$). This new enumeration corresponds to a countably infinite set. From there, we let $T \subset \mathbb{N}$ be the remaining labels in the enumeration after removing the repeated elements from the sequence. Then, $|T| = |A|$, and hence A is at most countably infinite. A cannot be finite as $A_1 \subset A$ and A_1 is not finite. Hence A is countably infinite. □

Lemma 2. The set of all finite sequences of elements from a countably infinite set B is also countably infinite.

Proof. Denote B_n the set of length n sequences consisting of elements from B. We show that B_n is countably infinite by induction.

Base Case: $|B| = |B_1|$, so B_1 is countably infinite.

Inductive Step: Suppose that B_k is countably infinite for $k \geq 1$. Then, define $B_{k,i}$ which is the set of $k+1$ length sequences, consisting of all sequences in B_k and then terminating with the ith element of B. $\left|B_{k,i}\right| = |B_k|$ so $B_{k,i}$ is countably infinite, and then $B_{k+1} = \bigcup_{i=1}^{\infty} B_{k,i}$ is countably infinite by Lemma 1. This concludes the argument by induction.

Finally, the set of all finite sequences of elements from B is simply $\bigcup_{k=1}^{\infty} B_k$, which is a union of countably infinite sets and therefore countably infinite again by Lemma '. □

Solution. A single-tape Turing machine can be uniquely specified by a finite integer sequence as follows. First, we state the number of states $|Q| = N$, then enumerate the states $q_s = -1, q_h = 0, q_1 = 1, \ldots, q_{N-2} = N-2$. We then state the size of the alphabet $|\Gamma| = L$ and enumerate the alphabet $\triangleright = -2, b = -1, 0 = 0, 1 = 1, \ldots, L-3 = L-3$. We then write the sequence of symbols appearing on the initial tape, starting from square 0, stopping when we encounter the last non-blank symbol. Finally, we list the program lines, each of which is a five-tuple of integers (as each of the states and symbols appearing in the program line we have associated to an integer, and the action s is one of $-1, 0, 1$). Between each part of this sequence we put a "space integer" -3.

As an example, the Turing machine in the text which computes the constant function (and let us say for the sake of example the tape starts with the number 7 on it, so 110 in binary followed by infinite blanks) has the associated finite integer sequence:

$$\underbrace{5}_{|Q|}, \underbrace{-1}_{q_s}, \underbrace{0}_{q_h}, \underbrace{1}_{q_1}, \underbrace{2}_{q_2}, \underbrace{3}_{q_3}, -3, \underbrace{4}_{|\Sigma|}, \underbrace{-2}_{\triangleright}, \underbrace{-1}_{b}, 0, 1, -3, \underbrace{1,1,0}_{7}, -3, \underbrace{-1,-2,1,-2,1}_{\langle q_s,\triangleright,q_1,\triangleright,+1\rangle},$$

$$\underbrace{1,0,1,-1,1}_{\langle q_1,0,q_1,b,+1\rangle}, \underbrace{1,1,1,-1,1}_{\langle q_1,1,q_1,b,+1\rangle}, \underbrace{1,-1,2,-1,-1}_{\langle q_1,b,q_2,b,-1\rangle}, \underbrace{2,-1,2,-1,-1}_{\langle q_2,b,q_2,b,-1\rangle}, \underbrace{2,-2,3,-2,1}_{\langle q_2,\triangleright,q_3,\triangleright,+1\rangle}, \underbrace{3,-1,0,1,0}_{\langle q_3,b,q_h,1,0\rangle}$$

Because the number of states is finite, the size of the alphabet is finite, the number of non-blank squares on the initial tape is finite, and there is a finite number of lines in the program, the above algorithm produces a finite sequence of integers for a given Turing machine.

Now, observe that the integers are countably infinite, as:

$$\begin{array}{rcccl} f & : & \mathbb{N} & \longrightarrow & \mathbb{Z} \\ & & n & \longmapsto & \begin{cases} \frac{n}{2} & n \text{ even} \\ -\frac{n-1}{2} & n \text{ odd} \end{cases} \end{array} \tag{1}$$

is a bijection. So too then is the set of all finite integer sequences by the Lemma 2. Since the finite integer sequences describing Turing machines as produced by the above prescription are a subset of all finite integer sequences, we must therefore be able to associate any single-tape Turing machine to a unique natural number, as claimed. □

Exercise 3.3: Turing machine to reverse a bit string

($\star$) Describe a Turing machine which takes a binary number x as input, and outputs the bits of x in reverse order. (*Hint:* In this exercise and the next it may help to use a multi-tape Turing machine and/or symbols other than $\triangleright, 0, 1$ and the blank.)

Solution

Concepts Involved: Turing Machines.
For a two-tape Turing machine, recall that program lines are of the form $\langle q, x_1, x_2, q', x_1', x_2', s_1, s_2 \rangle$, with q the internal state of the machine, tape 1 (2) reads x_1 (x_2) at the current position, the internal state of the machine changes to q', x_1 (x_2) is overwritten with x_1' (x_2'), and tape 1 (2) is moved according to s_1 (s_2) being $0, +1$, or -1.

As the hint suggests, we consider a multi-tape construction, in particular a two-tape construction (as is discussed in the text, such a machine may be simulated by a single-tape machine). We suppose the input is on the first tape, which starts with $\triangleright$, followed by x written in binary (and blanks thereafter), and the second tape starts on $\triangleright$ and is otherwise completely blank. We will write the output on the second tape. The program is as follows (enumerated for clarity):

$$
\begin{aligned}
1 &: \langle q_s, \triangleright, \triangleright, q_1, \triangleright, \triangleright, +1, 0 \rangle \\
2 &: \langle q_1, 0, \triangleright, q_1, 0, \triangleright, +1, 0 \rangle \\
3 &: \langle q_1, 1, \triangleright, q_1, 1, \triangleright, +1, 0 \rangle \\
4 &: \langle q_1, b, \triangleright, q_2, b, \triangleright, -1, +1 \rangle \\
5 &: \langle q_2, 0, b, q_2, 0, 0, -1, +1 \rangle \\
6 &: \langle q_2, 1, b, q_2, 1, 1, -1, +1 \rangle \\
7 &: \langle q_2, \triangleright, b, q_h, \triangleright, b, 0, 0 \rangle
\end{aligned}
$$

Line 1 starts the machine and sets it into state q_1. Lines 2/3 then proceed to have the machine travel along the first tape without modifying it. Line 4 runs when the end of the input x is reached, putting the machine into state q_2. From there, lines 5/6 run, wherein the first tape is traversed in reverse, and its (reversed) bits are written onto the second tape. Line 7 runs when the start of the first tape is reached, at which point the program halts, with the reversed bitstring written onto the second tape. □

Exercise 3.4: Turing machine to add modulo 2

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

Solution

Concepts Involved: Turing Machines

We take the digits of x, y to be organized from most to least significant going from left to right. We use

a three tape construction (with instruction lines of the form $\langle q, x_1, x_2, x_3, q', x'_1, x'_2, x'_3, s_1, s_2, s_3 \rangle$); the first tape will hold the input (of the form described in the question-$\triangleright$, then the bits of x, then a blank, then the bits of y, then blanks), the second tape is used for convenience and will hold a copy of x (so that we can read the bits of x, y simultaneously), and on the third tape we write the output. Both the second/third tape start with $\triangleright$ and otherwise only contain blanks. The program is as follows (enumerated for clarity):

$$
\begin{aligned}
1 &: \langle q_s, \triangleright, \triangleright, \triangleright, q_1, \triangleright, \triangleright, \triangleright, +1, +1, 0 \rangle \\
2 &: \langle q_1, 0, b, \triangleright, q_1, 0, 0, \triangleright, +1, +1, 0 \rangle \\
3 &: \langle q_1, 1, b, \triangleright, q_1, 1, 1, \triangleright, +1, +1, 0 \rangle \\
4 &: \langle q_1, b, b, \triangleright, q_2, b, b, \triangleright, +1, 0, +1 \rangle \\
5 &: \langle q_2, 0, b, b, q_2, 0, b, b, +1, 0, +1 \rangle \\
6 &: \langle q_2, 1, b, b, q_2, 1, b, b, +1, 0, +1 \rangle \\
7 &: \langle q_2, b, b, b, q_3, b, b, b, -1, -1, 0 \rangle \\
8 &: \langle q_3, 0, 0, b, q_3, 0, 0, 0, -1, -1, -1 \rangle \\
9 &: \langle q_3, 1, 0, b, q_3, 1, 0, 1, -1, -1, -1 \rangle \\
10 &: \langle q_3, 0, 1, b, q_3, 0, 1, 1, -1, -1, -1 \rangle \\
11 &: \langle q_3, 1, 1, b, q_4, 1, 1, 0, -1, -1, -1 \rangle \\
12 &: \langle q_4, 0, 0, b, q_3, 0, 0, 1, -1, -1, -1 \rangle \\
13 &: \langle q_4, 1, 0, b, q_4, 1, 0, 0, -1, -1, -1 \rangle \\
14 &: \langle q_4, 0, 1, b, q_4, 0, 1, 0, -1, -1, -1 \rangle \\
15 &: \langle q_4, 1, 1, b, q_4, 1, 1, 1, -1, -1, -1 \rangle \\
16 &: \langle q_3, b, \triangleright, b, q_h, b, \triangleright, 0, 0, 0, 0 \rangle \\
17 &: \langle q_4, b, \triangleright, b, q_h, b, \triangleright, 1, 0, 0, 0 \rangle
\end{aligned}
$$

Line 1 starts the machine and sets it into state 2. Lines 2/3 proceed to have the machine copy x onto the second tape. Line 4 puts the machine into q_2, wherein Lines 5/6 move the tapehead on the first tape to the end/least significant digit of y (in tandem moving the tapehead of the third tape), until reaching a blank, wherein line 7 transitions the machine into state q_3. Note at this point the third tapehead has moved $n+1$ squares (for n-bit x, y), to account for the fact that the addition may produce a $n+1$ bit number.

Lines 8-15 are performing bitwise addition of x, y (from the second/first tape) and writing the result onto the third tape. Whether or not the carry bit is 0/1 is encoded into the internal state q_3/q_4. Finally, in lines 16-17, after going through all n bits of x, y, the machine writes 0/1 in the $n+1$th position (based on whether the state is q_3/q_4) and halts. □

Exercise 3.5: Halting problem with no inputs

($\star$) Show that given a Turing machine M there is no algorithm to determine whether M halts when the input to the machine is a blank tape.

Solution

Concepts Involved: Turing Machines, Halting Problem

Suppose for the sake of contradiction such an algorithm (called HALTBLANK) existed. Now, suppose we have a Turing machine that computes some function $x(y)$ (with corresponding Turing number x). Then, consider the family of functions $x_y(z)$ which compute $x(y)$ irrespective of the input z (in particular, even if the input is the blank tape/empty) - since there exist Turing machines x/valid programs for computing $x(y)$, there also exist Turing machines x_y/valid programs for computing $x_y(z)$. Then, HALTBLANK could be run on x_y to determine whether it halts when the input is a blank tape. But then we have an algorithm that determines whether or not the Turing machine with number x halts upon input of the number y, i.e. an algorithm that solves the Halting problem - contradiction. □

Exercise 3.6: Probabilistic halting problem

($\star$) Suppose we number the probabilistic Turing machines using a scheme similar to that found in Exercise 3.2 and define the probabilistic halting function $h_p(x)$ to be 1 if machine x halts on input of x with probability at least $1/2$ and 0 if machine x halts on input of x with probability less than $1/2$. Show that there is no probabilistic Turing machine which can output $h_p(x)$ with probability of correctness strictly greater than $1/2$ for all x.

Solution

Concepts Involved: Turing Machines, Halting Problem

Suppose for the take of a contradiction such a probabilistic Turing machine existed, which carried out the algorithm HALTP(x) which outputs $h_p(x)$ with probability $p_c > 1/2$. Then, we can consider an algorithm computing TURINGP(x), defined analogously to TURING(x) of the standard halting problem proof of Box 3.2:

```
TURINGP(x)

y = HALTP(x)
if y=0 then
 | halt
else
 | loop forever
end
```

This is a valid (probabilistic) program with some Turing number t. Suppose $h_p(t) = 1$ - this occurs if TURINGP halts on input t with probability $\geq 1/2$. But by inspecting TURINGP, halting happens if $y = \text{HALTP}(t) = 0$ which is outputted with probability $1 - p_c < 1/2$, contradiction. Conversely, suppose $h_p(t) = 0$ - this occurs if TURINGP halts on input t with probability $< 1/2$. But by inspecting TURINGP, halting happens if $y = \text{HALTP}(t) = 0$ which is outputted with probability $p_c \geq 1/2$, again a contradiction. Thus the assumption of the existence of a Turing machine that can output $h_p(x)$ with probability $p_c > 1/2$ for all x must be false. □

Exercise 3.7: Halting oracle

Suppose a *black box* is made available to us which takes a non-negative integer x as input, and then outputs the value of $h(x)$, where $h(\cdot)$ is the halting function defined in Box 3.2 on page 130. This type of black box is sometimes known as an *oracle* for the halting problem. Suppose we have a regular Turing machine which is augmented by the power to call the oracle. One way of accomplishing this is to use a two-tape Turing machine, and add an extra program instruction to the Turing machine which results in the oracle being called, and the value of $h(x)$ being printed on the second tape, where x is the current contents of the second tape. It is clear that this model for computation is more powerful than the conventional Turing machine model, since it can be used to compute the halting function. Is the halting problem for this model of computation undecidable? That is, can a Turing machine aided by an oracle for the halting problem decide whether a program for the Turing machine with oracle will halt on a particular input?

Solution

Concepts Involved: Turing Machines, Halting Problem

The answer is that the halting problem for this model is indeed undecideable/a Turing machine + halting oracle cannot decide whether a program for a Turing machine + oracle will halt on a particular input. This is seen by promoting all appearing functions/algorithms in Box 3.2 to be with Turing machine + halting oracle. In full gory detail, take the function:

$$h_O(x) = \begin{cases} 0 & \text{if machine } x + \text{halting oracle does not halt upon input of } x \\ 1 & \text{if machine } x + \text{halting oracle does halt upon input of } x \end{cases}$$

and suppose for the sake of contradiction that the halting problem for the Turing machine + halting oracle model of computation is decidable - then there exists some Turing machine + halting oracle algorithm, to evaluate $h_O(x)$, which we can call HALTO(x). Then, consider an algorithm computing TURINGO(x):

```
TURINGO(x)

y = HALTO(x)
if y=0 then
 | halt
else
 | loop forever
end
```

which is a valid program with associated Turing number t. Then $h(t) = 1$ iff TURINGO(t) halts on t, but upon inspection of the program it halts on input t if and only if $h(t) = 0$ - contradiction. Indeed, the addition of the halting oracle does not change the argument whatsoever! □

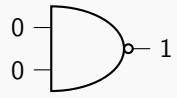
Exercise 3.8: Universality of `NAND`

Show that the `NAND` gate can be used to simulate the `AND`, `XOR`, and `NOT` gates, provides wires, ancilla bits and `FANOUT` are available.

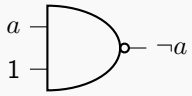
Solution

Concepts Involved: Logic Gates

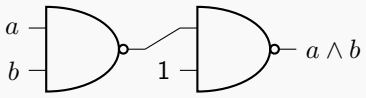
We start by showing how we can get a 1 qubit using two 0 ancilla bits and a NAND gate.



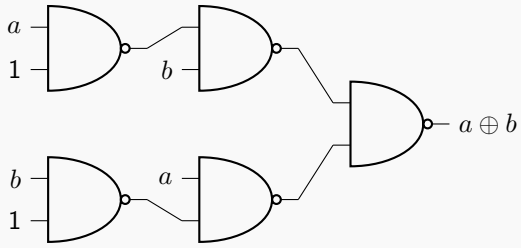
We will now show how to simulate the NOT, AND, and XOR gates. We note that we will use "1" to denote as shorthand a 1 bit constructed using two ancilla bits (as above). a/b represent the input bits. We start with the NOT gate.



Next, we simulate the AND gate.



For the XOR simulated gate, we note that we first use FANOUT twice to copy both input bits.



Having simulated the three gates using the NAND gate only, we conclude that the NAND is universal. □

Exercise 3.9

Prove that $f(n)$ is $O(g(n))$ if and only if $g(n)$ is $\Omega(f(n))$. Deduce that $f(n)$ is $\Theta(g(n))$ if and only if $g(n)$ is $\Theta(f(n))$.

Solution

Concepts Involved: Asymptotic Notation

Suppose $f(n)$ is $O(g(n))$. Then, there exists $c > 0$ such that for all $n > n_0$, $f(n) \leq cg(n)$. Therefore, we have $\frac{1}{c} > 0$ such that for all $n > n_0$, $\frac{1}{c}f(n) \leq g(n)$. Hence, $g(n)$ is $\Omega(f(n))$. Conversely, if $g(n)$ is $\Omega(f(n))$, there exists $c > 0$ such that for all $n > n_0$, $cf(n) \leq g(n)$. Hence, we have $\frac{1}{c} > 0$ such that for all $n > n_0$, $f(n) \leq \frac{1}{c}g(n)$ and hence $f(n)$ is $O(g(n))$.
Therefore, if $f(n)$ is $\Theta(g(n))$ then $f(n)$ is $O(g(n))$ and $\Omega(g(n))$, and by the above argument, $g(n)$ is $O(f(n))$ and $\Omega(f(n))$ and hence $g(n)$ is $\Theta(f(n))$. The converse holds in the same way. □

Exercise 3.10

Suppose $g(n)$ is a polynomial of degree k. Show that $g(n)$ is $O(n^l)$ for any $l \geq k$.

Solution

Concepts Involved: Asymptotic Notation

By assumption, $g(n) = a_0 + a_1 n^1 + a_2 n^2 + \ldots + a_k n^k$ with $a_k \neq 0$. For $n \geq 1$ we have that $n^l \geq n^k$ if $l \geq k$, and hence if $l \geq k$ we have that $a_i n^l \geq a_i n^i$ for all $i \in \{0, \ldots, k\}$. Therefore, we have that:

$$(a_0 + a_1 + \ldots + a_k) n^l \geq a_0 + a_1 n^1 + \ldots + a_k n^k = g(n)$$

for $n \geq 1$ and hence $g(n)$ is $O(n^l)$. □

Exercise 3.11

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

Solution

Concepts Involved: Asymptotic Notation

Let $k > 0$ and $c > 0$. By the definition of the exponential we have that:

$$\exp\left(cn^k\right) = \sum_{j=0}^{\infty} \frac{(an^k)^j}{j!} = \sum_{j=0}^{\infty} \frac{a^{kj} n^{kj}}{j!}$$

now, there exists some $j_0 \in \mathbb{Z}$ for which $kj_0 > 1$. Since for $n \geq 0$ the terms in the above sum are non-negative, we find:

$$\exp\left(cn^k\right) \geq \frac{c^{kj_0} n^{kj_0}}{j_0!}$$

Now, choose c sufficiently large such that $c^{kj_0} \geq j_0!$. We then find that:

$$\exp\left(cn^k\right) \geq \frac{c^{kj_0} n^{kj_0}}{j_0!} \geq n^{kj_0}$$

Then for $n \geq 1$ it follows that $n^{kj_0} \geq n$ as $kj_0 \geq 1$ and so:

$$\exp\left(cn^k\right) \geq n$$

Since the logarithm is monotonic, we may take the log of both sides and preserve the inequality:

$$cn^k \geq \log n$$

So we have shown that for any $k > 0$, there exists $c > 0$ such that for all $n > 1$, $cn^k \geq \log n$. Hence, $\log n$ is $O(n^k)$ for any $k > 0$. □

Exercise 3.12: $n^{\log n}$ is super-polynomial

Show that n^k is $O(n^{logn})$ for any k, but that $n^{\log n}$ is never $O(n^k)$.

Solution

Concepts Involved: Asymptotic Notation

First, note that for any k, $e^k \leq n$ for sufficiently large $n > n_0$ and so $k \leq \log n$ by monotonicity of the logarithm. Therefore, for $n > n_0$ it follows by monotonicity (of exponentiation) that $n^k \leq n^{\log n}$ and so n^k is $O(n^{\log n})$.
Now, consider an arbitrary $a > 0$. It still follows for sufficiently large $n > n_0$ that $e^{ak} \leq n$ and so $ak \leq \log n$ and $n^a n^k \leq n^{\log n}$. But for any $c > 0$ $n^a > c$ for sufficiently large n and so:

$$cn^k \leq n^{\log n}$$

So since for any $c > 0$ there exists some n_0' for which $n > n_0'$ implies $cn^k \leq n^{\log n}$, it follows that $n^{\log n}$ is never $O(n^k)$. □

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

($\star$) Show that c^n is $\Omega(n^{\log n})$ for any $c > 1$, but that $n^{\log n}$ is never $\Omega(c^n)$.

Solution

Concepts Involved: Asymptotic Notation

First note from Exercise 1.11 that $\log n$ is $O(n^k)$ for any $k > 0$. Specifically, take $k = 1/2$; then there exists $a > 0$ such that for $n > n_0$:

$$an^{1/2} \geq \log n$$

and therefore squaring both sides:

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

Now for any $c > 1$, we can define $a' = \frac{a^2}{\log c} > 0$ and write:

$$na' \log c = \log c^{na'} \geq \log n^{\log n}$$

Exponentiating both sides preserves the inequality, and so:

$$c^{na'} = c^{a'} c^n \geq n^{\log n}$$

and so there exists a constant $\frac{1}{c^{a'}} > 0$ such that for $n > n_0$, $c^n \geq \frac{1}{c^{a'}} n^{\log n}$ and therefore c^n is $\Omega(n^{\log n})$.
Now, let $b > 0$ be some arbitrarily small constant. For sufficiently large n, we have that $na' \log c + \log b > 0$

and so for sufficiently large n it further follows that:

$$\log b + na' \log c = \log\left(bc^{na'}\right) \geq \log n^{\log n}$$

where a' is defined as it was previously. Therefore exponentiating both sides:

$$bc^{na'} = bc^{a'}c^n \geq n^{\log n}$$

so for sufficiently large n, for any arbitrarily small constant b' it follows that $b'c^n \geq n^{\log n}$ and so $n^{\log n}$ is never $\Omega(c^n)$. □

Exercise 3.14

Suppose $e(n)$ is $O(f(n))$ and $g(n)$ is $O(h(n))$. Show that $e(n)g(n)$ is $O(f(n)h(n))$.

Solution

Concepts Involved: Asymptotic Notation

By assumption, we have that $e(n) \leq c_1 f(n)$ for some $c_1 > 0$ and for all $n > n_1$ and that $g(n) \leq c_2 h(n)$ for some $c_2 > 0$ and for all $n > n_2$. Let $n_0 = \max n_1, n_2$. We then have that for $n > n_0$ that:

$$e(n)g(n) \leq c_1 f(n) c_2 h(n) = (c_1 c_2)(f(n)h(n))$$

so therefore $e(n)g(n)$ is $O(f(n)h(n))$. □

Exercise 3.15: Lower bound for compare-and-swap based sorts

($\star$) Suppose an n element list is sorted by applying some sequence of compare-and-swap operations to the list. There are $n!$ possible initial orderings of the list. Show that after k of the compare-and-swap operations have been applied, at most 2^k of the possible initial orderings will have been sorted into the correct order. Conclude that $\Omega(n \log n)$ compare and swap operations are required to sort all possible initial orderings into the correct order.

Solution

Concepts Involved: Asymptotic Notation, Compare-and-Swap

We prove the first statement by induction. After 0 steps, we have that $1 = 2^0$ out of the $n!$ possible orderings are already sorted. Let $k \in \mathbb{N}, k \geq 0$ and suppose that after k swaps, at most 2^k of the initial orderings have been sorted into the correct order. We now consider the state of the list after the $k+1$th swap. Each of the 2^k initial orderings from the previous step are correctly sorted already (so the swap does nothing), and there are a further 2^k initial orderings that are one swap away from the 2^k from the previous step, and hence the $k+1$th swap will put 2^k more initial orderings into the correct order. Therefore, after 2^{k+1} compare and swaps, there are at most $2^k + 2^k = 2^{k+1}$ possible initial orderings that are sorted into the correct order. This proves the claim.
Using the above fact, we have that in order to have all $n!$ possible initial orderings correct after k

steps that $2^k \geq n!$. Taking logarithms on both sides, we have that $\log(2^k) \geq \log(n!)$ and hence $k \geq \log(n!)$. Using Stirling's approximation for factorials (`https://en.wikipedia.org/wiki/Stirling%27s_approximation`), we have that:

$$k \geq n \log n - n \log e + O(\log n)$$

from which we conclude that k is $\Omega(n \log n)$ and hence $\Omega(n \log n)$ compare and swap operations are required to sort all possible initial orderings into the correct order. □

Exercise 3.16: Hard-to-compute functions exist

($\star\star$) Show there exist Boolean functions on n inputs which require at least $2^n / \log n$ logic gates to compute.
The correct bound should be $2^n/n$ logic gates. Alternatively, the assumption that the number of wires is fixed to be n throughout the computation can be added, in which case the bound is indeed $2^n / \log n$ (see remark).

Solution

Concepts Involved: Logic Gates, Boolean Functions, Computational Complexity

A Boolean function on n inputs has 2^n possible input bitstrings and 2 possible outputs for each bitstring. Thus there are a total of 2^{2^n} Boolean functions on n inputs.
Let $C(m)$ be the number of possible circuits consisting of m logical operations. Suppose we have a logical gate that acts on f wires and outputs $f+1$ wires; at the ith step (for $i = 1, \ldots m$) there are $\binom{n+i-1}{f}$ choices of wires where this logical gate could act, and so a total number of circuits:

$$C(m) = \prod_{i=1}^{m} \binom{n+i-1}{f} = O((n+m)^{fm})$$

If we have g available types of logic gates, at each step of the above we have g choices, and so the total number of possible circuits becomes:

$$C(m) = O((n+m)^{fm} g^m)$$

We need m large enough such that $C(m) \geq 2^{2^n}$ in order to be able to compute all possible Boolean functions, and so:

$$O((n+m)^{fm} g^m) \geq 2^{2^n}$$

taking logarithms of both sides:

$$O(fm \log(n+m) + m \log(g)) \geq 2^n \log(2)$$

Neglecting constants and considering that $m \gg n$, this requires:

$$m \log(m) = \Omega(2^n)$$

which holds so long as:

$$m = \Omega(\frac{2^n}{n}).$$

Since this is the number of gates required to compute all possible Boolean functions on n inputs, it must be the case that some Boolean functions require this gate count, and the claim follows. □

Remark: Note that if we only consider logic gates that preserve the number of wires from input to output, then there are $\binom{n}{f}^g = O(n^{fg})$ possible choices of gates in the circuit and so $O(n^{fgm})$ possible circuits of m gates. Then, requiring $O(n^{fgm}) \geq 2^{2^n}$ we find that $m = \Omega(\frac{2^n}{\log n})$, so the original problem statement holds with the assumption that the number of wires is fixed to be n throughout the circuit.

Exercise 3.17

($\star$) Prove that a polynomial-time algorithm for finding the factors of a number m exists if and only if the factoring decision problem is in $\mathbf{P}$.

Solution

Concepts Involved: Factoring, Computability, Computational Complexity

$\boxed{\Longrightarrow}$: Supposed a polynomial time algorithm for finding the factors of a number m exists. Then, this solves the factoring decision problem in polynomial time - for a given $l < m$, simply take the minimum of all obtained prime factors (of which there are $O(\log m)$ many) and check whether it is less than (YES) or greater than (NO) l.

$\boxed{\Longleftarrow}$: Suppose the factoring decision problem is in $\mathbf{P}$, so we can compute $\texttt{Factoring}(m, l)$ in poly time. From this, we can construct a polynomial time algorithm for finding all prime factors of a number m as follows. We look for the smallest prime factor via binary search:

(1) Compute $\texttt{Factoring}(m, m/2)$.

(2) If YES, then we know that the prime factor is in $[0, m/2)$, so then compute $\texttt{Factoring}(m, m/4)$. If NO, we know that the prime factor is in $[m/2, m]$, so then compute $\texttt{Factoring}(m, 3m/4)$.

(3) Repeat the above step (halving the size of the set each time) until the smallest factor p_1 of m is obtained, and store it.

(4) Set $m \to m/p_1$.

(5) Repeat the above steps until the residual m is no longer composite, at which point the prime factorization of m has been fully determined.

The binary search requires $O(\log m)$ calls to Factoring, and we solve for $O(\log m)$ factors, so overall the algorithm involves $O(\log^2 m)$ calls to Factoring (which runs in poly time) and thus the entire algorithm runs in polynomial time, as desired. □

Exercise 3.18

($\star$) Prove that if $\mathbf{coNP} \neq \mathbf{NP}$ then $\mathbf{P} \neq \mathbf{NP}$.

Solution

Concepts Involved: Computational Complexity

We prove this via the contrapositive. Suppose $\mathbf{P} = \mathbf{NP}$.
Let $L \in \mathbf{NP}$; since $L \in \mathbf{P}$, there exists a Turing machine T that decides L in poly time. Now, let M be a Turing machine that simulates T on all inputs (and ignores any input after a blank), and accepts if T rejects, and rejects if T accepts. Then, for $x \notin L$, there exists a witness string w - in particular *any* dummy string (say, $w = x$ itself) and then M will accept in poly time (since T rejects x in poly time) when started in $x -$ blank $- w$. For $x \in L$, for all strings w, M rejects in poly time when started in $x -$ blank $- w$ (as T accepts x in poly time). Thus, $L \in \mathbf{coNP}$, and so **NP** $\subseteq$ **coNP**.
The argument for **coNP** $\subseteq$ **NP** is analogous (swap **NP** $\leftrightarrow$ **coNP** and $x \in L \leftrightarrow x \notin L$ in the argument above), and so we conclude **coNP** $=$ **NP**. □

Exercise 3.19

$(\star)$ The `Reachability` problem is to determine whether there is a path between two specified vertices in a graph. Show that `Reachability` can be solved using $O(n)$ operations if the graph has n vertices. Use the solution to `Reachability` to show that it is possible to decide whether a graph is connected in $O(n^2)$ operations.
The corrected exercise is that `Reachability` can be solved using $O(n^2)$ operations if the graph has n vertices, and that this can be used to decide whether a graph is connected in $O(n^3)$ operations.

Solution

Concepts Involved: Graph Theory, Reachability

`Reachability` can be solved using a graph search algorithm, for example breadth-first-search, which we give below:

```
BFSREACH(G, V1, V2)

if V1 = V2 then
  return True
end
tovisit = []
visited = []
current = V1
for v adj to current in G do
  tovisit.append(v)
end
while tovisit ≠ ∅ do
  visited.append(curr)
  current = tovisit[0]
  tovisit.remove(tovisit[0])
  if current = V2 then
    return True
  end
  for v adj to current in G do
    if v ∉ visited then
      tovisit.append(v)
    end
  end
end
return False
```

The above algorithm explores all vertices reachable from V_1 (starting by looking at adjacent vertices to V_1, then vertices adjacent to those vertices, hence "breadth first"), halting if V_2 has been reached. Note that no vertex/edge is explored twice, since visited vertices are marked and ignored in further iterations. In the worst case scenario, the above algorithm explores every vertex and edge of the graph, and thus requires $O(|V| + |E|)$ operations. A dense graph can have $|E| = O(|V|^2)$ edges (e.g. $\frac{|V|\cdot(|V|-1)}{2}$ for a fully connected graph), and so the above algorithm solves `Reachability` in $O(|V|^2) = O(n^2)$ operations. Since reachability is transitive (i.e. if there exists a path between V_1, V_2 and V_2, V_3, then there exists a path between V_1, V_3 by connecting the two paths), to show that a graph is connected (i.e. that any vertex can be reached by any other vertex), it suffices to show that from one vertex V every other vertex V' can be reached. Thus, to verify a graph is connected, we can run BFSREACH(G, V, V′) a total of $n - 1 = O(n)$ times, for a fixed vertex V and ranging over all other vertices V'. If we get TRUE for all runs then we conclude that the graph is connected, and if we get FALSE for any of the runs, we conclude that the graph is not connected. Since we run an $O(n^2)$ algorithm $O(n)$ times, the verification requires $O(n^3)$ operations. □

Exercise 3.20: Euler's theorem

($\star$) Prove Euler's theorem. In particular, if each vertex has an even number of incident edges, give a constructive procedure for finding an Euler cycle.

Solution

Concepts Involved: Graph Theory, Euler Cycles.
An Euler cycle of a graph $G = (V, E)$ is an ordering of the edges E such that every edge in the graph is visited exactly once.

Suppose $G = (V, E)$ is fully connected and each vertex has an incident number of edges. Then, the following algorithm (due to Hierholzer) constructs an Euler cycle (using the RANDOMCYCLE algorithm as a subroutine):

```
RANDOMCYCLE(start, E)

path = []
choose random e = (start, next) ∈ E incident to start
path.append(e)
E.remove(e)
current = next
while current ≠ start do
    choose random e = (current, next) ∈ E incident to start
    path.append(e)
    E.remove(e)
    current = next
end
return path

EULERCYCLE(V, E)

pick random vertex v from V
cycle = RANDOMCYCLE(v, E)
for e in cycle do
    if e[0] has unremoved edges in E then
        path = RANDOMCYCLE(e[0], E)
        cycle.append(path, location=before edge e)
        return to start of for loop
    end
end
return cycle
```

Let us discuss why this algorithm indeed succeeds, by walking through a given run of it:

(1) Pick at random $v \in V$ to be the starting vertex.

(2) Pick a random path of incident edges, removing each edge from E as it is added to the path. Continue this procedure until you return to v. This is guaranteed as the path cannot terminate anywhere else (since every vertex in the graph has an even number of incident edges, if a vertex is entered, there will always exist an exit edge).

(3) At the conclusion of the previous step, we are guaranteed to have a cycle, but not a Euler cycle (there may remain unexplored edges). Thus, let us iterate through the vertices appearing in the already-constructed cycle; if we encounter a vertex v' that has unexplored edges (i.e. edges unremoved from

E), then run step 2 again, using v' as the start/return vertex. Insert the path obtained this way to the cycle obtained from the previous step.

(4) Run step 3 until no vertices on the cycle contain unexplored incident paths. Since the graph is connected, this will mean that no unexplored edges remain on the cycle. At this point, the cycle contains every edge once (each edge is removed as a candidate after it is added to the path) and is thus a Euler cycle. Thus Euler's theorem is proven by construction.

□

Exercise 3.21: Transitive property of reduction

($\star$) Show that if a language L_1 is reducible to the language L_2 and the language L_2 is reducible to L_3 then the language L_1 is reducible to the language L_3.

Solution

Concepts Involved: Reducibility, Turing Machines

By assumption, there exist Turing machines T, T' operating in polynomial time such that if T is given input x it outputs $R(x)$ with $x \in L_2$ if and only if $R(x) \in L_1$ and if T' is given input x it outputs $R'(x)$ with $x \in L_3$ if and only if $R'(x) \in L_2$.
Define T'' to be the Turing machine which given input x, sequentially simulates T' and then T. This can be done as follows. First, ensure that the programs of T', T do not contain any shared internal states; if they do share common states, simply rename all states of T' such that this is no longer the case. Next, set the program of T'' to be simply all program lines of T, T' combined. The last subtlety is to ensure the correct start, halting, and transition behaviour. Next, the starting state $q_s^{T''}$ of T'' to be that of T', $q_s^{T'}$. Set the halting state $q_s^{T''}$ to be that of T, q_h^T. Finally, replace the halting state $q_h^{T'}$ of T' with a transition state q_{trans}, and add the following lines to the program:

$$\langle q_{\text{trans}}, 0, q_{\text{trans}}, 0, -1\rangle$$
$$\langle q_{\text{trans}}, 1, q_{\text{trans}}, 1, -1\rangle$$
$$\langle q_{\text{trans}}, \triangleright, q_s^T, \triangleright, 0\rangle$$

The step of making sure T', T do not share internal states ensures that the two simulations can run independently sequentially. T'' starts by simulating T' with input x, until it reaches what would have been the halting state $q_h^{T'}$ (having written $R'(x)$ onto the tape), now q_{trans}. Then, the above three lines of the program run, which ensure that the Turing machine returns to the starting square, $\triangleright$, and then starts the simulation of T on input $R'(x)$, which leads to output $R(R'(x))$.
That $R(R'(x)) \in L_1 \iff x \in L_3$ follows from $R'(x) \in L_2 \iff x \in L_3$ and $R''(x) \in L_1 \iff x \in L_2$. It remains to show that T'' runs in polynomial time. The T' part of T'' runs in polynomial time, and produces (as an intermediate) $R'(x)$, the size of which is bounded by some polynomial in $|x|$, say $p(|x|)$. Thus the T part of T'' runs in polynomial time in $p(|x|)$, which is just another polynomial in $|x|$ (composition of polynomials yields another polynomial) - thus T'' runs in polynomial time in $|x|$, and we conclude L_1 is reducible to L_3. □

Exercise 3.22

Suppose L is complete for a complexity class, and L' is another language in the complexity class such that L reduces to L'. Show that L' is complete for the complexity class.

Solution

Concepts Involved: Reducibility, Completeness

Let L'' be any language in the complexity class. Since L is complete for the complexity class, L'' reduces to L, and since L reduces to L', by transitivity of reduction (see previous exercise) L'' reduces to L'. Thus L' is complete for the complexity class. □

Exercise 3.23

($\star$) Show that SAT is $\mathbf{NP}$-complete by first showing that SAT is in $\mathbf{NP}$, and then showing that CSAT reduces to SAT. (*Hint:* for the reduction it may help to represent each distinct wire in an instance of CSAT by different variables in a Boolean formula.)

Solution

Concepts Involved: Computational Complexity, Reducibility, Completeness

If φ is a Boolean formula, then $\varphi \in$ SAT if φ is satisfiable. Then, consider the Turing machine M which given input φ-blank-x (for an assignment x), outputs $\varphi(x)$.
If $\varphi \in$ SAT, then there exists a witness string w, namely the satisfying assignment, for which M accepts in polynomial time when given φ-blank-w as input. If $\varphi \notin$ SAT, then for any string w it will be the case that M rejects in polynonimal time when given φ-blank-w as input (as there is no satisfying assignment). Thus, SAT $\in \mathbf{NP}$.
To see that CSAT reduces to SAT, we require a Turing machine operating in polynonimal time that when given input x outputs $R(x)$, and $x \in$ CSAT if and only if $R(x) \in$ SAT. This is realized by a Turing machine which maps each input wire $w_1, \ldots, w_n$ in a Boolean circuit to a Boolean variable $x_1, \ldots, x_n$, and maps each AND(i_1, i_2) gate (with i_1, i_2 the input wires) to $(i_1 \wedge i_2)$, each OR(i_1, i_2) gate to $(i_1 \vee i_2)$, and each NOT(i) gate to $\neg i$. The Boolean formula is then output as the composition of these translated gate operations. This runs in poly time in the circuit size, since each gate in the circuit has constant size input/output and thus translates to Boolean clauses in constant time. Thus CSAT reduces to SAT.
Taking these two facts together, we conclude SAT is $\mathbf{NP}$-complete. □

Exercise 3.24: 2SAT has an efficient solution

($\star\star$) Suppose φ is a Boolean formula in conjunctive normal form, in which each clause contains only two literals.

(1) Construct a (directed) graph $G(\varphi)$ with directed edges in the following way: the vertices of G correspond to variables x_k and their negations $\neg x_j$ in φ. There is a (directed) edge (α, β) in G if and only if the clause $(\neg\alpha \vee \beta)$ or the clause $(\beta \vee \neg\alpha)$ is present in φ. Show that φ is not satisfiable if and only if there exists a variable x such that there are paths from x to $\neg x$ and from $\neg x$ to x in $G(\varphi)$.

(2) Show that given a directed graph G containing n vertices it is possible to determine whether two vertices v_1 and v_2 are connected in polynomial time.

(3) Find an efficient algorithm to solve 2SAT.

Solution

Concepts Involved: Computational Complexity, Boolean Logic, Graph Theory

(1) $\boxed{\Longleftarrow}$: Suppose there exists some variable x such that there are paths from x to $\neg x$ and $\neg x$ to x. This implies that there exist paths of vertices $x, y_1, y_2, \ldots, y_n, \neg x$ and $\neg x, y_1', y_2', \ldots, y_m', \neg x'$ with y_i, y_i' either variables or their negations, and so $(\neg x \vee y_1), (\neg y_1 \vee y_2), \ldots, (\neg y_n \vee \neg x)$ and $(x \vee y_1'), (\neg y_1' \vee y_2), \ldots (\neg y_m' \vee x)$ appear as clauses in φ.

To better understand what this means, consider the following Boolean expression:

$$(x \vee y) \wedge (\neg y \vee z)$$

and suppose we wish it to be true/evaluate to 1. If $y = 0$, then the first clause is false, so it must be that $z = 1$. If $y = 1$, then the second clause is false, so it must be that $x = 1$. In particular, for this to be true it must be that $(x \vee z) = 1$.

Applying this argument repeatedly for the two sets of clauses we found above, we have:

$$(\neg x \vee y_1) \wedge (\neg y_1 \vee y_2) \wedge \ldots \wedge (\neg y_n \vee \neg x) = 1 \implies (\neg x \vee \neg x) = 1 \implies \neg x = 1$$

$$(x \vee y_1') \wedge (\neg y_1' \vee y_2) \wedge \ldots \wedge (\neg y_m' \vee x) = 1 \implies (x \vee x) = 1 \implies x = 1$$

But then since φ contains both sets of these clauses, it must be that for $\varphi = 1$ we have that $x = 1$ and $\neg x = 1$ simultaneously; this is not possible, so φ is not satisfiable.

$\boxed{\Longrightarrow}$: Suppose now that there are no such variables x for which there are paths from x to $\neg x$ and $\neg x$ to x. Then, a satisfying assignment for φ can be constructed as follows.

(a) Choose variable x_1. If there is a path from x_1 to $\neg x_1$, then to have a satisfying assignment this implies (from the previous argument) that $(\neg x \vee \neg x) = \neg x = 1$ so set $x_1 = 0$ and alternatively if there is a path from $\neg x_1$ to x_1, then set $x_1 = 1$ (by assumption, it cannot be the case that both paths exist). If neither exists, set $x_1 = 0, 1$ randomly.

(b) Consider all directed paths starting from x_1/terminating at $\neg x_1$, or starting from $\neg x_1$/terminating at x_1, which in order for φ to be satisfied imply $(\neg x_1 \vee y)/(y \vee \neg x_1) = 1$ or $(x_1 \vee y)/(y \vee x_1) = 1$ for y variables or negations. If x_1 was set to be 0, then consider all such $(x_1 \vee y)/(y \vee x_1) = 1$ and set $y = 1$, and if x_1 was set to be 0, then consider all such $(\neg x_1 \vee y)/(y \vee \neg x_1) = 1$ and set $y = 1$.

This assignment of ys is guaranteed to be possible - WLOG suppose $x_1 = 1$ and suppose for the sake of contradiction there were paths from x_1 to y and x_1 to $\neg y$. But then there would also exist a path from y to $\neg x_1$, and by transitivity, a path from x_1 to $\neg x_1$. If in step (a) we had set $x_1 = 1$ because there was a path from $\neg x_1$ to x_1, we find that there exist both paths x_1 to $\neg x_1$ and $\neg x_1$ to x_1 - contradiction. If in step (a) we had set $x_1 = 1$ because there were no paths of the form x_1 to $\neg x_1$ or $\neg x_1$ to x_1, we again get a contradiction as we have just constructed a path from x_1 to $\neg x_1$. Thus, we have shown any y assigned in the (b) step are guaranteed to have a consistent assignment.

(c) Continue down the list of variables $x_2, x_3, \ldots$ until an unfixed variable x_j is encountered, and repeat steps (a)/(b) using x_j. Repeat until all variables have been assigned.

Since the above steps specify variables in such a way that that all clauses of φ are true, φ is thus satisfiable.

(2) In Ex. 3.19, we constructed an $O(n^2)$ search algorithm (with n the number vertices) to deduce if two vertices were connected in an undirected graph. By modifying the algorithm to only consider adjacent vertices connected by a directed edge, and also by running it twice (swapping $v_1 \leftrightarrow v_2$ as start/endpoints), the algorithm generalizes to the directed case, still running in $O(n^2)$ (polynomial) time.

(3) Constructing $G(\varphi)$ takes $O(|V|+|E|)$ operations, with number of vertices $|V| = 2n$ (n variables and their negations) and number of edges $|E|$ at most $|V|(|V|-1) = 2n(2n-1)$ for a fully connected directed graph, and so requires $O(n^2)$ operations. Then, for any given variable x to check $x, \neg x$ are connected we can run the search algorithm of (2) in $O(n^2)$ time. We do this for each of the n variables, which takes $n \cdot O(n^2) = O(n^3)$ time. If for all variables x, x and $\neg x$ are not connected, then we conclude that φ is satisfiable. If for any of the x we find that x and $\neg x$ are connected, we conclude that φ is satisfiable. Hence, we have solved 2SAT in $O(n^3)$ (polynomial) time.

□

Exercise 3.25: $\mathbf{PSPACE} \subseteq \mathbf{EXP}$

The complexity class $\mathbf{EXP}$ (for *exponential time*) contains all decision problems which may be decided by a Turing machine running in exponential time, that is time $O(2^{n^k})$, where k is any constant. Prove that $\mathbf{PSPACE} \subseteq \mathbf{EXP}$. (*Hint:* If a Turing machine has l internal states, an m letter alphabet, and uses space $p(n)$, argue that the machine can exist in one of at most $lm^{p(n)}$ different states, and that if the Turing machine is to avoid infinite loops then it must halt before revisiting a state.)

Solution

Concepts Involved: Computational Complexity, Turing Machines

Following the hint, if a Turing machine has a m letter alphabet and uses space $p(n)$, then the number of possible states of the memory is $m^{p(n)}$. The total number of states is the product of the number of internal states and memory states, and so is $lm^{p(n)}$. Since a Turing machine is deterministic, if it revisits an internal state, then it will loop (infinitely many times). Thus, for a decision problem to be in $\mathbf{PSPACE}$, the Turing machine must halt/decide before revisiting a state, and in particular this means that it must halt/decide in $\leq lm^{p(n)}$ time (else by the Pigeonhole principle, there will be an internal state that is visited more than once).
For any polynomial $p(n)$, there exists some sufficiently large k for which $lm^{p(n)} \leq 2^{n^k}$ (this can be seen by taking logarithms, wherein $lp(n)\log m \leq n^k$ follows for sufficiently large k - see Ex. 3.10), and so if a decision problem is in $\mathbf{PSPACE}$, it is also in $\mathbf{EXP}$. Thus $\mathbf{PSPACE} \subseteq \mathbf{EXP}$. □

Exercise 3.26: $\mathbf{L} \subseteq \mathbf{P}$

The complexity class $\mathbf{L}$ (for *logarithmic space*) contains all decision problems which may be decided by a Turing machine running in logarithmic space, that is, in space $O(\log(n))$. More precisely, the class $\mathbf{L}$ is defined using a two-tape Turing machine. The first tape contains the problem instance, of size n, and is a read-only tape, in the sense that only program lines which don't change the contents of the first tape are allowed. The second tape is a working tape which initially contains only blanks. The logarithmic space requirement is imposed on the second, working tape only. Show that $\mathbf{L} \subseteq \mathbf{P}$.

Solution

Concepts Involved: Computational Complexity, Turing Machines

The argument is analogous to the last problem. The number of possible states of a Turing machine with l internal states, an m letter alphabet, and $\log(n)$ memory is $lm^{\log(n)}$. For a decision problem to be in $\mathbf{L}$, the Turing machine must halt/decide in $\leq lm^{O(\log(n))}$ time.
We recall that $\mathbf{P} = \bigcup_k \mathbf{TIME}(n^k)$. We can always find a sufficiently large k for which $lm^{O(\log(n))} \leq n^k$ (this is perhaps easiest to see after taking logarithms of both sides, where we see $\log(l) + O(\log(n))\log(m) \leq k\log(n)$ for sufficiently large k), and so if a decision problem is in $\mathbf{L}$, it is also in $\mathbf{P}$; thus $\mathbf{L} \subseteq \mathbf{P}$. □

Exercise 3.27: Approximation algorithm for VERTEX COVER

Let $G = (V, E)$ be an undirected graph. Prove that the following algorithm finds a vertex cover for G that is within a factor of two of being a minimial vertex cover.

$VC = \emptyset$
$E' = E$
while $E' \neq \emptyset$ **do**
  let (α, β) be any edge of E'
  $VC = VC \cup \{\alpha, \beta\}$
  remove from E' every edge incident on α or β
end
return VC

Solution

Concepts Involved: Graph Theory, Vertex Covers.
Recall that a vertex cover of a graph $G = (V, E)$ is a set of vertices V' such that every edge in the graph has one or both vertices contained in V'.

It is clear that the above algorithm outputs a vertex cover of G; for any $e \in E$, e must be removed from E' at some step k, by which it is incident to either α_k or β_k (or both), which are contained in the output vertex set VC.
To see that the VC output from the algorithm is within a factor of two of the minimal vertex cover (which we call MVC), consider that at step k of the loop, the algorithm adds vertices α_k, β_k to VC corresponding to edge $e_k = (\alpha_k, \beta_k)$. Because MVC is a vertex cover, for any $e \in E$, either one or both of its vertices are contained in MVC, and in particular this is true for each e_k. Thus, at least one of α_k, β_k is in MVC - in comparison, the algorithm adds both α_k, β_k to VC, and thus VC is at most a factor of two larger than MVC. □

Exercise 3.28: Arbitrariness of the constant in the definition of BPP

Suppose k is a fixed constant, $1/2 < k \leq 1$. Suppose L is a language such that there exists a Turing machine M with the property that whenever $x \in L$, M accepts x with probability at least k, and whenever $x \notin L$, M rejects x with probability at least k. Show that $L \in \mathbf{BPP}$.

Solution

Concepts Involved: Computational Complexity, Turing Machines, Chernoff Bound

If $k \geq 3/4$, then automatically $L \in \mathbf{BPP}$ as if $x \in L$ then M accepts with probability at least $k \geq 3/4$ and if $x \notin L$ then M rejects with probability at least $k \geq 3/4$.
If $k < 3/4$, then consider a Turing machine M' which simulates M n times (with n a function of k), and then accepts/rejects based on majority vote of the n outcomes. Writing $k = \frac{1}{2} + \epsilon$, by the Chernoff bound (see Box 3.4), if $x \in L$ ($x \notin L$) M' accepts (rejects) with probability at least $1 - e^{-2\epsilon^2 n}$. Setting:

$$1 - e^{-2\epsilon^2 n} \geq \frac{3}{4}$$

we can take logarithms of both sides to obtain:

$$n \geq \log(4)/2\epsilon^2$$

so taking n to be large enough to satisfy the above bound, we find that if $x \in L$ ($x \notin L$), M' accepts (rejects) with probability at least $3/4$, so $L \in \mathbf{BPP}$. □

Exercise 3.29: Fredkin gate is self-inverse

Show that applying two consecutive Fredkin gates gives the same outputs as inputs.

Solution

Concepts Involved: Fredkin Gates. Recall the input/output table of the Fredkin gate:

Inputs			Outputs		
a	b	c	a'	b'	c'
0	0	0	0	0	0
0	0	1	0	0	1
0	1	0	0	1	0
0	1	1	1	0	1
1	0	0	1	0	0
1	0	1	0	1	1
1	1	0	1	1	0
1	1	1	1	1	1

We check for all possible 8 input states that applying the Fredkin gate returns the original input state.

$$\begin{aligned}
F[F[(0,0,0)]] &= F[(0,0,0)] = (0,0,0) \\
F[F[(0,0,1]] &= F[(0,0,1)] = (0,0,1) \\
F[F[(0,1,0]] &= F[(0,1,0)] = (0,1,0) \\
F[F[(0,1,1)]] &= F[(1,0,1)] = (0,1,1) \\
F[F[(1,0,0)]] &= F[(1,0,0)] = (1,0,0) \\
F[F[(1,0,1)]] &= F[(0,1,1)] = (1,0,1) \\
F[F[(1,1,0)]] &= F[(1,1,0)] = (1,1,0) \\
F[F[(1,1,1)]] &= F[(1,1,1)] = (1,1,1)
\end{aligned}$$

We conclude that the Fredkin gate is self-inverse. □

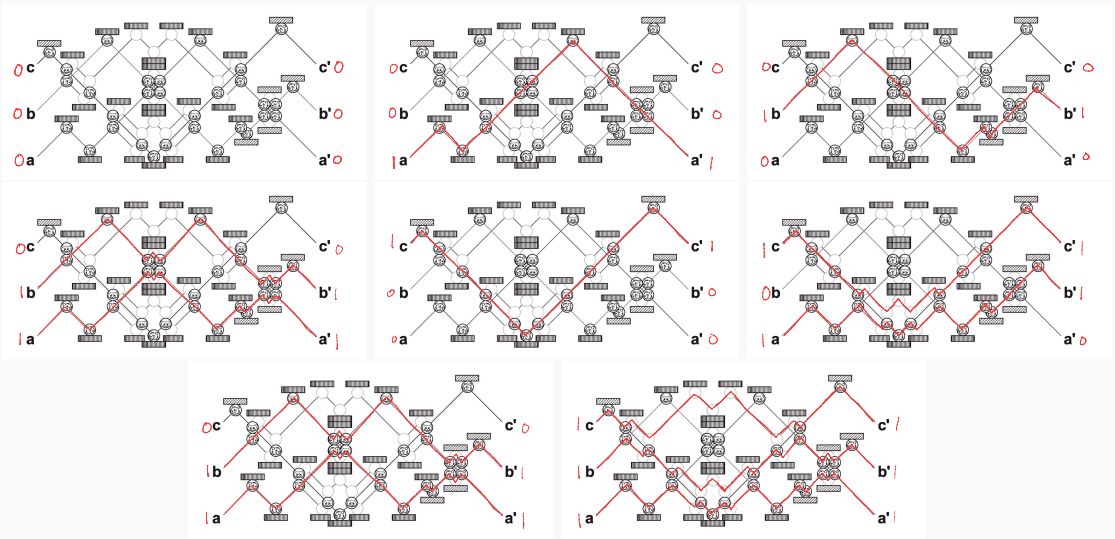
Exercise 3.30

Verify that the billiard ball computer in Figure 3.14 computes the Fredkin gate.

Solution

Concepts Involved: Fredkin Gates

By tracing out the billiard trajectories, we can confirm that the Fredkin gate truth table of Fig 3.15 (also see Ex. 3.29) is reproduced.



□

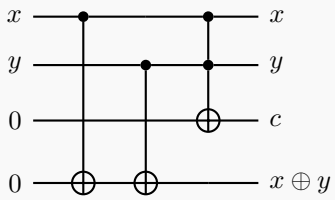
Exercise 3.31: Reversible half-adder

Construct a reversible circuit which, when two bits x and y are input, outputs $(x, y, c, x \oplus y)$, where c is the carry bit when x and y are added.

Solution

Concepts Involved: Logic Gates, Reversible Computation

The following circuit accomplishes this task:



x, y are left invariant since they are only acted on by controls. The two controlled-NOT gates compute $x \oplus y = x + y \pmod 2$ on the fourth bit, and the Toffoli sets the carry (third) bit to 1 if both x, y are 1. The circuit is reversible, since both controlled-NOT and Toffoli gates are reversible. □

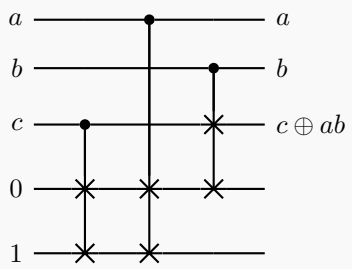
Exercise 3.32: From Fredkin to Toffoli and back again

($\star\star$) What is the smallest number of Fredkin gates needed to simulate a Toffoli gate? What is the smallest number of Toffoli gates needed to simulate a Fredkin gate?

Solution

Concepts Involved: Fredkin Gates, Toffoli Gates

Toffoli from Fredkin. First, note that since Fredkin gates conserve the number of 1s while Toffolis do not, it is impossible to construct a Toffoli from a Fredkin without the use of (non-resettable) ancilla. If we allow for such ancilla, we get a three-Fredkin construction (with a/b the controls and c the target of the simulated Toffoli):



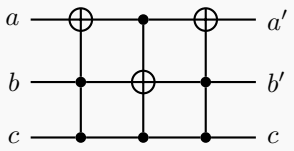
The action of the circuit is:

$$(a,b,c,0,1) \to (a,b,c,c,\bar{c}) \to \begin{cases} (a,b,c,c,\bar{c}) & a=0 \\ (a,b,c,\bar{c},c) & a=1 \end{cases} \to \begin{cases} (a,b,c,c,\bar{c}) & a=0,b=0 \\ (a,b,c,\bar{c},c) & a=1,b=0 \\ (a,b,c,c,\bar{c}) & a=0,b=1 \\ (a,b,\bar{c},c,c) & a=1,b=1 \end{cases}$$

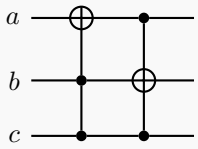
From which we can see that a,b are always left invariant and c is flipped only if $a=1,b=1$ - this is precisely the Toffoli.

So three Fredkins is sufficient. That one Fredkin is not enough can be seen from the fact that the Toffoli needs to control off of two bits (and act trivially on them) while the Fredkin only controls off of one bit. For two Fredkins, since the Toffoli needs to control off of both qubits a/b, this forces the control of the first Fredkin to be on a and the second to be on b (or vise versa). Further, since a/b cannot be acted on nontrivially, this further forces one part of the controlled-SWAP of the Fredkin to be on c, while the other must act on an ancilla. This means at most what two Fredkin's can accomplish is swapping the state of c with pre-determined ancilla bits, depending on the state of a/b. But since the ancilla are pre-determined (while the output bit of a Toffoli depends on the input), it is impossible to make a Toffoli this way from two Fredkins. Thus three Fredkins are necessary/are the minimum, which the above construction saturates.

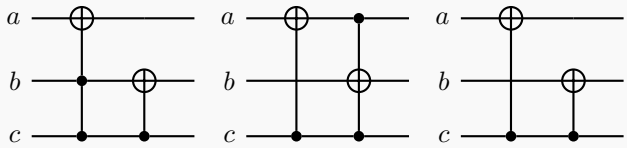
Fredkin from Toffoli. Since the Toffoli gate is a controlled-CNOT gate, and the Fredkin gate is a controlled-SWAP gate, we can use adapt the SWAP-as-three-CNOT construction of Figure 1.7 to make a Fredkin out of three Toffolis:



So three Toffolis is sufficient. That one Toffoli is not enough immediately follows from the fact that the Fredkin gate acts nontrivially on two bits, while the Toffoli only acts nontrivially on a single bit. For two Toffolis, one of the Toffolis must have target on bit a while the other must have target on bit b if both are to be acted on nontrivially. Further, one of the controls for both Toffolis must be on c, as c controls the action on bits a/b for the Fredkin. The only other control could be the (non-target) a/b bit, and so the only possibility is:



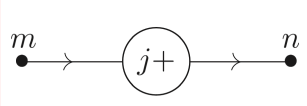
up to a swap of the order of the two gates. But this sends $(1,0,1) \to (1,1,1)$ which is not the correct action of the Fredkin (and in the case where we swapped the order, we would find $(0,1,1) \to (1,1,1)$ which again is not the correct action). The only other possibility is that the non-c control is tied to an external ancilla; if the control is tied to a 0 ancilla, the gate drops out. If the control is tied to a 1 ancilla, the Toffoli becomes a `CNOT` with control on c and target on a/b. If we do this, we get one of (depending on whether we remove the control on a, b or both):



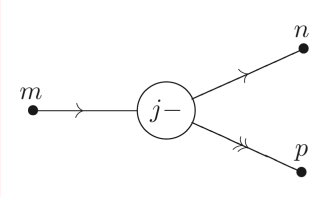
The first sends $(1,0,1) \to (1,1,1)$, the second sends $(1,0,1) \to (0,0,1)$, and the third sends $(0,0,1) \to (1,1,1)$, none of which are the correct action. A similar analysis goes through for the cases with swapped order, and having exhausted the possibilities, we conclude that 2 Toffolis cannot form a Fredkin. Thus 3 Toffolis are necessary/the minimum (which our construction saturates). □

Problem 3.1: Minsky machines

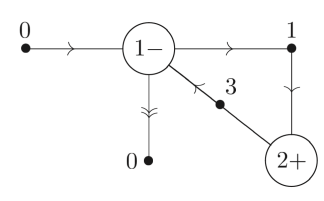
($\star\star\star$) A *Minsky machine* consists of a finite set of *registers*, $r_1, r_2, \ldots, r_k$, each capable of holding an arbitrary non-negative integer, and a *program*, made up of *orders* of one of two types. The first type has the form:



The interpretation is that at point m in the program register r_j is incremented by one, and execution proceeds to point n in the program. The second type of order has the form:



The interpretation is that at point m in the program, register r_j is decremented if it contains a positive integer, and execution proceeds to point n in the program. If register r_j is zero then execution simply proceeds to point p in the program. The *program* for the Minsky machine consists of a collection of such orders, of a form like:



The starting and all possible halting points for the program are conventionally labeled zero. This program takes the contents of register r_1 and adds them to register r_2, while decrementing r_1 to zero.

(1) Prove that all (Turing) computable functions can be computed on a Minsky machine, in the sense that given a computable function $f(\cdot)$ there is a Minsky machine program that when the registers start in the state $(n, 0, \ldots, 0)$ gives as output $(f(n), 0, \ldots, 0)$.

(2) Sketch a proof that any function which can be computed on a Minsky machine, in the sense just defined, can also be computed on a Turing machine.

Problem 3.2: Vector games

($\star\star$) A *vector game* is specified by a finite list of vectors, all of the same dimension, and with integer co-ordinates. The game is to start with a vector $\mathbf{x}$ of non-negative integer co-ordinates and to add to $\mathbf{x}$ the first vector from the list which preserves the non-negativity of all the components, and to repeat this process until it is no longer possible. Prove that for any computable function $f(\cdot)$ there is a vector game which when started with the vector $(n, 0, \ldots, 0)$ reaches $(f(n), 0, \ldots, 0)$ (*Hint:* Show that a vector game in $k + 2$ dimensions can simulate a Minsky machine containing k registers.)

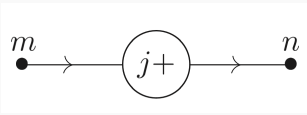
Solution

Concepts Involved: Computability, Minsky Machines

Since we showed in the previous exercise that all Turing computable functions can be computed on a Minsky machine, it suffices to show that any Minsky machine may be simulated by a vector game.
Following the hint, for a Minsky machine with k registers, we consider the following vector game in $k + 2$ dimensions. The state vector is of the form:

$$\mathbf{x} = (r_1, r_2, \ldots, r_k, q_1, q_2)$$

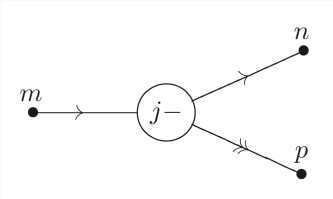
with coordinates $1, \ldots, k$ simulating the state of registers $1, \ldots, k$, and coordinates $k+1, k+2$ simulating the point in the program.
The orders are translated as follows; for each order of the first type (increment):



include the vectors:

$$(0, \ldots, 0, \underbrace{1}_{j}, 0, \ldots, 0, \underbrace{-m}_{k+1}, \underbrace{n}_{k+2})$$
$$(0, \ldots, 0, \underbrace{n}_{k+1}, \underbrace{-n}_{k+2})$$

And for each order of the second type (decrements):



include the vectors:

$$(0, \ldots, 0, \underbrace{-1}_{j}, 0, \ldots, 0, \underbrace{-m}_{k+1}, \underbrace{n}_{k+2})$$
$$(0, \ldots, 0, \underbrace{n}_{k+1}, \underbrace{-n}_{k+2})$$

and:

$$(0, \ldots, 0, \underbrace{-m}_{k+1}, \underbrace{p}_{k+2})$$
$$(0, \ldots, 0, \underbrace{p}_{k+1}, \underbrace{-p}_{k+2})$$

In both of these cases, the first vector of the pair has the action of doing the incrementation/decrementation/passing of register j according to the state (as encoded in the $k+1$th place), and the second vector has the action of transitioning the state by moving the stored value from the $k+2$nd place to the $k+1$th place.
Add these vectors into the list of vectors of the vector game, ordered from most to least negative in the $k+1$th and $k+2$th entries (this is done to ensure that, e.g. if the $k+1$th entry of $\mathbf{x}$ is $-m$, we do not run any instructions corresponding to points $n < m$). Playing this vector game with input $(r_1, \ldots, r_n, 0, 0)$ will then simulate the Minsky machine with input $(r_1, \ldots, r_n)$.

□

Problem 3.3: Fractran

($\star$) A *Fractran* program is defined by a list of positive rational numbers $q_1, \ldots, q_n$. It acts on a positve integer m by replacing it by $q_i m$ where i is the least number such that $q_i m$ is an integer. If there is ever a time when there is no i such that $q_i m$ is an integer, then execution stops. Prove that for any computable function $f(\cdot)$ there is a Fractran program which when started with 2^n reaches $2^{f(n)}$ without going through any intermediate powers of 2. (*Hint:* use the previous problem.)

Solution

Concepts Involved: Computability, Vector Games

From the previous problem, we know that for any computable $f(\cdot)$ there exists a vector game that whens started with $(n, 0, \ldots, 0)$ reaches $(f(n), 0, \ldots, 0)$. Let $\mathbf{v}_1, \ldots \mathbf{v}_n$ be the list of such vectors defining the above vector game, all of length k. We construct a Fractran program that simulates this vector game as follows; to each $\mathbf{v}_i = (v_i^1, v_i^2, v_i^3, \ldots, v_i^k)$, associate:

$$q_i = 2^{v_i^1} 3^{v_i^2} 5^{v_i^3} \ldots p_k^{v_i^k}$$

where p_k is the kth largest prime number. These q_i (which are rational as $v_i^k \in \mathbb{Z}$/the vectors have integer coordinates) make up the Fractran program.

The progression of the vector game that $\mathbf{x} \to \mathbf{x} + \mathbf{v}_i$ where i is the least index such that $\mathbf{x} + \mathbf{v}_i$ has non-negative components maps exactly onto the progression of the Fractran program, where $m \to q_i m$ where i is the least number such that $q_i m$ is an integer (negative components in the vector game $\implies$ prime factors in the denominator of $q_i m$, making it non-integral). Thus, if we start the Fractran program with 2^n, we reach $2^{f(n)}$. No intermediate power of 2 is reached - if this was the case and some intermediate power 2^m was reached, this implies that the simulated vector game reaches $(m, 0, \ldots, 0)$ before $(f(n), 0, \ldots, 0)$, which implies the simulated Minsky machine reaches the $(m, 0, \ldots, 0)$ state - but based on our construction in Ex. 3.1 this Minsky machine would halt and give $(m, 0, \ldots, 0)$ as output, never reaching $(f(n), 0, \ldots, 0)$ - contradiction. □

Problem 3.4: Undecidability of dynamical systems

($\star\star$) A Fractran program is essentially just a very simple dynamical system taking positive integers to positive integers. Prove that there is no algorithm to decide whether such a dynamical system ever reaches 1.

Solution

Concepts Involved: Computability, Fractran, Halting Problem

It will be useful to recall Lemma 2 from our solution to Exercise 3.2, that the set of all finite sequences of elements from a countably infinite set is also countably infinite.

First, note that $\mathbb{Q}$ (the set of rational numbers) is countable (to see this, consider the formal definition of the rationals as:

$$\mathbb{Q} = \{(m, n) : m \in \mathbb{Z}, n \in \mathbb{N}\} / \sim$$

where $\sim$ denotes the equivalence relation $(m_1, n_2) \sim (m_2, n_2)$ if $m_1 n_2 = m_2 n_1$. This is of course just the formalization of the familiar notion of rational numbers as the ratio of two integers, and two rationals being identified via cross multiplication of denominators. Lemma 2 then implies that $\mathbb{Q}$ is countably infinite, as it is the set of length 2 sequences (modulo an equivalence relation) with elements taken from $\mathbb{Z}, \mathbb{N}$, both countably infinite sets.
Then, recalling that a Fractran program consists of a finite list of rationals/non-integer fractions $q_1, \ldots, q_n$, we can again apply Lemma 2 to conclude that the set of Fractran programs is again countably infinite. Thus, each Fractran program can be associated with a natural number n in a way that uniquely specifies the program.
From Problem 3.3, for any computable function $f(\cdot)$, there exists a Fractran program which when started with 2^n reaches $2^{f(n)}$ without going through intermediate powers of 2. Now, suppose for the sake of contradiction that there was an algorithm to decide whether a Fractram program reaches 1. Then, the function:

$$r(n) = \begin{cases} 0 & \text{program } n \text{ started on } 2^n \text{ does not reach 1} \\ 1 & \text{program } n \text{ started on } 2^n \text{ reaches 1.} \end{cases}$$

is computable, and thus there exists a Fractran program F with number f that when started on 2^n reaches $2^{r(n)}$ without going through intermediate values of 2. Now, consider whether F started on 2^f reaches 1.

- If yes, then it reaches $1 = 2^{r(f)}$ without reaching intermediate values of 2 (i.e. as a final state), but $1 = 2^{r(f)} \implies r(f) = 0$, which is only true if f does not reach 1; contradiction.
- If no, then $r(f) = 0$, so it reaches $2^{r(f)} = 2^0 = 1$, again a contradiction.

Thus we get a contradiction, and we conclude that no algorithm to decide whether a Fractran program reaches 1 exists. □

Problem 3.5: Non-universality of two bit reversible logic

($\star\star$) Suppose we are trying to build circuits using only one and two bit reversible logic gates, and ancilla bits. Prove that there are Boolean functions which cannot be computed in this fashion. Deduce that the Toffoli gate cannot be simulated using one and two bit reversible gates, even with the aid of ancilla bits.

Solution

Concepts Involved: Logic Gates, Boolean Functions, Toffoli Gates

Consider the constant Boolean function $f(x) = 1$. This cannot be computed using one/two-bit reversible gates, since circuits composed of reversible circuits can only compute bijective functions. However, it is also not a useful example in the sense that *any* reversible circuit (not just reversible circuits consisting of 1/2-bit reversible logic gates) would not be able to compute it.
A more useful characterization follows from the fact that circuits of one/two-bit reversible gates can only compute *affine* functions, i.e. functions of the form $f(x) = Ax + b$ for a binary and reversible matrix A and binary vector b. Our solution is an adaptation of the very elegant discussion of Cyclone, as can be found on StackExchange.

First we claim that any bijective function $f(x) : \{0,1\}^2 \to \{0,1\}^2$ if affine:

$$f(x) = Ax + b$$

for $A \in \mathsf{GL}_{2\times 2}(\mathbb{Z}_2)$ and b a binary length 2-vector.
To see this, first consider that there are $4! = 24$ such total possible bijective functions (4 choices of outputs for $x = (0,0)^T$, then 3 remaining choices for $(0,1)^T$, 2 for $(1,0)^T$, and 1 for $(1,1)^T$).
Then, let us count the possible affine functions. There are 4 choices for b - $(0,0)^T, (0,1)^T, (1,0)^T, (1,1)^T$ and 6 choices for A:

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

making for $6 \cdot 4 = 24$ apparent choices of affine f; let us verify that each such f is unique. Letting $f(x) = Ax + b, f'(x) = A'x + b'$, if $f = f'$ then for all inputs x:

$$Ax + b = A'x + b' \implies (A - A')x + (b - b') = 0$$

setting $x = (0,0)^T$, we find $b = b'$, and then considering $x = (1,0)^T$ we have:

$$\begin{bmatrix}a_{00} - a'_{00} & a_{01} - a'_{01}\\ a_{10} - a'_{10} & a_{11} - a'_{11}\end{bmatrix}\begin{bmatrix}1 & 0\end{bmatrix} = 0 \implies \begin{bmatrix}a_{00} - a'_{00}\\ a_{10} - a'_{10}\end{bmatrix} = 0 \implies a_{00} = a'_{00}, a_{10} = a'_{10}$$

and an analogous calculation for $x = (0,1)^T$ shows that $a_{01} = a'_{01}, a_{11} = a'_{11}$ and so $A = A'$. Thus each f is unique, and so indeed the 24 affine functions specified by the 6 A and 4 b coincide with the 24 possible bijective 2-bit binary functions.
The above shows that the action of any reversible one/two-bit gate is affine. The action of such a gate when embedded into a circuit on multiple bits (and ancilla) is still affine - simply consider the larger matrix A/larger vector b that has trivial action on the other bits. Further, the composition of many such gates is seen to be affine, since the composition of two affine functions $f_1(x) = A_1x + b_1, f_2(x) = A_2x + b_2$ is again affine:

$$f_2 \circ f_1(x) = f_2(A_1x + b_1) = A_2(A_1x + b_1) + b_2 = A_2A_1x + (A_2b_1 + b_2)$$

with matrix A_2A_1 and vector $A_2b_1 + b_2$.
Thus, we have shown that any function computable by one/two bit reversible logic gates (and ancilla) must be affine - any non-affine Boolean function cannot be computed in this fashion.
Before moving onto the next part of the function, note that the sum of any two affine functions is again affine:

$$f_1(x) + f_2(x) = A_1x + b_1 + A_2x + b_2 = (A_1 + A_2)x + (b_1 + b_2).$$

and analogously for the difference of two affine functions.
Now, let us recall the Toffoli gate; translating the truth table of Figure 3.17 into vectorial form, the Toffoli

gate is a function $T : \{0,1\}^3 \mapsto \{0,1\}^3$ such that:

$$T(x) = T(\begin{bmatrix} a \\ b \\ c \end{bmatrix}) = \begin{bmatrix} a \\ b \\ c \oplus ab \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} x + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} x$$

Suppose for the sake of contradiction that T was affine; clearly the identity function $I(x) = x$ is affine, so by our earlier observation, $T - I$ must be affine, and hence the function:

$$D(x) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} x$$

must be affine. but the above matrix is non-invertible, and hence D cannot be affine - contradiction. We thus conclude that the Toffoli cannot be simulated with one/two bit reversible gates. □

Remark: This result is in stark contrast to the quantum case, where (for example) $\{\texttt{CNOT}, T, H\}$ forms a universal gateset, consisting of only one/two qubit reversible gates!

Problem 3.6: Hardness of approximation of TSP

Let $r \geq 1$ and suppose that there is an approximation algorithm for TSP which is guaranteed to find the shortest tour among n cities to within a factor r. Let $G = (V, E)$ be any graph on n vertices. Define an instance of TSP by identifying cities with vertices in V, and define the distance between cities i and j to be 1 if (i, j) is an edge of G, and to be $\lceil r \rceil |V| + 1$ otherwise. Show that if the approximation algorithm is applied to this instance of TSP then it returns a Hamiltonian cycle for G if one exists, and otherwise returns a tour of length more than $\lceil r \rceil |V|$. From the **NP**-completeness of HC it follows that no such approximation algorithm can exist unless $\mathbf{P} = \mathbf{NP}$.

Solution

Concepts Involved: Graph Theory, Reducibility, Computational Complexity, Hamiltonian Cycles

Suppose G has a Hamiltonian cycle - then, a tour taking the Hamiltonian cycle is of length $n = |V|$, which is optimal/minimal (as all n cities must be visited on the tour, and the minimum distance between cities is 1). Any other tour requires at least taking one path between cities that is not an edge of G, and so has a total distance of at least $(\lceil R \rceil |V| + 1) + |V| - 1 = \lceil r \rceil |V| + |V| = (\lceil r \rceil + 1)|V|$, which is not within the shortest tour to within a factor r. Hence, te approximation algorithm must return the Hamiltonian cycle if it exists.

Conversely, suppose no Hamiltonian cycle for G exists. Then, any tour involves at least one path that is not an edge of G, which has total distance at least $(\lceil r \rceil + 1)|V| > \lceil r \rceil |V|$.

Thus the approximation algorithm finds a Hamiltonian cycle for G if it exists, solving HC in polynomial time. Thus, by the **NP**-completeness of HC, no such approximation algorithm can exist unless $\mathbf{P} = \mathbf{NP}$. □

Problem 3.7: Reversible Turing machines

($\star\star\star$)

(1) Explain how to construct a reversible Turing machine that can compute the same class of functions as is computable on an ordinary Turing machine. (*Hint:* It may be helpful to use a multi-tape construction.)

(2) Give general space and time bounds for the operation of your reversible Turing machine, in terms of the time $t(x)$ and space $s(x)$ required on an ordinary single-tape Turing machine to compute a function $f(x)$.

Problem 3.8: Find a hard-to-compute class of functions (Research)

($\star\star\star$) Find a natural class of functions on n inputs which requires a super-polynomial number of Boolean gates to compute.

Solution

Concepts Involved: Boolean Functions, Computability, Circuit Lower Bounds

To the authors' knowledge, this remains an open problem. Indeed, it's worth noting that showing such a bound for any NP-complete function would imply P $\neq$ NP. An interesting result by Razborov and Rudich (Natural Proofs. Journal of Computer and System Sciences, 55(1):24-35, 1997.) is that any boolean functions that are "natural" (satisfying a constructivity and largeness condition) cannot be used to show super-polynomial lower bounds. □

Problem 3.9: Reversible $\mathbf{PSPACE} = \mathbf{PSPACE}$

($\star\star$) It can be shown that the problem 'quantified satisfiability', or QSAT, is **PSPACE**-complete. That is, every other language in **PSPACE** can be reduced to QSAT in polynomial time. The language QSAT is defined to consist of all Boolean formulae φ in n variables $x_1, \ldots, x_n$, and in conjunctive normal form, such that:

$$\exists_{x_1} \forall_{x_2} \exists_{x_3} \ldots \forall_{x_n} \varphi \text{ if } n \text{ is even;}$$
$$\exists_{x_1} \forall_{x_2} \exists_{x_3} \ldots \exists_{x_n} \varphi \text{ if } n \text{ is odd.}$$

Prove that a reversible Turing machine operating in polynomial space can be used to solve QSAT. Thus, the class of languages decidable by a computer operating reversibly in polynomial space is equal to **PSPACE**.

Problem 3.10: Ancilla bits and efficiency of reversible computation

($\star\star\star$) Let p_m be the mth prime number. Outline the construction of a reversible circuit which, upon the input of m and n such that $n > m$, outputs the product $p_m p_n$, that is $(m, n) \mapsto (p_m p_n, g(m, n))$ where $g(m, n)$ is the final state of the ancilla bits used by the circuit. Estimate the number of ancilla qubits your circuit requires. Prove that if a polynomial (in $\log n$) size reversible circuit can be found that uses $O(\log(\log n))$ ancilla bits then the problem of factoring a product of two prime numbers is in **P**.

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

($\star$) 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 $|v_1\rangle, \ldots, |v_k\rangle$ be the eigenvectors with eigenvalue 1 and $|v_{k+1}\rangle, \ldots, |v_n\rangle$ 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:

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

By Euler's identity we have:

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

Grouping terms, we obtain:

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

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:

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

□

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:

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

□

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:

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

which verifies equation (4.8). □

Exercise 4.6: Bloch sphere interpretation of rotations

$(\star)$ 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:

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

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:

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

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

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

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

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:

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

Note that in the second last equality we use that $\cos\left(\frac{\pi}{2}\right) = 0$ and hence $\frac{2}{\sqrt{2}}\cos\left(\frac{\pi}{2}\right) = \cos\left(\frac{\pi}{2}\right)$. 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

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

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

($\star\star$) 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

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

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 $\hat{m}$. □

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:

$$\begin{aligned}A &= R_z(\pi)R_y(-\pi/4)\\ B &= R_y(\pi/4)R_z(-\pi/2)\\ C &= R_z(-\pi/2)\end{aligned}$$

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:

$$\begin{aligned}HXH &= \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}2 & 0\\ 0 & -2\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = Z\\ HYH &= \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}0 & -2i\\ 2i & 0\end{bmatrix} = -\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} = -Y\\ HZH &= \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}0 & 2\\ 2 & 0\end{bmatrix} = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} = X\end{aligned}$$

□

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.

$$\begin{aligned}HXH &:= \frac{1}{2}(X+Z)X(X+Z) = \frac{1}{2}(I+ZX)(X+Z) = \frac{1}{2}(X+Z+Z+XZX) = Z\\ HYH &:= \frac{1}{2}(X+Z)Y(X+Z) = \frac{1}{2}(XY+ZY)(X+Z) = \frac{1}{2}(XYX+ZXY+ZYX+ZYZ) = -Y\end{aligned}$$

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

($\star$) 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 $R_{\hat{\mathbf{n}}_{12}}(\beta_{12})$.

$$\begin{aligned} R_{\hat{\mathbf{n}}_2}(\beta_2)R_{\hat{\mathbf{n}}_1}(\beta_1) &= (c_2 I - is_2\hat{\mathbf{n}}_2\cdot\boldsymbol{\sigma})\cdot(c_1 I - is_1\hat{\mathbf{n}}_1\cdot\hat{\sigma}) \\ &= c_2c_1 I - i\left(c_1s_2\hat{\mathbf{n}}_2\cdot\boldsymbol{\sigma} + c_2s_1\hat{\mathbf{n}}_1\cdot\boldsymbol{\sigma}\right) - s_2s_1 \underbrace{(\hat{\mathbf{n}}_2\cdot\boldsymbol{\sigma})\cdot(\hat{\mathbf{n}}_1\cdot\boldsymbol{\sigma})}_{(\hat{\mathbf{n}}_2\cdot\hat{\mathbf{n}}_1)I + i(\hat{\mathbf{n}}_2\times\hat{\mathbf{n}}_1)\cdot\boldsymbol{\sigma}} \\ &= \left[c_2c_1 - s_2s_1\left(\hat{\mathbf{n}}_2\cdot\hat{\mathbf{n}}_1\right)\right] I - i\left[c_1s_2\hat{\mathbf{n}_2} + c_2s_1\hat{\mathbf{n}_1} + s_2s_1\left(\hat{\mathbf{n}}_2\times\hat{\mathbf{n}}_1\right)\right]\cdot\boldsymbol{\sigma} \end{aligned}$$

Identifying this operation to a single rotation $R_{\hat{\mathbf{n}}_{12}}(\beta_{12}) \equiv c_{12}I - is_{12}\hat{\mathbf{n}}_{12}\cdot\boldsymbol{\sigma}$, we arrive at the required relations (up to a presumable typesetting error)

$$\begin{aligned} c_{12} &= c_2c_1 - s_2s_1\left(\hat{\mathbf{n}}_2\cdot\hat{\mathbf{n}}_1\right) \\ s_{12}\hat{\mathbf{n}}_{12} &= c_1s_2\hat{\mathbf{n}_2} + c_2s_1\hat{\mathbf{n}_1} + s_2s_1\left(\hat{\mathbf{n}}_2\times\hat{\mathbf{n}}_1\right) \end{aligned}$$

(2) Setting $\beta_1 = \beta_2$ and $\hat{\mathbf{n}}_1 = \hat{\mathbf{z}}$ in the formulas proven above combined with the fact that $c = c_1 = \cos(\beta_1/2) = \cos(\beta_2/2) = c_2$ (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

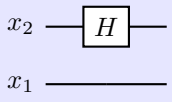
$$\begin{aligned} a_i\sigma_i b_j\sigma_j &= a_ib_j\delta_{ij} + i\left(a_ib_j\epsilon_{ijk}\right)\sigma_k \\ &= (\mathbf{a}\cdot\mathbf{b})\,I + i\,(\mathbf{a}\times\mathbf{b})_k\,\sigma_k, \end{aligned}$$

where we have used standard Einstein index notation. Thus in matrix form, we have

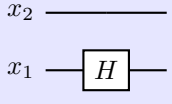
$$(\mathbf{a}\cdot\boldsymbol{\sigma})\cdot(\mathbf{b}\cdot\boldsymbol{\sigma}) = (\mathbf{a}\cdot\mathbf{b})\,I + i\,(\mathbf{a}\times\mathbf{b})\cdot\boldsymbol{\sigma}\,.$$

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

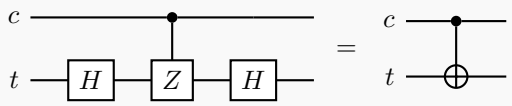
Construct a CNOT gate from one controlled-Z gate, that is, the gate whose action in the computational basis is specified by the unitary matrix

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

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:

$$\begin{aligned}(I_1 \otimes H_2)(CZ_{1,2})(I_1 \otimes H_2) &= \frac{1}{2}\begin{bmatrix}1 & 1 & 0 & 0\\ 1 & -1 & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 1 & -1\end{bmatrix}\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\end{bmatrix}\begin{bmatrix}1 & 1 & 0 & 0\\ 1 & -1 & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 1 & -1\end{bmatrix}\\ &= \frac{1}{2}\begin{bmatrix}2 & 0 & 0 & 0\\ 0 & 2 & 0 & 0\\ 0 & 0 & 0 & 2\\ 0 & 0 & 2 & 0\end{bmatrix}\\ &= CX_{1,2}\end{aligned}$$

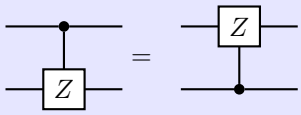
□

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:

$$\rho \cong \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

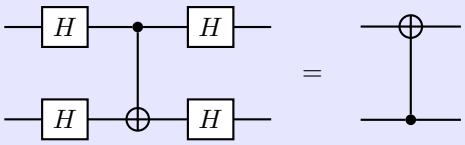
Studying the action of the `CNOT` gate on this density matrix, we calculate:

$$\begin{aligned} CX_{1,2}\,\rho\,CX_{1,2} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & a_{14} & a_{13} \\ a_{21} & a_{22} & a_{24} & a_{23} \\ a_{41} & a_{42} & a_{44} & a_{43} \\ a_{31} & a_{32} & a_{34} & a_{33} \end{bmatrix} \end{aligned}$$

□

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 $\text{CNOT}_{1,2}$ 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:

$$\begin{aligned}|+\rangle|+\rangle &\mapsto |+\rangle|+\rangle\\ |-\rangle|+\rangle &\mapsto |-\rangle|+\rangle\\ |+\rangle|-\rangle &\mapsto |-\rangle|-\rangle\\ |-\rangle|-\rangle &\mapsto |+\rangle|-\rangle\end{aligned}$$

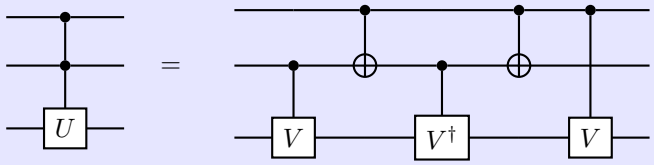
which is exactly what we wanted to prove. □

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 $VV = U$. □

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.

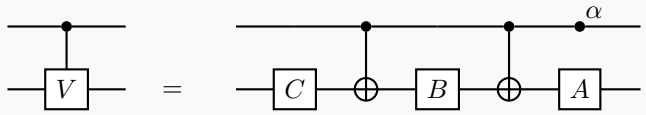
Exercise 4.22

($\star$) 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.

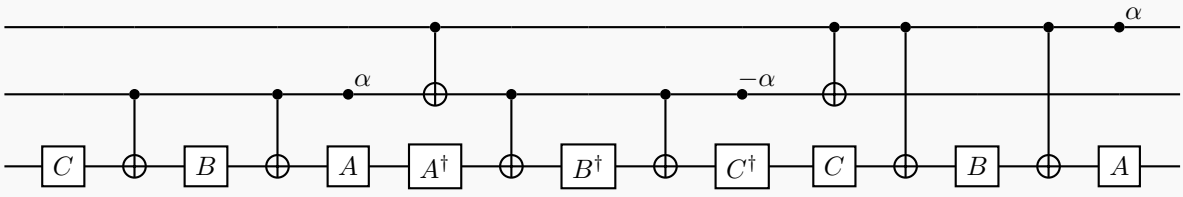
Solution

Concepts Involved: Unitary Operators, Controlled Operations, Gate Decomposition

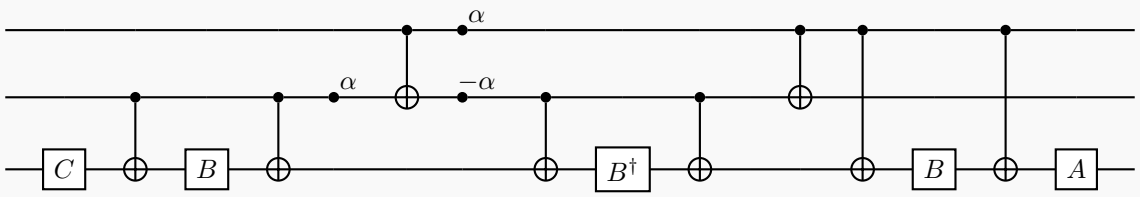
We use Figure 4.8 as our starting point (as depicted in Ex. 4.21). We then replace the controlled-V operations with the construction of Fig. 4.6:



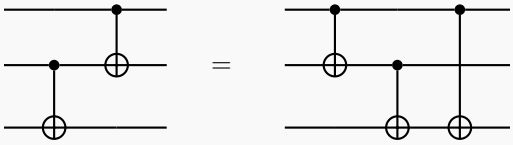
with the decomposition $V = \exp(i\alpha)AXBXC$ and $V^\dagger = \exp(-i\alpha)C^\dagger XB^\dagger XA^\dagger$. This gives the $C^2(U)$ circuit:



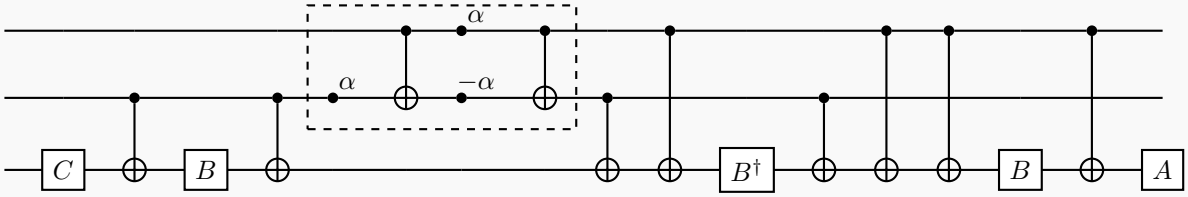
On the last qubit we have $AA^\dagger = C^\dagger C = I$, which removes these gates. Further, the α-phase gates commute through the controls, so let us move these towards the front of the circuit:



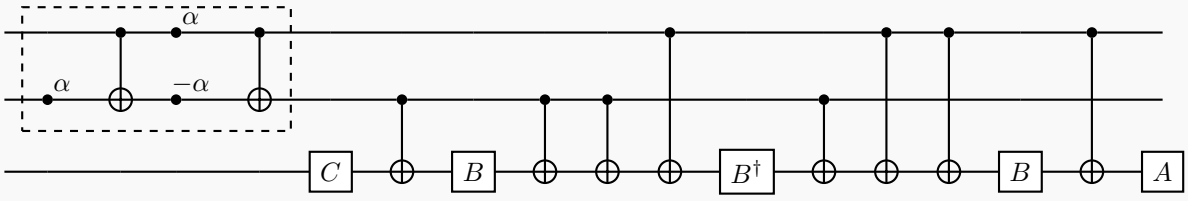
Then we make use of the identity:



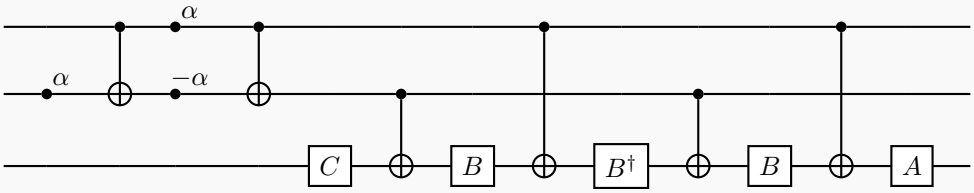
which follows from the fact that both circuits leave the first qubit unchanged, both circuits apply X to the second qubit only if the first qubit is $|1\rangle$, and both circuits apply X to the third qubit only if the second qubit is $|1\rangle$.
With this identity, we move the sixth `CNOT` past the 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 `CNOT`s mutually cancel, leaving us with:



which is an implementation of $C^2(U)$ only using eight one-qubit gates and six `CNOT`s, as claimed. □

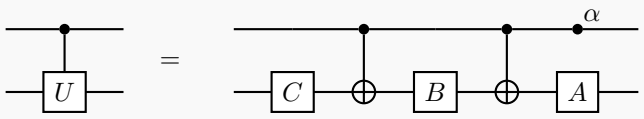
Exercise 4.23

($\star$) 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:

$$R_y(\theta) = e^{i\cdot 0}R_z(0)R_y(\theta)R_z(0) = e^{i\cdot 0}R_y(\theta/2)XR_y(-\theta/2)XI$$

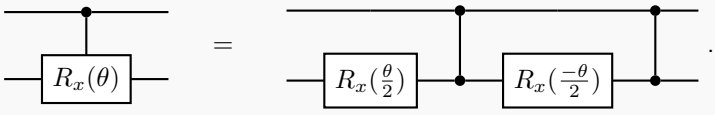
So we have the $C^1(R_y(\theta))$ gate as depicted above with $\alpha = 0$, $A = R_y(\theta/2)$, $B = R_y(-\theta/2)$, $C = I$. This is a construction with only two single qubit gates required.
For $R_x(\theta)$, we can write:

$$R_x(\theta) = R_z(\frac{\pi}{2})R_y(\theta)R_z(-\frac{\pi}{2}) = e^{i0}R_z(\frac{\pi}{2})R_y(\frac{\theta}{2})XR_y(-\frac{\theta}{2})XR_z(-\frac{\pi}{2})$$

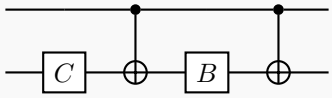
So we we have the $C^1(R_x(\theta))$ gate as depicted above with $\alpha = 0, A = R_z(\frac{\pi}{2})R_y(\frac{\theta}{2}), B = R_y(-\frac{\theta}{2}), C = R_z(-\frac{\pi}{2})$. This requires three single-qubit gates.
It is impossible to realize a $C^1(R_x(\theta))$ gate using using only two CNOTs and two single-qubit gates - we prove this at the end of this solution. However, we note that if we had access to two controlled-Z gates instead, then we could do a two CZ and two single-qubit gate construction. This follows from the observation:

$$R_x(\theta) = R_x(\frac{\theta}{2})ZR_x(-\frac{\theta}{2})Z$$

and so a $C^1(R_x(\theta))$ could be realized as:



Impossibility proof for constructing $C^1(R_x(\theta))$ with 2 CNOTs + 2 single-qubit gates. Since the state of the first qubit must be left invariant, no non-trivial single-qubit gates can act on the first qubit and both targets of the CNOT gates must act on the second qubit. Therein, we have three choices of single-qubit gate locations (before the two CNOTs, in-between, and after). One of the single-qubit gates must be placed in between (else, the CNOTs would cancel, as they are self-inverse). By symmetry, it suffices to consider the case where one of the single qubit gates comes before (call it C) and the other after (call it B), i.e. we have:



Suppose for the sake of contradiction that the above realizes $C^1(R_x(\theta))$ for arbitrary θ.
If the control qubit is 0, we have:

$$CB = I$$

and hence $B = C^\dagger$, and so $BC = C^\dagger C = I$ as well.
Conversely, if the control qubit is 1, we have:

$$CXBX = R_x(\theta)$$

wherein right multiplying both sides by X and taking the trace, we obtain:

$$\mathrm{Tr}(CXB) = \mathrm{Tr}\big(R_x(\theta+\pi)\big) \implies \mathrm{Tr}(BCX) = \mathrm{Tr}(X) = \mathrm{Tr}\big(R_x(\theta+\pi)\big)$$

where we have used that $X^2 = I$ and the cyclicity of the trace on the LHs. Since X is traceless, it must be the case that $\mathrm{Tr}\big(R_x(\theta+\pi)\big)$ vanishes, and so:

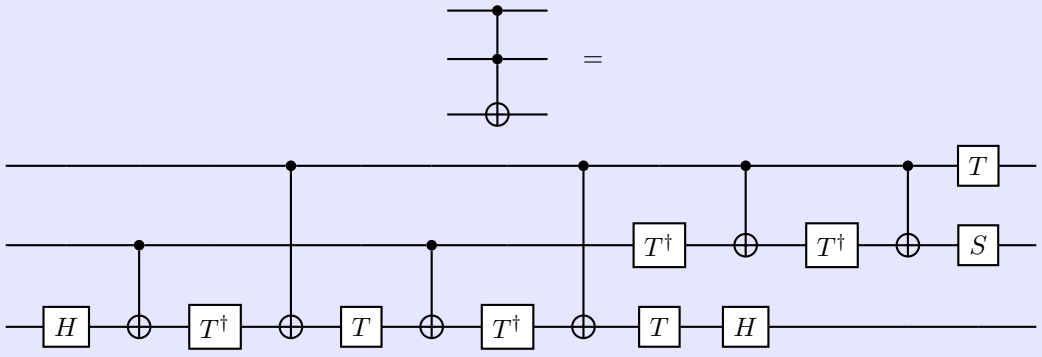
$$\mathrm{Tr}\big(R_x(\theta+\pi)\big) = \mathrm{Tr}\left(\cos\left(\frac{\theta+\pi}{2}\right)I - i\sin\left(\frac{\theta+\pi}{2}\right)X\right) = 2\cos\left(\frac{\theta+\pi}{2}\right) = 0 \implies \theta = 2n\pi, n \in \mathbb{Z}$$

So only $R_x(2n\pi) = I$ can be implemented in this way - contradiction. □

Remark: Thank you to Sam Buercklin for pointing out an error in the $C^1(R_x(\theta))$ construction of the original version of this solution, as well as for the proof sketch for the impossiblility of realizing $C^1(R_x(\theta))$ using 2 `CNOT`s + 2 single-qubit gates.

Exercise 4.24

($\star$) 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

($\star$) Recall that the Fredkin (controlled-swap) gate performs the transform

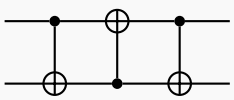
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

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?

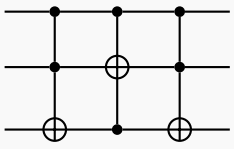
Solution

Concepts Involved: Controlled Operations

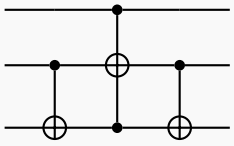
1. We recall the SWAP gate composed of three `CNOT` gates (Figure 1.7 in N&C):



   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:

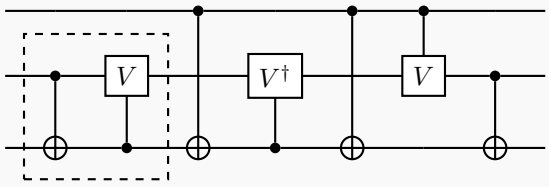


2. We claim that:



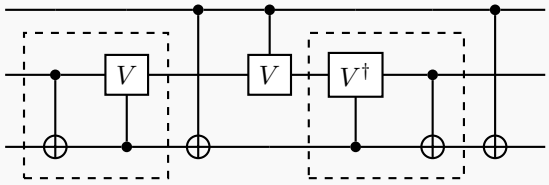
is also the Fredkin gate. In both circuits, the first qubit is left invariant (only acts as a control). The action on the second qubit is unchanged by removing the two Toffolis. The action on the third qubit requires a little more thought. If the first qubit is $|0\rangle$, then in the original circuit 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:

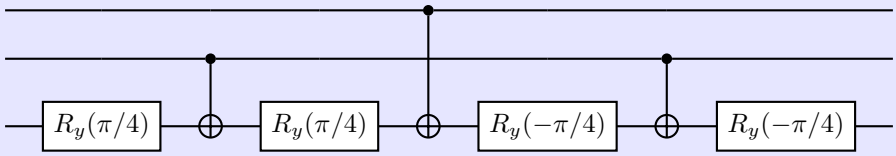


The second dashed box can also be viewed as a single two-qubit gate, and thus we can reduce the two-qubit gate count to five.

□

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

($\star\star$) 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:

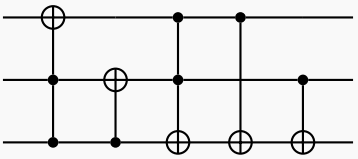
$$\begin{aligned} CX_{1,3} &= (56)(78) \\ CX_{2,1} &= (37)(48) \\ CX_{2,3} &= (34)(78) \\ CX_{3,1} &= (26)(48) \\ CX_{3,2} &= (24)(68) \end{aligned}$$

$$\mathsf{Toffoli}_1 = (48)$$
$$\mathsf{Toffoli}_2 = (68)$$
$$\mathsf{Toffoli}_3 = (78)$$

The transformation given in the question is the permutation (2345678) and now we can reason from the permutations directly. In particular, observe that:

$$\begin{aligned}(2345678) &= (34)(56)(78)(2468) \\ &= (34)(56)(78)(24)(68)(48) \\ &= (34)(56)(78)^2(78)(24)(68)(48) \\ &= \big[(34)(78)\big]\,\big[(56)(78)\big]\,\big[(78)\big]\,\big[(24)(68)\big]\,\big[(48)\big] \\ &= CX_{2,3}CX_{1,3}\mathsf{Toffoli}_3CX_{3,2}\mathsf{Toffoli}_1\end{aligned}$$

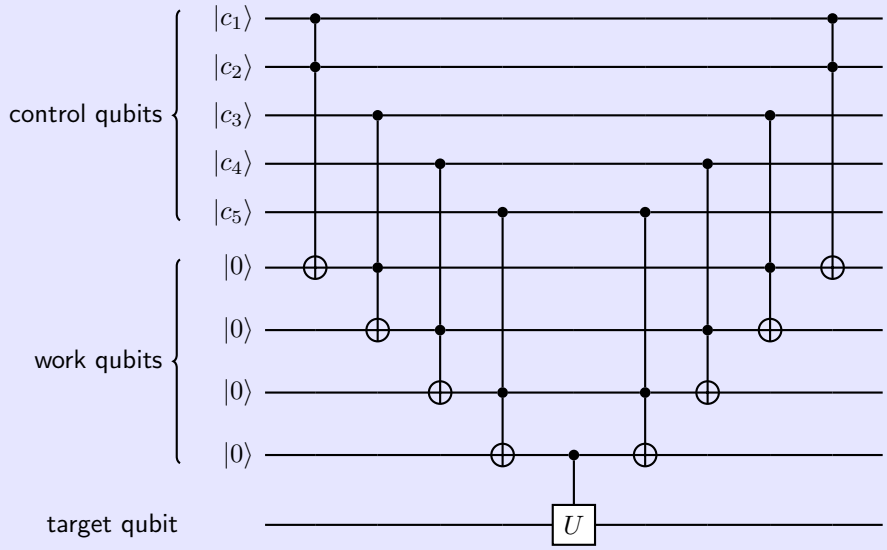
where the first two lines follow by permutation composition, the third follows by the fact that cycles square to the identity, and the fourth by the fact that disjoint cycles can be freely commuted. As a circuit:



□

Exercise 4.28

($\star\star$) 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.

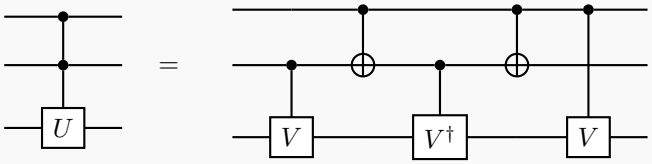


This question is impossible as stated - a proof based on determinants of the available gates (see this stackexchange post for details) shows that it is not possible.

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

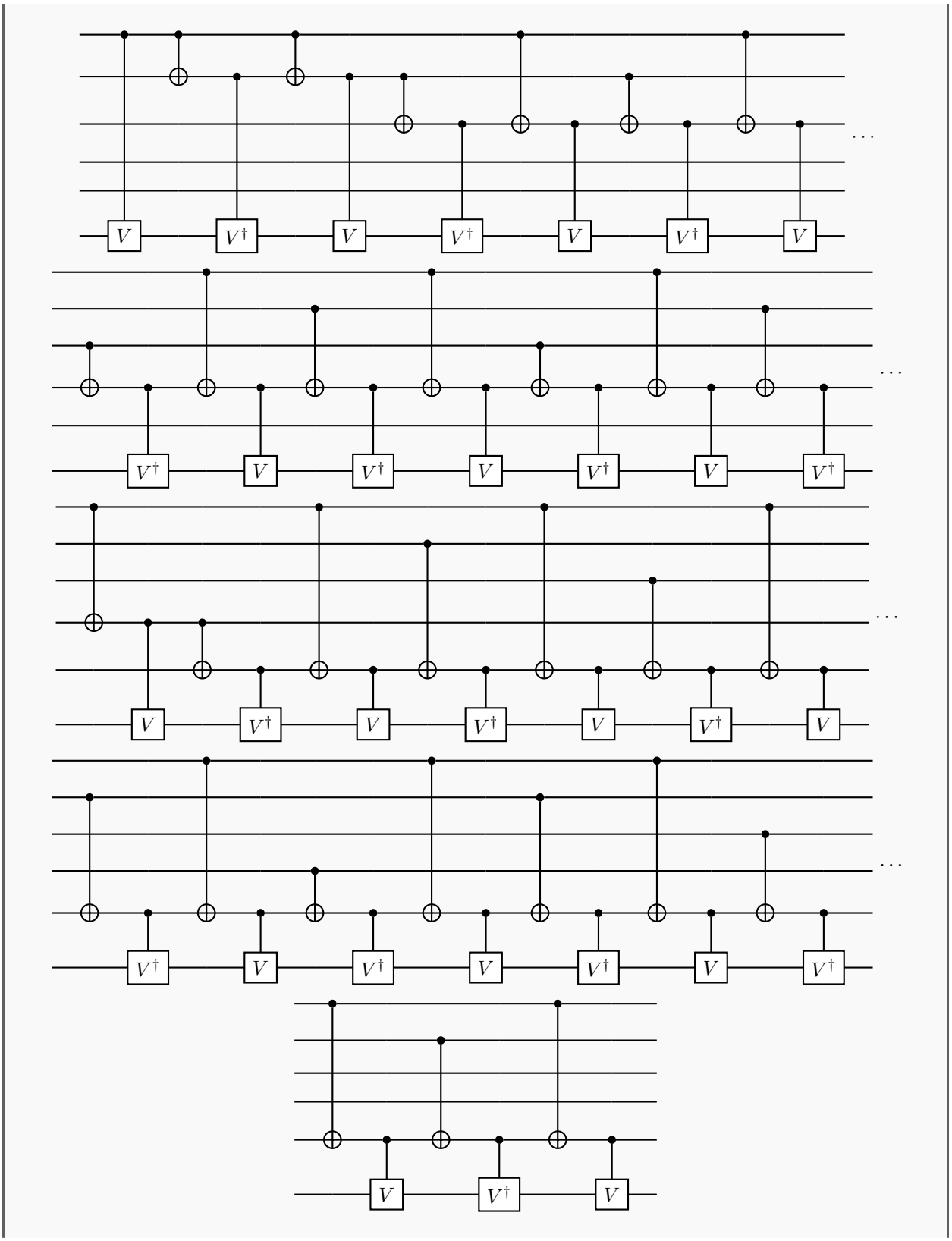
$$\underbrace{\sum_{k_1} x_{k_1}}_{V} - \underbrace{\sum_{k_1<k_2} (x_{k_1} \oplus x_{k_2})}_{V^\dagger} + \underbrace{\sum_{k_1<k_2<k_3} (x_{k_1} \oplus x_{k_2} \oplus x_{k_3})}_{V} - \ldots + (-1)^{n-1}(x_1 \oplus x_2 \oplus \ldots \oplus x_n) = 2^{n-1}(x_1 \wedge x_2 \wedge \ldots \wedge x_n)$$

(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

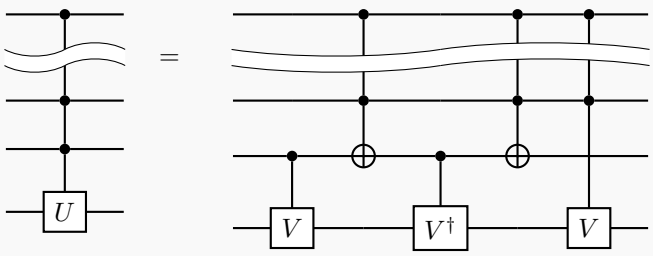
number of CNOT gates - in particular allowing us to implement the operation in 2^{n-1} $C^1(V)$, $2^{n-1}-1$ $C^1(V^\dagger)$, and 2^n-2 CNOT gates. Fig 4.8 is the $n=2$ version of this identity, where the first step realizes $+x_2$ (implementing V), the second step realizes $-x_1 \oplus x_2$ (implementing $V^\dagger$) 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†
V
V†
V
V†
V
...
V†
V
V†
V
V†
V
V†
...
V
V†
V
V†
V
V†
V
...
V†
V
V†
V
V†
V
V†
V
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

($\star\star\star$) 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

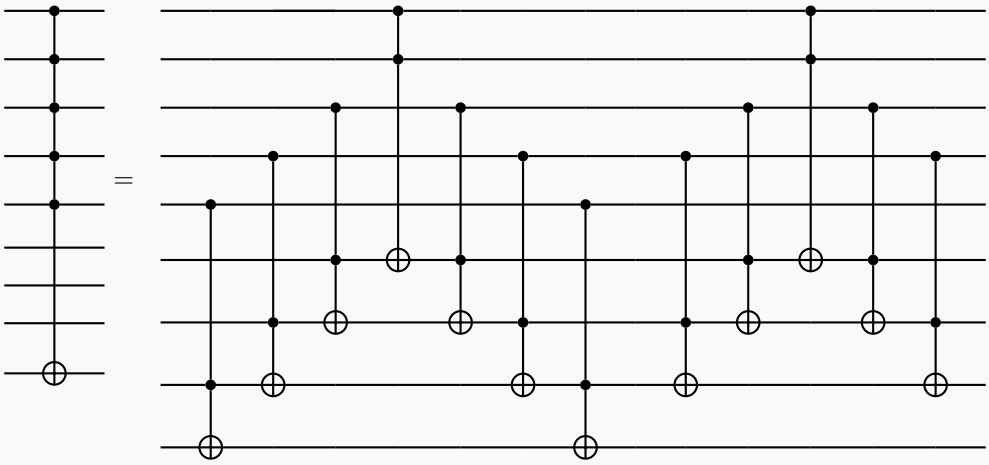
($\star\star\star$) Suppose U is a single qubit unitary operation. Find a circuit containing $O(n^2)$ Toffoli, `CNOT` and single qubit gates which implements a $C^n(U)$ gate (for $n > 3$), using no work qubits.

Solution

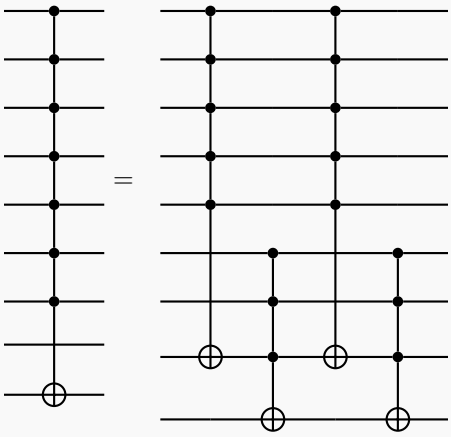
Concepts Involved: Controlled Operations, Gate Decomposition

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 $n = 5$:

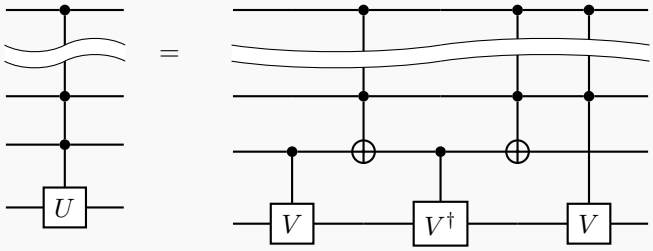


Looking at the first part of the circuit on the LHS, we can see that for $m \in \{1, \dots n-2\}$, the $n+m$th bit is flipped if the first m bits are 1. Hence if the first n bits are 1 then this correctly flips the target $2n-1$th qubit (the desired $C^n(X)$ action), at the expense of flipping work qubits $n+1 \dots 2n-2$. 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

$$\begin{aligned}\rho' &= p_0\rho_0 + p_1\rho_1 &(2)\\ &= \mathrm{Tr}(\rho P_0)\frac{P_0\rho P_0}{\mathrm{Tr}(\rho P_0)} + \mathrm{Tr}(\rho P_1)\frac{P_1\rho P_1}{\mathrm{Tr}(\rho P_1)} &(3)\\ &= P_0\rho P_0 + P_1\rho P_1 &(4)\end{aligned}$$

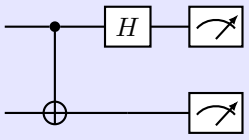
Finally, we have

$$\begin{aligned}\mathrm{Tr}_2(\rho') &= \mathrm{Tr}_2(P_0\rho P_0 + P_1\rho P_1)\\ &= \mathrm{Tr}_2(\rho P_0^2) + \mathrm{Tr}_2(\rho P_1^2)\\ &= \mathrm{Tr}_2(\rho P_0) + \mathrm{Tr}_2(\rho P_1)\\ &= \mathrm{Tr}_2(\rho(P_0 + P_1))\\ &= \mathrm{Tr}_2(\rho)\end{aligned}$$

□

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

$$M'_{00} = |00\rangle \left(\langle 00| \left(H_1 \otimes I_2\right) \texttt{CNOT}\right) = \frac{1}{\sqrt{2}}|00\rangle(\langle 00| + \langle 11|) = |00\rangle\langle\beta_{00}|$$
$$M'_{01} = |01\rangle \left(\langle 01| \left(H_1 \otimes I_2\right) \texttt{CNOT}\right) = \frac{1}{\sqrt{2}}|01\rangle(\langle 01| + \langle 10|) = |01\rangle\langle\beta_{01}|$$
$$M'_{10} = |10\rangle \left(\langle 10| \left(H_1 \otimes I_2\right) \texttt{CNOT}\right) = \frac{1}{\sqrt{2}}|10\rangle(\langle 00| - \langle 11|) = |10\rangle\langle\beta_{10}|$$
$$M'_{11} = |11\rangle \left(\langle 11| \left(H_1 \otimes I_2\right) \texttt{CNOT}\right) = \frac{1}{\sqrt{2}}|11\rangle(\langle 01| - \langle 10|) = |11\rangle\langle\beta_{11}|$$

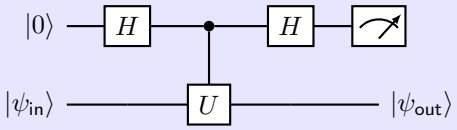
Therefore, it is easy to see that $\left(M'_k\right)^\dagger M'_k$ are projections onto the Bell states

$$\left(M'_{00}\right)^\dagger M_{00} = |\beta_{00}\rangle\langle\beta_{00}|$$
$$\left(M'_{01}\right)^\dagger M_{01} = |\beta_{01}\rangle\langle\beta_{01}|$$
$$\left(M'_{10}\right)^\dagger M_{10} = |\beta_{10}\rangle\langle\beta_{10}|$$
$$\left(M'_{11}\right)^\dagger M_{11} = |\beta_{11}\rangle\langle\beta_{11}|\,.$$

□

Exercise 4.34: Measuring an operator

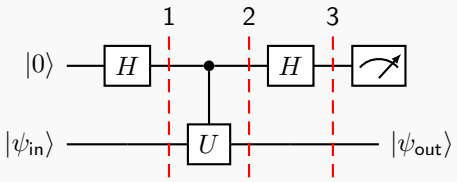
Suppose we have a single qubit operator U with eigenvalues ± 1, so that U is both Hermitian and unitary, so it can be regarded as both an observable and a quantum gate. Suppose we wish to measure the observable U. That is, we desire to obtain a measurement result indicating one of the two eigenvalues, 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, $|\psi_{\text{in}}\rangle = |u_\pm\rangle$.



$$|\Psi(t_1)\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)|u_\pm\rangle \quad .$$

$$\begin{aligned}|\Psi(t_2)\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle\,|u_\pm\rangle + |1\rangle U|u_\pm\rangle\right)\\ &= \frac{1}{\sqrt{2}}\left(|0\rangle\,|u_\pm\rangle \pm |1\rangle|u_\pm\rangle\right)\\ &= |\pm\rangle|u_\pm\rangle\end{aligned}$$

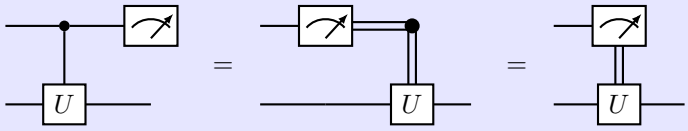
$$|\Psi(t_3)\rangle = |0/1\rangle|u_\pm\rangle$$

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 $|x, y\rangle \mapsto |x, x + y \bmod 4\rangle$.

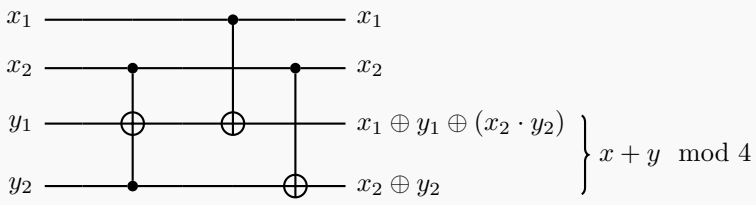
Solution

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

($\star$) 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

($\star$) 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

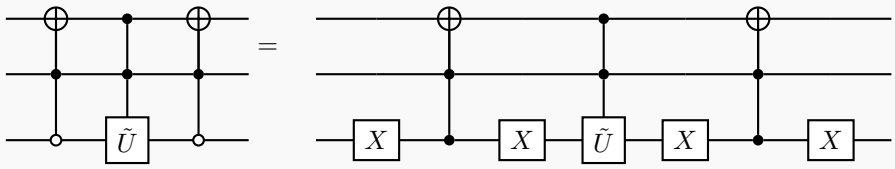
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & 0 & c \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & b & 0 & 0 & 0 & 0 & d \end{bmatrix}$$

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$(|010\rangle, |111\rangle)$ 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

$$E(R_n(\alpha), R_n(\alpha+\beta)) = \left|1 - \exp\big(i\beta/2\big)\right|,$$

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

$$\begin{aligned} \|I - R_n(\beta)\|_{\text{op}} &= \max\left(|1 - e^{i\frac{\beta}{2}}|, |1 - e^{-i\frac{\beta}{2}}|\right) \\ &= |1 - e^{i\frac{\beta}{2}}| \end{aligned}$$

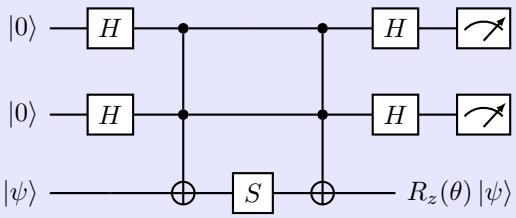
since both are equal. Thus, the error

$$E(R_n(\alpha), R_n(\alpha+\beta)) = |1 - e^{i\frac{\beta}{2}}|.$$

□

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.

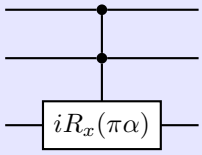
Solution

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

($\star\star$) 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 $G = \mathrm{CC}(iR_x(\pi\alpha))$ 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 $iR_x(\pi\alpha) \in \mathrm{SU}(2)$ 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 $\mathrm{SU}(2)$ 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.

$$\Rightarrow\ G \text{ is universal for quantum computation whenever } \alpha\notin\mathbb{Q}.$$

□

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:

$$\begin{aligned}&\text{Hermitian: } \rho = \rho^\dagger,\\ &\text{Trace One: } \mathrm{Tr}(\rho) = 1,\\ &\text{Positive Semidefinite: } \langle v|\rho|v\rangle \geq 0 \text{ for all vectors } |v\rangle.\end{aligned}$$

3. Counting Parameters

A. **Hermitian Condition:** A $2^n \times 2^n$ Hermitian matrix has:

$$\begin{aligned}&2^n \text{ real parameters for the diagonal,}\\ &\frac{(2^n)(2^n-1)}{2} \text{ complex parameters for the off-diagonal elements, contributing}\\ &(2^n)(2^n-1) \text{ real parameters.}\end{aligned}$$

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:

$$e^{-i\sum_{k=1}^{L} H_k t} = e^{-i(\sum_{k=1}^{L-1} H_k + H_L)t} = e^{-i\sum_{k=1}^{L-1} H_k t} e^{-iH_L t} = \prod_{k=1}^{L-1} e^{-iH_k t} e^{-iH_L t} = \prod_{k=1}^{L} e^{-iH_k t} \tag{5}$$

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

($\star$) 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,

$$E(U_{\Delta t}^m, e^{-2miH\Delta t}) \leq m\alpha\Delta t^3,$$

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

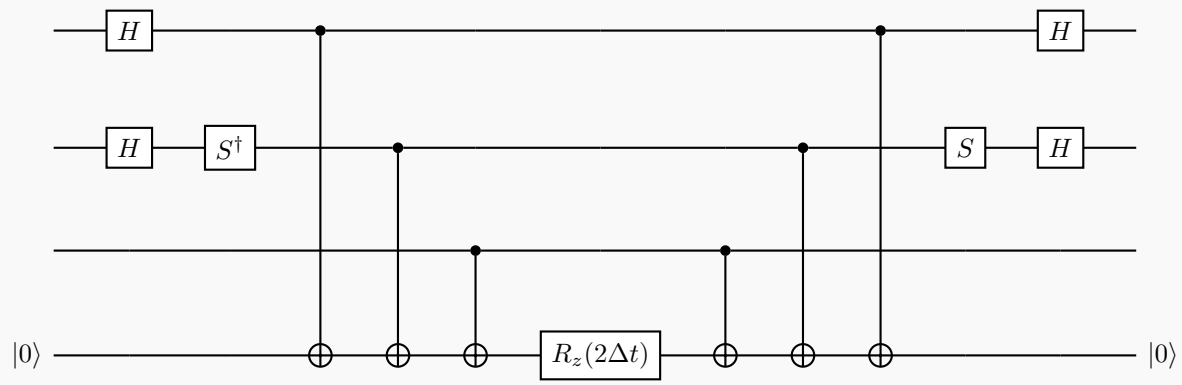
$$H = X_1 \otimes Y_2 \otimes Z_3$$

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 simultation 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

($\star\star$) 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 $|\tilde{1}\rangle = \mathrm{QFT}_n|1\rangle$ 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

$$2T + n.$$

□

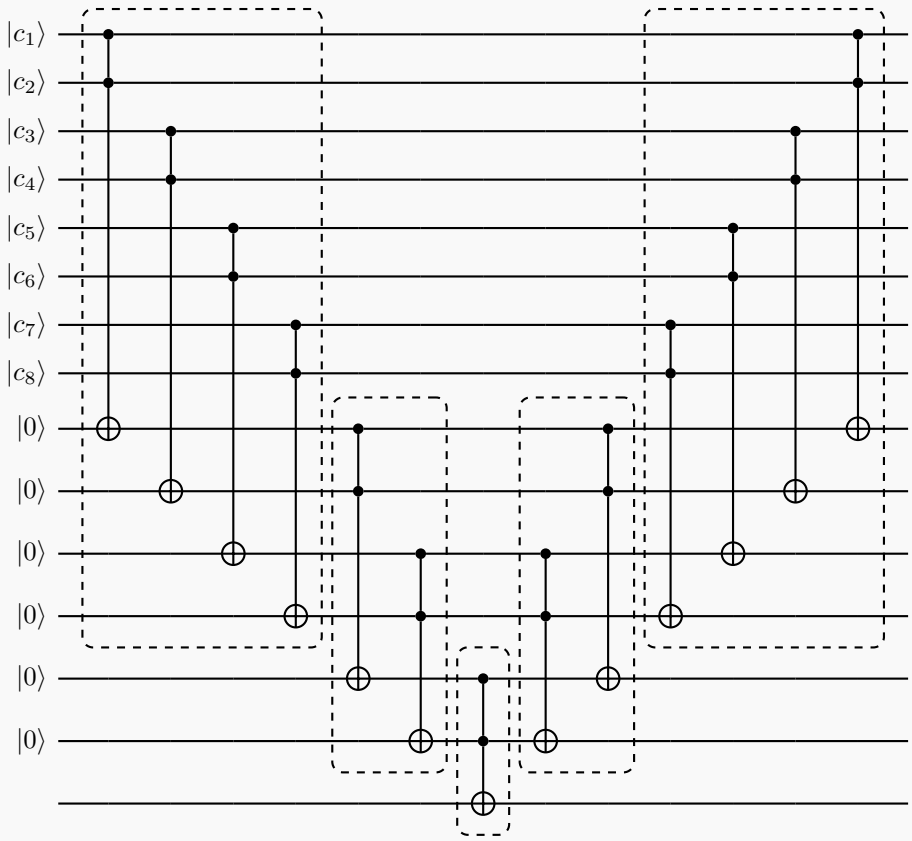
Problem 4.2

($\star$) 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

($\star\star$) 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 $\exp\bigl(-ih_g g\Delta\bigr)$ may be implemented using $O(n)$ one and two qubit operations.

(4) Show that

$$\exp(-iH\Delta) = \prod_g \exp\bigl(-ih_g g\Delta\bigr) + O(4^n\Delta^2)$$

where the product is taken with respect to any fixed ordering of the n-fold tensor products of Pauli matrices, g.

(5) Show that

$$U = \left[\prod_g \exp\bigl(-ih_g h\Delta\bigr)\right]^k + O(4^n\Delta)$$

(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

$$U = V\mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_{2^n}})V^\dagger \Rightarrow H = V\mathrm{diag}(\theta_1, \ldots, \theta_{2^n})V^\dagger,$$

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)

($\star\star\star$) 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) Yu, Duan, and Ying show in arXiv:1301.3372 that 5 two-qubit gates are necessary, so the construction given in Fig 4.8 of the text is optimal.

(2) Shende and Markov's construction in arXiv:0803.2316 shows that the minimal number of necessary `CNOT`s is 6, but optimizing the one-qubit gate count 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)

($\star\star\star$) 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: Hamiltonian Simulation, Quantum Complexity

To the authors' knowledge, this problem remains open. The problem can be sharpened by pointing out which families of Hamiltonians have known efficient simulation techniques (note: the following list is not nearly exhaustive, and quantum simulation remains a rich and active research area!):

- Before going to the realm of quantumly efficient, the Gottesman-Knill theorem (as discussed in Chapter 10, also in arXiv:quant-ph/9807006) shows that evolution of stabilizer states under stabilizer Hamiltonians is *classically* efficient. An analogous result for fermions due to Terhal and DiVincenzo (arXiv:quant-ph/0108010) shows that similarly, simulation of free-fermions is classically efficient.
- As discussed in the text, for local Hamiltonians one can use Trotter simulation (see, e.g. Lloyd science.273.5278.1073).
- More generically, sparse Hamiltonians admit efficient simulation (for example, the algorithm of arXiv:quant-ph/0508139 due to Berry, Ahokas, Cleve, and Sanders).

□

Remark: Although not quite the same setting, it is also worth highlighting the recent result of Anshu, Breuckmann, and Nirkhe `https://arxiv.org/abs/2206.13228`arXiv:2206.13228. They resolving the NLTS (No Low-lying Trivial States) conjecture by showing that a set of frustration free, local commuting Hamiltonians (based on constructions of good quantum LDPC codes) have low-energy states that all require $\Omega(\log n)$ circuit depth to prepare.

Problem 4.6: Universality with prior entanglement

($\star\star\star$) 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:

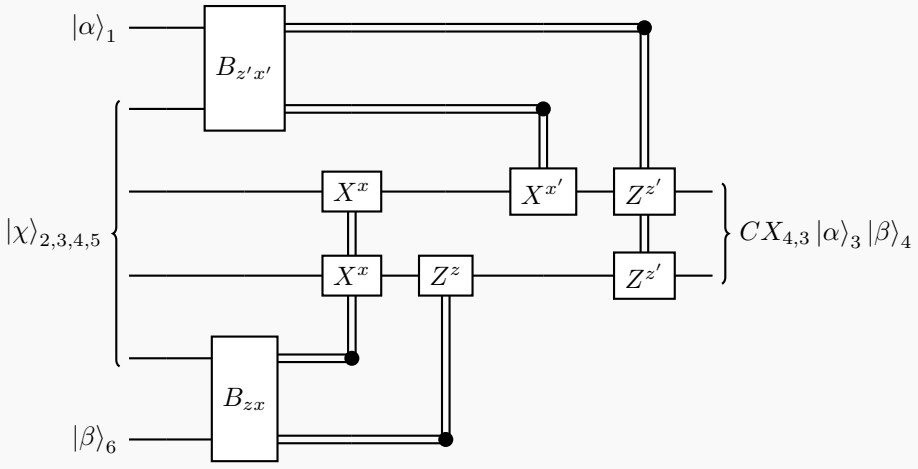
We follow the construction of Gottesman and Chuang, given in arXiv:quant-ph/9908010. Throughout, it will be helpful to recall the definition of the bell states:

$$|B_{zx}\rangle = \frac{|0x\rangle + (-1)^z |1\bar{x}\rangle}{\sqrt{2}} = Z_1^z X_1^x |B_{00}\rangle = Z_1^z X_1^x \frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

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:

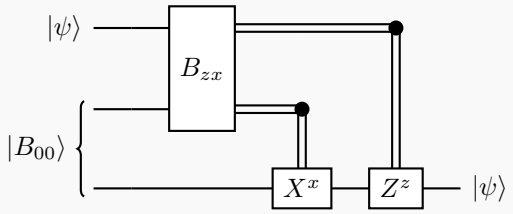
$$|\chi\rangle = \frac{(|00\rangle + |11\rangle)|00\rangle + (|01\rangle + |10\rangle)|11\rangle}{2} = \frac{|B_{00}\rangle|00\rangle + |B_{01}\rangle|11\rangle}{\sqrt{2}}$$

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, as proposed by Raussendorf and Briegel (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.

5 The quantum Fourier transform and its applications

Exercise 5.1

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

Solution

Concepts Involved: Unitary matrices, Roots of Unity

Let the linear transformation U be defined on the computational basis by

$$U_{kj} = \frac{1}{\sqrt{N}} e^{2\pi i jk/N}$$

To show that U is unitary, we compute the matrix product $U^\dagger U$. The matrix elements are

$$(U^\dagger U)_{j\ell} = \sum_{k=0}^{N-1} U_{kj}^* U_{k\ell} = \frac{1}{N} \sum_{k=0}^{N-1} e^{2\pi i k(\ell - j)/N}$$

This is a finite geometric series of N-th roots of unity. Its sum is

$$\sum_{k=0}^{N-1} e^{2\pi i k(\ell-j)/N} = \begin{cases} N & \text{if } j = \ell \\ 0 & \text{if } j \neq \ell \end{cases}$$

Therefore,

$$(U^\dagger U)_{j\ell} = \delta_{j\ell}$$

and so $U^\dagger U = I$, proving that U is unitary. □

Exercise 5.2

Explicitly compute the Fourier transform of the n qubit state $|00\ldots0\rangle$.

Solution

Concepts Involved: Quantum Fourier Transform (QFT), Computational Basis.

Let $N = 2^n$ and consider the n-qubit state $|0\rangle = |00\ldots0\rangle$. The QFT acts on computational basis states $|x\rangle$ as:

$$\text{QFT}_N |x\rangle = \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} e^{2\pi i xy/N} |y\rangle$$

For $x = 0$, we have $e^{2\pi i \cdot 0 \cdot y/N} = 1$, so:

$$\text{QFT}_N |0\rangle = \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} |y\rangle = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} |y\rangle$$

□

Remark: This is identical to applying Hadamard gates to each qubit

$$H^{\otimes n}|0\rangle = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} |y\rangle$$

Exercise 5.3: (Classical fast Fourier transform)

Suppose we wish to perform a Fourier transform of a vector containing 2^n complex numbers on a classical computer. Verify that the straightforward method for performing the Fourier transform, based upon direct evaluation of Equation (5.1) requires $\Theta(2^{2n})$ elementary arithmetic operations. Find a method for reducing this to $\Theta(n2^n)$ operations, based upon Equation (5.4).

Solution

Concepts Involved: Discrete Fourier Transform (DFT); computational complexity; Fast Fourier Transform (FFT); divide-and-conquer; binary decomposition.

Direct evaluation of the DFT from Equation (5.1)

$$y_k = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{2\pi i j k/N}$$

requires $\Theta(N)$ operations per output y_k, and there are $N = 2^n$ outputs, so total cost is $\Theta(N^2) = \Theta(2^{2n})$. Using the FFT idea inspired by Equation (5.4), we recursively split the input into even/odd parts. This leads to the recurrence

$$T(N) = 2T(N/2) + \Theta(N)$$

which solves to $T(N) = \Theta(N \log N) = \Theta(n2^n)$. □

Remark: The FFT is a classical algorithm that reduces exponential DFT cost to quasi-linear by exploiting symmetry in complex exponentials. It is structurally analogous to the QFT circuit decomposition.

Exercise 5.4

Give a decomposition of the controlled-R_k gate into single qubit and `CNOT` gates.

Solution

Concepts Involved: Controlled phase gate CR_k; single-qubit phase rotation; gate conjugation; universal gate set.

The controlled-R_k gate acts as

$$CR_k = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes R_k \quad \text{where} \quad R_k = \begin{bmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{bmatrix}$$

To decompose this using only single-qubit and CNOT gates, observe that any controlled-unitary CU can be written as

$$CU = (I \otimes V) \cdot \text{CNOT} \cdot (I \otimes V^\dagger) \cdot \text{CNOT} \quad \text{if } U = VXV^\dagger X$$

For diagonal gates R_k, this identity holds with $V = R_k^{1/2}$. Therefore

$$CR_k = (I \otimes R_k^{1/2}) \cdot \text{CNOT} \cdot (I \otimes R_k^{-1/2}) \cdot \text{CNOT}$$

This construction ensures that the target qubit is unaffected when the control is $|0\rangle$, and acquires a phase R_k when the control is $|1\rangle$, as desired. □

Remark: This decomposition exploits the identity $R_k = R_k^{1/2} X R_k^{-1/2} X$, and uses only 2 CNOTs and 2 single-qubit phase gates. It generalizes to any controlled diagonal gate.

Exercise 5.5

Give a quantum circuit to perform the inverse quantum Fourier transform.

Solution

Concepts Involved: Inverse quantum Fourier transform (QFT^{-1}), Hermitian adjoint, qubit reversal.

The inverse QFT on n qubits is the adjoint of the QFT:

$$\text{QFT}^{-1}|k\rangle = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{2^n-1} e^{-2\pi i jk/2^n} |j\rangle$$

To implement this in a quantum circuit:

(1) Reverse the order of the QFT circuit.

(2) Replace each gate with its Hermitian conjugate:

- $H^\dagger = H$
- $CR_k^\dagger =$ controlled-R_k^{-1}, i.e., rotation by $e^{-2\pi i/2^k}$

(3) Add SWAP gates at the end if needed to reverse qubit order.

□

Remark: This circuit uses $\Theta(n^2)$ gates. For practical implementations, small-angle controlled rotations can often be omitted to obtain an approximate inverse QFT with depth $\Theta(n \log n)$.

Exercise 5.6: Approximate quantum Fourier transform

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

Solution

Concepts Involved: Quantum Fourier transform (QFT)operator norm, approximation error.

Let U be the ideal QFT on n qubits, and V an approximation where each controlled-R_k gate is implemented with precision $\Delta = 1/p(n)$ for some polynomial $p(n)$. The QFT uses $\Theta(n^2)$ such gates.
By Equation (4.63) in the book, the total error from replacing each U_j by V_j satisfies:

$$\|U - V\| \leq \sum_j \|U_j - V_j\|$$

If each gate contributes at most Δ error, then:

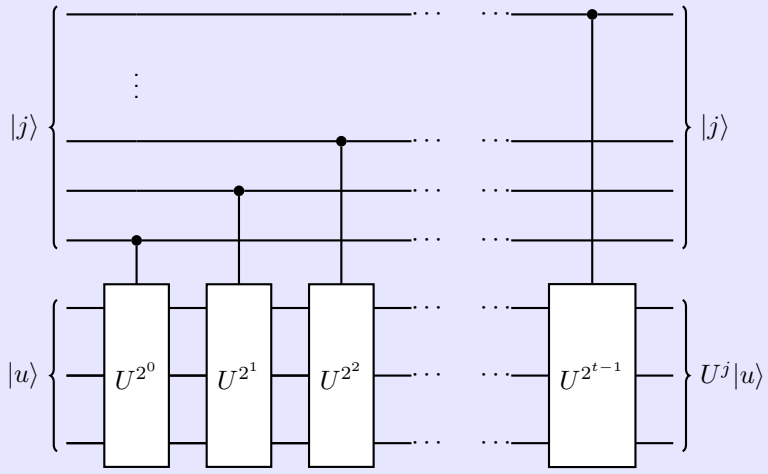
$$E(U,V) = \max_{|\psi\rangle} \|(U-V)|\psi\rangle\| = \Theta(n^2 \cdot \Delta) = \Theta\left(\frac{n^2}{p(n)}\right)$$

□

Remark: This shows that polynomial precision per gate yields overall polynomial accuracy. In practice, one may truncate small-angle controlled-R_k gates for large k, further simplifying the implementation without exceeding acceptable error bounds.

Exercise 5.7

Additional insight into the circuit in Figure 5.2 may be obtained by showing, as you should now do, that the effect of the sequence of controlled-U operations like that in Figure 5.2 is to take the state $|j\rangle|u\rangle$ to $|j\rangle U^j|u\rangle$. (Note that this does not depend on $|u\rangle$ being an eigenstate of U.)



Solution

Concepts Involved: Phase Estimation, Controlled Operations

The control register takes the form $|j\rangle = |j_{t-1}j_{t-2}\ldots j_1 j_0\rangle = |j_{t-1}\rangle\,|j_{t-2}\rangle\ldots|j_1\rangle\,|j_0\rangle$ (with j written out in binary). The first gate then applies $U^{(2^0)^{j_0}}$ to $|u\rangle$, the second gate then applies $U^{(2^1)^{j_1}}$ to $|u\rangle$, and so on until the tth gate applies $U^{(2^{t-1})^{j_{t-1}}}$ to $|u\rangle$. The net effect is:

$$|j\rangle\,|u\rangle \mapsto |j\rangle \prod_{i=0}^{t-1} U^{(2^i)^{j_i}}\,|u\rangle = |j\rangle\,U^j\,|u\rangle$$

as claimed. □

Exercise 5.8

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

Solution

Concepts Involved: Phase estimation, measurement accuracy, linearity of quantum circuits.

Given input $|0\rangle\sum_u c_u|u\rangle$, the phase estimation algorithm outputs $\sum_u c_u|\tilde{\varphi}_u\rangle|u\rangle$. For each eigenstate $|u\rangle$, the probability of measuring an estimate $\tilde{\varphi}_u$ accurate to n bits is at least $1-\varepsilon$, provided $t =$

$n + \lceil \log(2 + 1/2\varepsilon) \rceil$ (Equation 5.35). By linearity, the total probability of measuring such an estimate and collapsing to $|u\rangle$ is at least $|c_u|^2(1-\varepsilon)$. □

Remark: This ensures that approximate phase estimation remains reliable even on superposition inputs, with per-eigenstate success scaling with $|c_u|^2$.

Exercise 5.9

Let U be a unitary transform with eigenvalues ± 1, which acts on a state $|\psi\rangle$. Using the phase estimation procedure, construct a circuit to collapse $|\psi\rangle$ into one or the other of the two eigenspaces of U, giving also a classical indicator as to which space the final state is in. Compare your result to Exercise 4.34.

Solution

Concepts Involved: Phase Estimation, Quantum Measurement

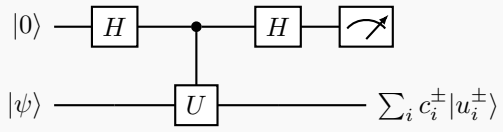
If U has eigenvalues ± 1 then it is also Hermitian, and thus its eigenvalues form an orthonormal basis $\left\{|u_j\rangle\right\}_j$. Any $|\psi\rangle$ can thus be expanded as:

$$|\psi\rangle = \sum_i c_i |u_i\rangle = \sum_i c_i^+ |u_i^+\rangle + \sum_{i'} c_{i'}^- |u_{i'}^-\rangle$$

Where we have divided the sum into the $+1$ eigenspace and the -1 eigenspace.
We go through the quantum phase estimation procedure with $e^{2\pi i\varphi_0} = 1 \implies \varphi_0 = 0, e^{2\pi i\varphi_1} \implies \varphi_1 = 1/2$, both which can be specified with a single decimal/bit worth of precision. So, using a single/$t = 1$ qubit initialized to $|0\rangle$, we have:

$$\begin{aligned}
|0\rangle|\psi\rangle &\overset{H_1}{\mapsto} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|\psi\rangle \\
&\overset{C(U)}{\mapsto} \frac{1}{\sqrt{2}}(|0\rangle|\psi\rangle + |1\rangle U|\psi\rangle) \\
&= \frac{1}{\sqrt{2}}\left(|0\rangle\left(\sum_i c_i^+|u_i^+\rangle + \sum_{i'} c_{i'}^-|u_{i'}^-\rangle\right) + |1\rangle\left(\sum_i c_i^+|u_i^+\rangle - \sum_{i'} c_{i'}^-|u_{i'}^-\rangle\right)\right) \\
&\overset{H_1}{\mapsto} |0\rangle\sum_i c_i^+|u_i^+\rangle + |1\rangle\sum_{i'} c_{i'}^-|u_i^-\rangle
\end{aligned}$$

Where we note the quantum fourier transform on a single qubit is just the Hadamard gate. If we now measure the first qubit, the second qubit collapses into the ± 1 eigenspace of U, with the measurement outcome as the classical indicator of the subspace. In circuit notation:



We note that the circuit of Ex. 4.34 is identical to this circuit (just with $|\psi\rangle$ a single qubit state), i.e. it was just the single-qubit phase estimation algorithm. □

Exercise 5.10

Show that the order of $x = 5$ modulo $N = 21$ is 6.

Solution

Concepts Involved: Modular arithmetic, Euler's theorem.

We seek the smallest r such that $5^r \equiv 1 \mod 21$. Compute successive powers of 5 modulo 21:

$$\begin{aligned} 5^1 &\equiv 5 \mod 21 \\ 5^2 &\equiv 25 \equiv 4 \mod 21 \\ 5^3 &\equiv 20 \mod 21 \\ 5^4 &\equiv 100 \equiv 16 \mod 21 \\ 5^5 &\equiv 80 \equiv 17 \mod 21 \\ 5^6 &\equiv 85 \equiv 1 \mod 21 \end{aligned}$$

Since 6 is the smallest such integer, the order of $5 \mod 21$ is 6. □

Remark: This confirms that 5 generates a cyclic subgroup of order 6 in $\mathbb{Z}_{21}^*$, consistent with $\varphi(21) = 12$, since $r = 6 \mid \varphi(21)$.

Exercise 5.11

Show that the order of x satisfied $r \leq N$.

Solution

Concepts Involved: Modular arithmetic, Euler's theorem, Lagrange's theorem, Group theory.

Let $x \in \mathbb{Z}_N^*$, and let r be the order of $x \mod N$, i.e., the smallest positive integer such that $x^r \equiv 1 \mod N$.
Then x lies in the multiplicative group $\mathbb{Z}_N^*$, which has $\varphi(N)$ elements. By Lagrange's theorem, the order r must divide $\varphi(N)$, so:

$$r \leq \varphi(N)$$

Since $\varphi(N) < N$ for all $N > 2$, it follows that:

$$r \leq \varphi(N) < N \Rightarrow r \leq N$$

□

Remark: This bound is useful when searching for the order in algorithms such as Shor's, as it ensures the period must lie in the finite range $1 \leq r \leq N$.

Exercise 5.12

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

Solution

Concepts Involved: Modular Arithmetic, Unitary Operators, Permutations.
U is defined by:

$$U\left|y\right\rangle \equiv \left|xy(\text{mod}N)\right\rangle .$$

When $N \leq y \leq 2^L - 1$, the convention is used that $xy \pmod N$ is just y again.

For $N \leq y \leq 2^L - 1$, U just acts as the identity. For $y \in \{0, \ldots, N-1\}$, since x, N are coprime, multiplication of y by x modulo N maps y to a unique $y' \in \{0, \ldots, N-1\}$. This can be observed as follows; we consider $yx, y'x \mod N$ for $y, y' \in \{0, \ldots, N-1\}$. If $yx \equiv y'x \pmod N$, then $(y-y')x = kN$ for some integer k. But since x, N are coprime, this can only be true if N divides $y-y'$; but $y-y' \leq N$ so this cannot be the case unless $y = y'$, so we conclude that all $yx \pmod N$ are unique.
The conclusion from the above is that U acts as a permutation operator on the computational basis states. In particular, U is real (satisfying $U^\dagger = U^T$), and (when represented in the computational basis) has one entry of 1 per row/column and is zero everywhere else. From this property, it follows that:

$$(UU^T)_{ij} = \sum_k U_{ik}U^T_{kj} = \sum_k U_{ik}U_{jk} = \delta_{ij}.$$

so $U^T = U^{-1}$ (this is true of all permutation matrices!). Thus, $U^\dagger U = U^T U = I$ and U is unitary. □

Remark: The hint points us towards the fact that we can consider U^{-1} defined by $U^{-1}\left|y\right\rangle \equiv \left|x^{-1}y(\text{mod } N)\right\rangle$ with x^{-1} the multiplicative inverse of x modulo N. Since U, U^{-1} act as real permutations, this implies $U^{-1} = U^\dagger$ (and thus, unitarity).

Exercise 5.13

Prove (5.44) (*Hint:* $\sum_{s=0}^{r-1} \exp\left(-2\pi isk/r\right) = r\delta_{k0}$.) In fact, prove that

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi isk/r} \left|u_s\right\rangle = \left|x^k \bmod N\right\rangle$$

Solution

Concepts Involved: Order Finding, Phase Estimation
$\left|u_s\right\rangle$ is defined to be:

$$\left|u_s\right\rangle \equiv \frac{1}{\sqrt{r}} \sum_{l=0}^{r-1} \exp\left[\frac{-2\pi isl}{r}\right] \left|x^l \text{mod } N\right\rangle$$

and (5.44) is:

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

Substituting in the definition off $|u_s\rangle$, we have:

$$\begin{aligned}\frac{1}{\sqrt{r}}\sum_{s=0}^{r-1} e^{2\pi isk/r}\,|u_s\rangle &= \frac{1}{\sqrt{r}}\sum_{s=0}^{r-1} e^{2\pi isk/r}\left(\frac{1}{\sqrt{r}}\sum_{l=0}^{r-1} e^{-2\pi isl/r}\,|x^l \bmod N\rangle\right)\\ &= \frac{1}{r}\sum_{l=0}^{r-1}|x^l \bmod N\rangle\sum_{s=0}^{r-1} e^{-2\pi is(l-k)/r}\\ &= \frac{1}{r}\sum_{l=0}^{r-1}|x^l \bmod N\rangle\, r\delta_{kl}\\ &= |x^k \bmod N\rangle\end{aligned}$$

where we apply the hint (sum over roots of unity) in the third equality. (5.44) follows by setting $k = 0$ on both sides of the above expression. □

Exercise 5.14

($\star\star$) The quantum state produced in the order-finding algorithm, before the inverse Fourier transform, is

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

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

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

and start the second register in the state $|0\rangle$. Also show how to construct V using $O(L^3)$ gates.

Solution

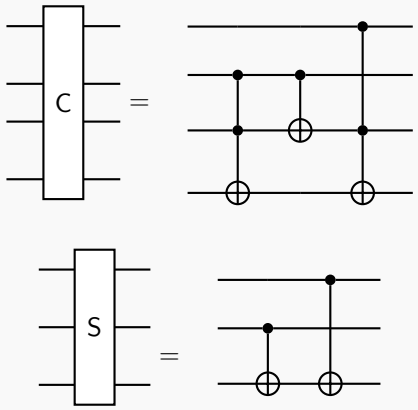
Concepts Involved: Order Finding, Phase Estimation, Modular Arithmetic

If we start the second register in the $|0\rangle$ state, then:

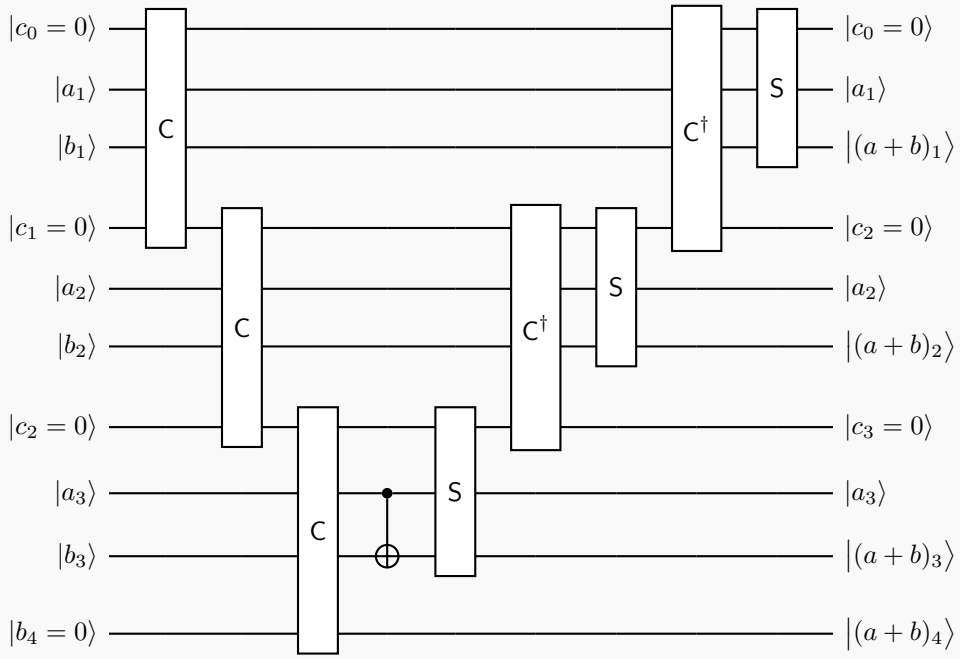
$$V\sum_{j=0}^{2^{t-1}-1}|j\rangle\,|0\rangle = \sum_{j=0}^{2^{t-1}-1} V\,|j\rangle\,|0\rangle = \sum_{j=0}^{2^{t-1}}|j\rangle\;|x^j \bmod N\rangle$$

which is the state we need.

Let us show how to construct V; the key additional ingridient required will be a reversible adder, which can be constructed from reversible classical gates, using the ingredients of Section 3.2.5. This presentation is adapted from Vedral, Barenco, and Ekert (arXiv:quant-ph/9511018). In particular, by combining the following bit carry (C) and sum (S) operations:



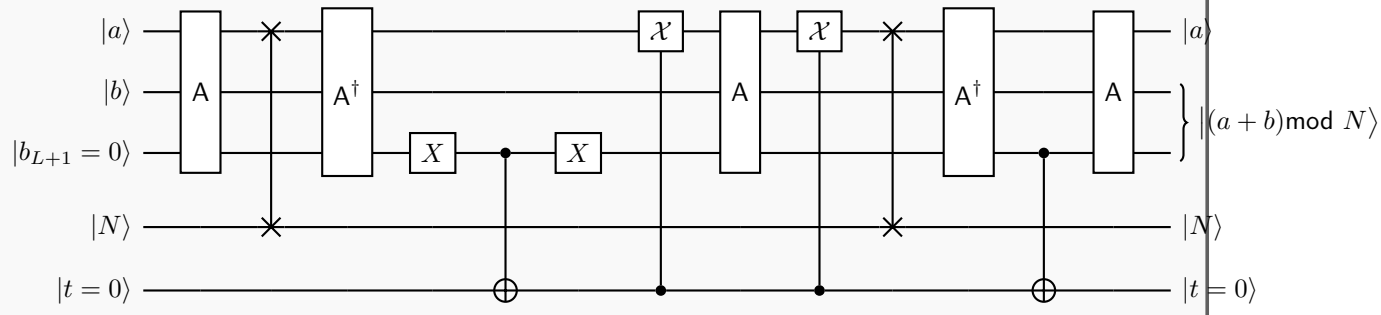
two L bit numbers $a = a_L a_{L-1} \ldots a_2 a_1$, $b = b_L b_{L-1} \ldots b_2 b_1$ can be added with the use of L ancilla (Which are initialized to and reset to 0) - below we show the $L = 3$ construction, which generalizes to arbitrary L by extending the cascade:



Decomposing the Toffolis in C into CNOT/H/T gates, by the construction of Fig. 4.9 each $C/C^\dagger$ requires 33 elementary gates. Since using the above reversible adder for adding two L bit number requires L C gates, $L-1$ $C^\dagger$ gates, L S gates, and one CNOT, we the number of gates required is:

$$33L + 33(L-1) + 2L + 1 = 68L - 32 = O(L).$$

Using the above circuit for a reversible adder (A), we can then construct a reversible modular adder as follows, using $\log_2(N) + 1$ ancilla (in addition to the L ancilla required for the standard addition operations):



The top two lines in the above are registers of multiple qubits holding a, b, and the fourth line is also a multi-qubit register holding N. The third line is a qubit holding the value of the $L+1$th bit of the addition (i.e. the overflow bit), and the fifth line is a temporary ancilla qubit $|t\rangle$ used to perform operations based on the value of the overflow.

The first A adds a, b. Then, the register SWAP + $A^\dagger$ subtracts N from the result. Then, based on the value of the overflow qubit, the temporary qubit $|t\rangle$ gets toggled. If it is toggled, the controlled-$\mathcal{X}$ gates fire. These are gates that toggle between $N \leftrightarrow 0$ (using $\log_2(N)$ `CNOT` gates), leaving the first register with value N if there is an overflow, and resetting to 0 if there is no overflow. In conjunction with these controlled-$\mathcal{X}$ gates, the third A adds N (or 0) back into the second register. The SWAP gate then resets the $|N\rangle$ register into its original state. The last three gates have the action of resetting the $|t\rangle$ ancilla into its original state - first by subtracting a (via the $\mathsf{A}^\dagger$) from the second register (giving $\big|(a+b)\text{mod } \mathsf{N} - a\big\rangle$, wherein the $|b_{L+1}\rangle$ signifies overflow in the subtraction, which is used to reset the $|t\rangle$ qubit via the `CNOT`. The final A adds back the subtracted a, leaving the register in the state $\big|(a+b)\text{mod } N\big\rangle$, and the ancilla qubits in their original state, as desired.

Each of the $\mathsf{A}/\mathsf{A}^\dagger$, register SWAPs, and controlled-$\mathcal{X}$ operations are computed with $O(L)$ operations, so the total cost of the reversible adder is again $7 \cdot O(L) + 4 = O(L)$.

Using the modular adder (which we call ADDN) and the established modular exponentiation/U^j gate, we can construct V as follows:

$$\begin{aligned} V_{1\to 2}\,|j\rangle_1\,|k\rangle_2\,|1\rangle_3 &= U^{j\dagger}_{1\to 3}\mathsf{ADDN}_{2\to 3}U^{j}_{1\to 3}\,|j\rangle_1\,|k\rangle_2\,|1\rangle_3 \\ &= U^{j\dagger}_{1\to 3}\mathsf{ADDN}_{2\to 3}\,|j\rangle_1\,|k\rangle_2\,|x^j \text{mod } N\rangle_3 \\ &= U^{j\dagger}_{1\to 3}\,|j\rangle_1\,|k + x^j \text{mod } N\rangle_2\,|x^j \text{mod } N\rangle_3 \\ &= |j\rangle_1\,|k + x^j \text{mod } N\rangle_2\,|1\rangle_3 \end{aligned}$$

Since U^j requires $O(L^3)$ gates and ADDM requires $O(L)$ gates, the total gate count is $2O(L^3) + O(L) = O(L^3)$. □

Remark: Although it is somewhat glossed over in Nielsen and Chuang/Box 5.2, the concrete circuits for performing the modular exponentiation unitary U^j indeed makes use of the reversible modular adder described - see Sections C/D and Figures 5/6 of arXiv:quant-ph/9511018 for details.

Exercise 5.15

($\star\star$) Show that the least common multiple of positive integers x and y is $xy/\gcd(x, y)$, and thus may be computed in $O(L^2)$ operations if x and y are L bit numbers.

Solution

Concepts Involved: Divisibility, Fundamental Theorem of Arithmetic, Arithmetic Algorithms

By the fundamental theorem of arithmetic, we can decompose x, y into their prime factors:

$$x = p_1^{x_1} p_2^{x_2} \dots p_n^{x_n} y \qquad\qquad = p_1^{y_1} p_2^{y_2} \dots p_n^{y_n}$$

wherein:

$$xy = p_1^{x_1+y_1} p_2^{x_2+y_2} \dots p_n^{x_n+y_n}.$$

We then observe that:

$$\begin{aligned}\gcd(x,y) &= p_1^{\min(x_1,y_1)} p_2^{\min(x_2,y_2)} \dots p_n^{\min(x_n,y_n)} \\ \operatorname{lcm}(x,y) &= p_1^{\max(x_1,y_1)} p_2^{\max(x_2,y_2)} \dots p_n^{\max(x_n,y_n)}\end{aligned}$$

wherein using that $\min(a,b) + \max(a,b) = a + b$:

$$\begin{aligned}\gcd(x,y) \cdot \operatorname{lcm}(x,y) &= p_1^{\min(x_1,y_1)+\max(x_1,y_1)} p_2^{\min(x_2,y_2)+\max(x_2,y_2)} \dots p_n^{\min(x_n,y_n)+\max(x_n,y_n)} \\ &= p_1^{x_1+y_1} p_2^{x_2+y_2} \dots p_n^{x_n+y_n} \\ &= xy\end{aligned}$$

and so $xy = \gcd(x,y) \cdot \operatorname{lcm}(x,y)$, and dividing both sides by $\gcd(x,y)$ the claim follows.
Multiplying x, y and dividing by $\gcd(x,y)$ takes $O(L^2)$ operations using the elementary school methos of multiplication, so the last thing to analyze is the time complexity of computing $\gcd(x,y)$. The analysis of Euclid's algorithm for the GCD done in Appendix 4 is not sufficiently optimal for our purposes ($O(L)$ division steps of $O(L^2)$ cost each, for a total cost of $O(L^3)$ operations). We thus consider the binary GCD algorithm, due to Stein. The algorithm is built on four observations:

(1) $\gcd(x,0) = x$

(2) $\gcd(2x,2y) = \gcd(x,y)$ (2 is a common divisor)

(3) $\gcd(x,2y) = \gcd(x,y)$ for x odd; 2 is not a common divisor

(4) $\gcd(x,y) = \gcd(x,y-x)$ if x, y are odd and $x \leq y$

The GCD algorithm thus involves the looping of (2)-(4) until the terminal condition (1) is reached. In pseudocode:

BINARYGCD(x, y):

if $x = 0$ *or* $y = 0$ **then**
  return $x + y$
end
while x, y *both even* **do**
  $x, y = x/2, y/2$
end
while $y \neq 0$ **do**
  while x, y *both not odd* **do**
    if x *even* **then**
      $x = x/2$
    else
      $y = y/2$
    end
  end
  if $x \geq y$ **then**
    $x, y = y, x$
  end
  $y = y - x$
end
return x

Note that the bitwise addition/subtraction take $O(L)$ operations, checking for odd/evenness takes $O(1)$ operations (just looking at the last bit), checking that a number is 0 requires $O(L)$ operations (looking at each bit) and division by 2 requires $O(L)$ operations (bitshift by one to the right). So, each application of one of the reduction conditions (1)-(4) requires $O(L)$ operations, and we at most need to reduce $O(L)$ times, for a total cost of $O(L^2)$ operations to find $\gcd(x, y)$. □

Exercise 5.16

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

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

and thus that (5.58) holds.

Solution

Concepts Involved: Definite integrals; monotonic decreasing functions, Series convergence.

For $x \geq 2$,

$$\int_x^{x+1} \frac{1}{y^2} dy = \frac{1}{x} - \frac{1}{x+1} = \frac{1}{x(x+1)} \geq \frac{2}{3x^2}$$

since $\frac{1}{x(x+1)} \geq \frac{2}{3x^2}$ holds iff $x \geq 2$.

Thus, for all integers $q \geq 2$,

$$\frac{1}{q^2} \leq \frac{3}{2}\int_q^{q+1} \frac{1}{y^2}dy \Rightarrow \sum_{q=2}^{\infty} \frac{1}{q^2} \leq \frac{3}{2}\int_2^{\infty} \frac{1}{y^2}dy = \frac{3}{2}\cdot\frac{1}{2} = \frac{3}{4}$$

□

Remark: This confirms the bound in equation (5.58), ensuring that the tail of the $\sum 1/q^2$ series contributes a controlled, finite error in approximating quantum Fourier transforms.

Exercise 5.17

($\star\star$) Suppose N is L bits long. The aim of this exercise is to find an efficient classical algorithm to determine whether $N = a^b$ for some integers $a \geq 1$ and $b \geq 2$. This may be done as follows:

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

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

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

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

Exercise 5.18: Factoring 91

Suppose we wish to factor $N = 91$. Confirm that steps 1 and 2 are passed. For step 3, suppose we choose $x = 4$, which is co-prime to 91. Compute the order r of x with respect to N, and show that $x^{r/2} \mod 91 = 64 \neq -1 \pmod{91}$, so the algorithm succeeds, giving $\gcd(64-1, 91) = 7$.
It is unlikely that this is the most efficient method you've seen for factoring 91. Indeed, if all computations had to be carried out on a classical computer, this reduction would not result in an efficient factoring algorithm, as no efficient method is known for solving the order-finding problem on a classical computer.

Solution

Concepts Involved: Factoring, Modular Arithmetic

Step 1 is passed as 91 is not even.
Step 2 is passed as 91 is not an integer power of another integer - here this is obvious by inspection as 91 is small; in principle we could run the algorithm from the previous exercise to check this, but it is also easy to enumerate and see that $3^2 = 9, 3^3 = 27, 3^4 = 81, 3^5 = 243 > 91, 5^2 = 25, 5^3 = 125 > 91, 7^1 = 7, 7^2 = 49, 7^3 = 343 > 91, 11^2 = 121 > 91$ at which we point we stop (note that powers of even integers are covered by step 1, and powers of 9 are of course powers of 3).
Step 3: We "choose" 4.

Step 4: Since the numbers are small, we can just do this by hand:

$$\begin{aligned}
4^1 &= 4 \pmod{91} \\
4^2 &= 16 \pmod{91} \\
4^3 &= 64 \pmod{91} \\
4^4 &= 256 = 74 \pmod{91} \\
4^5 &= 1024 = 23 \pmod{91} \\
4^6 &= 4096 = 1 \pmod{91}
\end{aligned}$$

so the the order of 4 w.r.t 91 is 6.
Step 5: 6 is even, and then we note that $4^{6/2} = 4^3 = 64 \neq -1 \pmod{91}$. We then find that $\gcd(64 - 1, 91) = \gcd(63, 91) = 7$ is a nontrivial factor (this could be done generically via Euclid's algorithm, here can be done by inspection) so the algorithm succeeds. Note that $\gcd(64 + 1, 91) = 13$ would have also given us a nontrivial factor (the other one, as $91 = 7 \cdot 13$). □

Exercise 5.19

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

Solution

Concepts Involved: Factoring

We proceed by enumeration of the odd numbers up to 15. 1 is not composite, $3, 5, 7$ are prime, $9 = 3^2$, $11, 13$ are prime, and so $15 = 3 \cdot 5$ is the smallest composite odd number that is not a power of a smaller integer. □

Exercise 5.20

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

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

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

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

The corrected formula is:

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

i.e. without the square root. We also instead consider:

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

Solution

Concepts Involved: Fourier Transforms

Let us break up the sum into groups, over $\{0, r, 2r, \ldots, N-r\}$, $\{1, r+1, 2r+1, \ldots, N-r+1\}, \ldots \{r-1, 2r-1, 3r-1, \ldots, N-1\}$ (i.e. multiples of r, shifted by $s = 0, \ldots r-1$):

$$\begin{aligned} \hat{f}(l) &= \frac{1}{N} \sum_{s=0}^{r-1} \sum_{x \in \{s, r+s, \ldots, N-r+s\}} e^{-2\pi i l x/N} f(x) \\ &= \frac{1}{N} \sum_{s=0}^{r-1} \sum_{k \in \{0,r,\ldots,N-r\}} e^{-2\pi i l (k+s)/N} f(k+s) \\ &= \frac{1}{N} \sum_{s=0}^{r-1} \sum_{k \in \{0,r,\ldots,N-r\}} e^{-2\pi i l s/N} e^{-2\pi i l k/N} f(s) \\ &= \frac{1}{N} \sum_{s=0}^{r-1} e^{-2\pi i l s/N} f(s) \sum_{k \in \{0,r,\ldots,N-r\}} e^{-2\pi i l k/N} \end{aligned}$$

where in the second equality we re-index the sum to be over $k = x - s$ and in the third equality we use that f is periodic in r to say that $f(k+s) = f(s)$ for all $k \in \{0, r, \ldots, N-r\}$. By the given identity (which applies just as well to a sum over $e^{-2\pi i k l/N}$ as it does to $e^{2\pi i k l/N}$, by taking the complex conjugate of both sides of the formula), if l is not an integer multiple of N/r, this vanishes, and if l is an integer

multiple of N/r, then the inner sum evaluates to N/r and so:

$$\begin{aligned}\hat{f}(l) &= \frac{1}{N}\sum_{s=0}^{r-1} e^{-2\pi i l s/N} f(s)\frac{N}{r} \\ &= \frac{1}{r}\sum_{s=0}^{r-1} e^{-2\pi i l s/N} f(s).\end{aligned}$$

(5.63) is precisely this result on the level of quantum states. □

Exercise 5.21: Period-finding and phase estimation

($\star\star$) Suppose you are given a unitary operator U_y which performs the transformation $U_y |f(x)\rangle = |f(x+y)\rangle$, for the periodic function described above.

(1) Show that the eigenvectors of U_y are $|\hat{f}(l)\rangle$, and calculate their eigenvalues.

(2) Show that given $|f(x_0)\rangle$ for some x_0, U_y can be used to realize a black box which is as useful as U in solving the period-finding problem.

We don't believe (2) is correct as stated - we instead suppose that we have access to a black box that is capable of performing the controlled-$U_y^{2^n}$ operation (as in the phase estimation algorithm), for $n = 0, 1, \ldots, L-1$.

Solution

Concepts Involved: Fourier Transforms, Period Finding, Phase Estimation

(1) We can check that the $|\hat{f}(l)\rangle$ are eigenvectors (and the associated eigenvalues) directly by applying U_y:

$$\begin{aligned}U_y |\hat{f}(l)\rangle &= \frac{1}{\sqrt{r}}\sum_{x=0}^{r-1} e^{-2\pi i l x/r} U_y |f(x)\rangle \\ &= \frac{1}{\sqrt{r}}\sum_{x=0}^{r-1} e^{-2\pi i l x/r} |f(x+y)\rangle \\ &= \frac{1}{\sqrt{r}}\sum_{x'=y}^{r-1+y} e^{-2\pi i l (x'-y)/r} |f(x')\rangle \\ &= e^{2\pi i l y/r}\frac{1}{\sqrt{r}}\sum_{x'=y}^{r-1+y} e^{-2\pi i l x'/r} |f(x')\rangle\end{aligned}$$

where in the third equality we introduce $x' = x + y$ and re-index the sum in terms of x'.

Since both $e^{-2\pi i l x'/r}$, $f(x')$ are periodic with period r, and the sum is over one (shifted by y) period,

it is equivalent to a sum from $x' = 0$ to $x' = r-1$, and so:

$$U_y \left|\hat{f}(l)\right\rangle = e^{2\pi i l y/r} \frac{1}{\sqrt{r}} \sum_{x'=0}^{r-1} e^{-2\pi i l x'/r} \left|f(x')\right\rangle = e^{2\pi i l y/r} \left|\hat{f}(l)\right\rangle$$

so the $\left|\hat{f}(l)\right\rangle$ (for each $l = 0, 1, \ldots, r-1$) are eigenvectors of U_y with eigenvalues $e^{2\pi i l y/r}$.

(2) As stated, we don't believe U_y can be used in the same way to solve the period finding problem - a single application can at most generate $U_y \left|f(x_0)\right\rangle = \left|f(x_0+y)\right\rangle$, which is insufficient to find the period, and even using repeated applications, we have no way to generate superpositions of $\left|f(x)\right\rangle$; thus, we assume as we had for quantum phase estimation that we have access to a black box that can apply controlled-$U_y^{2^n}$ gates for $n = 0, 1, \ldots, L-1$.

Given this, we can construct (as we saw for phase estimation) a circuit C of the form:

$$\left|j\right\rangle \left|u\right\rangle \overset{C}{\to} \left|j\right\rangle U_y^j \left|u\right\rangle$$

using the controlled-$U_y^{2^n}$ black boxes (see the construction of Fig. 5.2 and Ex. 5.7). Then, we can run the period finding procedure, but now starting with initial state $\left|0\right\rangle \left|f(x_0)\right\rangle$ and applying C instead of U:

$$\begin{aligned}
&\left|0\right\rangle \left|f(x_0)\right\rangle && \text{initial state} \\
&\overset{H^{\otimes t}}{\mapsto} \frac{1}{\sqrt{2^t}} \sum_{x=0}^{2^t-1} \left|x\right\rangle \left|f(x_0)\right\rangle && \text{create superposition} \\
&\overset{C}{\mapsto} \frac{1}{\sqrt{2^t}} \sum_{x=0}^{2^t-1} \left|x\right\rangle U_y^x \left|f(x_0)\right\rangle && \text{apply } C \\
&= \frac{1}{\sqrt{2^t}} \sum_{x=0}^{2^t-1} \left|x\right\rangle \left|f(x_0+xy)\right\rangle && \\
&\approx \frac{1}{\sqrt{r2^t}} \sum_{l=0}^{r-1} \sum_{x=0}^{2^t-1} e^{2\pi i l(x_0+xy)/r} \left|x\right\rangle \left|\hat{f}(l)\right\rangle && \\
&= \frac{1}{\sqrt{r2^t}} \sum_{l=0}^{r-1} e^{2\pi i l x_0} \sum_{x=0}^{2^t-1} e^{2\pi i l x y/r} \left|x\right\rangle \left|\hat{f}(l)\right\rangle && \\
&\overset{QFT^{-1}}{\mapsto} \frac{1}{\sqrt{r}} \sum_{l=0}^{r-1} \left|\widetilde{ly/r}\right\rangle \left|\hat{f}(l)\right\rangle && \text{apply inverse Fourier transform to first register} \\
&\to \widetilde{ly/r} && \text{measure first register} \\
&\to r/\gcd(ly, r) && \text{apply continued fractions algorithm}
\end{aligned}$$

Thus the blackbox + $O(L^2)$ applications allows us to find $r/\gcd(ly, r)$.
Note that we don't get r directly from this due to the extra factor arising from y, but instead a divisor of r of the form $d_i = r/\text{gcd}(l_i y, r)$ for a random $l_i \in \{0, 1, \ldots, r-1\}$. However, we can obtain r with high probability by repeating this procedure $m = O(L)$ times and considering $r = \text{lcm}(d_1, d_2, \ldots d_m)$.
Let's see that this works. Letting $r = p_1^{\alpha_1} \ldots p_n^{\alpha_n}$ be the prime factorization of r, we want $p_j^{\alpha_j}$ to be

contained in some $d_i = r/\gcd(l_i y, r)$, which is the case if $p_j^{\alpha_j} \nmid l_i y$. In other words, we have failure if $p_j^{\alpha_j - \beta} | l_i$ where p^β is the prime factor contained in y. This occurs with probability $1/p^{\alpha_j - \beta} \leq 1/p \leq 1/2$. Thus taking the LCM over m runs, the chance of failing to include $p_j^{\alpha_j}$ is $O(2^{-m})$, and for at most $O(L)$ prime factors via union bound the chance of failing to capture all prime factors is $O(L \cdot 2^{-m})$. Taking $m = O(L)$ this is exponentially small in L. Thus, with $O(L^3)$ operations ($O(L)$ runs of the $O(L^2)$ algorithm), we are able to (with high probability) obtain r.
We again remark that access to U_y (and a $|f(x_0)\rangle$ is not as powerful as having access to U - we require controlled-$U_y^{2^n}$ operations, and even with these, require $O(L^3)$ as opposed to $O(L^2)$ operations to find r.
□

Exercise 5.22

Show that

$$|\hat{f}(l_1, l_2)\rangle = \sum_{x_1=0}^{r-1} \sum_{x_2=0}^{r-1} e^{-2\pi i (l_1 x_1 + l_2 x_2)/r} |f(x_1, x_2)\rangle = \frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} e^{-2\pi i l_2 j/r} |f(0, j)\rangle ,$$

and we are constrained to have $l_1/s - l_2$ be an integer multiple of r for this expression to be non-zero.
In order for the pre-factors to work out, some adjustment is needed to the provided formulas; we thus show:

$$|\hat{f}(l_1, l_2)\rangle = \frac{1}{r} \sum_{x_1=0}^{r-1} \sum_{x_2=0}^{r-1} e^{-2\pi i (l_1 x_1 + l_2 x_2)/r} |f(x_1, x_2)\rangle = \sum_{j=0}^{r-1} e^{-2\pi i l_2 j/r} |f(0, j)\rangle .$$

We also replace the condition with "$l_1 - l_2 s$ is an integer multiple of r".

Solution

Concepts Involved: Fourier Transforms, Discrete Logarithms.
Recall that $f(x_1, x_2) = b^{x_1} a^{x_2} \pmod N = a^{s x_1} a^{x_2} \pmod N$

The first/third expressions are equal by definition, so it suffices to show that the second/third are equivalent.
In the second expression, let $j = x_2 + s x_1 \mod r$ so $x_2 = j - s x_1$. Then, note that:

$$f(x_1, x_2) = f(x_1, j - s x_1) = a^{s x_1} a^{j - s x_1} = a^0 a^j = f(0, j)$$

and so re-indexing the second sum in terms of j:

$$\begin{aligned}\frac{1}{r}\sum_{x_1=0}^{r-1}\sum_{x_2=0}^{r-1} e^{-2\pi i(l_1x_1+l_2x_2)/r}\left|f(x_1,x_2)\right\rangle &= \frac{1}{r}\sum_{x_1=0}^{r-1}\sum_{j=sx_1 \bmod r}^{r-1+sx_1 \bmod r} e^{-2\pi i x_1(l_1-sl_2)(l_1x_1+l_2(j-sx_1))}\left|f(0,j)\right\rangle \\ &= \frac{1}{r}\sum_{x_1=0}^{r-1}\sum_{j=0}^{r-1} e^{-2\pi i l_2 j/r}e^{-2\pi i(l_1-sl_2)x_1/r}\left|f(0,j)\right\rangle \\ &= \frac{1}{r}\sum_{j=0}^{r-1} e^{-2\pi i l_2 j/r}\left|f(0,j)\right\rangle\left(\sum_{x_1=0}^{r-1} e^{-2\pi i(l_1-sl_2)x_1/r}\right)\end{aligned}$$

The sum in brackets equals r if $l_1 - l_2 s$ is an integer multiple of r, and is zero otherwise, so considering the former case:

$$\begin{aligned}\frac{1}{r}\sum_{x_1=0}^{r-1}\sum_{x_2=0}^{r-1} e^{-2\pi i(l_1x_1+l_2x_2)/r}\left|f(x_1,x_2)\right\rangle &= \frac{1}{r}\sum_{j=0}^{r-1} e^{-2\pi i l_2 j/r}\left|f(0,j)\right\rangle\cdot r \\ &= \sum_{j=0}^{r-1} e^{-2\pi i l_2 j/r}\left|f(0,j)\right\rangle\end{aligned}$$

which was the desired relation. □

Exercise 5.23

Compute

$$\frac{1}{r}\sum_{l_1=0}^{r-1}\sum_{l_2=0}^{r-1} e^{-2\pi i(l_1x_1+l_2x_2)/r}\left|\hat{f}(l_1,l_2)\right\rangle$$

using (5.70), and show that the result is $f(x_1, x_2)$.
This should read $\left|f(x_1,x_2)\right\rangle$ (the output should be a state), and there is also a typo in the coefficient, where we should have $e^{+2\pi i(l_1x_1+l_2x_2)/r}$ for the inverse transform.

Solution

Concepts Involved: Fourier Transforms, Discrete Logarithms

Substituting in our expression for $|\hat{f}(l_1, l_2)\rangle$ from the previous exercise, we find:

$$\begin{aligned}
&\frac{1}{r}\sum_{l_1=0}^{r-1}\sum_{l_2=0}^{r-1} e^{2\pi i(l_1x_1+l_2x_2)/r}\,|\hat{f}(l_1,l_2)\rangle \\
&= \frac{1}{r}\sum_{l_1=0}^{r-1}\sum_{l_2=0}^{r-1} e^{2\pi i(l_1x_1+l_2x_2)/r}\left(\frac{1}{r}\sum_{x_1=0}^{r-1}\sum_{x_2=0}^{r-1} e^{-2\pi i(l_1x_1'+l_2x_2')/r}\,|f(x_1',x_2')\rangle\right) \\
&= \frac{1}{r^2}\sum_{x_1'=0}^{r-1}\sum_{x_2'=0}^{r-1} |f(x_1',x_2')\rangle\left(\sum_{l_1=0}^{r-1} e^{-2\pi i l_1(x_1'-x_1)}\right)\left(\sum_{l_2=0}^{r-1} e^{-2\pi i l_2(x_2'-x_2)}\right) \\
&= \frac{1}{r^2}\sum_{x_1'=0}^{r-1}\sum_{x_2'=0}^{r-1} |f(x_1',x_2')\rangle\,(r\delta_{x_1x_1'})(r\delta_{x_2x_2'}) \\
&= |f(x_1,x_2)\rangle
\end{aligned}$$

□

Exercise 5.24

Construct the generalized continued fractions algorithm needed in step 6 of the discrete logarithm algorithm to determine s from estimates of sl_2/r and l_2/r.

Solution

Concepts Involved: Continued Fractions, Discrete Logarithm

Running continued fractions on $\widetilde{l_2/r}$, we obtain r (note that the success probability is technically not quite $O(1)$ of obtaining r depending on its prime factorization - we technically get $r/\text{gcd}(r, l_2)$. This can be remedied by doing many runs of the algorithm with randomized l_2 and taking the least common multiples of $i = 1, \ldots, \log(N)$ such $r/\text{gcd}(r, l_2^i)$s, e.g. as is discussed in Ex. 5.21). We can then extract s as the integer approximation to:

$$s \approx \frac{\widetilde{sl_2/r} \cdot r}{l_2}$$

and verify that $b = a^s \pmod N$. □

Exercise 5.25

Construct a quantum circuit for the black box U used in the quantum discrete logarithm algorithm, which takes a and b as parameters, and performs the unitary transform $|x_1\rangle|x_2\rangle|y\rangle \mapsto |x_1\rangle|x_2\rangle|y \oplus b^{x_1}a^{x_2}\rangle$. How many elementary operations are required?

Solution

Concepts Involved: Discrete Logarithm, Controlled Operations, Modular Arithmetic

We can essentially combine the modular exponentiation unitary of Box 5.2:

$$U_x \left|z\right\rangle \left|y\right\rangle = \left|z\right\rangle \left|x^z y \pmod N\right\rangle$$

and the modular addition unitary of Ex. 5.14:

$$V_x \left|j\right\rangle \left|k\right\rangle = \left|j\right\rangle \left|k + x^j \pmod N\right\rangle$$

where here we take $N = 2$. We then observe:

$$\begin{aligned} U_{b,1\to 3} V_{a,2\to 3} U_{b,1\to 3}^\dagger \left|x_1\right\rangle \left|x_2\right\rangle \left|y\right\rangle &= U_{b,1\to 3} V_{a,2\to 3} \left|x_1\right\rangle \left|x_2\right\rangle \left|b^{-x_1} y \pmod 2\right\rangle \\ &= U_{b,1\to 3} \left|x_1\right\rangle \left|x_2\right\rangle \left|b^{-x_1} y + a^{x_2} \pmod 2\right\rangle \\ &= \left|x_1\right\rangle \left|x_2\right\rangle \left|b^{x_1} b^{-x_1} y + b^{x_1} a^{x_2} \pmod 2\right\rangle \\ &= \left|x_1\right\rangle \left|x_2\right\rangle \left|y + b^{x_1} a^{x_2} \pmod 2\right\rangle \\ &= \left|x_1\right\rangle \left|x_2\right\rangle \left|y \oplus b^{x_1} a^{x_2}\right\rangle \end{aligned}$$

Thus, the composition of modular exponentiation/addition of $U_b V_a U_b^\dagger$ realizes the desired black box for the quantum discrete logarithm algorithm. Since both U_b, V_a require $O(L^3)$ gates, in total the black box requires $3 \cdot O(L^3) = O(L^3)$ gates. □

Exercise 5.26

($\star$) Since K is a subgroup of G, when we decompose G into a product of cyclic groups of prime power order, this also decomposes K. Re-express (5.77) to show that determining l_i' allows one to sample from the corresponding cyclic subgroup K_{p_i} of K.

Exercise 5.27

($\star\star$) Of course, the decomposition of a general finite Abelian group G into a product of cyclic groups of prime power order is usually a difficult problem (at least as hard as factoring integers, for example). Here, quantum algorithms come to the rescue again: explain how the algorithms in this chapter can be used to efficiently decompose G as desired.

Exercise 5.28

($\star\star$) Write out a detailed specification of the quantum algorithm to solve the hidden subgroup problem, complete with runtime and success probability estimates, for finite Abelian groups.

Exercise 5.29

($\star\star$) Give quantum algorithms to solve the Deutsch and Simon problems listed in Figure 5.5, using the framework of the hidden subgroup problem.

Problem 5.1

($\star\star\star$) Construct a quantum circuit to perform the quantum Fourier transform

$$|j\rangle \mapsto \frac{1}{\sqrt{p}} \sum_{k=0}^{p-1} e^{2\pi i j k/p} |k\rangle$$

where p is prime.

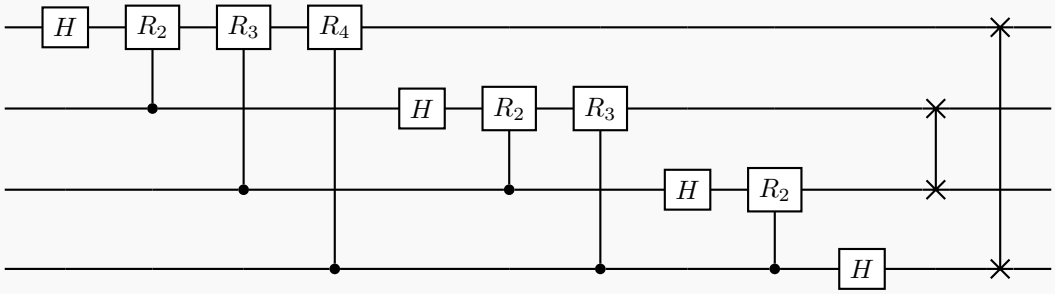
Problem 5.2: Measured quantum Fourier transform

($\star\star$) Suppose the quantum Fourier transform is performed as the last step of a quantum computation, followed by a measurement in the computational basis. Show that the combination of quantum Fourier transform and measurement is equivalent to a circuit consisting entirely of *one* qubit gates and measurement, with classical control, and no two qubit gates. You may find the discussion of Section 4.4 useful.
We find the exercise is not quite correct as stated, and instead we consider the *inverse* quantum Fourier transform.

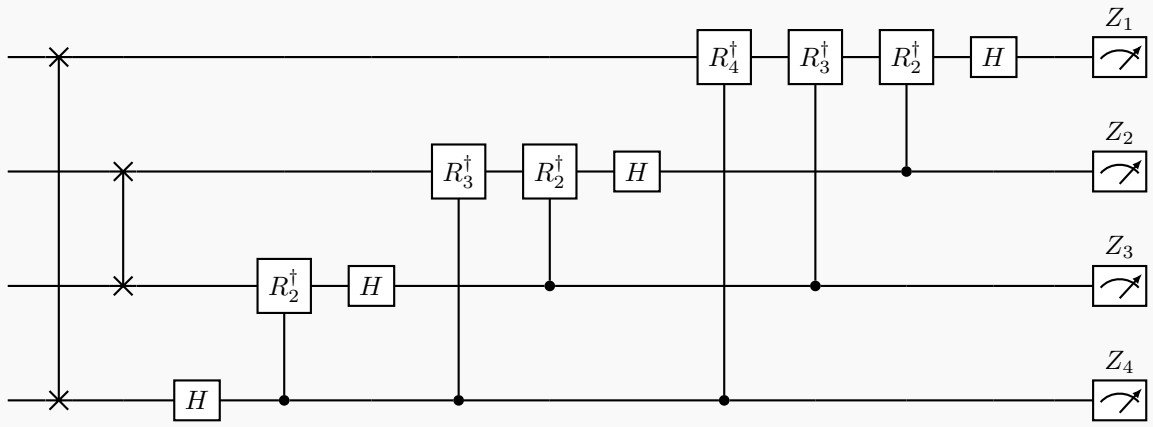
Solution

Concepts Involved: Quantum Fourier Transform, Quantum Measurement, Controlled Operations, Principle of Deferred Measurement

Throughout, we work with 4-qubit circuits for the convenience of drawing, but the argument is fully generic. Denoting $R_k = \text{diag}(1, e^{2\pi i/2^k})$, the QFT is given by (from Figure 5.1):

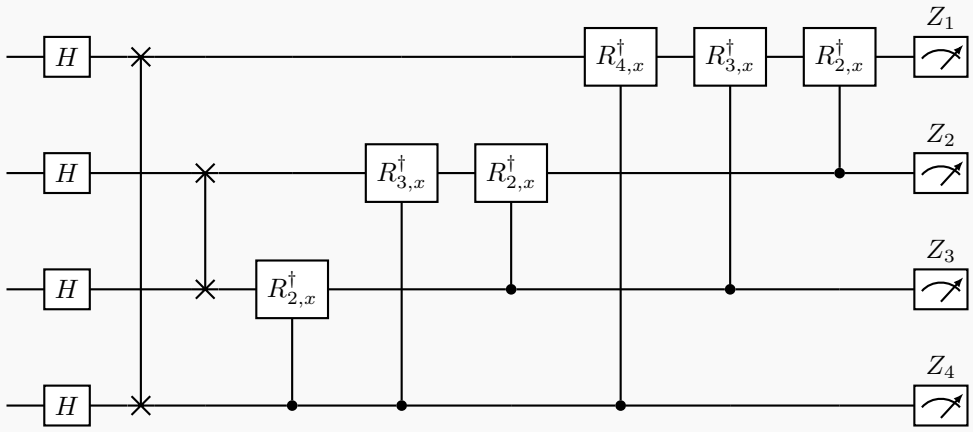


From this circuit, it becomes clear why the original formulation of the question is not quite correct. The issue is as follows - we wish to use the technique of commuting the computational basis measurement through controls (as shown in Ex. 4.35) to replace the two-qubit controlled-R_k^x gates by single qubit gates controlled by the classical measurement outcomes. However, for every qubit in the circuit appearing above, they act as controls *before* they are acted on nontrivially by other gates - in order to pullback the computational basis measurement, we would want to have the opposite, where the qubits act as controls at the end.
Thus, we believe the correct formulation of the question is to instead consider the inverse quantum Fourier transform, followed by a computational basis measurement. The circuit for $\text{QFT}^\dagger$ can be obtained by taking the conjugate of the gates appearing in the above QFT circuit and reversing the order. We consider a computational basis measurement at the end of this circuit:

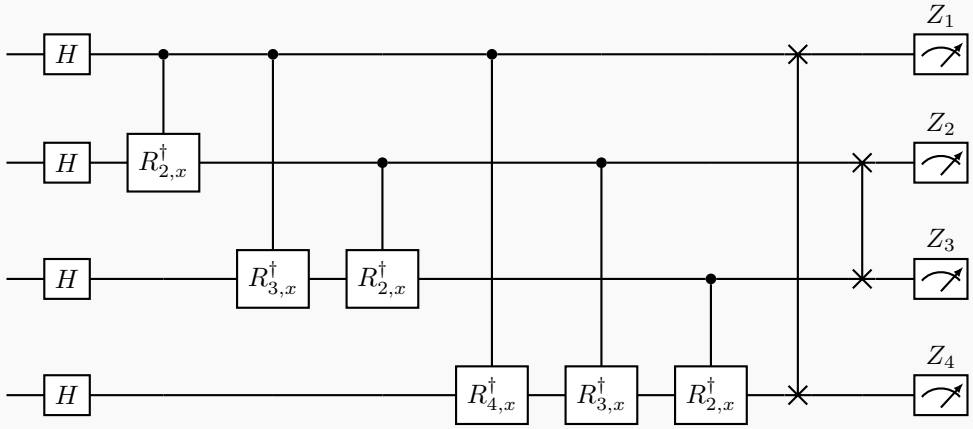


we can see that this indeed has the desired setup for commuting the measurements past the controls, as each qubit acts as a control right before the computational basis measurement. Before doing this, let's remove the SWAPs.

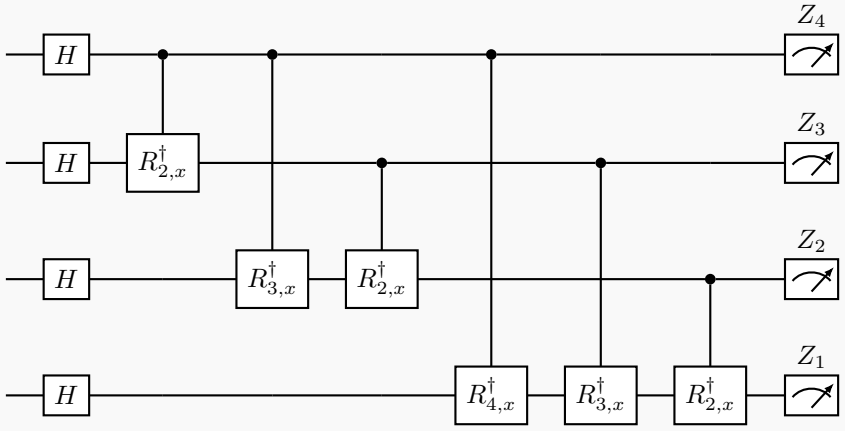
We will find it convenient to push all Hadamards to the start of the circuit - using that $H^2 = I$ and $HR_z(\theta)H = R_x(\theta)$, commuting the Hs across the z-rotations changes them to x-rotations. Further, a pair of Hadamards (moreover, a pair of any single qubit gates) commutes across a SWAP, so:



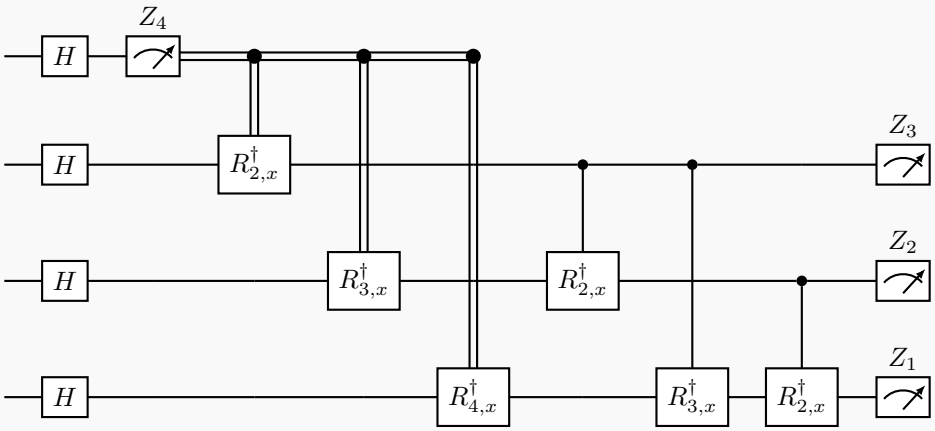
Now, let us push the SWAP gates to the very end of the circuit, which swaps the qubits on which the controlled-$R^\dagger$ gates act when we push them through:



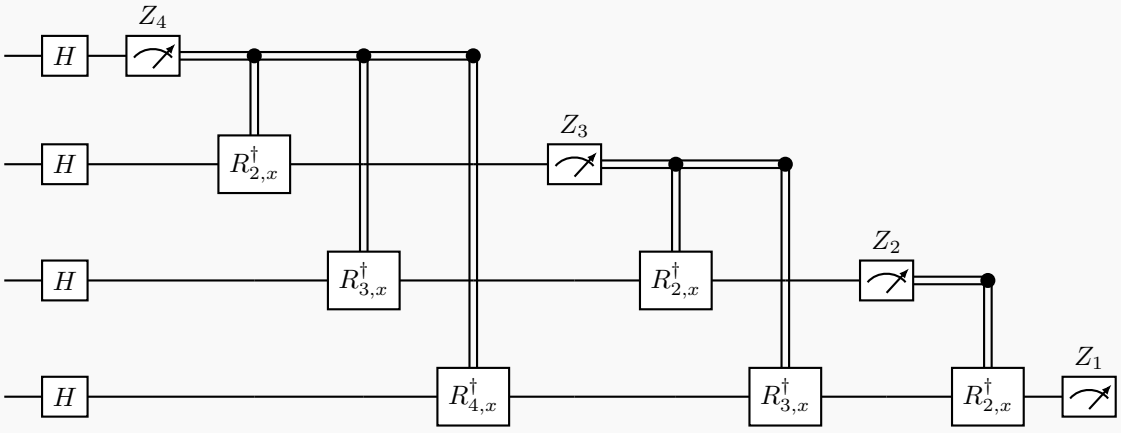
We can now push the SWAPs past the measurement, which amounts to a relabelling of the classical measurement outcomes:



Now, we can use the result of Ex. 4.35 to commute the computational basis measurements with the controls, replacing the controlled rotation gates with single qubit rotations conditioned on the (classical) measurement outcome. Starting with the top qubit:



We can do the same for the second/third qubit:

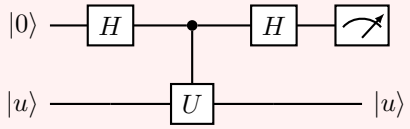


Thus we have rewritten the inverse QFT as a circuit consisting solely of one qubit gates and measurements, with classical control and no two qubit gates. For the n qubit case, this is easily seen to generalize to n sequential measurements which classically control the single-qubit rotations on the subsequent qubits. □

Remark: Note that most algorithms (such as phase estimation and order finding) use the *inverse* Fourier transform before the final computational basis measurement, so our restated result (rather than the original statement) is also the more useful one.

Problem 5.3: Kitaev's algorithm

($\star\star$) Consider the quantum circuit



where $|u\rangle$ is an eigenstate of U with eigenvalue $e^{2\pi i\varphi}$. Show that the top qubit is measured to be 0 with probability $p \equiv \cos^2(\pi\varphi)$. Since the state $|u\rangle$ is unaffected by the circuit it may be reused; if U can be replaced by U^k, where k is an arbitrary integer under your control, show that by repeating this circuit and increasing k appropriately, you can efficiently obtain as many bits of p as desired, and thus, of φ. This is an alternative to the phase estimation algorithm.

Solution

Concepts Involved: Phase Estimation, Quantum Measurement, Controlled Operations, Chernoff Bound

Analyzing the given circuit, we have:

$$\begin{aligned}
|0\rangle|u\rangle &\overset{H_1}{\mapsto} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|u\rangle \\
&\overset{C(U)}{\mapsto} \frac{1}{\sqrt{2}}(|0\rangle|u\rangle + |1\rangle U|u\rangle) \\
&= \frac{1}{\sqrt{2}}(|0\rangle|u\rangle + e^{2\pi i\varphi}|1\rangle|u\rangle) \\
&= \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i\varphi}|1\rangle)|u\rangle \\
&\overset{H_1}{\mapsto} \left(\frac{1+e^{2\pi i\varphi}}{2}|0\rangle + \frac{1-e^{2\pi i\varphi}}{2}|1\rangle\right)|u\rangle
\end{aligned}$$

Thus the top qubit is measured to be in state 0 with probability:

$$p = p_0 = \left|\frac{1+e^{2\pi i\varphi}}{2}\right|^2 = \frac{1+\cos(2\pi\varphi)}{2} = \cos^2(\pi\varphi)$$

as claimed.

We can analogously compute $p_1 = \frac{1-\cos(2\pi\varphi)}{2}$, which gives a bias $\cos(2\pi\varphi)$ - thus by repeatedly applying this circuit, we can attempt to extract this bias (and thus φ). Using the Chernoff bound (see Box 3.4), estimating the bias within error ϵ with probability at least $1-\delta$ requires $\frac{\log\left(\frac{1}{\delta}\right)}{\epsilon^2}$ samples. If our estimate is to be within m bits of accuracy, i.e. $\epsilon \sim \pm 2^{-m}$, we thus require $\exp(m)$ samples/runs of the above circuit.
If we have access to CU^k for arbitrary integer k, then such an exponential runtime in m can be avoided as follows. First, note that if $|u\rangle$ is an eigenstate of U with eigenvalue $e^{2\pi i\varphi}$ then it is an eigenstate of U^k with eigenvalue $e^{2\pi ik\varphi}$. Thus the bias of the above circuit using U^k instead of U becomes $\cos(2\pi k\varphi)$. Taking $k = 2^{n-1}$ for $n = 1, 2, 3, \ldots$ then allows us to obtain the nth decimal place of the binary expansion of φ (as each successive k shifts the decimal expansion over by one place), and thus for each k we can take $\epsilon \sim 2^{-1}$/sufficient accuracy to just extract the first decimal place of $k\varphi$. Each such run only requires $O(1)$ circuit runs, and thus $O(n)$ runs total for obtaining n bits of φ. □

Problem 5.4

($\star$) The runtime bound $O(L^3)$ we have given for the factoring algorithm is not tight. Show that a better upper bound of $O(L^2 \log L \log\log L)$ operations can be achieved.

Solution

Concepts Involved: Order Finding, Factoring, Modular Exponentiation, Continued Fractions

Analyzing the factoring algorithm (and in particular the order finding subroutine), we find that the $O(L^3)$ entered in the analysis in two places - in the modular exponentiation algorithm (see Box 5.2, where we have $O(L)$ multiplications of $O(L^2)$ cost (using elementary longhand binary multiplication)) and in the continued fractions algorithm (where we had $O(L)$ split and invert steps of $O(L^2)$ arithmetic cost).
The arithmetic cost in both of these steps can be reduced by using the Schönhage–Strassen algorithm, which is an algorithm with asymptotic cost of $O(L \log L \log\log L)$ operations/gates for the multiplication of L bit numbers, making use of recursive applications of the fast Fourier transform. A detailed analysis of this multiplication algorithm can be found, e.g. in Knuth's Art of Computer Programming Volume 2, Section 4.3.3.
Using this algorithm brings the operations required for modular exponentiation and continued fractions to $O(L^2 \log L \log\log L)$, and since these were the the most expensive aspects of the algorithm, we thus obtain the improved upper bound on the factoring algorithm to $O(L^2 \log L \log\log L)$, as claimed. □

Remark: Note that there is also the more recent work of Harvey, Hoeven, and Lecerf (arXiv:1407.3360) that gives an $O(L \log L)$ algorithm for L-bit multiplication, which brings down the bound to $O(L^2 \log L)$. This algorithm is galactic however, so in practice this is not a meaningful improvement.
We also point the reader to to the recent construction by Regev (arXiv:2308.06572), which gives an upper bound (up to logarithmic factors) of $O(L^{3/2})$ operations, though at the expense of requiring $O(L^{3/2})$ qubits (as opposed to $O(L)$).

Problem 5.5: Non-Abelian hidden subgroups (Research)

($\star\star\star$) Let f be a function on a finite group G to an arbitrary finite range X, which is promised to be constant and distinct on distinct left cosets of a subgroup K. Start with the state

$$\frac{1}{\sqrt{|G|^m}} \sum_{g_1,\ldots,g_m} |g_1,\ldots,g_m\rangle \, |f(g_1),\ldots,f(g_m)\rangle$$

and prove that picking $m = 4\log|G| + 2$ allows K to be identifies with probability at least $1 - 1/|G|$. Note that G does not necessarily have to be Abelian, and being able to perform a Fourier transform over G is not required. This result shows that one can produce (using only $O(\log|G|)$ oracle calls) a final result in which the pure state outcomes corresponding to different possible hidden subgroups are nearly orthogonal. However, it is unknown whether a POVM exists or not which allows the hidden subgroup to be identified *efficiently* (i.e. using $\mathrm{poly}(\log|G|)$ operations) from this final state.

Problem 5.6: Addition by Fourier transforms

($\star\star$) Consider the task of constructing a quantum circuit to compute $|x\rangle \mapsto |x + y \mod 2^n\rangle$, where y is a fixed constant, and $0 \leq x < 2^n$. Show that one efficient way to do this, for values of y such as 1, is to first perform a quantum Fourier transform, then to apply single qubit phase shifts, then an inverse Fourier transform. What values of y can be added easily this way, and how many operations are required?

Solution

Concepts Involved: Quantum Fourier Transform

For $x = x_1 x_2 \ldots x_n$ (in binary representation), the QFT of $|x\rangle$ is - using the product representation:

$$|x\rangle \overset{\text{QFT}}{\mapsto} \frac{(|0\rangle_1 + e^{2\pi i 0.x_n}|1\rangle_1)(|0\rangle_2 + e^{2\pi i 0.x_{n-1}x_n}|1\rangle_2)\ldots(|0\rangle_n + e^{2\pi i 0.x_1 x_2 \ldots x_n}|1\rangle_n)}{2^{n/2}}$$

To increment x_1 (which appears once, in the first place of the nth qubit) by 1 in the above expression, we apply a phase shift of $e^{i\pi}$ to the last (nth) qubit, i.e. apply $\mathrm{diag}(1, e^{i\pi})_n \cong R_z^n(\pi)$ (Z-rotation up to an unimportant global phase). To increment x_2 (which appears twice, in the first place of the $n-1$th qubit and in the second place of the nth qubit) by 1, we can apply $\mathrm{diag}(1, e^{i\pi})_{n-1} \cong R_z^{n-1}(\pi)$ to the $n-1$th qubit and $\mathrm{diag}(1, e^{i\pi/2})_n \cong R_z^n(\pi/2)$ to the nth qubit. Continuing this pattern, to increment x_k by 1, we apply the phase shifts/rotations $R_z^{n-k+1}(\pi) \cdot R_z^{n-k+1+1}(\pi/2)\ldots R_z^{n-1}(\pi/2^{k-2})R_z^n(\pi/2^{k-1})$. If we let $y = y_1 y_2 \ldots y_n$ in binary form, we can do this procedure for $y_1, \ldots y_n$ (incrementing x_k by 1 if $y_k = 1$, and doing nothing otherwise). We then obtain the state:

$$\frac{(|0\rangle_1 + e^{2\pi i 0.(x_n+y_n)}|1\rangle_1)(|0\rangle_2 + e^{2\pi i 0.(x_{n-1}+y_{n-1})(x_n+y_n)}|1\rangle_2)\ldots(|0\rangle_n + e^{2\pi i 0.(x_1+y_1)(x_2+y_2)\ldots(x_n+y_n)}|1\rangle_n)}{2^{n/2}}$$

where the organization of phases ensures that the carry bits from the additions transfer to the next significant digit in the correct manner. If we then take the inverse quantum Fourier transform, we obtain:

$$\overset{\text{QFT}^{-1}}{\mapsto} |(x_1 + y_1)(x_2 + y_2)\ldots(x_n + y_n) \mod 2^n\rangle = |x + y \mod 2^n\rangle$$

Note that any $y = y_1 y_2 \ldots y_n$ can be added via this procedure. All y are "easy" in the sense they can

be added with polynomial operations (as we discuss in the next paragraph), but adding more significant digits is easier both in the sense that less significant digits require more rotation gates (incrementing x_k requires k z-rotations) and also (exponentially) more precise rotation gates (incrementing x_k requires implementing a $R_z(\pi/2^{k-1})$ rotation).

The addition via phase shifts as described takes $O(n^2)$ operations ($O(n)$ phase shifts for up to $O(n)$ nonzero digits of y), but can be reduced to $O(n)$ by taking all layers of the phase shifts/rotations and collapsing them into a single z-rotation gate per qubit. However, both the quantum fourier transform and its inverse (see Ex. 5.5) require $\Theta(n^2)$ gates, so the total operation cost of the addition is still $\Theta(n^2)$. □

6 Quantum search algorithms

Exercise 6.1

Show that the unitary operator corresponding to the phase shift in the Grover iteration is $2\left|0\right\rangle\!\left\langle 0\right| - I$.

Solution

Concepts Involved: Unitary matrices, Grover Search

We wish to find the unitary phase shift operator U whose action on a computational basis state $\left|x\right\rangle$ is:

$$U\left|x\right\rangle = -(-1)^{\delta_{x0}}\left|x\right\rangle$$

Let's check that $2\left|0\right\rangle\!\left\langle 0\right| - I$ does the trick:

$$\left(2\left|0\right\rangle\!\left\langle 0\right| - I\right)\left|x\right\rangle = 2\left\langle 0|x\right\rangle\left|0\right\rangle - \left|x\right\rangle = 2\delta_{x0}\left|0\right\rangle - \left|x\right\rangle = \left(2\delta_{x0} - 1\right)\left|x\right\rangle = -(-1)^{\delta_{x0}}\left|x\right\rangle.$$

□

Exercise 6.2

Show that the operation $\left(2\left|\psi\right\rangle\!\left\langle\psi\right| - I\right)$ applied to a general state $\sum_k \alpha_k\left|k\right\rangle$ produces

$$\sum_k \left[-\alpha_k + 2\left\langle\alpha\right\rangle\right]\left|k\right\rangle$$

where $\left\langle\alpha\right\rangle \equiv \sum_k \alpha_k/N$ is the mean value of the α_k. For this reason, $\left(2\left|\psi\right\rangle\!\left\langle\psi\right| - I\right)$ is sometimes reffered to as the *inversion about mean* operation.

Solution

Concepts Involved: Grover Search

Appplying the operator yields:

$$\begin{aligned}
\left(2\left|\psi\right\rangle\!\left\langle\psi\right| - I\right)\sum_k \alpha_k\left|k\right\rangle &= 2\sum_k \alpha_k\left\langle\psi|k\right\rangle\left|\psi\right\rangle - \sum_k \alpha_k\left|k\right\rangle \\
&= 2\sum_k \alpha_k \frac{1}{N^{1/2}}\left|\psi\right\rangle - \sum_k \alpha_k\left|k\right\rangle \\
&= 2\left\langle\alpha\right\rangle N^{1/2}\left|\psi\right\rangle - \sum_k \alpha_k\left|k\right\rangle \\
&= 2\left\langle\alpha\right\rangle N^{1/2}\frac{1}{N^{1/2}}\sum_k\left|k\right\rangle - \sum_k \alpha_k\left|k\right\rangle \\
&= \sum_k[-\alpha_k + 2\left\langle\alpha\right\rangle]\left|k\right\rangle
\end{aligned}$$

□

Exercise 6.3

($\star$) Show that in the $|\alpha\rangle, |\beta\rangle$ basis, we may write the Grover iteration as

$$G = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

where θ is a real number in the range 0 to $\pi/2$ (assuming for simplicity that $M \leq N/2$; this limitation will be lifted shortly), chosen so that

$$\sin\theta = \frac{2\sqrt{M(N-M)}}{N}.$$

Solution

Concepts Involved: Grover Search

The action of the oracle O is to leave non-solutions invariant and attach a minus sign to the solutions, so $|\alpha\rangle \to |\alpha\rangle$ and $|\beta\rangle \to -|\beta\rangle$, hence in the $|\alpha\rangle, |\beta\rangle$ basis:

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

Now if we express:

$$|\psi\rangle = \sqrt{\frac{N-M}{N}}\,|\alpha\rangle + \sqrt{\frac{M}{N}}\,|\beta\rangle$$

We find:

$$|\psi\rangle\langle\psi| = \frac{N-M}{N}\,|\alpha\rangle\langle\alpha| + \frac{\sqrt{M(N-M)}}{N}(|\alpha\rangle\langle\beta| + |\beta\rangle\langle\alpha|) + \frac{M}{N}\,|\beta\rangle\langle\beta|$$

Hence:

$$2\,|\psi\rangle\langle\psi| - I \cong \begin{bmatrix} 2\frac{N-M}{N} - 1 & 2\frac{\sqrt{M(N-M)}}{N} \\ 2\frac{\sqrt{M(N-M)}}{N} & 2\frac{M}{N} - 1 \end{bmatrix} = \begin{bmatrix} \frac{N-2M}{N} & 2\frac{\sqrt{M(N-M)}}{N} \\ 2\frac{\sqrt{M(N-M)}}{N} & -\frac{N-2M}{N} \end{bmatrix}$$

Therefore:

$$G = (2\,|\psi\rangle\langle\psi| - I)O \cong \begin{bmatrix} \frac{N-2M}{N} & 2\frac{\sqrt{M(N-M)}}{N} \\ 2\frac{\sqrt{M(N-M)}}{N} & -\frac{N-2M}{N} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \frac{N-2M}{N} & -2\frac{\sqrt{M(N-M)}}{N} \\ 2\frac{\sqrt{M(N-M)}}{N} & \frac{N-2M}{N} \end{bmatrix}$$

Choosing θ such that:

$$\sin\theta := 2\frac{\sqrt{M(N-M)}}{N}$$

We find the cosine to be:

$$\cos\theta = \sqrt{1-\sin^2\theta} = \sqrt{\frac{N^2}{N^2} - 4\frac{M(N-M)}{N^2}} = \sqrt{\frac{(N-2M)^2}{N^2}} = \frac{N-2M}{N}$$

Therefore we can conclude:

$$G \cong \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

□

Exercise 6.4

Give explicit steps for the quantum search algorithm, as above, but for the case of multiple solutions $(1 < M < N/2)$.

Solution

Concepts Involved: Grover Search

Algorithm: Quantum Search (Multiple Solutions)

- **Inputs:** (1) A black box oracle O which performs the transformation $O|x\rangle|q\rangle = |x\rangle|q \oplus f(x)\rangle$ where $f(x) = 0$ for all $0 \leq x < 2^n$ except for $x \in S = \{x_0, x_1, \ldots x_M\}$ $(1 < M < N/2)$ for which $f(x) = 1$. (2) $n+1$ qubits in the state $|0\rangle$.
- Outputs: One solution $x_0 \in S = \{x_0, x_1, \ldots, x_M\}$
- Runtime: $O(\sqrt{2^n/M})$ operations. Succeeds with probability $O(1)$.
- Procedure:

  (1) $|0\rangle^{\otimes n}|0\rangle$ (initial state)

  (2) $\frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1}|x\rangle\left[\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right]$ (apply $H^{\otimes n}$ to first n qubits, HX to the last qubit)

  (3) $[(2|\psi\rangle\langle\psi| - 1)O]^R \frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1}|x\rangle\left[\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right] \approx |\beta\rangle\left[\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right]$ (Apply the Grover iteration $R \approx \lceil\frac{\pi}{4}\sqrt{\frac{2^n}{M}}\rceil$ times. $|\beta\rangle$ is a uniform superposition of solution states.)

  (4) $\rightarrow x_0 \in S$ (Measure the first n qubits, yielding one of the solutions.)

□

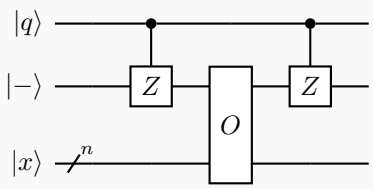
Exercise 6.5

Show that the augmented oracle O' may be constructed using one application of O, and elementary quantum gates, using the extra qubit $|q\rangle$.

Solution

Concepts Involved: Grover Search

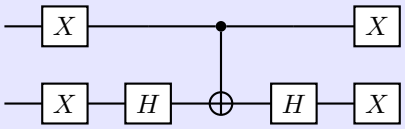
The circuit:



accomplishes the task of being the augmented oracle O'. Suppose $|q\rangle = |0\rangle$ - then the controlled-Z operations have no effect/drop out of the circuit, and the has the desired action of $O|x\rangle|-\rangle \to (-1)^{f(x)}|-\rangle$. Suppose instead that $|q\rangle = |1\rangle$. Then the controlled-Z flips $|-\rangle \to |+\rangle$, and therein the oracle has no effect $O|x\rangle|+\rangle = |x\rangle|+\rangle$ (with the $|+\rangle$ being flipped back to $|-\rangle$ by the final controlled Z). Hence, we have shown that the given O' marks an item only if x is a solution to the search problem and the extra bit q is set to zero. □

Exercise 6.6

Verify that the gates in the dotted box in the second figure of Box 6.1 perform the conditional phase shift operation $2|00\rangle\langle 00| - I$, up to an unimportant global phase factor.



Solution

Concepts Involved: Grover Search

Let us deduce the net unitary performed by the gates in the dotted box (we do the matrix multiplications in block form):

$$
\begin{aligned}
(X_1 \otimes X_2)(I_1 \otimes H_2)\mathsf{CNOT}_{1,2}(I_1 \otimes H_2)(X_1 \otimes X_2) &\cong \begin{bmatrix} 0 & X \\ X & 0 \end{bmatrix}\begin{bmatrix} H & 0 \\ 0 & H \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix}\begin{bmatrix} H & 0 \\ 0 & H \end{bmatrix}\begin{bmatrix} 0 & X \\ X & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & XH \\ XH & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & X \end{bmatrix}\begin{bmatrix} 0 & HX \\ HX & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & XH \\ XH & 0 \end{bmatrix}\begin{bmatrix} 0 & HX \\ XHX & 0 \end{bmatrix} \\
&= \begin{bmatrix} XHXHX & 0 \\ 0 & XHHX \end{bmatrix}
\end{aligned}
$$

Using that $X^2 = H^2 = I$ (Exercises **??**, **??**), the bottom right entry is:

$$XHHX = XIX = XX = I$$

And using that $HXH = Z$ (Exercise **??**) and $XZ = -ZX$ (Exercise **??**):

$$XHXHX = XZX = -ZXX = -Z$$

We thus find:

$$(X_1 \otimes X_2)(I_1 \otimes H_2)\mathsf{CNOT}_{1,2}(I_1 \otimes H_2)(X_1 \otimes X_2) \cong \begin{bmatrix} -Z & 0 \\ 0 & I \end{bmatrix} = -\mathsf{diag}(1,-1,-1,-1)$$

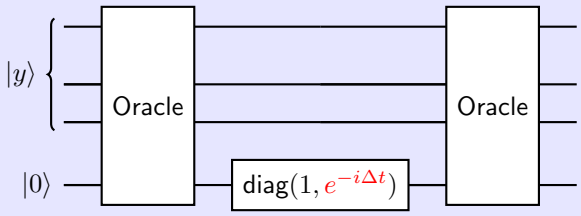
Which agrees with:

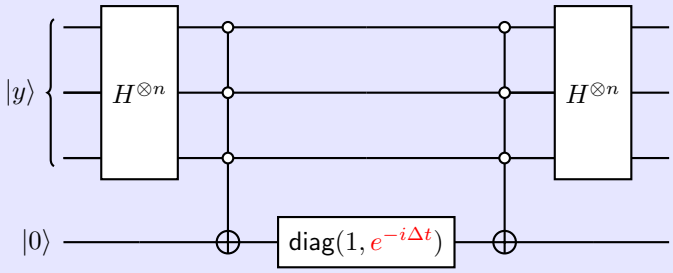
$$2\,|00\rangle\langle 00| - I \cong \mathsf{diag}(1,-1,-1,-1)$$

up to a global phase. □

Exercise 6.7

($\star$) Verify that the circuits shown in Figures 6.4 and 6.5 (reproduced below) implement the operations $\exp\big(-i\,|x\rangle\langle x|\,\Delta t\big)$ and $\exp\big(-i\,|\psi\rangle\langle\psi|\,\Delta t\big)$, respectively, with $|\psi\rangle$ as in (6.24).
The circuits appearing in the text appear to have an error in the phase acting on the oracle workspace qubit, which we correct in the diagrams below.





Solution

Concepts Involved: Grover Search, Quantum Simulation, Controlled Operations

For each of the two circuits, it suffices to verify that the circuit has the claimed action in some basis. For

Fig 6.4, let $|y\rangle$ (on qubits $1, \dots n$) be some state in the computational basis. Then we have:

$$\begin{aligned} O\begin{bmatrix}1 & 0\\ 0 & e^{-i\Delta t}\end{bmatrix}_{n+1} O\,|y\rangle_{1,\dots n}\,|0\rangle_{n+1} &= O\begin{bmatrix}1 & 0\\ 0 & e^{-i\Delta t}\end{bmatrix}_{n+1} |y\rangle_{1,\dots n}\,\big|0 \oplus f(y)\big\rangle_{n+1} \\ &= O\,|y\rangle_{1\dots n}\, e^{-i\Delta t\delta_{1f(y)}}\,\big|f(y)\big\rangle_n \\ &= e^{-i\Delta t\delta_{1f(y)}} O\,|y\rangle_{1\dots n}\,\big|f(y)\big\rangle_n \\ &= e^{-i\Delta t\delta_{1f(y)}}\,|y\rangle_{1\dots n}\,\big|f(y)\oplus f(y)\big\rangle_n \\ &= e^{-i\Delta t\delta_{1f(y)}}\,|y\rangle_{1\dots n}\,|0\rangle_n \\ &= e^{-i\Delta t\delta_{xy}}\,|y\rangle_{1\dots n}\,|0\rangle_n \end{aligned}$$

So the action on the state is $e^{-i\Delta t}$ if $x = y$, and 1 otherwise - hence the action of the overall circuit is $\exp\big(-i\Delta t\,|x\rangle\langle x|\big)$, as claimed.
For Fig 6.5, instead let $|y\rangle$ be some state in the $|\pm\rangle$/Pauli-X eigenbasis, or equivalently of states of the form $|y\rangle = H^{\otimes n}\,|z\rangle$ for computational basis states $|z\rangle$. Then we have:

$$\begin{aligned} &H^{\otimes n}\bar{C}^{1\dots n}(X_n)\begin{bmatrix}1 & 0\\ 0 & e^{-i\Delta t}\end{bmatrix}_{n+1} \bar{C}^{1\dots n}(X_n)H^{\otimes n}\,|y\rangle_{1,\dots n}\,|0\rangle_{n+1} \\ &= H^{\otimes n}\bar{C}^{1\dots n}(X_n)\begin{bmatrix}1 & 0\\ 0 & e^{-i\Delta t}\end{bmatrix}_{n+1} \bar{C}^{1\dots n}(X_n)\,|z\rangle_{1,\dots,n}\,|0\rangle_n \\ &= H^{\otimes n}\bar{C}^{1\dots n}(X_n)\begin{bmatrix}1 & 0\\ 0 & e^{-i\Delta t}\end{bmatrix}_{n+1} |z\rangle\,|\delta_{z\mathbf{0}}\rangle \\ &= H^{\otimes n}\bar{C}^{1\dots n}(X_n)\,|z\rangle\, e^{-i\Delta t\delta_{z\mathbf{0}}}\,|\delta_{z\mathbf{0}}\rangle \\ &= e^{-i\Delta t\delta_{z\mathbf{0}}} H^{\otimes n}\,|z\rangle\,|\delta_{z\mathbf{0}}\oplus\delta_{z\mathbf{0}}\rangle \\ &= e^{-i\Delta t\delta_{z\mathbf{0}}}\,|y\rangle\,|0\rangle \end{aligned}$$

So the action of the state is $e^{-i\Delta t}$ if $z = \mathbf{0}$, i.e. if $|y\rangle = H^{\otimes n}\,|0\rangle^{\otimes n} = |+\rangle^{\otimes n} = |\psi\rangle$, and 1 otherwise. Hence, the action of the overall circuit is $\exp\big(-i\Delta t\,|\psi\rangle\langle\psi|\big)$, as claimed. □

Exercise 6.8

Suppose the simulation step is performed to accuracy $O(\Delta t^r)$. Show that the number of oracle calls required to simulate H to reasonable accuracy is $O(N^{r/2(r-1)})$. Note that as r becomes large the exponent of N approaches $1/2$.

Solution

Concepts Involved: Grover Search, Accuracy

The number of total steps is:

$$O(t/\Delta t) = O(\sqrt{N}/\Delta t)$$

and hence the total error is:

$$O(\Delta t^r \cdot \sqrt{N}/\Delta t) = O(\Delta t^{r-1}\sqrt{N})$$

We want the error to be $O(1)$, and hence we choose $\Delta t = \Theta(N^{-\frac{1}{2(r-1)}})$. The total number of oracle calls is then $O(N^{\frac{r}{2(r-1)}})$, as claimed. □

Exercise 6.9

Verify Equation (6.25). (*Hint:* see Exercise 4.15.):

$$\begin{aligned} U(\Delta t) = &\left(\cos^2\left(\frac{\Delta t}{2}\right) - \sin^2\left(\frac{\Delta t}{2}\right)\hat{\boldsymbol{\psi}}\cdot\hat{\mathbf{z}}\right) I \\ &- 2i\sin\left(\frac{\Delta t}{2}\right)\left(\cos\left(\frac{\Delta t}{2}\right)\frac{\hat{\boldsymbol{\psi}}+\hat{\mathbf{z}}}{2} + \sin\left(\frac{\Delta t}{2}\right)\frac{\hat{\boldsymbol{\psi}}\times\hat{\mathbf{z}}}{2}\right)\cdot\boldsymbol{\sigma} \end{aligned}$$

Solution

Concepts Involved: Composition of single qubit operations

We have:

$$U(\Delta t) = \exp\left(-i\left|\psi\right\rangle\!\left\langle\psi\right|\Delta t\right)\exp\left(-\left|x\right\rangle\!\left\langle x\right|\Delta t\right) = \exp\left(-i\frac{I+\hat{\boldsymbol{\psi}}\cdot\boldsymbol{\sigma}}{2}\Delta t\right)\exp\left(-i\frac{I+\hat{\mathbf{z}}\cdot\boldsymbol{\sigma}}{2}\Delta t\right)$$

This is a rotation about axis $\hat{\mathbf{n}}_1 = \hat{\mathbf{z}}$ by angle $\beta_1 = \Delta t$ followed about axis $\hat{\mathbf{n}}_2 = \hat{\boldsymbol{\psi}}$ by the same angle $\beta_2 = \Delta t$, so we can apply the result of **??**(b) to write it as a single rotation through angle β_{12} about axis $\hat{\boldsymbol{n}}_{12}$ satisfying:

$$\cos\left(\frac{\beta_{12}}{2}\right) = \cos^2\left(\frac{\Delta t}{2}\right) - \sin^2\left(\frac{\Delta t}{2}\right)\hat{\mathbf{z}}\cdot\hat{\boldsymbol{\psi}}$$

$$\sin\left(\frac{\beta_{12}}{2}\right)\hat{\mathbf{n}}_{12} = \sin\left(\frac{\Delta t}{2}\right)\cos\left(\frac{\Delta t}{2}\right)(\hat{\mathbf{z}}+\hat{\boldsymbol{\psi}}) - \sin^2\left(\frac{\Delta t}{2}\right)\hat{\boldsymbol{\psi}}\times\hat{\mathbf{z}}$$

And so:

$$\begin{aligned} U(\Delta t) &= \exp\left(-i\frac{I+\hat{\boldsymbol{n}}_{12}\cdot\boldsymbol{\sigma}}{2}\beta_{12}\right) \\ &= \cos\left(\frac{\beta_{12}}{2}\right)I - i\sin\left(\frac{\beta_{12}}{2}\right)\hat{\boldsymbol{n}}_{12}\cdot\boldsymbol{\sigma} \\ &= \left(\cos^2\left(\frac{\Delta t}{2}\right) - \sin^2\left(\frac{\Delta t}{2}\right)\hat{\boldsymbol{\psi}}\cdot\hat{\mathbf{z}}\right) I \\ &\quad - 2i\sin\left(\frac{\Delta t}{2}\right)\left(\cos\left(\frac{\Delta t}{2}\right)\frac{\hat{\boldsymbol{\psi}}+\hat{\mathbf{z}}}{2} + \sin\left(\frac{\Delta t}{2}\right)\frac{\hat{\boldsymbol{\psi}}\times\hat{\mathbf{z}}}{2}\right)\cdot\boldsymbol{\sigma} \end{aligned}$$

where we have used the result of Ex. 4.5 in the second equality. □

Exercise 6.10

($\star$) Show that by choosing Δt appropriately we can obtain a quantum search algorithm which uses $O(\sqrt{N})$ queries, and for which the final state is $|x\rangle$ exactly, that is, the algorithm works with probability 1, rather than with some smaller probability.

Solution

Concepts Involved: Grover Search, Quantum Simulation

Suppose we wish to rotate from $|\psi\rangle\langle\psi|$ to $|x\rangle\langle x|$ exactly, i.e. by a total rotation angle θ_{tot}. This rotation angle can be found by taking the inner product:

$$\hat{\boldsymbol{\psi}} \cdot \hat{\mathbf{z}} = (2\alpha\beta, 0, \alpha^2 - \beta^2) \cdot (0, 0, 1) = \alpha^2 - \beta^2 = \cos(\theta_{\text{tot}})$$

and thus:

$$\theta_{\text{tot}} = \arccos\left(\alpha^2 - \beta^2\right) = \arccos\left(\left(\sqrt{\frac{1}{N}}\right)^2 - \left(\sqrt{\frac{N-1}{N}}\right)^2\right) = \arccos\left(1 - \frac{1}{N}\right) \tag{6}$$

which for large N is $\approx \pi$.
Subdivide this total rotation angle into n equal rotations so $\theta_{\text{tot}} = n\theta \implies \theta = \frac{\theta_{\text{tot}}}{n}$. Then, Eq. (6.28) reads (as the rotation accomplished by a single application of $U(\Delta t)$):

$$\cos\left(\frac{\theta_{\text{tot}}}{2n}\right) = 1 - \frac{2}{N}\sin^2\left(\frac{\Delta t}{2}\right)$$

So to get this exact rotation angle, we can rearrange the above to choose Δt to be:

$$\Delta t = 2\arcsin\left(\sqrt{\frac{N}{2}\left(1 - \cos\left(\frac{\theta_{\text{tot}}}{2n}\right)\right)}\right) \tag{7}$$

However, we must choose the subdivisions (and hence number of queries) n large enough such that Eq. (6.28) holds, which requires that:

$$\sin^2\left(\frac{\Delta t}{2}\right) = \frac{N}{2}\left(1 - \cos\left(\frac{\theta_{\text{tot}}}{2n}\right)\right) \leq 1 \tag{8}$$

Which expanding for large subdivisions n we have:

$$\frac{N}{2}\left(1 - \left(1 - \frac{1}{2}\left(\frac{\theta_{\text{tot}}}{2n}\right)^2\right)\right) \leq 1 \implies \frac{\theta_{\text{tot}}^2}{16}\frac{N}{n^2} \leq 1 \tag{9}$$

so the algorithm requires $O(\sqrt{N})$ queries. □

Exercise 6.11: Multiple solution continuous quantum search

Guess a Hamiltonian with which one may solve the continuous time search problem in the case where the search problem has M solutions.

Solution

Concepts Involved: Grover Search, Quantum Simulation

We guessed the Hamiltonian for the single solution search problem as that which rotated the equal weight superposition $|\psi\rangle = |+\rangle^{\otimes n} = \frac{\sum_x |x\rangle}{\sqrt{N}}$ into the solution $|x\rangle$, which yielded (Eqs. (6.18), (6.19)):

$$H = |x\rangle\langle x| + |\psi\rangle\langle\psi|\,, \quad H' = |x\rangle\langle\psi| + |\psi\rangle\langle x| \tag{10}$$

Thus if we have multiple (M) solutions $|y_1\rangle, \ldots |y_M\rangle$, we can rotate the equal weight superposition $|\psi\rangle$ into a vector solely made out of solution states $|\text{sol}\rangle = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} |y_i\rangle$, and candidate Hamiltonians would be:

$$H = |\text{sol}\rangle\langle\text{sol}| + |\psi\rangle\langle\psi|\,, H' = |\text{sol}\rangle\langle\psi| + |\psi\rangle\langle\text{sol}|\,. \tag{11}$$

□

Exercise 6.12: Alternative Hamiltonian for quantum search

($\star$) Suppose

$$H = |x\rangle\langle\psi| + |\psi\rangle\langle x|$$

(1) Show that it takes time $O(1)$ to rotate from the state $|\psi\rangle$ to the state $|x\rangle$, given an evolution according to the Hamiltonian H.

(2) Explain how a quantum simulation of the Hamiltonian H may be performed, and determine the number of oracle calls your simulation technique requires to obtain the solution with high probability.

Solution

Concepts Involved: Grover Search, Quantum Simulation

(1) We work in the 2D subspace spanned by the solution $|x\rangle$ and the equal weight superposition $|\psi\rangle = \frac{\sum_x |x\rangle}{\sqrt{N}} = \alpha\,|x\rangle + \beta\,|y\rangle = \frac{1}{\sqrt{N}}\,|x\rangle + \sqrt{\frac{1-N}{N}}\,|y\rangle$ with $|y\rangle$ orthogonal to $|x\rangle$. In the $|x\rangle, |y\rangle$ basis we may write the Hamiltonian as:

$$H = \begin{bmatrix} \alpha & \beta \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \alpha & 0 \\ \beta & 0 \end{bmatrix} = \begin{bmatrix} 2\alpha & \beta \\ \beta & 0 \end{bmatrix} = \alpha(I+Z) + \beta X = \alpha I + (\beta X + \alpha Z)$$

Then we observe:

$$\exp(-iHt)\,|\psi\rangle = \exp(-i\alpha t)\left[\cos(t)\,|\psi\rangle - i\sin(t)(\beta X + \alpha Z)\,|\psi\rangle\right]$$

We can neglect the global phase factor $e^{-i\alpha t}$ and further observe that $[\beta X + \alpha Z]\,|\psi\rangle = |x\rangle$, so:

$$\exp(-iHt) \cong \cos(t)\,|\psi\rangle - i\sin(t)\,|x\rangle$$

The probability of measuring $|x\rangle$ occurs as $t = \frac{\pi}{2}(n+1)$ for integer n, and hence it takes $t = \frac{\pi}{2} = O(1)$ time for time evolution by H to rotate $|\psi\rangle$ to $|x\rangle$.

(2) In the text, we simulated the time evolution of $H' = |x\rangle\langle x| + |\psi\rangle\langle\psi|$ by constructing circuits for $\exp\left(-i\,|x\rangle\langle x|\,\Delta t\right)$ and $\exp\left(-i\,|\psi\rangle\langle\psi|\,\Delta t\right)$ and alternating them in turn. This strategy does not apply for the given H because the individual terms are not Hermitian:

$$(|x\rangle\langle\psi|)^\dagger = |\psi\rangle\langle x| \neq |x\rangle\langle\psi|\,.$$

However, note that in the subspace spanned by $|x\rangle, |\psi\rangle$ that H' took the form (also expressed in the $|x\rangle, |y\rangle$ basis):

$$H' = I + \alpha(\beta X + \alpha Z) \cong \alpha(\beta X + \alpha Z)$$

from which we see that $H = \frac{H'}{\alpha}$ (up to identity terms). Thus, we see that the circuits simulating H' for time $\alpha\Delta t$ generate the desired time evolution for H for time Δt. The resource requirements to obtain the solution with high probability is thus $O(\sqrt{N})$ oracle calls, as discussed in the text.

□

Exercise 6.13

($\star$) Consider a classical algorithm for the counting problem which samples uniformly and independently k times from the search space, and let $X_1, \ldots, X_k$ be the results of the oracle calls, that is, $X_j = 1$ if the j th oracle call revealed a solution to the problem, and $X_j = 0$ if the j th oracle call did not reveal a solution to the problem. This algorithm returns the estimate $S \equiv N \times \sum_j X_j/k$ for the number of solutions to the search problem. Show that the standard deviation in S is $\Delta S = \sqrt{M(N-M)/k}$. Prove that to obtain a probability at least $3/4$ of estimating M correctly to within an accuracy $\sqrt{M}$ for all values of M we must have $k = \Omega(N)$.

Solution

Concepts Involved: Probability, Expectation, Standard Deviation

For each X_j (for $j = 1, \ldots, k$) we have that the probability that $X_j = 1$ (jth call found a solution) or $X_j = 0$ (jth call did not find a solution) given that there are M solutions out of a total search space of N, is:

$$P(X_j = 1) = \frac{M}{N}, \quad P(X_j = 0) = 1 - P(X_j = 1) = \frac{N-M}{N}$$

Thus the expectation value of a given X_j is:

$$\mathbf{E}(X_j) = 1 \cdot P(X_j = 1) + 0 \cdot P(X_j = 0) = \frac{M}{N}$$

by the linearity of expectation (Ex. A1.4), the expectation value of S is:

$$\mathbf{E}(S) = \mathbf{E}(\frac{N}{k}\sum_j X_j) = \frac{N}{k}\sum_j \mathbf{E}(X_j) = \frac{N}{k}\sum_j \frac{M}{N} = \frac{N}{k}(k \cdot \frac{M}{N}) = M.$$

Next, we calculate the variance in S. Recall that:

$$\mathsf{Var}(S) = \mathbf{E}[(S - \mathbf{E}(S))^2] = \mathbf{E}(S^2) - \mathbf{E}(S)^2$$

So we also need to calculate the expectation value of S^2:

$$S^2 = (\frac{N}{k}\sum_i X_i)(\frac{N}{k}\sum_j X_j) = \frac{N^2}{k^2}\sum_{ij} X_i X_j = \frac{N^2}{k^2}\left(\sum_{i\neq j}\sum_j X_i X_j + \sum_j X_j^2\right)$$

Note that X_i, X_j are independent if $i \neq j$, so $\mathbf{E}(X_i X_j) = \mathbf{E}(X_i)\mathbf{E}(X_j)$ for $i \neq j$ (by Ex. A1.5). If $i = j$, we instead have:

$$\mathbf{E}(X_j^2) = 1^2 \cdot P(X_j = 1) + 0^2 \cdot P(X_j = 0) = \frac{M}{N}$$

so computing $\mathbf{E}(S^2)$ via linearity:

$$\begin{aligned}
\mathbf{E}(S^2) &= \frac{N^2}{k^2}\left(\sum_{i\neq j}\sum_j \mathbf{E}(X_i X_j) + \sum_j \mathbf{E}(X_j)^2\right) \\
&= \frac{N^2}{k^2}\left(\sum_{i\neq j}\sum_j \mathbf{E}(X_i)\mathbf{E}(X_j) + \sum_j \mathbf{E}(X_j)^2\right) \\
&= \frac{N^2}{k^2}\left(\sum_{i\neq j}\sum_j \frac{M}{N}\frac{M}{N} + \sum_j \frac{M}{N}\right) \\
&= \frac{N^2}{k^2}\left((k^2-k)\frac{M}{N}\frac{M}{N} + k\frac{M}{N}\right) \\
&= (1 - \frac{1}{k})M^2 + \frac{MN}{k}
\end{aligned}$$

Thus computing the variance of S:

$$\mathsf{Var}(S) = \mathbf{E}(S^2) - \mathbf{E}(S)^2 = (1 - \frac{1}{k})M^2 + \frac{MN}{k} - M^2 = \frac{MN - M^2}{k} = \frac{M(N-M)}{k}$$

and thus we find the standard deviation:

$$\boxed{\Delta S = \sqrt{\mathsf{Var}(S)} = \sqrt{\frac{M(N-M)}{k}}}.$$

We want a probability at least 3/4 of estimating M to within accuracy $\sqrt{M}$, so:

$$P(|S-M|<\sqrt{M})\geq\frac{3}{4}\implies P(|S-M|\geq\sqrt{M})\leq\frac{1}{4}$$

To bound this, we use Chebychev's inequality (Ex. A1.6), which says for any random variable X with finite variance, and for any $\lambda>0$:

$$P(\left|X-\mathbf{E}(X)\right|\geq\lambda\Delta(X))\leq\frac{1}{\lambda^2}$$

In particular we have $\mathbf{E}(S)=M$, and for $\lambda\Delta(S)=\sqrt{M}$ we take $\lambda=\frac{\sqrt{M}}{\Delta S}$, so Chebychev gives:

$$P(|S-M|\geq\sqrt{M})\leq\frac{1}{\left(\frac{\sqrt{M}}{\Delta S}\right)^2}=\frac{\mathrm{Var}(S)}{M}=\frac{N-M}{k}.$$

We require $\frac{N-M}{k}\leq\frac{1}{4}$, and in the worst case where the LHS is largest with $M=1$, we require:

$$N-1\leq\frac{k}{4}$$

and so $k=\Omega(N)$ as claimed. □

Exercise 6.14

($\star\star$) Prove that *any* classical counting algorithm with a probability at least $3/4$ for estimating M correctly to within an accuracy $c\sqrt{M}$ for some constant c and for all values of M must make $\Omega(N)$ oracle calls.

Exercise 6.15

Use the Cauchy-Schwartz inequality to show that for any normalized state vector $|\psi\rangle$ and set of N orthonormal basis vectors $|x\rangle$,

$$\sum_x\|\psi-x\|^2\geq 2N-2\sqrt{N}.$$

Solution

Concepts Involved: Inner Products

We can write:

$$\begin{aligned}\sum_x\|\psi-x\|^2&=\sum_x\left(\langle\psi|-\langle x|\right)\left(|\psi\rangle-|x\rangle\right)\\&=\sum_x\left(\langle\psi|\psi\rangle+\langle x|x\rangle-\langle\psi|x\rangle-\langle x|\psi\rangle\right)\\&=\sum_x\left(2-2\mathrm{Re}(\langle\psi|x\rangle)\right)\end{aligned}$$

Now define $|y\rangle = \sum_x |x\rangle$ (for which $\langle y|y\rangle = N$ and take the sum over x (using the sesquilinearity of the inner product), by which:

$$\begin{aligned}\sum_x \|\psi - x\|^2 &= 2N - 2\text{Re}(\langle\psi|y\rangle) \\ &\geq 2N - 2\big|\langle\psi|y\rangle\big| \\ &\geq 2N - 2\sqrt{\langle\psi|\psi\rangle\,\langle y|y\rangle} \\ &= 2N - 2\sqrt{1\cdot N} \\ &= 2N - 2\sqrt{N}\end{aligned}$$

where the first inequality follows as $|z| > \text{Re}(z)$ and in the second we apply Cauchy-Shwartz. □

Exercise 6.16

Suppose we merely require that the probability of an error being made is less than $1/2$ when averaged uniformly over the possible values for x, instead of for all values of x. Show that $O(\sqrt{N})$ oracle calls are still required to solve the search problem.
Since we are providing a lower bound, we show that $\Omega(\sqrt{N})$ calls are required.

Solution

Concepts Involved: Grover Search, Cauchy-Shwartz

The first part of the optimality proof of showing $D_k \leq 4k^2$ proceeds identically as it appears in the text. We go through the $D_k = \Omega(N)$ of the proof with the modified assumption; instead of requiring $\left|\langle x|\psi_k^x\rangle\right|^2 \geq 1/2$ for all x, we instead require $\frac{1}{N}\sum_x \left|\langle x|\psi_k^x\rangle\right|^2 \geq 1/2$ or:

$$\sum_x \left|\langle x|\psi_k^x\rangle\right|^2 \geq \frac{N}{2}$$

With this, we can bound E_k as is done in the text:

$$\begin{aligned}E_k = \sum_x \|\psi_k^x - x\|^2 &= \sum_x \left[2 - \langle x|\psi_k^x\rangle - \langle\psi_k^x|x\rangle\right] \\ &= 2N - 2\sum_x \left|\langle x|\psi_k^x\rangle\right| \\ &\leq 2N - 2\sum_x \left|\langle x|\psi_k^x\rangle\right|^2 \\ &\leq 2N - 2\frac{N}{2} \\ &= N\end{aligned}$$

where in the second line we use that WLOG $\langle x|\psi_k^x\rangle = \left|\langle x|\psi_k^x\rangle\right|$, in the third line we use that $\left|\langle x|\psi_k^x\rangle\right| \leq 1$ so $\left|\langle x|\psi_k^x\rangle\right|^2 \leq \left|\langle x|\psi_k^x\rangle\right|$, and in the fourth line we use the bound on the error probability.

Now, we can go through the identical manipulations as is done in the text to obtain Eq. (6.51):

$$D_k \geq (\sqrt{F_k} - \sqrt{E_k})^2$$

with $F_k = \sum_x \|x - \psi_k\|^2 \geq 2N - 2\sqrt{N}$ from the previous exercise, and so:

$$D_k \geq (\sqrt{2N - 2\sqrt{N}} - \sqrt{N})^2 \geq (\sqrt{N})^2 = N$$

Combining this with $D_k \leq 4k^2$, we find that $k \geq \sqrt{N/4}$ and we thus require $\Omega(\sqrt{N})$ oracle calls. □

Exercise 6.17: Optimality for multiple solutions

($\star\star$) Suppose the search problem has M solutions. Show that $O(\sqrt{N/M})$ oracle applications are required to find a solution.
Since we are providing a lower bound, we show that $\Omega(\sqrt{N/M})$ calls are required.

Solution

Concepts Involved: Grover Search, Cauchy-Shwartz

The solution follows the structure of the argument of the single-solution case. We adapt the helpful presentation of cppascalinux' blog post.
Assumption on M. WLOG, let us suppose that $M \leq \frac{N}{4}$ - else if $M > \frac{N}{4}$, then $\frac{N}{M} < \frac{N}{N/4} = 4$. Trivially, $\Omega(1)$ (a constant number of) oracle calls are required to find a solution for any search algorithm, which automatically implies $\Omega(\sqrt{N/M})$ calls. Hence showing an $\Omega(\sqrt{N/M})$ lower bound is not interesting in this case.
Oracle for multiple solutions. In particular, denote by $\mathcal{M}$ a set of M solutions (i.e. if $x \in \mathcal{M}$ then $f(x) = 1$, and $f(x) = 0$ otherwise). Then, consider the oracle:

$$O_{\mathcal{M}} = I - 2 \sum_{x \in \mathcal{M}} |x\rangle\langle x| .$$

Suppose our algorithm starts in the state $|\psi\rangle$, and applies $O_{\mathcal{M}}$ exactly k times, interleaved with unitaries $U_1, \ldots, U_k$. We can then define:

$$|\psi_k^{\mathcal{M}}\rangle = U_k O_{\mathcal{M}} U_{k-1} U_{k-2} O_{\mathcal{M}} \ldots U_1 O_{\mathcal{M}} |\psi\rangle$$
$$|\psi_k\rangle = U_k U_{k-1} U_{k-2} \ldots U_1 |\psi\rangle$$

our goal is to bound the deviation between these two (resulting from the oracle):

$$D_k = \sum_{\mathcal{M}} \left\| \psi_k^{\mathcal{M}} - \psi_k \right\|^2$$

note that the sum here is over all M-sized subsets of the solution space $S = \{1, 2, \ldots, N\}$ (i.e. over all possible solution sets of this size).
Upper bound. We give an inductive proof that $D_k \leq 4\binom{N-1}{M-1}k^2$. The $k = 0$ case is automatically satisfied,

as $\left|\psi_0^{\mathcal{M}}\right\rangle = |\psi_0\rangle$ so $D_0 = 0$. Now, suppose the statement holds for $k \geq 0$, and consider:

$$\begin{aligned} D_{k+1} &= \sum_{\mathcal{M}} \left\| O_{\mathcal{M}} \psi_k^{\mathcal{M}} - \psi_k \right\|^2 \\ &= \sum_{\mathcal{M}} \left\| O_{\mathcal{M}} (\psi_k^{\mathcal{M}} - \psi_k) + (O_{\mathcal{M}} - I)\psi_k \right\|^2 \end{aligned}$$

For the first term:

$$\sum_{\mathcal{M}} \left\| O_{\mathcal{M}} (\psi_k^{\mathcal{M}} - \psi_k) \right\|^2 = \sum_{\mathcal{M}} \left\| \psi_k^{\mathcal{M}} - \psi_k \right\|^2 = D_k \leq 4 \binom{N-1}{M-1} k^2$$

where the first equality follows by unitary invariance of the norm, and the inequality follows by the inductive hypothesis.
Now, looking at the second term:

$$\begin{aligned} \sum_{\mathcal{M}} \left\| (O_{\mathcal{M}} - I)\psi_k \right\|^2 &= \sum_{\mathcal{M}} \left\| (-2 \sum_{x \in \mathcal{M}} |x\rangle\langle x|)\psi_k \right\|^2 \\ &= \sum_{\mathcal{M}} \left\| \sum_{x \in \mathcal{M}} \langle x|\psi_k\rangle \, |x\rangle \right\|^2 \\ &= 4 \sum_{\mathcal{M}} \sum_{x \in \mathcal{M}} \left| \langle \psi_k | x \rangle \right|^2 \end{aligned}$$

Now, consider how many times a given $x \in S = \{1, 2, \ldots, N\}$ appears in the above sum. If we consider solution sets of size $|\mathcal{M}| = M$, fixing a given x to be in the solution set, there are $N-1$ choices to fill the remaining $M-1$ slots, and hence $\binom{N-1}{M-1}$ occurences of a particular x in the double sum $\sum_{\mathcal{M}} \sum_{x \in \mathcal{M}}$ Thus:

$$\sum_{\mathcal{M}} \left\| (O_{\mathcal{M}} - I)\psi_k \right\|^2 = 4 \binom{N-1}{M-1} \sum_{x \in S} \left| \langle \psi_k | x \rangle \right|^2 = 4 \binom{N-1}{M-1}$$

where the sum is now over all squared coefficients of $|\psi_k\rangle$, which is one by normalization.
A brief lemma before proceeding - consider for vectors b_i, c_i (with $i = 1, \ldots, n$) that:

$$\begin{aligned} \sum_i \|b_i + c_i\|^2 &\leq \sum_i (\|b_i\|^2 + 2\|b_i\|\|c_i\| + \|c_i\|^2) \\ &\leq \sum_i \|b_i\|^2 + \sum_i \|c_i\|^2 + 2\sqrt{\sum_i \|b_i\|^2} \sqrt{\sum_i \|c_i\|^2} \\ &= \left(\sqrt{\sum_i \|b_i\|^2} + \sqrt{\sum_i \|c_i\|^2} \right)^2 \end{aligned}$$

where the first inequality is the triangle inequality, and the second inequality follows from Cauchy-Schwartz applied to the vectors $|B\rangle = (\|b_1\|, \|b_2\|, \ldots, \|b_n\|)^T, |C\rangle = (\|c_1\|, \|c_2\|, \ldots, \|c_n\|)^T$. Applying this lemma

with $b_i = O_{\mathcal{M}}(\psi_k^{\mathcal{M}} - \psi_k), c_i = (O_{\mathcal{M}} - I)\psi_k$, we obtain:

$$\begin{aligned} D_{k+1} &\leq \left(\sqrt{\sum_{\mathcal{M}} \left\| (O_{\mathcal{M}} - I)\psi_k \right\|^2} + \sqrt{\sum_{\mathcal{M}} \left\| (O_{\mathcal{M}} - I)\psi_k \right\|^2} \right)^2 \\ &\leq \left(4\sqrt{\binom{N-1}{M-1} k^2} + \sqrt{4\binom{N-1}{M-1}} \right)^2 \\ &\leq \left(2(k+1)\sqrt{\binom{N-1}{M-1}} \right)^2 \\ &= 4(k+1)^2 \binom{N-1}{M-1} \end{aligned}$$

which completes the induction.

Lower bound. Define $P_{\mathcal{M}} = \sum_{x \in \mathcal{M}} |x\rangle\langle x|$ to be the projection onto the solution space, and then let us write:

$$D_k = \sum_{\mathcal{M}} \left\| \psi_k^{\mathcal{M}} - \psi_k \right\|^2 = \sum_{\mathcal{M}} \left\| (I - P_{\mathcal{M}})\psi_k^{\mathcal{M}} + (P_{\mathcal{M}}\psi_k^{\mathcal{M}} - \psi_k) \right\|^2$$

We suppose $\langle \psi_k^{\mathcal{M}} | P_{\mathcal{M}} | \psi_k^{\mathcal{M}} \rangle \geq 1/2$ for all $\mathcal{M}$, such that an observation yields a solution to the search problem with probability at least one half. Then:

$$\begin{aligned} \left\| (I - P_{\mathcal{M}})\psi_k^{\mathcal{M}} \right\|^2 &= \langle \psi_k^{\mathcal{M}} | (I - P_{\mathcal{M}})^\dagger (I - P_{\mathcal{M}}) | \psi_k^{\mathcal{M}} \rangle \\ &= \langle \psi_k^{\mathcal{M}} | (I - 2P_{\mathcal{M}} + P_{\mathcal{M}}^2) | \psi_k^{\mathcal{M}} \rangle \\ &= \langle \psi_k^{\mathcal{M}} | (I - P_{\mathcal{M}}) | \psi_k^{\mathcal{M}} \rangle \\ &= \langle \psi_k^{\mathcal{M}} | \psi_k^{\mathcal{M}} \rangle - \langle \psi_k^{\mathcal{M}} | P_{\mathcal{M}} | \psi_k^{\mathcal{M}} \rangle \\ &\leq 1 - \frac{1}{2} \\ &= \frac{1}{2} \end{aligned}$$

where we note the hermicity and idempotency of the projector $P_{\mathcal{M}}$ in the second/third lines. Summing over all possible $\binom{N}{M}$ solution sets $\mathcal{M}$, we can bound the first term:

$$\sum_{\mathcal{M}} \left\| (I - P_{\mathcal{M}})\psi_k^{\mathcal{M}} \right\|^2 \leq \frac{1}{2}\binom{N}{M}.$$

We can also find:

$$\begin{aligned}\left\|(P_{\mathcal{M}}\psi_k^{\mathcal{M}}-\psi_k)\right\|^2 &= \langle\psi_k^{\mathcal{M}}|\,P_{\mathcal{M}}^\dagger P_{\mathcal{M}}\,|\psi_k^{\mathcal{M}}\rangle - \langle\psi_k^{\mathcal{M}}|\,P_{\mathcal{M}}^\dagger\,|\psi_k\rangle - \langle\psi_k|\,P_{\mathcal{M}}\,|\psi_k^{\mathcal{M}}\rangle + \langle\psi_k|\psi_k\rangle \\ &= \langle\psi_k^{\mathcal{M}}|\,P_{\mathcal{M}}\,|\psi_k^{\mathcal{M}}\rangle - 2\,\mathrm{Re}(\,\langle\psi_k^{\mathcal{M}}|\,P_{\mathcal{M}}^\dagger P_{\mathcal{M}}\,|\psi_k\rangle) + 1 \\ &\geq \frac{1}{2} - 2\|P_{\mathcal{M}}\psi_k\|^2 + 1 \\ &= \frac{3}{2} - 2\,\langle\psi_k|\,P_{\mathcal{M}}^\dagger P_{\mathcal{M}}\,|\psi_k\rangle \\ &= \frac{3}{2} - 2\,\langle\psi_k|\,P_{\mathcal{M}}\,|\psi_k\rangle\end{aligned}$$

where we use the $\langle\psi_k^{\mathcal{M}}|\,P_{\mathcal{M}}\,|\psi_k^{\mathcal{M}}\rangle \geq 1/2$ assumption in the inequality.
Summing over $\mathcal{M}$, we can use this to bound the second term:

$$\begin{aligned}\sum_{\mathcal{M}}\left\|(P_{\mathcal{M}}\psi_k^{\mathcal{M}}-\psi_k)\right\|^2 &\geq \sum_{\mathcal{M}}\left(\frac{3}{2} - 2\,\langle\psi_k|\,P_{\mathcal{M}}\,|\psi_k\rangle\right) \\ &= \frac{3}{2}\binom{N}{M} - 2\sum_{\mathcal{M}}\langle\psi_k|\,P_{\mathcal{M}}\,|\psi_k\rangle\end{aligned}$$

in our argument for the upper bound, we already saw that the second term evaluates to $\binom{N-1}{M-1}$, and combining this with the binomial coefficient identity of $\binom{N-1}{M-1} = \frac{M}{N}\binom{N}{M}$, we get:

$$\sum_{\mathcal{M}}\left\|(P_{\mathcal{M}}\psi_k^{\mathcal{M}}-\psi_k)\right\|^2 \geq \left(\frac{3}{2} - 2\frac{M}{N}\right)\binom{N}{M}$$

Now since we assume that $M \leq \frac{N}{4}$, $\frac{M}{N} \leq \frac{1}{4}$ and so:

$$\sum_{\mathcal{M}}\left\|(P_{\mathcal{M}}\psi_k^{\mathcal{M}}-\psi_k)\right\|^2 \geq (\frac{3}{2} - 2\frac{1}{4})\binom{N}{M} = \binom{N}{M}.$$

We can combine these two bounds by considering the following inequality, which is analogous to one we showed in the upper bound case:

$$\begin{aligned}\sum_i \|b_i + c_i\|^2 &\geq \sum_i (\|b_i\|^2 - 2\|b_i\|\|c_i\| + \|c_i\|^2) \\ &\geq \sum_i \|b_i\|^2 + \sum_i \|c_i\|^2 - 2\sqrt{\sum_i \|b_i\|^2}\sqrt{\sum_i \|c_i\|^2} \\ &= \left(\sqrt{\sum_i \|b_i\|^2} - \sqrt{\sum_i \|c_i\|^2}\right)^2\end{aligned}$$

where the first inequality follows from the (reverse) triangle inequality and the second inequality from Cauchy-Shwartz. Setting $c_i = (I - P_{\mathcal{M}})\psi_k^{\mathcal{M}}, b_i = (P_{\mathcal{M}}\psi_k^{\mathcal{M}} - \psi_k)$ and using the bounds we showed for

the two terms, we get:

$$\begin{aligned} D_k &\geq \left(\sqrt{\sum_{\mathcal{M}} \left\| (P_{\mathcal{M}} \psi_k^{\mathcal{M}} - \psi_k) \right\|^2} - \sqrt{\sum_{\mathcal{M}} \left\| (I - P_{\mathcal{M}}) \psi_k^{\mathcal{M}} \right\|^2} \right)^2 \\ &\geq \left(\sqrt{\binom{N}{M}} - \sqrt{\frac{1}{2}\binom{N}{M}} \right)^2 \\ &= \left(1 - \frac{1}{\sqrt{2}}\right)^2 \binom{N}{M} \\ &= \left(\frac{3}{2} - \sqrt{2}\right) \binom{N}{M} \end{aligned}$$

Bound on k. Combining the upper and lower bounds on D_k:

$$\left(\frac{3}{2} - \sqrt{2}\right) \binom{N}{M} \leq D_k \leq 4k^2 \binom{N-1}{M-1} \implies \frac{\left(\frac{3}{2} - \sqrt{2}\right) \binom{N}{M}}{4\binom{N-1}{M-1}} \leq k^2 \implies \sqrt{\frac{\left(\frac{3}{2} - \sqrt{2}\right)}{4} \frac{N}{M}} \leq k$$

where we use that $\binom{N}{M} / \binom{N-1}{M-1} = \frac{N}{M}$. We conclude that:

$$k = \Omega(\frac{N}{M}).$$

□

Exercise 6.18

($\star$) Prove that the minimum degree polynomial representing a Boolean function $F(X)$ is unique.

Solution

Concepts Involved: Boolean Functions, Method of Polynomials

Suppose for the sake of contradiction that p, p' are distinct minimum degree (multilinear) polynomials representing F. Then, $(p - p')(X) = F(X) - F'(X) = 0$ for all $X \in \{0,1\}^N$, but $p - p' \neq 0$, so there exist some monomials with nonzero coefficients in $p - p'$. In particular, let $\prod_{k \in K} X_k$ be the minimum degree such monomial (with $K \subseteq \{1, \ldots N\}$ being the index set). Then, consider $Y \in \{0,1\}^N$, defined such that $Y_k = 1$ if $k \in K$ and $Y_k = 0$ otherwise. Then, $(p - p')(Y)$ does not vanish (the $\prod_{k \in K} X_k$ term is nonzero, and any potentially larger-degree terms are vanish since we chose the minimal index set for which Y is one) - contradiction. □

Exercise 6.19

Show that $P(X) = 1 - (1 - X_0)(1 - X_1) \ldots (1 - X_{N-1})$ represents OR.

Solution

Concepts Involved: Boolean Functions, Method of Polynomials

We can see that $P(X) = 0$ if $X_1 = X_2 = \ldots = X_{N-1} = 0$ and $P(X) = 1$ otherwise (if any of the $X_i = 1$). This is exactly the definition of the OR function. □

Exercise 6.20

($\star\star$) Show that $Q_0(\text{OR}) \geq N$ by constructing a polynomial which represents the OR function from the output of a quantum circuit which computes OR with zero error.
We found this problem statement to be quite misleading - to issue a clarification, (we believe) the intention of the problem is *not* to explicitly construct a quantum circuit which computes OR with zero error (such a realization would of course only provide an upper bound on $Q_0(\text{OR})$). Instead, the strategy is to posit (without explicit construction) such a circuit, and then lower bound the number of queries it must make by N.

Solution

Concepts Involved: Method of Polynomials, Quantum Query Complexity

We follow the construction of Beals, Buhrman, Cleve, Mosca, and de Wolf of arXiv:quant-ph/9802049. Consider a quantum circuit that computes OR with zero-error using $T = Q_0(\text{OR})$ queries. As discussed in this section, the output of such a circuit looks like:

$$|\varphi^X\rangle = \sum_{k=0}^{2^n-1} c_k(X)\,|k\rangle$$

where c_k are polynomials of degree at most T in the variables $X_0, X_1, \ldots, X_{N-1}$.
Let B the set of all computational basis states ending in 10, such that the output is the answer 0. Then, for $k \in B$ we have $c_k(X) = 0$ if $X \neq \mathbf{0}$, else the probability of getting the incorrect answer on $|\varphi^X\rangle$ is nonzero, contradicting the zero-error assumption. Additionally, there exists some $k' \in B$ such that $c_{k'}(\mathbf{0}) \neq 0$, since the probability of getting the correct answer for $|\varphi^{X=\mathbf{0}}\rangle$ must also be nonzero.
Now consider:

$$p(X) = \text{Re}(1 - \frac{c_{k'}(X)}{c_{k'}(\mathbf{0})}).$$

This polynomial has degree at most T (since $c_{k'}$ has degree at most T), and represents OR (as by construction it evaluates to 0 if $X = \mathbf{0}$ and 1 otherwise). But p must have degree at least $\deg(\text{OR}) = N$, from which we conclude that $T = Q_0(\text{OR}) \geq N$, as claimed. □

Problem 6.1: Finding the minimum

($\star\star\star$) Suppose $x_1, \ldots x_N$ is a database of numbers held in memory, as in Section 6.5. Show that only $O(\log(N)\sqrt{N})$ accesses to the memory are required on a quantum computer, in order to find the smallest element on the list, with probability at least one-half.

Problem 6.2: Generalized quantum searching

($\star\star\star$) Let $|\psi\rangle$ be a quantum state, and define $U_{|\psi\rangle} = I - 2|\psi\rangle\langle\psi|$. That is $U_{|\psi\rangle}$ gives the state $|\psi\rangle$ a -1 phase, and leaves states orthogonal to $|\psi\rangle$ invariant.

(1) Suppose we have a quantum circuit implementing a unitary operator U such that $U|0\rangle^{\otimes n} = |\psi\rangle$. Explain how to implement $U_{|\psi\rangle}$.

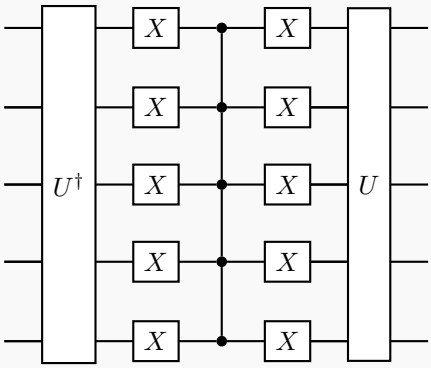
(2) Let $|\psi_1\rangle = |1\rangle$, $|\psi_2\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$, $|\psi_3\rangle = (|0\rangle - i|1\rangle)/\sqrt{2}$. Suppose an unknown oracle O is selected from the set $U_{|\psi_1\rangle}, U_{|\psi_2\rangle}, U_{|\psi_3\rangle}$. Give a quantum algorithm which identifies the oracle with just *one* application of the oracle. (*Hint:* consider superdense coding.)

(3) **(Research):** More generally, given k states $|\psi_1\rangle, \ldots, |\psi_k\rangle$, and an unknown oracle O selected from the set $U_{|\psi_1\rangle}, \ldots, U_{|\psi_k\rangle}$, how many oracle applications are required to identify the oracle, with high probability?

Solution

Concepts Involved: Grover Search, Superdense Coding, Bell Basis

(1) First, note that if we are given a quantum circuit implementing U, we can obtain $U^\dagger$ by individually taking the adjoint of each gate. With this in mind, we can realize $U_{|\psi\rangle}$ using the circuit below (drawn for 5 qubits, but the structure is identical for arbitrary n):



Let's verify that this circuit has the claimed action. If we input $|\psi\rangle$, we obtain:

$$\begin{aligned}
(UX^{\otimes n}C^n(Z)X^{\otimes n}U^\dagger)|\psi\rangle &= (UX^{\otimes n}C^n(Z)X^{\otimes n}U^\dagger)U|0\rangle^{\otimes n} \\
&= UX^{\otimes n}C^n(Z)X^{\otimes n}|0\rangle^{\otimes n} \\
&= UX^{\otimes n}C^n(Z)|1\rangle^{\otimes n} \\
&= UX^{\otimes n}(-1)|1\rangle^{\otimes n} \\
&= (-1)U|0\rangle^{\otimes n} \\
&= (-1)|\psi\rangle
\end{aligned}$$

Suppose we input any state $|\psi^\perp\rangle$ orthogonal to $|\psi\rangle$, - then, since unitaries preserve orthogonality, $U^\dagger$ maps $|\psi^\perp\rangle$ to a state orthogonal to $|0\rangle^{\otimes n}$, i.e. a superposition over bitstrings not containing

$|0\rangle^{\otimes n}$, so:

$$\begin{aligned}(UX^{\otimes n}C^n(Z)X^{\otimes n}U^\dagger)\,|\psi^\perp\rangle &= UX^{\otimes n}C^n(Z)X^{\otimes n}(\sum_{\mathbf{x}\neq\mathbf{0}} c_{\mathbf{x}}\,|\mathbf{x}\rangle)\\ &= UX^{\otimes n}C^n(Z)\sum_{\mathbf{x}\neq\mathbf{0}} c_{\mathbf{x}}\,|\bar{\mathbf{x}}\rangle\\ &= UX^{\otimes n}\sum_{\mathbf{x}\neq\mathbf{0}} c_{\mathbf{x}}\,|\bar{\mathbf{x}}\rangle\\ &= U\sum_{\mathbf{x}\neq\mathbf{0}} c_{\mathbf{x}}\,|\mathbf{x}\rangle\\ &= |\psi^\perp\rangle\end{aligned}$$

where $\bar{\mathbf{x}}$ denotes the complement bitstring of $\mathbf{x}$. Notably, $C^n(Z)$ acts trivially here as the sum does not contain the all-1 state.

Thus, this circuit gives $|\psi\rangle$ a -1 phase and leaves states orthogonal to $|\psi\rangle$ invariant, and we conclude it represents $U_{|\psi\rangle} = 1 - 2\,|\psi\rangle\langle\psi|$.

(2) Actually, we can do even better than the question proposes, and add $U_0 = I$ to the set. The algorithm ORACLEIDENTIFY(O) is:

ORACLEIDENTIFY(O):

Initialize $|\beta_{00}\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$
Apply O on the first qubit
Measure in the Bell basis, and obtain $|\beta_{xz}\rangle$.
Output $n = xz$.

Where we recall one way of representing the Bell basis is given by $|\beta_{xz}\rangle = X_1^x Z_1^z\,|\beta_{00}\rangle$. This works because $U_0 = I, U_1 = I - 2\,|1\rangle\langle 1| = Z, U_2 = I - 2\,|-\rangle\langle -| = X, U_3 = I - 2\,|-i\rangle\langle -i| = Y$, and so applying the oracle O_n to the first qubit of $|\beta_{00}\rangle$ simply maps us to $|\beta_{n=xz}\rangle$. Which Bell state we have (and thus which Oracle we applied) can then be deduced with certainty from a Bell basis measurement, in a way that only requires a single application of the oracle. The above algorithm is of course just the superdense coding algorithm reframed as an oracle identification problem.

(3) The problem is quite open ended, and we are unaware if there exists a useful general result. However, we review some results in the spirit of the problem:

- The query complexity for *Boolean* oracles was introduced/studied by Ambainis/Iwama/Kawachi/Masuda/Putra/Yamashita in arXiv:quant-ph/0403056, subsequently refined by Ambainis/Iwama/Kawachi/Raymond/Yamashita in arXiv:quant-ph/0411204, and is fully resolved by Kathari in arXiv:1311.7685. Given the task of identifying an oracle of an N-bit string, selected from a subset of size M out of the 2^N possible, the (optimal) quantum query complexity is $\Theta(\sqrt{M})$ for $M \leq N$ and $\Theta(\frac{N\log M}{\log(N/\log M)+1})$ for $N < M \leq 2^N$ (compared to the optimal classical query complexity of $\Theta(\min\{N, M\})$.

- For the unitary case, Copeland and Pommersheim arXiv:1812.09428 have some results in the case where the unitaries form a group structure, e.g. showing that $\Omega(n)$ queries are required to identify a given (permutation) oracle in S_n. Huang and Li place a bound on the query complexity in the case of distinguishing two unitary oracles arXiv:2012.02944.

□

Problem 6.3: Database retrieval

($\star\star\star$) Given a quantum oracle that returns $|k, y \oplus X_k\rangle$ given an n qubit query (and one scratchpad qubit) $|k, y\rangle$, show that with high probability, all $N = 2^n$ bits of X can be obtained using only $N/2 + \sqrt{N}$ queries. This implies the general upper bound $Q_2(F) \leq N/2 + \sqrt{N}$ for any F.

Problem 6.4: Quantum searching and cryptography

($\star\star$) Quantum searching can, potentially, be used to speed up the search for cryptographic keys. The idea is to search through the space of all possible keys for decryption, in each case trying the key, and checking to see whether the decrypted message makes 'sense'. Explain why this idea doesn't work for the Vernam cipher (Section 12.6). When might it work for cryptosystems such as DES? (For a description of DES see, for example, [MvOV96] or [Sch96a].)

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 $|\alpha\rangle$:

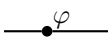
$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle$$

Applying a to $|\alpha\rangle$ and using that $a\,|n\rangle = \sqrt{n}\,|n\rangle$, we have:

$$\begin{aligned}a\,|\alpha\rangle &= a e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \\ &= e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} a\,|n\rangle \\ &= e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} \sqrt{n}\,|n-1\rangle \\ &= e^{-|\alpha|^2/2} \alpha \sum_{n=0}^{\infty} \frac{\alpha^{n-1}}{\sqrt{(n-1)!}} |n-1\rangle \\ &= \alpha e^{-|\alpha|^2/2} \sum_{n'=0}^{\infty} \frac{\alpha^{n'}}{\sqrt{(n')!}} |n'\rangle \\ &= \alpha\,|\alpha\rangle\end{aligned}$$

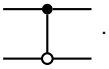
where in the second-to-last inequality we re-index the sum. We conclude that $|\alpha\rangle$ is an eigenstate of a with eigenvalue α. □

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:

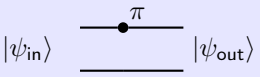
 .

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

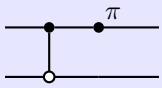
First, we note that $P = e^{i\Delta a^\dagger a}$ and so $P|n\rangle = e^{i\Delta n}$. Therefore:

$$\begin{aligned}
P|\alpha\rangle &= Pe^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle \\
&= e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} P|n\rangle \\
&= e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} e^{i\Delta n} |n\rangle \\
&= e^{-\frac{|\alpha|^2}{2}} \sum_{n=0}^{\infty} \frac{(\alpha e^{i\Delta})^n}{\sqrt{n!}} |n\rangle \\
&= e^{-\frac{\left|\alpha e^{i\alpha}\right|^2}{2}} \sum_{n=0}^{\infty} \frac{(\alpha e^{i\Delta})^n}{\sqrt{n!}} |n\rangle \\
&= |\alpha e^{i\Delta}\rangle
\end{aligned}$$

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, $|01\rangle \mapsto (|01\rangle + |10\rangle)/\sqrt{2}$ and $|10\rangle \mapsto (|01\rangle - |10\rangle)/\sqrt{2}$ 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 $|01\rangle, |10\rangle$ 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:

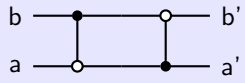
$$U = B_{\theta=\pi/4} P(\pi) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = -H$$

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 $\pi/4 \to -\pi/4$. □

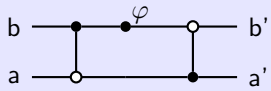
Exercise 7.10: Mach–Zehnder interferometer

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 $D(\gamma) = \exp\left(\gamma a^\dagger - \gamma^* a\right)$ 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:

$$B |\alpha\rangle_a |\beta\rangle_b = \left|\tfrac{1}{\sqrt{2}}(\alpha+\beta)\right\rangle_{a'} \left|\tfrac{1}{\sqrt{2}}(-\alpha+\beta)\right\rangle_{b'}.$$

In terms of the original mode labels (since the names of the physical modes don't change), this becomes

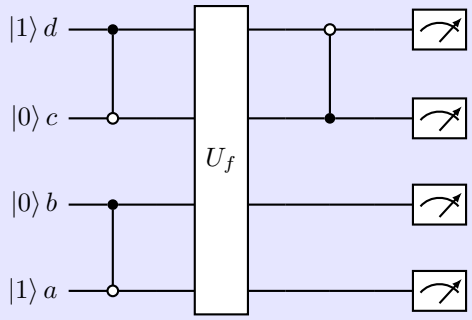
$$B |\alpha\rangle_a |\beta\rangle_b = \left|\tfrac{1}{\sqrt{2}}(\alpha+\beta)\right\rangle_{a} \left|\tfrac{1}{\sqrt{2}}(-\alpha+\beta)\right\rangle_{b}.$$

□

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 $U_f |x,y\rangle = U_f \left|x, y+f(x)\right\rangle$.

   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

$$K = \exp\left(i\chi L\, a^\dagger a\, b^\dagger b\right)$$

and acts on $|\alpha\rangle_a |n\rangle_b$ as

$$K|\alpha\rangle|n\rangle = \exp\left(i\chi L\, n\, a^\dagger a\right)|\alpha\rangle|n\rangle = \left|\alpha e^{i\chi Ln}\right\rangle|n\rangle\,.$$

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 $R_1 = R_2 = 0.9$.

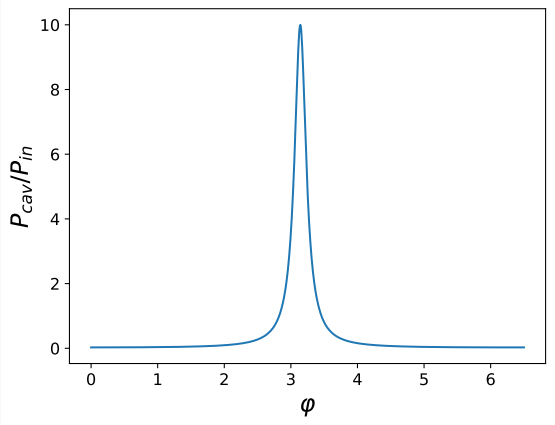
Solution

Concepts Involved: Plotting

A plot of the power of the cavity internal field:

$$\frac{P_{\text{cav}}}{P_{\text{in}}} = \frac{1 - R_1}{\left|1 + e^{i\varphi}\sqrt{R_1 R_2}\right|^2}$$

with the given parameters yields:



We observe a peak in the power profile at detuning $\varphi = \pi$. □

Exercise 7.16: Electric dipole selection rules

($\star\star$) 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:

$$\int Y^*_{l_1,m_2} Y_{1m} Y_{l_2 m_2} d\Omega$$

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:

$$\begin{aligned}
[r_i, L^2] &= [r_i, L_n L_n] \\
&= L_n [r_i, L_n] + [r_i, L_n] L_n \\
&= L_n (i\epsilon_{ink} r_k) + (i\epsilon_{ink} r_k) L_n \\
&= i\epsilon_{ink}(L_n r_k + r_k L_n) \\
&= i\epsilon_{ink}([r_k, L_n] + 2 r_k L_n) \\
&= i\epsilon_{ink}(-i\epsilon_{knl} r_l + 2 r_k L_n) \\
&= \epsilon_{ink}\epsilon_{knl} r_l + 2i\epsilon_{ink} r_k L_n \\
&= -\epsilon_{ink}\epsilon_{lnk} r_l - 2i\epsilon_{ikn} r_k L_n \\
&= -2\delta_{il} r_l - 2i\epsilon_{ikn} r_k L_n \\
&= -2(r_i + i\epsilon_{ikn} r_k L_n)
\end{aligned}$$

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:

$$\begin{aligned}
[r_x, L^2] &= 2i(L_y r_z - r_y L_z) = 2i(r_z L_y - L_z r_y) \\
[r_y, L^2] &= 2i(L_z r_x - r_z L_x) = 2i(r_x L_z - L_x r_z) \\
[r_z, L^2] &= 2i(L_x r_y - r_x L_y) = 2i(r_y L_x - L_y r_x)
\end{aligned}$$

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:

$$\begin{aligned}
&2(l_1(l_1+1) + l_2(l_2+1)) \, \langle l_1, m_1 | \, r_i \, | l_2, m_2 \rangle \\
&= \Big[(l_1(l_1+1))^2 - 2(l_1(l_1+1)l_2(l_2+1)) + (l_2(l_2+1))^2 \Big] \, \langle l_1, m_1 | \, r_i \, | l_2, m_2 \rangle \\
&= (l_1(l_1+1) - l_2(l_2+1))^2 \, \langle l_1, m_1 | \, r_i \, | l_2, m_2 \rangle
\end{aligned}$$

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 $|$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,

$$H\left|\chi_n\right\rangle = \frac{1}{\sqrt{2}}\left(H\left|n,1\right\rangle + H\left|n+1,0\right\rangle\right) = g\sqrt{n+1}\left|\chi_n\right\rangle,$$

$$H\left|\overline{\chi}_n\right\rangle = \frac{1}{\sqrt{2}}\left(H\left|n,1\right\rangle - H\left|n+1,0\right\rangle\right) = -g\sqrt{n+1}\left|\overline{\chi}_n\right\rangle.$$

Thus, $\left|\chi_n\right\rangle$ and $\left|\overline{\chi}_n\right\rangle$ 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 $\{\left|n,1\right\rangle, \left|n+1,0\right\rangle\}$ 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 $\{\left|1,0\right\rangle, \left|0,1\right\rangle\}$, 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 $\left|\psi(0)\right\rangle = \left|0,1\right\rangle$ (atom excited, no photons), we find

$$\left|\psi(t)\right\rangle = U(t)\left|0,1\right\rangle = \cos(gt)\left|0,1\right\rangle - i\sin(gt)\left|1,0\right\rangle.$$

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?

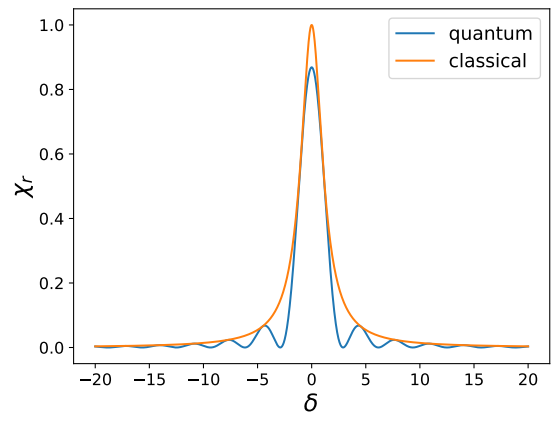
Solution

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

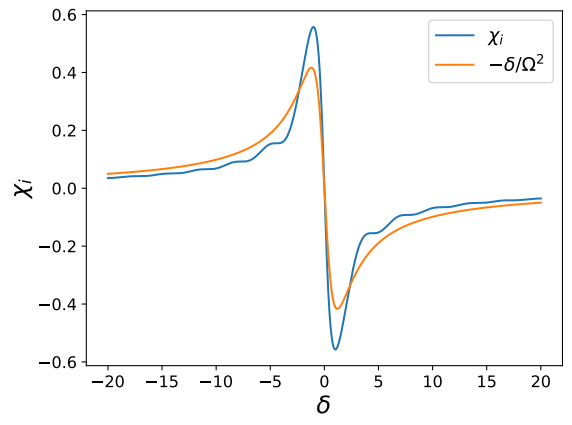
of the field, i.e.:

$$\chi_i = \arg(\langle 01| U |01\rangle) - \arg(\langle 00| U |00\rangle)$$

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

($\star$) 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):

$$\begin{aligned} U_{H_1} = &(\cos\Omega_a t - i\frac{\delta}{\Omega_a}\sin\Omega_a t)\,|100\rangle\langle 100| + (\cos\Omega_a t + i\frac{\delta}{\Omega_a}\sin\Omega_a t)\,|001\rangle\langle 001| \\ &- i\frac{g_a}{\Omega_a}\sin\Omega_a t(|100\rangle\langle 001| + |001\rangle\langle 100|) \\ &+ (\cos\Omega_b t - i\frac{\delta}{\Omega_b}\sin\Omega_b t)\,|010\rangle\langle 010| + (\cos\Omega_b t + i\frac{\delta}{\Omega_b}\sin\Omega_b t)\,|002\rangle\langle 002| \\ &- i\frac{g_b}{\Omega_b}\sin\Omega_b t(|010\rangle\langle 002| + |002\rangle\langle 010|) \end{aligned}$$

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 $\text{diag}(1, e^{-i\varphi_b})$ on the first qubit and the phase gate $\text{diag}(1, e^{-i\varphi_a})$ 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

($\star$) 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

($\star\star$) 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 $|F, m_F\rangle$ for which $F^2 |F, m_F\rangle = F(F+1) |F, m_F\rangle$ and $f_z |F, m_F\rangle = m_F |F, m_F\rangle$.

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:

$$
\begin{aligned}
[i_x, i_y] &= \frac{1}{4}\left(\begin{bmatrix} 3i & 0 & -2\sqrt{3}i & 0 \\ 0 & i & 0 & -2\sqrt{3}i \\ 2\sqrt{3}i & 0 & -i & 0 \\ 0 & 2\sqrt{3}i & 0 & -3i \end{bmatrix} - \begin{bmatrix} -3i & 0 & -2\sqrt{3}i & 0 \\ 0 & -i & 0 & -2\sqrt{3}i \\ 2\sqrt{3}i & 0 & i & 0 \\ 0 & 2\sqrt{3}i & 0 & 3i \end{bmatrix}\right) \\
&= i\frac{1}{2}\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \\
&= ii_z
\end{aligned}
$$

$$
\begin{aligned}
[i_y, i_z] &= \frac{1}{4}\left(\begin{bmatrix} 0 & -i\sqrt{3} & 0 & 0 \\ 3i\sqrt{3} & 0 & 2i & 0 \\ 0 & 2i & 0 & 3i\sqrt{3} \\ 0 & 0 & -i\sqrt{3} & 0 \end{bmatrix} - \begin{bmatrix} 0 & -3i\sqrt{3} & 0 & 0 \\ i\sqrt{3} & 0 & -2i & 0 \\ 0 & -2i & 0 & i\sqrt{3} \\ 0 & 0 & -3i\sqrt{3} & 0 \end{bmatrix}\right) \\
&= i\frac{1}{2}\begin{bmatrix} 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & 2 & 0 & \sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{bmatrix} \\
&= ii_x
\end{aligned}
$$

$$
\begin{aligned}
[i_z, i_x] &= \frac{1}{4}\left(\begin{bmatrix} 0 & 3\sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & -2 & 0 & -\sqrt{3} \\ 0 & 0 & -3\sqrt{3} & 0 \end{bmatrix} - \begin{bmatrix} 0 & \sqrt{3} & 0 & 0 \\ 3\sqrt{3} & 0 & -2 & 0 \\ 0 & 2 & 0 & -3\sqrt{3} \\ 0 & 0 & -\sqrt{3} & 0 \end{bmatrix}\right) \\
&= i\frac{1}{2}\begin{bmatrix} 0 & -i\sqrt{3} & 0 & 0 \\ i\sqrt{3} & 0 & -2i & 0 \\ 0 & 2i & 0 & -i\sqrt{3} \\ 0 & 0 & i\sqrt{3} & 0 \end{bmatrix} \\
&= ii_y
\end{aligned}
$$

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

($\star$) 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:

$$\begin{aligned} P_{\text{decay}} &= \frac{1}{(2\pi c)^3}\frac{8\pi}{3}\int_0^\infty \omega^2 \frac{\omega_0^2}{2\hbar\omega\epsilon_0 c^2}\left|\langle 0|\,\boldsymbol{\mu}\,|1\rangle\right|^2 \frac{4\sin^2(\frac{1}{2}(\omega-\omega_0)t)}{(\omega-\omega_0)^2} d\omega \\ &= \frac{2\omega_0^2\left|\langle 0|\,\boldsymbol{\mu}\,|1\rangle\right|^2}{3\pi^2\hbar\epsilon_0 c^5}\int_0^\infty \frac{\omega\sin^2(\frac{1}{2}(\omega-\omega_0)t)}{(\omega-\omega_0)^2} d\omega \end{aligned}$$

The above integral is of the form $\int_0^\infty \frac{\sin^2(x)}{x}$ which is divergent. This is an ultraviolet/UV divergence arising from large frequencies/energies, or equivalently short length scales. Many techniques from quantum field theory have been developed to tame these infinities - such as regularization and

renormalization. Deferring the details to a different solutions manual for a quantum field theory textbook, let us assume some large frequency cutoff ω_c to the above integral:

$$P_{\text{decay}} \approx \frac{2\omega_0^2 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2}{3\pi^2 \hbar \epsilon_0 c^5} \int_0^{\omega_c} \frac{\omega \sin^2(\frac{1}{2}(\omega-\omega_0)t)}{(\omega-\omega_0)^2} d\omega \tag{12}$$

Since the integrand is sharply peaked at $\omega = \omega_0$ (and hence practically zero for $\omega < 0$), we may replace the lower limit with $\omega = -\omega_c$:

$$P_{\text{decay}} \approx \frac{2\omega_0^2 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2}{3\pi^2 \hbar \epsilon_0 c^5} \int_{-\omega_c}^{\omega_c} \frac{\omega \sin^2(\frac{1}{2}(\omega-\omega_0)t)}{(\omega-\omega_0)^2} d\omega$$

Now let us make the substitution $x = \frac{1}{2}(\omega - \omega_0)t$. The upper limits become $x = \frac{1}{2}(\pm\omega_c - \omega_0)t \approx \pm\frac{\omega_c t}{2}$, and so:

$$\begin{aligned} P_{\text{decay}} &= \frac{2\omega_0^2 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2}{3\pi^2 \hbar \epsilon_0 c^5} \int_{-\omega_c t/2}^{\omega_c t/2} \frac{\frac{2}{t}(x+\omega_0)\sin^2(x)}{(\frac{2}{t}x)^2} dx \\ &= \frac{\omega_0^2 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2 t}{3\pi^2 \hbar \epsilon_0 c^5} \int_{-\omega_c t/2}^{\omega_c t/2} \frac{(x+\omega_0)\sin^2(x)}{x^2} dx \end{aligned}$$

The $\frac{\sin^2(x)}{x}$ term is odd and hence vanishes, leaving us with:

$$P_{\text{decay}} = \frac{\omega_0^3 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2 t}{3\pi^2 \hbar \epsilon_0 c^5} \int_{-\omega_c t/2}^{\omega_c t/2} \frac{\sin^2(x)}{x^2} dx$$

Taking the limit $\omega_c \to \infty$, the remaining integral evaluates to π, and hence:

$$P_{\text{decay}} = \frac{\omega_0^3 \left|\langle 0| \boldsymbol{\mu} |1\rangle\right|^2 t}{3\pi \hbar \epsilon_0 c^5}$$

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 $\omega_0/2\pi \approx 10^{15}\text{Hz}$. 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 $\omega_0 = (2\pi)\cdot 10^{15}\text{Hz}$, $q = 1.60\times 10^{-19}\text{C}$, $a_0 = 5.29\times 10^{-11}\text{m}$, $\hbar = 1.05\times 10^{-34}\text{kgm}^2\text{s}^{-1}$, $\epsilon_0 = 8.85\times 10^{-12}\text{C}^2\text{kg}^{-1}\text{m}^{-3}\text{s}^2$ and $c = 3.00\times 10^8\text{ms}^{-1}$ we find:

$$\gamma_{\text{red}}^{\text{ed}} = 7.5\times 10^7\text{s}^{-1}$$

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:

$$H = e^{i\pi/2}R_z(\pi)R_y(-\pi/2)R_z(0)$$

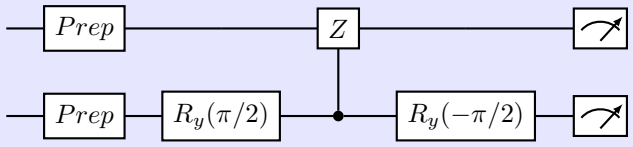
and combining this with the result of Ex. 4.10 (conjugating both sides by Hadamards) we conclude:

$$H = e^{i\pi/2}R_x(\pi)R_y(\pi/2)$$

□

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

($\star$) 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:

$$\begin{aligned}
H_{\text{rot}} &= e^{i\omega Zt/2}H_{\text{lab}}e^{-i\omega Zt/2} - \frac{\omega}{2}Z \\
&= e^{i\omega Zt/2}\left(\frac{\omega}{2}Z + c_1(t)X + c_2(t)Y\right)e^{-i\omega Zt/2} - \frac{\omega}{2}Z \\
&= c_1(t)e^{i\omega Zt/2}Xe^{-i\omega Zt/2} + c_2(t)e^{i\omega Zt/2}Ye^{-i\omega Zt/2} \\
&= c_1(t)(X\cos(\omega t) - Y\sin(\omega t)) + c_2(t)(Y\cos(\omega t) + X\sin(\omega t)) \\
&= \left(c_1(t)\cos(\omega t) + c_2(t)\sin(\omega t)\right)X + \left(-c_1(t)\sin(\omega t) + c_2(t)\cos(\omega t)\right)Y
\end{aligned}$$

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:

$$\hbar\omega = \mu_B B \implies \omega = \frac{\mu_B B}{\hbar} = \frac{(5\times10^{-27}\text{JT}^{-1})(11.8\text{T})}{1.05\times10^{-34}\text{Js}} = 562\text{rad}\cdot\text{s}^{-1}$$

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 $T_{\pi/2} = 10\times10^{-6}$s we find:

$$B_1 = \frac{(1.05\times10^{-34}\text{Js})\cdot\pi}{2\cdot(5\times10^{-27}\text{JT}^{-1})\cdot(1\times10^{-5}\text{s})} = 3.3\times10^{-3}\text{T} = 33\text{Gauss}$$

□

Exercise 7.35: Motional narrowing

($\star$) 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:

$$\begin{aligned}\left\langle H^D_{1,2}\right\rangle &= \frac{\gamma_1\gamma_2\hbar}{4r^3}\frac{1}{4\pi}\int d\Omega[1-3\cos^2\theta] \\ &= \frac{\gamma_1\gamma_2\hbar}{4r^3}\frac{1}{4\pi}\int_0^{2\pi} d\varphi\int_0^{\pi}\sin\theta d\theta[1-3\cos^2\theta] \\ &= \frac{\gamma_1\gamma_2\hbar}{8r^3}\left(\left(-\cos\theta\big|_0^{\pi}\right)+\left(\cos^3\theta\Big|_0^{\pi}\right)\right) \\ &= \frac{\gamma_1\gamma_2\hbar}{8r^3}(2-2) \\ &= 0\end{aligned}$$

hence the claim is proven. □

Exercise 7.36: Thermal equilibrium NMR state

For $n=1$ show that the thermal equilibrium state is

$$\rho \approx 1 - \frac{\hbar\omega}{2k_BT}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$

and for $n=2$ (and $\omega_A \approx 4\omega_B$),

$$\rho \approx 1 - \frac{\hbar\omega_B}{4k_BT}\begin{bmatrix}5 & 0 & 0 & 0\\ 0 & 3 & 0 & 0\\ 0 & 0 & -3 & 0\\ 0 & 0 & 0 & -5\end{bmatrix}$$

The above formulas are incorrect - by inspection they violate the normalization condition $\mathrm{Tr}(\rho)=1$. The corrected formula for $n=1$ is:

$$\rho \approx \frac{1}{2}\left(1 - \frac{\hbar\omega}{2k_BT}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\right)$$

and the corrected formula for $n=2$ is:

$$\rho \approx \frac{1}{4}\left(1 - \frac{\hbar\omega_B}{2k_BT}\begin{bmatrix}5 & 0 & 0 & 0\\ 0 & 3 & 0 & 0\\ 0 & 0 & -3 & 0\\ 0 & 0 & 0 & -5\end{bmatrix}\right)$$

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:

$$\rho \approx \frac{1}{2}\left(1-\frac{\hbar\omega}{2k_BT}\begin{bmatrix}1&0\\0&-1\end{bmatrix}\right) \tag{13}$$

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:

$$\rho \approx \frac{1}{4}\left(1-\frac{\hbar\omega_B}{2k_BT}\begin{bmatrix}5&0&0&0\\0&3&0&0\\0&0&-3&0\\0&0&0&-5\end{bmatrix}\right)$$

□

Exercise 7.37: NMR spectrum of coupled spins

($\star$) 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

$$\begin{aligned}V(t)&=V_0\,\mathrm{Tr}\left[e^{iHt}\rho e^{-iHt}(iX_1+Y_1)\right]\\&=V_0\,\mathrm{Tr}\left[e^{iJZ_1Z_2t}e^{i\pi Y_1/4}\frac{1}{4}[1-\beta\hbar\omega_0(Z_1+Z_2)]e^{-i\pi Y_1/4}e^{-iJZ_1Z_2t}(iX_1+Y_1)\right]\end{aligned}$$

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 $\sim \mathrm{Tr}\left[(iX_1+Y_1)\right]$ 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:

$$V(t) = -\frac{V_0\beta\hbar\omega_0}{4}\,\mathrm{Tr}\Big[e^{iJZ_1Z_2t}e^{i\pi Y_1/4}Z_1e^{-i\pi Y_1/4}e^{-iJZ_1Z_2t}(iX_1+Y_1)\Big]$$

Since Y, Z anticommute we have that

$$e^{i\pi Y_1/4}Z_1e^{-i\pi Y_1/4} = e^{i\pi Y_1/4}e^{i\pi Y_1/4}Z_1 = e^{i\pi Y_1/2}Z_1 = \left(\cos\left(\frac{\pi}{2}\right)I_1 + i\sin\left(\frac{\pi}{2}\right)Y_1\right)Z_1 = iY_1Z_1 = -X_1$$

Now we study $e^{iJZ_1Z_2t}X_1e^{-iJZ_1Z_2t}$, which is composed of four terms after expanding the exponentials into sines/cosines

$$e^{iJZ_1Z_2t}X_1e^{-iJZ_1Z_2t} = \cos^2(Jt)X_1 + i\cos(Jt)\sin(Jt)(Z_1Z_2X_1 - X_1Z_1Z_2) + \sin^2(Jt)Z_1Z_2X_1Z_1Z_2$$

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

$$Z1Z_2X_1Z_1Z_2 = Z_1Z_2(-Z_1)Z_2X_1 = -X_1$$

Hence in the trace expression we are left with:

$$V(t) = \frac{V_0\beta\hbar\omega_0}{4}\,\mathrm{Tr}\Big[(\cos^2(Jt) - \sin^2(Jt))X_1(iX_1+Y_1)\Big]$$

The X_1 term gives $\mathrm{Tr}\big[X_1^2\big] = \mathrm{Tr}[I] = 4$ and the Y_1 term gives $\mathrm{Tr}[X_1Y_1] = \mathrm{Tr}[iZ_1] = 0$, 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

$$V(t)$$
$$= \frac{V_0}{16}\,\mathrm{Tr}\Big[e^{iJZ_1(Z_2+Z_3+Z_4)t}e^{i\pi Y_1/4}[1 - \beta\hbar\omega_0(Z_1+Z_2+Z_3+Z_4)]e^{-i\pi Y_1/4}e^{-iJZ_1(Z_2+Z_3+Z_4)t}(iX_1+Y_1)\Big]$$

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:

$$V(t) = \frac{V_0\beta\hbar\omega_0}{16}\,\mathrm{Tr}\Big[e^{iJZ_1(Z_2+Z_3+Z_4)t}X_1e^{-iJZ_1(Z_2+Z_3+Z_4)t}(iX_1+Y_1)\Big]$$

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

$$\begin{aligned}
V(t) = \frac{V_0\beta\hbar\omega_0}{16}\,\mathrm{Tr}\Big[\Big(&\cos^6(Jt)X_1 + \cos^4(Jt)\sin^2(Jt)(Z_1Z_2X_1Z_1Z_2 + Z_1Z_3X_1Z_1Z_3 + Z_1Z_4X_1Z_1Z_4) \\
&+ \cos^2(Jt)\sin^4(Jt)(Z_1Z_2Z_1Z_3X_1Z_1Z_2Z_1Z_3 + Z_1Z_2Z_1Z_4X_1Z_1Z_2Z_1Z_4 + Z_1Z_3Z_1Z_4X_1Z_1Z_3Z_1Z_4) \\
&+ \sin^6(Jt)(Z_1Z_2Z_1Z_3Z_1Z_4)X_1(Z_1Z_2Z_1Z_3Z_1Z_4)\Big)\,(iX_1 + Y_1)\Big] \\
= \frac{V_0\beta\hbar\omega_0}{16}&(\cos^6(Jt) - 3\cos^4(Jt)\sin^2(Jt) + 3\cos^2(Jt)\sin^4(Jt) - \sin^6(Jt))\,\mathrm{Tr}\big[(X_1(iX_1 + Y_1)\big] \\
= \frac{V_0\beta\hbar\omega_0}{16}&(\cos^6(Jt) - 3\cos^4(Jt)\sin^2(Jt) + 3\cos^2(Jt)\sin^4(Jt) - \sin^6(Jt))\,\mathrm{Tr}[iI + iZ_1] \\
= \frac{V_0\beta\hbar\omega_0}{16}&(\cos^6(Jt) - 3\cos^4(Jt)\sin^2(Jt) + 3\cos^2(Jt)\sin^4(Jt) - \sin^6(Jt))(i16 + 0) \\
= iV_0\beta\hbar\omega_0&(\cos^6(Jt) - 3\cos^4(Jt)\sin^2(Jt) + 3\cos^2(Jt)\sin^4(Jt) - \sin^6(Jt)) \\
= iV_0\beta\hbar\omega_0&(\cos^2(Jt) - \sin^2(Jt))^3 \\
= iV_0\beta\hbar\omega_0&\cos^3(2Jt) \\
= iV_0\beta\hbar\omega_0&\left(\frac{e^{i2Jt} + e^{-i2Jt}}{2}\right)^3 \\
= \frac{iV_0\beta\hbar\omega_0}{8}&(e^{i6Jt} + 3e^{i2Jt} + 3e^{-i2Jt} + e^{-i6Jt})
\end{aligned}$$

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:

$$\begin{aligned} R_{x1}^2 e^{-iaZ_1 t} R_{x1}^2 &= \left(\cos\left(\frac{\pi}{2}\right) I - i\sin\left(\frac{\pi}{2}\right) X\right) \left(\cos(at) I - i\sin(at) Z\right) \left(\cos\left(\frac{\pi}{2}\right) I - i\sin\left(\frac{\pi}{2} X\right)\right) \\ &= (-iX)\left(\cos(at) I - i\sin(at) Z\right)(-iX) \\ &= \left(\cos(at) I + i\sin(at) Z\right)(-iX)(-iX) \\ &= -\left(\cos(at) I + i\sin(at) Z\right) \\ &= -e^{iaZ_1 t}. \end{aligned}$$

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

($\star$) What set of pulses can be used to refocus evolution under *any* single spin Hamiltonian $H^{\text{sys}} = \sum_k c_k \sigma_k$?

We don't believe that there is a single set of pulses that works for a generic single-spin Hamiltonian, so we interpret the question as designing a spin sequence given knowledge of H^{sys}.

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

($\star\star$) 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^2_{zi}, for which we observe that (similar to Ex. 7.38):

$$R^2_{zi}e^{-iaX_it}R^2_{zi} = (-iZ)e^{-iaX_it}(-iZ) = e^{iaX_it}(-iZ)^2 = -e^{iaX_it}$$

$$R^2_{zi}e^{-iaY_it}R^2_{zi} = (-iZ)e^{-iaY_it}(-iZ) = e^{iaX_it}(-iZ)^2 = -e^{iaY_it}$$

where we have used the anticommutativity of the Pauli matrices. Of course, $R^2_{zi}e^{-iaZ_it}R^2_{zi} = -e^{-iaZ_it}$ as Z commutes with itself.
Thus, we find:

$$\begin{aligned} &e^{-iHt}R^2_{z1}e^{-iHt}R^2_{z1} \\ &= e^{-iHt}R^2_{z1}e^{-i(\sum_{i\in\{x,y,z\}}\sum_{j\in\{x,y,z\}}H_{ij}\sigma_i^{(1)}\sigma_j^{(2)})t}R^2_{z1} \\ &= -e^{-iHt}e^{+i(\sum_{i\in\{x,y\}}\sum_{j\in\{x,y,z\}}H_{ij}\sigma_i^{(1)}\sigma_j^{(2)})t}e^{-i(\sum_{j\in\{x,y,z\}}H_{zj}\sigma_z^{(1)}\sigma_j^{(2)})t} \\ &= -e^{-i(\sum_{i\in\{x,y,z\}}\sum_{j\in\{x,y,z\}}H_{ij}\sigma_i^{(1)}\sigma_j^{(2)})t}e^{+i(\sum_{i\in\{x,y\}}\sum_{j\in\{x,y,z\}}H_{ij}\sigma_i^{(1)}\sigma_j^{(2)})t}e^{-i(\sum_{j\in\{x,y,z\}}H_{zj}\sigma_z^{(1)}\sigma_j^{(2)})t} \\ &= -e^{-i(\sum_{j\in\{x,y,z\}}H_{zj}\sigma_z^{(1)}\sigma_j^{(2)})2t} \end{aligned}$$

and analogously for R^2_{z2}. 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:

$$\begin{aligned}\hat{\mathbf{m}} &= \frac{\sin\left(-\frac{\pi}{2}(\frac{b}{c}+1)\right)\cos\left(\frac{\pi}{2}\right)\hat{\mathbf{z}} + \cos\left(-\frac{\pi}{2}(\frac{b}{c}+1)\right)\sin\left(\frac{\pi}{2}\right)\hat{\mathbf{n}} - \sin\left(-\frac{\pi}{2}(\frac{b}{c}+1)\right)\sin\left(\frac{\pi}{2}\right)\hat{\mathbf{n}}\times\hat{\mathbf{z}}}{\sin(\beta/2)} \\ &= \frac{\cos\left(\frac{\pi}{2}(\frac{b}{c}+1)\right)\hat{\mathbf{n}} + \sin\left(\frac{\pi}{2}(\frac{b}{c}+1)\right)\frac{\hat{\mathbf{y}}}{\sqrt{2}}}{\sin(\beta/2)}\end{aligned}$$

□

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).

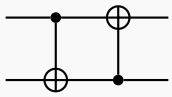
Solution

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

($\star$) 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:

$$(26)(37) = (26)(48)^2(37) = (26)(48)(37)(48) = CX_{3,1}CX_{2,1}$$

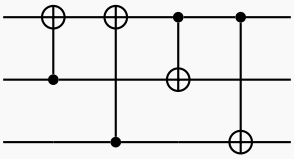
and further that:

$$CX_{1,3}CX_{1,2} = (56)(78)(57)(68) = (58)(67)$$

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

($\star\star$) 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

($\star$) 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 $1 \leftrightarrow 2$:

$$e^{iHt}(iX_2 + Y_2)e^{-iHt} \sim X_2 + Y_2 + Z_1Y_2 + Z_1X_2 \tag{14}$$

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

($\star\star\star$) 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 $M_5 = R_{y1}R_{y2}R_{x3}$:

$$\begin{aligned} V_5 \sim \mathrm{Tr}\big(\rho\,(&Z_1 + X_1 + Z_1X_2 + Y_1X_2 + Z_1Y_3 + Y_1Y_3 + Z_1X_2Y_3 + Y_1X_2Y_3 \\ &+ Z_2 + Y_2 + X_1Z_2 + X_1Y_2 + Z_2Y_3 + Y_2Y_3 + X_1Z_2Y_3 + X_1Y_2Y_3 \\ &+ X_3 + Z_3 + X_1X_3 + X_1Z_3 + X_2X_3 + X_2Z_3 + X_1X_2X_3 + X_1X_2Z_3)\big) \end{aligned}$$

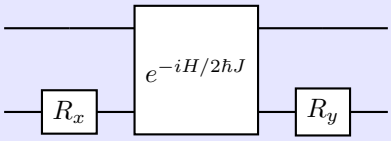
Finally for $M_6 = R_{y1}R_{y2}R_{y3}$:

$$\begin{aligned} V_6 \sim \mathrm{Tr}\big(\rho\,(&Z_1 + Y_1 + Z_1X_2 + Y_1X_2 + Z_1X_3 + Y_1Z_3 + Z_1X_2X_3 + Y_1X_2X_3 \\ &+ Z_2 + Y_2 + X_1Z_2 + X_1Y_2 + Z_2Y_3 + Y_2X_3 + X_1Z_2X_3 + X_1Y_2X_3 \\ &+ \ Z_3 + Y_3 + X_1Z_3 + X_1Y_3 + X_2Z_3 + X_2Y_3 + X_1X_2Z_3 + X_1X_2Y_3)\big) \end{aligned}$$

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:

$$U = e^{iH/2\hbar J} = e^{i\pi Z_1 Z_2/4} = \cos\left(\frac{\pi}{4}\right) I + i \sin\left(\frac{\pi}{4}\right) Z_1 Z_2 = \mathrm{diag}(e^{i\pi/4}, e^{-i\pi/4}, e^{-i\pi/4}, e^{i\pi/4})$$

For convenience, let us factor out a global phase:

$$U \cong \mathrm{diag}(1, -i, -i, 1)$$

We can then calculate:

$$R_{y2}UR_{x2} = (\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \otimes \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix})\text{diag}(1,-i,-i,1)(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \otimes \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -i\\ -i & 1\end{bmatrix})$$
$$= \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & -i & 0 & 0\\ 0 & 0 & 0 & -1\\ 0 & 0 & -i & 0\end{bmatrix}$$

which we can see has the action:

$$|00\rangle \mapsto |00\rangle$$
$$|01\rangle \mapsto -i\,|01\rangle$$
$$|10\rangle \mapsto -i\,|11\rangle$$
$$|11\rangle \mapsto -\,|10\rangle$$

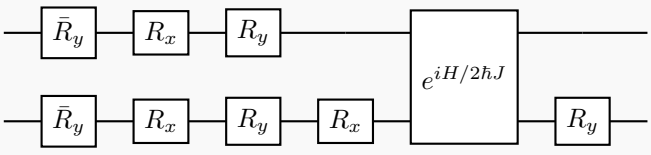
i.e up to single qubit phases we indeed have a `CNOT` gate. We require the use one-qubit Z rotations that have the net action:

$$|00\rangle \mapsto |00\rangle$$
$$|01\rangle \mapsto i\,|01\rangle$$
$$|10\rangle \mapsto i\,|10\rangle$$
$$|11\rangle \mapsto -\,|11\rangle$$

before the faulty `CNOT` gate in order to cancel the phases. We observe that $R_z = \text{diag}(e^{-i\pi/4}, e^{i\pi/4}) \cong \text{diag}(1, i)$ acting on both qubits has precisely this action, so the proper `CNOT` is given by:

$$CX_{1,2} = R_{y2}UR_{x2}R_{z1}R_{z2}$$

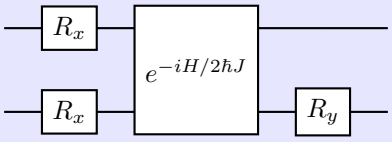
Noticing that $R_y\bar{R}_x\bar{R}_y = \text{diag}(e^{-i\pi/4}, e^{i\pi/4}) = R_z$ (with $\bar{R} = R^\dagger$), we thus have that the circuit below has the action (up to a global phase) of a `CNOT`:



□

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. □

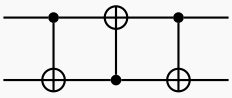
Exercise 7.49: NMR swap gate

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

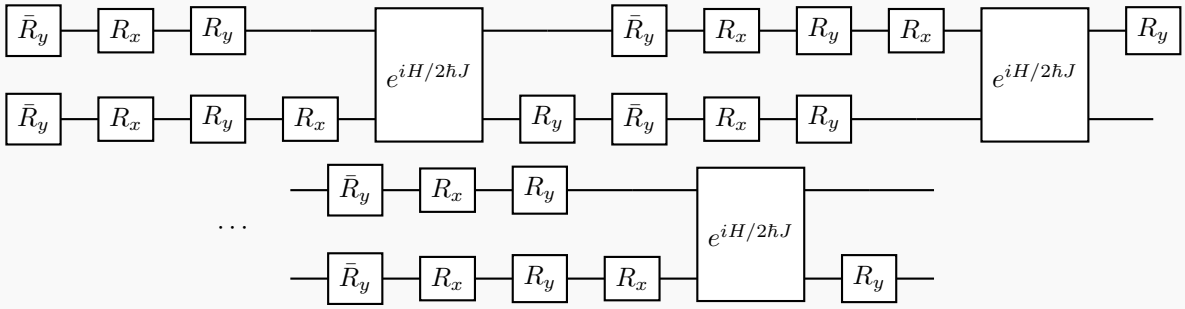
Solution

Concepts Involved: Controlled Operations, Rotations, Gate Decomposition

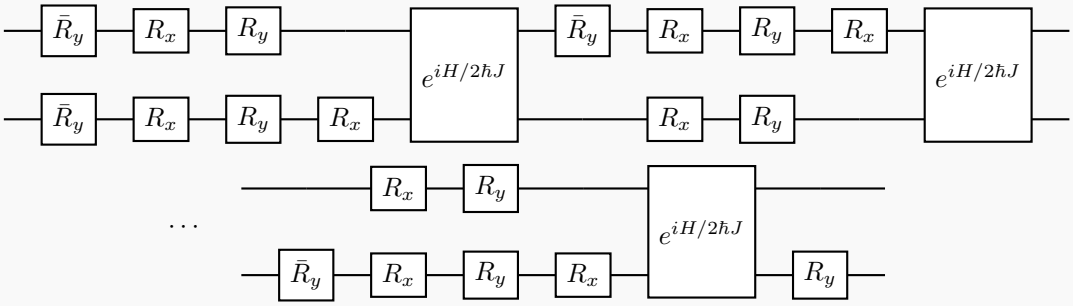
We can use the circuit of Figure 1.7:



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:

$$\bar{R}_{z1}\bar{R}_{z2}\tau = R_{y1}R_{x1}\bar{R}_{y1}R_{y2}R_{x2}\bar{R}_{y2}\tau = \text{diag}(1, -1, -1, -1) = P \cong \text{diag}(-1, 1, 1, 1) = O_0$$

$$R_{z1}\bar{R}_{z2}\tau = R_{y1}\bar{R}_{x1}\bar{R}_{y1}R_{y2}R_{x2}\bar{R}_{y2}\tau = \mathrm{diag}(1,-1,1,1) = O_1$$

$$\bar{R}_{z1}R_{z2}\tau = R_{y1}R_{x1}\bar{R}_{y1}R_{y2}\bar{R}_{x2}\bar{R}_{y2}\tau = \mathrm{diag}(1,1,-1,1) = O_2$$

$$R_{z1}R_{z2}\tau = R_{y1}\bar{R}_{x1}\bar{R}_{y1}R_{y2}\bar{R}_{x2}\bar{R}_{y2}\tau = \mathrm{diag}(1,1,1,-1) = O_3$$

□

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

($\star$) 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:

$$\begin{aligned}
\sqrt{\texttt{SWAP}} &= \exp\left(-i\pi\mathbf{S}_1 \cdot \mathbf{S}_2/2\right) \\
&= \exp\left(-i\pi X_1X_2/8\right)\exp\left(-i\pi Y_1Y_2/8\right)\exp\left(-i\pi Z_1Z_2/8\right) \\
&= \left(\cos\left(\frac{\pi}{8}\right)I - i\sin\left(\frac{\pi}{8}\right)X_1X_2\right)\left(\cos\left(\frac{\pi}{8}\right)I - i\sin\left(\frac{\pi}{8}\right)Y_1Y_2\right)\left(\cos\left(\frac{\pi}{8}\right)I - i\sin\left(\frac{\pi}{8}\right)Z_1Z_2\right) \\
&= \begin{bmatrix} e^{i\pi/8} & 0 & 0 & 0 \\ 0 & e^{i\pi/8}\frac{1+i}{2} & e^{i\pi/8}\frac{1-i}{2} & 0 \\ 0 & e^{i\pi/8}\frac{1-i}{2} & e^{i\pi/8}\frac{1+i}{2} & 0 \\ 0 & 0 & 0 & e^{i\pi/8} \end{bmatrix} \\
&\cong \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1+i}{2} & \frac{1-i}{2} & 0 \\ 0 & \frac{1-i}{2} & \frac{1+i}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

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

($\star\star\star$) 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 Section 4.2.4 of Quantum Algorithms, due to Beth and Rötteler. We make the observation that permutations which operate via linear transformations on the names of the bitstrings/binary words of length n (of which cyclic permutations are a subset) can be associated with with invertible binary $n \times n$ 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

($\star\star$) 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-1} pairs of unary states/wires corresponding to the qubit in the $|0/1\rangle$ 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:

$$
\begin{aligned}
|000\rangle &\leftrightarrow |00000001\rangle = |0\rangle \\
|001\rangle &\leftrightarrow |00000010\rangle = |1\rangle \\
|010\rangle &\leftrightarrow |00000100\rangle = |2\rangle \\
|011\rangle &\leftrightarrow |00001000\rangle = |3\rangle \\
|100\rangle &\leftrightarrow |00010000\rangle = |4\rangle \\
|101\rangle &\leftrightarrow |00100000\rangle = |5\rangle \\
|110\rangle &\leftrightarrow |01000000\rangle = |6\rangle \\
|111\rangle &\leftrightarrow |10000000\rangle = |7\rangle
\end{aligned}
$$

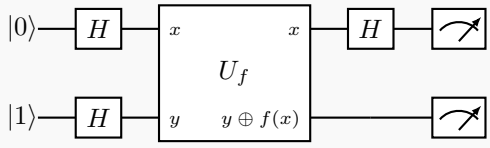
   To apply an effective Z-rotation on the second qubit, we would apply phase shifters of $P(-\theta/2)$ to $|0\rangle, |1\rangle, |4\rangle, |5\rangle$ and $P(\theta/2)$ to $|2\rangle, |3\rangle, |6\rangle, |7\rangle$. To apply an effective Y-rotation on the second qubit, we would apply beamsplitters B_θ 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:

$$|00\rangle \leftrightarrow |0001\rangle$$
$$|01\rangle \leftrightarrow |0010\rangle$$
$$|10\rangle \leftrightarrow |0100\rangle$$
$$|11\rangle \leftrightarrow |1000\rangle$$

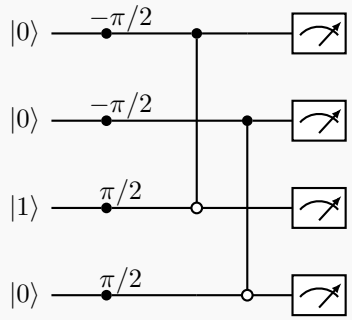
and denote $|0001\rangle$ as the top optical line down to $|1000\rangle$ as the bottom optical line.

The 1-qubit Deutsch-Jozsa circuit to reproduce looks like (in the binary representation):

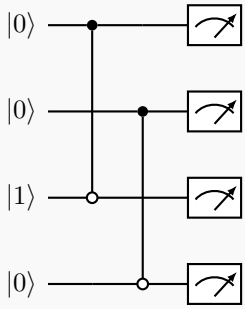


It will also be useful to recall the decomposition of the Hadamard gate as $H = R_y(\pi/2)Z = R_y(\pi/2)R_z(\pi)$. Let us look at the four possible cases of $f(x)$.

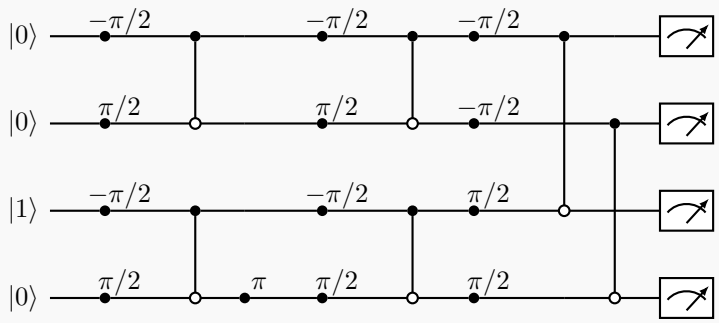
- $f(x) = 0$. in this case, U_f is the identity, so the Hadamard on the top qubit cancels, leaving just the single Hadamard on the bottom qubit. In unary, the resulting circuit takes the form:



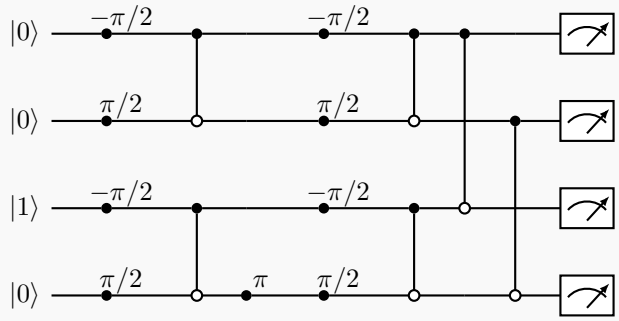
- $f(x) = 1$. In this case $U_f = X_2$. Again the top two Hadamards cancel, and on the second qubit we have $XH = (HZH)H = HZ = R_y(\pi/2)ZZ = R_y(\pi/2)$. Thus in the unary representation:



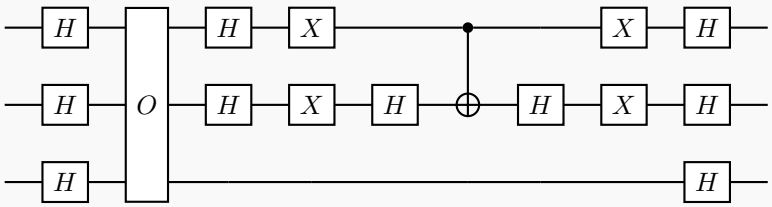
- $f(x) = x$. In this case, $U_f = CX_{1,2} = H_2CZ_{1,2}H_2$ so two Hadamards on the bottom qubit cancel and the resulting (binary) circuit is $H_1H_2CZ_{1,2}H_1$. In Unary, this takes the form:



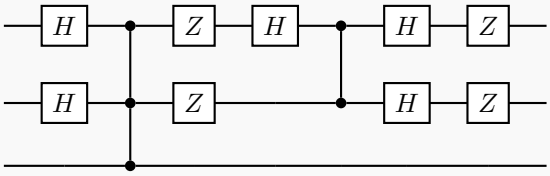
- $f(x) = x \oplus 1$. In this case $U_f = CX_{1,2}X_2 = H_2CZ_{1,2}H_2H_2Z_2H_2 = H_2CZ_{1,2}Z_2H_2$ and again the two Hadamards on the bottom qubits cancel, leaving us with $H_1H_2CZ_{1,2}Z_2H_1 = H_1R_{y,2}(\pi/2)Z_2Z_2CZ_{1,2}H_1 = H_1R_{y,2}(\pi/2)CZ_{1,2}H_1$. In Unary, this takes the form:



3. We adopt the 2-qubit search circuit given in Box 6.1:

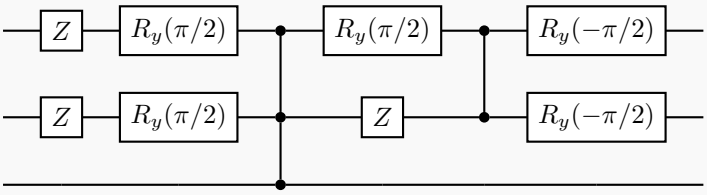


For simplicity, we take the case where $O = CCX_{1,2,3}$ (Toffoli with controls on the first two qubits and the target as the third qubit). The other three cases can be handled by appropriate conjugation by X on the control qubits. We remove Xs from our circuit via Hadamard conjguation via inserting $H^2 = I$, which yields:

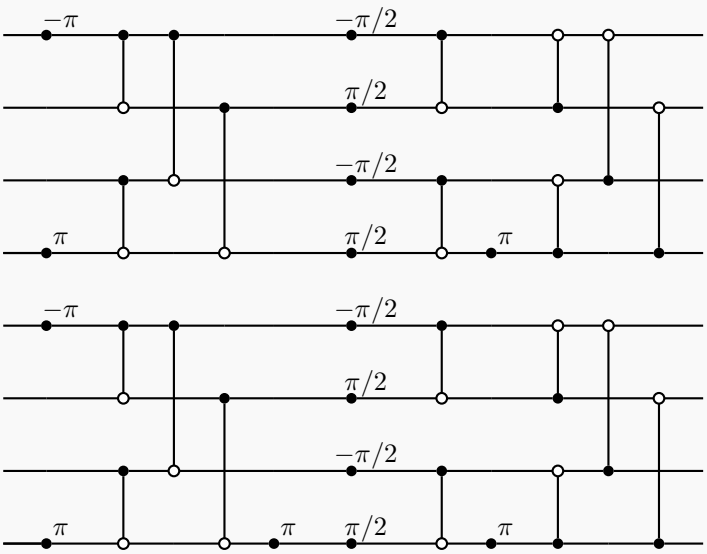


the only possibly foreign gate is the CCZ gate, which has the action $CCZ\,|111\rangle = -|111\rangle$ and identity on all other computational basis states. In the unary representation, it is easily implemented by a phase shift by π on the bottom/$|7\rangle$ state.

Finally, Let us use $H = R_y(\pi/2)Z = ZR_y(-\pi/2)$ to obtain the final circuit to convert:



As a unary circuit:



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

($\star\star\star$) 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

$$A_n = e^{-|\alpha|^2}\frac{|\alpha|^2}{n!}\begin{bmatrix} \cos\left(\theta\sqrt{n}\right) & \frac{i\sqrt{n}}{\alpha}\sin\left(\theta\sqrt{n}\right) \\ \frac{i\alpha}{\sqrt{n+1}}\sin\left(\theta\sqrt{n+1}\right) & \cos\left(\theta\sqrt{n+1}\right)\end{bmatrix}$$

The results of Exercise 7.17 may be helpful.

There is an error in the above expression, the correct expression should be:

$$A_n = e^{-|\alpha|^2/2}\frac{\alpha^n}{\sqrt{n!}}\begin{bmatrix} \cos\left(\theta\sqrt{n}\right) & \frac{i\sqrt{n}}{\alpha}\sin\left(\theta\sqrt{n}\right) \\ \frac{i\alpha}{\sqrt{n+1}}\sin\left(\theta\sqrt{n+1}\right) & \cos\left(\theta\sqrt{n+1}\right)\end{bmatrix}$$

2. 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 Approxiamtion, Fidelity

1. We recall from Exercise 7.17 that:

$$H\,|\chi_n\rangle = \sqrt{n+1}\,|\chi_n\rangle$$
$$H\,|\overline{\chi}_n\rangle = -\sqrt{n+1}\,|\overline{\chi}_n\rangle$$

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 distri-

bution

$$p_n \approx e^{-x}\frac{x^n}{\sqrt{2\pi n}n^n e^{-n}}$$

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

$$\begin{aligned} p_L &\approx \frac{e^{-L\sqrt{x}}}{\sqrt{2\pi}}\exp\left(\log\left(\left(\frac{x}{x-L\sqrt{x}}\right)^{x-L\sqrt{x}}\right)\right) \\ &= \frac{1}{\sqrt{2\pi}}\exp\Big(-L\sqrt{x}+(x-L\sqrt{x})[\log(x)-\log\big(x-L\sqrt{x}\big)]\Big) \end{aligned}$$

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

$$\begin{aligned} p_L &\approx \frac{1}{\sqrt{2\pi}}\exp\left(-L\sqrt{x}+(x-L\sqrt{x})[\log(x)-(\log(x)-\frac{L}{\sqrt{x}}-\frac{1}{2}\frac{L^2}{x})]\right) \\ &= \frac{1}{\sqrt{2\pi}}\exp\left(-L\sqrt{x}+(x-L\sqrt{x})[\frac{L}{\sqrt{x}}+\frac{1}{2}\frac{L^2}{x})]\right) \\ &= \frac{1}{\sqrt{2\pi}}\exp\left(-L\sqrt{x}+L\sqrt{x}+\frac{1}{2}L^2-L^2-\frac{1}{2}\frac{L^3}{\sqrt{x}}\right) \\ &= \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{L^2}{2}-\frac{1}{2}\frac{L^3}{\sqrt{x}}\right) \end{aligned}$$

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:

$$\begin{bmatrix} \cos\left(\theta\sqrt{n}\right) & \frac{i\sqrt{n}}{\alpha}\sin\left(\theta\sqrt{n}\right) \\ \frac{i\alpha}{\sqrt{n+1}}\sin\left(\theta\sqrt{n+1}\right) & \cos\left(\theta\sqrt{n+1}\right) \end{bmatrix} = \begin{bmatrix} \cos(\theta a\alpha) & ia\sin(\theta a\alpha) \\ (i/b)\sin(\theta b\alpha) & \cos(\theta b\alpha) \end{bmatrix}$$

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:

$$A_L^\dagger A_L \approx \frac{e^{-L^2/2}}{\sqrt{2\pi}} \begin{bmatrix} \cos^2(\varphi) + \sin^2(\varphi) & i\cos(\varphi)\sin(\varphi) - i\cos(\varphi)\sin(\varphi) \\ -i\cos(\varphi)\sin(\varphi) + i\cos(\varphi)\sin(\varphi) & \cos^2(\varphi) + \sin^2(\varphi) \end{bmatrix}$$
$$= \frac{e^{-L^2/2}}{\sqrt{2\pi}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

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 $L = O(1)$; $a := \sqrt{n}/\alpha = \sqrt{1+Ly}$, $b := \sqrt{n+1}/\alpha = \sqrt{1+Ly+y^2}$; $U = e^{i\varphi\sigma_x} = \left(\begin{smallmatrix} \cos\varphi & i\sin\varphi \\ i\sin\varphi & \cos\varphi \end{smallmatrix}\right)$, $\varphi = \alpha\theta$; $M_L = \left(\begin{smallmatrix} \cos(a\varphi) & i\,a\,\sin(a\varphi) \\ i\,b^{-1}\sin(b\varphi) & \cos(b\varphi) \end{smallmatrix}\right)$; $K_L := U^\dagger M_L - I$; $|g(L)|^2 = (2\pi)^{-1/2} e^{-L^2/2}$.

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 $Q_L(\mathbf{r}) = \mathbf{r}^\top A_L\,\mathbf{r} + \text{const}$ 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

($\star\star$) 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,

$$\mathcal{Y}_{\hat{\mathbf{n}}}^{\text{blue}}(\theta) = \exp\left[\left(e^{i\varphi}\,|00\rangle\langle 11| + e^{-i\varphi}\,|11\rangle\langle 00| + e^{i\varphi}\sqrt{2}\,|01\rangle\langle 12| + e^{-i\varphi}\,|12\rangle\langle 01| + \ldots\right)\frac{i\theta}{2}\right],$$

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 $a^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle$ 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 $|0\rangle$.

(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 $|00\rangle, |01\rangle, |10\rangle$, and $|11\rangle$, 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 $|0\rangle$?

Solution

Concepts Involved: Controlled Operations, Operator Functions, Rotations

(1) The red sideband couples $|1,n\rangle \leftrightarrow |0,n{+}1\rangle$; with motion initially $|0\rangle$ the active subspace is $\mathcal{S} = \text{span}\{|1,0\rangle, |0,1\rangle\}$ 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

$$\mathcal{Y}_{\hat{n}}^{\text{red}}(\theta)\big|_{\mathcal{S}} = \exp\Big(\tfrac{i\theta}{2}A\Big) = \cos\tfrac{\theta}{2}\,I + i\sin\tfrac{\theta}{2}\,A.$$

For a π pulse,

$$S^j \equiv \mathcal{Y}_{\hat{n}}^{\text{red},\,j}(\pi)\big|_{\mathcal{S}} = iA = i\begin{pmatrix} 0 & e^{-i\varphi} \\ e^{+i\varphi} & 0 \end{pmatrix},$$

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 $\mathcal{Y}_\gamma^{\mathrm{blue}}(\cdot)$ acts block-diagonally as an $\mathrm{SU}(2)$ rotation. Assume $\mathcal{Y}_{\hat{n}}^{\mathrm{blue}}(\theta)$ 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

$$W = \mathcal{Y}_\alpha^{\mathrm{blue}}(\beta), \qquad U = W^\dagger\,\mathcal{Y}_{\hat{n}}^{\mathrm{blue}}(\theta)\,W.$$

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 $\mathcal{B}_0 = \mathrm{span}\{|00\rangle, |11\rangle\}$, 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 $\mathcal{B}_1 = \mathrm{span}\{|01\rangle, |12\rangle\}$ the blue pulse is a scalar $-I$ (hence $|01\rangle \mapsto -|01\rangle$ and no leakage), while on $\mathcal{B}_0 = \mathrm{span}\{|00\rangle, |11\rangle\}$ it is an $\mathrm{SU}(2)$ rotation by angle $\pi\sqrt{2}$ about some in-plane axis. Choose

$$U \;=\; \mathcal{Y}_{\hat{y}}^{\mathrm{blue}}\big(-\tfrac{\pi}{2}\big)\;\mathcal{Y}_{\hat{x}}^{\mathrm{blue}}\Big(\pi\sqrt{2}\Big)\;\mathcal{Y}_{\hat{y}}^{\mathrm{blue}}\big(\tfrac{\pi}{2}\big)\,.$$

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 $\{|00\rangle\,, |01\rangle\,, |10\rangle\,, |11\rangle\}$,

$$\boxed{U = \mathrm{diag}\,\big(e^{-i\pi/\sqrt{2}},\ -1,\ 1,\ e^{+i\pi/\sqrt{2}}\big)}$$

is diagonal as required.

(5) Let

$$U = \mathrm{diag}\big(e^{-i\gamma},\,-1,\,1,\,e^{+i\gamma}\big), \qquad \gamma = \tfrac{\pi}{\sqrt{2}},$$

acting on (internal⊗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 $|0\rangle$),

$$G = \mathrm{diag}\big(e^{-i\gamma},\ 1,\ -1,\ e^{+i\gamma}\big) \quad \text{in the basis } \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}.$$

Choose single-qubit Z phases

$$Z_c = \mathrm{diag}\big(1,\,e^{-i(\pi+\gamma)}\big), \qquad Z_t = \mathrm{diag}\big(1,\,e^{-i\gamma}\big).$$

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

$$\frac{\mathcal{E}_m(\rho)}{\mathrm{tr}\left(\mathcal{E}_m(\rho)\right)}$$

Also show that the probability of obtaining this measurement result is $p(m) = \mathrm{tr}\left(\mathcal{E}_m(\rho)\right)$.

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:

$$\mathrm{tr}\big(\mathcal{E}_m(\rho)\big) = \mathrm{tr}\Big(M_m \rho M_m^\dagger\Big) = \mathrm{tr}\Big(M_m |\psi\rangle\langle\psi| M_m^\dagger\Big) = \langle\psi| M_m^\dagger M_m |\psi\rangle = p(m)$$

where we have used that $\mathrm{tr}\big(|\alpha\rangle\langle\beta|\big) = \langle\alpha|\beta\rangle$. 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:

$$\frac{M_m |\psi\rangle}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}} \left(\frac{\langle\psi| M_m^\dagger}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}} \right) = \frac{M_m |\psi\rangle\langle\psi| M_m^\dagger}{\langle\psi| M_m^\dagger M_m |\psi\rangle} = \frac{M_m \rho M_m^\dagger}{p(m)} = \frac{\mathcal{E}_m(\rho)}{\mathrm{tr}\big(\mathcal{E}_m(\rho)\big)}$$

which is what we wished to show. □

Exercise 8.3

($\star$) 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: Operator-Sum 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

($\star$) 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

($\star$) 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 Kraus operators, ensuring that the process remains reversible in the extended system-environment space. □

Exercise 8.9: Measurement model

($\star$) 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

($\star$) 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

$$W_{jk} = \mathrm{Tr}(E_j^\dagger E_k).$$

Since W is constructed from inner products, it satisfies

$$W_{jk}^* = \mathrm{Tr}((E_j^\dagger E_k)^*) = \mathrm{Tr}(E_k^\dagger E_j) = W_{kj},$$

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,

$$\mathrm{rank}(W) \leq d^2.$$

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 $\{F_j\}$. □

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 $W_{ij} = \mathrm{Tr}\left(E_j'^\dagger E_k'\right)$ as in the last problem, wherein W now satisfies $\mathrm{rank}(W) \leq dd'$ 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 $\det\left(O'\right) = 1$ 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

($\star$) 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 $\rho' = U\rho U^\dagger = \frac{1}{2}\left(I + \mathbf{r}' \cdot \boldsymbol{\sigma}\right)$ 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

$$U(\boldsymbol{\sigma}\cdot\mathbf{r})U^\dagger = (c^2-s^2)\,\boldsymbol{\sigma}\cdot\mathbf{r} + 2cs\,\boldsymbol{\sigma}\cdot(\hat{n}\times\mathbf{r}) + 2s^2(\hat{n}\cdot\mathbf{r})\,\boldsymbol{\sigma}\cdot\hat{n}$$
$$= \boldsymbol{\sigma}\cdot\Big(\mathbf{r}\cos\theta + (\hat{n}\times\mathbf{r})\sin\theta + \hat{n}(\hat{n}\cdot\mathbf{r})(1-\cos\theta)\Big).$$

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. 8.12). □

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

($\star$) 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 $K^\dagger K = |0\rangle\langle 0| + \lambda^2\,|1\rangle\langle 1| \le I$, 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

$$x' = \frac{\lambda x}{p}, \qquad y' = \frac{\lambda y}{p}, \qquad z' = \frac{(1-\lambda^2)+(1+\lambda^2)z}{(1+\lambda^2)+(1-\lambda^2)z}. \tag{$*$}$$

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 ($*$):

$$r_1' = (0,0,1), \qquad r_2' = (0,0,-1), \qquad \bar r' = \Big(0,0,\frac{1-\lambda^2}{1+\lambda^2}\Big).$$

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

$$\begin{aligned}
\mathrm{Tr}\left([\mathcal{E}_p(\rho)]^k\right) &= \sum_{m=0}^{k}\binom{k}{m}(1-p)^{k-m}p^m\,\mathrm{Tr}\left(\rho^{k-m}\right)\left(\frac{I}{d}\right)^m \\
&\le \mathrm{Tr}\left(\rho^k\right)\sum_{m=0}^{k}\binom{k}{m}(1-p)^{k-m}p^m\left(\frac{1}{d}\right)^m && \text{as } \mathrm{Tr}(\rho^k)\ge\mathrm{Tr}(\rho^{k-m})\ \forall m\ge 0 \\
&= \mathrm{Tr}\left(\rho^k\right)(1-p+\frac{p}{d})^k \\
&\le \mathrm{Tr}\left(\rho^k\right)(1-p+p)^k && \forall\, d\ge 1 \\
&= \mathrm{Tr}\left(\rho^k\right)
\end{aligned}$$

□

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

$$\mathcal{E}(\rho) = (1-p)\rho + \frac{p}{d}I,$$

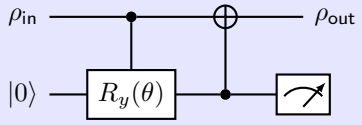
where $0 \le p \le 1$ is the depolarizing probability and I is the identity operator in the d-dimensional space. The operator-sum (Kraus) representation for this channel is

$$\begin{aligned}
E_0 &= \sqrt{1-p}\,I \\
E_{i,j} &= \sqrt{\frac{p}{d}}\,|i\rangle\langle j| \qquad i,j = 1,\ldots,d.
\end{aligned}$$

□

Exercise 8.20

Show that the circuit in Figure 8.13 models the amplitude damping quantum operation, with $\sin^2(\theta/2) = \gamma$.



Solution

Concepts Involved: Kraus Representation, 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

$$\begin{aligned} E_0 &\equiv \langle 0_E|U|0_E\rangle = P_0\langle 0_E|P_0|0_E\rangle + P_1\langle 0_E|P_0 R_y(\theta)|0_E\rangle = P_0 + P_1.\cos\big(\theta/2\big) \\ E_1 &\equiv \langle 1_E|U|0_E\rangle = X P_1\langle 1_E|P_1 R_y(\theta)|0_E\rangle = X P_1.\sin\big(\theta/2\big) \end{aligned}$$

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

($\star$) 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 $|n\rangle \equiv |n_a\rangle$,

$$U\,|n,0\rangle_{ab} = \frac{(U a^\dagger U^\dagger)^n}{\sqrt{n!}}\;|0,0\rangle = \frac{(a^\dagger \cos\theta - i\, b^\dagger \sin\theta)^n}{\sqrt{n!}}\,|0,0\rangle\,.$$

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

$$E_0^{\mathsf{dr}} = \sqrt{1-\gamma}\, I$$
$$E_1^{\mathsf{dr}} = \sqrt{\gamma}\left[|00\rangle\langle 01| + |00\rangle\langle 10|\right]$$

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

$$K_0 \otimes K_0\,|\psi\rangle = \sqrt{1-\gamma}\,|\psi\rangle,$$
$$K_1 \otimes K_0\,|\psi\rangle = a\sqrt{\gamma}\,|00\rangle\,,$$
$$K_0 \otimes K_1\,|\psi\rangle = b\sqrt{\gamma}\,|00\rangle\,,$$
$$K_1 \otimes K_1\,|\psi\rangle = 0.$$

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

$$E_0^{\text{dr}} = \sqrt{1-\gamma}I,$$
$$E_1^{\text{dr}} = \sqrt{\gamma}\left(|00\rangle\langle 01| + |00\rangle\langle 10|\right).$$

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:

$$p_0 = \frac{e^{-E_0/k_BT}}{\mathcal{Z}},$$
$$p_1 = \frac{e^{-E_1/k_BT}}{\mathcal{Z}},$$

where the partition function is:

$$\mathcal{Z} = e^{-E_0/k_BT} + e^{-E_1/k_BT}.$$

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:

$$p_1 = \frac{e^{-E_1/k_BT}}{e^{-E_0/k_BT} + e^{-E_1/k_BT}}$$
$$= \frac{e^{-\Delta E/k_BT}}{1 + e^{-\Delta E/k_BT}}.$$

Thus, the ratio of the probabilities satisfies:

$$\frac{p_1}{p_0} = e^{-\Delta E/k_BT}.$$

Given a state ρ_∞ with diagonal elements $(\rho_\infty)_{00} = p$ and $(\rho_\infty)_{11} = 1-p$, we can determine its temperature by solving:

$$\frac{1-p}{p} = e^{-\Delta E/k_BT}.$$

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

$$\tilde{E}_0 = \begin{pmatrix} \cos\big(\theta/2\big) & 0 \\ 0 & \cos\big(\theta/2\big) \end{pmatrix}, \quad \tilde{E}_1 = \begin{pmatrix} \sin\big(\theta/2\big) & 0 \\ 0 & -\sin\big(\theta/2\big) \end{pmatrix}.$$

Using the transform equations,

$$\tilde{E}_0 = u_{00}E_0 + u_{01}E_1$$

yields $u_{00} = \cos\big(\theta/2\big)$ and $u_{01} = \sin\big(\theta/2\big)$.
Similar equations for $\tilde{E}_1$ gives us, $u_{10} = \sin\big(\theta/2\big)$ and $u_{11} = -\cos\big(\theta/2\big)$. 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

$$\begin{aligned}
\mathcal{E}_{PD}(I) &= E_0 I E_0^\dagger + E_1 I E_1^\dagger \\
&= E_0 E_0^\dagger + E_1 E_1^\dagger \\
&= \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\lambda} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\lambda} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\lambda} \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1-\lambda \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \lambda \end{bmatrix} \\
&= I
\end{aligned}$$

so it is also unital. For the amplitude damping channel we have

$$\begin{aligned}
\mathcal{E}_{\text{AD}}(I) &= E_0 I E_0^\dagger + E_1 I E_1^\dagger \\
&= E_0 E_0^\dagger + E_1 E_1^\dagger \\
&= \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix} + \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ \sqrt{\gamma} & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1-\gamma \end{bmatrix} + \begin{bmatrix} \gamma & 0 \\ 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1+\gamma & 0 \\ 0 & 1-\gamma \end{bmatrix} \\
&\neq I
\end{aligned}$$

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

($\star$) 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*:

$$\mathcal{E}_m(\rho) = \sum_k A_{m,k}\,\rho\,A_{m,k}^\dagger, \qquad \sum_k A_{m,k}^\dagger A_{m,k} \;\le\; I.$$

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

($\star\star$) 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 $\|\mathbf{r}\| \leq 1$. 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 $3m \geq 12$, i.e. $m \geq 4$.

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. □

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

($\star$) 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.

$$\chi = \begin{pmatrix} 1-\frac{\gamma}{2}+s & 0 & 0 & \frac{\gamma}{2} \\ 0 & \frac{\gamma}{2} & -\frac{i\gamma}{2} & 0 \\ 0 & \frac{i\gamma}{2} & \frac{\gamma}{2} & 0 \\ \frac{\gamma}{2} & 0 & 0 & 1-\frac{\gamma}{2}-s \end{pmatrix}_{\{I,X,Y,Z\}/\sqrt{2}}, \qquad s = \sqrt{1-\gamma}.$$

This χ is positive semidefinite and satisfies the trace preserving constraint in the normalized Pauli basis. □

Problem 8.1: Lindblad form to quantum operation

($\star$) 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

($\star$) 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

$$\mathcal{R}_m\left(\frac{\tilde{\mathcal{E}}_m(\rho)}{\mathrm{tr}\left[\tilde{\mathcal{E}}_m(\rho)\right]}\right) = \rho,$$

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:

$$\tilde{\rho}_{123}^{(m)} = \frac{(E_m^{(12)} \otimes I_3)\rho_{123}(E_m^{(12)} \otimes I_3)^\dagger}{\mathsf{Tr}[(E_m^{(12)} \otimes I_3)\rho_{123}(E_m^{(12)} \otimes I_3)^\dagger]}.$$

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:

$$\tilde{\mathcal{E}}_m(\rho_1) = \mathsf{Tr}_{12}\left[(E_m^{(12)} \otimes I_3)(\rho_1 \otimes |\Psi\rangle_{23}\langle\Psi|)(E_m^{(12)} \otimes I_3)^\dagger\right].$$

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:

$$\mathcal{R}_m\left(\frac{\tilde{\mathcal{E}}_m(\rho_1)}{\mathsf{Tr}[\tilde{\mathcal{E}}_m(\rho_1)]}\right) = \rho_1,$$

where the division by $\mathsf{Tr}[\tilde{\mathcal{E}}_m(\rho_1)]$ ensures normalization.

□

Problem 8.3: Random unitary channels

($\star\star\star$) 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

$$\mathcal{E}(X) = \frac{1}{2}\mathsf{Tr}(X)I - \frac{1}{2}X^T$$

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

$$\mathcal{E}(X) = A_1 X A_1^\dagger + A_2 X A_2^\dagger + A_3 X A_3^\dagger$$

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:

$$D\left(\sum_i p_i\rho_i, \sigma\right) \leq \sum_i p_i D(\rho_i, \sigma).$$

□

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:

$$D\left(\sum_i p_i\rho_i, \sum_i p_i\sigma_i\right) \leq \sum_i p_i D(\rho_i, \sigma_i),$$

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. □

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:

$$\|A\|_1 = \operatorname{Tr}\sqrt{A^\dagger A}.$$

Let $B := A^\dagger A$, which is a positive operator. Then:

$$\|UAU^\dagger\|_1 = \operatorname{Tr}\sqrt{(UAU^\dagger)^\dagger(UAU^\dagger)} = \operatorname{Tr}\sqrt{UA^\dagger U^\dagger UAU^\dagger} = \operatorname{Tr}\sqrt{UA^\dagger AU^\dagger}.$$

Now apply the fact that for any positive operator B, $\sqrt{UBU^\dagger} = U\sqrt{B}U^\dagger$. So:

$$\|UAU^\dagger\|_1 = \operatorname{Tr}(U\sqrt{B}U^\dagger) = \operatorname{Tr}(\sqrt{B}) = \|A\|_1,$$

using cyclicity of the trace. Hence, $\|UAU^\dagger\|_1 = \|A\|_1$. □

Exercise 9.15

($\star$) Show that

$$F(\rho,\sigma) = \max_{|\varphi\rangle} \left|\langle\psi|\varphi\rangle\right|$$

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

$$\max_{|\varphi\rangle} |\,\langle\psi|\varphi\rangle\,| = \max_{U} \left|\langle\psi|\,(I\otimes U)\,|\tilde{\varphi}\rangle\right|.$$

Uhlmann's theorem states that

$$F(\rho,\sigma) = \max_{U} \left|\langle\psi|\,(I\otimes U)\,|\tilde{\varphi}\rangle\right|,$$

for any fixed purifications $|\psi\rangle$ of ρ and $|\tilde{\varphi}\rangle$ of σ.
Thus, we conclude

$$F(\rho,\sigma) = \max_{|\varphi\rangle} \left|\langle\psi|\varphi\rangle\right|.$$

□

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

$$\mathrm{tr}\Big(A^\dagger B\Big) = \langle m|(A\otimes B)|m\rangle$$

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 $\{|i_Q\rangle\}$. Define the (unnormalized) maximally entangled state:

$$|m\rangle = \sum_{i=1}^{d} |i_R\rangle \otimes |i_Q\rangle \in \mathcal{H}_R \otimes \mathcal{H}_Q.$$

Let A be an operator on R, and B an operator on Q. Then:

$$\begin{aligned}
\langle m|\,(A\otimes B)\,|m\rangle &= \left(\sum_i \langle i_R| \otimes \langle i_Q|\right)(A\otimes B)\left(\sum_j |j_R\rangle \otimes |j_Q\rangle\right) \\
&= \sum_{i,j} \langle i_R|\,A\,|j_R\rangle \cdot \langle i_Q|\,B\,|j_Q\rangle \\
&= \sum_{i,j} A_{ij} B^*_{ij} \\
&= \sum_{i,j} (A^\dagger)_{ji} B_{ij} \\
&= \mathrm{tr}\Big(A^\dagger B\Big).
\end{aligned}$$

□

Remark: This relation plays a key role in the Choi–Jamiołkowski isomorphism and quantum process tomography.

Exercise 9.17

Show that $0 \le A(\rho,\sigma) \le \pi/2$, 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

- $A(\rho, \sigma) = 0 \iff F(\rho, \sigma) = 1 \iff \rho = \sigma$,
- $A(\rho, \sigma) = \dfrac{\pi}{2} \iff F(\rho, \sigma) = 0 \iff$ 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:

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

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,

$$F\left(\sum_i p_i\rho_i, \sigma\right) \geq \sum_i p_i F(\rho_i, \sigma).$$

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,

$$\begin{aligned} F\left(\sum_i p_i\rho_i,\ \sigma\right) &\geq |\langle\Psi|\Phi\rangle| \\ &= \left|\sum_i p_i \langle\psi_i|\varphi\rangle\right| \\ &\geq \sum_i p_i |\langle\psi_i|\varphi\rangle| \\ &= \sum_i p_i F(\rho_i, \sigma). \end{aligned}$$

□

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,

$$1 - F(|\psi\rangle, \sigma)^2 \leq D(|\psi\rangle, \sigma).$$

□

Exercise 9.22: Chaining property for fidelity measures

($\star$) 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

$$E(VU, \mathcal{F}\circ\mathcal{E}) \leq E(U,\mathcal{E}) + E(V,\mathcal{F}).$$

□

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

$$\bar{F} = 1 \quad \Longleftrightarrow \quad \mathcal{E}(\rho_j) = \rho_j \quad \text{for all } j \text{ such that } p_j > 0.$$

($\Rightarrow$) 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,

$$\mathcal{E}(\rho_j) = \rho_j \quad \text{for all } j \text{ with } p_j > 0.$$

($\Leftarrow$) 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

($\star\star$) 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

$$|\operatorname{Tr}(\sqrt{\rho}\sqrt{\sigma}U)| \le \sqrt{\operatorname{Tr}(\rho P)}\,\sqrt{\operatorname{Tr}(\sigma P^{-1})}.$$

Taking the supremum over U yields

$$\operatorname{Tr}\sqrt{\sqrt{\rho}\,\sigma\,\sqrt{\rho}} \;\le\; \sqrt{\operatorname{Tr}(\rho P)\,\operatorname{Tr}(\sigma P^{-1})}.$$

Squaring and infimizing over $P > 0$ gives

$$F(\rho,\sigma) \;\le\; \inf_{P>0} \operatorname{Tr}(\rho P)\operatorname{Tr}\Big(\sigma P^{-1}\Big).$$

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

($\star$) 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

($\star\star$) 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,

$$F(\rho,\mathcal{E}) \;\geq\; 1-\eta-\sum_{k}\sum_{m\neq k}\lambda_k\lambda_m\,\langle m|\mathcal{E}(|k\rangle\langle k|)|m\rangle. \tag{15}$$

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 $k\neq 1$,

$$\sum_{m\neq k}\lambda_k\lambda_m\langle m|\mathcal{E}(|k\rangle\langle k|)|m\rangle \;\leq\; \lambda_k\lambda_1\,\eta. \tag{16}$$

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

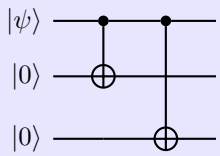
$$F(\rho,\mathcal{E}) \;\geq\; 1-\tfrac{3}{2}\,\eta.$$

□

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. □

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

($\star$) 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):

$$\boxed{F_{\text{min}} = \sqrt{|\psi|\rho\,|\psi\rangle} = \sqrt{\frac{(1-p)^3 + 3p(1-p)^2 + 3p^2(1-p) + p^3}{2}} = \frac{1}{\sqrt{2}}}$$

□

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 $Z_1 Z_2 Z_3$.

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 $Z_1 Z_2 Z_3(|000\rangle \pm |111\rangle) = |000\rangle \pm (-1)^3\,|111\rangle = |000\rangle \mp |111\rangle$, we have:

$$\begin{aligned}&Z_1 Z_2 Z_3\left(\alpha\frac{(|000\rangle-|111\rangle)(|000\rangle+|111\rangle)(|000\rangle+|111\rangle)}{2\sqrt{2}} + \beta\frac{(|000\rangle+|111\rangle)(|000\rangle-|111\rangle)(|000\rangle-|111\rangle)}{2\sqrt{2}}\right) \\ &= \alpha\frac{(|000\rangle+|111\rangle)(|000\rangle+|111\rangle)(|000\rangle+|111\rangle)}{2\sqrt{2}} + \beta\frac{(|000\rangle-|111\rangle)(|000\rangle-|111\rangle)(|000\rangle-|111\rangle)}{2\sqrt{2}} \\ &= |\psi_L\rangle\end{aligned}$$

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 $PE_i^\dagger E_j P$ 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:

$$F(|0\rangle, \mathcal{E}(|0\rangle\langle 0|)) = \sqrt{\langle 0|\left((1-\frac{2p}{3})\,|0\rangle\langle 0| + \frac{2p}{3}\,|1\rangle\langle 1|\right)|0\rangle} = \sqrt{1-\frac{2p}{3}} \tag{17}$$

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:

$$\langle\psi| X |\psi\rangle = \text{Tr}\big(\rho_\psi X\big) = \text{Tr}\left(\frac{X + r_x I + r_y YX + r_z ZX}{2}\right) = r_x$$

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

($\star$) 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 columnn 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:

$$G = \begin{bmatrix} \mathbf{e}_{1,k} & \mathbf{e}_{2,k} & \cdots & \mathbf{e}_{k,k} \\ -\mathbf{w}_1 & -\mathbf{w}_2 & \cdots & -\mathbf{w}_n \end{bmatrix} = \begin{bmatrix} I_k \\ -A \end{bmatrix}$$

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

($\star\star$) 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

$$\frac{k}{n} \geq 1 - H_2\left(\tfrac{2t}{n}\right),$$

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 $\binom{n}{\alpha n} \leq 2^{nH_2(\alpha)}$ (See Ex. 12.27 for the proof of this bound and a related lower bound) and the inequality $B(n,\alpha n) \leq (\alpha n+1)\binom{n}{\alpha n}$ (this follows from taking $i = \alpha n$ in every term of the sum for $B(n,\alpha n)$, and the monotonicity of $\binom{n}{i}$ for $i \leq n/2$), 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 $C \subseteq C^\perp$.

$\boxed{\Longrightarrow}$: 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-NOTs.

Solution

Concepts Involved: Parity Check Matrices, Controlled Operations

By the definition of matrix multiplication:

$$(Hx)_i = \sum_j H_{ij} x_j \tag{18}$$

So, if we want the ith bit in the second register to become $\big|(Hx)_i\big\rangle$, 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

($\star$) 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 $\textbf{CSS}(C_1, C_2)$ in the sense that they have the same error-correcting properties. These codes, which we'll refer to as $\textbf{CSS}_{u,v}(C_1, C_2)$ 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:

$$\begin{aligned} Z^u X^v |x + C_2\rangle &= \frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} (-1)^{u \cdot y} (-1)^{u \cdot (x+y)} |x + y\rangle \\ &= \frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} (-1)^{u \cdot x} |x + y\rangle \\ &= (-1)^{u \cdot x} \sum_{y \in C_2} |x + y\rangle \end{aligned}$$

Up to the phase factor $(-1)^{u \cdot x}$, the states appearing on the RHS are identical to the states defining the code $\textbf{CSS}(C_1, C_2)$ (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 $\textbf{CSS}_{u,v}(C_1, C_2)$ has the same error correcting properties as $\textbf{CSS}(C_1, C_2)$. □

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. □

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. □

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

$\boxed{\Longrightarrow}$: 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

$$r(g) = (a_1, \ldots, a_n \mid b_1, \ldots, b_n) \in \mathbb{F}_2^{2n} \tag{19}$$

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

$$\Lambda = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \tag{20}$$

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,

$$g \text{ and } g' \text{ commute} \iff r(g)\Lambda r(g')^T = 0 \tag{21}$$

□

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.

□

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 $\mathbb{C}^{4\times4}$ over field $\mathbb{C}$, and since $\dim \mathbb{C}^{4\times4} = 16$ they form a basis) we can write any operator $O \in \mathbb{C}^{4\times4}$ 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 $O \in \mathbb{C}^{4\times4}$, which is only possible if $U = V$. □

Exercise 10.39

Verify (10.91).

Solution

Concepts Involved: Unitary Operators

We verify by matrix multiplication:

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

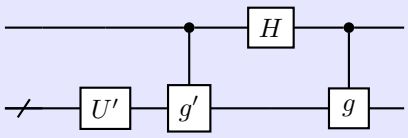
$$SZS^\dagger = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -i \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & (-i)^2 \end{bmatrix} =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

□

Exercise 10.40

($\star\star$) 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.



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 $2, \ldots, n{+}1$. 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 $Z_1 \mapsto Z_1 Z_j$ 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 $X_1 \mapsto X_1 X_j$ 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 $\{H, S, \mathrm{CNOT}\}$. □

Exercise 10.41

Verify Equations (10.92) through (10.95).

Solution

Concepts Involved: Controlled Operations

Toffoli projector form. Let $U = \mathrm{CCNOT}_{1,2\to3}$ 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.

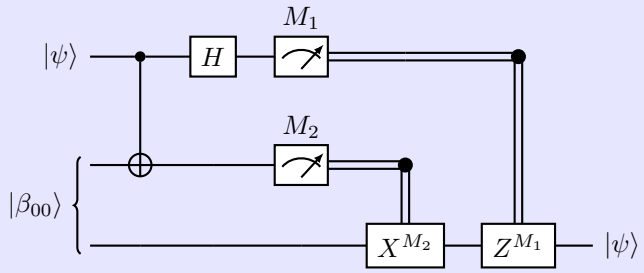
(5) $UZ_3U^\dagger = \dfrac{1}{2}\, Z_3\,(I+Z_1+Z_2-Z_1Z_2)$. We have $\{X_3-I_3, Z_3\} = -2Z_3$ and $(X_3-I_3)Z_3(X_3-I_3) = 0$, and Z_3 commutes with Π. Therefore

$$UZ_3U^\dagger = Z_3 + \{A, Z_3\} + AZ_3A = Z_3 + \Pi \otimes (-2Z_3) + 0 = Z_3\,(I - 2\Pi).$$

Since $I - 2\Pi = \frac{1}{2}\,(I + Z_1 + Z_2 - Z_1Z_2)$, the claim follows. □

Exercise 10.42

($\star\star$) 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.



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

($\star$) 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

$$g_1 = X_1X_2, \qquad g_2 = X_2X_3.$$

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	$ZZIIIIIII$
g_2	$IZZIIIIII$
g_3	$IIIZZIIII$
g_4	$IIIIZZIII$
g_5	$IIIIIIZZI$
g_6	$IIIIIIIZZ$
g_7	$XXXXXXIII$
g_8	$IIIXXXXXX$
$\bar{Z}$	$XXXXXXXXX$
$\bar{X}$	$ZZZZZZZZZ$

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

($\star$) 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

($\star$) 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

($\star$) 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

$$G_Z^{(\mathrm{enc})} = \big[\, 0\ 0\ 0 \ \big|\ A_2^{\mathsf{T}}\ 0\ I_k \,\big].$$

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

($\star$) 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

$$G_Z = \big[\, 0\ 0\ 0 \ \big|\ A_2^{\mathsf{T}}\ 0\ I_k \,\big], \qquad G_X = \big[\, 0\ E^{\mathsf{T}}\ I_k \ \big|\ C^{\mathsf{T}}\ 0\ 0 \,\big].$$

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

($\star$) 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.

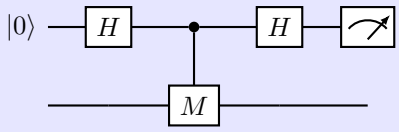


Figure 10.13 (Measuring M).

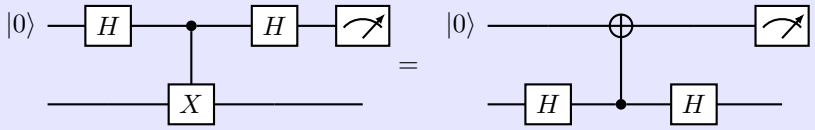


Figure 10.14 (Measuring X).

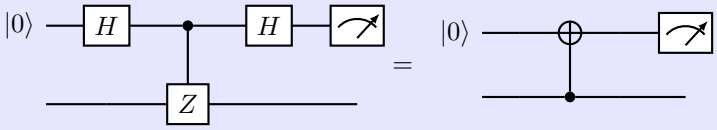


Figure 10.15 (Measuring Z).

Solution

Concepts Involved: Quantum Measurement, Controlled Operations

The verification of Figure 10.13 was already done in Ex. 4.34. For the claimed equivalences:

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

($\star$) 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.

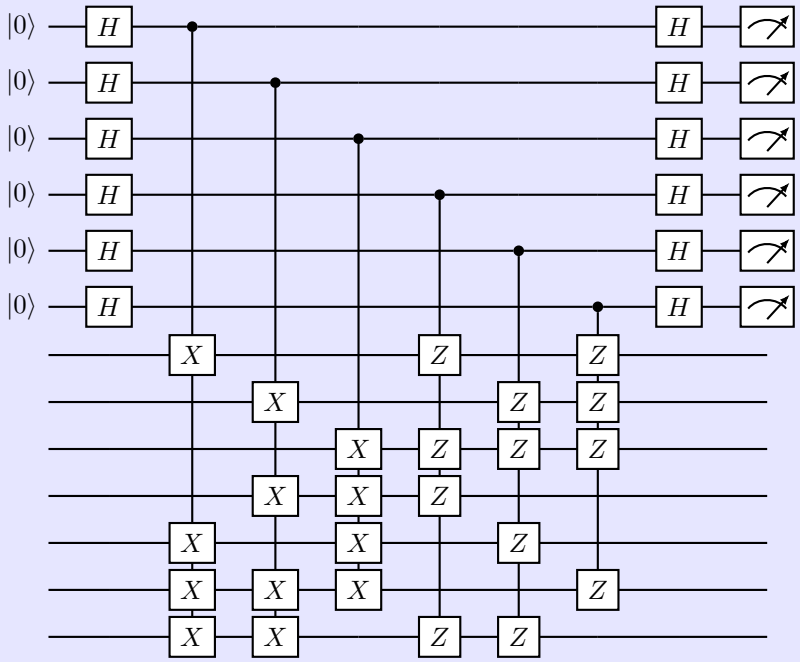


Figure 10.16.

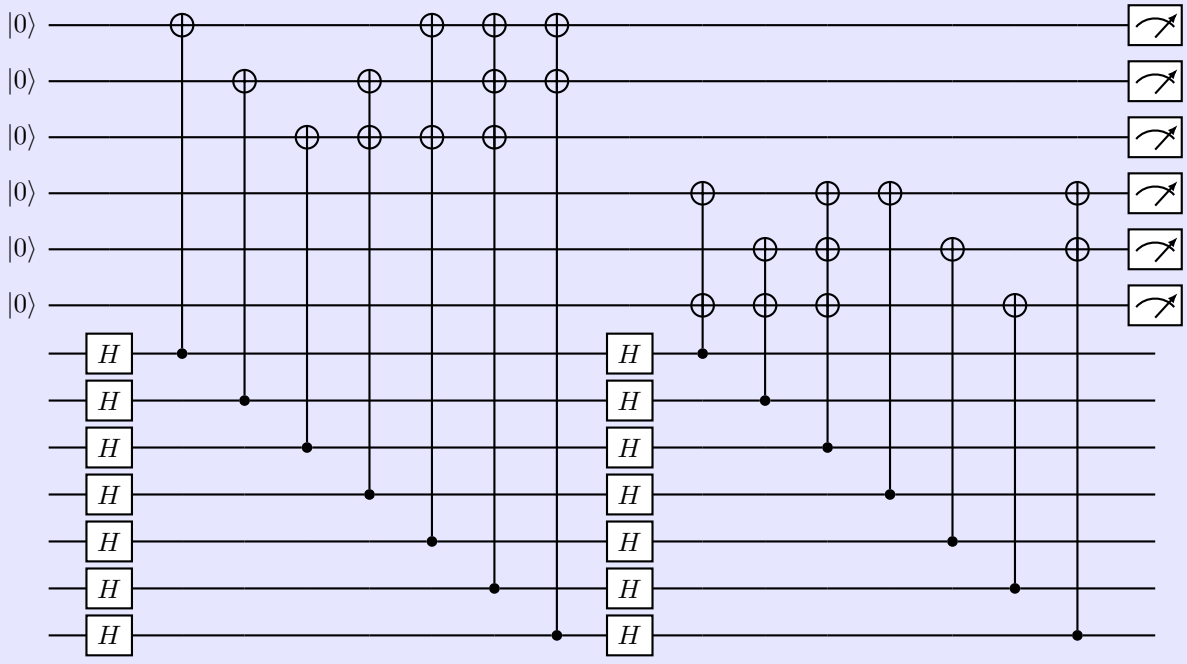


Figure 10.17.

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

($\star$) 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

$$\begin{aligned}
g_1:&\ |+\rangle_a,\ \mathrm{CNOT}_{a\to1},\mathrm{CNOT}_{2\to a},\mathrm{CNOT}_{3\to a},\mathrm{CNOT}_{a\to4}, && \text{measure } X_a,\\
g_2:&\ |+\rangle_a,\ \mathrm{CNOT}_{a\to2},\mathrm{CNOT}_{3\to a},\mathrm{CNOT}_{4\to a},\mathrm{CNOT}_{a\to5}, && \text{measure } X_a,\\
g_3:&\ |+\rangle_a,\ \mathrm{CNOT}_{a\to1},\mathrm{CNOT}_{a\to3},\mathrm{CNOT}_{4\to a},\mathrm{CNOT}_{5\to a}, && \text{measure } X_a,\\
g_4:&\ |+\rangle_a,\ \mathrm{CNOT}_{1\to a},\mathrm{CNOT}_{a\to2},\mathrm{CNOT}_{a\to4},\mathrm{CNOT}_{5\to a}, && \text{measure } X_a.
\end{aligned}$$

□

Exercise 10.61

($\star$) 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):

$$E^\dagger = \begin{cases} I, & \text{if } s_X = 000,\ s_Z = 000, \\ X_j, & \text{if } s_X = \mathrm{col}_j(H),\ s_Z = 000, \\ Z_j, & \text{if } s_X = 000,\ s_Z = \mathrm{col}_j(H), \\ Y_j, & \text{if } s_X = \mathrm{col}_j(H) = s_Z. \end{cases}$$

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 $\to$ qubit map. The 3-bit syndromes correspond to the columns of H as follows

j	1	2	3	4	5	6	7
$\mathrm{col}_j(H)$	$\begin{bmatrix}1\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\\1\end{bmatrix}$

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):

$$E^\dagger = \begin{cases} I, & \text{if } s_X = 000,\ s_Z = 000, \\ X_j, & \text{if } s_X = \mathrm{col}_j(H),\ s_Z = 000, \\ Z_j, & \text{if } s_X = 000,\ s_Z = \mathrm{col}_j(H), \\ Y_j, & \text{if } s_X = \mathrm{col}_j(H) = s_Z. \end{cases}$$

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 $\to$ qubit map. The 3-bit syndromes correspond to the columns of H as follows:

j	1	2	3	4	5	6	7
$\mathrm{col}_j(H)$	$\begin{bmatrix}1\\0\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\\1\end{bmatrix}$	$\begin{bmatrix}0\\1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\\1\end{bmatrix}$

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

($\star\star$) 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 $S_{\mathrm{out}} \subset \mathcal{P}_{n_1}$ and chosen logical Paulis $\overline{X}_{\mathrm{out}}, \overline{Z}_{\mathrm{out}} \in N(S_{\mathrm{out}}) \setminus S_{\mathrm{out}}$ and C_{in} be an $[n_2, 1]$ stabilizer code with stabilizer $S_{\mathrm{in}} \subset \mathcal{P}_{n_2}$ and logical Paulis $\overline{X}_{\mathrm{in}}, \overline{Z}_{\mathrm{in}} \in N(S_{\mathrm{in}}) \setminus S_{\mathrm{in}}$.
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"

$$\varphi : \mathcal{P}_{n_1} \longrightarrow \mathcal{P}_{n_1 n_2}, \qquad \varphi\left(\bigotimes_{j=1}^{n_1} P_j\right) = \bigotimes_{j=1}^{n_1} \overline{P_j}_{\mathrm{in}}^{(j)},$$

where $\overline{I}_{\mathrm{in}} := I^{\otimes n_2}$, $\overline{X}_{\mathrm{in}}, \overline{Z}_{\mathrm{in}}$ are the inner logicals, $\overline{Y}_{\mathrm{in}} := i\,\overline{X}_{\mathrm{in}}\overline{Z}_{\mathrm{in}}$, and the superscript $^{(j)}$ indicates action on block j only. Because $\overline{X}_{\mathrm{in}}, \overline{Z}_{\mathrm{in}}$ 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:

$$U\left|0_L\right\rangle \mapsto \frac{\left|0_L\right\rangle + \left|1_L\right\rangle}{\sqrt{2}}, \quad U\left|1_L\right\rangle \mapsto \frac{\left|0_L\right\rangle - \left|1_L\right\rangle}{\sqrt{2}}$$

up to a possible global phase. □

Exercise 10.64: Back propogation of errors

($\star$) 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. $U\left|x, y\right\rangle = \left|x,\, y \oplus x\right\rangle$.

(A) Stabilizer / Heisenberg-picture propagation. We simply compute Pauli conjugations

$$\begin{aligned} UX_cU^\dagger \left|x, y\right\rangle &= UX_c\left|x,\, y\oplus x\right\rangle = U\left|x\oplus 1,\, y\oplus x\right\rangle \\ &= \left|x\oplus 1,\, (y\oplus x)\oplus(x\oplus 1)\right\rangle = \left|x\oplus 1,\, y\oplus 1\right\rangle = (X_cX_t)\left|x, y\right\rangle, \end{aligned}$$

so $UX_cU^\dagger = X_cX_t$ (a control-X propagates to the target).

$$\begin{aligned} UZ_tU^\dagger \left|x, y\right\rangle &= U\big((-1)^y\left|x,\, y\oplus x\right\rangle\big) = (-1)^{y\oplus x}\left|x, y\right\rangle \\ &= (-1)^x(-1)^y\left|x, y\right\rangle = (Z_cZ_t)\left|x, y\right\rangle, \end{aligned}$$

so $UZ_tU^\dagger = Z_cZ_t$ (a target-Z kicks back to the control).
Thus the CNOT conjugation table is

$$X_c \mapsto X_cX_t, \quad Z_c \mapsto Z_c, \quad X_t \mapsto X_t, \quad Z_t \mapsto Z_cZ_t.$$

(B) Direct circuit identities (projector form). Use

$$\mathrm{CNOT}_{c\to t} = \left|0\right\rangle\!\left\langle 0\right|_c \otimes I_t + \left|1\right\rangle\!\left\langle 1\right|_c \otimes X_t.$$

Control-X propagation.

$$\begin{aligned} (X_c\otimes I_t)\,\mathrm{CNOT} &= (X\otimes I)\big(\left|0\right\rangle\!\left\langle 0\right|\otimes I + \left|1\right\rangle\!\left\langle 1\right|\otimes X\big) \\ &= \left|1\right\rangle\!\left\langle 0\right|\otimes I + \left|0\right\rangle\!\left\langle 1\right|\otimes X \\ &= \big(\left|0\right\rangle\!\left\langle 0\right|\otimes I + \left|1\right\rangle\!\left\langle 1\right|\otimes X\big)(X\otimes X) \\ &= \mathrm{CNOT}\,(X_c\otimes X_t). \end{aligned}$$

Hence X_c pushed through CNOT picks up an X_t.

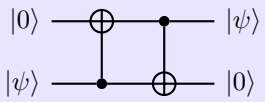
Target-Z back-propagation.

$$\begin{aligned}(I_c\otimes Z_t)\,\mathrm{CNOT} &= (I\otimes Z)\big(\,|0\rangle\langle 0|\otimes I + |1\rangle\langle 1|\otimes X\big)\\ &= |0\rangle\langle 0|\otimes Z + |1\rangle\langle 1|\otimes ZX\\ &= |0\rangle\langle 0|\otimes Z - |1\rangle\langle 1|\otimes XZ\\ &= \big(\,|0\rangle\langle 0|\otimes I + |1\rangle\langle 1|\otimes X\big)\,(Z_c\otimes Z_t)\\ &= \mathrm{CNOT}\,(Z_c\otimes Z_t),\end{aligned}$$

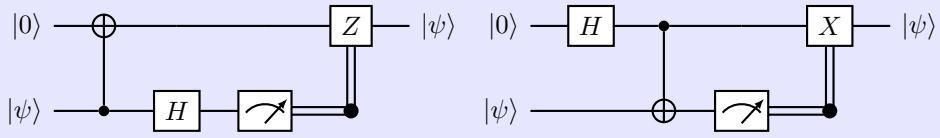
where we used $ZX = -XZ$ and $Z_c\,|0\rangle\langle 0| = +\,|0\rangle\langle 0|$, $Z_c\,|1\rangle\langle 1| = -\,|1\rangle\langle 1|$. Thus pushing Z_t through CNOT produces an extra Z_c on the control.
Both derivations establish X on the control spreads to the target, and Z on the target kicks back 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 $m \in \{0, 1\}$:

$$m = 0 : \ \alpha |0\rangle_A + \beta |1\rangle_A = |\psi\rangle_A , \qquad m = 1 : \ \beta |0\rangle_A + \alpha |1\rangle_A = X |\psi\rangle_A .$$

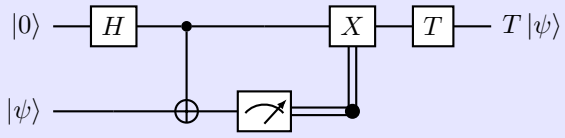
Applying X^m on A gives $|\psi\rangle_A$ deterministically. Again, B is known (in $|m\rangle$) and can be reset to $|0\rangle$. Hence, in both circuits the net effect is

$$|0\rangle_A |\psi\rangle_B \longmapsto |\psi\rangle_A |0\rangle_B ,$$

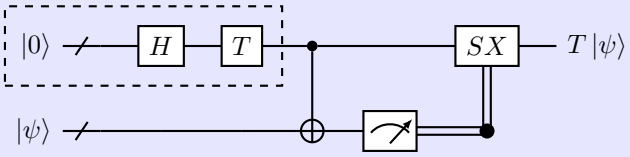
i.e., the unknown state is transferred to the top wire using one CNOT plus measurement and a classically controlled Pauli. □

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.



There is an error in the problem statement, and the corrected first relation is that $TX = e^{-i\pi/4}SXT$.

Solution

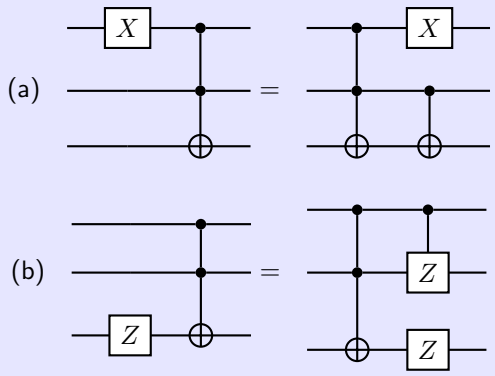
Concepts Involved: Quantum Measurement, Controlled Operations

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 $|a, b, c\rangle$:

$$\text{LHS}: \quad UX_1\,|a, b, c\rangle = U\,|a \oplus 1, b, c\rangle = \big|a \oplus 1, b,\ c \oplus (a \oplus 1)b\big\rangle\,.$$

$$\text{RHS}: \quad X_1\,\mathrm{CX}_{2\to 3}\,U\,|a, b, c\rangle = X_1\,\mathrm{CX}_{2\to 3}\,|a, b, c \oplus ab\rangle = X_1\,|a, b,\ c \oplus ab \oplus b\rangle = |a \oplus 1, b,\ c \oplus ab \oplus b\rangle\,.$$

Since $(a \oplus 1)b = ab \oplus b$, the third components coincide; hence LHS = RHS for all a, b, c, so

$$UX_1 \;=\; X_1\,\mathrm{CX}_{2\to 3}\,U.$$

(b) Z on the target.

$$\text{LHS}: \quad UZ_3\,|a, b, c\rangle = U\big((-1)^c\,|a, b, c\rangle\big) = (-1)^c\,|a, b,\ c \oplus ab\rangle\,.$$

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

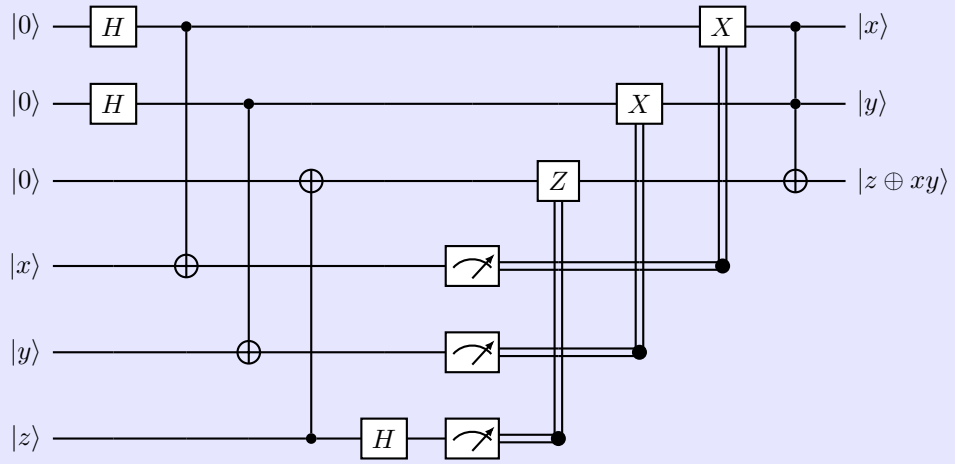
$$UZ_3 \;=\; Z_3\,\mathrm{CZ}_{1,2}\,U.$$

□

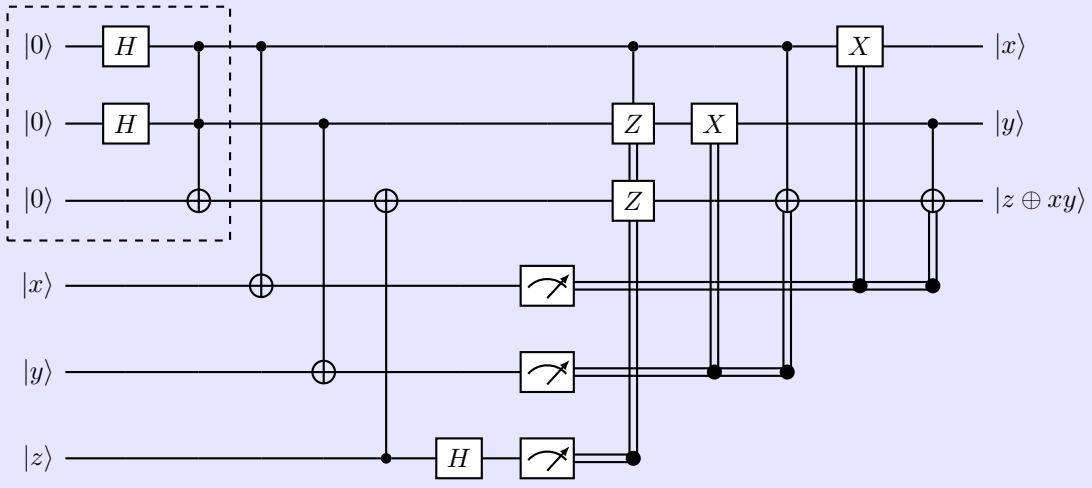
Exercise 10.68: Fault-tolerant Toffoli gate construction

($\star$) 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 $\mathrm{CNOT}_{a\to i}$,

$$\mathrm{CNOT}_{a\to i}\, Z_a\, \mathrm{CNOT}_{a\to i}^{\dagger} = Z_a.$$

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

($\star$) 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

$$M = e^{-i\pi/4}\,SX \;=\; e^{-i(\pi/4)Z}X \;=\; \tfrac{1}{\sqrt{2}}(X+Y). \tag{22}$$

Thus $M^\dagger = M$ and $M^2 = I$, so the eigenvalues are ± 1.

Controlled-M. Using projector decompositions,

$$\mathrm{CNOT}_{a\to d} = \big(|0\rangle\langle 0|_a \otimes I_d\big) + \big(|1\rangle\langle 1|_a \otimes X_d\big), \tag{23}$$
$$\mathrm{CS}_{a\to d} = \big(|0\rangle\langle 0|_a \otimes I_d\big) + \big(|1\rangle\langle 1|_a \otimes S_d\big), \tag{24}$$
$$\mathrm{CZ}_{a\to d} = \big(|0\rangle\langle 0|_a \otimes I_d\big) + \big(|1\rangle\langle 1|_a \otimes Z_d\big), \tag{25}$$
$$T_a = |0\rangle\langle 0|_a + e^{i\pi/4}\,|1\rangle\langle 1|_a. \tag{26}$$

Hence

$$T_a\,\mathrm{CZ}_{a\to d}\,\mathrm{CS}_{a\to d}\,\mathrm{CNOT}_{a\to d} = |0\rangle\langle 0|_a \otimes I_d \;+\; |1\rangle\langle 1|_a \otimes \big(e^{i\pi/4}ZSX\big)_d \tag{27}$$

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 $\bar{Y} = i\bar{X}\bar{Z} = iX^{\otimes 7}Z^{\otimes 7} = i(-iY)^{\otimes 7} = i(-i)^7Y^{\otimes 7} = -Y^{\otimes 7}$. So indeed this is the correct action.

Measurement scheme. Prepare a verified m-qubit cat ancilla

$$|\mathrm{cat}_m\rangle \;=\; \tfrac{1}{\sqrt{2}}\big(|0^m\rangle + |1^m\rangle\big), \tag{28}$$

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

$$U \;=\; \prod_j \mathrm{C}M_{a_j\to d_j} \;=\; |0^m\rangle\langle 0^m| \otimes I \;+\; |1^m\rangle\langle 1^m| \otimes M_L, \tag{29}$$

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,

$$[Z_a, \mathrm{CNOT}_{a\to d}] = 0, \qquad [Z_a, \mathrm{CS}_{a\to d}] = 0, \tag{30}$$

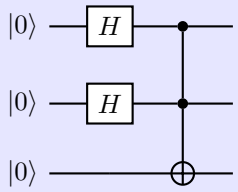
so an ancilla Z never propagates to data and only flips the X-basis readout. Ancilla X/Y faults can propagate to the *paired* data qubit but to no others. Verified cat preparation bounds non-Z ancilla faults to at most one, and repeating the measurement with majority vote suppresses readout flips. Hence the procedure both measures M correctly and satisfies the single-fault–single-data-error criterion. □

Exercise 10.72: Fault-tolerant Toffoli ancilla state construction

($\star\star$) 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

Computing the stabilizers explicitly from the given circuit, we find:

$$\langle Z_1, Z_2, Z_3\rangle \overset{H_1\otimes H_2}{\to} \langle X_1, X_2, Z_3\rangle \overset{CCX}{\to} \langle X_1 CNOT_{2\to 3}, X_2 CNOT_{1\to 3}, Z_3 CZ_{1\to 2}\rangle$$

where in the last step we use Ex. 10.67.
The three (mutually commuting) stabilizers of the state are non-Pauli, but can still be measured fault-tolerantly. To see this, let us work with the Steane code, and within that encoding we notice that all three can be applied transversally as gates (as Paulis and controlled-NOTs/Controlled-Zs are transversal). Thus we can use the previous construction of controlled operations on a verified cat state & majority vote of a parity measurement to fault-tolerantly measure all three sequentially. If we measure the $+1$ outcome for all three, we have precisely the desired stabilizer state and are done. If we measure the -1 outcome, we can apply an encoded $Z_1/Z_2/X_3$ operation (transversally, and thus fault-tolerantly) to correct the state to be a $+1$ eigenvalue of the measured operator(s).

□

Exercise 10.73: Fault-tolerant encoded state construction

($\star$) 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.

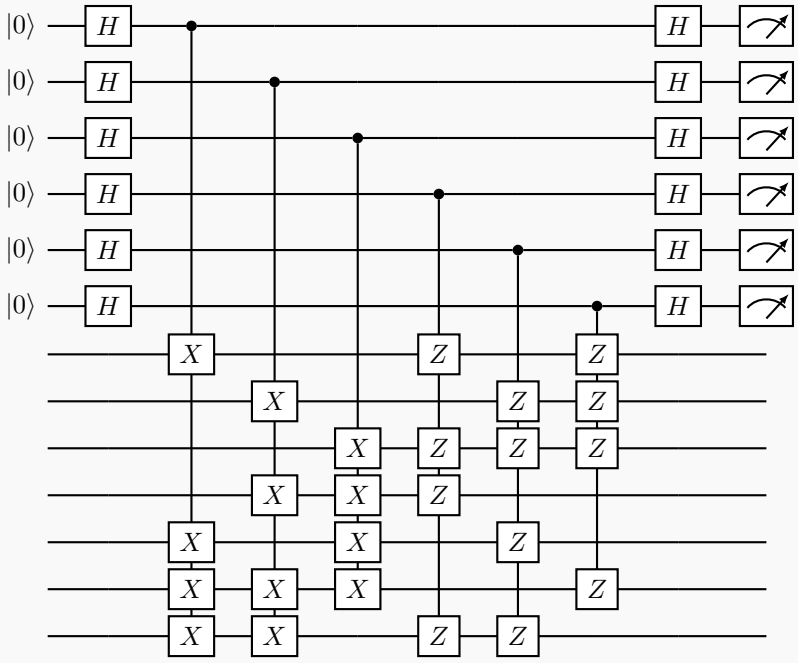
(3) Calculate the error probability of this circuit, and of the circuit when the generator measurements are repeated three times and majority voting is done.

(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: Fault Tolerance

(1) The circuit is too large to draw in full, but simply replace each (of the 6) ancilla in the circuit of Fig 10.16 (depicted below)

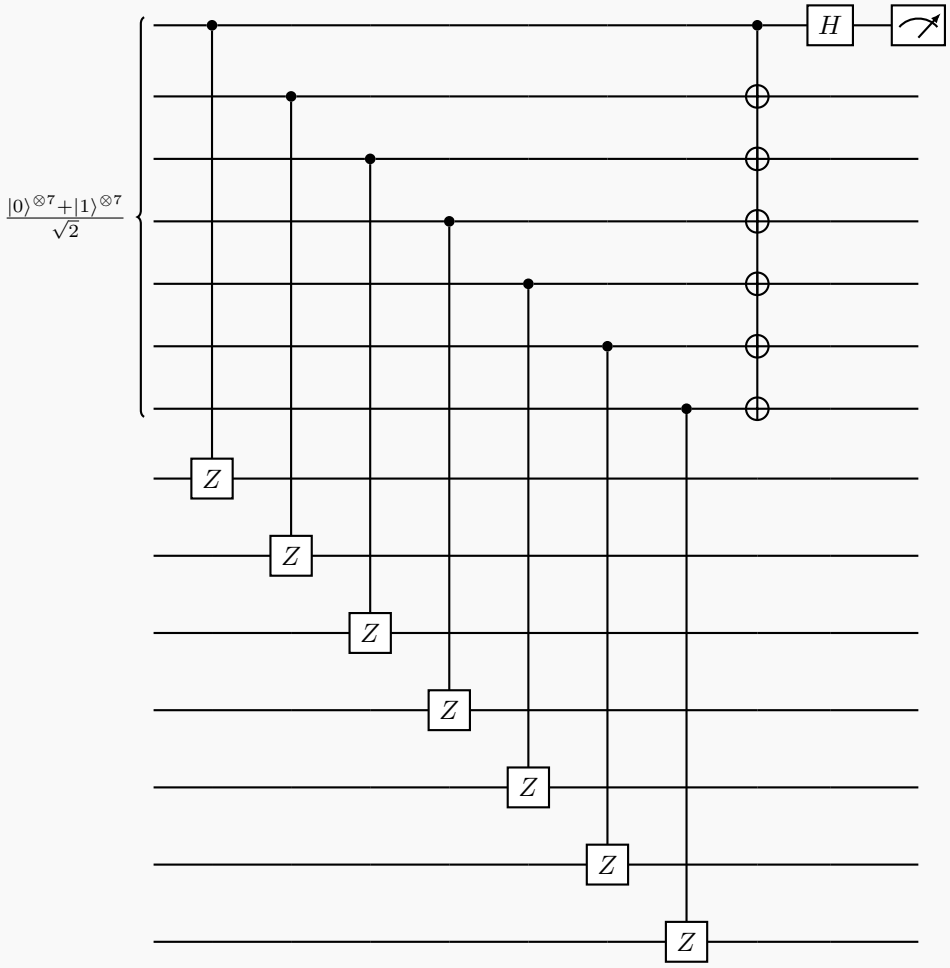


with a 4-qubit cat state $\frac{|0000\rangle+|1111\rangle}{\sqrt{2}}$, replacing each $CX^{\otimes 4}/CZ^{\otimes 4}$ gatex with 4 CX/CZ between ancilla/physical qubit pairs + $CX^{\otimes 3}$ across the ancilla.

(2) Recalling that:

$$\bar{Z} = Z_1 Z_2 Z_3 Z_4 Z_5 Z_6 Z_7$$

we can fault tolerantly measure this via:



appending this to the circuit at the end of the previous step.

(3) Due to the controls being on different qubits, errors do not propagate in the code block (recall Ex. 10.69 and 10.70). Hence, each component has error probability $O(p)$, which can be independently summed to yield $O(p)$ error probability for each stabilizer generator measurement and of the whole circuit. Repeating the measurements three times and taking a majority vote, the error is reduced to $O(p^2)$.

(4) The recovery operations corresponding to different error syndromes/generator measurement outcomes is provided in Ex. 10.61. The recovery operations are simply single-qubit Paulis (which do not propagate). Since the Steane code can tolerate single-qubit errors, this recovery is fault-tolerant.

□

Exercise 10.74

($\star$) 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: Fault Tolerance

The procedure is identical to the steps of Ex. 10.73.

(1) Simply repeat the procedure with the above Steane code circuit with the generators $g_1 = XZZXI, g_2 = IZZX, g_3 = XIXZZ, g_4 = ZXIXZ$.

(2) Again, repeat the above procedure, with $\bar{Z} = ZZZZZ$.

(3) The error probability of the circuit is $O(p)$, and $O(p^2)$ with repeated rounds + majority vote, as above.

(4) Studying what single-qubit errors anticommute with the above stabilizer generators, for syndrome $g_1g_2g_3g_4$ (with $g_i = 0$ if measured to be $+1$, $g_i = 1$ if measured to be -1), we get the table:

Syndrome	Correction Operator
0000	I
0001	X_1
1000	X_2
1100	X_3
0110	X_4
0011	X_5
1010	Z_1
0101	Z_2
0010	Z_3
1001	Z_4
0100	Z_5
1011	Y_1
1101	Y_2
1110	Y_3
1111	Y_4
0111	Y_5

.

As before, recovery is fault-tolerant as the correction operations act on single qubits.

□

Problem 10.1

($\star$) 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:

$$\begin{aligned} C'_m P'_m \mathcal{E}_2(\rho) P'_m C'^\dagger_m &= (V^\dagger C_m U^\dagger)(UP_mU^\dagger)U\mathcal{E}_1(V\rho V^\dagger)U^\dagger(UP_mU^\dagger)(V^\dagger C_m U^\dagger)^\dagger \\ &= V^\dagger C_m U^\dagger UP_mU^\dagger U\mathcal{E}_1(V\rho V^\dagger)U^\dagger UP_mU^\dagger UC_m^\dagger V \\ &= V^\dagger C_m P_m \mathcal{E}_1(V\rho V^\dagger) P_m C_m^\dagger V \\ &\propto V^\dagger V \rho V^\dagger V \\ &= \rho \end{aligned}$$

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

($\star\star\star$) 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

$$\log_2 V(n, 2t) \;\leq\; k_1 \;\leq\; n - \log_2 V(n, 2t). \tag{31}$$

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

$$\Pr\left[\exists\, w \in \mathcal{W} \cap C_2^\perp\right] \;\leq\; V(n, 2t)\, 2^{-k_2},$$

so there exists $C_2 \subseteq C_1$ with $\mathrm{dist}(C_2^\perp) \geq 2t+1$ whenever

$$k_2 \;\geq\; \log_2 V(n, 2t). \tag{32}$$

Combining (31)–(32), 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

($\star\star\star$) 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

$$G = \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & I \end{array}\right]$$

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: Stabilizer Codes, Logicals

We present a simplified version of Gottesman's construction, given in 4.2 of arXiv:quant-ph/9705052 - the construction there is a bit more involved as the $\bar{X}$ logicals are encoded explicitly - here, we get a little more for free.

The goal is to start from the unencoded $|0\rangle^{\otimes n}$ state and to prepare the logical $\left|\bar{\mathbf{0}}\right\rangle = \left|\bar{0}_1 \dots \bar{0}_k\right\rangle$ state, which is a $+1$ eigenstate/stabilizer state of stabilizers $g_1 \dots g_{n-k}$ and Z-logicals $\bar{Z}_1 \dots \bar{Z}_k$, with the stabilizers and logicals in standard form:

$$\begin{array}{ccc|ccc} \overbrace{I}^{r} & \overbrace{A_1}^{n-k-r} & \overbrace{A_2}^{k} & \overbrace{B}^{r} & \overbrace{0}^{n-k-r} & \overbrace{C}^{k} & \\ 0 & 0 & 0 & D & I & E & \} r \\ 0 & 0 & 0 & A_2^T & 0 & I & \} n-k-r \\ & & & & & & \} k \end{array}$$

Since all stabilizers and logicals commute, the encoding network can be applied in any order. To this end, note that the last $n-k-r$ stabilizers and all of the $\bar{Z}$ logicals are purely Z-type in standard form. Thus, the initial state $|0\rangle^{\otimes n}$ is already stabilized by these operators, so for these we are already done.

The only nontrivial aspect of the encoding is thus the first r stabilizers. These stabilizers all have X in the ith position (from the I in the top left $r \times r$ block of the standard form stabilizer matrix), so the first step of the network is to apply a Hadamard H to each of the $i = 1, \dots r$ qubits. Then, for each $i = 1, \dots r$ let g_i' be the stabilizer g_i restricted to qubits $[1, \dots i-1] \cup [i+1, \dots n]$ (leaving off qubit i), and to the X-type part of the stabilizer. Using qubit i as control, apply the operation $C(g_i')$ (which can be decomposed into individual CNOT gates, one for each X appearing in g_i'). The ith qubit can be ignored as the X_i term in g_i was already implemented by the Hadamard step. The Z-type part of the stabilizer can be ignored in this step as again, $|0\rangle^{\otimes n}$ is already stabilized by it (even if we had included the controlled-Z gates for the B, C terms, the action would have been trivial).

That this step encodes the g_i stabilizers can be easily be seen from (taking $i = 1$ as an example):

$$\begin{aligned} g_1 C_{1\to[2,\dots,n]}(g_1') H_1 |0\rangle^{\otimes n} &= g_1' X_1 C_{1\to[2,\dots,n]}(g_1') |+\rangle_1 \otimes |0\rangle^{\otimes n-1} \\ &= g_1' g_1' C_{1\to[2,\dots,n]}(g_1') X_1 |+\rangle_1 \otimes |0\rangle^{\otimes n-1} \\ &= C_{1\to[2,\dots,n]}(g_1') |+\rangle_1 \otimes |0\rangle^{\otimes n-1} \\ &= C_{1\to[2,\dots,n]}(g_1') H_1 |0\rangle^{\otimes n} \end{aligned}$$

where we have used the familiar result that commuting the X_1 across the control toggles the gate and produces g_1', which cancels with the g_1' being multiplied.

Thus this step encodes the g_i stabilizers for $i = 1, \dots r$, and the resulting state is stabilized by all stabilizer generators and Z-logicals. Thus our encoding of $\left|\bar{\mathbf{0}}\right\rangle = \left|\bar{0}_1 \dots \bar{0}_k\right\rangle$ state is complete! We may then apply

the X-logicals $\bar{X}_1, \ldots, \bar{X}_k$ according to their standard form $\left[0E^T I|C^T 00\right]$ to this encoded state to obtain any of the basis states $|\bar{c}_1 \ldots \bar{c}_k\rangle$, i.e.:

$$|\bar{c}_1 \ldots \bar{c}_k\rangle = \bar{X}_1^{c_1} \ldots \bar{X}_k^{c_k} \left|\bar{0}_1 \ldots \bar{0}_k\right\rangle.$$

□

Problem 10.4: Encoding by teleportation

($\star\star\star$) 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

$$|\Phi^+_{1\text{-}L}\rangle \;=\; \frac{|0\rangle|0_L\rangle + |1\rangle|1_L\rangle}{\sqrt{2}}. \tag{33}$$

On $q_0 \cup B$, this is the unique $+1$ common eigenstate of the commuting set

$$\mathcal{S}_\Phi \;=\; \left\langle\, S_1, \ldots, S_{n-1},\, Z_{q_0}\bar{Z}_B,\, X_{q_0}\bar{X}_B \,\right\rangle. \tag{34}$$

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 $X_\psi X_{q_0}$.

$$\text{Record } m_Z, m_X \in \{0,1\} \text{ such that } Z_\psi Z_{q_0} = (-1)^{m_Z}, \quad X_\psi X_{q_0} = (-1)^{m_X}. \tag{35}$$

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,

$$|\psi\rangle_\psi \otimes |\Phi^+_{1\text{-}L}\rangle_{q_0,B} \xrightarrow{\text{measure } ZZ,XX \text{ on } (\psi,q_0)} \bar{X}^{\,m_X} \bar{Z}^{\,m_Z} |\psi\rangle_L \text{ on } B. \tag{36}$$

Thus apply the logical correction on B

$$(m_Z, m_X) = (0,0) \Rightarrow I, \quad (1,0) \Rightarrow \bar{Z}, \quad (0,1) \Rightarrow \bar{X}, \quad (1,1) \Rightarrow \bar{Z}\bar{X}, \tag{37}$$

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:

$$P_{\text{fail}} \;=\; C\,p^2 + O(p^3), \tag{38}$$

where C counts malignant fault pairs (gadget- and code-dependent). For general distance d with gadgets respecting distance, letting $t = \lfloor (d-1)/2 \rfloor$,

$$P_{\text{fail}} \;=\; \tilde{C}\,p^{\,t+1} + O\Big(p^{\,t+2}\Big)\,. \tag{39}$$

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

($\star\star\star$) 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: Fault Tolerance, Stabilizer Codes, Logicals, Controlled Operations

If $C(S)$ is a CSS code, then a bitwise application of controlled-NOTs between every qubit of the code block suffices (as discussed in Section 10.6.2 of the text, for the Steane code), as this is transversal (and thus fault-tolerant), transforms the logicals in the desired fashion, and preserves the stabilizers, i.e. simply permutes elements of $S \times S$ where S is the stabilizer group of one of the code blocks. However, this does *not* work for all $[n, 1]$ codes - for example take the 5-qubit code presented in the text, which has stabilizer generator $g_1 = XZZXI$. A bit-wise CNOT across two logical qubits takes $g_1 \otimes IIIII \mapsto g_1 \otimes XIIXI$, but $XIIXI$ is not a stabilizer, so this is not a valid operation that preserves the codespace.
We instead follow the construction of Gottesman (which can be found in 5.5 of arXiv:quant-ph/9705052). Consider the four-qubit operation U that maps:

$$XIII \overset{U}{\mapsto} XXXI$$
$$IXII \overset{U}{\mapsto} IXXX$$
$$IIXI \overset{U}{\mapsto} XIXX$$
$$IIIX \overset{U}{\mapsto} XXIX$$
$$ZIII \overset{U}{\mapsto} ZZZI$$
$$IZII \overset{U}{\mapsto} IZZZ$$
$$IIZI \overset{U}{\mapsto} ZIZZ$$
$$IIIZ \overset{U}{\mapsto} ZZIZ$$

U is a normalizer operation on 4-qubits, as it maps Pauli strings to other Pauli strings.
Now, consider any four blocks of any stabilizer code, and apply U bitwise across the four blocks. Then, the stabilizers g of the code map as:

$$g \otimes I \otimes I \otimes I \overset{U^{\otimes n}}{\mapsto} g \otimes g \otimes g \otimes I$$
$$I \otimes g \otimes I \otimes I \overset{U^{\otimes n}}{\mapsto} I \otimes g \otimes g \otimes g$$
$$I \otimes I \otimes g \otimes I \overset{U^{\otimes n}}{\mapsto} g \otimes I \otimes g \otimes g$$
$$I \otimes I \otimes I \otimes g \overset{U^{\otimes n}}{\mapsto} g \otimes g \otimes I \otimes g$$

here $I = I^{\otimes n}$ on the code block. $U^{\otimes n}$ thus preserves the stabilizer (it permutes elements of $S \times S \times S \times S$), and consists of normalizer gates applied transversally (bitwise across code blocks). This operation works for any stabilizer code encoding a single qubit, and furthermore, is fault-tolerant (under the assumption

that the code can correct errors on a single qubit).
Now, suppose we fault-tolerantly prepare the single qubit logical states $|\bar{\psi}\rangle_1$, $|\bar{\varphi}\rangle_2$ of $C(S)$, and further fault-tolerantly prepare the $+1$ $\bar{Z}$ eigenstate $|\bar{0}\rangle_3$, $|\bar{0}\rangle_4$ of $C(S)$ on two logical ancilla. Then, we may consider the logicals:

$$\begin{aligned}
\bar{X}_1 &= \bar{X} \otimes I \otimes I \otimes I \\
\bar{X}_2 &= I \otimes \bar{X} \otimes I \otimes I \\
\bar{Z}_1 &= \bar{Z} \otimes I \otimes I \otimes I \\
\bar{Z}_2 &= I \otimes \bar{Z} \otimes I \otimes I \\
\bar{Z}_3 &= I \otimes I \otimes \bar{Z} \otimes I \\
\bar{Z}_4 &= I \otimes I \otimes I \otimes \bar{Z}
\end{aligned}$$

noting that $\bar{Z}_3, \bar{Z}_4$ stabilize the state. Then, applying $U^{\otimes n}$, these map to:

$$\begin{aligned}
\bar{X}_1 &\overset{U^{\otimes n}}{\mapsto} \bar{X}_1' = \bar{X} \otimes \bar{X} \otimes \bar{X} \otimes I \\
\bar{X}_2 &\overset{U^{\otimes n}}{\mapsto} \bar{X}_2' = I \otimes \bar{X} \otimes \bar{X} \otimes \bar{X} \\
\bar{Z}_1 &\overset{U^{\otimes n}}{\mapsto} \bar{Z}_1' = \bar{Z} \otimes \bar{Z} \otimes \bar{Z} \otimes I \\
\bar{Z}_2 &\overset{U^{\otimes n}}{\mapsto} \bar{Z}_2' = I \otimes \bar{Z} \otimes \bar{Z} \otimes \bar{Z} \\
\bar{Z}_3 &\overset{U^{\otimes n}}{\mapsto} \bar{Z}_3' = \bar{Z} \otimes I \otimes \bar{Z} \otimes \bar{Z} \\
\bar{Z}_4 &\overset{U^{\otimes n}}{\mapsto} \bar{Z}_4' = \bar{Z} \otimes \bar{Z} \otimes I \otimes \bar{Z}
\end{aligned}$$

Logicals can be mapped to equivalent logicals via multiplication by a stabilizer, so in anticipation of the next step, let us multiply $\bar{Z}_1', \bar{Z}_2'$ by the stabilizers $\bar{Z}_3', \bar{Z}_4'$ as follows:

$$\begin{aligned}
\bar{Z}_1' \to \bar{Z}_1'' &= \bar{Z}_1' \bar{Z}_3' \bar{Z}_4' = \bar{Z} \otimes I \otimes I \otimes I \\
\bar{Z}_2' \to \bar{Z}_2'' &= \bar{Z}_2' \bar{Z}_3' = \bar{Z} \otimes \bar{Z} \otimes I \otimes I
\end{aligned}$$

Now, let us fault-tolerantly measure $\bar{X}_3 = I \otimes I \otimes \bar{X} \otimes I$ and $\bar{X}_4 = I \otimes I \otimes I \otimes \bar{X}$, i.e. measure the ancilla qubits in the logical X-basis (which is permitted by assumption). We can see that this measurement commutes with all of $\bar{X}_1', \bar{X}_2', \bar{Z}_1'', \bar{Z}_2''$ and hence these logicals are preserved under this measurement. This justifies the previous step, as $\bar{Z}_1', \bar{Z}_2'$ before stabilizer multiplication had $\bar{Z}$ terms on the ancilla qubits, which would not commute with the measurement. If we then discard the ancillas, the logical $\bar{X}/\bar{Z}$ on (logical) qubits $1/2$ look like:

$$\begin{aligned}
\bar{X}_1' &= \bar{X} \otimes \bar{X} \\
\bar{X}_2' &= I \otimes \bar{X} \\
\bar{Z}_1' &= \bar{Z} \otimes I \\
\bar{Z}_2' &= \bar{Z} \otimes \bar{Z}
\end{aligned}$$

So the net action of this procedure has been the following map on the logical Pauli operators of the two

qubits:

$$\bar{X} \otimes I \to \bar{X} \otimes \bar{X}$$
$$I \otimes \bar{X} \to I \otimes \bar{X}$$
$$\bar{Z} \otimes I \to \bar{Z} \otimes I$$
$$I \otimes \bar{Z} \to \bar{Z} \otimes \bar{Z}$$

which comparing with the result of Ex. 4.31, is precisely the action of a logical controlled-NOT gate with the first qubit as the control. Thus, we have used fault-tolerant stabilizer state preparation (to create the single-qubit encoded states $|\bar{\psi}\rangle_1$, $|\bar{\varphi}\rangle_2$, $|0\rangle_3$, $|0\rangle_4$), fault-tolerant measurement of elements of the stabilizer (measuring $\bar{X}_3, \bar{X}_4$) and tranversally applied normalizer gates ($U^{\otimes n}$) to implement a fault-tolerant logical controlled-NOT gate that works for any $[n, 1]$ stabilizer code capable of correcting single-qubit errors. □

11 Entropy and information

Exercise 11.1: Simple calculations of entropy

What is the entropy associated with the toss of a fair coin? With the roll of a fair die? How would the entropy behave if the coin or die were unfair?

Solution

Concepts Involved: Shannon Entropy

The entropy associated with a fair coin toss X_{coin} (with $p_H = p_T = \frac{1}{2}$) is:

$$H(X_{\text{coin}}) = -\frac{1}{2}\log\left(\frac{1}{2}\right) - \frac{1}{2}\log\left(\frac{1}{2}\right) = \log(2) = 1$$

and with a fair die X_{die} (with $p_i = \frac{1}{6}$ for $i \in \{1, 2, 3, 4, 5, 6\}$) is:

$$H(X_{\text{die}}) = 6 \cdot (-\frac{1}{6}\log\left(\frac{1}{6}\right)) = \log(6).$$

The entropy decreases if the coin/die are unfair; intuitively, this is because there is less uncertainty before we learn the result than in the fair case. For example in the limiting case, we can see that for a maximally unfair coin/die where one of the probabilities was one and the others were zero (wherein there is no uncertainty of the result whatsoever) that the entropy would be zero. For the case of the binary entropy (which we can view as a coin with weighted probability of heads p) we prove in Ex. 11.3 that indeed the fair coin of $p = 1/2$ maximizes the entropy. □

Exercise 11.2: Intuitive justification for the definition of entropy

($\star$) Suppose we are trying to quantify how much information is provided by an event E which may occur in a probabilistic experiment. We do this using an 'information function' $I(E)$ whose value is determined by the event E. Suppose we make the following assumptions about this function:

- $I(E)$ is a function only of the probability of the event E, so we may write $I = I(p)$, where p is a probability in the range 0 to 1.
- I is a smooth function of probability.
- $I(pq) = I(p) + I(q)$ when $p, q > 0$. (*Interpretation:* The information gained when two independent events occur with individual probabilities p and q is the sum of the information gained from each event alone.)

Show that $I(p) = k \log p$, for some constant, k. It follows that the average information gain when one of a mutually exclusive set of events with probabilities $p_1, \ldots, p_n$ occurs is $k \sum_i p_i \log p_i$, which is just the Shanon entropy, up to a constant factor.

Solution

Concepts Involved: Shannon Entropy

Since $I(p)$ is a smooth function of p, it is differentiable (an arbitrary number of times) w.r.t. p. Now, let us consider $I(pq) = I(p) + I(q)$ and differentiate w.r.t. p, using the chain rule:

$$I'(pq) \cdot q = I'(p)$$

Now we differentiate the above w.r.t. q, using the product rule:

$$I'(pq) + I''(pq)pq = 0 \implies I''(pq) = -\frac{I'(pq)}{pq}$$

Defining $x = pq$, we obtain a differential equation for $I'(x)$:

$$I''(x) = -\frac{I'(x)}{x} \implies \frac{dI'(x)}{dx} = -\frac{I'(x)}{x} \implies \frac{dI'(x)}{I'(x)} = -\frac{dx}{x}$$

Integrating both sides, we obtain:

$$\ln\big(I'(x)\big) = -\ln(x) + c_1 \implies I'(x) = e^{c_1} e^{\ln\left(\frac{1}{x}\right)} = \frac{e^{c_1}}{x}$$

Integrating once more:

$$I(x) = e^{c_1} \ln(x) + c_2 = k \ln(x) + c_2$$

where we have defined $e^{c_1} = k$. Now, using that $I(pq) = I(p) + I(q)$ for $p = q = 1$ gives us:

$$I(1) = I(1) + I(1) \implies k\ln(1) + c_2 = k\ln(1) + c_2 + k\ln(1) + c_2 \implies c_2 = 2c_2$$

this sets $c_2 = 0$, so we conclude:

$$I(p) = k \ln(p) = k' \log(p).$$

where the change of base of the logarithm can be absorbed into the constant. □

Exercise 11.3

Prove that the binary entropy $H_{bin}(p)$ attains its maximum value of one at $p = 1/2$

Solution

Concepts Involved: Shannon Entropy, Binary Entropy

Taking the derivative w.r.t. p using the product rule, we find:

$$\begin{aligned}\frac{\partial H_{\text{bin}}(p)}{\partial p} &= -\log(p) - p\frac{1}{p\ln(2)} - (-1)\log(1-p) - (1-p)\frac{-1}{(1-p)\ln(2)} \\ &= -(\log(p) - \log(1-p)) \\ &= -\log\left(\frac{p}{1-p}\right)\end{aligned}$$

We have an extremum when the derivative is zero, which occurs when the argument of the logarithm is one, which occurs when $p = 1/2$. To verify that this is indeed a maximum, we look at the second derivative:

$$\frac{\partial^2 H_{\text{bin}}(p)}{\partial p^2} = -\frac{1}{p(1-p)\ln(2)}$$

which evaluated at $p = 1/2$ yields $-\frac{4}{\ln(2)} < 0$, so indeed $p = 1/2$ is a maximum. For $p = 1/2$ we find that $H_{\text{bin}}(p = 1/2) = 1$ (see Ex. 11.1), thus completing the proof. □

Exercise 11.4: Concavity of the binary entropy

($\star$) From Figure 11.1, it appears that the binary entropy is a concave function. Prove that this is so, that is:

$$H_{bin}(px_1 + (1-p)x_2) \geq pH_{bin}(x_1) + (1-p)H_{bin}(x_2),$$

where $0 \leq p, x_1, x_2 \leq 1$. Prove in addition that the binary entropy is strictly concave, that is, the above inequality is an equality only for the trivial cases $x_1 = x_2$ or $p = 0$ or $p = 1$.

Solution

Concepts Involved: Shannon Entropy, Binary Entropy, Concavity/Convexity

We first prove that for a general twice-differentiable real function f, $f''(x) \leq 0$ implies f concave, in that for any $0 \leq p \leq 1$:

$$f(px_1 + (1-p)x_2) \geq pf(x_1) + (1-p)f(x_2).$$

for any x_1, x_2 in the domain of f.
WLOG, consider $x_1 \leq x_2$. Define $y = px_1 + (1-p)x_2$, which satisfies $x_1 \leq y \leq x_2$. By the Mean value theorem, there exists $y_1 \in (x_1, y)$ and $y_2 \in (y, x_2)$ such that:

$$f(y) - f(x_1) = (y - x_1)f'(y_1)$$

$$f(x_2) - f(y) = (x_2 - y)f'(y_2)$$

Then:

$$\begin{aligned} pf(x_1) + (1-p)f(x_2) &= p\left(f(y) - (y-x_1)f'(y_1)\right) + (1-p)\left(f(y) + (x_2-y)f'(y_2)\right) \\ &= f(y) + p(1-p)(x_2-x_1)[f'(y_2) - f'(y_1)] \end{aligned}$$

which rearranging:

$$f(y) - pf(x_1) + (1-p)f(x_2) = p(1-p)(x_2-x_1)[f'(y_1) - f'(y_2)]$$

Let's study the RHS. $p(1-p) \geq 0$ as $0 \leq p \leq 1$ and $x_2 - x_1 \geq 0$ since $x_2 \geq x_1$. For the last term, note that $y_2 \geq y_1$, and so $(f'(y_1) - f'(y_2)) \geq 0$ since $f''(x) \leq 0$ which implies f' is decreasing. Thus:

$$f(y) - pf(x_1) + (1-p)f(x_2) \geq 0$$

and we conclude f is convex.
Further, note that if $f''(x) < 0$ (as opposed to ≤ 0), then f' is strictly decreasing and so $(f'(y_1) - f'(y_2)) > 0$ (so long as $y_2 > y_1$). Then, the only way that the inequality becomes equality is if $p(1-p)(x_2-x_1)$ vanishes, which occurs if $x_1 = x_2$ or $p = 0$ or $p = 1$. Thus, if f has strictly negative second derivative then f is strictly concave.
Now, observe that the binary entropy $H_{\text{bin}}(p)$ has (strictly) negative second derivative for all $p \in [0,1]$ (See Ex. 11.3) and so by the above theorem, it is strictly convex. □

Exercise 11.5: Subadditivity of the Shannon entropy

Show that $H(p(x,y)\|p(x)p(y)) = H(p(x)) + H(p(y)) - H(p(x,y))$. Deduce that $H(X,Y) \leq H(X) + H(Y)$, with equality if and only if X and Y are independent random variables.

Solution

Concepts Involved: Shannon Entropy, Relative Entropy

From the definition, we have:

$$\begin{aligned} H(p(x,y)\|p(x)p(y)) &= \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \\ &= \sum_{x,y} p(x,y) \log\big(p(x,y)\big) - \sum_{x,y} p(x,y) \log\big(p(x)\big) - \sum_{x,y} p(x,y) \log\big(p(y)\big) \\ &= \sum_{x,y} p(x,y) \log\big(p(x,y)\big) - \sum_{x} p(x) \log\big(p(x)\big) - \sum_{y} p(y) \log\big(p(y)\big) \\ &= -H(p(x,y)) + H(p(x)) + H(p(y)) \end{aligned}$$

where in the second equality we apply log laws, and in the third equality we sum over y in the second term and x in the third term, using that $\sum_y p(x,y) = p(x)$ and $\sum_x p(x,y) = p(y)$ (probabilities sum to one). Now, using Theorem 11.1 (non-negativity of the relative entropy), we find:

$$H(p(x,y)\|p(x)p(y)) = -H(p(x,y)) + H(p(x)) + H(p(y)) \geq 0$$

and so:

$$H(X,Y) \leq H(X) + H(Y).$$

Further, the relative entropy $H(p(x,y)\|p(x)p(y))$ is zero iff $p(x,y) = p(x)p(y)$ for all x, y, i.e. X, Y are independent random variables. □

Exercise 11.6: Proof of strong subadditivity

($\star$) Prove that $H(X,Y,Z) + H(Y) \leq H(X,Y) + H(Y,Z)$, with equality if and only if $Z \to Y \to X$ forms a Markov chain.

Solution

Concepts Involved: Shannon Entropy, Markov Chains, Subadditivity, Conditional Probability

Consider:

$$\begin{aligned}
\sum_{x,y,z} p(x,y,z) \ln \frac{p(x,y)p(y,z)}{p(x,y,z)p(y)} &\leq \frac{1}{\ln 2} \sum_{x,y,z} p(x,y,z) \left(\frac{p(x,y)p(y,z)}{p(x,y,z)p(y)} - 1 \right) \\
&= \frac{1}{\ln 2} \sum_{x,y,z} \left(\frac{p(x,y)p(y,z)}{p(y)} - p(x,y,z) \right) \\
&= \frac{1}{\ln 2} \left(\sum_{y} \frac{p(y)p(y)}{p(y)} - 1 \right) \\
&= \frac{1}{\ln 2} \left(\sum_{y} p(y) - 1 \right) \\
&= \frac{1}{\ln 2} (1 - 1) \\
&= 0
\end{aligned}$$

where we use $\log x \leq (x-1)/\ln 2$ in the first line, and in the subsequent lines use the fact that sums over probabilities over a given variable yield 1. Splitting up the logarithm in the initial expression using log laws, the strong subadditivity inequality follows.
Note that $\log(x) = (x-1)/\ln 2$ if and only if $x = 1$, which applied to this case yields:

$$\frac{p(x,y)p(y,z)}{p(x,y,z)p(y)} = 1$$

if we massage this expression by using the definition of the conditional probability:

$$\begin{aligned} 1 &= \frac{p(x|y)p(y)p(y,z)}{p(x,y,z)p(y)} \\ &= \frac{p(x|y)p(y,z)}{p(x,y,z)} \\ &= \frac{p(x|y)p(y,z)}{p(x|y,z)p(y,z)} \\ &= \frac{p(x|y)}{p(x|y,z)} \end{aligned}$$

and so:

$$p(x|y,z) = p(x|y)$$

so we conclude equality in the subadditivity inequality iff $Z \to Y \to X$ is a Markov chain. □

Exercise 11.7

In Exercise 11.5 you implicitly showed that the mutual information $H(X:Y)$ could be expressed as the relative entropy of two probability distributions, $H(X:Y) = H(p(x,y)\|p(x)p(y))$. Find an expression for the conditional entropy $H(Y \mid X)$ as a relative entropy between two probability distributions. Use this expression to deduce that $H(Y \mid X) \geq 0$ and to find the equality conditions.
This exercise is incorrect, see pseudo-solution for details.

Solution

Concepts Involved: Shannon Entropy, Relative Entropy, Conditional Entropy, Mutual Information

Using the definition of the conditional entropy:

$$\begin{aligned} H(Y|X) &= H(X,Y) - H(X) \\ &= -\sum_{x,y} p(x,y)\log p(x,y) + \sum_{x} p(x)\log p(x) \\ &= -\sum_{x,y} p(x,y)\log p(x,y) + \sum_{x} p(x,y)\log p(x) \\ &= -\sum_{x,y} p(x,y)\log \frac{p(x,y)}{p(x)} \\ &\cong -H(p(x,y)\|p(x)) \end{aligned}$$

Note the choice of $\cong$ notation because this is not a valid expression for the relative entropy, since $p(x,y), p(y)$ are not probability distributions over the same index set. We do not believe there is a well-defined expression of $H(Y|X)$ as a relative entropy.
Note this also means that the suggestion of the last line of the exercise to use that the relative entropy is non-negative to prove that $H(Y|X)$ is non-negative is not valid (indeed, this would actually lead to the incorrect conclusion that $H(Y|X)$ is non-positive!) Instead, one can follow the argument of Theorem

11.3(2) in the text and observe that $p(x,y) \leq p(x)$ for all y, and so $\log \frac{p(x,y)}{p(x)} \leq 0$ for each term of the sum, leading to $H(Y|X)$ to be positive (and zero only if Y is a function of X in which case $p(x,y) = p(x)$). □

Exercise 11.8: Mutual information is not always subadditive

Let X and Y be independent identically distributed random variables taking the values 0 and 1 with probability 1/2. Let $Z \equiv X \oplus Y$, where $\oplus$ denotes addition modulo 2. Show that the mutual information in this case is not subadditive,

$$H(X, Y : Z) \nleq H(X : Z) + H(Y : Z).$$

Solution

Concepts Involved: Shannon Entropy, Mutual Information, Subadditivity

The single and joint probabilities are easily seen to be:

$$p(X = 0/1) = P(Y = 0/1) = P(Z = X \oplus Y = 0/1) = \frac{1}{2}$$

$$p(X = 0/1, Y = 0/1) = p(X = 0/1, Z = 0/1) = p(Y = 0/1, Z = 0/1) = \frac{1}{4}$$

$$\begin{aligned} &p(X = 0, Y = 0, Z = 0) = p(X = 0, Y = 1, Z = 1) \\ &= p(X = 1, Y = 0, Z = 1) = p(X = 1, Y = 1, Z = 0) = \frac{1}{4} \end{aligned}$$

$$\begin{aligned} &p(X = 0, Y = 0, Z = 1) = p(X = 0, Y = 1, Z = 0) \\ &= p(X = 1, Y = 0, Z = 0) = p(X = 1, Y = 1, Z = 1) = 0 \end{aligned}$$

from this we can compute the entropies:

$$H(X) = H(Y) = H(Z) = \log 2 = 1$$

$$H(X, Y) = H(X, Z) = H(Y, Z) = \log 4 = 2$$

$$H(X, Y, Z) = \log 4 = 2$$

and so computing the mutual information:

$$H(X, Y : Z) = H(X, Y) + H(Z) - H(X, Y, Z) = 2 + 1 - 2 = 1$$

$$H(X:Z) = H(X) + H(Z) - H(X,Z) = 1 + 1 - 2 = 0$$

$$H(Y:Z) = H(Y) + H(Z) - H(Y,Z) = 1 + 1 - 2 = 0$$

and thus:

$$H(X,Y:Z) \nleq H(X:Z) + H(Y:Z).$$

□

Exercise 11.9: Mutual information is not always superadditive

Let X_1 be a random variable taking values 0 or 1 with respective probabilities of $1/2$ and $X_2 \equiv Y_1 \equiv Y_2 \equiv X_1$. Show that the mutual information in this case is not superadditive,

$$H(X_1:Y_1) + H(X_2:Y_2) \nleq H(X_1,X_2:Y_1,Y_2).$$

Solution

Concepts Involved: Shannon Entropy, Mutual Information

The single variable probabilities are:

$$p(X_1 = 0/1) = \frac{1}{2}$$

with the same for X_2, Y_1, Y_2. The joint probabilities are:

$$p(X_1 = X_2 = 0) = p(X_1 = X_2 = 1) = \frac{1}{2}$$

$$p(X_1 = 0, X_2 = 1) = p(X_1 = 1, X_2 = 0) = 0$$

and the same for the joint probabilities of $(Y_1,Y_2),(X_1,Y_1),(X_2,Y_2)$. Finally, for the joint probability of four variables:

$$p(X_1 = X_2 = Y_1 = Y_2 = 0) = p(X_1 = X_2 = Y_1 = Y_2 = 1) = \frac{1}{2}$$

and zero otherwise.
Thus we can compute the entropies:

$$H(X_1) = H(X_2) = H(Y_1) = H(Y_2) = \log 2 = 1$$

$$H(X_1,X_2) = H(Y_1,Y_2) = H(X_1,Y_1) = H(X_2,Y_2) = \log 2 = 1$$

$$H(X_1,X_2,Y_1,Y_2) = \log 2 = 1.$$

Finally, computing the mutual information:

$$H(X_1 : Y_1) = H(X_1) + H(Y_1) - H(X_1, Y_1) = 1 + 1 - 1 = 1$$

$$H(X_2 : Y_2) = H(X_2) + H(Y_2) - H(X_2, Y_2) = 1 + 1 - 1 = 1$$

$$H(X_1, X_2 : Y_1, Y_2) = H(X_1, X_2) + H(Y_1, Y_2) - H(X_1, X_2, Y_1, Y_2) = 1 + 1 - 1 = 1$$

from which we conclude:

$$H(X_1 : Y_1) + H(X_2 : Y_2) \not\leq H(X_1, X_2 : Y_1, Y_2)$$

□

Exercise 11.10

($\star$) Show that if $X \to Y \to Z$ is a Markov chain then so is $Z \to Y \to X$.

Solution

Concepts Involved: Markov Chains, Conditional Probability

We use the shorthand that $P(Z = z) = p(z)$ (and the same for x, y).
If $X \to Y \to Z$ is a Markov chain, then:

$$p(z|y, z) = p(z|y)$$

Then using the definition of the conditional probability twice:

$$\begin{aligned} p(z|x, y) &= \frac{p(z, y, x)}{p(y, x)} \\ &= \frac{p(z, y, x)}{p(x|y)p(y)} \end{aligned}$$

And also:

$$p(z|y) = \frac{p(z, y)}{p(y)}$$

So cancelling out $p(y)$ from both sides, the Markov chain condition reduces to:

$$\frac{p(z, y, x)}{p(x|y)} = p(z, y)$$

and so:

$$p(z, y, x) = p(x|y)p(z, y)$$

Now, to show $Z \to Y \to X$ is a Markov chain, let us evaluate:

$$p(x|y,z) = \frac{p(x,y,z)}{p(y,z)}.$$

Substituting in the expression for $p(z,y,x) = p(x,y,z)$ obtained from the $X \to Y \to Z$ Markov chain condition:

$$\begin{aligned} p(x|y,z) &= \frac{p(x|y)p(z,y)}{p(z,y)} \\ &= p(x|y) \end{aligned}$$

and thus $Z \to Y \to X$ is a Markov chain. □

Exercise 11.11: Example calculations of entropy

Calculate $S(\rho)$ for

$$\rho = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}; \ \rho = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}; \ \rho = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Solution

Concepts Involved: Density Matrices, Von Neumann Entropy

For the first density matrix:

$$\rho = 1\,|0\rangle\langle 0| + 0\,|1\rangle\langle 1|$$

and so:

$$S(\rho) = -(1\log(1) + 0\log(0)) = 0 \tag{40}$$

For the second density matrix:

$$\rho = 1\,|+\rangle\langle +| + 0\,|-\rangle\langle -|$$

and so:

$$S(\rho) = -(1\log(1) + 0\log(0)) = 0$$

Finally for the third density matrix:

$$\rho = \frac{3+\sqrt{5}}{6}\,|v_+\rangle\langle v_+| + \frac{3-\sqrt{5}}{6}\,|v_-\rangle\langle v_-|$$

with $|v_+\rangle = \frac{1}{\sqrt{10+2\sqrt{5}}}\begin{bmatrix}1+\sqrt{5}\\2\end{bmatrix}$ and $|v_-\rangle = \frac{1}{\sqrt{10+2\sqrt{5}}}\begin{bmatrix}-2\\1+\sqrt{5}\end{bmatrix}$. Then:

$$\begin{aligned}S(\rho) &= -\left(\frac{3+\sqrt{5}}{6}\log\left(\frac{3+\sqrt{5}}{6}\right)+\frac{3-\sqrt{5}}{6}\log\left(\frac{3-\sqrt{5}}{6}\right)\right)\\ &= -\frac{3+\sqrt{5}}{6}\log\left(3+\sqrt{5}\right)-\frac{3-\sqrt{5}}{6}\log\left(3-\sqrt{5}\right)+\log(6)\end{aligned}$$

□

Exercise 11.12: Comparison of quantum and classical entropies

Suppose $\rho = p\,|0\rangle\langle 0| + (1-p)\,|+\rangle\langle +|$. Evaluate $S(\rho)$. Compare the value of $S(\rho)$ to $H(p, 1-p)$.

Solution

Concepts Involved: Von Neumann Entropy, Binary Entropy

We can write ρ as:

$$\rho = \frac{1}{2}\begin{bmatrix}1+p & 1-p\\1-p & 1-p\end{bmatrix}$$

Which has eigenvalues:

$$\det(\rho - \lambda\mathbb{I}) = 0 \implies \lambda_\pm = \frac{1 \pm \sqrt{1+2p^2-2p}}{2}$$

Thus:

$$\begin{aligned}S(\rho) &= -\frac{1+\sqrt{1+2p^2-2p}}{2}\log\left(\frac{1+\sqrt{1+2p^2-2p}}{2}\right)-\frac{1-\sqrt{1+2p^2-2p}}{2}\log\left(\frac{1-\sqrt{1+2p^2-2p}}{2}\right)\\ &= -\frac{1+\sqrt{1+2p^2-2p}}{2}\log\left(1+\sqrt{1+2p^2-2p}\right)-\frac{1-\sqrt{1+2p^2-2p}}{2}\log\left(1-\sqrt{1+2p^2-2p}\right)\\ &\quad +\log(2)\end{aligned}$$

Next, calculating $H(p, 1-p)$:

$$H(p, 1-p) = -p\log(p) - (1-p)\log(1-p)$$

Comparing the two, we find that $S(\rho) \leq H(p, 1-p)$ with equality only when $p = 0, 1$. Intuitively this makes sense because $H(p, 1-p)$ is the entropy of a mixed state of two orthogonal states with weights/probabilities $p, 1-p$, while ρ is a mixture of two non-orthogonal states. □

Exercise 11.13: Entropy of a tensor product

Use the joint entropy theorem to show that $S(\rho \otimes \sigma) = S(\rho) + S(\sigma)$. Prove this result directly from the definition of the entropy.

Solution

Concepts Involved: Von Neumann Entropy, Tensor Products

ρ is Hermitian, and thus by the spectral theorem can be diagonalized as $\rho = \sum_i \lambda_i |i\rangle\langle i|$ with the $|i\rangle$ orthogonal. Further, $\mathrm{Tr}(\rho) = 1$ enforces $\sum_i \lambda_i = 1$. Thus, we may write:

$$\begin{aligned} S(\rho \otimes \sigma) &= S(\sum_i \lambda_i |i\rangle\langle i| \otimes \sigma) \\ &= H(\lambda_i) + \sum_i \lambda_i S(\sigma) \\ &= S(\rho) + S(\sigma) \end{aligned}$$

where in the second equality we apply the joint entropy theorem, in the third equality we use $\sum_i \lambda_i = 1$ and the fact that the definitions of the Shannon entropy and Von Neumann entropy coincide when viewing the eigenvalues of a density operator as a set of probabilities.
Next, we prove things directly from the definition of the entropy. Diagonalize ρ, σ as:

$$\rho = \sum_i \lambda_i^\rho |v_i^\rho\rangle\langle v_i^\rho|, \quad \sigma = \sum_j \lambda_j^\rho |v_j^\sigma\rangle\langle v_j^\sigma|$$

with $\lambda_i^\rho, \lambda_j^\sigma$ the eigenvalues, with $\sum_i \lambda_i^\rho = \sum_j \lambda_j^\sigma = 1$ from normalization.
Then, $\rho \otimes \sigma$ can be diagonalized as:

$$\rho \otimes \sigma = \sum_{ij} \lambda_i^\rho \lambda_j^\sigma (|v_i^\rho\rangle \otimes |v_i^\sigma\rangle)(\langle v_i^\rho| \otimes \langle v_i^\sigma|)$$

with $\lambda_i^\rho \lambda_j^\sigma$ the eigenvalues. Then from the definition of the Von Neumann entropy:

$$\begin{aligned} S(\rho \otimes \sigma) &= -\sum_{ij} \lambda_i^\rho \lambda_j^\sigma \log\left(\lambda_i^\rho \lambda_j^\sigma\right) \\ &= -\sum_{ij} \lambda_i^\rho \lambda_j^\sigma \log(\lambda_i^\rho) - \sum_{ij} \lambda_i^\rho \lambda_j^\sigma \log(\lambda_i^\sigma) \\ &= -\sum_i \lambda_i^\rho \log(\lambda_i^\rho) - \sum_j \lambda_j^\sigma \log\left(\lambda_j^\sigma\right) \\ &= S(\rho) + S(\sigma) \end{aligned}$$

where in the third equality we sum over j in the first sum and i in the second (both of which yield factors of 1). □

Exercise 11.14: Entanglement and negative conditional entropy

Suppose $|AB\rangle$ is a pure state of a composite system belonging to Alice and Bob. Show that $|AB\rangle$ is entangled if and only if $S(B \mid A) < 0$.

Solution

Concepts Involved: Von Neumann Entropy, Conditional Entropy, Entanglement

We will prove the equivalent statement that $|AB\rangle$ is unentangled if and only if $S(B|A) \geq 0$.
$\boxed{\Longrightarrow}$: Suppose $|AB\rangle$ is unentangled. Then, $\rho_{AB} = \rho_A \otimes \rho_B$. So:

$$S(B|A) = S(A,B) - S(A) = S(A) + S(B) - S(A) = S(B) \geq 0$$

well in the first equality we use the definition of the conditional entropy, and in the second equality we use the result of Ex. **??**.
$\boxed{\Longleftarrow}$: Suppose $S(B|A) \geq 0$. Then:

$$0 \leq S(B|A) = S(A,B) - S(A) = -S(A) \tag{41}$$

where we have used that $S(A,B) = 0$ since the composite system $|AB\rangle$ is a pure state. Since the entropy is non-negative, this implies that:

$$0 = S(A)$$

and so ρ_A is pure, and hence by Ex. 2.78 $|AB\rangle$ is unentangled. □

Exercise 11.15: Generalised measurements can decrease entropy

Suppose a qubit is in the state ρ is measured using the measurement operators $M_1 = |0\rangle\langle 0|$ and $M_2 = |0\rangle\langle 1|$. If the result of the measurement is unknown to us then the state of the system afterwards is $M_1 \rho M_1^\dagger + M_2 \rho M_2^\dagger$. Show that this procedure can *decrease* the entropy of the qubit.

Solution

Concepts Involved: Von Neumann Entropy, Quantum Measurement

Suppose we start with $\rho = I/2$, the maximally mixed state of a qubit. Then, $S(\rho) = \log 2$. After the measurement, the state is:

$$\begin{aligned}
\rho \to \rho' &= M_1 \rho M_1^\dagger + M_2 \rho M_2^\dagger \\
&= |0\rangle\langle 0| \, (I/2) \, |0\rangle\langle 0| + |0\rangle\langle 1| \, (I/2) \, |1\rangle\langle 0| \\
&= |0\rangle \, \langle 0|0\rangle \, \langle 0| \, /2 + |0\rangle \, \langle 1|1\rangle \, \langle 0| \, /2 \\
&= |0\rangle\langle 0|
\end{aligned}$$

so the post-measurement state is a pure state with $S(\rho') = 0$, so we can see that the entropy has decreased. □

Exercise 11.16: Equality conditions for $S(A,B) \geq S(B) - S(A)$

($\star\star$) Let $\rho^{AB} = \sum_i \lambda_i |i\rangle\langle i|$ is a spectral decomposition for ρ^{AB}. Show that $S(A,B) = S(B) - S(A)$ if and only if the operators $\rho_i^A \equiv \mathrm{Tr}_B(|i\rangle\langle i|)$ have a common eigenbasis and the $\rho_i^B \equiv \mathrm{Tr}_A(|i\rangle\langle i|)$ have orthogonal support.
We believe there is an error with this exercise - the condition should be that the ρ_i^A are identical, not that they just have a common eigenbasis.

Solution

Concepts Involved: Von Neumann Entropy, Subadditivity

First, let us show that the original conditions are insufficient. Consider qutrits/three-level systems A, B and consider the equal weight mixture of the three pure states:

$$|\psi_1\rangle = |0\rangle_A |0\rangle_B$$

$$|\psi_2\rangle = |0\rangle_A |1\rangle_B$$

$$|\psi_3\rangle = |1\rangle_A |2\rangle_B$$

Wherein we can observe $\rho = \sum_i \frac{1}{3} |\psi_i\rangle\langle\psi_i|$ has eigenstates $|\psi_i\rangle$ with eigenvalue $\frac{1}{3}$ each. Calculating the reduced density operators for A, we find:

$$\rho_1^A = |0\rangle\langle 0|_A, \quad \rho_1^A = |0\rangle\langle 0|_A, \quad \rho_3^A = |1\rangle\langle 1|_A$$

which share a common eigenbasis $\{|0\rangle_A, |1\rangle_A, |2\rangle_A\}$. Meanwhile, looking at the reduced density operators for B:

$$\rho_1^B = |0\rangle\langle 0|_B, \quad \rho_1^B = |1\rangle\langle 1|_B, \quad \rho_3^B = |2\rangle\langle 2|_B$$

which all have orthogonal support.
Computing the joint and individual entropies, we find:

$$S(A,B) = -\sum \lambda_i \log \lambda_i = 3 \cdot -\frac{1}{3} \log \frac{1}{3} = \log 3$$

$$S(A) = S(\frac{2}{3} |0\rangle\langle 0|_A + \frac{1}{3} |1\rangle\langle 1|_A) = -\frac{2}{3} \log\left(\frac{2}{3}\right) - \frac{1}{3} \log\left(\frac{1}{3}\right) > 0$$

$$S(B) = S(\frac{1}{3} |0\rangle\langle 0|_A + \frac{1}{3} |1\rangle\langle 1|_B + \frac{1}{3} |2\rangle\langle 2|_B) = -3 \cdot \frac{1}{3} \log\left(\frac{1}{3}\right) = \log 3$$

Hence $S(A) + S(B) > \log 3 = S(A,B)$ so equality does not hold, which tells us that the conditions as originally posed in the problem are too weak.
Let us slightly rewind the proof of the triangle inequality in the text. Let R be the system that purifies

AB. Then $S(A,B) = S(R)$ and $S(A,R) = S(B)$ and so:

$$S(A,B) = S(B) - S(A) \iff S(A) + S(R) = S(A,R)$$

But this is true iff $\rho^{AR} = \rho^A \otimes \rho^R$, i.e. A is unentangled with the purifying system.
The purified state can be written as:

$$|ABR\rangle = \sum_i \sqrt{\lambda_i}\, |i\rangle_{AB}\, |i\rangle_R$$

So:

$$\rho^{AR} = \mathrm{Tr}_B(\sum_{ij} \sqrt{\lambda_i \lambda_j}\, |i\rangle\langle j|_{AB} \otimes |i\rangle\langle j|_R) = \sum_{ij} \sqrt{\lambda_i \lambda_j} \rho^A_{ij} \otimes |i\rangle\langle j|_R$$

with $\rho^A_{ij} = \mathrm{Tr}_B(|i\rangle\langle j|_{AB})$.
Then calculating the reduced density operators:

$$\rho^A = \mathrm{Tr}_R(\rho^{AR}) = \sum_{ij} \sqrt{\lambda_i \lambda_j} \rho^A_{ij} \,\mathrm{Tr}\big(|i\rangle\langle j|\big) = \sum_{ij} \sqrt{\lambda_i \lambda_j} \rho^A_{ij} \delta_{ij} = \sum_i \lambda_i \rho^A_{ii} = \sum_i \lambda_i \rho^A_i$$

$$\begin{aligned} \rho^R = \mathrm{Tr}_A(\rho^{AR}) = \sum_{ij} \sqrt{\lambda_i \lambda_j}\, \mathrm{Tr}_A(\rho^A_{ij})\, |i\rangle\langle j|_R &= \sum_{ij} \sqrt{\lambda_i \lambda_j}\, \mathrm{Tr}_{AB}(|i\rangle\langle j|_{AB})\, |i\rangle\langle j|_R \\ &= \sum_{ij} \sqrt{\lambda_i \lambda_j} \delta_{ij}\, |i\rangle\langle j|_R \\ &= \sum_i \lambda_i\, |i\rangle\langle i|_R \end{aligned}$$

So:

$$\rho^A \otimes \rho^R = \sum_{ij} \lambda_i \lambda_j \rho^A_i \otimes |j\rangle\langle j|_R$$

Comparing this expression to that of ρ^{AR}, the two expressions are equal iff:

(1) $\rho^A_{ij} = 0$ for $i \neq j$, as no $i \neq j$ terms appear in $\rho^A \otimes \rho^R$.

(2) $\sum_i \lambda_i \rho^A_i = \rho^A_j$ for every j.

Condition (2) holds iff each of the ρ^A_i are identical. Condition (1) requires a bit of massaging. It says that:

$$\mathrm{Tr}_B(|i\rangle\langle j|_{AB}) = \sum_n \langle n|_B\, |i\rangle\langle j|_{AB}\, |n\rangle_B = 0$$

if $i \neq j$, where $\big\{|n\rangle\big\}_n$ is an orthonormal basis of subsystem B.
Schmidt decomposing the eigenvectors $|i\rangle_{AB}$ of ρ^{AB}, we have:

$$|i\rangle_{AB} = \sum_k \mu_{ik}\, |i_k\rangle_A \otimes |i_k\rangle_B$$

Substituting this into the vanishing trace condition:

$$\sum_{kln} \mu_{ik}\mu_{jl}^* \langle n|i_k\rangle_B \langle j_l|n\rangle_B |i_k\rangle\langle j_l|_A = 0$$

We can rearrange this to find:

$$\begin{aligned}\sum_{kln} \mu_{ik}\mu_{jl}^* \langle n|i_k\rangle_B \langle j_l|n\rangle_B |i_k\rangle\langle j_l|_A &= \sum_{kl} \mu_{ik}\mu_{jl}^* \langle j_l| \left(\sum_n |n\rangle\langle n|\right) |i_k\rangle_B |i_k\rangle\langle j_l|_A \\ &= \sum_{kl} \mu_{ik}\mu_{jl}^* \langle j_l| I |i_k\rangle_B |i_k\rangle\langle j_l|_A \\ &= \sum_{kl} \mu_{ik}\mu_{jl}^* \langle j_l|i_k\rangle_B |i_k\rangle\langle j_l|_A \\ &= 0\end{aligned}$$

This is true iff $\langle j_l|i_k\rangle_B = 0\ \forall k,l$ and $\forall i,j$. But $\rho_i^B = \mathrm{Tr}_A(|i\rangle\langle i|) = \sum_k \mu_{ik} |i_k\rangle\langle i_k|_B$, so this is equivalent to the condition that the ρ_i^B have orthogonal support.
Thus we have shown that for a state $\rho^{AB} = \sum_i \lambda_i |i\rangle\langle i|$, $S(A,B) = S(B) - S(A)$ iff $\rho_i^A = \mathrm{Tr}_B(|i\rangle\langle i|)$ are identical and $\rho_i^B = \mathrm{Tr}_A(|i\rangle\langle i|)$ haev orthogonal support. □

Exercise 11.17

($\star$) Find an explicit non-trivial example of a mixed state ρ for AB such that $S(A,B) = S(B) - S(A)$.

Solution

Concepts Involved: Von Neumann Entropy, Subadditivity

A trivial example is obtained by considering a product state of a pure state of subsystem A and a mixed state of subystem B, for which:

$$S(A,B) = S(A) + S(B) = S(B) = S(B) - S(A)$$

where the first equality follows from Ex. 11.13 and we have then used that $S(A) = 0$ for a pure state. As a concrete example, we could take $\rho^A = |0\rangle\langle 0|$ and $\rho^B = I/2$, $\rho^{AB} = \rho^A \otimes \rho^B$, for which $S(\rho^{AB}) = S(\rho^B) = \log 2$ and $S(\rho^A) = 0$.
We can construct the first non-trivial example (i.e. an example where A, B are actually entangled) inspired by the above, by considering a case where all of the mixedness is still localized to subsystem B. Namely, consider A holding 1 qubit and B holding two qubits, with:

$$\rho^{AB} = |B_{00}\rangle\langle B_{00}|_{AB_1} \otimes I_{B_2}/2$$

i.e. A, B sharing a Bell pair $|B_{00}\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$ and B separately holding a maximally mixed qubit. It is then clear to see that (introducing a purification system R) $\rho^{AR} = \rho^A \otimes \rho^R$ (as A is maximally entangled with B and all of the mixedness of ρ^A arises from correlations with B), and hence that $S(A,B) = S(B) - S(A)$.

More concretely, we can calculate:

$$S(A,B) = S(\rho^{AB}) = S(\rho^{AB_1B_2}) = S(\rho^{AB_1} \otimes \rho^{B_2}) = S(|B_{00}\rangle\langle B_{00}|_{AB_1}) + S(I_{B_2}/2) = 0 + \log 2 = \log 2$$

$$S(A) = S(I_{A_1}/2) = \log 2$$

$$S(B) = S(\rho^{B_1B_2}) = S(\rho^{B_1} \otimes \rho^{B_2}) = S(I_{B_1}/2) + S(I_{B_2}/2) = \log 2 + \log 2 = \log 4$$

where we have again used Ex. 11.13 for the entropies of tensor products, as well the result of Ex. 2.75 that the reduced density operators for Bell pairs are just the maximally mixed state on qubits.
From the above calculations, we easily see that:

$$S(A,B) = S(B) - S(A).$$

□

Exercise 11.18

Prove that equality holds in the concavity inequality (11.79) if and only if all the ρ_is are the same.

Solution

Concepts Involved: Von Neumann Entropy, Concavity, Subadditivity

From the proof of subadditivity, we know that equality in the subadditivity conditions holds only when the state is unentangled across the bipartition. Since concavity is proven using subadditivity, equality holds if and only if:

$$\rho^{AB} = \sum_i p_i \rho_i \otimes |i\rangle\langle i|$$

is unentangled. Thus, we show that ρ^{AB} is unentangled if and only if all the ρ_is are the same.
$\boxed{\Longrightarrow}$: Suppose ρ^{AB} is unentalged. Then, $\rho^{AB} = \rho^A \otimes \rho^B$ with $\rho^A = \mathrm{Tr}_B(\rho^{AB}), \rho^B = \mathrm{Tr}_A(\rho^{AB})$. We then find:

$$\rho^A = \mathrm{Tr}_B(\rho^{AB}) = \sum_j \langle j|_B \sum_i p_i \rho_i^A \otimes |i\rangle\langle i|_B |j\rangle_B = \sum_{ij} p_i \rho_i^A \delta_{ij} = \sum_i p_i \rho_i^A$$

and:

$$\rho^B = \mathrm{Tr}_A(\sum_i p_i \rho^A \otimes |i\rangle\langle i|_B) = \sum_i p_i \,\mathrm{Tr}\Big(\rho^A\Big)\, |i\rangle\langle i|_B = \sum_i p_i \, |i\rangle\langle i|_B$$

We then have:

$$\rho^{AB} = \rho^A \otimes \rho^B \implies \sum_i p_i \rho_i \otimes |i\rangle\langle i| = \sum_{ij} p_i p_j \rho_i \otimes |j\rangle\langle j|$$

From which it must be the case that (since the sums must agree on $|j\rangle\langle j|_B$ terms):

$$p_j\rho_j = \sum_i p_i p_j \rho_i \implies \rho_j = \sum_i p_i \rho_i.$$

Since this holds for all j, it must be the case that all ρ_is are the same.
$\boxed{\Longleftarrow}$: Suppose all the ρ_is are the same, say $\rho = \rho_i$. Then:

$$\rho^{AB} = \sum_i p_i \rho_i \otimes |i\rangle\langle i| = \sum_i p_i \rho \otimes |i\rangle\langle i| = \rho \otimes \left(\sum_i p_i |i\rangle\langle i| \right)$$

so ρ^{AB} is unentangled. □

Exercise 11.19

($\star\star$) Show that there exists a set of unitary matrices U_j and a probability distribution p_j such that for any matrix A,

$$\sum_i p_i U_i A U_i^\dagger = \mathrm{Tr}(A)\frac{I}{d},$$

where d is the dimension of the Hilbert space A lives in. Use this observation and the strict concavity of the entropy to give an alternate proof that the completely mixed state I/d on a space of d dimensions is the unique state of maximal entropy.

Solution

Concepts Involved: Unitary Operators, Concavity, Von Neumann Entropy

For inspiration, we first study the single-qubit case, and recall the discussion of the depolarizing channel and its operator-sum representation. In particular, in Ex. 8.17 we studied the channel:

$$\mathcal{E}(A) = \frac{A + XAX + YAY + ZAZ}{4}$$

which satisfies $\mathcal{E}(I) = I, \mathcal{E}(X) = \mathcal{E}(Y) = \mathcal{E}(Z) = 0$ and so for a general operator (acting on $\mathbb{C}^2$) with $A = aI + bX + cY + dZ$ (with $a = \mathrm{Tr}(A)/d$) we find that

$$\mathcal{E}(A) = \mathrm{Tr}(A)\frac{I}{2}$$

So taking $p_i = \frac{1}{4}$ and $U_i = \sigma_i$ for the Pauli matrices (and $\sigma_0 = I$) gives the claimed result for qubits.
For qudits, we can generalize the above observation by introducing the generalized Pauli operators (also known as Weyl-Heisenberg operators), which for a d-dimensional system with computational basis states $|0\rangle, \ldots |d-1\rangle$ are defined as:

$$\mathcal{X} = \sum_{i=0}^{d-1} |i\rangle\langle i-1|, \quad \mathcal{Z} = \sum_{i=0}^{d-1} \omega^i |i\rangle\langle i|$$

where $\omega = \frac{2\pi i}{d}$ is the dth root of unity and $j-1$ is understood to be subtraction modulo d. These are known as the shift and clock operators, respectively. We observe that these are both unitary. Further, from the definition, we can compute:

$$\mathcal{Z}\mathcal{X} = \sum_{i,j=0}^{d-1} \omega^i \left|i\right\rangle\!\left\langle j-1\right| \delta_{i,j} = \sum_{i=0}^{d-1} \omega^i \left|i\right\rangle\!\left\langle i-1\right|$$

$$\mathcal{X}\mathcal{Z} = \sum_{i,j=0}^{d-1} \omega^j \left|i\right\rangle\!\left\langle j\right| \delta_{i-1,j} = \sum_{i=0}^{d-1} \omega^{i-1} \left|i\right\rangle\!\left\langle i-1\right|$$

So:

$$\mathcal{Z}\mathcal{X} = \omega\mathcal{X}\mathcal{Z}$$

and by repeated application:

$$\mathcal{Z}^l\mathcal{X}^m = \omega^{lm}\mathcal{X}^m\mathcal{Z}^l.$$

If we define:

$$U_{l,m} = \mathcal{Z}^l\mathcal{X}^m$$

we find (via twicefold application of the above commutation relation):

$$U_{l',m'}U_{l,m}U_{l',m'}^\dagger = \omega^{(l'm-lm')}U_{l,m}U_{l'm'}U_{l',m'}^\dagger = \omega^{(l'm-lm')}U_{l,m}.$$

We observe that there are d^2 $U_{l,m}$ operators (letting $l, m = 0, \ldots, d-1$), and that these are orthogonal under the Hilbert-Schmidt Inner product (see Ex. 2.39):

$$\mathrm{Tr}\Big(U_{l,m}^\dagger U_{l',m'}\Big) = d\delta_{l,l'}\delta_{m,m'}$$

Let us derive the above. First, note that:

$$U_{l,m}^\dagger U_{l',m'} = \omega^{(l'-l)m}\mathcal{Z}^{l'-l}\mathcal{X}^{m'-m}.$$

If $m'-m \not\equiv 0 \pmod d$ then the $\mathcal{X}^{m'-m}$ term ensures that the matrix has no diagonal terms and it is traceless. If $m'-m \equiv 0 \pmod d$ but $l'-l \not\equiv 0 \pmod d$, the $\mathcal{Z}^{l'-l}$ operator is a diagonal operator with (non-unit) roots of unity along the diagonal, wherein taking the trace yields zero (sums of roots of unity are zero). Thus, the trace of the above is only nonzero if $l = l', m = m' \pmod d$, in which case we just have the identity operator, with trace d.

Given the above, we can see that the $U_{l,m}$ form a complete basis of $L_{\mathbb{C}^d}$. Thus for any operator A we may expand:

$$A = \sum_{l,m} c_{l,m}U_{l,m}$$

for complex coefficients $c_{l,m}$ (note that the coefficients are generically complex as the $U_{l,m}$ are not Hermitian). Note that $c_{0,0} = \mathrm{Tr}(A)/d$, as $U_{0,0} = I$ is the only term in the expansion with non-vanishing trace.

Now, let $U_i = U_{l',m'}$ for $l', m' = 0, \ldots, d-1$ and $p_i = p_{l',m'} = \frac{1}{d^2}$ for all l', m'. We then have for any operator A that:

$$\begin{aligned}\sum_i p_i U_i A U_i^\dagger &= \frac{1}{d^2} \sum_{l',m'=0}^{d-1} U_{l',m'} \left(\sum_{l,m=0}^{d-1} c_{l,m} U_{l,m} \right) U_{l',m'}^\dagger \\ &= \frac{1}{d^2} \sum_{l,m} c_{l,m} \sum_{l',m'} U_{l',m'} U_{l,m} U_{l',m'}^\dagger \\ &= \frac{1}{d^2} \sum_{l,m} c_{l,m} \sum_{l',m'} \omega^{(l'm - lm')} U_{l,m}\end{aligned}$$

Now, unless $l = m = 0$, the sum over l', m' corresponds to d-fold sums over dth roots of unity, which evaluates to zero. Thus, only the $U_{0,0} = I$ term survives (with d^2 copies), so:

$$\sum_i p_i U_i A U_i^\dagger = \frac{1}{d^2} c_{00} d^2 = \frac{1}{d^2} \frac{\mathrm{Tr}(A)}{d} d^2 I = \mathrm{Tr}(A) \frac{I}{d}$$

which is what we wished to show.

Before proving the second claim, we note the obvious fact that unitaries do not change the entropy of a state ρ. To see, this, consider the spectral decomposition $\rho_i = \sum_i \lambda_i |i\rangle\langle i|$. Then, $U \rho_i U^\dagger = \sum_i \lambda_i U |i\rangle\langle i| U^\dagger = \sum_{i'} \lambda_{i'} \left|i'\right\rangle\left\langle i'\right|$ with the eigenvectors rotated to $\left|i'\right\rangle = U |i\rangle$ and the eigenvalues unchanged. Thus, $S(\rho) = -\sum_i \lambda_i \log \lambda_i = S(U \rho U^\dagger)$.

With this, we apply the derived channel to a quantum state ρ of a qudit:

$$\sum_i p_i U_i \rho U_i^\dagger = \mathrm{Tr}(\rho) \frac{I}{d} = \frac{I}{d} \tag{42}$$

where we have used normalization. Taking the entropy of both sides:

$$\log d = S(\frac{I}{d}) = S(\sum_i p_i U_i \rho U_i^\dagger) \geq \sum_i p_i S(U_i \rho U_i^\dagger) = \sum_i p_i S(\rho) = S(\rho)$$

where the inequality follows from concavity, the second-to-last equality from the invariance of the entropy under unitaries, and the last from $\sum_i p_i = 1$.

Since the above inequality only holds with equality when all $U_i \rho U_i^\dagger$ are identical, which only holds if $\rho = I/d$ (and there are no other Pauli terms in the expansion), we conclude that the complete mixed state I/d is the unique state of maximal entropy in d-dimensional Hilbert space. □

Remark: In the literature, it is common to define $D_{l,m} = \omega^{-lm/2} Z^l Z^m$ to be the phase-space displacement operator (with the extra phase factor to ensure $D_{l,m}^\dagger = D_{-l,-m}$), which is the finite-dimensional analogue of the displacement operators we encountered in Chapter 7.

Further, note that much like the expansion of a qubit density matrix ρ in the Pauli basis had the geometrical interpretation of the Bloch vector, the expansion of a qudit density matrix ρ into the generalized Pauli

matrices similarly has a geometric interpretation, where we may write $\rho = \frac{I + \boldsymbol{\xi} \cdot \boldsymbol{D}}{d}$ for a $d^2 - 1$ dimensional complex vector $\boldsymbol{\xi}$.

Exercise 11.20

($\star$) Let P be a projector and $Q = I - P$ the complementary projector. Prove that there are unitary operators U_1 and U_2 and a probability p such that for all ρ, $P\rho P + Q\rho Q = pU_1\rho U_1^\dagger + (1-p)U_2\rho U_2^\dagger$. Use this observation to give an alternate proof of Theorem 11.9 based on concavity.

Solution

Concepts Involved: Projectors, Projective Measurement, Unitary Operators, Von Neumann Entropy, Concavity

First, observe that $P+Q = I$ is unitary. Further, from the defining properties of projectors $P^\dagger = P^2 = P$, $Q^\dagger = Q^2 = Q$, and the complementarity conditions $PQ = QP = 0$, we observe that $P - Q$ is unitary:

$$(P-Q)^\dagger(P-Q) = (P-Q)^2 = P^2 - QP - PQ + Q^2 = P^2 + Q^2 = P + Q = I.$$

The interpretation of this operator is that it acts as an identity on the subspace spanned by P, and -1 on the complementary subspace.
We then observe that, for any ρ:

$$\begin{aligned}
&\frac{1}{2}(P+Q)\rho(P+Q) + \frac{1}{2}(P-Q)\rho(P-Q) \\
&= \frac{1}{2}(P\rho P + P\rho Q + Q\rho P + Q\rho Q) + \frac{1}{2}(P\rho P - P\rho Q - Q\rho P + Q\rho Q) \\
&= P\rho P + Q\rho Q
\end{aligned}$$

so the claim holds with $U_1 = P + Q = I$, $U_2 = P - Q$, and $p = \frac{1}{2}$.
To provide an alternative proof of Theorem 11.9, we generalize the above result to a non-binary projective measurement. Consider a complete set of projectors $\{P_i\}_{i=1}^n$. Then, the post-mesurement state is given by $\rho' = \sum_i P_i \rho P_i$. We can instead realize this via a probabilistic sum of unitaries, as in the binary case. Consider bitstrings $\mathbf{b}$ of length $n-1$ (of which there are 2^{n-1} unique strings). We can then write down a corresponding unitary:

$$U_{\mathbf{b}} \equiv P_1 + \sum_{i=2}^{n} (-1)^{\mathbf{b}_{i-1}} P_i.$$

We can easily check that $U_{\mathbf{b}}$ is unitary:

$$\begin{aligned} U_{\mathbf{b}}^\dagger U_{\mathbf{b}} &= (P_1 + \sum_{i=2}^{n} (-1)^{\mathbf{b}_{i-1}} P_i)^\dagger (P_1 + \sum_{i=2}^{n} (-1)^{\mathbf{b}_{i-1}} P_i) \\ &= (P_1 + \sum_{i=2}^{n} (-1)^{\mathbf{b}_{i-1}} P_i)^2 \\ &= P_1^2 + \sum_{i=2}^{n} P_i^2 \\ &= \sum_{i=1}^{n} P_i \\ &= I \end{aligned}$$

where we have used the hermiticity (second line), orthogonality (third line), idempotency (fourth line) and completeness (final line) of the projectors. Further, we observe that:

$$\frac{1}{2^{n-1}} \sum_{\mathbf{b}} U_{\mathbf{b}} \rho U_{\mathbf{b}}^\dagger = \sum_i P_i \rho P_i = \rho'$$

where all cross terms $P_i \rho P_j$ for $i \neq j$ vanish from the relative minus signs between bitstring terms (which are summed over), while the "diagonal" terms $P_i \rho P_i$ all survive (of which there are 2^{n-1} in the sum, in total). Thus summing over the family of unitaries $U_{\mathbf{b}}$ with uniform probability $p = \frac{1}{2^{n-1}}$ yields the same action as the projective measurement.
Finally, to our alternative proof! Let ρ' be the post-measurement state after a projective meausrement described by projectors $\{P_i\}_i$. Then:

$$S(\rho') = S(\sum_i P_i \rho P_i) = S(\sum_{\mathbf{b}} \frac{1}{2^{n-1}} U_{\mathbf{b}} \rho U_{\mathbf{b}}^\dagger) \geq \sum_{\mathbf{b}} \frac{1}{2^{n-1}} S(U_{\mathbf{b}} \rho U_{\mathbf{b}}^\dagger) = \sum_{\mathbf{b}} \frac{1}{2^{n-1}} S(\rho) = S(\rho)$$

where the inequality follows from concavity and the second-to-last equality follows from the invariance of the entropy under unitaries (as shown in the previous exercise). Thus projective measurements increase entropy. □

Exercise 11.21: Concavity of the Shannon entropy

Use the concavity of the von Neumann entropy to deduce that the Shannon entropy is concave in probability distributions.

Solution

Concepts Involved: Von Neumann Entropy, Shannon Entropy, Concavity

Consider $i = 1, \ldots n$ sets of probabilities labelled by i, so $\{q_k^i\}_{i=1,\ldots n}$ with q_k^i real and non-negative and $\sum_k q_k^i = 1$ for each i. Take the index set k to be the same for each distribution (such that adding probability distributions is well defined).
To each such probability distribution, we can associate a density operator $\rho_i = \sum_k q_k^i |k\rangle\langle k|$. By construc-

tion we can see that:

$$H(q_i) = -\sum_k q_k^i \log\Big(q_k^i\Big) = S(\rho_i)$$

Let p_i be another probability distribution, from which we can construct a probability distribution from the linear combination to be $q'_k = \sum_i p_i q_k^i$. To this we can associate $\rho' = \sum_i p_i \rho_i$. We then observe:

$$H(q') = H(\sum_i p_i q_i) = S(\sum_i p_i \rho_i) \geq \sum_i p_i S(\rho_i) = \sum_i p_i H(q_i)$$

and thus the Shannon entropy is concave in probability distributions. □

Exercise 11.22: Alternate proof of concavity

($\star\star$) Define $f(p) \equiv S(p\rho+(1-p)\sigma)$. Argue that to show concavity it is sufficient to prove that $f''(p) \leq 0$. Prove that $f''(p) \leq 0$, first for the case where ρ and σ are invertible, and then for the case where they are not.

Solution

Concepts Involved: Von Neumann Entropy, Concavity

We showed that $f''(p) \leq 0$ (for real, twice differentiable functions) suffices to show concavity in Ex. 11.4. Let us thus compute the second derivative of:

$$f(p) = S(p\rho + (1-p)\sigma) = -\operatorname{Tr}\Big((p\rho + (1-p)\sigma)\log\big(p\rho + (1-p)\sigma\big)\Big)$$

First, note that since the trace is linear (a finite sum), it commutes with differentiation. Thus:

$$f'(p) = -\frac{\mathrm{d}}{\mathrm{d}p}\operatorname{Tr}\Big((p\rho + (1-p)\sigma)\log\big(p\rho + (1-p)\sigma\big)\Big) = -\operatorname{Tr}\left(\frac{\mathrm{d}}{\mathrm{d}p}\Big[(p\rho + (1-p)\sigma)\log\big(p\rho + (1-p)\sigma\big)\Big]\right)$$

Now, we break up the calculation into cases, as suggested in the question.
$\boxed{\rho, \sigma \text{ invertible}}$: Since ρ, σ are positive and invertible - thus positive definite, their convex sum $p\rho+(1-p)\sigma$ is also positive definite and invertible. Thus, we can evaluate:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}p}\Big[(p\rho + (1-p)\sigma)\log\big(p\rho + (1-p)\sigma\big)\Big] &= \frac{\mathrm{d}}{\mathrm{d}p}(p\rho + (1-p)\sigma)\log\big(p\rho + (1-p)\sigma\big) \\
&\quad + (p\rho + (1-p)\sigma)\frac{\mathrm{d}}{\mathrm{d}p}\log\big(p\rho + (1-p)\sigma\big) \\
&= (\rho - \sigma)\log\big(p\rho + (1-p)\sigma\big) \\
&\quad + (p\rho + (1-p)\sigma)(p\rho + (1-p)\sigma)^{-1}(\rho - \sigma) \\
&= (\rho - \sigma)[I + \log\big(p\rho + (1-p)\sigma\big)]
\end{aligned}$$

where we have used the product rule, chain rule, and the known derivative of the logarithm, taking it on

faith that these derivative rules carry over in the case of matrix functions. Thus:

$$f'(p) = -\operatorname{Tr}\Big((\rho-\sigma)[I + \log\big(p\rho + (1-p)\sigma\big)]\Big)$$

Taking the second derivative:

$$\begin{aligned} f''(p) &= -\frac{\mathrm{d}}{\mathrm{d}p}\operatorname{Tr}\Big((\rho-\sigma)[I + \log\big(p\rho + (1-p)\sigma\big)]\Big) \\ &= -\operatorname{Tr}\left(\frac{\mathrm{d}}{\mathrm{d}p}\left[(\rho-\sigma)[I + \log\big(p\rho + (1-p)\sigma\big)]\right]\right) \\ &= -\operatorname{Tr}\Big((\rho-\sigma)[0 + (p\rho + (1-p)\sigma)^{-1}(\rho-\sigma)]\Big) \\ &= -\operatorname{Tr}\Big((\rho-\sigma)^2(p\rho + (1-p)\sigma)^{-1}\Big) \end{aligned}$$

Now, $(\rho-\sigma)^2$ is positive and $(p\rho + (1-p)\sigma)^{-1}$ are positive definite, and so their product has non-negative eigenvalues (this is seen as follows - suppose A is positive and B is positive definite; then there exists $B^{1/2}, B^{-1/2}$, from which we can write $AB = B^{-1/2}(B^{1/2}AB^{1/2})B^{1/2}$ so AB is related via similarity transformation to $B^{1/2}AB^{1/2}$, which is Hermitian and positive and hence has non-negative real eigenvalues. Thus AB also has non-negative real eigenvalues). Thus, the negative of the trace must be non-positive, and we conclude:

$$f''(p) \leq 0.$$

$\boxed{\rho, \sigma \text{ not invertible}}$: Now, it is not the case that $p\rho + (1-p)\sigma$ is invertible. However, by convention we take $0 \log 0 = 0$, so when taking the trace of $\operatorname{Tr}\Big((p\rho + (1-p)\sigma)\log\big(p\rho + (1-p)\sigma\big)\Big)$, such zero eigenvalues are ignored. This means that $\operatorname{Tr}\Big((p\rho + (1-p)\sigma)\log\big(p\rho + (1-p)\sigma\big)\Big)$ can be evaluated on the subspace for which $(p\rho + (1-p)\sigma)$ has nonzero eigenvalues, i.e. on the subspace $\operatorname{supp}(\rho) \cup \operatorname{supp}(\sigma)$ (or more precisely the span of $\operatorname{supp}(\rho) \cup \operatorname{supp}(\sigma)$). Restricting to this subspace, the previous argument goes through, as we can write down an inverse for $(p\rho + (1-p)\sigma)$ on this subspace (as it is positive definite on this subspace).
The above argument works for $p \in (0,1)$, with $p = 0, 1$ being subtle cases as the support of $p\rho + (1-p)\sigma$ jumps discontinuously at this point. But, it is also clear that at $p = 0, 1$ that the concavity inequality holds with equality, without having to do the derivative calculation. □

Exercise 11.23: Joint concavity implies concavity in each input

Let $f(A, B)$ be a jointly concave function. Show that $f(A, B)$ is concave in A, with B held fixed. Find a function of two variables that is concave in each of its inputs, but is not jointly concave.

Solution

Concepts Involved: Concavity, Joint Concavity

The claim follows easily from the definition; if $f(A, B)$ is jointly concave, then for all $\lambda \in [0, 1]$ we have:

$$\begin{aligned} f(\lambda A_1 + (1-\lambda)A_2, B) &= f(\lambda A_1 + (1-\lambda)A_2, \lambda B + (1-\lambda)B) \\ &\geq \lambda f(A_1, B) + (1-\lambda) f(A_2, B) \end{aligned}$$

so $f(A, B)$ is concave in A.
The argument of Ex. 11.4 can be applied in the multi-variable case to see that $f(x_1, \ldots, x_n)$ is concave in input x_i if $\partial_{x_i}^2 f \leq 0$ with $\partial_{x_i}^2$ the partial derivative w.r.t. x_i (second derivative. w.r.t. x_i with other arguments held fixed). With this, we can see that:

$$f(x, y) = xy$$

is concave in the individual inputs x, y, as $\partial_x^2 f = \partial_y^2 f = 0 \leq 0$.
However, consider the direction $x = y$, in which case: $f(x, x) = x^2$ - the second derivative along this direction is $2 > 0$ and so f is not jointly concave - more concretely from the definition, if we take $A_1 = B_1 = 0$ and $A_2 = B_2 > 0$, then:

$$f(\lambda A_1 + (1-\lambda)A_2, \lambda B_1 + (1-\lambda)B_2) = f((1-\lambda)A_2, (1-\lambda)A_2) = (1-\lambda)^2 A_2^2$$

while:

$$\lambda f(A_1, B_1) + (1-\lambda) f(A_2, B_2) = \lambda \cdot 0 + (1-\lambda)A_2^2 = (1-\lambda)A_2^2.$$

For any $0 < \lambda < 1$, $(1-\lambda)^2 < 1 - \lambda$ and so:

$$f(\lambda A_1 + (1-\lambda)A_2, \lambda B_1 + (1-\lambda)B_2) < \lambda f(A_1, B_1) + (1-\lambda) f(A_2, B_2)$$

so f is not jointly concave. □

Remark: The general conclusion as the above example illustrates is that element-wise concavity only implies concavity in specific directions, while joint concavity generically requires concavity in every direction.

Exercise 11.24

We obtained strong subadditivity as a consequence of the inequality $S(A) + S(B) \leq S(A, C) + S(B, C)$. Show that this inequality can be obtained as a consequence of strong subadditivity.

Solution

Concepts Involved: Von Neumann Entropy, Strong Subadditivity

Let R be a system that purifies ABC. Applying strong subadditivity to ACR, we obtain:

$$S(A, C, R) + S(A) \leq S(A, C) + S(A, R)$$

but since $ABCR$ is pure, $S(A, C, R) = S(B)$ and $S(A, R) = S(B, C)$ and so the above becomes:

$$S(B) + S(A) \leq S(A, C) + S(B, C)$$

which is the desired inequality. □

Exercise 11.25

($\star$) We obtained strong subadditivity as a consequence of the concavity of the conditional ~~information~~ entropy, $S(A|B)$. Show that the concavity of the conditional entropy may be deduced from strong subadditivity. (*Hint:* You may need to introduce an auxiliary system into the problem.)

Solution

Concepts Involved: Von Neumann Entropy, Strong Subadditivity, Conditional Entropy, Concavity.

Let ρ_0, ρ_1 be density operators on a system AB, and consider $\rho = \lambda\rho_0 + (1-\lambda)\rho_1$ for $\lambda \in [0,1]$.
We follow the hint and introduce a qubit auxiliary system C into the problem, and consider the following state on ABC:

$$\rho^{ABC} = \lambda\rho_0^{AB} \otimes |0\rangle\langle 0|^C + (1-\lambda)\rho_1^{AB} \otimes |1\rangle\langle 1|^C$$

Now, from strong subadditivity, we have:

$$S(A,B,C) + S(B) \leq S(A,B) + S(B,C)$$

which we can rearrange to obtain:

$$S(A,B,C) - S(B,C) \leq S(A,B) - S(B) \implies S(A|B,C) \leq S(A|B).$$

This is Theorem 11.15(1)/the statement that conditioning reduces entropy.
Now, notice $\rho^{AB} = \mathrm{Tr}_C(\rho^{ABC}) = \rho$, and so the RHS is just $S(A|B) = S(A|B)_\rho$. For the LHS, notice that subsystem C plays the role of a classical random variable with $P(C=0) = \lambda, P(C=1) = (1-\lambda)$, so:

$$S(A|BC) = P(C=0)S(A|B,C=0)_\rho + P(C=1)S(A|B,C=1)_\rho = \lambda P(A|B)_{\rho_0} + (1-\lambda)P(A|B)_{\rho_1}$$

putting this into the equation of Theorem 11.15(1):

$$\lambda S(A|B)_{\rho_0} + (1-\lambda)S(A|B)_{\rho_1} \leq S(A|B)_\rho$$

so $S(A|B)$ is concave. □

Exercise 11.26

Prove that $S(A:B) + S(A:C) \leq 2S(A)$. Note that the corresponding inequality for Shannon entropies holds since $H(A:B) \leq H(A)$. Find an example where $S(A:B) > S(A)$.

Solution

Concepts Involved: Von Neumann Entropy, Strong Subadditivity, Mutual Information

From the definition of mutual information, we have:

$$S(A:B)+S(A:C)=S(A)+S(B)-S(A,B)+S(A)+S(C)-S(A,C)$$

so:

$$S(A:B)+S(A:C)-2S(A)=S(B)+S(C)-S(A,C)-S(A,B)$$

The RHS is ≤ 0 from the first SSA inequality (Eq. (11.107)) and so we obtain:

$$S(A:B)+S(A:C)\leq 2S(A)$$

as claimed.
Taking $\rho^{AB}=|B_{00}\rangle\langle B_{00}|$ a Bell state, we have that $\rho^A=\rho^B=I/2$. $S(A,B)=0$ as ρ^{AB} is pure, while $S(A)=S(B)=\log 2$. Thus $S(A:B)=S(A)+S(B)-S(A,B)=\log 4>\log 2=S(A)$. Generically, this $S(A:B)>S(A)$ will hold if ρ^{AB} is an entangled pure state. □

Problem 11.1: Generalized Klein's inequality

($\star\star\star$) Suppose $f(\cdot)$ is a convex function from real numbers to real numbers. Then f induces a natural function $f(\cdot)$ on Hermitian operators, as described in Section 2.1.8 on page 75. Prove that

$$\mathrm{Tr}\big(f(A)-f(B)\big)\geq\mathrm{Tr}\big((A-B)f'(B)\big).$$

Use this result to show that the relative entropy is non-negative.

Solution

Concepts Involved: Convexity, Matrix functions, Trace.
Recall the definition of matrix functions acting on Hermitian operators as as $f(A)=f(\sum_a a\,|a\rangle\langle a|)=\sum_i f(a)\,|a\rangle\langle a|$ where $\{|a\rangle\}$ is the eigenbasis of A (which is guaranteed to exist via the spectral theorem).

Convexity of trace functions. We first prove a useful Lemma, namely that if $t\mapsto f(t)$ is convex as a function from reals to reals, then so is the function $A\mapsto\mathrm{Tr}\big(f(A)\big)$ mapping Hermitian operators to real numbers.
Consider Hermitian A,B and $\lambda A+(1-\lambda)B$ (which is also Hermitian) for $\lambda\in[0,1]$. Let $\{|i\rangle\}_{i=1}^n$ be the

eigenbasis of $\lambda A + (1-\lambda)B$, with associated eigenvalues μ_i. Then, consider:

$$\begin{aligned}
\lambda \operatorname{Tr}\big(f(A)\big) + (1-\lambda)\operatorname{Tr}\big(f(B)\big) &= \lambda \sum_i \langle i| \big(\sum_a f(a)\,|a\rangle\langle a|\big)\,|i\rangle + (1-\lambda)\sum_i \langle i| \big(\sum_b f(b)\,|b\rangle\langle b|\big)\,|i\rangle \\
&= \lambda \sum_i \sum_a f(a)\big|\langle i|a\rangle\big|^2 + (1-\lambda)\sum_i \sum_b f(b)\big|\langle i|b\rangle\big|^2 \\
&\geq \lambda \sum_i f\big(\sum_a a\big|\langle i|a\rangle\big|^2\big) + (1-\lambda)\sum_i f\big(\sum_b b\big|\langle i|b\rangle\big|^2\big) \\
&= \lambda \sum_i f(\langle i|\,A\,|i\rangle) + (1-\lambda)\sum_i f(\langle i|\,B\,|i\rangle) \\
&= \sum_i (\lambda f(\langle i|\,A\,|i\rangle) + (1-\lambda) f(\langle i|\,B\,|i\rangle)) \\
&\geq \sum_i f(\lambda\,\langle i|\,A\,|i\rangle + (1-\lambda)\,\langle i|\,B\,|i\rangle) \\
&= \sum_i f(\langle i|\,(\lambda A + (1-\lambda)B)\,|i\rangle) \\
&= \sum_i f(\mu_i) \\
&= \operatorname{Tr} f(\lambda A + (1-\lambda)B)
\end{aligned}$$

where in the first equality we take the trace in the $\big\{|i\rangle\big\}$ basis and write out $f(A), f(B)$ in terms of the eigenbasis of A, B. The inequalities on the third/sixth lines are due to the convexity of f. We thus conclude that $A \mapsto \operatorname{Tr}\big(f(A)\big)$ is convex.

Generalized Klein's inequality. Again consider Hermitian A, B and consider $\lambda A + (1-\lambda)B = B + \lambda C$ for $C = B - A$ and $\lambda \in [0,1]$. Now define $F(\lambda) = \operatorname{Tr}\big(f(B+\lambda C)\big)$, which is convex as trace functions are convex, due to the above Lemma. Thus for any $\lambda \in [0,1]$, convexity tells us that:

$$\lambda F(1) + (1-\lambda)F(0) \geq F(\lambda)$$

which rearranging:

$$F(1) - F(0) \geq \frac{F(\lambda) - F(0)}{\lambda}$$

In particular this holds in the limit where we take $\lambda \to 0$ on the RHS, which we recognize as the derivative of F at zero, and so:

$$F(1) - F(0) \geq F'(0)$$

We can identify:

$$F(1) = \operatorname{Tr}\big(f(B+C)\big) = \operatorname{Tr}\big(f(B+A-B)\big) = \operatorname{Tr}\big(f(A)\big), \quad F(0) = \operatorname{Tr}\big(f(B)\big)$$

and:

$$F'(\lambda) = \operatorname{Tr}\big(C f'(B+\lambda C)\big)$$

from the chain rule, and in particular:

$$F'(0) = \mathrm{Tr}\big(Cf'(B)\big) = \mathrm{Tr}\big((A-B)f'(B)\big)$$

and so substituting these into the inequality:

$$\mathrm{Tr}\big(f(A)\big) - \mathrm{Tr}\big(f(B)\big) = \mathrm{Tr}\big(f(A) - f(B)\big) \geq \mathrm{Tr}\big((A-B)f'(B)\big)$$

where we have used the linearity of the trace. □

Problem 11.2: Generalized relative entropy

($\star$) The definition of the relative entropy may be extended to apply to any two positive operators r and s,

$$S(r\|s) \equiv \mathrm{Tr}(r\log r) - \mathrm{Tr}(r\log s).$$

The earlier argument proving joint convexity of the relative entropy goes directly through for this generalized definition:

(1) For $\alpha, \beta > 0$ show that

$$S(\alpha r\|\beta s) = \alpha S(r\|s) + \alpha\,\mathrm{Tr}(r)\log\big(\alpha/\beta\big).$$

(2) Prove that the joint convexity of the relative entropy implies the subadditivity of the relative entropy,

$$S(r_1 + r_2\|s_1 + s_2) \leq S(r_1\|s_1) + S(r_2\|s_2).$$

(3) Prove that subadditivity of the relative entropy implies joint convexity of the relative entropy.

(4) Let p_i and q_i be probability distributions over the same set of indices. Show that

$$S\left(\sum_i p_i r_i \Big\| \sum_i q_i s_i\right) \leq \sum_i p_i S(r_i\|s_i) + p_i\,\mathrm{Tr}(r_i)\log\big(p_i/q_i\big).$$

In the case where the r_i are density operators so $\mathrm{Tr}(r_i) = 1$, this reduces to the pretty formula

$$S\left(\sum_i p_i r_i \Big\| \sum_i q_i s_i\right) \leq \sum_i p_i S(r_i\|s_i) + H(p_i\|q_i),$$

where $H(\cdot\|\cdot)$ is the Shannon relative entropy.

Solution

Concepts Involved: Von Neumann Entropy, Shannon Entropy, Relative Entropy, Joint Convexity, Subadditivity

(1) First, note that the expression is sensible, since if r, s are positive then $\alpha r, \beta s$ are positive so long

as $\alpha, \beta > 0$. Using the definition, log laws, and the linearity of the trace:

$$\begin{aligned} S(\alpha r \| \beta s) &= \operatorname{Tr}(\alpha r \log \alpha r) - \operatorname{Tr}(\alpha r \log \beta s) \\ &= \operatorname{Tr}\big(\alpha r (\log \alpha + \log r)\big) - \operatorname{Tr}\big(\alpha r (\log \beta + \log s)\big) \\ &= \alpha \log \alpha \operatorname{Tr}(r) + \alpha \operatorname{Tr}(r \log r) - \alpha \log \beta \operatorname{Tr}(r) - \alpha \operatorname{Tr}(r \log s) \\ &= \alpha(\operatorname{Tr}(r \log r) - \operatorname{Tr}(r \log s)) + \alpha \operatorname{Tr}(r) \log\big(\alpha / beta\big) \\ &= \alpha S(r \| s) + \alpha \operatorname{Tr}(r) \log\big(\alpha/\beta\big). \end{aligned}$$

(2) First note that for any $\alpha > 0$:

$$S(\alpha r \| \alpha s) = \alpha S(r \| s) + \alpha \operatorname{Tr}(r) \log\big(\alpha/\alpha\big) = \alpha S(r \| s) \qquad (*)$$

since $\log(1) = 0$. We will use this identity for (2) and (3).

Suppose $S(\cdot\|\cdot)$ is jointly convex. Then, take any $\lambda \in (0,1)$, and consider:

$$\begin{aligned} S(r_1 + r_2 \| s_1 + s_2) &= S(\lambda \frac{r_1}{\lambda} + (1-\lambda)\frac{r_2}{1-\lambda} \| \lambda \frac{s_1}{\lambda} + (1-\lambda)\frac{s_2}{1-\lambda}) \\ &\leq \lambda S(\frac{r_1}{\lambda} \| \frac{s_1}{\lambda}) + (1-\lambda) S(\frac{r_2}{1-\lambda} \| \frac{s_2}{1-\lambda}) \\ &= \lambda \frac{1}{\lambda} S(r_1 \| s_1) + (1-\lambda)\frac{1}{1-\lambda} S(r_2 \| s_2) \\ &= S(r_1 \| s_1) + S(r_2 \| s_2) \end{aligned}$$

where the inequality is from joint convexity, and the third line is using $(*)$. Thus we have subadditivity.

(3) For any $\lambda \in (0,1)$:

$$\begin{aligned} S(\lambda r_1 + (1-\lambda) r_2 \| \lambda s_1 + (1-\lambda) s_2) &\leq S(\lambda r_1 \| \lambda s_1) + S((1-\lambda) r_2 \| (1-\lambda) s_2) \\ &= \lambda S(r_1 \| s_1) + (1-\lambda) S(r_2 \| s_2) \end{aligned}$$

where the inequality follows from subadditivity and the second line is using $(*)$. If $\lambda = 0, 1$ then evidently the above inequality holds with equality. Thus we have joint convexity.

(4) We have:

$$\begin{aligned} S\left(\sum_i p_i r_i \| \sum_i q_i s_i\right) &\leq \sum_i S(p_i r_i \| q_i s_i) \\ &= \sum_i p_i S(r_i \| s_i) + p_i \operatorname{Tr}(r_i) \log\big(p_i/q_i\big) \end{aligned}$$

where the inequality is due to subadditivity and the second line is using (1). If $\mathrm{Tr}(r_i) = 1$, then:

$$S\left(\sum_i p_i r_i \| \sum_i q_i s_i\right) \leq \sum_i p_i S(r_i \| s_i) + p_i \log(p_i/q_i)$$
$$= \sum_i p_i S(r_i \| s_i) + H(p_i \| q_i)$$

as claimed.

□

Problem 11.3: Analogue of the triangle inequality for conditional entropy

($\star\star$)

(1) Show that $H(X,Y|Z) \geq H(X|Z)$.

(2) Show that it is not always true that $S(A,B|C) \geq S(A|C)$.

(3) Prove the conditional version of the triangle inequality,

$$S(A,B|C) \geq S(A|C) - S(B|C).$$

The claim of (3) is false. We give a counterexample and then show instead that:

$$S(A,B) \geq \left|S(A|C) - S(B|C)\right|$$

Solution

Concepts Involved: Conditional Entropy, Shannon Entropy, Von Neumann Entropy

(1) From the definitions of the conditional entropy, we have

$$H(X,Y|Z) = H(X,Y,Z) - H(Z), \quad H(X|Z) = H(X,Z) - H(Z).$$

From Theorem 11.3(3), we know that $H(W,V) \geq H(W)$ (with equality iff W is a function of V), so taking W to be the joint random variable (X,Z) and V to be Y, we find:

$$H(X,Y,Z) \geq H(X,Z)$$

and so combining this with the above definitions of the conditional entropy, we conclude:

$$H(X,Y|Z) \geq H(X|Z)$$

(2) Consider the three qubit GHZ state $|\psi\rangle_{ABC} = \frac{|0\rangle_A|0\rangle_B|0\rangle_C + |1\rangle_A|1\rangle_B|1\rangle_C}{\sqrt{2}}$. Then we observe that

$S(A,B,C) = 0$ (as $|\psi\rangle_{\text{ABC}}$ is pure) but:

$$\rho_{AC} = \text{Tr}_B(|\psi\rangle\langle\psi|_{ABC}) = \frac{|00\rangle\langle 00|_{AC} + |11\rangle\langle 11|_{AC}}{2}$$

and so:

$$S(A,C) = S(\rho_{AC}) = -\frac{1}{2}\log\left(\frac{1}{2}\right) - \frac{1}{2}\log\left(\frac{1}{2}\right) = \log(2) = 1$$

thus $S(A,C) > S(A,B,C)$, and so:

$$S(A,C) - S(C) > S(A,B,C) - S(C) \implies S(A|C) > S(A,B|C).$$

(3) Consider again the three qubit GHZ state, where we note that $\rho_C = \frac{|0\rangle\langle 0|_C + |1\rangle\langle 1|_C}{2} = I/2$ and so $S(C) = \log(2) = 1$.

We thus find $S(A,B|C) = S(A,B,C) - S(C) = 0 - 1 = -1$, $S(A|C) = S(A,C) - S(C) = 1 - 1 = 0$ (having computed $S(A,C) = 1$ in the previous part). By symmetry, $S(B|C) = 0$ identically, and since $-1 \not\geq 0$:

$$S(A,B|C) \not\geq \big|S(A|C) - S(B|C)\big|$$

and so the claim is false.

Instead, let us show:

$$S(A,B) \geq \big|S(A|C) - S(B|C)\big|$$

From strong subadditivity, we have:

$$S(A,B) + S(B,C) \geq S(A,B,C) + S(B)$$

and then applying the triangle inequality to the RHS, we get:

$$S(A,B) + S(B,C) \geq S(A,C) - S(B) + S(B) = S(A,C)$$

Thus:

$$S(A,B) \geq S(A,C) - S(B,C)$$

The argument is symmetric under interchange of $A \leftrightarrow B$ and so $S(A,B) \geq S(B,C) - S(A,C)$ as well, and hence:

$$S(A,B) \geq \big|S(A,C) - S(B,C)\big| = \big|(S(A,C) - S(C)) - (S(B,C) - S(C))\big| = \big|S(A|C) - S(B|C)\big|.$$

as claimed.

□

Problem 11.4: Conditional forms of strong subadditivity

($\star$)

(1) Prove that $S(A,B,C|D) + S(B|D) \leq S(A,B|D) + S(B,C|D)$.

(2) Show by explicit example that it is not always true that $H(D|A,B,C) + H(D|B) \leq H(D|A,B) + H(D|B,C)$.

Solution

Concepts Involved: Von Neumann Entropy, Shannon Entropy, Strong Subadditivity, Conditional Entropy

(1) Consider the strong subadditivity inequality:

$$S(E,F,G) + S(F) \leq S(E,F) + S(F,G)$$

Now let $E = A, F = BD, G = C$, so:

$$S(A,B,C,D) + S(B,D) \leq S(A,B,D) + S(B,C,D)$$

subtracting $2S(D)$ from both sides:

$$(S(A,B,C,D) - S(D)) + (S(B,D) - S(D)) \leq (S(A,B,D) - S(D)) + (S(B,C,D) - S(D))$$

and thus from the definition of conditional entropy:

$$S(A,B,C|D) + S(B|D) \leq S(A,B|D) + S(B,C|D).$$

(2) Let A, B be coinflips with $P(A=0) = P(A=1) = P(B=0) = P(B=1) = \frac{1}{2}$, and $D = C = A$. A, C are identical and fully specify D, implying $H(D|A,B,C) = H(D|A,B) = H(D|B,C) = 0$. Since D, B are uncorrelated coinflips, $H(D|B) = H(B,D) - H(B) = 2\log 2 - \log 2 = \log 2 = 1$, and hence:

$$H(D|A,B,C) + H(D|B) = 1 \not\leq 0 = H(D|A,B) + H(D|B,C)$$

so the inequality is violated.

□

Problem 11.5: Strong subadditivity – Research

($\star\star\star$) Find a simple proof of the strong subadditivity inequality for quantum entropies.

Solution

Concepts Involved: Von Neumann Entropy, Subadditivity

One such proof is given by Nielsen in arXiv:quant-ph/0408130, elaborating on the construction of Petz

(Quasi-entropies for finite quantum systems. Rep. Math. Phys., 23(1):57–65, 1986.). Two elementary proofs are also provided by Ruskai, in arXiv:quant-ph/0404126 and arXiv:quant-ph/0604206. □

12 Quantum information theory

Exercise 12.1

Suppose $|\psi\rangle$ and $|\varphi\rangle$ are two orthogonal quantum states of a single qubit. Design a quantum circuit with two input qubits (the 'data' and the 'target' qubits), with the data qubit in either the state $|\psi\rangle$ or $|\varphi\rangle$, and the target qubit prepared in the standard state $|0\rangle$, which produces as output $|\psi\rangle|\psi\rangle$ or $|\varphi\rangle|\varphi\rangle$, depending on whether $|\psi\rangle$ or $|\varphi\rangle$ was input to the data qubit.

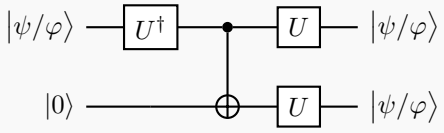
Solution

Concepts Involved: No-Cloning Theorem, Unitary Operators, Controlled Operations

Consider the single qubit unitary:

$$U = |\psi\rangle\langle 0| + |\varphi\rangle\langle 1|$$

where evidently $U\,|0/1\rangle = |\psi/\varphi\rangle$. Then the cloning circuit takes the form:



□

Exercise 12.2

Define U_y to be the unitary operator acting on system M whose action on a basis is $U_y\,|y'\rangle \equiv |y' + y\rangle$, where the addition is done modulo $n+1$. Show that $\left\{\sqrt{E_y} \otimes U_y\right\}$ is a set of operation elements defining a trace-preserving quantum operation $\mathcal{E}$ whose action on states of the form $\sigma \otimes |0\rangle\langle 0|$ agreed with (12.8).

Solution

Concepts Involved: Operator-Sum Representation, Quantum Operations

To check that $\mathcal{E}$ is trace-preserving, we let $F_y = \sqrt{E_y} \otimes U_y$ and confirm that $\sum_y F_y^\dagger F_y = I$. To this end:

$$\begin{aligned}
\sum_y F_y^\dagger F_y &= \sum_y (\sqrt{E_y} \otimes U_y)^\dagger (\sqrt{E_y} \otimes U_y) \\
&= \sum_y (\sqrt{E_y^\dagger} \otimes U_y^\dagger)(\sqrt{E_y} \otimes U_y) \\
&= \sum_y \sqrt{E_y E_y} \otimes U_y^\dagger U_y \\
&= \sum_y E_y \otimes I \\
&= I \otimes I \\
&= I
\end{aligned}$$

In the second equality we use that the E_y are positive and so the dagger and square root can be interchanged, in the third equality we use that the $E_y (= M_y^\dagger M_y)$ are Hermitian, and in the fifth equality we use the completeness relation of the POVM.
Finally, we check:

$$\begin{aligned}
\mathcal{E}(\sigma \otimes |0\rangle\langle 0|) &= \sum_y F_y (\sigma \otimes |0\rangle\langle 0|) F_y^\dagger \\
&= \sum_y \sqrt{E_y} \sigma \sqrt{E_y^\dagger} \otimes U_y |0\rangle\langle 0| U_y^\dagger \\
&= \sum_y \sqrt{E_y} \sigma \sqrt{E_y} \otimes |y\rangle\langle y|
\end{aligned}$$

which is the claimed result. □

Exercise 12.3

($\star$) Use the Holevo bound to argue that n qubits can not be used to transmit more than n bits of classical information.

Solution

Concepts Involved: Holevo Bound, Mutual Information, Von Neumann Entropy

Consider a scenario where Alice tries to transmit information to Bob by sending him n qubits. Bob's accessible information is the maximum over the mutual information $\max H(X : Y)$ over all possible measurement schemes Bob may preform. But the mutual information for any measurement scheme is bounded by Holevo - If Alice sends a state $\rho = \sum_x p_x \rho_x$:

$$H(X : Y) \leq S(\rho) - \sum_x p_x S(\rho_x)$$

and so:

$$\max H(X:Y) \leq S(\rho) - \sum_x p_x S(\rho_x)$$

Now, observe that the RHS/χ-quantity is maximized by taking ρ to be the maximally mixed state on n qubits as a uniform mixture of pure states drawn from an orthonormal basis - i.e.:

$$\rho = \frac{I}{2^n} = \frac{1}{2^n} \sum_{i=0}^{2^n-1} |0\rangle\langle 0|$$

The maximally mixed state is the unique state of maximal entropy (shown in Ex. 11.19 and in the text), and moreover $\sum_i p_i S(\rho_i) \geq 0$ (as a sum of non-negative quantities), but here $\rho_i = |i\rangle\langle i|$ pure so $S(\rho_i) = 0$ vanishes for each term and the entire sum is zero. Taken together, this shows the χ quantity is maximized by considering the maximally mixed state, and in particular:

$$S(I/2^n) - \sum_i \frac{1}{2^n} S(|i\rangle\langle i|) = \log(2^n) - \sum_i \frac{1}{2^n} \cdot 0 = n$$

This upper bounds the χ-quantity of any other state ρ that Alice may send, and so:

$$\max H(X:Y) \leq S(\rho) - \sum_x p_x S(\rho_x) \leq n$$

and so the accessible information is bounded:

$$\max H(X:Y) \leq n$$

Showing that n qubits cannot transmit more than n classical bits of classical information. □

Exercise 12.4

($\star\star$) Suppose Alice sends Bob an equal mixture of the four pure states

$$\begin{aligned}
|X_1\rangle &= |0\rangle \\
|X_2\rangle &= \sqrt{\frac{1}{3}}\left[|0\rangle + \sqrt{2}|1\rangle\right] \\
|X_3\rangle &= \sqrt{\frac{1}{3}}\left[|0\rangle + \sqrt{2}e^{2\pi i/3}|1\rangle\right] \\
|X_4\rangle &= \sqrt{\frac{1}{3}}\left[|0\rangle + \sqrt{2}e^{4\pi i/3}|1\rangle\right]
\end{aligned}$$

Show that the maximum mutual information between Bob's measurement and Alice's transmission is less than one bit. A POVM which achieves ≈ 0.415 bits is known. Can you construct this or better yet, one which achieves the Holevo bound?

Solution

Concepts Involved: Holevo Bound, Mutual Information, Von Neumann Entropy, POVM Measurement

The first claim already follows from the previous exercise - Alice sends only one qubit, and so cannot transmit more than one bit of classical information. But let's go through the calculation, for completeness. Alice sends the state:

$$\begin{aligned}
\rho &= \sum_i \frac{1}{4} |X_i\rangle\langle X_i| \\
&= \frac{1}{4} |0\rangle\langle 0| + \frac{1}{12}(|0\rangle\langle 0| + \sqrt{2}\,|0\rangle\langle 1| + \sqrt{2}\,|1\rangle\langle 0| + 2\,|1\rangle\langle 1|) \\
&+ \frac{1}{12}(|0\rangle\langle 0| + \sqrt{2}e^{-2\pi i/3}\,|0\rangle\langle 1| + \sqrt{2}e^{2\pi i/3}\,|1\rangle\langle 0| + 2\,|1\rangle\langle 1|) \\
&+ \frac{1}{12}(|0\rangle\langle 0| + \sqrt{2}e^{-4\pi i/3}\,|0\rangle\langle 1| + \sqrt{2}e^{4\pi i/3}\,|1\rangle\langle 0| + 2\,|1\rangle\langle 1|) \\
&= \frac{1}{2} |0\rangle\langle 0| + \frac{1}{2} |1\rangle\langle 1| \\
&= \frac{I}{2}
\end{aligned}$$

where we have used that sums of roots of unity are zero. The maximal mutual information can be bounded by Holevo:

$$\max H(X:Y) \leq S(\rho) - \sum_{i=1}^{4} \frac{1}{4} S(\,|X_i\rangle\langle X_i|) = S(\frac{I}{2}) - \sum_{i=1}^{4} \frac{1}{4} \cdot 0 = \log 2 = 1$$

where we have used that the entropy of a pure state is zero. Thus the maximal mutual information between the measurement and transmission is less than one bit.

Moving to the second part of the question, a decent first guess would be to construct a POVM from the $|X_i\rangle$ themselves - however, since none of the $|X_i\rangle$ are mutually orthogonal, we wouldn't expect this to be very effective at distinguishing Alice's state (indeed, taking $E_i = \frac{1}{2}\,|X_i\rangle\langle X_i|$ in the proceeding calculation gives a less optimal result of $H(X:Y) \approx 0.208$). A better guess is then to construct a POVM from states $|Y_i\rangle$ orthogonal to the $|X_i\rangle$ (i.e. $\langle Y_i|X_i\rangle = 0$), as if Bob were to measure $|Y_i\rangle$ he would know that Alice did not send $|X_i\rangle$. The $|Y_i\rangle$ states are:

$$\begin{aligned}
|Y_1\rangle &= |1\rangle \\
|Y_2\rangle &= \sqrt{\frac{1}{3}}\left[\sqrt{2}\,|0\rangle - |1\rangle\right] \\
|Y_3\rangle &= \sqrt{\frac{1}{3}}\left[\sqrt{2}\,|0\rangle - e^{-2\pi i/3}\,|1\rangle\right] \\
|Y_4\rangle &= \sqrt{\frac{1}{3}}\left[\sqrt{2}\,|0\rangle - e^{-4\pi i/3}\,|1\rangle\right]
\end{aligned}$$

Analagous to the calculation for Alice's state, we find that $\sum_i \frac{1}{4}\,|Y_i\rangle\langle Y_i| = \frac{I}{2}$ so the appropriate POVM elements to satisfy completeness (choosing the coefficients to be uniform) are $E_i = \frac{1}{2}\,|Y_i\rangle\langle Y_i|$.

Recall from chapter 11 that the classical mutual information can be written as:

$$\begin{aligned} H(X:Y) &= H(p(x,y)\|p(x)p(y)) \\ &= \sum_{x,y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \\ &= \sum_{x,y} p(y|x)p(x)\log\frac{p(y|x)p(x)}{p(x)p(y)} \\ &= \sum_{x,y} p(y|x)p(x)\log\frac{p(y|x)}{p(y)} \end{aligned}$$

Since Alice/Bob pick one of the four $|X_i\rangle$ / $|Y_j\rangle$ to measure with equal probability $p_i = p_j = \frac{1}{4}, \forall i,j \in \{1,2,3,4\}$, the only thing left to determine in the above expression are the conditional probabilities $p(y|x) = \mathrm{Tr}(E_j\rho_i) = \frac{1}{2}\left|\langle Y_j|X_i\rangle\right|^2$. By construction, the $|X_i\rangle, |Y_j\rangle$ are orthogonal for $i=j$ so $p(x|x) = 0$. For the other 12 pairs, we find that $\left|\langle Y_i|X_j\rangle\right|^2 = \frac{2}{3}$ for $i \neq j$ and so:

$$p(y|x) = \begin{cases} 0 & y = x \\ \frac{1}{3} & \text{otherwise} \end{cases}$$

Thus computing the mutual information, we have 4 terms that contribute nothing and 12 terms that contribute with equal weight with $p(y|x) = 1/3$ and $p(y) = p(x) = 1/4$:

$$H(X:Y) = 12 \cdot \frac{1}{3}\cdot\frac{1}{4}\log\left(\frac{\frac{1}{3}}{\frac{1}{4}}\right) = \log\left(\frac{4}{3}\right) \approx 0.415$$

which is the desired result.
A POVM that saturates the Holevo bound does not exist in this case, as this requires the states in Alice's ensemble to be mutually orthogonal, which they are not. □

Exercise 12.5: Variable-length zero error data compression

Consider the following heuristic for a variable length data compression scheme. Let $x_1, \ldots, x_n$ be the output from n uses of an i.i.d. source with entropy rate $H(X)$. If $x_1, \ldots, x_n$ is typical, then send a $H(X)$ bit index indicating which typical sequence it is. If $x_1, \ldots, x_n$ is atypical, send an uncompressed $\log d^n$ bit index for the sequence (recall that d is the alphabet size). Turn this heuristic into a rigorous argument that the source can be compressed to an average of R bits per source symbol, for any $R > H(X)$, with zero probability of error.

Solution

Concepts Involved: Shannon Entropy, Typical Sequences, Data Compression

The proof takes after the proof of Shannon's noiseless channel coding theorem.
Let $R > H(X)$, and choose $\epsilon > 0$ such that $H(X) + \epsilon < R$. Consider the set $T(n,\epsilon)$ of ϵ-typical sequences. For any $\delta > 0$ and for n sufficiently large, there are at most $2^{n(H(x)+\epsilon)} < 2^{nR}$ sequences, and

the probability of producing such a sequence is at least $1-\delta$ by the theorem of typical sequences. For any such typical sequence, we send an nR bit index which uniquely specifies the sequence. Otherwise, we send the uncompressed $\log d^n$ bit index for the sequence. The probability of error is zero as in the typical case (where we apply compression) the specifying bit index is unique, and otherwise we simply send the entire message.
Since the probability of producing a typical sequence is at least $1-\delta$, the average length of the sequence L is upper bounded by:

$$L \leq (1-\delta)nR + \delta n \log d$$

and since we always send at least nR bits, we also obtain the lower bound:

$$nR \leq L \leq (1-\delta)nR + \delta n \log d.$$

Dividing out by n, we obtain bounds on the average rate of compression of the source:

$$R \leq \frac{L}{n} \leq (1-\delta)R + \delta \log d$$

Since $\delta > 0$ is arbitrary, it follows that $\frac{L}{n} = R$, as claimed. □

Exercise 12.6

($\star\star$) In the notation of Box 12.4, give an explicit expression for C_X in terms of X. Also, describe how to construct a quantum circuit to perform U_n for arbitrary n. How many elementary operations do you require, as a function of n?
It appears to us that Eq. (12.60) in the text is incorrect as the block of n qubits is described as a pure state, while the i.i.d. quantum source is characterized by the mixed state $\rho = \frac{1}{4}\begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}$. We believe the intended expression was:

$$\sum_{X=\left\{\bar{0}\bar{0}\ldots\bar{0},\bar{0}\ldots\bar{0}\bar{1},\ldots,\bar{1}\bar{1}\ldots\bar{1}\right\}} C_X \left|X\right\rangle\!\left\langle X\right|.$$

Exercise 12.7: Data compression circuit

($\star\star$) Outline the construction of a circuit to reliably compress a qubit source with $\rho = p|0\rangle\langle 0| + (1-p)|1\rangle\langle 1|$ into nR qubits for any $R > S(\rho) = H(p)$.

Exercise 12.8: Compression of an ensemble of quantum states

Suppose that instead of adopting the definition of a quantum source based on a single density matrix ρ and the entanglement fidelity, we instead adopted the following *ensemble* definition, that an (i.i.d.) quantum source is specified by an ensemble $\{p_j, |\psi_j\rangle\}$ of quantum states, and that consecutive used of the source are independent and produce a state $|\psi_j\rangle$ with probability p_j. A compression-decompression scheme $(\mathcal{C}^n, \mathcal{D}^n)$ is said to be reliable in this definition if the *ensemble average fidelity* approaches 1 as $n \to \infty$:

$$\bar{F} \equiv \sum_J p_{j_1} \dots p_{j_n} F(\rho_J, (\mathcal{D}^n \circ \mathcal{C}^n)(\rho_J))^2,$$

Where $J = (j_1, \dots, j_n)$ and $\rho_J \equiv |\psi_{j_1}\rangle\langle\psi_{j_1}| \otimes \dots \otimes |\psi_{j_n}\rangle\langle\psi_{j_n}|$. Define $\rho \equiv \sum_j p_j |\psi_j\rangle\langle\psi_j|$ and show that provided $R > S(\rho)$ there exists a reliable compression scheme of rate R with respect to this definition of fidelity.

Solution

Concepts Involved: Von Neumann Entropy, Quantum Compression, Fidelity

Using the result from Chapter 9.3 (see Eq. (9.141) and its derivation), we can bound the average ensemble fidelity using the entanglement fidelity:

$$F(\sum_J p_{j_1} \dots p_{j_n} \rho_J, \mathcal{D}^n \circ \mathcal{C}^n) \leq \bar{F}$$

We can then identify:

$$\sum_J p_{j_1} \dots p_{j_n} \rho_J = \bigotimes_{i=1}^n \sum_{j_i} p_{j_i} |\psi_{j_i}\rangle\langle\psi_{j_i}| = \rho^{\otimes n}$$

Then using $(\mathcal{C}^n, \mathcal{D}^n)$ as constructed in Schumacher's noiseless channel coding theorem as proven in the text (see Eqs. (12.51)-(12.53), we have:

$$1 - 2\delta \leq F(\rho^{\otimes n}, \mathcal{D}^n \circ \mathcal{C}^n) \leq \bar{F}$$

where δ can be made arbitrarily small for large n, and hence $\bar{F} \to 1$ as $n \to \infty$. Thus, provided $R > S(\rho)$, there exists a reliable compression scheme of rate R with respect to this definition of fidelity - the exact same compression scheme that is used in Schumacher's noiseless channel coding theorem! □

Exercise 12.9

($\star$) The *erasure channel* has two inputs, 0 and 1, and three outputs, $0, 1$ and e. With probability $1-p$. the input is left alone. With probability p the input is 'erased', and replaced by e.

(1) Show that the capacity of the erasure channel is $1-p$.

(2) Prove that the capacity of the erasure channel is greater than the capacity of the binary symmetric channel. Why is this result intuitively plausible?

The claim of (2) is only true if we restrict the range of p-see solution for details.

Solution

Concepts Involved: Channel Capacity, Shannon's Noisy Channel Coding Theorem, Binary Symmetric Channel

Recall that the binary symmetric channel is parameterized by $0 \leq p \leq 1$, and has the action of flipping the transmitted bit with probability p while leaving it invariant with probability $1-p$. It has capacity $1 - H(p)$.

(1) We calculate the capacity of the erasure channel using Shannon's noisy channel coding theorem. Consider an input distribution $p(0) = q$, $p(1) = 1-q$ for some $0 \leq q \leq 1$ (which we will optimize). We study the mutual information:

$$H(X:Y) = H(Y) - H(Y|X) = H(Y) - \sum_x p(x) H(Y|X=x)$$

For fixed a given x, we have that with probability $1-p$ the output of the channel is x and with probability p the output of the channel is e (erased) and thus for all x, $H(Y|X=x) = H(p)$, and thus $\sum_x p(x)H(Y|X=x) = H(p)\sum_x p(x) = H(p)$, and so:

$$H(X:Y) = H(Y) - H(p)$$

analogous to the binary symmetric channel analysis. Now, let's consider the output random variable Y, which for p, q has distribution:

$$p(Y=0) = q(1-p), \quad p(Y=1) = (1-q)(1-p), \quad p(Y=e) = p$$

and thus:

$$\begin{aligned}
H(Y) &= -\sum_y p(y)\log\big(p(y)\big) \\
&= -q(1-p)\log\big(q(1-p)\big) - (1-q)(1-p)\log\big((1-q)(1-p)\big) - p\log(p) \\
&= -q(1-p)\log(q) - q(1-p)\log(1-p) \\
&\quad - (1-q)(1-p)\log(1-q) - (1-q)(1-p)\log(1-p) - p\log(p) \\
&= (1-p)(-q\log(q) - (1-q)\log(1-q)) - (1-p)\log(1-p) - p\log p \\
&= (1-p)H(q) + H(p)
\end{aligned}$$

Thus:

$$H(X:Y) = (1-p)H(q) + H(p) - H(p) = (1-p)H(q)$$

To find $C(\mathcal{N})$ we take the maximum over q, and since $H(q)$ is maximized at $q = 1/2$ for which $H(q = 1/2) = 1$, we obtain:

$$C(\mathcal{N}) = 1 - p$$

as claimed.

(2) Since $H(p)$ is concave (as shown in Ex. **??**), it lies above any chord that connects two points. In particular, noting that $H(0) = 0$ and $H(1/2) = 1$, we can consider the chord $y(p) = 2p$ connecting these two points, and thus for any $p \in [0, 1/2]$:

$$H(p) \geq 2p \geq p$$

and so recalling that the binary symmetric channel has capacity $1 - H(p)$, we have for $p \in [0, 1/2]$ that:

$$C(\mathcal{N}_{\text{bin}}) = 1 - H(p) \leq 1 - p = C(\mathcal{N}_{\text{erasure}})$$

This is intuitively plausible that the erasure channel capacity is larger than the binary symmetric channel, because with the binary symmetric channel there is uncertainty about whether an error has occurred, while the erasure channel signals/outputs e when an error has occurred.

However, the claim does not hold for all $p \in [0, 1]$. Mathematically, this follows from the fact that $H(p)$ is symmetric about $p = 1/2$ and so the capacity $C(\mathcal{N}_{\text{bin}})$ increases over $p \in [1/2, 1]$ back to 1, while the capacity of the erasure channel continues to decrease over $p \in [1/2, 1]$ to 0 at $p = 1$. Numerically, we can find a crossover point at $p_c = 0.7729$ where the binary symmetric channel capacity becomes greater than the erasure channel. The intuition for this result is that for $p > 1/2$, for the binary symmetric channel we can interchange the roles of 0 and 1 (apply a bitflip on the output of the channel), wherein the channel acts just like a binary symmetric channel with $p' = p - 1/2$, while the erasure channel only gets worse as the erasure probability is increased.

□

Exercise 12.10

($\star$) Suppose $\mathcal{N}_1$ and $\mathcal{N}_2$ are two discrete memoryless channels such that the input alphabet of $\mathcal{N}_2$ is equal to the output alphabet of $\mathcal{N}_1$. Show that

$$C(\mathcal{N}_2 \circ \mathcal{N}_1) \leq \min(C(\mathcal{N}_1), C(\mathcal{N}_2)).$$

Find an example where the inequality is strict.

Solution

Concepts Involved: Channel Capacity, Shannon's Noisy Channel Coding Theorem, Binary Symmetric Channel

Consider the channel $\mathcal{N}_2 \circ \mathcal{N}_1$; let X be the random channel Let Y be the random variable at the output of channel $\mathcal{N}_1$ and Z be the random variable at the output of $\mathcal{N}_2 \circ \mathcal{N}_1$. Since $\mathcal{N}_1, \mathcal{N}_2$ are memoryless, it follows that $X \to Y \to Z$ is a Markov chain, and from the Data processing inequality (Theorem 11.5) we have:

$$H(X:Y) \geq H(X:Z).$$

Additionally, from Ex. 11.10 we know that $Z \to Y \to X$ is also a Markov chain, and so we also find from the data processing inequality:

$$H(Z:Y) = H(Y:Z) \geq H(Z:X) = H(X:Z)$$

We can then identify $C(\mathcal{N}_2 \circ \mathcal{N}_1) = \max_{p(x)} H(X:Z)$ and $C(\mathcal{N}_1) = \max_{p(x)} H(X:Y)$, $C(\mathcal{N}_2) = \max_{p(y)} H(Y:Z)$, and from the two inequalities above it follows that:

$$C(\mathcal{N}_2 \circ \mathcal{N}_1) \leq \min(C(\mathcal{N}_1), C(\mathcal{N}_2))$$

As an example, consider the composition of two identical binary symmetric channels $\mathcal{N}_1(p) = \mathcal{N}_2(p)$ with $0 < p < 1/2$, which were shown to have $C(\mathcal{N}(p)) = 1 - H(p)$ and so $\min(C(\mathcal{N}_1(p)), C(\mathcal{N}_2(p))) = 1 - H(p)$. The composite channel $\mathcal{N}_2(p) \circ \mathcal{N}_1(p)$ maps an input $x = 0$ to an output $z = 0$ with probability:

$$\begin{aligned} p(z=0|x=0) &= p(z=0|y=0)p(y=0|x=0) + p(z=0|y=1)p(y=1|x=0) \\ &= p^2 + (1-p)^2 \\ &= 1 - 2p + 2p^2 \end{aligned}$$

and maps input $x = 0$ to output $z = 1$ with probability

$$\begin{aligned} p(z=1|x=0) &= p(z=1|y=0)p(y=0|x=0) + p(z=1|y=1)p(y=1|x=0) \\ &= (1-p)p + p(1-p) \\ &= 2p - 2p^2 \end{aligned}$$

and analogously for $p(z=1|x=1), p(z=0|x=1)$. In other words $\mathcal{N}_2(p) \circ \mathcal{N}_1(p)$ is just another binary symmetric channel with probability $q = 2p - 2p^2$, and thus with capacity $C(\mathcal{N}_2 \circ \mathcal{N}_1) = 1 - H(q)$.
We then note that:

$$q - p = 2p - 2p^2 - p = p(1-2p)$$

which is strictly positive for all $0 < p < 1/2$, so q is greater than p for this range. Since the binary entropy is a strictly increasing function on $(0, 1/2)$, we have:

$$H(q) > H(p) \implies 1 - H(q) < 1 - H(p)$$

and so the inequality is strict in this case. □

Exercise 12.11

($\star$) Show that the maximum in the expression (12.71) may be achieved using an ensemble of pure states. Show further that it suffices to consider only ensembles of at most d^2 pure states, where d is the dimension of the input to the channel.

Solution

Concepts Involved: Von Neumann Entropy, Pure States, Quantum Operations. (12.71) reads:

$$\chi(\mathcal{E}) = \max_{\{p_j,\rho_j\}} \left[S\left(\mathcal{E}\left(\sum_j p_j \rho_j \right) \right) - \sum_j p_j S(\mathcal{E}(\rho_j)) \right]$$

Consider any $\rho = \sum_i p_i \rho_i$; by the linearity of $\mathcal{E}$ and the concavity of the Von Neumann Entropy (see section 11.3.5), we obtain:

$$S(\mathcal{E}(\rho)) = S(\mathcal{E}(\sum_i p_i \rho_i)) = S(\sum_i p_i \mathcal{E}(\rho_i)) \geq \sum_i p_i \mathcal{E}(\rho_i)$$

Now, consider the ensemble $E = \{(p_1, \rho_1), \ldots, (p_{n-1}, \rho_{n-1}), (p_n, \rho_n)\}$ and suppose ρ_n is a mixed state. Then, let us write $\rho_n = \sum_{i=1}^m p_{n,i} \rho_{n,i}$ for $\rho_{n,i}$ pure, and consider the ensemble $E' = \{(p_1, \rho_1), \ldots (p_{n-1}, \rho_{n-1}), (p_n p_{n,1}, \rho_{n,1}), \ldots (p_n p_{n,m}, \rho_{n,m})\}$.
We can then see that the state $\sum_j p_j \rho_j$ computed for E, E' are identical, so the first term in the expression of (12.71) is the same across the two ensembles. However, by the above concavity argument, the second term is smaller for E' (thus less negative) and hence the overall expression is larger. For any ensemble containing mixed states, we can thus replace it with a pure state ensemble in this fashion, wherein the expression increases. Since for any ensemble containing mixed states we can construct a pure state ensemble which upper bounds the mixed state ensemble expression, we conclude that the maximum may be achieved with a pure state ensemble.
For the second part of the question - if d is the dimension of the input to the channel, consider some set of linearly independent and pure $\rho_1, \ldots, \rho_{d^2}$ which form a basis of the space of $d \times d$ dimensional Hermitian positive operators (wherein any ρ can be written as $\rho = \sum_{i=1}^{d^2} p_i \rho_i$).
Now, consider any ensemble $F' = \{(p'_1, \rho'_1), \ldots (p'_m, \rho'_m)\}$, with $m \geq d^2$. Decomposing every density matrix appearing in this ensemble as:

$$\rho'_j = \sum_{i=1}^{d^2} p_{j,i} \rho_i$$

We can then consider the ensemble $F = \{(p_{1,1}, \rho_1), \ldots, (p_{1,d^2}, \rho_{d^2}), \ldots, (p_{m,1}, \rho_1), \ldots, (p_{m,d^2}, \rho_{d^2})\}$ wherein by the same concavity argument as the first part we have that the expression is larger for F than F'. We can then observe that F is equivalent to the d^2-element ensemble:

$$\left\{ (\sum_{j=1}^m p_{j,1}, \rho_i), \ldots, (\sum_{j=1}^m p_{j,d^2}, \rho_{d^2}) \right\}.$$

Since for any $m \geq d^2$-element ensemble we are able to construct a d^2-element ensemble which upper bounds the m-element ensemble expression, we conclude that it suffices to consider only ensembles of at most d^2 pure states to achieve the maximum. □

Exercise 12.12

($\star$) Adapt the proof of the HSW theorem to find a proof of Shannon's noisy channel coding theorem, simplifying the proof wherever possible.

Exercise 12.13

Show that the entropy exchange is concave in the quantum operation $\mathcal{E}$.

Solution

Concepts Involved: Von Neumann Entropy, Quantum Operations, Entropy Exchange, Purifications, Concavity

By definition, the entropy exchange for a quantum operation $\mathcal{E}$ for input ρ is given by:

$$S(\rho, \mathcal{E}) \equiv S(R', Q')$$

where Q contains ρ, R is the reference system that initially purifies Q, and primes denote that $\mathcal{E}$ has been applied. In particular, given the above definition, we can write:

$$S(\rho, \mathcal{E}) = S(\rho^{Q'R'}) = S(\mathcal{E}^Q \otimes \mathcal{I}^R(|\psi^{QR}\rangle\langle\psi^{QR}|))$$

with $|\psi^{QR}\rangle$ the purification of ρ and $\mathcal{I}_R$ denoting the identity map on the refernece system. Now, consider $\lambda \in [0, 1]$ and any two quantum operations $\mathcal{E}^1, \mathcal{E}^2$. Then:

$$\begin{aligned} S(\rho, \lambda\mathcal{E}^1 + (1-\lambda)\mathcal{E}^2) &= S(([\lambda\mathcal{E}^1 + (1-\lambda)\mathcal{E}^2]_Q \otimes \mathcal{I}_R)(|\psi^{QR}\rangle\langle\psi^{QR}|)) \\ &= S(\lambda\mathcal{E}^1_Q \otimes \mathcal{I}_R(|\psi^{QR}\rangle\langle\psi^{QR}|) + (1-\lambda)\mathcal{E}^2_Q \otimes \mathcal{I}_R(|\psi^{QR}\rangle\langle\psi^{QR}|)) \\ &\geq \lambda S(\mathcal{E}^1_Q \otimes \mathcal{I}_R(|\psi^{QR}\rangle\langle\psi^{QR}|)) + (1-\lambda)S(\mathcal{E}^2_Q \otimes \mathcal{I}_R(|\psi^{QR}\rangle\langle\psi^{QR}|)) \\ &= \lambda S(\rho, \mathcal{E}^1) + (1-\lambda)S(\rho, \mathcal{E}^2) \end{aligned}$$

where the second inequality follows from linearity, and the inequality follows from the concavity of the Von Neumann entropy as shown in section 11.3.5. □

Exercise 12.14

Show that the condition $\rho^{RQ_1} = \rho^R \otimes \rho^{Q_1}$ is also *sufficient* to be able to correct errors on the subsystem Q_1.

Solution

Concepts Involved: Reversibility, Quantum Error Correction, Quantum Data Processing Inequality

We have Q an n qubit system, Q_1 a subsystem of Q, R a reference system that purifies Q, and E the environment.
Suppose $\rho^{RQ_1} = \rho^R \otimes \rho^{Q_1}$. Then, an interaction of Q_1 with the environment E (i.e. errors on the subset Q_1 of the qubits) maps:

$$\rho^{RQ_1} \otimes \rho^E = \rho^R \otimes \rho^{Q_1} \otimes \rho^E \mapsto \rho^{R'} \otimes \rho^{Q_1'E'} = \rho^R \otimes \rho^{Q_1'E'}$$

with the reference system R left untouched, and Q_1 and the environment now generically correlated. In particular, it is the case that $\rho^{R'E'} = \rho^{R'} \otimes \rho^{E'}$, which is the information theoretic reversibility condition (from the quantum data processing inequality) that the noise can be reversed, i.e. the errors on subsystem Q_1 is correctable. □

Exercise 12.15

($\star$) Apply all possible combinations of the subadditivity and strong subadditivity inequalities to deduce other inequalities for the two stage quantum process $\rho \mapsto \rho' = \mathcal{E}_1(\rho) \mapsto \rho'' = (\mathcal{E}_2 \circ \mathcal{E}_1)(\rho)$, expressing the results whenver possible in terms of entropy excahnges and the entropies $S(\rho), S(\rho'), S(\rho'')$. When it is not possible to express a quantity appearing in such an inequality in these terms, give a prescription for calculating the quantity only using a knowledge of ρ and operation elements $\{E_j\}$ for $\mathcal{E}_1$ and $\{F_k\}$ for $\mathcal{E}_2$.

Exercise 12.16

Show that in the case where $\mathcal{R}$ perfectly corrects $\mathcal{E}$ for the input ρ, the inequality

$$S(\rho) - S(\rho') + S(\rho', \mathcal{R}) \geq 0$$

must actually be satisfied with equality.

Solution

Concepts Involved: Von Neumann Entropy, Subadditivity, Reversibility, Entropy Exchange, Quantum Error Correction

Studying how the inequality was derived, for a general quantum operation $\mathcal{E}$ we have (Eq. (12.143)):

$$S(\rho) = S(R) = S(R') = S(Q', E') \leq S(Q') + S(E') = S(\mathcal{E}(\rho)) + S(\rho, \mathcal{E})$$

with the quantum system Q, the reference system R (initially purifying Q) and the environment E, with primes denoting the state of system/reference/environment after the unitary interaction of Q, E mocking up the operation $\mathcal{E}$. The inequality is due to subadditivity.
Applying the above to initial state $\rho' = \mathcal{E}(\rho)$ and the recovery operation $\mathcal{R}$, we find:

$$S(\rho') = S(R) = S(R') = S(Q', E') \leq S(Q') + S(E') = S(\mathcal{R}(\rho')) + S(\rho', \mathcal{R}) = S(\rho) + S(\rho', \mathcal{R})$$

where we have used in the last equality that $\mathcal{R}(\rho') = \mathcal{R}(\mathcal{E}(\rho)) = \rho$ from the correction condition.
Since it is the case that $\mathcal{R}$ perfectly corrects $\mathcal{E}$ for the input ρ, it must be the case that $\rho^{Q',E'} = \rho^{Q'} \otimes \rho^{E'}$ after the correction operation - this is because the correction operation perfectly recovers

ρ, which was initially uncorrelated with the environment system. As we saw in Chapter 11.3.4, the subadditivity inequality holds with equality iff the systems are uncorrelated, and thus:

$$S(\rho') = S(\rho) + S(\rho', \mathcal{R}) \implies S(\rho) - S(\rho') + S(\rho;, \mathcal{R}) = 0.$$

□

Exercise 12.17

($\star$) Show that $x \prec y$ if and only if for all real t, $\sum_{j=1}^{d} \max(x_j - t, 0) \leq \sum_{j=1}^{d} \max(y_j - t, 0)$, and $\sum_{j=1}^{d} x_j = \sum_{j=1}^{d} y_j$.

Solution

Concepts Involved: Majorization

Note that $\sum_{j=1}^{d} x_j = \sum_{j=1}^{d} y_j$ is part of the definition of $x \prec y$, so the nontrivial equivalence we are showing is that:

$$\forall k \in \{1, \dots d\}, \sum_{j=1}^{k} x_j \leq \sum_{j=1}^{k} y_j \iff \forall t, \sum_{j=1}^{d} \max(x_j - t, 0) \leq \sum_{j=1}^{d} \max(y_j - t, 0).$$

Throughout, WLOG we assume x, y are sorted in descending order.
$\boxed{\implies}$: Let $t \in \mathbb{R}$. Then, let $j = k$ be the largest index such that $x_j \geq t$ and $j = k'$ be the largest index such that $y_j \geq t'$. Then:

$$\sum_{j=1}^{d} \max(x_j - t, 0) = \sum_{j=1}^{k} (x_j - t), \quad \sum_{j=1}^{d} \max(y_j - t, 0) = \sum_{j=1}^{k'} (y_j - t)$$

as all terms after k, k' are set to zero by the max.
There are then three cases:

- $\boxed{k = k'}$: In this case:

  $$\sum_{j=1}^{k} (y_j - t) \geq \sum_{j=1}^{k} (x_j - t)$$

  immediately from $x \prec y$.

- $\boxed{k' > k}$ In this case:

  $$\sum_{j=1}^{k'} (y_j - t) \geq \sum_{j=1}^{k} (y_j - t) \geq \sum_{j=1}^{k} (x_j - t)$$

  where the first inequality follows from the omission of (non-negative) terms of the sum and the second from $x \prec y$.

- $\boxed{k > k'}$ In this case

$$\sum_{j=1}^{k'}(y_j - t) \geq \sum_{j=1}^{k}(y_j - t) \geq \sum_{j=1}^{k}(x_j - t)$$

where the first inequality comes from the fact that $(y_j - t)$ for $j = k' + 1, k' + 2, \ldots k$ are non-positive terms by assumption, and hence adding them to the sum can only make the sum smaller. The second inequality is yet again just from $x \prec y$.

Thus in all cases $\sum_{j=1}^{k}(x_j - t) \leq \sum_{j=1}^{k'}(y_j - t)$, and so:

$$\sum_{j=1}^{d}\max(x_j - t, 0) \leq \sum_{j=1}^{d}\max(y_j - t, 0).$$

$\boxed{\Longleftarrow}$: Consider any $k \in \{1, \ldots d\}$, and let $t = y_k$. Then, by assumption we have:

$$\sum_{j=1}^{d}\max(x_j - y_k, 0) + ky_k \leq \sum_{j=1}^{d}\max(y_j - y_k, 0) + ky_k \qquad (*)$$

where we have added ky_k to both sides of the inequality. Then, first note that:

$$\sum_{j=1}^{d}\max(y_j - y_k, 0) + ky_k = \sum_{j=1}^{k}\max(y_j - y_k, 0) + ky_k = \sum_{j=1}^{k}(y_j - y_k) + ky_k = \sum_{j=1}^{k} y_j$$

where we have used the fact that $y_j - y_k < 0$ for $j > k$ (and so the max picks out the zero) and $y_j - y_k \geq 0$ (and so the max picks out $y_j - y_k$) for $j \leq k$. Next note that:

$$\sum_{j=1}^{d}\max(x_j - y_k, 0) + ky_k \geq \sum_{j=1}^{k}\max(x_j - y_k, 0) + ky_k \geq \sum_{j=1}^{k}(x_j - y_k) + ky_k = \sum_{j=1}^{k} x_j$$

where the first inequality follows from the fact that we omit non-negative ($j > k$) terms from the sum, and the second inequality follows from the fact that choosing any element from a maximum will always be less than the maximum. Substituting in these two observations into either side of $(*)$, we obtain:

$$\sum_{j=1}^{k} x_j \leq \sum_{j=1}^{k} y_j$$

as desired. □

Exercise 12.18

Use the previous exercise to show that the set of x such that $x \prec y$ is convex.

Solution

Concepts Involved: Convexity, Majorization

It will be easier to show the result directly from the definition of majorization. Suppose $x' \prec y$ and $x'' \prec y$, and WLOG suppose x', x'', y are all sorted in descending order.
Then, consider $x = \lambda x' + (1-\lambda)x''$ for $0 \leq \lambda \leq 1$ - if x', x'' are sorted in descending order, then so shall be x.
Since $x' \prec y$ and $x'' \prec y$, we have for all $1 \leq k \leq d$ that:

$$\sum_{j=1}^{k} x'_j \leq \sum_{j=1}^{k} y_j, \quad \sum_{j=1}^{k} x''_j \leq \sum_{j=1}^{k} y_j$$

with equality if $k = d$. Multiplying the left inequality on both sides by λ and the right inequality on both sides with $(1-\lambda)$, we get:

$$\sum_{j=1}^{k} \lambda x'_j \leq \lambda \sum_{j=1}^{k} y_j, \quad \sum_{j=1}^{k} (1-\lambda) x''_j \leq (1-\lambda) \sum_{j=1}^{k} y_j$$

And adding the two inequalities together:

$$\sum_{j=1}^{k} (\lambda x'_j + (1-\lambda)x''_j) = \sum_{j=1}^{k} x_j \leq \lambda \sum_{j=1}^{k} y_j + (1-\lambda) \sum_{j=1}^{k} y_j = \sum_{j=1}^{k} y_j$$

again with equality when $k = d$. Thus we conclude $x \prec y$ and so the set of x majorized by y is convex. □

Exercise 12.19

($\star$) Verify that $x' \prec y'$.

Solution

Concepts Involved: Majorization

We show that $\sum_{i=1}^{d} x'_i = \sum_{i=1}^{d} y'_i$ and that $\sum_{i=1}^{k} x'_i \leq \sum_{i=1}^{k} y'_i$ for $k = 1, \ldots d$.
For the first statement (equality for the full sum), we note that:

$$\sum_i x'_i = \sum_{i=2}^{d+1} x_i = \sum_{i=1}^{d+1} x_i - x_1$$

$$\sum_i y'_i = \sum_{i=2}^{d+1} Dy = \sum_{i=1}^{d+1} (Dy)_i - x_1$$

then, recalling that $D = tI + (1-t)T$ (for T the permutation swapping the 1st and jth elements):

$$\sum_{i=1}^{d+1}(Dy)_i = t\sum_{i=1}^{d+1}(Iy)_i + (1-t)\sum_{i=1}^{d+1}(Ty)_i = t\sum_{i=1}^{d+1}y_i + (1-t)\sum_{i=1}^{d+1}y_i = \sum_{i=1}^{d+1}y_i$$

where we have used that the sum over all elements of a permuted vector is equal to the sum over all elements of the original vector. Hence:

$$\sum_i y'_i = \sum_{i=1}^{d+1} y_i - x_1.$$

Since $x \prec y$, $\sum_{i=1}^{d+1} x_i = \sum_{i=1}^{d+1} y_i$, and hence:

$$\sum_i x'_i = \sum_{i=2}^{d+1} x_i = \sum_{i=1}^{d+1} x_i - x_1 = \sum_{i=1}^{d+1} y_i - x_1 = \sum_i y'_i.$$

For the second statement (inequality of partial sums), we consider two cases:
case 1: $k < j-1$ For this case, we note that $x_1 \le y_{j-1}$, and so $x'_i \le x_1 \le y'_i$ for all $i = 1, \ldots, j-2$, and hence (noting that x' is already sorted in descending order):

$$\sum_{i=1}^{k}(x'_i)^{\downarrow} = \sum_{i=1}^{k} x'_i \le \sum_{i=1}^{k} y'_i \le \sum_{i=1}^{k}(y'_i)^{\downarrow}$$

case 2: $k \ge j-1$ In this case, we have:

$$\begin{aligned}
\sum_{i=1}^{k}(x'_i)^{\downarrow} = \sum_{i=1}^{k} x'_i = \sum_{i=2}^{k+1} x_i = \sum_{i=1}^{k+1} x_i - x_1 &\le \sum_{i=1}^{k+1} y_i - x_1 \\
&= \sum_{i=2}^{k+1} y_i + (y_1 - x_1) \\
&= \sum_{i=2}^{k+1} y_i + (y_1 - (ty_1 + (1-t)y_j)) \\
&= \sum_{i=2}^{k+1} y_i + (1-t)y_1 + ty_j - y_j \\
&= \sum_{i=1}^{k} y'_i \\
&\le \sum_{i=1}^{k}(y'_i)^{\downarrow}
\end{aligned}$$

where the inequality in the first line follows from $x \prec y$, we substitute in $x_1 = ty_1 + (1-t)y_j$ in the third line, and in the second-to-last line we observe that the $(1-t)y_1 + ty_j$ replaces y_j in y' (and otherwise y_{i+1}, y'_i are identical) and since $\sum_{i=2}^{k+1} y_i$ contains y_j by assumption, we can replace the sum over y_i with the sum over y_i.

Having covered both cases, we conclude $x' \prec y'$. □

Exercise 12.20

($\star$) Show that the assumption that ρ_ψ is invertible may be removed from the proof of the converse part of Theorem 12.15.

Solution

Concepts Involved: Majorization, Density Operators, Measurement Operators

Since ρ_ψ is Hermitian, we can write it in its diagonal form as (taking it to be a d-dimensional state):

$$\rho_\psi = \sum_{i=1}^{d} \lambda_i \, |i\rangle\langle i|$$

with $\left\{|i\rangle\right\}_{i=1}^{d}$ the eigenvectors. In the case that ρ_ψ is invertible all the λ_i are nonzero, but in the general case it is possible that some of the λ_i are zero. In this case, suppose $\lambda_1, \ldots, \lambda_m$ are the nonzero eigenvalues, and we can write:

$$\rho_\psi = \sum_{i=1}^{m} \lambda_i \, |i\rangle\langle i| \left(+ \sum_{i=m+1}^{d} 0 \cdot |i\rangle\langle i| \right)$$

We can then define the pseudo-inverse of ρ_ψ to be:

$$\bar{\rho}_\psi = \sum_{i=1}^{m} \frac{1}{\lambda_i} \, |i\rangle\langle i|$$

which we can see satisfies:

$$\bar{\rho}_\psi \cdot \rho_\psi = \sum_{i=1}^{m} |i\rangle\langle i| = I_m$$

i.e. they multiply to the identity on the support of ρ_ψ.
Now, let's run through the modified version of the (converse) proof. As is in the text $\lambda_\psi \prec \lambda_\varphi$ so $\rho_\psi \prec \rho_\varphi$, and so there exist probabilities p_j and unitary operators U_j such that $\rho_\psi = \sum_j p_j U_j \rho_\varphi U_j^\dagger$ (for $j = 1, \ldots, J$). Then, define:

$$M_j = \sqrt{p_j \rho_\varphi} U_j^\dagger \bar{\rho}_\psi^{1/2}$$

for each j and also add the additional operator:

$$M_{J+1} = \Pi_\perp = \sum_{i=m+1}^{d} |i\rangle\langle i|$$

i.e. the projector onto the subspace for which ρ_ψ has no support. Then, we can see that:

$$\begin{aligned}\sum_{j=1}^{J+1} M_j^\dagger M_j &= \bar{\rho}_\psi^{1/2}\left(\sum_{j=1}^{J} p_j U_j \rho_\varphi U_j^\dagger\right)\bar{\rho}_\psi^{1/2} + \Pi_\perp^\dagger \Pi_\perp \\ &= \bar{\rho}_\psi^{1/2}\rho_\psi\bar{\rho}_\psi^{1/2} + \Pi_\perp \\ &= \sum_{i=1}^{m} |i\rangle\langle i| + \sum_{i=m+1}^{d} |i\rangle\langle i| \\ &= I\end{aligned}$$

so the M_j satisfy the completeness relation and thus define a measurement.
Suppose Alice performs a measurement described by these operators, obtaining the outcome j and corresponding $|\psi_j\rangle \propto M_j |\psi\rangle$ (note that outcome $j = J+1$ cannot be measured, due to ρ_ψ and $\Pi_\perp$ having orthogonal support). Then, if ρ_j is Alice's reduced density matrix corresponding to $|\psi_j\rangle$, we have:

$$\begin{aligned}\rho_j \propto M_j \rho_\psi M_j^\dagger &= p_j \sqrt{\rho_\varphi} U_j^\dagger \bar{\rho}_\psi^{1/2} \rho_\psi \bar{\rho}_\psi^{1/2} U_j \sqrt{\rho_\varphi} \\ &= p_j \sqrt{\rho_\varphi} U_j^\dagger I_m U_j \sqrt{\rho_\varphi}\end{aligned}$$

Since the support of $U_j^\dagger$ is that of ρ_ψ and the range of U_j is that of ρ_ψ (as $\rho_\psi = \sum_j p_j U_j \rho_\psi U_j^\dagger$), it follows that $U_j^\dagger I_m U_j = U_j^\dagger U_j = I$ and so:

$$\rho_j \propto p_j \rho_\varphi$$

so $\rho_j = \rho_\varphi$, and by Exercise 2.81, Bob may convert $|\psi_j\rangle$ into $|\varphi\rangle$ via a suitable unitary V_j. □

Exercise 12.21: Entanglement catalysis

Suppose Alice and Bob share a pair of four level systems in the state $|\psi\rangle = \sqrt{0.4}\,|00\rangle + \sqrt{0.4}\,|11\rangle + \sqrt{0.1}\,|22\rangle + \sqrt{0.1}\,|33\rangle$. Show that it is not possible for them to convert this state by LOCC to the state $|\varphi\rangle = \sqrt{0.5}\,|00\rangle + \sqrt{0.25}\,|11\rangle + \sqrt{0.25}\,|22\rangle$. Imagine, however, that a friendly bank is willing to offer them the loan of a *catalyst*, an entangled pair of qubits in the state $|c\rangle = \sqrt{0.6}\,|00\rangle + \sqrt{0.4}\,|11\rangle$. Show that it is possible for Alice and Bob to convert the state $|\psi\rangle|c\rangle$ to $|\varphi\rangle|c\rangle$ by local operations and classical communication, returning the catalyst $|c\rangle$ to the bank after the transformation is complete.

Solution

Concepts Involved: Majorization, LOCC

Calculating the reduced density operators for Alice's qubit for $|\psi\rangle, |\varphi\rangle$:

$$\rho_\psi = \mathrm{Tr}_B(|\psi\rangle\langle\psi|) = 0.4\,|0\rangle\langle 0| + 0.4\,|1\rangle\langle 1| + 0.1\,|2\rangle\langle 2| + 0.1\,|3\rangle\langle 3|$$

$$\rho_\varphi = \mathrm{Tr}_B(|\varphi\rangle\langle\varphi|) = 0.5\,|0\rangle\langle 0| + 0.25\,|1\rangle\langle 1| + 0.25\,|2\rangle\langle 2|$$

Since ρ_ψ, ρ_φ are diagonal, we can read off the eigenvalues (in descending order) $\lambda_\psi^\downarrow = (0.4, 0.4, 0.1, 0.1)$ and $\lambda_\varphi^\downarrow = (0.5, 0.25, 0.25, 0)$. Since $\sum_{i=1}^2 \lambda_\psi^{\downarrow i} = 0.4 + 0.4 = 0.8 \not\leq 0.75 = \sum_{i=1}^2 \lambda_\varphi^{\downarrow i}$, $\lambda_\psi \not\prec \lambda_\varphi$ and so by Theorem 12.15, $|\psi\rangle$ cannot be transformed into $|\varphi\rangle$ via LOCC.
Now, consider $|\psi\rangle|c\rangle$ and $|\varphi\rangle|c\rangle$:

$$\begin{aligned}|\psi\rangle|c\rangle = &\sqrt{0.24}|00\rangle|00\rangle + \sqrt{0.24}|11\rangle|00\rangle + \sqrt{0.06}|22\rangle|00\rangle + \sqrt{0.06}|33\rangle|00\rangle \\ &+ \sqrt{0.16}|00\rangle|11\rangle + \sqrt{0.16}|11\rangle|11\rangle + \sqrt{0.04}|22\rangle|11\rangle + \sqrt{0.04}|33\rangle|11\rangle\end{aligned}$$

$$\begin{aligned}|\varphi\rangle|c\rangle = &\sqrt{0.3}|00\rangle|00\rangle + \sqrt{0.15}|11\rangle|00\rangle + \sqrt{0.15}|22\rangle|00\rangle \\ &+ \sqrt{0.2}|00\rangle|11\rangle + \sqrt{0.1}|11\rangle|11\rangle + \sqrt{0.1}|22\rangle|11\rangle\end{aligned}$$

Suppose Alice/Bob take the first/second qubit of $|c\rangle$ respectively, and consider $\rho_{\psi c}$ and $\rho_{\varphi c}$ obtained by tracing out Bob's initial qubit and the second qubit of $|c\rangle$:

$$\begin{aligned}\rho_{\psi c} &= \mathrm{Tr}_{BC_2}(|\psi\rangle|c\rangle\langle c|\langle\psi|) \\ &= 0.24|00\rangle\langle 00| + 0.24|10\rangle\langle 10| + 0.06|20\rangle\langle 20| + 0.06|30\rangle\langle 30| \\ &+ 0.16|01\rangle\langle 01| + 0.16|11\rangle\langle 11| + 0.04|21\rangle\langle 21| + 0.04|31\rangle\langle 31|\end{aligned}$$

$$\begin{aligned}\rho_{\varphi c} &= \mathrm{Tr}_{BC_2}(|\varphi\rangle|c\rangle\langle\varphi|\langle c|) \\ &= 0.3|00\rangle\langle 00| + 0.15|10\rangle\langle 10| + 0.15|20\rangle\langle 20| \\ &+ 0.2|01\rangle\langle 01| + 0.1|11\rangle\langle 11| + 0.1|21\rangle\langle 21|\end{aligned}$$

Both of these are diagonal, and so we can read off the eigenvalues (in descending order):

$$\lambda_{\psi c}^\downarrow = (0.24, 0.24, 0.16, 0.16, 0.06, 0.06, 0.04, 0.04)$$

$$\lambda_{\varphi c}^\downarrow = (0.3, 0.2, 0.15, 0.15, 0.1, 0.1, 0, 0)$$

and so we can see that $\lambda_{\psi c} \prec \lambda_{\varphi c}$ as:

$$\begin{aligned}
\lambda_{\psi c}^{\downarrow 1} &= 0.24 \leq 0.3 = \lambda_{\varphi c}^{\downarrow 1} \\
\sum_{i=1}^{2} \lambda_{\psi c}^{\downarrow i} &= 0.48 \leq 0.5 = \sum_{i=1}^{2} \lambda_{\varphi c}^{\downarrow i} \\
\sum_{i=1}^{3} \lambda_{\psi c}^{\downarrow i} &= 0.64 \leq 0.65 = \sum_{i=1}^{3} \lambda_{\varphi c}^{\downarrow i} \\
\sum_{i=1}^{3} \lambda_{\psi c}^{\downarrow i} &= 0.8 \leq 0.8 = \sum_{i=1}^{3} \lambda_{\varphi c}^{\downarrow i} \\
\sum_{i=1}^{3} \lambda_{\psi c}^{\downarrow i} &= 0.86 \leq 0.9 = \sum_{i=1}^{3} \lambda_{\varphi c}^{\downarrow i} \\
\sum_{i=1}^{3} \lambda_{\psi c}^{\downarrow i} &= 0.92 \leq 1 = \sum_{i=1}^{3} \lambda_{\varphi c}^{\downarrow i} \\
\sum_{i=1}^{3} \lambda_{\psi c}^{\downarrow i} &= 0.96 \leq 1 = \sum_{i=1}^{3} \lambda_{\varphi c}^{\downarrow i} \\
\sum_{i=1}^{8} \lambda_{\psi c}^{\downarrow i} &= 1 = 1 = \sum_{i=1}^{8} \lambda_{\varphi c}^{\downarrow i}.
\end{aligned}$$

So, by Theorem 12.15 $|\psi\rangle|c\rangle$ can be transformed to $|\varphi\rangle|c\rangle$ by LOCC. Since $|c\rangle$ is left invariant (and in particular, still unentangled with the rest of the state) after this conversion, Alice/Bob can return it to the bank after. □

Exercise 12.22: Entanglement conversion without communication

($\star$) Suppose Alice and Bob are trying to convert a pure state $|\psi\rangle$ into a pure state $|\varphi\rangle$ using local operations only - no classical communciation. Show that this is possible if and only if $\lambda_\psi \cong \lambda_\varphi \otimes x$ where x is some real vector with non-negative entries summing to 1, and '$\cong$' means that the vectors on the left and right have identical non-zero entries.
We believe that the correct condition should be $\lambda_\varphi \cong \lambda_\psi \otimes x$.

Solution

Concepts Involved: Schmidt Decomposition

First we show a Lemma that $\lambda_\psi \cong \lambda_\varphi$ (where λ_ψ is the vector of eigenvalues of $\rho_\psi^A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|)$) implies $|\psi\rangle, |\varphi\rangle$ are related by local unitaries. As a note, the converse is also true by the fact that (local) unitaries preserve the spectrum of the reduced density matrices. Writing the Schmidt decomposition of the two states, we have:

$$|\psi\rangle = \sum_i \sqrt{\lambda_i}\,|a_i\rangle\,|b_i\rangle\,, \quad |\varphi\rangle = \sum_i \sqrt{\lambda_i}\,|a_i'\rangle\,|b_i'\rangle$$

wherein we can define local unitaries $U_A = \sum_i |a_i'\rangle\langle a_i|, U_B = \sum_i |b_i'\rangle\langle b_i|$, for which $U_A \otimes U_B |\psi\rangle = |\varphi\rangle$. Armed with this Lemma, let us start with the converse proof:

$\boxed{\Longleftarrow}$: First, Alice can begin by appending a pure ancilla state $\sigma_x^A = |\sigma_x\rangle\langle\sigma_x|^A$ with spectrum x to her subsystem $\rho^A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|)$. Her resulting state $\rho^A \otimes \sigma_x^A$ has eigenvalues $\lambda_\psi \otimes x$ (and the overall state $|\psi\rangle \otimes |\sigma_x\rangle^A$ remains pure). Then, applying the above Lemma, there exist local unitaries U_A, U_B that Alice/Bob may apply to convert the overall state into $|\varphi\rangle$. Both steps only required local operations, without classical communication.

$\boxed{\Longrightarrow}$: Suppose $|\psi\rangle$ is converted to $|\varphi\rangle$ using only local operations. The most general purity-preserving local operation that can be carried out by Alice and Bob is appending a pure state ancilla σ^A/σ^B and applying general local unitary U_A/U_B, so:

$$|\varphi\rangle\langle\varphi|^{AB} = (U_A \otimes U_B)(\,|\psi\rangle\langle\psi|^{AB} \otimes \sigma^A \otimes \sigma^B)(U_A^\dagger \otimes U_B^\dagger).$$

Note that the appended ancillas σ^A, σ^B cannot be mixed, as the unitaries U_A/U_B preserve the purity of the state and hence the output of the state is mixed. Similarly, we are not permitted to entangle the ancillas with the system and then trace out the ancillas, as this would again result in a mixed state output (as the reduced density operator of an entangled pure state is mixed). The case where unentangled ancillas are traced out is equivalent to the case of not appending them in the first place, in which case we simply obtain $\rho_\varphi^A = U_A \rho_\psi^A U_A^\dagger$ so $\lambda_\varphi = \lambda_\psi$, corresponding to the trivial case with $x = (1)$. Finally, although the case of local measurements is already handled by the above points about entangling + tracing out ancilla (by the arguments of Section 8.2.2 in the text), it is worth explicitly pointing out that measurements by Alice/Bob are not permitted, as without classical communication the post-measurement state becomes a probabilistic mixture over the possible local measurement outcomes, which again results in a mixed state as output.

Returning back to the most general case, let us look at Alice's local subsystem:

$$\begin{aligned}
\rho_\varphi^A &= \mathrm{Tr}_B(\,|\varphi\rangle\langle\varphi|^{AB}) \\
&= \mathrm{Tr}_B(U_A \otimes U_B(\,|\psi\rangle\langle\psi|^{AB} \otimes \sigma^A \otimes \sigma^B)U_A^\dagger \otimes U_B^\dagger) \\
&= U_A\,\mathrm{Tr}_B(I \otimes U_B(\,|\psi\rangle\langle\psi|^{AB} \otimes \sigma^A \otimes \sigma^B)I \otimes U_B^\dagger)U_A^\dagger \\
&= U_A\,\mathrm{Tr}_B(\,|\psi\rangle\langle\psi|^{AB} \otimes \sigma^A \otimes \sigma^B)U_A^\dagger \\
&= U_A(\rho_\psi^A \otimes \sigma^A)U_A^\dagger
\end{aligned}$$

where we use that U_A commutes with the partial trace operation in the third line, we use cyclicity in the fourth line to cancel the U_Bs, and the trace over B (both Bob's appended ancilla and his part of $|\psi\rangle\langle\psi|^{AB}$). Unitaries preserve eigenvalues, so the eigenvalues of ρ_φ^A are simply that of $\rho_\psi^A \otimes \sigma^A$. Hence, the non-zero eigenvalues of ρ_φ^A are given by $\lambda_\psi \otimes x$ where x are the eigenvalues of σ^A, which are real, non-negative, and sum to one by the properties of the eigenvalues of a density operator. Thus $\lambda_\varphi \cong \lambda_\psi \otimes x$ and the claim is proven. □

Exercise 12.23

Prove that the procedure for entanglement distillation we have described is optimal.

It seems as the authors already present an argument for this in the text, as they mix up distillation and dilution in the preceding paragraph - we just present arguments for why both presented protocols are optimal, here; the arguments are completely symmetric.

Solution

Concepts Involved: Entanglement Distillation, Entanglement Dilution

The text discusses a protocol that uses LOCC to distill a copy of $|\psi\rangle$ into $S(\rho_\psi)$ Bell states, and to dilute $S(\rho_\psi)$ Bell states into a copy of $|\psi\rangle$. Central to our argument will be the fact proven in the next exercise (Ex. 12.24) that the number of Bell states cannot be increased by LOCC.

Suppose a more efficient distillation protocol existed that could distill $|\psi\rangle$ into $S > S(\rho_\psi)$ Bell states. Then, starting with $S(\rho_\psi)$ Bell states, we could dilute them (using the known protocol) to create $|\psi\rangle$, which we then distill to obtain $S > S(\rho_\psi)$ Bell states. We have thus increased the number of Bell states only using LOCC - contradiction.

Suppose on the other hand a more efficient dilution protocol existed that could dilute $|\psi\rangle$ from $S < S(\rho_\psi)$ Bell states. But then starting from S Bell states, we could create $|\psi\rangle$, and then use the known dilution protocol to form $S(\rho_\psi) > S$ Bell states. We have thus increased the number of Bell states only using LOCC, again a contradiction.

We conclude that the provided protocols for distillation/dilution are both optimal. □

Exercise 12.24

($\star$) Recall that the Schmidt number of a bi-partite pure state is the number of non-zero Schmidt components. Prove that the Schmidt number of a pure quantum state cannot be increased by local operations and classical communication. Use this result to argue that the number of Bell states shared between Alice and Bob cannot be increased by local operations and classical communication.

Solution

Concepts Involved: Schmidt Decomposition, Schmidt Numbers, Majorization, LOCC

Recall from Problem 2.2 that the Schmidt number of a bipartite pure state $|\psi\rangle_{AB}$ is equal to the rank of $\rho_\psi = \mathrm{Tr}_B(|\psi\rangle\langle\psi|)$, i.e. the number of nonzero eigenvalues.

Consider bipartite pure states $|\psi\rangle, |\varphi\rangle$ such that $\mathrm{Sch}(\psi) = m < \mathrm{Sch}(\varphi) = n$; then, λ_ψ has m nonzero entries and λ_φ has n nonzero entries. Then, consider the normalization conditions (summing over only the nonzero eigenvalues):

$$\sum_{i=1}^{m} \lambda_i^\psi = 1$$

$$\sum_{i=1}^{m} \lambda_i^\varphi < \sum_{i=1}^{n} \lambda_i^\varphi = 1$$

so in particular:

$$\sum_{i=1}^{m} \lambda_i^\psi \not\leq \sum_{i=1}^{m} \lambda_i^\varphi$$

and so $\lambda_\psi \not\prec \lambda_\varphi$. Thus by Theorem 12.15, $|\psi\rangle$ cannot be transformed to $|\varphi\rangle$ by LOCC, and we conclude

that Schmidt number cannot be increased by LOCC.
A bell state $|B_{00}\rangle = \frac{1}{\sqrt{2}} |0\rangle_A |0\rangle_B + \frac{1}{\sqrt{2}} |1\rangle_A |1\rangle_B$ has Schmidt number 2 (and the same for its variants), and in particular generating a Bell state from a product state $|\psi\rangle_A \otimes |\psi\rangle_B$ (with Schmidt number 1) increases the Schmidt number, so LOCC cannot increase the number of shared Bell states. □

Exercise 12.25

Consider a system with n users, any pair of which would like to be able to communicate privately. Using public key cryptography how many keys are required? Using private key cryptography how many keys are required?

Solution

Concepts Involved: Cryptography

In the public case, each of the n users has a public key which they publish (and that others use to encrypt their messages to send to them) and a secret key which only they know that they use to easily perform the decryption. Thus, $2n = O(n)$ keys are required.
In the private case, it is required that every pair of users shares a key pair (encoding/decoding) only known to them; for n total users there are $\frac{n(n-1)}{2}$ pairs, and hence required number of keys. is $2 \cdot \frac{n(n-1)}{2} = n(n-1) = O(n^2)$; thus private key cryptography is "more expensive" in the number of keys. □

Exercise 12.26

Let a'_k be Bob's measurement result of qubit $|\psi_{a_k b_k}\rangle$, assuming a noiseless channel with no eavesdropping. Show that when $b'_k \neq b_k$, a'_k is random and completely uncorrelated with a_k. But when $b'_k = b_k$, $a'_k = a_k$.

Solution

Concepts Involved: Quantum Measurement

Alice sends bob states $|\psi_{a_k b_k}\rangle$ of the form:

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

Bob measures in basis b'_k (with $b'_k = 0$ being the Z basis $\{|0\rangle, |1\rangle\}$, and $b'_k = 1$ the X basis $\{|+\rangle, |-\rangle\}$).
From the above, we can immediately see that, for $b'_k = 0, b_k = 1$:

$$\left|\langle 0|\psi_{01}\rangle\right|^2 = \left|\langle 0|\psi_{11}\rangle\right|^2 = \left|\langle 1|\psi_{01}\rangle\right|^2 = \left|\langle 1|\psi_{11}\rangle\right|^2 = \frac{1}{2}$$

and for $b'_k = 1, b_k = 0$:

$$\left|\langle +|\psi_{00}\rangle\right|^2 = \left|\langle +|\psi_{10}\rangle\right|^2 = \left|\langle -|\psi_{00}\rangle\right|^2 = \left|\langle -|\psi_{10}\rangle\right|^2 = \frac{1}{2}$$

so if $b'_k \neq b_k$, Bob's outcome a'_k is random and uncorrelated with a_k. Conversely, for $b'_k = 0, b_k = 0$:

$$\left|\langle 0|\psi_{00}\rangle\right|^2 = \left|\langle 1|\psi_{10}\rangle\right|^2 = 1$$
$$\left|\langle 1|\psi_{00}\rangle\right|^2 = \left|\langle 0|\psi_{10}\rangle\right|^2 = 0$$

and for $b'_k = 1, b_k = 1$:

$$\left|\langle +|\psi_{01}\rangle\right| = \left|\langle -|\psi_{11}\rangle\right| = 1$$
$$\left|\langle -|\psi_{01}\rangle\right| = \left|\langle +|\psi_{11}\rangle\right| = 0$$

so we can see that when $b'_k = b_k$, $a'_k = a_k$. □

Exercise 12.27: Randomized sampling tests

($\star\star\star$) The random test of n of $2n$ check bits allows Alice and Bob to place an upper bound on the number of errors in their untested bits, with high probability. Specifically, for any $\delta > 0$, the probability of obtainining less than δn errors on the check bits, and more than $(\delta + \epsilon)n$ errors on the remanining n bits is asymptotically less than $\exp\left[-O(\epsilon^2 n)\right]$, for large n. We prove this claim here.

(1) Without loss of generality, you may assume that there are μn errors in the $2n$ bits, where $0 \leq \mu \leq 2$. Now, if there are δn errors on the check bits, and $(\delta + \epsilon)n$ errors on the rest, then $\delta = (\mu - \epsilon)/2$. The two conditional statements in the claim thus imply the following:

$$\begin{aligned} &< \delta n \text{ errors on check bits} &&\implies < \delta n \text{ errors on check bits} \\ &> (\delta + \epsilon)n \text{ errors on rest} &&\implies > (\mu - \delta)n \text{ errors on rest,} \end{aligned}$$

and in fact, the top claim on the right implies the bottom one on the right. Using this, show that the probability p which we would like to bound satisfies

$$p < \binom{2n}{n}^{-1} \binom{\mu n}{\delta n} \binom{(2-\mu)n}{(1-\delta)n} \delta n.$$

(2) Show that for large n, you can bound

$$\frac{1}{an+1} 2^{anH(b/a)} \leq \binom{an}{bn} \leq 2^{anH(b/a)}$$

where $H(\cdot)$ is the binary entropy function, Equation (11.8). Apply this to the above bound for p.

(3) Apply the bound $H(x) < 1 - 2(x - 1/2)^2$ to obtain the final result, $p < \exp\left[-O(\epsilon^2 n)\right]$. You may replace μ by a constant which expresses the worst possible case.

(4) Compare this result with the Chernoff bound, Box 3.4. Can you come up with a different way to derive an upper bound on p? For this step, we seemed to require a more powerful/general version Chernoff bound than was provided in the text.

Solution

Concepts Involved: Shannon Entropy, Binary Entropy, Probability, Chernoff Bound

(1) Since the top claim on the right implies the bottom one on the right, we find that our probability is bounded by:

$$p = P(< \delta n \text{ errors on check bits} \cup > (\delta+\epsilon)n \text{ errors on rest}) \leq P(< \delta n \text{ errors on check bits})$$

Thus, let us calculate the probability of obtaining $< \delta n$ errors on the check bits. There are $\binom{2n}{n}$ ways to choose n of the total $2n$ bits to be the check bits. Supposing there are μn errors on all of the bits, and k of these errors occur on the check bits, there are $\binom{\mu n}{k}$ choices of error bits to land in the check set, and $\binom{(2-\mu)n}{n-k}$ choices of correct bits to land in the remainder of the check set. Summing over $k = 0, 1, \ldots, \delta n - 1$ (i.e. all $k < \delta n$), we obtain:

$$p \leq \frac{\sum_{k=0}^{\delta n-1} \binom{\mu n}{k}\binom{(2-\mu)n}{n-k}}{\binom{2n}{n}}$$

We now wish to bound this sum; looking at ratios of successive terms, we find:

$$\frac{\binom{\mu n}{k+1}\binom{(2-\mu)n}{n-k-1}}{\binom{\mu n}{k}\binom{(2-\mu)n}{n-k}} = \frac{(\mu n - k)(n-k)}{(k+1)((1-\mu)n + k + 1)}$$

wherein the terms are increasing when the above ratio is > 1, and we have the crossover/maximum point when:

$$(\mu n - k)(n-k) = (k+1)((1-\mu)n + k + 1) \approx k((1-\mu)n + k)$$

where we neglect the $+1$s in the large n limit. Studying the above, equality is achieved at $k = \frac{\mu n}{2}$, and so the terms of the sum increase up to this k. In particular, note that for all $k = 0, 1, \ldots, \delta n - 1$, we have $k < \delta n = \frac{\mu - \epsilon}{2} n < \frac{\mu n}{2}$ for $\epsilon > 0$, and hence we can bound the sum by replacing every term with one where $k = \delta n$:

$$p < \frac{\delta n \binom{\mu n}{\delta n}\binom{(2-\mu)n}{n-\delta n}}{\binom{2n}{n}}$$

and so:

$$p < \binom{2n}{n}^{-1}\binom{\mu n}{\delta n}\binom{(2-\mu)n}{(1-\delta)n}\delta n.$$

(2) Upper bound: We adapt the helpful presentation of Qiaochu Yan from MathOverflow. From the binomial theorem, we have for any $x, y > 0$:

$$(x+y)^n = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k} \geq \binom{n}{k}x^k y^{n-k}$$

where we use that the sum is of positive terms and hence upper bounds any given term. Taking

$x = p, y = 1 - p$ we obtain:

$$\binom{n}{k} p^k (1-p)^{n-k} \leq 1 \implies \binom{n}{k} \leq \frac{1}{p^k (1-p)^{n-k}}$$

and taking logarithms we find:

$$\log \binom{n}{k} \leq -k \log p - (n-k) \log(1-p)$$

in particular taking $p = \frac{k}{n}$ (which minimizes the RHS) the above becomes:

$$\begin{aligned} \log \binom{n}{k} &\leq -k \log \frac{k}{n} - (n-k) \log\left(1 - \frac{k}{n}\right) \\ &= n \left(-\frac{k}{n} \log \frac{k}{n} - (1 - \frac{k}{n}) \log\left(1 - \frac{k}{n}\right) \right) \\ &= nH(\frac{k}{n}) \end{aligned}$$

and so:

$$\binom{n}{k} \leq 2^{nH(k/n)}$$

Taking $n \to an, k \to bn$, we obtain the upper bound:

$$\boxed{\binom{an}{bn} \leq 2^{anH(b/a)}}.$$

Lower bound: We adapt the elegant proof of Marcus M from Math StackExchange. Consider the integral:

$$I(n,k) = \int_0^1 p^k (1-p)^{n-k} dp = \frac{1}{(n+1)\binom{n}{k}}$$

the formula for $I(n,k)$ follows by induction on k; for $k = 0$ we have:

$$I(n,0) = \int_0^1 p^0 (1-p)^{n-0} dp = \int_0^1 (1-p)^n = -\frac{1}{n+1}(1-p)^{n+1}\Big|_0^1 = \frac{1}{n+1} = \frac{1}{n+1\binom{n}{0}}$$

Now supposing the formula holds for some $k-1 \geq 0$, we have:

$$\begin{aligned} I(n,k) &= \int_0^1 p^k(1-p)^{n-k}dp \\ &= -\frac{p^k(1-p)^{n-k+1}}{n-k+1}\bigg|_0^1 - \int_0^1 -\frac{kp^{k-1}(1-p)^{n-k+1}}{n-k+1}dp \\ &= 0 + \frac{k}{n-k+1}I(n,k-1) \\ &= \frac{k}{n+1-k}\frac{1}{(n+1)\binom{n}{k-1}} \\ &= \frac{1}{(n+1)\binom{n}{k}} \end{aligned}$$

where in the second line we use integration by parts, in the fourth line we apply the inductive hypothesis, and in the last line we use the binomial reccurence relation of $\binom{n}{k} = \frac{n-k+1}{k}\binom{n}{k-1}$. This completes the induction.

Now, noting that $f(p) = p^k(1-p)^{n-k}$ is maximized at $p = \frac{k}{n}$, we can bound the integral:

$$I(n,k) \leq (1-0)f(\frac{k}{n}) = \left(\frac{k}{n}\right)^k \left(1-\frac{k}{n}\right)^{n-k}$$

the RHS we can already recognize from the upper bound calculation to be $2^{-nH(k/n)}$, and so:

$$\frac{1}{(n+1)\binom{n}{k}} \leq 2^{-nH(k/n)} \implies \frac{1}{n+1}2^{nH(k/n)} \leq \binom{n}{k}$$

Then taking $n \to an, k \to bn$, we obtain:

$$\boxed{\frac{1}{an+1}2^{anH(b/a)} \leq \binom{an}{bn}}.$$

Applying this to the bound from (1), we find (using the lower bound to the $\binom{2n}{n}^{-1}$ factor: and the upper bound for the $\binom{\mu n}{\delta n}, \binom{(2-\mu)n}{(1-\delta)n}$ factors:

$$\begin{aligned} p &< \left(\frac{1}{2n+1}2^{2nH(1/2)}\right)^{-1} 2^{\mu nH(\delta/\mu)}2^{(2-\mu)nH((1-\delta)/(2-\mu))}\delta n \\ &= \frac{\delta n(2n+1)}{2^{2n}}2^{\mu nH(\delta/\mu)}2^{(2-\mu)nH((1-\delta)/(2-\mu))} \end{aligned}$$

(3) Let us study the exponential terms; using the provided entropy bound, we have:

$$\mu nH(\frac{\delta}{\mu}) \leq \mu n\left[1-2\left(\frac{\delta}{\mu}-\frac{1}{2}\right)^2\right] = \mu n\left[1-2\left(\frac{2\delta-\mu}{2\mu}\right)^2\right] = \mu n\left[1-2\left(-\frac{\epsilon}{2\mu}\right)^2\right]$$

$$(2-\mu)nH(\frac{1-\delta}{2-\mu}) \leq (2-\mu)n\left[1-2\left(\frac{1-\delta}{2-\mu}-\frac{1}{2}\right)^2\right] = (2-\mu)n\left[1-2\left(\frac{\mu-2\delta}{2(2-\mu)}\right)^2\right]$$
$$= (2-\mu)n\left[1-2\left(\frac{\epsilon}{2(2-\mu)}\right)^2\right]$$

where in the last equality we have used that $\delta = \frac{\mu-\epsilon}{2}$. Thus, combining all the exponential terms (including the 2^{-2n}), we have:

$$\leq \mu n\left[1-2\left(-\frac{\epsilon}{2\mu}\right)^2\right] + (2-\mu)n\left[1-2\left(\frac{\epsilon}{2(2-\mu)}\right)^2\right] - 2n$$
$$= -2\mu n\left(-\frac{\epsilon}{2\mu}\right)^2 - 2(2-\mu)n\left(\frac{\epsilon}{2(2-\mu)}\right)^2$$
$$= -\frac{\epsilon^2 n}{2\mu} - \frac{\epsilon^2 n}{2(2-\mu)}$$
$$= -\frac{\epsilon^2 n}{\mu(2-\mu)}$$

The worst case is when $\frac{1}{\mu(2-\mu)}$ is minimized, i.e. $\mu = 1$, wherein the exponential is bounded by $-\epsilon^2 n$, so:

$$p < \delta n(2n+1)2^{-\epsilon^2 n} = O(\exp\Big(-\epsilon^2 n\Big))$$

where we can neglect the polynomial terms as the bound is dominated by the decaying exponential.

(4) Comparing, we see that the Chernoff bound is also an upper bound of the form $p < O(-\epsilon^2 n)$, though applied to the sum of n independent random variables, which is different from the approach we took here. This suggests an alternative route for deriving an upper bound on p, as follows.

Let us imagine that each bit of the $2n$ total is independently included in the n check bits, with probability $1/2$. The total number of errors in the check bits is then:

$$\sum_{i=1}^{\mu n} X_i$$

with X_i independent and identically distributed random Bernoulli variables taking 1 (error) with probability 1/2 and 0 (no error) with probability 1/2.

We consider a more powerful/general version of the Chernoff bound, as e.g. discussed in these lecture notes - the bound as provided in Box 3.4 is not useful for unbiased Bernoulli variables, or for bounding the sum away from a value that is not exactly the mean, both of which we require here. In particular, let us use the lower tail bound from the lecture notes:

$$P(\sum_{i=1} X_i \leq (1-\alpha)m) \leq e^{-m\alpha^2/2}$$

with $m = \mathbf{E}(\sum_i X_i)$ and $\alpha > 0$ quantifying the deviation away from the mean.

For our case, we wish to bound the probability of $\sum_{i=1}^{\mu n} X_i < \delta n = \frac{\mu-\epsilon}{2}n$. The mean is given by

$m = \mathbf{E}(\sum_{i=1}^{\mu n} X_i) = \sum_{i=1}^{\mu n} \mathbf{E}(X_i) = \sum_{i=1}^{\mu n} \frac{1}{2} = \frac{\mu n}{2}$, so taking $\alpha = \frac{\epsilon}{\mu}$ we have:

$$P(\sum_i X_i \leq (1 - \frac{\epsilon}{\mu})\frac{\mu n}{2}) = P(\sum_i X_i \leq \frac{\mu - \epsilon}{2} n) \leq e^{-\frac{\mu n}{2} \cdot (\frac{\epsilon}{\mu})^2/2} = e^{-\frac{\epsilon^2 n}{4\mu}} = O(\exp\Big(-\epsilon^2 n\Big))$$

so we again find the desired upper bound on p - this approach is much quicker!

□

Exercise 12.28

Show that when $b = 1$, then a and a' are perfectly correlated with each other.
The problem statement should be reworded to "perfectly *anti*-correlated".

Solution

Concepts Involved: Quantum Measurement

If $b = 1$, then Bob has post-measurement state $|1\rangle$ if $a' = 0$ or $|-\rangle$ if $a' = 1$.
Since $\left|\langle a' = 0, b = 1 | a = 0\rangle\right|^2 = \left|\langle 1|0\rangle\right|^2 = 0$, if $a' = 0, b = 1$ it cannot have been the case that $a = 0$ and so we must have $a = 1$ (and Alice sent Bob the $|+\rangle$ state).
Further, since $\left|\langle a' = 1, b = 1 | a = 1\rangle\right|^2 = \left|\langle -|+\rangle\right|^2 = 0$, if $a' = 1, b = 1$ then it cannot have been the case that $a = 1$ and so we must have $a = 0$ (and Alice sent Bob the $|0\rangle$ state).
Hence, if $b = 1$, then either $a' = 0, a = 1$ or $a' = 1, a = 0$ and so a, a' are perfectly anti-correlated. □

Exercise 12.29

Give a protocol using six states, the eigenstates of X, Y and Z, and argue why it is also secure. Discuss the sensitivity of this protocol to noise and eavesdropping, in comparison with that of BB84 and B92.

Solution

Concepts Involved: BB84, B92

The protocol essentially is identical to that of the BB84 protocol, just using six states instead of four.

(1) Alice chooses $(6 + \delta)n$ random data bits. Note the change from $(4 + \delta)n$ to account for the fact that we now use 6 states, so a larger number of data bits are required to ensure that $2n$ bits are left in step (6).

(2) Alice chooses a random $(6 + \delta)n$-trit (a trit being 0, 1, or 2) string t. She encodes each data bit as $\{|0\rangle, |1\rangle\}$ if the corresponding bit of t is 0, $\{|+\rangle, |-\rangle\}$ if t is 1, and $\{|+i\rangle, |-i\rangle\}$ if t is 2.

(3) Alice sends the resulting state to Bob.

(4) Bob receives the $(6 + \delta)n$ qubits, announces this fact, and measures each basis in the X, Z, Y basis at random.

(5) Alice announces t.

(6) Alice and Bob discard any bits where Bob measured a different basis than Alice prepared. With high probability, there are at least $2n$ bits left (if not, abort the protocol). They keep $2n$ bits.

(7) Alice selects a subset of n bits that will serve as a check on Eve's interference, and tells Bob which bits she selected.

(8) Alice and Bob compare the values of the n check bits. If more than an acceptable number disagree, they abort the protocol. Alice and Bob perform information reconciliation and privacy amplification on the remaining n bits to obtain m shared key bits.

In a straightforward generalization of Ex. 12.26 which includes the Y-eigenstates, it is easy to see that when Bob's measurement basis t'_k disagrees with Alice's encoding basis t_k that his measurement outcome a'_k is uncorrelated with Alice's bit a_k, but when $t'_k = t_k$, then $a'_k = a_k$. The protocol has the exact same sensitivity to noise/eavesdropping as BB84/B92, as the procedure of running the check for intereference/errors is identical. □

Exercise 12.30

Simplify (12.199) to obtain the expression for $S(\rho)$ given in the statement of the Lemma.

Solution

Concepts Involved: Von Neumann Entropy

The expression of (12.199) is:

$$S(\rho_{\text{max}}) = -(1-2^{-s})\log\left(1-2^{-s}\right) - 2^{-s}\log\frac{2^{-s}}{2^{2n}-1}$$

which, splitting up the last term and then going through a sequence of rewrites:

$$\begin{aligned}
S(\rho_{\text{max}}) &= -(1-2^{-s})\log\left(1-2^{-s}\right) - 2^{-s}(\log\left(2^{-s}\right) - \log\left(2^{2n}-1\right)) \\
&= -(1-2^{-s})\log\left(1-2^{-s}\right) - 2^{-s}(-s - \log\left(2^{2n}-1\right)) \\
&= (2^{-s}-1)\log\left(1-2^{-s}\right) + 2^{-s}(s + \log\left(2^{2^n}-1\right)) \\
&\leq (2^{-s}-1)\frac{(1-2^{-s})-1}{\ln 2} + 2^{-s}(s + \log\left(2^{2n}\right)) \\
&= (2^{-s}-1)\frac{-2^{-s}}{\ln 2} + 2^{-s}(s+2n) \\
&= (2n+s+\frac{1}{\ln 2})2^{-s} - \frac{2^{-2s}}{\ln 2} \\
&= (2n+s+\frac{1}{\ln 2})2^{-s} + O(2^{-2s})
\end{aligned}$$

Where in the inequality we use that $\log x \leq \frac{x-1}{\ln 2}$ and that $\log(x-1) \leq \log(x)$. □

Exercise 12.31

It may be unclear why $S(\rho)$ bounds Eve's mutual information with Alice and Bob's measurement results. Show that this follows from assuming the worst about Eve, giving her *all* the control over the channel.

Solution

Concepts Involved: Mutual Information, Holevo Bound

If Eve has *all* the control over the channel, she can intercept the $|\beta_{00}\rangle^{\otimes n}$ state and subsequently choose to send whatever state $\rho = \sum_k p_k \rho_k$ she likes to Alice/Bob. In this case, Holevo tells us that the mutual information between what Eve sends and what Alice/Bob measures is bounded by:

$$H_{\text{eve:alice+bob}} \leq \chi = S(\rho) - \sum_k p_k S(\rho_k) \leq S(\rho)$$

□

Exercise 12.32

Note that the local measurements that Alice and Bob perform, such as $I \otimes X$ and $X \otimes I$, do *not* commute with the Bell basis. Show that despite this, the statistics which Alice and Bob compile from their measurements are the same as those which they would have obtained had they actually measured Π_{bf} and Π_{pf}.

Solution

Concepts Involved: Bell Basis, Quantum Measurement

The observation made in the question is easily seen from the fact $|\beta_{zx}\rangle = Z_1^z X_1^x |\beta_{00}\rangle = X_2^x Z_2^z |\beta_{00}\rangle$, from which we can deduce that local Paulis do not commute with the Bell state projectors; for example:

$$X_1 |\beta_{00}\rangle\langle\beta_{00}| = |\beta_{01}\rangle\langle\beta_{00}| \neq |\beta_{00}\rangle\langle\beta_{01}| = |\beta_{00}\rangle\langle\beta_{00}| X_1.$$

However, Alice and Bob can obtain the same measurement statistics from local Pauli measurements a measuring $\Pi_{\text{bf}}, \Pi_{\text{pf}}$. The single-qubit projectors are:

$$\Pi_{\pm}^{P_1} = \frac{I \otimes I \pm P_1 \otimes I}{2}, \quad \Pi_{\pm}^{P_2} = \frac{I \otimes I \pm I \otimes P_2}{2}$$

for $P_i = X_i, Z_i$. We can then see that:

$$\begin{aligned}
\Pi_+^{X_1}\Pi_+^{X_2} + \Pi_-^{X_1}\Pi_-^{X_2} &= \frac{I \otimes I + X_1 \otimes I + I \otimes X_2 + X_1 \otimes X_2}{4} + \frac{I \otimes I - X_1 \otimes I - I \otimes X_2 + X_1 \otimes X_2}{4} \\
&= \frac{I \otimes I + X_1 \otimes X_2}{2} \\
&= \Pi_{\text{bf}}
\end{aligned}$$

$$\begin{aligned}\Pi_+^{X_1}\Pi_-^{X_2} + \Pi_-^{X_1}\Pi_+^{X_2} &= \frac{I\otimes I + X_1\otimes I - I\otimes X_2 - X_1\otimes X_2}{4} + \frac{I\otimes I - X_1\otimes I + I\otimes X_2 - X_1\otimes X_2}{4} \\ &= \frac{I\otimes I - X_1\otimes X_2}{2} \\ &= 1 - \Pi_{\text{bf}}.\end{aligned}$$

So if Alice/Bob measure $X_1\otimes I, I\otimes X_2$, they get the same measurement statistics as measuring $\Pi_{\text{bf}}, 1-\Pi_{\text{bf}}$, where they measure $X_1\otimes X_2 = +1$ if their measurements have the same parity and $X_1\otimes X_2 = -1$ if their measurements have the opposite parity.
Analogous statements hold for Π_{pf}, replacing every X appearing above with a Z. □

Exercise 12.33

Let $\{M_1, M_2, \ldots, M_n\}$ be a set of measurement observables which produce respective results X_i when an input state ρ is measured. Argue that the random variables X_i obey classical probability arguments if $[M_i, M_j] = 0$, that is, they commute with each other.

Solution

Concepts Involved: Probability, Quantum Measurement

We recall that if $[M_i, M_j] = 0$ for all i, j then $\{M_1, \ldots, M_n\}$ are simultaneously diagonalizable. Thus writing $M_i = \sum_k m_i^k |k\rangle\langle k|$ and $\rho = \sum_k p_k |k\rangle\langle k| + (\text{off-diagonals})$ for the joint eigenbasis $\{|k\rangle\}_k$, we can see that the X_i obey classical probability arguments, with distributions given by $P(X_i = m_i^k) = p_k$. □

Exercise 12.34: Entanglement distillation by error-correction

($\star\star$) In Section 10.5.8, we saw that codewords of an $[n, m]$ qubit stabilizer code can be constructed by measuring its generators $g_1, \ldots, g_{n-m}$ on an *arbitrary* n qubit quantum state, then applying Pauli operations to change the result to be a simultaneous $+1$ eigenstate of the generators. Using that idea, show that if we start out with n EPR pairs in the state $|\beta_{00}\rangle^{\otimes n}$, and perform identical generator measurements on the two n qubit halves of the pairs, followed by Pauli operations to correct for *differences* in the measurement results between the pairs, then we obtain an encoded $|\beta_{00}\rangle^{\otimes m}$ state. Also show that if the stabilizer code corrects up to δn errors, then even if δn errors are suffered by an n qubit half, we still obtain $|\beta_{00}\rangle^{\otimes m}$.
We don't believe it suffices to correct for differences in the measurement results - instead, we should correct for all -1 outcomes of the generator measurements. Correcting for differences in the measurement outcomes would result in a state for which $g_i \otimes g_i = +1$ (for $i = 1, \ldots, n-m$). To us this seems insufficient - in order to have $2m$ logically encoded qubits, it must be the case that we are in the $+1$ eigenspace of all $2(n-m)$ stabilizers of $g_i \otimes I, I \otimes g_i$ for $i = 1, \ldots, n-m$.

Solution

Concepts Involved: Bell Pairs, Quantum Measurements, Stabilizer codes, Standard Form

Suppose that we start out with n EPR pairs in the state $|\beta_{00}\rangle^{\otimes n}$, and perform measurements of

$g_1, \ldots, g_{n-m}$ on the two n qubit halves of the pairs. If we then apply Pauli corrections to correct for these outcomes such that $g_i \otimes I = +1$ and $I \otimes g_i = +1$ for all $i = 1, \ldots, n-m$ (as would be required for any $2m$-qubit state encoded by this stabilizer code), the output state is:

$$\prod_{i=1}^{n-m} (P_i^+ \otimes P_i^+) \left|\beta_{00}\right\rangle^{\otimes n}$$

where:

$$P_i^{\pm} = \frac{I \pm g_i}{2}$$

is the projector onto the ± 1 eigenspace of the stabilizer g_i.

WLOG, suppose that we have brought the stabilizer code into standard form, such that the check matrix for the Z operators is of the form $G_z = [000|A_2^T 0 I]$ and the check matrix for the encoded X operators is of the form $G_x = [0E^T I|C^T 00]$. The upshot of this is that all logical operators $\bar{X}_j, \bar{Z}_j$ (for $j = 1, \ldots, m$) of the code contain no Y-operators. In particular, this means that $\bar{X}_j^T = \bar{X}_j, \bar{Z}_j^T = \bar{Z}_j$ (this is because $I^T = I, X^T = X, Z^T = Z$ - in contrast, $Y^T = -Y$).

Since the output state is an encoded state $2m$-qubit state (it is stabilized by all $2(n-m)$ generators), to show that it is the logical $\left|\beta_{00}\right\rangle^{\otimes m}$ state, it suffices to show that it is the $+1$ eigenstate of all logical operator pairs $\bar{X}_j \otimes \bar{X}_j, \bar{Z}_j \otimes \bar{Z}_j$. We can see that this is indeed the case, as:

$$\begin{aligned}
(\bar{X}_j \otimes \bar{X}_j) \prod_{i=1}^{n-m} (P_i^+ \otimes P_i^+) \left|\beta_{00}\right\rangle^{\otimes n} &= \prod_{i=1}^{n-m} (P_i^+ \otimes P_i^+)(\bar{X}_j \otimes \bar{X}_j) \left|\beta_{00}\right\rangle^{\otimes n} \\
&= \prod_{i=1}^{n-m} (P_i^+ \otimes P_i^+)(\bar{X}_j \bar{X}_j^T \otimes I) \left|\beta_{00}\right\rangle^{\otimes n} \\
&= \prod_{i=1}^{n-m} (P_i^+ \otimes P_i^+)(\bar{X}_j^2 \otimes I) \left|\beta_{00}\right\rangle^{\otimes n} \\
&= \prod_{i=1}^{n-m} (P_i^+ \otimes P_i^+)(I \otimes I) \left|\beta_{00}\right\rangle^{\otimes n} \\
&= (+1) \prod_{i=1}^{n-m} (P_i^+ \otimes P_i^+) \left|\beta_{00}\right\rangle^{\otimes n}
\end{aligned}$$

where the first equality follows from the fact that the logical operators commute with all stabilizer generators, the second equality uses the bounce identity $(A \otimes I) \left|\beta_{00}\right\rangle = (I \otimes A^T) \left|\beta_{00}\right\rangle$ for Bell states (we provide the general proof for maximally entangled states in Ex. 12.36), and in the third equality we use that $\bar{X}_j^T = \bar{X}_j$ from the above standard form discussion. The argument is analogous for the $\bar{Z}_j \otimes \bar{Z}_j$ operators, and thus we conclude that the state is indeed the encoded $\left|\beta_{00}\right\rangle^{\otimes m}$ state.

The bounce identity is central to showing the second claim - WLOG suppose that a weight-δn error operator E affects second qubit halves of the initial state $\left|\beta_{00}\right\rangle^{\otimes n}$, i.e. we start with $(I \otimes E) \left|\beta_{00}\right\rangle^{\otimes n}$ (if the errors occur on the first half of the qubits, we may use the bounce identity so that they instead act on the second half). If we then measure $g_1, \ldots, g_{n-m}$ on the first halves of the qubits, obtaining outcomes

$g_i = \pm 1$, the post-measurement state is:

$$\prod_{i=1}^{n-m} (P_i^{\pm} \otimes I)(I \otimes E) \left|\beta_{00}\right\rangle^{\otimes n} = (I \otimes E) \prod_{i=1}^{n-m} (P_i^{\pm} \otimes I) \left|\beta_{00}\right\rangle^{\otimes n}$$

where we use that $P_i^{\pm}$ commutes with the error operator as they act on different halves.
Now, observe that $(P_i^{\pm})^T = P_i^{\pm}$ if g_i contains an even number of pauli-Ys, and $(P_i^{\pm})^T = P_i^{\mp}$ if g_i contains an odd number of Pauli-Ys. On the first qubit halves, let us apply Pauli correction operators (based on the measurement outcome, and the Y-parity of the stabilizer generators) such that the transposed projector has sign $+1$. Then, we obtain the state:

$$\begin{aligned}
(I \otimes E) \prod_{i=1}^{n-m} (P_i^{\pm} \otimes I) \left|\beta_{00}\right\rangle^{\otimes n} &= (I \otimes E) \prod_{i=1}^{n-m} ((P_i^{\pm})^2 \otimes I) \left|\beta_{00}\right\rangle^{\otimes n} \\
&= (I \otimes E) \prod_{i=1}^{n-m} (P_i^{\pm} \otimes (P_i^{\pm})^T) \left|\beta_{00}\right\rangle^{\otimes n} \\
&= (I \otimes E) \prod_{i=1}^{n-m} (P_i^{\pm} \otimes P_i^{+}) \left|\beta_{00}\right\rangle^{\otimes n}
\end{aligned}$$

where we have used the idempotency of projectors and the bounce identity. Now, let us apply appropriate Pauli corrections on the first qubit halves again, now to correct everything to $g_i \otimes I = +1$, yielding:

$$(I \otimes E) \prod_{i=1}^{n-m} (P_i^{+} \otimes P_i^{+}) \left|\beta_{00}\right\rangle^{\otimes n}$$

But this is just the encoded $\left|\beta_{00}\right\rangle^{\otimes m}$ state with a weight δn error on one of the (physical) qubit halves! Thus, if we measure $g_1, \ldots, g_{n-m}$ on the second qubit halves, and correct into the subspace where $I \otimes g_i = +1$ for all $i = 1, \ldots, n-m$ (which we can do by assumption on the code distance), we recover the state without the error, which is precisely $\left|\beta_{00}\right\rangle^{\otimes m}$. □

Exercise 12.35

Show that the states $|\xi_{v_k,z,x}\rangle$ defined in (12.202) form an orthonormal basis for a 2^n-dimensional Hilbert space, that is,

$$\sum_{v_k,z,x} |\xi_{v_k,z,x}\rangle\langle\xi_{v_k,z,x}| = I.$$

Hint: for C_1 an $[n, k_1]$ code, C_2 an $[n, k_2]$ code, and $m = k_1 - k_2$, note that there are 2^m distinct values of v_k, 2^{n-k_1} distinct x, and 2^{k_2} distinct z.

Solution

Concepts Involved: CSS Codes, Orthonormal Basis

From our analysis of Ex. 10.27, we know that $\mathbf{CSS}_{z,x}(C_1, C_2)$ has the same properties as $\mathbf{CSS}(C_1, C_2)$, and in particular the codeword states have the same orthogonality property. Further, from the hint there

are $2^m \cdot 2^{n-k_1} \cdot 2^{k_2} = 2^{n+m-(k_1-k_2)} = 2^{n+m-m} = 2^n$ distinct choices of (v_k, x, z), and hence 2^n distinct codeword states $\left|\xi_{v_k,z,x}\right\rangle$.. Since we have 2^n distinct mutually orthogonal states, we conclude that they form an orthonormal basis of a 2^n-dimensional Hilbert space. □

Exercise 12.36

($\star$) Verify Equation (12.203).

Solution

Concepts Involved: Unitary Operators, EPR Pair States

First, note the following "bounce" identity of any maximally entangled state $|\Phi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle |i\rangle$ for any operator A:

$$(A \otimes I) |\Phi\rangle = (I \otimes A^T) |\Phi\rangle$$

To see this, write $A = \sum_{jk} A_{jk} |j\rangle\langle k|$ and then we find:

$$\begin{aligned}
(A \otimes I) |\Phi\rangle &= \frac{1}{\sqrt{d}} \sum_i A |i\rangle \otimes |i\rangle \\
&= \frac{1}{\sqrt{d}} \sum_i (\sum_{jk} A_{jk} |k\rangle \langle j|i\rangle) \otimes |i\rangle \\
&= \frac{1}{\sqrt{d}} \sum_i (\sum_k A_{ik} |k\rangle) \otimes |i\rangle \\
&= \frac{1}{\sqrt{d}} \sum_k |k\rangle \otimes (\sum_i A_{ik} |i\rangle) \\
&= \frac{1}{\sqrt{d}} \sum_k |k\rangle \otimes (A^T |k\rangle) \\
&= (I \otimes A^T) |\Phi\rangle
\end{aligned}$$

Now, consider that $\left\{\left|\xi_{v_k,z,x}\right\rangle\right\}_{v_k,z,x}$ is an ONB (as shown in the previous exercise), and so there is a unitary U that maps from the computational basis to the $\left|\xi_{v_k,z,x}\right\rangle$ basis, i.e. we can write $\left|\xi_{v_k,z,x}\right\rangle = U |j\rangle$. In particular, we can then write:

$$\begin{aligned}
\sum_{v_k,z,x} \left|\xi_{v_k,z,x}\right\rangle \left|\xi_{v_k,z,x}\right\rangle &= \sum_{j=0}^{2^n} (U |j\rangle)(U |j\rangle) \\
&= (U \otimes U) |\beta_{00}\rangle^{\otimes n} \\
&= (I \otimes U^T U) |\beta_{00}\rangle^{\otimes n}
\end{aligned}$$

Since the $\left\{\left|\xi_{v_k,z,x}\right\rangle\right\}_{v_k,z,x}$ are real linear combinations of the computational basis states, this implies the unitary transformation U appearing above is real - in particular, this means that it is orthogonal, and so

$U^T U = I$, and so we conclude

$$\sum_{v_k,z,x} \left|\xi_{v_k,z,x}\right\rangle \left|\xi_{v_k,z,x}\right\rangle = |\beta_{00}\rangle^{\otimes n} = \sum_{j=0}^{2^n} |j\rangle\,|j\rangle$$

as claimed. □

Exercise 12.37

This is an alternative way to understand why Alice's measurements in steps 9 and 10 collapse EPR pairs into random qubits encoded in a random quantum code. Suppose Alice has an EPR pair $(|00\rangle + |11\rangle)/\sqrt{2}$. Show that if she measures the first qubit in the X basis, then the second qubit collapses into an eigenstate of X determined by the measurement result. Similarly, show that if she measures in the Z basis, then the second qubit is left in a Z eigenstate labeled by the measurement result. Using this observation and the results of section 10.5.8, conclude that Alice's measurements of $H_1, H_2^\perp$, and $\bar{Z}$ on her EPR pair halves result in a random codeword of $CSS_{z,x}(C_1, C_2)$ determined by her measurement results.

Solution

Concepts Involved: Bell Pairs, Quantum Measurements, Stabilizer codes, CSS codes

From:

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

Alice measuring the first qubit measuring X and hence applying $\Pi_\pm = |\pm\rangle\langle\pm|$ on the first qubit collapses the second qubit into the same X-eigenstate. Identically, measuring Z and applying $\Pi_{0/1} = |0/1\rangle\langle 0/1|$ on the first qubit collapses the second qubit into the same Z-eigenstate.
From our analysis in Section 10.5.8, Alice measuring $H_1, H_2^\perp, \bar{Z}$ on her EPR halves collapses her qubit halves into the state stabilized by $\pm g_1^z, \ldots, \pm g_{n-k_1}^z, \pm g_1^x, \ldots, \pm g_{k_2}^x, \pm\bar{Z}_1, \ldots, \pm\bar{Z}_m$ with H_1 corresponding to the Z-stabilizers, $H_2^\perp$ corresponding to the X-stabilizers, and $\bar{Z}_i$ corresponding to the Z-logicals, all of which are made purely up of strings of purely physical Z or X Paulis. Thus, by our above analysis, the state of the unmeasured EPR halves is also stabilized by $\pm g_1^z, \ldots, \pm g_{n-k_1}^z, \pm g_1^x, \ldots, \pm g_{k_2}^x, \pm\bar{Z}_1, \ldots, \pm\bar{Z}_m$. The H_1 measurement outcomes determining x, the $H_2^\perp$ outcomes determining z, and the $\bar{Z}$ outcomes determining v_k in the codeword state:

$$\left|\xi_{v_k,z,x}\right\rangle = \frac{1}{\sqrt{|C_2|}} \sum_{w \in C_2} (-1)^{z\cdot w} |v_k + w + x\rangle$$

i.e. her measurement results in a random codeword of $\mathbf{CSS}_{z,x}(C_1, C_2)$ on the unmeasured qubits based on her measurement outcomes. □

Exercise 12.38

($\star$) Show that if you had the ability to distinguish non-orthogonal states, then it would be possible to compromise the security of BB84, and indeed, all of the QKD protocols we have described.

Solution

Concepts Involved: Quantum Distinguishability, BB84

From Ex. 1.2, we know that the ability to distinguish non-orthogonal states is equivalent to the ability to clone. Let us thus play the role of an (all-powerful) Eavesdropper Eve who has the power to clone states, in the Secure BB84 protocol.

In step 3, Eve can intercept the qubits that Alice sends to Bob. She can then clone them twice, and then send the (unmodified) qubit to Bob. While Bob measures his qubits randomly in Z, X, Eve can measure her first copy of each qubit in Z, and the second copy of each qubit in X. Thus, when Alice publicly announces b in step 6, Eve can keep the outcomes which correspond to the measurement basis in the announced b, and keep the appropriate bits when Alice announces the check/data bits in step 7. In step 8, since Eve has not interfered with the state of the qubits (only cloning them), Alice and Bob's check bits agree and so they do not abort the protocol.

In the information reconcilation/privacy amplification step, we assume that Eve has also deduced the classical codes C_1, C_2 that are to be used by Alice/Bob (this information is not encoded/securely stored, and so we assume that Eve has access to it). Thus, when Alice announces $x - v_k$ (for randomly chosen $v_k \in C_1$, Eve (like Bob), holding $x(+\epsilon)$ van subtract this from her result to obtain v_k, and then compute the coset of $v_k + C_2$ in C_1 to obtain the key k, as Alice and Bob have.

The key aspect of the power Eve has enters in step 3, when she is able to intercept and clone the sent qubits arbitrarily, allowing her to extract all information while making it appear on Alice/Bob's side as if nothing has happened. Indeed, this compromising strategy works for all discussed QKD protocols - Eve can intercept the sent qubits from Alice, clone as many times as needed before passing on an untouched state to Bob (thus leaving her interference undetected), and measure in as many bases as is necessary to extract the key in parallel with Alice/Bob. □

Problem 12.1

($\star\star$) In this problem, we will work through an alternate proof of the Holevo bound. Define the *Holevo chi quantity*,

$$\chi \equiv S(\rho) - \sum_x p_x S(\rho_x)$$

(1) Suppose the quantum system consists of two parts, A and B. Show that

$$\chi_A \leq \chi_{AB}.$$

(*Hint:* Introduce an extra system which is correlated with AB, and apply strong subadditivity.)

(2) Let $\mathcal{E}$ be a quantum operation. Use the previous result to show that

$$\chi' \equiv S(\mathcal{E}(\rho)) - \sum_x p_x S(\mathcal{E}(\rho_x)) \leq \chi \equiv S(\rho) - \sum_x p_x S(p_x)$$

That is, the Holevo chi quantity decreases under quantum operations. This is an important and useful fact in its own right.

(3) Let E_y be a set of POVM elements. Augment the quantum system under consideration with an

'apparatus' system, M, with an orthonormal basis $|y\rangle$. Define a quantum operation by

$$\mathcal{E}(\rho \otimes |0\rangle\langle 0|) \equiv \sum_y \sqrt{E_y}\rho\sqrt{E_y} \otimes |y\rangle\langle y|$$

where $|0\rangle$ is some standard pure state of M. Prove that after the action of $\mathcal{E}$, $\chi_M = H(X:Y)$. Use this and the previous two results to show that

$$H(X:Y) \leq S(\rho) - \sum_x p_x S(\rho_x)$$

which is the Holevo bound.

Solution

Concepts Involved: Von Neumann Entropy, Strong Subadditivity, Quantum Operations, Kraus Representation, Mutual Information

(1) As is introduced in the proof of the Holevo bound given in Section 12.1.1, introduce a system C such that:

$$\rho^{ABC} = \sum_x \rho_x^{AB} \otimes p_x \left|x\right\rangle\!\left\langle x\right|^C$$

Wherein (Adapting Eq. (12.11) of the text):

$$S(A,B:C) = S(A,B) + S(C) - S(A,B,C) = S(\rho^{AB}) - \sum_x p_x S(\rho_x^{AB}) = \chi_{AB}$$

By strong subadditivity, we have:

$$S(A,B,C) + S(A) \leq S(A,B) + S(A,C) \implies S(A,B,C) \leq S(A,B) + S(A,C) - S(A)$$

and so combining this with the previous equation:

$$\chi_{AB} \geq S(A,B) + S(C) - S(A,B) - S(A,C) + S(A) = S(A) - S(A,C) + S(C) \geq S(A) - S(A,C)$$

where the last inequality follows from the fact that $S(C) \geq 0$ (entropies are non-negative).

Now, observe that:

$$\rho^{AC} = \mathrm{Tr}_B(\rho^{ABC}) = \sum_x \mathrm{Tr}_B(\rho_x^{AB}) \otimes p_x \left|x\right\rangle\!\left\langle x\right|^C = \sum_x p_x \rho_x^A \otimes \left|x\right\rangle\!\left\langle x\right|^C$$

wherein:

$$S(A,C) = S(\rho^{AC}) = \sum_x p_x S(\rho_x^A)$$

and hence:

$$\chi_{AB} \geq S(A) - S(A,C) = S(\rho^a) - \sum_x p_x S(\rho_x^A) = \chi_A$$

as claimed.

(2) Suppose Q is our quantum system and let E be the environment system such that $\mathcal{E}$ on Q is mocked up as a unitary operation on QE with the environment E in some fixed pure initial state (unentangled with Q). Then, we have that $\chi_{QE} = \chi_Q$ for the initial state. Further, we recall that unitary operations leave the Von Neumann entropy invariant, and hence leave the Holevo χ-quantity invariant:

$$\chi_{U\rho U^\dagger} = S(U\rho U^\dagger) - \sum_x p_x S(U\rho_x U^\dagger) = S(\rho) - \sum_x p_x S(\rho_x) = \chi_\rho$$

Thus, $\chi_{Q'E'} = \chi_{QE}$ (with primes denoting that the unitary operation between Q/E mocking up $\mathcal{E}$ has been performed). Thus:

$$\chi_{Q'} \leq \chi_{Q'E'} = \chi_{QE} = \chi_Q$$

where the inequality follows from part (1).

(3) Computing $\rho_{M'}$ after the application of $\mathcal{E}$, we find:

$$\begin{aligned}
\rho^{M'} &= \mathrm{Tr}_Q(\sum_y \sqrt{E_y}\rho\sqrt{E_y} \otimes |y\rangle\langle y|) \\
&= \sum_y \mathrm{Tr}\big(\rho E_y\big)\, |y\rangle\langle y| \\
&= \sum_y p(y)\, |y\rangle\langle y|
\end{aligned}$$

where the last equality (in more detail) can be seen from:

$$\mathrm{Tr}\big(\rho E_y\big) = \mathrm{Tr}\left(\sum_x p_x \rho_x E_y\right) = \sum_x p_x\, \mathrm{Tr}\big(\rho_x E_y\big) = \sum_x p_x p(y|x) = \sum_x p(x,y) = p(y).$$

Thus:

$$S(\rho^{M'}) = -\sum_y p(y) \log\big(p(y)\big) = H(Y)$$

For the other term in $\chi_{M'}$, we calculate:

$$\rho_x^{M'} = \mathrm{Tr}_Q(\sum_y \sqrt{E_y}\rho_x\sqrt{E_y} \otimes |y\rangle\langle y|) = \sum_y \mathrm{Tr}\big(\rho_x E_y\big)\, |y\rangle\langle y| = \sum_y p(y|x)\, |y\rangle\langle y|$$

Thus:

$$\sum_x p_x S(\rho_x^{M'}) = \sum_x p_x(-\sum_y p(y|x)\log\big(p(y|x)\big)) = \sum_x p_x H(Y|X=x) = H(Y|X)$$

Therefore:

$$\chi_{M'} = S(\rho^{M'}) - \sum_x p_x S(\rho_x^{M'}) = H(Y) - H(Y|X) = H(Y:X) = H(X:Y)$$

which proves the claim.

Combining the previous results, we note:

$$\chi_{M'} \leq \chi_{Q'M'} \leq \chi_{QM} = \chi_Q$$

where the first inequality follows from part (1), the second inequality from part (2), and the last equality from the fact that M is in a pure fixed/reference state before the operation. Thus:

$$H(X:Y) \leq S(\rho) - \sum_x p_x S(\rho_x)$$

as claimed.

□

Problem 12.2

($\star$) This result is an extension of the previous problem. Provide a proof of the no-cloning theorem by showing that a cloning process for non-orthogonal pure states would necessarily increase χ.

Solution

Concepts Involved: No-Cloning, Density Operators, Von Neumann Entropy

Let $\rho_\psi = |\psi\rangle\langle\psi|, \rho_\varphi = |\varphi\rangle\langle\varphi|$ be non-orthogonal pure states, and suppose for the sake of contradiction that there existed a cloning quantum operation $\mathcal{E}$ such that:

$$\mathcal{E}(\rho_\psi \otimes \sigma) = \rho_\psi \otimes \rho_\psi$$

$$\mathcal{E}(\rho_\varphi \otimes \sigma) = \rho_\varphi \otimes \rho_\varphi$$

for some reference pure state σ. Then, consider the mixture:

$$\rho = \frac{1}{2}\rho_\psi \otimes \sigma + \frac{1}{2}\rho_\varphi \otimes \sigma = \frac{\rho_\psi + \rho_\varphi}{2} \otimes \sigma$$

In problem 12.4, we will derive a contradiction by applying the cloning map separately to $\frac{1}{2}\rho_\psi \otimes \sigma + \frac{1}{2}\rho_\varphi \otimes \sigma$ and $\frac{\rho_\psi + \rho_\varphi}{2} \otimes \sigma$. For the present problem, we only assume ρ_ψ, ρ_φ can be cloned individually by $\mathcal{E}$, and so evaluate $\mathcal{E}(\rho)$ using linearity to be:

$$\mathcal{E}(\rho) = \mathcal{E}(\frac{1}{2}\rho_\psi \otimes \sigma + \frac{1}{2}\rho_\varphi \otimes \sigma) = \frac{1}{2}\mathcal{E}(\rho_\psi \otimes \sigma) + \frac{1}{2}\mathcal{E}(\rho_\varphi \otimes \sigma) = \frac{\rho_\psi \otimes \rho_\psi + \rho_\varphi \otimes \rho_\varphi}{2}$$

Computing the Holevo χ-quantity for the initial state ρ, we have:

$$\begin{aligned}
\chi &= S(\frac{\rho_\psi + \rho_\varphi}{2} \otimes \sigma) - \left(\frac{1}{2}S(\rho_\psi \otimes \sigma) + \frac{1}{2}S(\rho_\varphi \otimes \sigma)\right) \\
&= S(\frac{\rho_\psi + \rho_\varphi}{2}) + S(\sigma) - \frac{1}{2}\left(S(\rho_\psi) + S(\sigma) + S(\rho_\varphi) + S(\sigma)\right) \\
&= S(\frac{\rho_\psi + \rho_\varphi}{2})
\end{aligned}$$

where we use the entropy of a tensor product result from Ex. 11.13 and that the entropy of pure states is zero.
Computing the χ quantity for the post-cloning process we find:

$$\begin{aligned}
\chi' &= S(\mathcal{E}(\rho)) - \left[\frac{1}{2}S(\mathcal{E}(\rho_\psi \otimes \sigma)) + \frac{1}{2}S(\mathcal{E}(\rho_\varphi \otimes \sigma))\right] \\
&= S(\frac{\rho_\psi \otimes \rho_\psi + \rho_\varphi \otimes \rho_\varphi}{2}) - \frac{1}{2}\left[S(\rho_\psi \otimes \rho_\psi) + S(\rho_\varphi \otimes \rho_\varphi)\right] \\
&= S(\frac{\rho_\psi \otimes \rho_\psi + \rho_\varphi \otimes \rho_\varphi}{2}) - \frac{1}{2}\left[S(\rho_\psi) + S(\rho_\psi) + S(\rho_\varphi) + S(\rho_\varphi)\right] \\
&= S(\frac{\rho_\psi \otimes \rho_\psi + \rho_\varphi \otimes \rho_\varphi}{2})
\end{aligned}$$

We now compare χ and χ' by computing the entanglement entropies that appear. To this end let us compute the 2 (nonzero) eigenvalues of $\rho, \mathcal{E}(\rho)$. Denoting the overlap of $|\psi\rangle, |\varphi\rangle$ by $t = \langle\psi|\varphi\rangle$, we can write:

$$|\varphi\rangle = t\,|\psi\rangle + \sqrt{(1-|t|^2)}\,|\psi^\perp\rangle$$

Note that $0 < |t| < 1$, with the lower bound coming from the fact that $|\psi\rangle, |\varphi\rangle$ are non-orthogonal and the upper bound from the fact that they are not identical. We can then write:

$$\begin{aligned}
\rho &= \frac{1}{2}\rho_\psi + \frac{1}{2}\rho_\varphi \\
&\cong \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} |t|^2 & t\sqrt{(1-|t|^2)} \\ t^*\sqrt{(1-|t|^2)} & 1-|t|^2 \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 1+|t|^2 & t\sqrt{(1-|t|^2)} \\ t^*\sqrt{(1-|t|^2)} & 1-|t|^2 \end{bmatrix}
\end{aligned}$$

where we have represented the density operators in the $|\psi\rangle, |\psi^\perp\rangle$ subspace (where they have nonzero

support). To compute the eigenvalues, note that $\mathrm{Tr}(\rho) = \lambda_+ + \lambda_- = 1$, and that:

$$\begin{aligned}\det(\rho) &= \lambda_+ \cdot \lambda_- \\ &= \frac{1+|t|^2}{2} \cdot \frac{1-|t|^2}{2} - \frac{t\sqrt{1-|t|^2}}{2}\frac{t^*\sqrt{1-|t|^2}}{2} \\ &= \frac{1-|t|^2}{4}\end{aligned}$$

From which we obtain:

$$\lambda_\pm = \frac{1 \pm |t|}{2}$$

and the other eigenvalues are 0.
Let us next calculate the eigenvalues of $\mathcal{E}(\rho) = \frac{\rho_\psi \otimes \rho_\psi + \rho_\varphi \otimes \rho_\varphi}{2}$. Performing Gram-Schmidt on $|\varphi\rangle \otimes |\varphi\rangle$, we find:

$$\begin{aligned}|\varphi\rangle \otimes |\varphi\rangle &= t^2 |\psi\rangle \otimes |\psi\rangle + \sqrt{1-|t^2|^2}\left|\psi'^{\perp}\right\rangle \otimes \left|\psi'^{\perp}\right\rangle \\ &= t^2 |\psi\rangle \otimes |\psi\rangle + \sqrt{1-|t|^4}\left|\psi'^{\perp}\right\rangle \otimes \left|\psi'^{\perp}\right\rangle\end{aligned}$$

where the coefficient of $|\psi\rangle \otimes |\psi\rangle$ is the square of the overlap of $|\psi\rangle, |\varphi\rangle$. We can then write:

$$\begin{aligned}\mathcal{E}(\rho) &= \frac{1}{2}\rho_\psi \otimes \rho_\psi + \frac{1}{2}\rho_\varphi \otimes \rho_\varphi \\ &= \frac{1}{2}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + \frac{1}{2}\begin{bmatrix}|t|^4 & t^2\sqrt{(1-|t|^4)}\\ (t^*)^2\sqrt{(1-|t|^4)} & 1-|t|^4\end{bmatrix} \\ &= \begin{bmatrix}1+|t|^4 & t^2\sqrt{1-|t|^4}\\ (t^*)^2\sqrt{1-|t|^4} & 1-|t|^4\end{bmatrix}\end{aligned}$$

Noting that $\mathrm{Tr}\big(\mathcal{E}(\rho)\big) = \lambda_+ + \lambda_- = 1$ and:

$$\det(\rho) = \lambda_+ \cdot \lambda_- = \frac{1-|t|^4}{4}$$

we find:

$$\lambda_\pm = \frac{1 \pm |t|^2}{2}$$

With the eigenvalues in hand, we can compute:

$$\chi = S(\rho) = -\sum_i \lambda_i \log \lambda_i = H_{\mathsf{bin}}(\frac{1+|t|}{2})$$

$$\chi' = S(\rho) = -\sum_i \lambda_i \log \lambda_i = H_{\text{bin}}(\frac{1+|t|^2}{2})$$

where $H_{\text{bin}}(p)$ is the binary Shannon entropy, which is maximized at $p = 1/2$ and decreasing over $[1/2, 1]$. Since $0 < |t| < 1$, it follows that $\frac{1}{2} < \frac{1+|t|^2}{2} < \frac{1+|t|}{2}$ and so $\chi' > \chi$. This contradicts the result of problem 12.1(2), and thus we conclude that the cloning map $\mathcal{E}$ cannot exist.

□

Problem 12.3

($\star\star\star$) For a fixed quantum source and rate $R > S(\rho)$, design a quantum circuit implementing a rate R compression scheme.

Problem 12.4: Linearity forbids cloning

Supose we have a quantum machine with two slots, A and B. Slot A, the *data slot*, starts out in an unknown quantum state ρ. This is the state to be copied. Slot B, the *target slot*, starts out in some standard quantum state, σ. We will assume that any candidate copying procedure is *linear* in the initial state,

$$\rho \otimes \sigma \mapsto \mathcal{E}(\rho \otimes \sigma) = \rho \otimes \rho$$

where $\mathcal{E}$ is some linear function. Show that if $\rho_1 \neq \rho_2$ are density operations such that

$$\mathcal{E}(\rho_1 \otimes \sigma) = \rho_1 \otimes \rho_1$$
$$\mathcal{E}(\rho_2 \otimes \sigma) = \rho_2 \otimes \rho_2$$

then any mixture of ρ_1 and ρ_2 is not copied correctly by this procedure.

Solution

Concepts Involved: No-Cloning, Density Operators

Denote $\rho = p_1\rho_1 + p_2\rho_2$ an arbitrary mixture of ρ_1, ρ_2. By the linearity of $\mathcal{E}$, we have:

$$\begin{aligned}\mathcal{E}(\rho \otimes \sigma) &= \mathcal{E}((p_1\rho_1 + p_2\rho_2) \otimes \sigma) \\ &= p_1\mathcal{E}(\rho_1 \otimes \sigma) + p_2\mathcal{E}(\rho_2 \otimes \sigma) \\ &= p_1\rho_1 \otimes \rho_1 + p_2\rho_2 \otimes \rho_2\end{aligned}$$

Let us compare the above to the cloning of ρ directly, i.e. to $\mathcal{E}(\rho \otimes \sigma) = \rho \otimes \rho$:

$$\begin{aligned}\rho \otimes \rho &= (p_1\rho_1 + p_2\rho_2) \otimes (p_1\rho_1 + p_2\rho_2) \\ &= p_1^2\rho_1 \otimes \rho_1 + p_1p_2(\rho_1 \otimes \rho_2 + \rho_2 \otimes \rho_1) + p_2^2\rho_2 \otimes \rho_2\end{aligned}$$

The two expressions are only equivalent if the cross-terms vanish - $\rho_1 \otimes \rho_2 + \rho_2 \otimes \rho_1$ is nonvanishing as the finite sum of positive operators cannot be zero, and so this only occurs if $p_1p_2 = 0$, i.e. one of p_1, p_2 vanishes (and the other is one). Thus arbitrary mixtures cannot be cloned. □

Problem 12.5: Classical capacity of a quantum channel (Research)

($\star\star\star$) Is the product state capacity (12.71) the true capacity of a noisy quantum channel for classical information, that is, the capacity when entangled inputs to the channel are allowed?

Solution

Concepts Involved: Quantum Channels, Capacity

The answer turns out to be *no*, as shown by a random construction due to Hastings in arXiv:0809.3972 - entanglement *can* boost the classical capacity of a quantum channel. □

Problem 12.6: Methods for achieving capacity (Research)

($\star\star\star$) Find an efficient construction for codes achieving rates near the product state capacity (12.71) of a noisy quantum channel for classical information.

Solution

Concepts Involved: Quantum Channels, Capacity

Using polar codes (first introduced by Arikan - arXiv:0807.3917), Wilde and Guha (arXiv:1109.2591) presented an explicit construction saturating the product state capacity (this is in constrast to the HSW theorem, which uses a random construction). Their polar code construction has complexity $O(N \log N)$ (efficient) in the encoding - showing that the decoding is also efficient/$O(N \log N)$ (as in the purely classical case) for general quantum channels appears to remain an open question, though there is progress towards the results by Renes/Dupuis/Renner arXiv:1109.3195 and by Wilde/Landon-Cardinal/Hayden arXiv:1302.0398. □

Problem 12.7: Quantum channel capacity (Research)

($\star\star\star$) Find a method to evaluate the capacity of a given quantum channel $\mathcal{E}$ for the transmission of quantum information.

Solution

Concepts Involved: Quantum Channels, Capacity

Formally, the LSD (Lloyd - arXiv:quant-ph/9604015, Shor - arXiv:quant-ph/0304102, Devetak - arXiv:quant-ph/0304127) theorem yields such a formula:

$$Q(\mathcal{N}) = \lim_{n\to\infty} \max_\rho I(\rho, \mathcal{E}^{\otimes n})$$

with $I(\rho, \mathcal{E}^{\otimes n})$ the coherent information for n copies of the channel $\mathcal{E}$. However, this is an optimization carried out over $n \to \infty$ uses of $\mathcal{E}$, and is generally not computable.
Though there exist special cases, e.g. degradable channels (as discussed by Shor and Devetak, arXiv:quant-ph/0311131) for which the capacity reduces to the single-shot case $\max_\rho I(\rho, \mathcal{E})$, generically coherent information is not additive but can be superadditive (see DiVincenzo/Shor/Smolin's quant-ph/9706061

for a first example), and so this (intractable) optimization over n must be done - thus a general method as this problem poses seems to not exist. □

A1 Notes on basic probability theory

Exercise A1.1

Prove Bayes' rule.

Solution

Concepts Involved: Probability, Conditional Probability.
Recall that conditional probabilities were defined as:

$$P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)}$$

and also recall that Bayes' rule is given by:

$$p(x|y) = p(y|x)\frac{p(x)}{p(y)}$$

By the definition of conditional probability:

$$p(y|x)\frac{p(x)}{p(y)} = \frac{p(X = x, Y = y)}{p(x)}\frac{p(x)}{p(y)} = \frac{P(X = x, Y = y)}{p(y)} = p(x|y)$$

□

Exercise A1.2

Prove the law of total probability.

Solution

Concepts Involved: Probability, Conditional Probability.
Recall that the law of total probability is given by:

$$p(y) = \sum_x p(y|x)p(x)$$

Using the identity $p(Y = y) = \sum_x p(X = x, Y = y)$ and Bayes' rule, we have

$$p(y) = \sum_x p(x, y) = \sum_x p(y|x)p(x)$$

□

Exercise A1.3

Prove that there exists a value of $x \geq \mathbf{E}(X)$ such that $p(x) > 0$.

Solution

Concepts Involved: Probability, Expectation.
Recall that the expectation of a random variable X is defined by:

$$\mathbf{E}(X) = \sum_x p(x)x$$

Let $\tilde{x} = \max\{x : x \text{ is a possible value of } X\}$. This maximum exists as we assume X can only take on a finite set of values. We therefore have that:

$$\mathbf{E}(X) = \sum_x p(x)x \leq \sum_x p(x)\tilde{x} = \tilde{x}\sum_x p(x) = \tilde{x}$$

Where in the last equality we use that the sum over all probabilities must be 1. □

Exercise A1.4

Prove that $\mathbf{E}(X)$ is linear in X

Solution

Concepts Involved: Probability, Expectation

Let $a, b \in \mathbb{R}$ and X, Y be random variables. We then have that:

$$\begin{aligned}
\mathbf{E}(aX + bY) &= \sum_x \sum_y p(x,y)(ax + by) \\
&= \sum_x \sum_y p(x,y)ax + \sum_x \sum_y p(x,y)by \\
&= a\sum_x \left(\sum_y p(x,y)\right) x + b\sum_y \left(\sum_x p(x,y)\right) y \\
&= a\sum_x p(x)x + b\sum_y p(y)y \\
&= a\mathbf{E}(X) + b\mathbf{E}(y)
\end{aligned}$$

which shows that expectation is linear. □

Exercise A1.5

Prove that for independent random variables X and Y, $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$.

Solution

Concepts Involved: Probability, Expectation, Independent Random Variables.

Recall two random variables X, Y are independent if

$$p(X = x, Y = y) = p(X = x)p(Y = y)$$

We have that:

$$\mathbf{E}(XY) = \sum_x \sum_y xyp(x, y) = \sum_x \sum_y xyp(x)p(y) = \left(\sum_x p(x)x\right)\left(\sum_y p(y)y\right) = \mathbf{E}(x)\mathbf{E}(y)$$

□

Exercise A1.6

($\star\star$) Prove Chebyshev's inequality.

Solution

Concepts Involved: Probability, Expectation, Variance.
Recall the definition of the variance and standard deviaiton of a random variable X:

$$\mathrm{Var}(X) = \mathbf{E}[(X - \mathbf{E}(X))^2], \quad \Delta(X) = \sqrt{\mathrm{Var}(X)}$$

Also, recall that Chebyshev's inequality reads:

$$p(\left|X - \mathbf{E}(X)\right| \geq \lambda\Delta(X)) \leq \frac{1}{\lambda^2}$$

where $\lambda > 0$.

We first establish Markov's inequality for the expectation value $\mathbf{E}(X)$. Let $a > 0$, and then we have that:

$$\mathbf{E}(X) = \sum_x xp(x) = \sum_{x \geq a} xp(x) + \sum_{x < a} xp(x) \geq \sum_{x \geq a} ap(x) + 0 = ap(X \geq a)$$

Therefore, we obtain that:

$$p(X \geq a) \leq \frac{\mathbf{E}(X)}{a}$$

for any random variable X and $a > 0$. Next, substitute X with $(X - \mathbf{E}(X))^2$ and let $a = \lambda^2 \mathrm{Var}(X)$ for $\lambda > 0$. Markov's inequality then states that:

$$p((X - \mathbf{E}(X))^2 \geq \lambda^2 \mathrm{Var}(X)) \leq \frac{\mathbf{E}(X - \mathbf{E}(X))^2}{\lambda^2 \mathrm{Var}(X)}$$

Since $\mathbf{E}(X - \mathbf{E}(X))^2 = \mathrm{Var}(X)$, we have that:

$$p((X - \mathbf{E}(X))^2 \geq \lambda^2 \mathrm{Var}(X)) \leq \frac{1}{\lambda^2}$$

If $\lambda > 0$, then $p((X - \mathbf{E}(X))^2 \geq \lambda^2 \operatorname{Var}(X)) = p(\big|X - \mathbf{E}(X)\big| \geq \lambda \Delta(X))$ by taking square roots, so we obtain:

$$p(\big|X - \mathbf{E}(X)\big| \geq \lambda \Delta(X)) \leq \frac{1}{\lambda^2}$$

as desired. □

A2 Group theory

Exercise A2.1

Prove that for any element of g of a finite group, there always exists a positive integer r such that $g^r = e$. That is, every element of such a group has an order.

Solution

Concepts Involved: Group Axioms, Order

Suppose G is a finite group, and $g \in G$. Then, there exists some $r_1, r_2 \in \mathbb{N}$ such that $r_1 \neq r_2$ and $g^{r_1} = g^{r_2}$. If this was not the case, then g^n would be unique for each $n \in \mathbb{N}$, contradicting the finiteness of G. WLOG take $r_1 < r_2$, and let $r = r_2 - r_1 \in \mathbb{N}$. Using associativity, we then have that:

$$g^{r_1} = g^{r_2} = g^{r_1 + r} = g^{r_1} g^r$$

from which we conclude that $g^r = e$. □

Exercise A2.2

($\star\star$) Prove Lagrange's Theorem.

Solution

Concepts Involved: Group Axioms, Subgroups, Order, Equivalence Relations

Let H be a subgroup of a group G and define the relation $\sim$ by $a \sim b$ iff $a = bh$ for some $h \in H$. $\sim$ is reflexive as $a = ae$ (so $a \sim a$) where $e \in H$ is the identity element. $\sim$ is symmetric as if $a \sim b$, then $a = bh$ for some $h \in H$ so $b = ah^{-1}$ (so $b \sim a$) where $h^{-1} \in H$ as H is closed under inverses. Finally $\sim$ is transitive as if $a \sim b$ and $b \sim c$, there exist $h_1, h_2 \in H$ such that $a = bh_1$ and $b = ch_2$ so $a = ch_2h_1$. As $h_2h_1 \in H$ (H is closed under multiplication) it follows that $a \sim c$. Having shown $\sim$ to have these three properties, we conclude it is an equivalence relation. Then, the equivalence classes of $\sim$ partition G, where the equivalence class of $g \in G$ is $[g] = \{gh | h \in H\}$.
Now, let $g \in G$ and define the map $\varphi_g : H \to [g]$ as $\varphi_g(h) = gh$. φ_g is injective as if $\varphi_g(h_1) = \varphi_g(h_2)$ then $gh_1 = gh_2$ and multiplying by g^{-1} on both sides $h_1 = h_2$. φ_g is surjective as if $k \in [g]$, then there exists some $h \in H$ such that $k = gh$ by the definition of $\sim$. Hence φ_g is bijective.
As per our prior observation, the equivalence classes of $\sim$ partition G, so $G = \bigcup_{i=1}^n [g_i]$ and $|G| = \left|\bigcup_{i=1}^n [g_i]\right| = \sum_{i=1}^n \left|[g_i]\right|$. Further, there is a bijection φ_{g_i} from each equivalence class to H, so $\left|[g_i]\right| = |H|$ for all i. Thus $|G| = \sum_{i=1}^n |H| = nH$ and hence $|H|$ divides $|G|$, as desired. □

Exercise A2.3

Show that the order of an element $g \in G$ divides $|G|$.

Solution

Concepts Involved: Group Axioms, Subgroups, Order, Lagrange's Theorem

Let $g \in G$ with order r. Then, define $H = \{g^n | n \in \mathbb{N}\}$. We claim that H is a subgroup of G. First, $g^n \in G$ for any n as G is closed under multiplicaton, so $H \subset G$. Next, if $g^{n_1}, g^{n_2} \in H$ then $g^{n_1} \cdot g^{n_2} = g^{n_1+n_2} \in H$. Associativity is inherited from the associativity of multiplication in G. Since $g^r = e \in H$, H contains the identity. Finally, for $g^k \in H$ we have $g^{r-k} \in H$ such that $g^k g^{r-k} = g^{r-k} g^k = g^r = e$ so H is closed under inverses. Hence the claim is proven.
Next, we observe that $|H| = r$ as H contains the r elements $e, g, g^2, \ldots g^{r-1}$. Hence by Lagrange's Theorem r divides $|G|$. □

Exercise A2.4

Show that if $y \in G_x$ then $G_y = G_x$

Solution

Concepts Involved: Group Axioms, Conjugacy Classes

Suppose $y \in G_x$. Then there exists some $g \in G$ such that $g^{-1}xg = y$. Multiplying both sides on the left by g and on the right by g^{-1} we find that $x = gyg^{-1}$. We now show the two inclusions.
$\boxed{\subseteq}$ Suppose that $k \in G_x$. Then there exists some $g' \in G$ such that $k = g'^{-1}xg'$. Then using $x = gyg^{-1}$ we find $k = g'^{-1}gyg^{-1}g'$. Now, $g^{-1}g' \in G$ (by closure) and it has inverse $g'^{-1}g$, and hence $k = g'^{-1}gyg^{-1}g' \in G_y$. So, $G_x \subseteq G_y$.
$\boxed{\supseteq}$ Suppose that $l \in G_y$. Then there exists some $g'' \in G$ such that $l = g''^{-1}yg''$. Then with $g^{-1}xg = y$ we find $l = g''^{-1}g^{-1}xgg''$. Much like before, $gg'' \in G$ (by closure) with inverse $g''^{-1}g^{-1}$ so $l \in G_x$. So, $G_y \subseteq G_x$.
We conclude that $G_y = G_x$. □

Exercise A2.5

Show that if x is an element of an Abelian group G, then $G_x = \{x\}$.

Solution

Concepts Involved: Abelian Groups, Conjugacy Classes

Evidently $x = e^{-1}xe \in G_x$ so $\{x\} \subseteq G_x$. Next, if $k \in G_x$ then $k = g^{-1}xg$ for some $g \in G$, but since G is abelian, $g^{-1}x = xg^{-1}$ so $k = xg^{-1}g = xe = x$ so $k \in \{x\}$ and hence $G_x \subseteq \{x\}$. We conclude that $G_x = \{x\}$. □

Exercise A2.6

Show that any group of prime order is cyclic.

Solution

Concepts Involved: Order, Cyclic Groups

Suppose $|G| = p$ where p is prime. Since G is finite, every element of G has an order by Exercise A2.1. Since the order of any element $g \in G$ divides $|G| = p$ by Exercise A2.3, and since p is prime, the order of g is either 1 or p. Since $|G| > 1$, there exists at least one $g \in G$ with order p, and this g is a generator of G (with $g^1 = g, g^2, g^3, \ldots, g^p = e$ distinct and comprising all the elements of G. In fact this is true of any non-identity g). Hence G is cyclic. □

Exercise A2.7

($\star$) Show that every subgroup of a cyclic group is cyclic.

Solution

Concepts Involved: Group Axioms, Subgroups, Cyclic Groups, Euclid's Division Algorithm

First we prove a necessary Lemma, namely that any nonempty subset of the natural numbers contains a least element. We show this by proving the contrapositive. Suppose that $A \subseteq \mathbb{N}$ has no least element. Then $1 \notin A$ as then 1 would be the least element. Suppose then that $1, \ldots k-1 \notin A$; then $k \notin A$ as then k would be the least element. By strong induction, there exists no $k \in \mathbb{N}$ such that $k \in A$, i.e. A is empty. This concludes the proof of the lemma.
Let $G = \langle a \rangle$ be a cyclic group and H a subgroup of G. If $H = \{e\}$, it is trivially cyclic and we are done. If $H \neq \{e\}$, then there exists some $a^n \in H$ with $n \neq 0$. Since H is closed under inverses, $(a^n)^{-1} = a^{-n} \in H$ as well which ensures that H contains some positive power of a. Then consider the set $A = \{k \in \mathbb{N} | a^k \in H\}$. Any nonempty subset of the naturals has a minimum element; therefore let $d = \min A$. It is immediate that $\langle a^d \rangle$ is a subgroup of H as $a^d \in H$ and H is a group. To show the reverse containment, suppose that $g \in H$. Since H is a subgroup of the cyclic G, it follows that $g = a^p$ for some $p \in \mathbb{Z}$. We can then write $p = qd + r$ for $0 \leq r < d$ by Euclid's Division algorithm (see Appendix 4). We then have that $a^r = a^{p-qd} = a^p (a^d)^{-q} \in H$ by closure. Now, since d is the least positive integer for which $a^d \in H$ and $0 \leq r < d$, it must follow that $r = 0$. Therefore, $p = qd$ and hence $a^{qd} = (a^d)^q \in \langle a^d \rangle$. So, H is a subgroup of $\langle a^d \rangle$. We conclude that $H = \langle a^d \rangle$ and hence H is cyclic.
□

Exercise A2.8

Show that if $g \in G$ has finite order r, then $g^m = g^n$ if and only if $m = n \pmod{r}$.

Solution

Concepts Involved: Order, Modular Arithmetic, Euclid's Division Algorithm

Suppose $g \in G$ has finite order r.
$\boxed{\Longleftarrow}$ First suppose that $m = n \pmod{r}$. Then $m - n = kr$ for some $k \in \mathbb{N}$. Therefore $g^{m-n} = g^{kr}$. But $g^{kr} = (g^r)^k = e^k = e$, so $g^{m-n} = g^m g^{-n} = e$, and multiplying both sides by g^n we find $g^m = g^n$.
$\boxed{\Longrightarrow}$ Suppose $g^m = g^n$. Then multiplying both sides by g^{-n} we find $g^{m-n} = e$. By Euclid's Division algorithm there exist integers q, p such that $m - n = qr + p$ with $0 \leq p < r$. We then have that

$g^{m-n} = g^{qr+p} = g^{qr}g^p = e$. Furthermore, $g^{qr} = (g^r)^q = e^q = e$ so $g^p = e$. But since g has order r and $0 \leq p < r$, it follows that $p = 0$. Hence $m - n = qr$ and so $m \equiv n \pmod{r}$. □

Exercise A2.9

Cosets define an equivalence relation between elements. Show that $g_1, g_2 \in G$ are in the same coset of H in G if and only if there exists some $h \in H$ such that $g_2 = g_1 h$.

Solution

Concepts Involved: Equivalence Relations, Cosets

In Exercise A2.2 we showed that the relation $\sim$ on a group G defined by $g_1 \sim g_2$ iff $g_1 = g_2 h$ for some $h \in H$ was an equivalence relation. The equivalence classes of this equivalence relation were $\{gh | h \in H\}$, i.e. precisely the left cosets of H in G. So, g_1, g_2 are in the same coset of H in G if and only if $g_1 = g_2 h$ for some $h \in H$, which is exactly what we wished to show. □

Exercise A2.10

How many cosets of H are there in G?

Solution

Concepts Involved: Equivalence Relations, Cosets

We observe that the map $\varphi_g : H \to [g]$ defined in the solution of Exercise A2.2 is a map from H to a right coset of H in G defined by g. Since we showed that this map was bijective, this shows that $|H| = |Hg|$ for any $g \in G$. Furthermore, since the cosets define an equivalence relation between elements of G, the cosets of H in G partition G. So, we conclude that there are $|G|/|H|$ cosets of H in G, each of cardinality $|H|$. □

Exercise A2.11: Characters

$(\star)$ Prove the properties of characters given above.

Solution

Concepts Involved: Matrix Groups, Character (Trace)
Recall that the character of a matrix group $G \subset M_n$ is a function on the group defined by $\chi(g) = \mathrm{tr}(g)$ where tr is the trace function. It has the properties that (1) $\chi(I) = n$, (2) $|\chi(g)| \leq n$, (3) $|\chi(g)| = n$ implies $g = e^{i\theta} I$, (4) χ is constant on any given conjugacy class of G, (5) $\chi(g^{-1}) = \chi^*(g)$ and (6) $\chi(g)$ is an algebraic number for all g.

The six properties are proven below.

(1) $\chi(I) = \mathrm{tr}(I) = \sum_{k=1}^{n} 1 = n$.

(2) Let $g \in G$. Since G is finite, by Exercise A2.1 it follows that g has order r such that $g^r = I$. So, g may be diagonalized with roots of unity $e^{2\pi i j/r}$, $j \in \{0, 1, \ldots, r-1\}$ on the diagonal. We then

find using the triangle inequality that:

$$\left|\chi(g)\right| = \left|\mathrm{tr}(g)\right| = \left|\sum_{k=1}^{n} e^{2\pi i j_k/r}\right| \leq \sum_{k=1}^{n} \left|e^{2\pi i j_k/r}\right| = \sum_{i} 1 = n$$

which proves the claim.

(3) The (complex) triangle inequality $|z_1 + z_2| \leq |z_1| + |z_2|$ is saturated when $z_1 = kz_2$ for some $k \geq 0$. This can only occur in the above equation when every λ_i in the sum is identical (as distinct roots of unity are not related by a non-negative constant). If the λ_is are identical, then g is diagonal with diagonal entries of unit modulus, so $g = e^{i\theta} I$ as claimed.

(4) Let $G_x = \left\{g^{-1}xg | g \in G\right\}$ be the conjugacy class of x in G. We then have for any $h \in G_x$ that $\chi(h) = \chi(g^{-1}xg) = \mathrm{tr}\left(g^{-1}xg\right) = \mathrm{tr}\left(xgg^{-1}\right) = \mathrm{tr}(xI) = \mathrm{tr}(x)$, using the cyclicity of the trace. We conclude that χ is constant on the conjugacy class.

(5) By the same argument as (2), $g \in G$ can be diagonalized with roots of unity $e^{2\pi ij/r}$ on the diagonal:

$$g = \begin{bmatrix} e^{2\pi i j_1/r} & & & \\ & e^{2\pi i j_2/r} & & \\ & & \ddots & \\ & & & e^{2\pi i j_n/r} \end{bmatrix}$$

It then follows that g^{-1} is:

$$g^{-1} = \begin{bmatrix} e^{-2\pi i j_1/r} & & & \\ & e^{-2\pi i j_2/r} & & \\ & & \ddots & \\ & & & e^{-2\pi i j_n/r} \end{bmatrix}$$

So we have that:

$$\chi(g^{-1}) = \mathrm{tr}\left(g^{-1}\right) = \sum_{j=1}^{n} e^{-2\pi i j_k/r} = \sum_{j=1}^{n} (e^{2\pi i j_k/r})^* = \left(\sum_{j=1}^{n} e^{2\pi i j_k/r}\right)^* = (\mathrm{tr}(g))^* = \chi^*(g).$$

which proves the claim.

(6) $\chi(g)$ is the sum of r-th roots of unity, which are algebraic; hence $\chi(g)$ is algebraic as the sum of algebraic numbers.

□

Exercise A2.12: Unitary matrix groups

($\star\star$) A unitary matrix group is comprised solely of unitary matrices (those who which satisfy $U^\dagger U = I$). Show that every matrix group is equivalent to a unitary matrix group. If a representation of a group consists entirely of unitary matrices, we may refer to it as being a *unitary representation*.

Solution

Concepts Involved: Matrix Groups, Character (Trace), Equivalence, Unitary Operators.
Recall that two groups are equivalent if they are isomorphic (i.e. there is a bijection between the groups that respects the group multiplication) and the isomorphic element have the same character.

Let $G = \{A_1, \dots A_n\}$ be a finite matrix group. Then define

$$A = \sum_{i=1}^{n} A_i^\dagger A_i.$$

By Ex. 2.25 each term of the above sum is positive, and by Ex. 2.24 each term is Hermitian. The sum of Hermitian operators is Hermitian, so A is Hermitian. By Ex. 2.21, A is diagonalizable. Let U be the unitary matrix that diagonalizes A. We then have that D is a diagonal matrix, with:

$$D = UAU^\dagger.$$

Let $D^{1/2}$ be the matrix obtained by taking the square root of the diagonal entries of D. Then define $T = D^{1/2}U$. We then claim that $G_U = \{V_1, \dots V_n\}$ is a unitary matrix group equivalent to G, where:

$$V_i = TA_iT^{-1}.$$

We have three points to verify; (i) That the V_is are unitary, (ii) That $\varphi : G \to G_u$ defined by $\varphi(A_i) = TA_iT^{-1} = V_i$ is an isomorphism, and (iii) that the characters of A_i and V_i are equivalent.

(i) For any V_i, we have:

$$\begin{aligned}
V_i^\dagger V_i &= (TA_iT^{-1})^\dagger (TA_iT^{-1}) \\
&= (D^{1/2}UA_iU^\dagger D^{-1/2})^\dagger (D^{1/2}UA_iU^\dagger D^{-1/2}) \\
&= (D^{-1/2}UA_i^\dagger U^\dagger D^{1/2})(D^{1/2}UA_iU^\dagger D^{-1/2}) \\
&= (D^{-1/2}UA_i^\dagger U^\dagger)D(UA_iU^\dagger D^{-1/2}) \\
&= (D^{-1/2}UA_i^\dagger U^\dagger)(UAU^\dagger)(UA_iU^\dagger D^{-1/2}) \\
&= (D^{-1/2}UA_i^\dagger)A(A_i^\dagger U^\dagger D^{-1/2}) \\
&= (D^{-1/2}UA_i^\dagger)\left(\sum_{j=1}^{n} A_j^\dagger A_j\right)(A_iU^\dagger D^{-1/2}) \\
&= D^{-1/2}U\left(\sum_{j=1}^{n} (A_jA_i)^\dagger (A_jA_i)\right)U^\dagger D^{-1/2} \\
&= D^{-1/2}U\left(\sum_{k=1}^{n} A_k^\dagger A_k\right)U^\dagger D^{-1/2} \\
&= D^{-1/2}UAU^\dagger D^{-1/2} \\
&= D^{-1/2}DD^{-1/2} \\
&= I
\end{aligned}$$

Where in the sixth equality we use the unitarity of U, and in the ninth equality we use that $A_j A_i = A_k$ iterates over all the group elements as A_j iterates over all the group elements. To see that this is the case, it suffices to show that the map $\psi_i : M_n \to M_n$ defined by $\psi_i(A_j) = A_j A_i$ is a bijection. To see that it is injective, suppose that $\psi_i(A_{j_1}) = \psi_i(A_{j_2})$. Then it follows that $A_{j_1} A_i = A_{j_2} A_i$, and multiplying on the left by A_i^{-1} (which exists) we find that $A_{j_1} = A_{j_2}$. To see that it is surjective, suppose that $A_{j'} \in M_n$. Then, there exists $A_{j'} A_i^{-1} \in M_n$ such that $\psi_i(A_{j'} A_i^{-1}) = A_{j'} A_i^{-1} A_i = A_{j'}$. We conclude that ψ_i is bijective.

(ii) Firstly, φ is a homomorphism as for any A_i, A_j we have:

$$\varphi(A_i)\varphi(A_j) = V_i V_j = T A_i T^{-1} T A_j T^{-1} = T A_i A_j T^{-1} = \varphi(A_i A_j).$$

Next, φ is surjective by construction. Finally, it is injective; suppose that $V_i = V_j$. Then we have that:

$$T A_i T^{-1} = T A_j T^{-1}$$

And multiplying both sides on the left by T^{-1} on the left and T on the right we find that $A_i = A_j$. Hence we conclude that φ is a bijective homomorphism and hence an isomorphism.

(iii) This is immediate from the cyclicity of the trace:

$$\chi(V_i) = \mathrm{tr}\Big(T A_i T^{-1}\Big) = \mathrm{tr}\Big(T^{-1} T A_i\Big) = \mathrm{tr}(A_i) = \chi(A_i).$$

The claim is therefore proven. □

Exercise A2.13

($\star$) Show that every irreducible Abelian matrix group is one dimensional.

Solution

Concepts Involved: Matrix Groups, Schur's Lemma

First we prove Schur's second Lemma; if G is an irreducible matrix group $G \subset M_n$ and there exists an $n \times n$ matrix S such that $S g_i = g_i S \ \forall i$, then $S = \lambda I$ for some $\lambda \in \mathbb{C}$.
To see this, consider such an S as above - it has at least one eigenvalue $\lambda \in \mathbb{C}$, wherein $S - \lambda I$ is not invertible as $\det(S - \lambda I) = 0$. However, $S - \lambda I$ still commutes with every $g_i \in G$ (since S, I do). Thus, it must be the case that $S - \lambda I$ is the zero matrix, and therefore $S = \lambda I$.
Now, suppose G is abelian. Then, every $h \in G$ commutes with every other element $g \in G$, so we may apply Schur's second Lemma as proven above to see that $h = \lambda_h I$ for every element of G. But since G is irreducible, it must be one-dimensional (else, for $\dim(G) = n \geq 2$ it would be equivalent to another matrix group H of block diagonal form with $h' = \mathrm{diag}(m_1, m_2) \in H$ with $m_1 = \lambda_h I_{n-k}, m_2 = \lambda_h I_k$ for some $k \geq 1$, and not reducible). □

Exercise A2.14

($\star\star$) Show that if ρ is an irreducible representation of G, then $|G|/d_\rho$ is an integer.

Solution

Concepts Involved: Matrix Groups, Schur's Lemma, Fundamental Theorem of Representations

This is a standard theorem (sometimes called the "Frobenius Divisibility Theorem"), though is difficult to prove with the tools given in the chapter. Proofs can be found in many books discussing group theory/representation theory (e.g. Chapter 19.2 of Dummit and Foote's "Abstract Algebra"). We provide a standard proof here, for reference.

First construct the object:

$$S = \sum_{g \in C} \rho(g)$$

where the sum is over a conjugacy class C of G.

$$S\rho(g) = \rho(g)S$$

for all $g \in G$, as:

$$S\rho(g) = \sum_{h \in C} \rho(h)\rho(g) = \sum_{h \in C} \rho(ghg^{-1})\rho(g) = \sum_{h \in C} \rho(ghg^{-1}g) = \sum_{h \in C} \rho(g)\rho(h) = \rho(g)S$$

where we use that conjugacy classes are invariant under conjugation in the second equality, so the sum is left invariant.

By Schur's second lemma (as proven in Ex. A2.13), $S = \lambda I$ for some $\lambda \in \mathbb{C}$ (with I the identity matrix of size $d_\rho \times d_\rho$). Taking its trace, we find:

$$\mathrm{Tr}(S) = \sum_{g \in C(G)} \mathrm{Tr}\big(\rho(g)\big) \implies \lambda d_\rho = \sum_{g \in C} \chi^\rho(g)$$

by Ex. A2.11 characters are invariant over a conjugacy class, so each of the $|C|$ terms in the sum contributes $\chi^\rho(C)$, yielding:

$$\lambda = \frac{|C|}{d_\rho}\chi^\rho(C)$$

and thus:

$$S = \frac{|C|}{d_\rho}\chi^\rho(C)I$$

and note we can associate a S for each conjugacy class C_i, which we denote S_i.

Now, let us evaluate:

$$Q = \sum_i S_i(\chi^\rho(C_i))^*$$

If we evaluate this using our obtained expression for S_i, we find:

$$Q = \sum_i \frac{|C_i|}{d_\rho} \chi^\rho(C)(\chi^\rho(C))^* = \frac{1}{d_\rho} \sum_i |C_i|(\chi_i^\rho)^* \chi_i^\rho = \frac{|G|}{d_\rho}$$

where in the last equality we have used the orthogonality of characters proven in Ex. A2.15.
Any finite matrix group is equivalent to a unitary matrix group, so WLOG take the representation ρ to be finite. Then the eigenvalues λ of $\rho(g)$ are complex phases (Ex. 2.18) - since we consider finite G for all $g \in G$ we have $g^n = e$ for some n and therefore that $\rho(g)^n = \rho(g^n) = \rho(e) = I$, implying that $\lambda^n = 1$ and so the eigenvalues of $\rho(g)$ are roots of unity. Thus, the characters $\chi^\rho(g)$ (the sum of all eigenvalues of $\rho(g)$) we consider are sums of roots of unity and therefore algebraic integers. It then follows that each of the S_i and that Q is an algebraic integer (as sums/products of algebraic integers are again algebraic integers). Since Q is also rational is also rational (it is a ratio of two integers $\frac{|G|}{d_\rho}$), this means that Q is an integer. □

Exercise A2.15

($\star$) Using the Fundamental Theorem, prove that characters are orthogonal, that is:

$$\sum_{i=1}^{r} r_i (\chi_i^p)^* \chi_i^q = |G|\delta_{pq} \text{ and } \sum_{p=1}^{r} (\chi_i^p)^* \chi_j^p = \frac{|G|}{r_i}\delta_{ij}$$

where p, q, and δ_{pq} have the same meaning as in the theorem and χ_i^p is the value the character of the pth irreducible representation takes on the ith conjugacy class of G and r_i is the size of the ith conjugacy class of G and r_i is the size of the ith conjugacy class.

Solution

Concepts Involved: Fundamental Theorem of Representations, Characters

To prove the first equality, take $i = j$ and $k = l$ in the fundamental theorem to obtain:

$$\sum_{g \in G} [\rho^p(g)]_{ii}^{-1} [\rho^q(g)]_{kk} = \frac{|G|}{d_\rho} \delta_{ik}\delta_{ik}\delta_{pq}$$

where repeated indices are summed over. On the LHS we just have the traces of the irrep matrices (i.e. the characters), and on the RHS we note that $\delta_{ik}\delta_{ik} = \delta_{ii} = d_\rho$, so the above equation becomes:

$$\sum_{g \in G} (\chi^p(g))^* \chi^q(g) = |G|\delta_{pq}.$$

Note that the first term in the sum is obtained by the observation of $\mathrm{Tr}\big(\rho^{-1}(g)\big) = \mathrm{Tr}\big(\rho(g^{-1})\big) = \chi^*(g)$ (Using Ex. A2.11(5)). Then, using that characters are constant on conjugacy classes (Ex. A2.11(4)), the terms in the LHS of the above sum can be grouped via conjugacy class to obtain:

$$\sum_{i=1}^{r} r_i (\chi_i^p)^* \chi_i^q = |G|\delta_{pq}$$

which is the first character orthogonality relation.
To prove the second character relation, consider the matrix defined by:

$$U_{ip} = \sqrt{\frac{r_i}{|G|}}\chi_i^p$$

We then observe that the first orthogonality relation (rewritten):

$$\sum_{i=1}^{r}\left(\sqrt{\frac{r_i}{|G|}}\chi_i^p\right)^*\left(\sqrt{\frac{r_i}{|G|}}\chi_i^q\right) = \delta_{pq}$$

is the statement that the columns of U are orthonormal w.r.t. the complex inner product, or that $U^\dagger U = I$. But this means that U is unitary, satisfying $UU^\dagger = I$, which means that the rows of U are orthogonal, i.e.:

$$\sum_{p=1}^{r}\left(\sqrt{\frac{r_i}{|G|}}\chi_i^p\right)^*\left(\sqrt{\frac{r_i}{|G|}}\chi_j^p\right) = \delta_{ij}$$

which we can rearrange to get the second orthogonality relation:

$$\sum_{p=1}^{r}(\chi_i^p)^*\chi_j^p = \frac{|G|}{r_i}\delta_{ij}.$$

□

Exercise A2.16

($\star$) S_3 is the group of permutations of three elements. Suppose we order these as mapping 123 to: 123;231;312;213;132, and 321, respectively. Show that there exist two one-dimensional irreducible representations of S_3, one of which is trivial, and the other of which is 1,1,1,-1,-1,-1, corresponding in order to the six permutations given earlier. Also show that there exists a two dimensional irreducible representation, with the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix},$$
$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix}$$

Verify that the representations are orthogonal.

There is a slight error in the exercise, where the order of the last two matrices in the 2D irrep should be switched, with the correct order as:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix},$$
$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}$$

Solution

Concepts Involved: Fundamental Theorem of Representations

To verify that these are indeed valid representations, we check that they respect the group multiplication; the multiplication table of S_3 looks like:

$\cdot$	**123**	**231**	**312**	**213**	**132**	**321**
123	123	231	312	213	132	321
231	231	312	123	321	213	132
312	312	123	231	132	321	213
213	213	132	321	123	231	312
132	132	321	213	312	123	231
321	321	213	132	231	312	123

where we have $g' = g_{\text{row}} \cdot g_{\text{col}}$ (the column permutation is applied first). We can then verify that the representations respect this group multiplication structure. The trivial representation (mapping each element to 1) satisfies it trivially. That the sign representation respects the group multiplication structure is easily seen from the fact that $123, 231, 312$ are even permutations while $213, 132, 321$ are odd permutations, which are mapped to ± 1 respectively, which preserves the $(\text{even})^2 = (\text{odd})^2 = \text{even}$ and $\text{even} \cdot \text{odd} = \text{odd} \cdot \text{even} = \text{odd}$ multiplication structure of permutations.

For the proposed two-dimensional representation, let us label the matrices as I, R, R^2, S, SR, SR^2 in order.

We can then compute via matrix multiplication:

$$I \cdot A = A \cdot I = A$$

$$R \cdot R = \left(\frac{1}{2}\begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}\right) \cdot \left(\frac{1}{2}\begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix} = R^2$$

$$R^2 \cdot R = \left(\frac{1}{2}\begin{bmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix}\right) \cdot \left(\frac{1}{2}\begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$$S \cdot R = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \left(\frac{1}{2}\begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix} = SR$$

$$SR \cdot R = \left(\frac{1}{2}\begin{bmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & -1 \end{bmatrix}\right) \cdot \left(\frac{1}{2}\begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} = SR^2$$

$$R \cdot S = \left(\frac{1}{2}\begin{bmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}\right) \cdot \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} = SR^2$$

and using matrix multiplication associativity for the remainder of the relations, we find the table:

$\cdot$	I	R	R^2	S	SR	SR^2
I	I	R	R^2	S	SR	SR^2
R	R	R^2	I	SR^2	S	SR
R^2	R^2	I	R	SR	SR^2	S
S	S	SR	SR^2	I	R	R^2
SR	SR	SR^2	S	R^2	I	R
SR^2	SR^2	S	SR	R	R^2	I

where again $M = M_{\text{row}} \cdot M_{\text{col}}$. We see that the group multiplication law is obeyed, so this is indeed also a representation.
To check that the representations are irreducible, we can use Theorem A2.3 and verify via the characters; for the trivial irrep:

$$\frac{1}{|G|}\sum_{g\in G}\left|\chi(g)\right|^2 = \frac{1}{6}(6 \cdot |1|^2) = 1$$

for the sign irrep:

$$\frac{1}{|G|}\sum_{g\in G}\left|\chi(g)\right|^2=\frac{1}{6}(3\cdot|1|^2+|1|^2+3\cdot|-1|^2)=1$$

where the characters are equivalent to the representations in the 1D case. For the 2d irrep:

$$\begin{aligned}\frac{1}{|G|}\sum_{g\in G}\left|\chi(g)\right|^2&=\frac{1}{6}\left(\left|\mathrm{Tr}\left(\begin{bmatrix}1&0\\0&1\end{bmatrix}\right)\right|^2+\left|\mathrm{Tr}\left(\frac{1}{2}\begin{bmatrix}-1&-\sqrt{3}\\\sqrt{3}&-1\end{bmatrix}\right)\right|^2+\left|\mathrm{Tr}\left(\frac{1}{2}\begin{bmatrix}-1&\sqrt{3}\\-\sqrt{3}&-1\end{bmatrix}\right)\right|^2\right.\\&\left.+\left|\mathrm{Tr}\left(\begin{bmatrix}-1&0\\0&1\end{bmatrix}\right)\right|^2+\left|\mathrm{Tr}\left(\frac{1}{2}\begin{bmatrix}1&-\sqrt{3}\\-\sqrt{3}&-1\end{bmatrix}\right)\right|^2+\left|\mathrm{Tr}\left(\frac{1}{2}\begin{bmatrix}1&\sqrt{3}\\\sqrt{3}&-1\end{bmatrix}\right)\right|^2\right)\\&=\frac{1}{6}\left(|2|^2+|-1|^2+|-1|^2+|0|^2+|0|^2+|0|^2\right)\\&=1\end{aligned}$$

hence all three representations are indeed irreducible. That the representations are inequivalent can be seen from the fact that $\mathrm{Tr}\big(\chi^\rho(213)\big)$ is unequal across the three reprensentations. S_3 has three conjugacy classes $C_1=\{123\},C_2=\{231,312\},C_3=\{213,132,321\}$ and hence these three exhaust the possible irreps, by the fundamental theorem. We now verify in accordance with the theorem that the representations are orthogonal, i.e. they satisfy:

$$\sum_{g\in G}[\rho^p(g)]^{-1}_{ij}[\rho^q(g)]_{kl}=\frac{|G|}{d_\rho}\delta_{il}\delta_{jk}\delta_{pq}.$$

First checking the orthogonality between the inequivalent irreps:

$$\sum_{g\in G}[\rho^{\mathsf{triv}}(g)]^{-1}[\rho^{\mathsf{sign}}(g)]=3\cdot[1]^{-1}[1]+3\cdot[1]^{-1}[-1]=0$$

$$\begin{aligned}\sum_{g\in G}[\rho^{\mathsf{triv}}(g)]^{-1}[\rho^{\mathsf{2D}}(g)]&=\sum_{g\in G}[\rho^{\mathsf{2D}}(g)]\\&=\left(\begin{bmatrix}1&0\\0&1\end{bmatrix}+\frac{1}{2}\begin{bmatrix}-1&-\sqrt{3}\\\sqrt{3}&-1\end{bmatrix}+\frac{1}{2}\begin{bmatrix}-1&\sqrt{3}\\-\sqrt{3}&-1\end{bmatrix}\right.\\&\left.+\begin{bmatrix}-1&0\\0&1\end{bmatrix}+\frac{1}{2}\begin{bmatrix}1&-\sqrt{3}\\-\sqrt{3}&-1\end{bmatrix}+\frac{1}{2}\begin{bmatrix}1&\sqrt{3}\\\sqrt{3}&-1\end{bmatrix}\right)\\&=\begin{bmatrix}1-\frac{1}{2}-\frac{1}{2}-1+\frac{1}{2}+\frac{1}{2}&0-\frac{\sqrt{3}}{2}+\frac{\sqrt{3}}{2}+0-\frac{\sqrt{3}}{2}+\frac{\sqrt{3}}{2}\\0+\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}+0-\frac{\sqrt{3}}{2}+\frac{\sqrt{3}}{2}&1-\frac{1}{2}-\frac{1}{2}+1-\frac{1}{2}-\frac{1}{2}\end{bmatrix}\\&=\begin{bmatrix}0&0\\0&0\end{bmatrix}\end{aligned}$$

$$\begin{aligned}\sum_{g\in G}[\rho^{\text{sign}}(g)]^{-1}[\rho^{\text{2D}}(g)] &= \left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \frac{1}{2}\begin{bmatrix}-1 & -\sqrt{3}\\ \sqrt{3} & -1\end{bmatrix} + \frac{1}{2}\begin{bmatrix}-1 & \sqrt{3}\\ -\sqrt{3} & -1\end{bmatrix}\right.\\ &\quad \left. + -\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1 & -\sqrt{3}\\ -\sqrt{3} & -1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1 & \sqrt{3}\\ \sqrt{3} & -1\end{bmatrix}\right)\\ &= \begin{bmatrix}1-\frac{1}{2}-\frac{1}{2}+1-\frac{1}{2}-\frac{1}{2} & 0-\frac{\sqrt{3}}{2}+\frac{\sqrt{3}}{2}-0+\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}\\ 0+\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}-0+\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2} & 1-\frac{1}{2}-\frac{1}{2}-1+\frac{1}{2}+\frac{1}{2}\end{bmatrix}\\ &= \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}\end{aligned}$$

That the orthogonality relation also holds within each representation can be readily checked. For the 1D irreps:

$$\sum_{g\in G}[\rho^{\text{triv}}(g)]^{-1}[\rho^{\text{triv}}(g)] = 6\cdot[1]^{-1}[1] = 6 = \frac{6}{1}$$

$$\sum_{g\in G}[\rho^{\text{sign}}(g)]^{-1}[\rho^{\text{sign}}(g)] = 3\cdot[1]^{-1}[1] + 3\cdot[-1]^{-1}[-1] = 6 = \frac{6}{1}$$

For the 2D irrep, note that $R^{-1} = R^2$ and $S^{-1} = S$, so:

$$\sum_{g\in G}[\rho^{2D}(g)]^{-1}_{ij}[\rho^{2D}(g)]_{kl} = I_{ij}I_{kl} + (R^2)_{ij}R_{kl} + R_{ij}(R^2)_{kl} + S_{ij}S_{kl} + (SR)_{ij}(SR)_{kl} + (SR^2)_{ij}(SR^2)_{kl}$$

note that this expression is symmetric under $ij \leftrightarrow kl$. We then check:

$$\begin{aligned}\sum_{g\in G}[\rho^{2D}(g)]^{-1}_{00}[\rho^{2D}(g)]_{00} &= (1)(1) + (-\frac{1}{2})(-\frac{1}{2}) + (-\frac{1}{2})(-\frac{1}{2}) + (-1)(-1) + (\frac{1}{2})(\frac{1}{2}) + (\frac{1}{2})(\frac{1}{2})\\ &= 3 = \frac{6}{2}\end{aligned}$$

$$\begin{aligned}\sum_{g\in G}[\rho^{2D}(g)]^{-1}_{00}[\rho^{2D}(g)]_{01} &= (1)(0) + (-\frac{1}{2})(-\frac{\sqrt{3}}{2}) + (-\frac{1}{2})(\frac{\sqrt{3}}{2}) + (-1)(0) + (\frac{1}{2})(\frac{\sqrt{3}}{2}) + (\frac{1}{2})(-\frac{\sqrt{3}}{2})\\ &= 0\end{aligned}$$

$$\begin{aligned}\sum_{g\in G}[\rho^{2D}(g)]^{-1}_{00}[\rho^{2D}(g)]_{10} &= (1)(0) + (-\frac{1}{2})(\frac{\sqrt{3}}{2}) + (-\frac{1}{2})(-\frac{\sqrt{3}}{2}) + (-1)(0) + (\frac{1}{2})(\frac{\sqrt{3}}{2}) + (\frac{1}{2})(-\frac{\sqrt{3}}{2})\\ &= 0\end{aligned}$$

$$\sum_{g\in G}[\rho^{2D}(g)]^{-1}_{00}[\rho^{2D}(g)]_{11} = (1)(1)+(-\frac{1}{2})(-\frac{1}{2})+(-\frac{1}{2})(-\frac{1}{2})+(-1)(1)+(\frac{1}{2})(-\frac{1}{2})+(\frac{1}{2})(-\frac{1}{2})$$
$$= 0$$

$$\sum_{g\in G}[\rho^{2D}(g)]^{-1}_{01}[\rho^{2D}(g)]_{01} = (0)(0)+(-\frac{\sqrt{3}}{2})(-\frac{\sqrt{3}}{2})+(\frac{\sqrt{3}}{2})(\frac{\sqrt{3}}{2})+(0)(0)+(\frac{\sqrt{3}}{2})(\frac{\sqrt{3}}{2})+(-\frac{\sqrt{3}}{2})(-\frac{\sqrt{3}}{2})$$
$$= 3 = \frac{6}{2}$$

$$\sum_{g\in G}[\rho^{2D}(g)]^{-1}_{01}[\rho^{2D}(g)]_{10} = (0)(0)+(-\frac{\sqrt{3}}{2})(\frac{\sqrt{3}}{2})+(\frac{\sqrt{3}}{2})(-\frac{\sqrt{3}}{2})+(0)(0)+(\frac{\sqrt{3}}{2})(\frac{\sqrt{3}}{2})+(-\frac{\sqrt{3}}{2})(-\frac{\sqrt{3}}{2})$$
$$= 0$$

$$\sum_{g\in G}[\rho^{2D}(g)]^{-1}_{01}[\rho^{2D}(g)]_{11} = (0)(1)+(-\frac{\sqrt{3}}{2})(-\frac{1}{2})+(\frac{\sqrt{3}}{2})(-\frac{1}{2})+(0)(1)+(\frac{\sqrt{3}}{2})(-\frac{1}{2})+(-\frac{\sqrt{3}}{2})(-\frac{1}{2})$$
$$= 0$$

$$\sum_{g\in G}[\rho^{2D}(g)]^{-1}_{10}[\rho^{2D}(g)]_{10} = (0)(0)+(-\frac{\sqrt{3}}{2})(\frac{\sqrt{3}}{2})+(\frac{\sqrt{3}}{2})(-\frac{\sqrt{3}}{2})+(0)(0)+(\frac{\sqrt{3}}{2})(\frac{\sqrt{3}}{2})+(-\frac{\sqrt{3}}{2})(-\frac{\sqrt{3}}{2})$$
$$= 0$$

$$\sum_{g\in G}[\rho^{2D}(g)]^{-1}_{10}[\rho^{2D}(g)]_{11} = (0)(1)+(-\frac{\sqrt{3}}{2})(-\frac{1}{2})+(\frac{\sqrt{3}}{2})(-\frac{1}{2})+(0)(1)+(\frac{\sqrt{3}}{2})(-\frac{1}{2})+(-\frac{\sqrt{3}}{2})(-\frac{1}{2})$$
$$= 0$$

$$\sum_{g\in G}[\rho^{2D}(g)]^{-1}_{11}[\rho^{2D}(g)]_{11} = (1)(1)+(-\frac{1}{2})(-\frac{1}{2})+(-\frac{1}{2})(-\frac{1}{2})+(1)(1)+(-\frac{1}{2})(-\frac{1}{2})+(-\frac{1}{2})(-\frac{1}{2})$$
$$= 3 = \frac{6}{2}$$

So after this exhaustive computation, we have checked that the orthogonality relation holds (for all possible choices of irreps and matrix indices). □

Remark: The labelling of the matrices of the 2D irrep of S_3 make it clear that S_3 is isomorphic to D_3, the symmetries of the triangle (wherein we can interpret R is the 2D rotation matrix with angle $\pi/3$ and S the 2D reflection matrix about the x axis).

Exercise A2.17

Prove that the regular representation is faithful.

Solution

Concepts Involved: Regular Representations, Faithful Representations

The permutation matrices in the regular representation are specified by the group multiplication table, with (labelling the group elements as $\{g_i\}_{i=1}^{N}$):

$$[\rho^R(g_k)]_{ij} = \begin{cases} 1 & \text{if } g_k g_j = g_i \\ 0 & \text{otherwise} \end{cases}$$

then if $g_k g_l = g_m$, we have:

$$[\rho^R(g_m)]_{ij} = \begin{cases} 1 & \text{if } g_m g_j = g_i \\ 0 & \text{otherwise} \end{cases}$$

and:

$$\begin{aligned} [\rho^R(g_k)\rho^R(g_l)]_{ij} &= [\rho^R(g_k)]_{ia}[\rho^R(g_l)]_{aj} \\ &= \begin{cases} 1 & \text{if } g_k g_a = g_i \text{ and } g_l g_j = g_a \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 1 & \text{if } g_k g_l g_j = g_i \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 1 & \text{if } g_m g_j = g_i \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

thus the permutation matrices respect the group multiplication law of G, and ρ^R is indeed a homomorphism. Further, as the rows of the group multiplication table are unique to each element, there is a one-to-one mapping of group elements to their associated permutation matrices, and ρ^R is an isomorphism. Thus the regular representation is faithful. □

Exercise A2.18

Show that the character of the regular representation is zero except on the representation of the identity element, for which $\chi(I) = |G|$.

Solution

Concepts Involved: Regular Representations, Characters

As discussed in the previous exercise, we have:

$$[\rho^R(g_k)]_{ij} = \begin{cases} 1 & \text{if } g_k g_j = g_i \\ 0 & \text{otherwise} \end{cases}$$

Noting that $g_k g_i = g_i$ if and only if $g_k = I$, we thus have:

$$\chi^R(g_k) = [\rho^R(g_k)]_{ii} = \begin{cases} |G| & g_k = I \\ 0 & \text{otherwise} \end{cases}$$

as claimed. □

Exercise A2.19

Use Theorem A2.5 to show that the regular representation contains d_{ρ^p} instances of each irreducible representation ρ^p. Thus, if R denotes the regular representation, and $\hat{G}$ denotes the set of all inequivalent irreducible representations, then:

$$\chi_i^R = \sum_{\rho \in G} d_\rho \chi_i^\rho$$

Solution

Concepts Involved: Regular Representation, Characters

Take ρ to be the regular representation R in Theorem A2.5, so then:

$$c_p = \frac{1}{|G|} \sum_{i=1}^{r} r_i (\chi_i^p)^* \chi_i^R$$

χ_i^R vanishes on every conjugacy class i of G except that containing the identity (with $r_e = 1$), for which it evaluates to $\chi_i^R(e) = |G|$ (Ex. A2.18), so the above sum reduces to:

$$c_p = \frac{1}{|G|} \sum_{i=1}^{r} (\chi_i^p)^* |G| \delta_{ie} = (\chi_e^p)^* = d_{\rho^p}$$

where we use that the character of the identity for a representation ρ^p is just d_{ρ^p}. This proves the first claim, and thus we have:

$$\rho^R = \bigoplus_p d_{\rho^p} \rho^p$$

Evaluating the representations in the above equation on a particular conjugacy class i of G and taking the trace of both sides, we obtain:

$$\chi_i^R = \sum_{\rho \in G} d_\rho \chi_i^\rho$$

as claimed. □

Exercise A2.20

The character of the regular representation is zero except for the conjugacy class i containing e, the identity element in G. Show, therefore, that

$$\sum_{\rho \in G} d_\rho \chi^\rho(g) = N\delta_{ge}.$$

Solution

Concepts Involved: Regular Representations, Characters

Clearly:

$$\sum_{\rho \in \hat{G}} d_\rho \chi^\rho(g) = \chi_i^R(g) = N\delta_{ge}$$

where in the first equality we use the result of Ex. A2.19 and in the second equality we again use the fact that the character of the regular representation is zero except for the conjugacy class containing e (Ex. A2.18), which is a singleton, and for which $\chi_i^R(e) = N$ the order of the group. □

Exercise A2.21

Show that $\sum_{\rho \in \hat{G}} d_\rho^2 = |G|$.

Solution

Concepts Involved: Regular Representations, Characters

First, note that $d_\rho = \chi^\rho(e)$. Then:

$$\sum_{\rho \in \hat{G}} d_\rho^2 = \sum_{\rho \in \hat{G}} d_\rho \chi^\rho(e) = N\delta_{ee} = N = |G|$$

where in the second equality we have invoked the result of Ex. A2.20. □

Exercise A2.22

$(\star)$ Substitute (A2.10) into (A2.9) and prove that $\hat{f}(\rho)$ is obtained.

Solution

Concepts Involved: Group Fourier Transform, Fundamental Theorem of Representations

Taking (A2.10):

$$f(g) = \frac{1}{\sqrt{N}} \sum_{\rho \in \hat{G}} \sqrt{d_\rho} \, \mathrm{Tr}\Big(\hat{f}(\rho)\rho(g^{-1})\Big)$$

and substituting it into the RHS of (A2.9):

$$\hat{f}(\rho) = \sqrt{\frac{d_\rho}{N}} \sum_{g \in G} f(g)\rho(g)$$

we obtain:

$$\begin{aligned}\sqrt{\frac{d_{\rho^p}}{N}} \sum_{g \in G} f(g)\rho^p(g) &= \sqrt{\frac{d_{\rho^p}}{N}} \sum_{g \in G} \left(\frac{1}{\sqrt{N}} \sum_{\rho^q \in \hat{G}} \sqrt{d_{\rho^q}} \, \mathrm{Tr}\Big(\hat{f}(\rho^q)\rho^q(g^{-1})\Big) \right) \rho^p(g) \\ &= \frac{1}{N} \sum_{\rho^q \in \hat{G}} \sqrt{d_{\rho^p} d_{\rho^q}} \sum_{g \in G} \mathrm{Tr}\Big(\hat{f}(\rho^q)\rho^q(g^{-1})\Big)\rho^p(g)\end{aligned}$$

Writing the trace out in terms of the matrix indices, we have:

$$\mathrm{Tr}\Big(\hat{f}(\rho^q)\rho^q(g^{-1})\Big) = [\hat{f}(\rho^q)\rho^q(g^{-1})]_{jj} = [\hat{f}(\rho^q)]_{ji}[\rho^q(g^{-1})]_{ij} = [\hat{f}(\rho^q)]_{ji}[\rho^q(g)]^{-1}_{ij}$$

with repeated indices summed over. If we write our earlier equation out in terms of the matrix indices:

$$\begin{aligned}\left[\sqrt{\frac{d_{\rho^p}}{N}} \sum_{g \in G} f(g)\rho^p(g)\right]_{kl} &= \frac{1}{N} \sum_{\rho^q \in \hat{G}} \sqrt{d_{\rho^p} d_{\rho^q}} \sum_{g \in G} [\hat{f}(\rho^q)]_{ji}[\rho^q(g)]^{-1}_{ij}[\rho^p(g)]_{kl} \\ &= \frac{1}{N} \sum_{\rho^q \in \hat{G}} \sqrt{d_{\rho^p} d_{\rho^q}} [\hat{f}(\rho^q)]_{ji} \sum_{g \in G} [\rho^q(g)]^{-1}_{ij}[\rho^p(g)]_{kl}\end{aligned}$$

now we can apply the Fundamental Theorem of Representations to obtain:

$$\begin{aligned}\left[\sqrt{\frac{d_{\rho^p}}{N}} \sum_{g \in G} f(g)\rho^p(g)\right]_{kl} &= \frac{1}{N} \sum_{\rho^q \in \hat{G}} \sqrt{d_{\rho^p} d_{\rho^q}} [\hat{f}(\rho^q)]_{ji} \frac{N}{d_{\rho^p}} \delta_{il}\delta_{jk}\delta_{pq} \\ &= \frac{1}{N} \sqrt{d_{\rho^p} d_{\rho^p}} [\hat{f}(\rho^q)]_{kl} \frac{N}{d_{\rho^p}} \\ &= [\hat{f}(\rho^q)]_{kl}\end{aligned}$$

and thus the claim is proven. □

Exercise A2.23

Let us represent an Abelian group G by $g \in [0, N-1]$, with addition as the group operation, and define $\rho_h(g) = \exp\big[-2\pi igh/N\big]$ at the h representation of g. This representation is one-dimensional, so $d_\rho = 1$. Show that the Fourier transform relations for G are

$$\hat{f}(h) = \frac{1}{\sqrt{N}} \sum_{g=0}^{N-1} f(g) e^{-2\pi igh/N} \text{ and } f(h) = \frac{1}{\sqrt{N}} \sum_{g=0}^{N-1} \hat{f}(g) e^{2\pi igh/N}$$

Solution

Concepts Involved: Group Fourier Transform

Note in this case that we have N h-irreps, labelled by $h \in [0, N-1]$, all of which can be seen to obey the homomorphism property, and all satisfying:

$$\frac{1}{|N|} \sum_{g \in G} \big|\chi_h(g)\big|^2 = \frac{1}{|N|} \sum_{g \in G} \Big|\exp\big(-2\pi ihg/N\big)\Big| = \frac{1}{|N|} \sum_{g \in G} 1 = 1$$

so indeed all are irreducible.
Substituting in $d_\rho = 1$, $\rho_h(g) = \exp\big(-2\pi igh/N\big)$, and $g \in [0, \ldots N-1]$ into the group Fourier transform of Eq. (A2.9), we immediately obtain:

$$\hat{f}(h) = \frac{1}{\sqrt{N}} \sum_{g \in G} f(g) e^{-2\pi igh/N} = \frac{1}{\sqrt{N}} \sum_{g=0}^{N-1} f(g) e^{-2\pi igh/N}$$

Next, note that $\rho_h(g^{-1}) = (\rho_h(g))^{-1} = \exp\big(2\pi igh/N\big)$, so substituting this and $d_\rho = 1$ into the inverse Fourier transform of Eq. (A2.10) we find:

$$f(g) = \frac{1}{\sqrt{N}} \sum_{h \in \hat{G}} \mathrm{Tr}\Big(\hat{f}(h) e^{2\pi ihg/N}\Big) = \frac{1}{\sqrt{N}} \sum_{h=0}^{N-1} \hat{f}(h) e^{2\pi ihg/N}$$

where $\hat{f}(h)$ is just a scalar function in this case and so the trace can be ommitted (trace of a scalar is a scalar). The group element labels/irrep labels can be freely exchanged in this case (noting that $\rho_h(g) = \rho_g(h)$) so we are free to rewrite the above as:

$$f(h) = \frac{1}{\sqrt{N}} \sum_{g=0}^{N-1} \hat{f}(g) e^{2\pi ihg/N}.$$

□

Exercise A2.24

Using the results of Exercise A2.16, construct the Fourier transform over S_3 and express it as a 6x6 unitary matrix.

Solution

Concepts Involved: Group Fourier Transform

We recall the definition of the group Fourier transform (Eq. (A2.9)):

$$\hat{f}(\rho) = \sqrt{\frac{d_\rho}{N}} \sum_{g \in G} f(g)\rho(g)$$

For S_3, we can write $f(g), \hat{f}(\rho)$ as length $|S_3| = 6$ vectors:

$$f(g) \cong \begin{bmatrix} f(123) \\ f(231) \\ f(312) \\ f(213) \\ f(132) \\ f(321) \end{bmatrix}, \quad \hat{f}(\rho) = \begin{bmatrix} f(\rho^{\mathsf{triv}}) \\ f(\rho^{\mathsf{sign}}) \\ f(\rho^{\mathsf{2D}}_{00}) \\ f(\rho^{\mathsf{2D}}_{01}) \\ f(\rho^{\mathsf{2D}}_{10}) \\ f(\rho^{\mathsf{2D}}_{11}) \end{bmatrix}$$

and use the irrep matrices as determined/checked in Ex. A2.16 to construct the unitary matrix U_{S_3} - note that the matrices appearing there are all unitary and so shall be the matrix associated to the fourier transform. Thus reading off the matrix elements from the irrep matrices, we find:

$$U_{S_3} \cong \begin{bmatrix} \sqrt{\frac{d_{\mathsf{t}}}{6}}\rho^{\mathsf{t}}(123) & \sqrt{\frac{d_{\mathsf{t}}}{6}}\rho^{\mathsf{t}}(231) & \sqrt{\frac{d_{\mathsf{t}}}{6}}\rho^{\mathsf{t}}(312) & \sqrt{\frac{d_{\mathsf{t}}}{6}}\rho^{\mathsf{t}}(213) & \sqrt{\frac{d_{\mathsf{t}}}{6}}\rho^{\mathsf{t}}(132) & \sqrt{\frac{d_{\mathsf{t}}}{6}}\rho^{\mathsf{t}}(321) \\ \sqrt{\frac{d_{\mathsf{s}}}{6}}\rho^{\mathsf{s}}(123) & \sqrt{\frac{d_{\mathsf{s}}}{6}}\rho^{\mathsf{s}}(231) & \sqrt{\frac{d_{\mathsf{s}}}{6}}\rho^{\mathsf{s}}(312) & \sqrt{\frac{d_{\mathsf{s}}}{6}}\rho^{\mathsf{s}}(213) & \sqrt{\frac{d_{\mathsf{s}}}{6}}\rho^{\mathsf{s}}(132) & \sqrt{\frac{d_{\mathsf{s}}}{6}}\rho^{\mathsf{s}}(321) \\ \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{00}(123) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{00}(231) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{00}(312) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{00}(213) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{00}(132) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{00}(321) \\ \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{01}(123) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{01}(231) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{01}(312) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{01}(213) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{01}(132) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{01}(321) \\ \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{10}(123) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{10}(231) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{10}(312) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{10}(213) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{10}(132) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{10}(321) \\ \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{11}(123) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{11}(231) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{11}(312) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{11}(213) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{11}(132) & \sqrt{\frac{d_{\mathsf{2D}}}{6}}\rho^{\mathsf{2D}}_{11}(321) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2\sqrt{3}} \end{bmatrix}$$

□

A3 The Solovay-Kitaev theorem

Exercise A3.1

($\star$) In Chapter 4 we made use of the distance measure $E(U,V) = \max_{|\psi\rangle} \left\|(U-V)|\psi\rangle\right\|$, where the maximum is over all pure states $|\psi\rangle$. Show that when U and V are single qubit rotations, $U = R_{\hat{\mathbf{m}}}(\theta), V = R_{\hat{\mathbf{n}}}(\varphi)$, $D(U,V) = 2E(U,V)$, and thus it does not matter whether we use the trace distance or the measure $E(\cdot,\cdot)$ for the Solovay-Kitaev theorem.

Solution

Concepts Involved: Trace Distance, Rotations

First, note that both the measure $E(U,V)$ and the trace distance $D(U,V)$ are invariant under unitaries on both arguments, i.e.:

$$E(U,V) = \max_{|\psi\rangle} \left\|(U-V)|\psi\rangle\right\| = \max_{|\psi\rangle} \left\|U^\dagger(U-V)|\psi\rangle\right\| = \max_{|\psi\rangle} \left\|(I-U^\dagger V)|\psi\rangle\right\| = E(I,U^\dagger V)$$

and:

$$\begin{aligned} D(U,V) = \mathrm{Tr}\big(|U-V|\big) = \mathrm{Tr}\left(\sqrt{(U-V)^\dagger(U-V)}\right) &= \mathrm{Tr}\left(\sqrt{(U-V)^\dagger UU^\dagger(U-V)}\right) \\ &= \mathrm{Tr}\left(\left|U^\dagger(U-V)\right|\right) \\ &= \mathrm{Tr}\left(\left|I-U^\dagger V\right|\right) \\ &= D(I,U^\dagger V) \end{aligned}$$

Further, note that for single-qubit rotations that:

$$W = U^\dagger V = R_{\hat{\mathbf{m}}}(-\theta)R_{\hat{\mathbf{n}}}(\varphi) = R_{\hat{\mathbf{k}}}(\beta) = e^{-i\hat{\mathbf{k}}\cdot\boldsymbol{\sigma}\beta/2}$$

where the axis $\hat{\mathbf{k}}$ and the angle β are obtained from the rotation composition formulas derived in Ex. 4.15. Note also then that:

$$W^\dagger + W = e^{i\hat{\mathbf{k}}\cdot\boldsymbol{\sigma}\beta/2} + e^{-i\hat{\mathbf{k}}\cdot\boldsymbol{\sigma}\beta/2} = 2\cos\left(\frac{\beta}{2}\right)I$$

With this, we can calculate:

$$\begin{aligned}
E(U,V) &= E(I,W) \\
&= \max_{|\psi\rangle} \sqrt{\langle\psi| (I-W)^\dagger (I-W) |\psi\rangle} \\
&= \max_{|\psi\rangle} \sqrt{\langle\psi| (2I - W + W^\dagger) |\psi\rangle} \\
&= \max_{|\psi\rangle} \sqrt{2 - 2\cos\left(\frac{\beta}{2}\right) \langle\psi| I |\psi\rangle} \\
&= \max_{|\psi\rangle} \sqrt{2 - 2\cos\left(\frac{\beta}{2}\right)} \\
&= \sqrt{2 - 2\cos\left(\frac{\beta}{2}\right)} \\
&= \sqrt{4\sin^2(\frac{\beta}{4})} \\
&= 2\sin\left(\frac{\beta}{4}\right)
\end{aligned}$$

where we note that the maximization over $|\psi\rangle$ drops out as the result is independent of the state. Calculating the trace distance:

$$\begin{aligned}
D(U,V) &= D(I,W) \\
&= \mathrm{Tr}\sqrt{(I-W)^\dagger (I-W)} \\
&= \mathrm{Tr}\sqrt{2I - W^\dagger - W} \\
&= \mathrm{Tr}\sqrt{(2 - 2\cos\left(\frac{\beta}{2}\right))I} \\
&= 2\sin\left(\frac{\beta}{4}\right) \mathrm{Tr}\left(\sqrt{I}\right) \\
&= 2\sin\left(\frac{\beta}{4}\right) \mathrm{Tr}(I) \\
&= 4\sin\left(\frac{\beta}{4}\right)
\end{aligned}$$

so indeed $D(U,V) = 2E(U,V)$ as claimed. □

Exercise A3.2

Suppose A and B are Hermitian matrices such that $\mathrm{tr}\,|A|, \mathrm{tr}\,|B| \leq \epsilon$. Prove that for sufficiently small ϵ,

$$D([e^{-iA}, e^{-iB}]_{\mathsf{gp}}, e^{-[A,B]}) \leq d\epsilon^3$$

for some constant d, establishing Equation (A3.6). (*Comment:* for practical purposes it may be interesting to obtain good bounds on d.)

Solution

Concepts Involved: Trace Distance

Let us expand out the exponentials appearing in the above to quadratic order in A, B.

$$e^{\pm iA} = I \pm iA + \frac{1}{2}(iA)^2 = I \pm iA - \frac{1}{2}A^2 + O(A^3)$$

$$e^{-[A,B]} = I - [A,B] + O((A,B)^3)$$

The group commutator term thus looks like:

$$\begin{aligned}
[e^{-iA}, e^{-iB}]_{\mathsf{gp}} &= e^{-iA}e^{-iB}e^{iA}e^{iB} \\
&= (I - iA - \frac{1}{2}A^2)(I - iB - \frac{1}{2}B^2)(I + iA - \frac{1}{2}A^2)(I + iB - \frac{1}{2}B^2) + O((A,B)^3) \\
&= I - AB + A^2 + AB + BA + B^2 - AB - A^2 - B^2 + O((A,B)^3) \\
&= I - [A,B] + O((A,B)^3)
\end{aligned}$$

Thus:

$$\begin{aligned}
D([e^{-iA}, e^{-iB}]_{\mathsf{gp}}, e^{-[A,B]}) &= \mathrm{Tr}\Big(\Big|(I - [A,B] + O((A,B)^3)) - I - [A,B] + O((A,B)^3))\Big|\Big) \\
&= \mathrm{Tr}\Big(O((A,B)^3)\Big) \\
&\leq d\epsilon^3
\end{aligned}$$

for some constant d. A bound on d can be estimated by retaining the third order terms in the above calculation. □

Exercise A3.3

Let $\mathbf{x}$ and $\mathbf{y}$ be any two real vectors. Show that:

$$D(u(\mathbf{x}), u(\mathbf{y})) = 2\sqrt{2}\sqrt{1 - \cos(x/2)\cos(y/2) - \sin(x/2)\sin(y/2)\hat{\mathbf{x}} \cdot \hat{\mathbf{y}}},$$

where $x \equiv \|\mathbf{x}\|, y \equiv \|\mathbf{y}\|$, and $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the unit vectors in the $\mathbf{x}$ and $\mathbf{y}$ directions, respectively.

Solution

Concepts Involved: Trace Distance

From the definition of the trace distance we have:

$$\begin{aligned} D(u(\mathbf{x}),u(\mathbf{y})) &= \mathrm{Tr}\left(\sqrt{(u(\mathbf{x})-u(\mathbf{y}))^\dagger(u(\mathbf{x})-u(\mathbf{y}))}\right) \\ &= \mathrm{Tr}\left(\sqrt{2I-u(\mathbf{x})^\dagger u(\mathbf{y})-u(\mathbf{y})^\dagger u(\mathbf{x})}\right) \end{aligned}$$

Now noting that $u(\mathbf{x}) = \exp\big(-i\mathbf{x}\cdot\boldsymbol{\sigma}/2\big) = \exp\big(-i\hat{\mathbf{x}}\cdot\boldsymbol{\sigma}x/2\big)$, we find:

$$\begin{aligned} u(\mathbf{x})^\dagger u(\mathbf{y}) + u(\mathbf{y})^\dagger u(\mathbf{x}) &= \exp\big(i\hat{\mathbf{x}}\cdot\boldsymbol{\sigma}x/2\big)\exp\big(-i\hat{\mathbf{y}}\cdot\boldsymbol{\sigma}y/2\big) + \exp\big(i\hat{\mathbf{y}}\cdot\boldsymbol{\sigma}y/2\big)\exp\big(-i\hat{\mathbf{x}}\cdot\boldsymbol{\sigma}x/2\big) \\ &= (\cos\big(x/2\big)I + i\sin\big(x/2\big)\hat{\mathbf{x}}\cdot\boldsymbol{\sigma})(\cos\big(y/2\big)I - i\sin\big(y/2\big)\hat{\mathbf{y}}\cdot\boldsymbol{\sigma}) \\ &+ (\cos\big(y/2\big)I + i\sin\big(y/2\big)\hat{\mathbf{y}}\cdot\boldsymbol{\sigma})(\cos\big(x/2\big)I - i\sin\big(x/2\big)\hat{\mathbf{x}}\cdot\boldsymbol{\sigma}) \\ &= 2\cos\big(x/2\big)\cos\big(y/2\big)I + \sin\big(x/2\big)\sin\big(y/2\big)\{\hat{\mathbf{x}}\cdot\boldsymbol{\sigma},\hat{\mathbf{y}}\cdot\boldsymbol{\sigma}\} \\ &= 2\cos\big(x/2\big)\cos\big(y/2\big)I + 2\sin\big(x/2\big)\sin\big(y/2\big)\hat{\mathbf{x}}\cdot\hat{\mathbf{y}}I \\ &= (2\cos\big(x/2\big)\cos\big(y/2\big) + 2\sin\big(x/2\big)\sin\big(y/2\big)\hat{\mathbf{x}}\cdot\hat{\mathbf{y}})I \end{aligned}$$

where in the second-to-last equality we note the anticommutation property of the Pauli matrices from Ex. 2.40.
Thus:

$$\begin{aligned} D(u(\mathbf{x}),u(\mathbf{y})) &= \sqrt{2-\cos\big(x/2\big)\cos\big(y/2\big)-\sin\big(x/2\big)\sin\big(y/2\big)\hat{\mathbf{x}}\cdot\hat{\mathbf{y}}}\,\mathrm{Tr}\Big(\sqrt{I}\Big) \\ &= 2\sqrt{2}\sqrt{1-\cos\big(x/2\big)\cos\big(y/2\big)-\sin\big(x/2\big)\sin\big(y/2\big)\hat{\mathbf{x}}\cdot\hat{\mathbf{y}}} \end{aligned}$$

as claimed. □

Exercise A3.4

Show that in the case of $\mathbf{y} = \mathbf{0}$ the formula for $D(u(\mathbf{x}),u(\mathbf{y}))$ reduces to

$$D(u(\mathbf{x}),I) = 4\sin\left|\frac{x}{4}\right|$$

Solution

Concepts Involved: Trace Distance

If $\mathbf{y} = 0$, then $u(\mathbf{y}) = I$ and the last term in the $D(u(\mathbf{x}),u(\mathbf{y}))$ formula drops out, yielding:

$$D(u(\mathbf{x}),I) = 2\sqrt{2}\sqrt{1-\cos\big(x/2\big)} = 2\sqrt{2}\sqrt{1-(1-2\sin^2(x/4))} = 4\left|\sin\left(\frac{x}{4}\right)\right| = 4\sin\left|\frac{x}{4}\right|$$

where in the second equality we use the double angle identity $(2\theta) = 1-2\sin^2(\theta)$. □

Exercise A3.5

Show that when $x, y \leq \epsilon$,

$$D(u(\mathbf{x}), u(\mathbf{y})) = \|\mathbf{x} - \mathbf{y}\| + O(\epsilon^3)$$

This exercise appears to be incorrect, and we instead find:

$$D(u(\mathbf{x}), u(\mathbf{y})) = \|\mathbf{x} - \mathbf{y}\| + O(\frac{\epsilon^4}{\|\mathbf{x} - \mathbf{y}\|})$$

which reduces to the claim of the exercise if $\|\mathbf{x} - \mathbf{y}\| = \Omega(\epsilon)$, but this is not guaranteed. Note that if we instead work with squared quantities, we find:

$$D(u(\mathbf{x}), u(\mathbf{y}))^2 = \|\mathbf{x} - \mathbf{y}\|^2 + O(\epsilon^4).$$

This appears to present an error in the proof of Lemma A3.2 as it is given in the text. We spend the rest of the exercise with a proposed patch for the proof.

Solution

Concepts Involved: Trace Distance

When $x, y \leq \epsilon$, we have that $\cos(x/2) \approx 1 - (x/2)^2/2 + O(\epsilon^4)$ and $\sin(x/2) \approx x/2 + O(\epsilon^3)$ (and similarly for the y terms), so the expression of (A3.10) becomes:

$$\begin{aligned}
&D(u(\mathbf{x}), u(\mathbf{y})) \\
&= 2\sqrt{2}\sqrt{1 - (1 - (x/2)^2/2 + O(\epsilon^4))(1 - (y/2)^2/2 + O(\epsilon^4)) - (x/2 + O(\epsilon^3))(y/2 + O(\epsilon^3))\hat{\mathbf{x}} \cdot \hat{\mathbf{y}}} \\
&= 2\sqrt{2}\sqrt{(x/2)^2/2 + (y/2)^2/2 - (x/2)(y/2)\hat{\mathbf{x}} \cdot \hat{\mathbf{y}} + O(\epsilon^4)} \\
&= \sqrt{\mathbf{x}^2 + \mathbf{y}^2 - 2\mathbf{x} \cdot \mathbf{y} + O(\epsilon^4)} \\
&= \sqrt{\|\mathbf{x} - \mathbf{y}\|^2 + O(\epsilon^4)} \\
&= \|\mathbf{x} - \mathbf{y}\| \sqrt{1 + \frac{O(\epsilon^4)}{\|\mathbf{x} - \mathbf{y}\|}} \\
&= \|\mathbf{x} - \mathbf{y}\| + O(\frac{\epsilon^4}{\|\mathbf{x} - \mathbf{y}\|})
\end{aligned}$$

where in the last equality we taylor expand the square root, with $\sqrt{1 + x} = 1 + \frac{1}{2}x + O(x^2)$. If we square both the first and fourth lines, we instead obtain:

$$D(u(\mathbf{x}), u(\mathbf{y}))^2 = \|\mathbf{x} - \mathbf{y}\|^2 + O(\epsilon^4).$$

This result calls the proof of Lemma A3.2 into question. We attempt to correct it as follows - the first ingridient is Lipschitz continuity. Note that $u(\mathbf{x}) = \exp(-\mathbf{x} \cdot \boldsymbol{\sigma}/2)$ is smooth on all of $\mathbb{R}^3$ and has bounded derivative. Hence on any compact subset of $\mathbb{R}^3$, it will be Lipschitz continuous. In particular, for an ϵ-ball

of $\mathbb{R}^3$, so taking $\mathbf{x}, \mathbf{y}$ such that $\|\mathbf{x}\|, \|\mathbf{y}\| \leq \epsilon$:

$$D(u(\mathbf{x})u(\mathbf{y})) \leq K\|\mathbf{x} - \mathbf{y}\|$$

for some constant K.
The second ingredient is as follows; suppose $D(u(\mathbf{y}), u(\mathbf{y}_0)) < \epsilon^2 + O(\epsilon^5)$ and $\|\mathbf{y}\|, \|\mathbf{y}_0\| \leq \epsilon$. Then:

$$D(u(\mathbf{y}), u(\mathbf{y}_0))^2 < \epsilon^4 + O(\epsilon^7)$$

then using the result we found for the squared distance:

$$\|\mathbf{y} - \mathbf{y}_0\|^2 + O(\epsilon^4) < \epsilon^4 + O(\epsilon^7) \implies \|\mathbf{y} - \mathbf{y}_0\| < c\epsilon^2$$

for some constant c.
Using these two results, let us use them in place of where Ex. A3.5 was used in the text, namely in (A3.16)-(A3.21):

$$\begin{aligned}
D(U, [U_1, U_2]_{\mathsf{gp}}) &\leq D(u(\mathbf{x}), u(\mathbf{y}_0 \times \mathbf{z}_0)) + d'\epsilon^3 \\
&\leq K\|\mathbf{x} - \mathbf{y}_0 \times \mathbf{z}_0\| + d'\epsilon^3 \\
&= K\|\mathbf{y} \times \mathbf{z} - \mathbf{y}_0 \times \mathbf{z}_0\| + d'\epsilon^3 \\
&= K\big\|(\mathbf{y} - \mathbf{y}_0) \times \mathbf{z} + \mathbf{y}_0 \times (\mathbf{z} - \mathbf{z}_0)\big\| + d'\epsilon^3 \\
&\leq K(\|\mathbf{y} - \mathbf{y}_0\|\|\mathbf{z}\| + \|\mathbf{y}_0\|\|\mathbf{z} - \mathbf{z}_0\|) + d'\epsilon^3 \\
&\leq K(c\epsilon^2(\epsilon + O(\epsilon^5)) + (\epsilon + O(\epsilon^3))c'\epsilon^2) + d'\epsilon^3 \\
&\leq C\epsilon^3
\end{aligned}$$

where in the second line we use the Lipschitz result, in the fifth line we use the traingle inequality, and in the sixth line we use the bounds on $\|\mathbf{y} - \mathbf{y}_0\|, \|\mathbf{z} - \mathbf{z}_0\|$ from the bounds on the squared distances. Thus we recover the claimed $O(\epsilon^3)$ bound of the Lemma. □

Remark: Thank you to Alex Nachmann for helpful discussions about Lipschitz continuity.

Exercise A3.6

($\star\star$) Fixing the set of $\mathcal{G}$ of elementary gates, describe an algorithm which, given a description of a single qubit unitary gate U and a desired accuracy $\epsilon > 0$, efficiently computes a sequence of gates from $\mathcal{G}$ such that ϵ-approximates U.

Solution

Concepts Involved: Unitary Approximations

The point of the question is to adapt the proof of the Solovay-Kitaev theorem/Lemma A3.2 (which are nearly constructive) into a (classical) algorithm that produces the approximating unitary. We follow/adapt the presentation of Nielsen and Dawson in arXiv:quant-ph/0505030.
The iterative aspect of the proof will translate to an algorithm that is recursive. In pseudocode, the algorithm takes the form:

Solovay-Kitaev(U, n):
if $n = 0$ **then**
  return $U_0 =$ basic-approximation(U)
else
  $U_{n-1} =$ Solovay-Kitaev($U, n-1$)
  $\Delta = UU_{n-1}^\dagger$
  $V_1, V_2 =$ Group-Commutator(Δ)
  $W_1 =$ Solovay-Kitaev($V_1, n-1$)
  $W_2 =$ Solovay-Kitaev($V_2, n-1$)
  return $U_n = [W_1, W_2]_{\text{gp}} U_{n-1} = W_1 W_2 W_1^\dagger W_2^\dagger U_{n-1}$
end

The algorithm takes in as input U the gate to be approximated and n the order of approximation, with the output U_n approximating U to error ϵ_n, i.e. $D(U_n, U) = \epsilon_n$. We will soon discuss the condition on n for $\epsilon_n < \epsilon$ the dsired accuracy. The gate set $\mathcal{G}$ enters implicitly in the $n = 0$ step, as we will also discuss shortly.

Note a slight departure in notation from the appendix - there, $U_n \in \mathcal{G}_{5^n l_0}$ referred to the unitary that was appended at the nth step of the iterative proof, while here U_n is the nth order approximation to U.

Let us walk through the algorithm. The $n = 0$ step corresponds to the 0th order approximation to U, i.e. the best approximation to U consisting of length l_0 sequences of elementary gates from $\mathcal{G}$ (with l_0 sufficiently large such that $\mathcal{G}_{l_0}$ forms an ϵ_0-net, where ϵ_0 is a $\mathcal{G}$-independent constant, whose value we will soon discuss). This is done via a precomputation step. Before running the above algorithm, we first enumerate all $|G|^{l_0}$ possible sequences of gates from $\mathcal{G}$, and store them in memory. The basic-approximation(U) function then iterates over all $U_0 \in \mathcal{G}_{l_0}$ until finding the one for which $D(U_0, U) < \epsilon_0$. Note that this step requires exponential time in l_0, but is a constant in ϵ.

Let us walk through the more interesting part of the algorithm. U_{n-1} - the ϵ_{n-1}-approximation to U - is first obtained via a recursive call to Solovay-Kitaev. We then compute $\Delta = UU_{n-1}^\dagger$ which we note satisfies $D(\Delta, I) < \epsilon_{n-1}$.

In the next line, we apply the line of reasoning in the first part of the proof of Lemma A3.2 (where we replace each ϵ^2 appearing in that proof with an ϵ_{n-1}). Let $\Delta = u(\mathbf{x})$ with $\|\mathbf{x}\| \leq \epsilon_{n-1}$. Then, we choose any pair of vectors $\mathbf{y}, \mathbf{z}$ of length $O(\sqrt{\epsilon_{n-1}})$ such that $\mathbf{x} = \mathbf{y} \times \mathbf{z}$, which have associated unitaries $V_1 = u(\mathbf{y}), V_2 = u(\mathbf{z})$.

Then in the next two lines, we again recursively call Solovay-Kitaev to obtain ϵ_{n-1} approximations to V_1, V_2 of the form W_1, W_2. The algebra of the proof of Lemma A3.2 (Eqs. (A3.15-A3.21)) shows that $D(\Delta, [W_1, W_2]_{\text{gp}}) = O(\epsilon_{n-1}^{3/2})$, and so as $\Delta U_{n-1} = U$ we find that $D(U_n, U) = D([W_1, W_2]_{\text{gp}} U_{n-1}, U) = C\epsilon_{n-1}^{3/2}$ for some constant C.

We can analyze this algorithm to obtain the recurrences of the approximation error ϵ_n, gate sequence length l_n, and time t_n, which are:

$$\begin{aligned} \epsilon_n &= C\epsilon_{n-1}^{3/2} \\ l_n &= 5l_{n-1} \\ t_n &= 3t_{n-1} + \text{const.}. \end{aligned}$$

The error we have discussed above. l_n follows from the fact that $U_n = W_1 W_2 W_1^\dagger W_2^\dagger U_{n-1}$ consists of the product of 5 operators of length l_{n-1}. t_n follows from the fact that at each n, we recursively call

Solovay-Kitaev with $n-1$ thrice (plus a constant term for the time complexity of computing the group commutator, taking products of operators etc. which does not affect the asymptotic analysis). These recurrence relations can be solved to obtain:

$$\epsilon_n = \frac{1}{C^2}(\epsilon_0 C^2)^{\frac{3}{2}n}$$
$$l_n = O(5^n)$$
$$t_n = O(3^n)$$

In order for the error to decrease in n, we require that $\epsilon_0 C^2 < 1$. In arXiv:quant-ph/0505030 a rough estimate of the constants appearing in the algorithm is made such that $C \approx 4\sqrt{2}$ and so $\epsilon_0 < 1/32$. This also sets the error threshold/sequence length necessary for the initial basic approximation step (based on $\mathcal{G}$, l_0 must be chosen sufficiently large such that $d(U_0, U) < \epsilon_0 < 1/32$ for $U_0 \in \mathcal{G}_{l_0}$).
For a fixed target accuracy $\epsilon_n < \epsilon$, we require:

$$n = \left\lceil \frac{\log\left[\frac{\log(1/\epsilon C^2)}{\log(1/\epsilon_0 C^2)}\right]}{\log(3/2)} \right\rceil$$

from which we obtain:

$$l_\epsilon = O(\log^{\frac{\log(5)}{\log(3/2)}}(1/\epsilon))$$
$$t_\epsilon = O(\log^{\frac{\log(3)}{\log(3/2)}}(1/\epsilon))$$

i.e. the output sequence length and the runtime are both polylogarithmic in $1/\epsilon$ (and we note that we recover the power of the logarithm appearing in the Solovay-Kitaev theorem proof of the text). Thus we have an efficient classical algorithm which produces an approximation to a single-qubit unitary U to desired accuracy ϵ using a sequence of gates from an elementary set $\mathcal{G}$. □

Problem A3.1

($\star\star$) The following problem outlines a more elaborate construction that achieves $O(\log^2(1/\epsilon)\log^c(\log(1/\epsilon)))$ bound on the number of gates required to approximate to within ϵ of a desired target, for any $c > 2$.

(1) Suppose $\mathcal{N}$ is a δ-net in S_ϵ, for $0 < \delta < \epsilon \leq \epsilon_0$, ϵ_0 sufficiently small. Show that $[\mathcal{N}, \mathcal{N}]_{\text{gp}}$ is a $d\delta\epsilon$-net in S_{ϵ^2}, for some constant d. We appear to require the additional assumption that $\delta = \Omega(\epsilon^2)$.

(2) Suppose G_l is a δ-net in S_ϵ, for $0 < \delta < \epsilon \leq \epsilon_0$. Show that $\mathcal{G}_{4^k l}$ is a $d^k \delta \epsilon^{2^k - 1}$-net in $S_{\epsilon^{2^k}}$. We again require an additional assumption, namely that $d^{k'}\delta = \Omega(\epsilon^{2^{k'}+1})$ for $k' = 0, 1, 2 \ldots$.

(3) Suppose we define k by

$$k \equiv \left\lceil \log\left(\frac{\log(1/\epsilon)}{\log(1/\epsilon_0)}\right) \right\rceil,$$

and suppose we can find l such that $\mathcal{G}_l$ is a δ_0-net for S_{ϵ_0}, where

$$d^k \delta_0 = \epsilon_0.$$

Show that $\mathcal{G}_{4^k l}$ is an ϵ-net for $S_{\epsilon_0^{2^k}}$.

(4) Use the already-proved version of the Solovay-Kitaev theorem to show that choosing $l = O(k^c)$ suffices in the previous part of the problem, where $c = \log(5)/\log(3/2)$ is the constant appearing in the exponent in the already-proved version of the Solovay–Kitaev theorem.

(5) Combine the previous results to prove that $O(\log^2(1/\epsilon)\log^c(\log(1/\epsilon)))$ gates can be used to ϵ-approximate an arbitrary gate in $SU(2)$

(6) Show that any $c > 2$ can appear in the conclusion of the previous result.

The construction in this exercise is erroneous, based on private correspondence with Greg Kuperberg (and also as noted in arXiv:2306.13158). We sketch the construction/what we felt to be the "intended spirit of the solution", pointing out where we believe the construction to fail, in step (3).

Solution

Concepts Involved: Nets, Unitary Approximations

(1) We follow the (modified, see Ex. A3.5) proof of Lemma A3.2.

Let $U \in S_{\epsilon^2}$ and pick $\mathbf{x}$ such that $U = u(\mathbf{x})$. By Ex. A3.4 it follows that $\|\mathbf{x}\| \leq \epsilon^2 + O(\epsilon^6)$. Choose any pair of vectors $\mathbf{y}, \mathbf{z}$ of length most $\epsilon + O(\epsilon^5)$ such that $\mathbf{x} = \mathbf{y} \times \mathbf{z}$. $\mathcal{N}$ is a δ-net for S_ϵ, so choose U_1, U_2 in $\mathcal{N} \cap S_\epsilon$ such that:

$$D(U_1, u(\mathbf{y})) < \delta + O(\delta^3)$$
$$D(U_2, u(\mathbf{z})) < \delta + O(\delta^3)$$

and let $\mathbf{y}_0, \mathbf{z}_0$ be chosen such that $U_1 = u(\mathbf{y}_0), U_2 = u(\mathbf{z}_0)$. By Ex. A3.4 it follows that $\|\mathbf{y}_0\|, \|\mathbf{z}_0\| \leq \epsilon + O(\epsilon^3)$. We then wish to bound $D(U, [U_1, U_2]_{\text{gp}})$, and so use the triangle inequality:

$$D(U, [U_1, U_2]_{\text{gp}}) \leq D(U, u(\mathbf{y}_0 \times \mathbf{z}_0)) + D(u(\mathbf{y}_0 \times \mathbf{z}_0), [U_1, U_2]_{\text{gp}})$$

The second term is at most $d'\epsilon^3$ by Ex. A3.2 (with d' larger than the constant appearing in that exercise, due to the fact that $\|\mathbf{y}_0\|, \|\mathbf{z}_0\| \leq \epsilon + O(\epsilon^3)$). This is where the additional assumption is required, as unless $\delta = \Omega(\epsilon^2)$ then there is no way to guarantee that $O(\epsilon^3) = O(\delta\epsilon)$ - with the additional assumption, the second term can be bounded by $d'\delta\epsilon$.

Now looking at the first term, we substitute $U = u(\mathbf{x})$ and use the results from Lipschitz continuity

and distance bounds from our solution to Ex. A3.5:

$$\begin{aligned} D(U,[U_1,U_2]_{\text{gp}}) &\leq D(u(\mathbf{x}),u(\mathbf{y}_0\times\mathbf{z}_0))+d'\delta\epsilon \\ &\leq K\|\mathbf{y}\times\mathbf{z}-\mathbf{y}_0\times\mathbf{z}_0\|+d'\delta\epsilon \\ &= K\big\|(\mathbf{y}-\mathbf{y}_0)\times\mathbf{z}+\mathbf{y}_0\times(\mathbf{z}-\mathbf{z}_0)\big\|+d'\delta\epsilon \\ &\leq K(\|\mathbf{y}-\mathbf{y}_0\|\|\mathbf{z}\|+\|\mathbf{y}_0\|\|\mathbf{z}-\mathbf{z}_0\|)+d'\delta\epsilon \\ &\leq K(c\epsilon^2(\epsilon+O(\epsilon^5))+(\epsilon+O(\epsilon^3))c'\epsilon^2)+d'\delta\epsilon \\ &= K(c+c')\epsilon^3+d'\delta\epsilon+O(\epsilon^5) \\ &\leq d\delta\epsilon \end{aligned}$$

where in the last inequality we again use the additional assumption that $\delta=\Omega(\epsilon^2)$.

(2) We proceed by induction on k. The $k=0$ case immediately follows by assumption.

Now, suppose the claim holds for some $k\geq 0$. Then, we have that $G_{4^k l}$ is a $d^k\delta\epsilon^{2^k-1}$-net in $S_{\epsilon^{2^k}}$.

Then, by part (1), noting that $0<d^k\delta\epsilon^{2^k-1}<\epsilon^{2^k}\leq\epsilon_0$ and $d^k\delta\epsilon^{2^k-1}=\Omega(\epsilon^{2^{k+1}})$ (using the additional assumption), we find that $[G_{4^k l},G_{4^k l}]_{\text{gp}}$, and hence $G_{4^{k+1}l}$ is a $d(d^k\delta\epsilon^{2^k-1})\epsilon^{2^k}=d^{k+1}\delta\epsilon^{2^{k+1}-1}$-net in $S_{\epsilon^{2^{k+1}}}$, completing the induction.

(3) By assumption, G_l is a δ_0-net in S_{ϵ_0}. Further, we have that $0<\delta_0=\epsilon_0/d^k<\epsilon_0\leq\epsilon_0$ (so long as $d>1$) and we also require that $d^{k'}\delta_0=\Omega(\epsilon_0^{2^{k'}+1})$ for $k'=0,1,\ldots,k-1,k$ (each step of the induction).

We believe this is where the construction breaks down - namely, the lower bound condition fails. Let us see this:

$$d^{k'}\delta_0=\Omega(\epsilon_0^{2^{k'}+1})\implies d^{k'}\epsilon_0/d^k=\Omega(\epsilon_0^{2^{k'}+1})\implies d^{k'-k}=\Omega(\epsilon_0^{2^{k'}})$$

In particular, taking $k'=1$ we have:

$$d^{1-k}\geq C\epsilon_0^2\implies d^{k-1}\leq C/\epsilon_0^2$$

but for any fixed ϵ_0, as we take $\epsilon\to 0$, $k\to\infty$ and so the LHS diverges (as $d>1$), so the lower bound necessary for applying (2) fails. Let us continue on with the "spirit" of the problem, even though we believe at this point the construction to fail.

Applying (2), we have that $G_{4^k l}$ is a $d^k\delta_0\epsilon_0^{2^k-1}=\epsilon_0\epsilon_0^{2^k-1}=\epsilon_0^{2^k}$-net in $S_{\epsilon_0^{2^k}}$.

Now, since:

$$k\geq\log\left(\frac{\log(1/\epsilon)}{\log(1/\epsilon_0)}\right)\implies 2^k\geq\frac{\log(1/\epsilon)}{\log(1/\epsilon_0)}\implies\epsilon\geq\epsilon_0^{2^k}$$

Thus if $G_{4^k l}$ is an $\epsilon_0^{2^k}$-net in $S_{\epsilon_0^{2^k}}$, it is also a ϵ-net in $S_{\epsilon_0^{2^k}}$.

(4) The already proved Solovay-Kitaev theorem says that G_l is an δ_0-net in $SU(2)$, and hence S_{ϵ_0} for $l=O(\log^c(1/\delta_0))$ with $c=\log(5)/\log(3/2)$. Thus with $\delta_0=\epsilon_0/d^k$:

$$l=O(\log^c(d^k/\epsilon_0))=O(\log^c(d^k)-\log^c(\epsilon_0))=O(k^c\log^c(d)-\log^c(\epsilon_0))=O(k^c)$$

where $\log^c(d), \log^c(\epsilon_0)$ are constants that can be absorbed into the $O(\cdot)$.

(5) The previous results show that $\mathcal{G}_{4^k l}$ is an ϵ-net for $S_{\epsilon_0^{2^k}}$, so taking $l = O(k^c)$ we obtain a U-approximation consiting of:

$$\begin{aligned} O(4^k k^c) &= O(4^{\log\left(\frac{\log(1/\epsilon)}{\log(1/\epsilon_0)}\right)} \log^c\left(\frac{\log(1/\epsilon)}{\log(1/\epsilon_0)}\right)) \\ &= O((2^{\log\left(\frac{\log(1/\epsilon)}{\log(1/\epsilon_0)}\right)})^2 \log^c\left(\frac{\log(1/\epsilon)}{\log(1/\epsilon_0)}\right)) \\ &= O(\log^2(1/\epsilon)\log^c(\log(1/\epsilon))) \end{aligned}$$

where the constants coming from the $\log(1/\epsilon_0)$s can be absorbed into the $O(\cdot)$.

(6) This comes down to combining parts (4)/(5) - in (4), instead of using the version of the SK theorem proven in the appendix with gate sequence length $O(\log^c(1/\epsilon))$ to pick $l = O(k^c)$ for $c = \log(5)/\log(3/2)$, we can use the result of (5) with gate sequence length $O(\log^2(1/\epsilon)\log^c(\log(1/\epsilon))) = O(\log^{2+\kappa}(\log(1/\epsilon))$ for any $\kappa > 0$ to pick $l = O(k^{2+\kappa})$, which then (repeating the arguments of part (5)) gives a gate sequence length of $O(\log^2(1/\epsilon)\log^{2+\kappa}(\log(1/\epsilon))$.

□

Remark: Thank you to Greg Kuperberg for insightful discussions about this problem (and where it appears to not hang together) and various approaches to the Solovay-Kitaev theorem more generally.
As a comment, though it seems to us that the particular approach laid out in the problem does not work as stated, this does not necessarily rule out possibilities for improved exponents using the techniques provided in the text.

Problem A3.2: (Research)

($\star\star\star$) If it exists, find an approximation procedure asymptotically faster than the result found in the previous problem. Ideally, a procedure would (a) saturate the $\Omega(\log(1/\epsilon))$ lower bound on the number of gates required to perform the approximation, and (b) provide an efficient algorithm for constructing such an approximation sequences of gates.

Solution

Concepts Involved: Unitary Approximations

The construction given in the text involves $O(\log^c(1/\epsilon))$ gates, for $c = \log_{3/2}(5) \approx 4$. A construction with $c = 3 + \delta$ can be found in Kitaev, Shen, and Vyalyi's "Classical and Quantum Computation". The state-of-the-art is due to Kuperberg (in arXiv:2306.13158), which achieves $c = \log_\varphi(2) + \delta \approx 1.44$ for $\varphi = (1+\sqrt{5})/2$. Both algorithms are classically efficient (poly$(\log(1/\epsilon))$ time). □

Problem A3.3: (Research)

($\star\star\star$) Fix a finite set of single qubit gates $\mathcal{G}$ which can be performed fault-tolerantly and which generate a set dense in the single qubit gates; say the $\pi/8$ gate and the Hadamard gate. Develop an elegant, efficient, and reasonable tight method which, given an arbitrary single qubit gate U and some $\epsilon > 0$, produces a sequence of gates from the fault-tolerant set giving a ϵ-approximation to U, up to a global phase.

Solution

Concepts Involved: Unitary Approximations

For particular gate sets, a number-theoretic approach (going beyond the machinery of Solovay-Kitaev) yields algorithms for approximating single-qubit U with $O(\log(1/\epsilon))$ (optimal!) gates which run in poly$(\log(1/\epsilon))$ time:

- Selinger's (arXiv:1212.6253), Kliuchnikov/Maslov/Mosca's (arXiv:1212.6964), and Ross/Selinger's (arXiv:1403.2975) construction for Clifford + $\pi/8$.
- Bocharov/Gurevich/Svore's construction (arXiv:1303.1411) for the V-basis
- Kliuchnikov/Bocharov/Svore's construction. (arXiv:1310.4150) for braiding of Fibonacci anyons.
- Bocharov/Roetteler/Svore's construction (arXiv:1409.3552) for Clifford + $\pi/12$.
- Kliuchnikov/Bocharov/Roetteler/Yard's construction (arXiv:1510.03888) for gate sets derived from totally definite quaternion algebras.

□

A4 Number theory

Exercise A4.1: Transitivity

Show that if $a|b$ and $b|c$ then $a|c$.

Solution

Concepts Involved: Divisibility

We have $b = ak_1$ and $c = bk_2$ for $k_1, k_2 \in \mathbb{Z}$, so $c = bk_2 = (ak_1)k_2 = ak_1k_2$ for $k_1k_2 \in \mathbb{Z}$ and so $a|c$. □

Exercise A4.2

Show that if $d|a$ and $d|b$ then d also divides linear combinations of a and b, $ax + by$, where x and y are integers.

Solution

Concepts Involved: Divisibility

We have $a = dk_1$ and $b = dk_2$ for $k_1, k_2 \in \mathbb{Z}$, so for $x, y \in \mathbb{Z}$ we have $ax + by = a(dk_1) + b(dk_2) = (ak_1 + bk_2)d$ with $ak_1 + bk_2 \in \mathbb{Z}$ so $d|ax + by$. □

Exercise A4.3

Suppose a and b are positive integers. Show that if $a|b$ then $a \leq b$. Conclude that if $a|b$ and $b|a$ then $a = b$.

Solution

Concepts Involved: Divisibility

We have $b = ak$ for $k \in \mathbb{Z}$ and since $a, b > 0$ it must follow that $k > 0$ and hence $k \geq 1$. Thus $a \leq ak = b$. If $b|a$ also then $b \leq a$ and so combining the two relations $a = b$. □

Exercise A4.4

Find the prime factorizations of 697 and 36300.

Solution

Concepts Involved: Prime Factorization

By brute-force division (starting from the smallest primes until we find a clean factor) we find:

$$697 = 17 \cdot 41$$

$$36300 = 2^2 \cdot 3 \cdot 5^2 \cdot 11^2$$

□

Exercise A4.5

For p a prime prove that all integers in the range 1 to $p-1$ have multiplicative inverses modulo p. Which integers in the range 1 to p^2-1 do not have multiplicative inverses modulo p^2?

Solution

Concepts Involved: Modular Arithmetic

Since p is prime, all integers $a = 1, \ldots, p-1$ are trivially coprime with p. Thus by corollary A4.4 they have multiplicative inverses modulo p.
By the same corollary, integers in $1, \ldots, p^2-1$ which do not have multiplicative inverses modulo p^2 are those which are not coprime with p^2, i.e. have a common divisor greater than 1. The positive divisors of p^2 are $1, p, p^2$, and hence if $a < p^2$ is coprime with p^2 it must be that a has divisor p, i.e. $a = p, 2p, \ldots (p-1)p$.

□

Exercise A4.6

Find the multiplicative inverse of 17 modulo 24.

Solution

Concepts Involved: Modular Arithmetic

Since 17 is prime, 17 and 24 are coprime and the multiplicative inverse is guaranteed to exist. We then look for integers a^{-1}, b that satisfy:

$$17 \cdot a^{-1} + b \cdot 24 = 1 \tag{43}$$

In particular, by iterating through $\frac{17 \cdot a^{-1} - 1}{24}$ for $a^{-1} \in \mathbb{N}$ and seeing when we obtain an integer, we find $(17)^{-1} = 17$ and $b = 12$, so the multiplicative inverse is of 17 modulo 24 is 17. □

Exercise A4.7

$(\star)$ Find the multiplicative inverse of $n+1$ modulo n^2, where n is any integer greater than 1.

Solution

Concepts Involved: Modular Arithmetic

$n^2, n+1$ are coprime as they do not share any common factors that are not 1 (to see this, consider the prime factorization of n, which has no shared common factors with $n+1$ - else there would exist a prime p which divided their difference, $n+1-n = 1$ - contradiction. n^2 has the same prime factors as n (with doubled multiplicity), and hence again does not share any common factors with $n+1$). Thus we are

guaranteed the existence of a multiplicative inverse of $n+1$ modulo n^2. Thus there exist integers a^{-1}, b such that:

$$(n+1)a^{-1} + bn^2 = 1$$

Notice that $(n+1)(-n+1) = (-n^2+1)$ and so the above is satisfied by taking $a^{-1} = (-n+1), b = 1$. If we want a^{-1} to be positive, we can simply add n^2, wherein $a^{-1} = (n^2 - n + 1)$ and $b = -n$. Thus the multiplicative inverse of $n+1$ modulo n^2 is $n^2 - n + 1$. □

Exercise A4.8: Uniqueness of the Inverse

Suppose b and b' multiplicative inverses of a, modulo n. Prove that $b = b' \pmod{n}$

Solution

Concepts Involved: Modular Arithmetic

If b, b' are multiplicative inverses of a modulo n, there exist integers k, k' such that:

$$ab + kn = 1$$

$$ab' + k'n = 1$$

Subtracting the first equation from the second, we find:

$$a(b-b') + n(k-k') = 0 \implies a(b-b') = 0 \pmod{n}$$

Since $\gcd(a, n) = 1$ by Corollary A4.4, it must be that $(b-b')$ divides n, and hence $b = b' \pmod{n}$. □

Exercise A4.9

Explain how to find $\gcd(a, b)$ if the prime factorizations of a and b are known. Find the prime factorizations of 6825 and 1430, and use them to compute $\gcd(6825, 1430)$.

Solution

Concepts Involved: Modular Arithmetic, Prime Factorization

If the prime factorizations of a, b are known, their greatest common divisor can be easily found by multiplying together all common prime factors (and their multiplicities).
For example:

$$6825 = 3 \cdot 5^2 \cdot 7 \cdot 13$$

$$1430 = 2 \cdot 5 \cdot 11 \cdot 13$$

the shared prime factors are $5^1, 13^1$ and so $\gcd(6825, 1430) = 5 \cdot 13 = 65$. □

Exercise A4.10

What is $\varphi(187)$?

Solution

Concepts Involved: Prime Factorization, Euler φ function

The prime factorization of 187 is $187 = 11 \cdot 17$. Thus using the formula of Eq. (A4.23), we find:

$$\varphi(187) = \sum_{j=1}^{k} p_j^{\alpha_j - 1}(p_j - 1) = 11^{1-1}(11-1) \cdot 17^{1-1}(17-1) = 10 \cdot 16 = 160.$$

□

Exercise A4.11

($\star\star$) Prove that

$$n = \sum_{d|n} \varphi(d)$$

where the sum is over all positive divisors d of n, including 1 and n. (*Hint:* Prove the result for $n = p^\alpha$ first, then use the multiplicative property (A4.22) of φ to complete the proof.)

Solution

Concepts Involved: Prime Factorization, Euler φ function

We follow the hint and first show the result for $n = p^\alpha$. In this case the divisors of n are just p^m for $m = 0, 1, \ldots, \alpha$, so:

$$\sum_{d|p^\alpha} \varphi(d) = \sum_{m=0}^{\alpha} \varphi(p^m) = 1 + \sum_{m=1}^{\alpha} \left(p^{m-1}(p-1) \right) = 1 + \sum_{m=1}^{\alpha} (p^m - p^{m-1})$$

whre in the last equality we use Eq. (A4.21) for the Euler-φ function of prime powers (and peel off the $\varphi(p^0) = \varphi(1) = 1$ term). Now notice that the sum telescopes, and so only the p^{m-1} from the first term and the p^m from the last term survive:

$$\sum_{d|p^\alpha} \varphi(d) = 1 + p^\alpha - p^{1-1} = p^\alpha$$

which proves the claim for prime powers.
We proceed to the case of general n, where we can use the fundamental theorem of arithmetic to write:

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_n^{\alpha_n}$$

with $p_1, \ldots, p_n$ prime. A divisor of n takes the form:

$$d = p_1^{m_1} p_2^{m_2} \ldots p_n^{m_n}$$

where $m_1 = 0, \ldots, \alpha_1, m_2 = 0, \ldots, \alpha_2$, and so on. Thus the sum over all divisors can be written as:

$$\sum_{d|n} \varphi(d) = \sum_{m_1=0}^{\alpha_1} \sum_{m_2=0}^{\alpha_2} \ldots \sum_{m_n=0}^{\alpha_n} \varphi(p_1^{m_1} p_2^{m_2} \ldots p_n^{m_n})$$

Since $p_1, \ldots p_n$ are prime, they are trivially coprime, so we can use the multiplicative property of Eq. (A4.22) to write:

$$\sum_{d|n} \varphi(d) = \sum_{m_1=0}^{\alpha_1} \sum_{m_2=0}^{\alpha_2} \ldots \sum_{m_n=0}^{\alpha_n} \varphi(p_1^{m_1}) \varphi(p_2^{m_2}) \ldots \varphi(p_n^{m_n}) = \prod_{i=1}^{n} \sum_{m_i=0}^{\alpha_i} \varphi(p_i^{m_i})$$

From the $n = p^\alpha$ case, we saw that $\sum_{m_i}^{\alpha_i} \varphi(p_i^{m_i}) = p_i^{\alpha_i}$, and so:

$$\sum_{d|n} \varphi(d) = \prod_{i=1}^{n} p_i^{\alpha_i} = n$$

which completes the proof. □

Exercise A4.12

Verify that $\mathbb{Z}_n^*$ forms a group of size $\varphi(n)$ under the operation of multiplication modulo n.

Solution

Concepts Involved: Group Axioms, Modular Arithmetic Euler φ function

That $|\mathbb{Z}_n^*| = \varphi(n)$ is immediate from the definition of φ. We then check that $\mathbb{Z}_n^*$ satisfies the group axioms:

- Associativity: This follows immediately from the associativity of integer multiplication
- Identity: For any n, 1 is coprime to n and this is the multiplicative identity.
- Closure under products: If $a, b \in \mathbb{Z}_n^*$, then there exists multiplicative inverses $a^{-1}, b^{-1} \in \mathbb{Z}$. It then follows that $ab \in \mathbb{Z}_n^*$ as this has multplicative inverse $b^{-1}a^{-1}$.
- Closure under inverses: If $a \in \mathbb{Z}_n^*$, then there exists $a^{-1} \in \mathbb{Z}$ such that $a^{-1}a = 1$. But then it follows that $a^{-1} \in \mathbb{Z}_n^*$ as it has a multiplicative inverse a.

□

Exercise A4.13

Let a be an arbitrary element of $\mathbb{Z}_n^*$. Show that $S \equiv \{1, a, a^2, \ldots\}$ forms a subgroup of $\mathbb{Z}_n^*$, and that the size of S is the least value of r such that $a^r = 1 \pmod{n}$.

Solution

Concepts Involved: Group Axioms, Subgroups, Cyclic Groups, Order

$\mathbb{Z}_n^*$ is a finite group (with order $\varphi(n)$, from the previous exercise), so by Ex. A2.1, $a \in \mathbb{Z}_n^*$ has some finite order r, i.e. such that $a^r = 1 \pmod{n}$. This implies that $|S| = r$.
We then check that the subgroup obeys the group axioms.

- Associativity: Automatic (inherited from the group)
- Identity: Contains 1.
- Closure under products: If $a^n, a^m \in S$ then $a^n \cdot a^m = a^{n+m} \in S$.
- Closure under inverses: For all $a^n \in S$, there exists $a^{r-n} \in S$ such that $a^n \cdot a^{r-n} = a^r = 1 \pmod{n}$, so $a^{r-n} = (a^n)^{-1} \in S$.

□

Exercise A4.14

Suppose g is generator for $\mathbb{Z}_n^*$. Show that g must have order $\varphi(n)$.

Solution

Concepts Involved: Cyclic Groups, Order

If g is a generator for $\mathbb{Z}_n^*$, then every $x \in \mathbb{Z}_n^*$ may be written as $x = g^k \pmod{n}$, i.e. $\mathbb{Z}_n^*$ is cyclic. By Ex. A4.12 we know that $|\mathbb{Z}_n^*| = \varphi(n)$, and cyclic groups have order equal to the order of their generator, so g has order $\varphi(n)$. □

Exercise A4.15

Lagrange's theorem (Theorem A2.1 on page 610) is an elementary result of group theory stating that the size of a subgroup must divide the order of the group. Use Lagrange's theorem to provide an alternate proof of Theorem A4.9, that is, show that $a^{\varphi(n)} = 1 \pmod{n}$ for any $a \in \mathbb{Z}_n^*$.

Solution

Concepts Involved: Lagrange's Theorem, Order

Consider any $a \in \mathbb{Z}_n^*$ - then a is coprime to n. As in Ex. A4.13, consider the subgroup $S_a = \{1, a, a^2, \ldots a^{r-1}\}$ where r is the order of a. From Lagrange's theorem, we therefore find that $|S_a| = r$ divides $|\mathbb{Z}_n^*| = \varphi(n)$, i.e. that $\varphi(n) = kr$ for some $k \in \mathbb{N}$. Thus, $a^{\varphi(n)} = a^{kr} = (a^r)^k = 1^r = 1 \pmod{n}$.
□

Exercise A4.16

Use Theorem A4.9 to show that the order of x modulo N must divide $\varphi(N)$.

Solution

Concepts Involved: Modular Arithmetic, Order

The order $r \in \mathbb{N}$ of x modulo N is the least positive integer such that $x^r = 1 \pmod{N}$. By assumption, x is coprime to N, and so by theorem A4.9 we also find that $x^{\varphi(N)} = 1 \pmod{N}$. Since r is the least postive integer to satisfy this relation, $r \leq \varphi(N)$.
Suppose that r did not divide $\varphi(N)$ - then $\varphi(N) = kr + m$ for some $k \in \mathbb{N}$ and $r = 1, \dots k-1$. But then $x^{\varphi(N)} = x^{kr+m} = x^m \neq 1 \pmod{N}$, contradicting Theorem A4.9. Thus r must divide $\varphi(N)$. □

Remark: Though we use Theorem A4.9 as directed, the result follows much more immediately by considering an intermediate argument of the previous exercise. Simply consider the subgroup $S_x = \{1, x, x^2, \dots, x^{r-1}\}$ of $\mathbb{Z}_N^*$ and use Lagrange's theorem to conclude that $|S_x| = r$ divides $\left|\mathbb{Z}_N^*\right| = \varphi(N)$.

Exercise A4.17: Reduction of order-finding to factoring

($\star\star$) We have seen that an efficient order-finding algorithm allows us to factor efficiently. Show that an efficient factoring algorithm would allow us to efficiently find the order modulo N of any x co-prime to N.

Solution

Concepts Involved: Prime Factorization, Order, Euler-φ function

Let r be the order modulo N of x (coprime to N) which we wish to find.
Using the efficient factoring algorithm, first compute the prime factorization of $N = p_1^{\alpha_1} \dots p_k^{\alpha_k}$. Then, we can (efficiently) compute $\varphi(N)$ using Eq. (A4.23):

$$\varphi(N) = \prod_{j=1}^{k} p_j^{\alpha_j - 1}(p_j - 1)$$

Then, use the factoring algorithm again to (efficiently) find the prime factorization of $\varphi(N)$ as $\varphi(N) = q_1^{\beta_1} \dots q_l^{\beta_l}$.
From the previous exercise, we know that r divides $\varphi(N)$, and so $r = q_1^{\beta'_1} \dots q_l^{\beta'_l}$. To find r, proceed as follows - start by setting $r_0 = \varphi(N) = q_1^{\beta_1} \dots q_l^{\beta_l}$. At the nth step (where we are checking the prime q_i in the expansion), we verify if if $x^{r_n/q_i} = 1 \pmod{N}$ - if so, update $r_{n+1} = r_n/q_i$ (and then repeat), if not, keep $r_{n+1} = r_n$ and move onto the next prime q_{i+1} in the expansion. Proceed until we have iterated through all primes appearing in the expansion, at which point the procedure terminates with $r = q_1^{\beta'_1} \dots q_l^{\beta'_l}$ with $\beta'_1 \dots \beta'_l$ the minimal values such that $x^r = 1 \pmod{N}$, i.e. the order r. This procedure is efficient as $\varphi(N) \leq N$ and we go through $O(\log(\varphi(N))) = O(\log(N))$ prime factors steps in constructing r from reducing $\varphi(N)$. □

Exercise A4.18

Find the continued fraction expansion for $x = 19/17$ and $x = 77/65$.

Solution

Concepts Involved: Continued Fractions

We go through the procedure as outlined in the text, which is repeating the steps of splitting (into integer/fractional parts) and inverting (the fractional part) until inversion becomes unncessary (i.e. a 1 appears in the numerator without inversion):

$$\begin{aligned} x = \frac{19}{17} &= 1 + \frac{2}{17} \\ &= 1 + \frac{1}{\frac{17}{2}} \\ &= 1 + \frac{1}{8 + \frac{1}{2}} \end{aligned}$$

$$\begin{aligned} x = \frac{77}{65} &= 1 + \frac{12}{65} \\ &= 1 + \frac{1}{\frac{65}{12}} \\ &= 1 + \frac{1}{5 + \frac{5}{12}} \\ &= 1 + \frac{1}{5 + \frac{1}{\frac{12}{5}}} \\ &= 1 + \frac{1}{5 + \frac{1}{2 + \frac{2}{5}}} \\ &= 1 + \frac{1}{5 + \frac{1}{2 + \frac{1}{\frac{5}{2}}}} \\ &= 1 + \frac{1}{5 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2}}}} \end{aligned}$$

□

Exercise A4.19

Show that $q_n p_{n-1} - p_n q_{n-1} = (-1)^n$ for $n \geq 1$. Use this fact to conclude that $\gcd(p_n, q_n) = 1$. (*Hint:* Induct on n.)

Solution

Concepts Involved: Continued Fractions

We follow the hint and proceed by induction. For the $n = 1$ base case, we have:

$$q_1 p_{1-1} - p_1 q_{1-1} = q_1 p_0 - p_1 q_0 = a_1 \cdot a_0 - (1 + a_0 a_1) \cdot 1 = -1 = (-1)^1$$

where we have used the definitions of p_0, q_0, p_1, q_1 given in the statement of Theorem A4.15.

Suppose now the claim holds for some $k \geq 1$. Then, checking the $k+1$ case:

$$\begin{aligned} q_{k+1}p_k - p_{k+1}q_k &= (a_{k+1}q_k + q_{k-1})p_k - (a_{k+1}p_k + p_{k-1})q_k \\ &= p_k q_{k-1} - q_k p_{k-1} \\ &= (-1)(q_k p_{k-1} - p_k q_{k-1}) \\ &= (-1)(-1)^k \\ &= (-1)^{k+1} \end{aligned}$$

where we have used the inductive definitions for p_{k+1}, q_{k+1} (which hold for $n \geq k+1 \leq N$) in the first equality and have used the inductive hypothesis in the second-to-last equality. Thus the claim is proven. Having proven the claim, note that multiplying the obtained formula by $(-1)^n$ yields:

$$q_n((-1)^n p_{n-1}) + p_n((-1)^{n+1} q_{n-1}) = 1$$

which by Theorem A4.2 shows that $\gcd(p_n, q_n) = 1$. □

Problem A4.1: Prime number estimate

($\star\star\star$) Let $\pi(n)$ be the number of prime numbers which are less than n. A difficult-to-prove result known as the *prime number theorem* asserts that $\lim_{n\to\infty} \pi(n)\log(n)/n = 1$ and this $\pi(n) \approx n/\log(n)$. This problem gives poor man's version of prime number theorem which gives a pretty good lower bound on the distribution of prime numbers.

(1) Prove that $n \leq \log\binom{2n}{n}$

(2) Show that

$$\log\binom{2n}{n} \leq \sum_{p \leq 2n} \left\lfloor \frac{\log(2n)}{\log p} \right\rfloor \log p$$

where the sum is over all primes p less than or equal to $2n$.

(3) Use the previous two results to show that

$$\pi(2n) \geq \frac{n}{\log(2n)}$$

Solution

Concepts Involved:

(1) First, note that:

$$\binom{2n}{n} = \frac{(2n)!}{n!(2n-n)!} = \frac{(2n)!}{n!n!} = \frac{\prod_{j=1}^{2n} j}{(\prod_{i=1}^{n} i)^2} = \prod_{j=1}^{n} \frac{j+n}{j} = \prod_{j=1}^{n} \left(1 + \frac{n}{j}\right)$$

Now note that $j \leq n$ for each term in the sum, so $1 + \frac{n}{j} \geq 2$ for each term, and thus:

$$\binom{2n}{n} \geq \prod_{j=1}^{n} 2 = 2^n$$

Taking logarithms on both sides we conclude:

$$n \leq \log \binom{2n}{n}$$

(2) As a Lemma, let us prove *Legendre's formula*. Let $\nu_p(n)$ be the exponent of the largest power of a prime p that divides n. Then:

$$\nu_p(n!) = \sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor.$$

The proof is as follows - The first term counts the factors of p in $n!$ coming from all multiples of p less than n. The second term counts the additional factors of p coming from multiples of p^2 less than n, and so on until we obtain all factors of p by taking the sum over all powers of p.

With the formula established, we can compute ν_p for $\binom{2n}{n}$:

$$\nu_p\binom{2n}{n} = \nu_p(\frac{(2n)!}{n!n!}) = \nu_p((2n)!) - 2\nu_p(n!) = \sum_{i=1}^{\infty} \left(\left\lfloor \frac{2n}{p^i} \right\rfloor - 2\left\lfloor \frac{n}{p^i} \right\rfloor \right)$$

Further, the terms of the sum become zero after $p^i > 2n$, i.e. for $i > \left\lfloor \frac{\log(2n)}{\log p} \right\rfloor$, so:

$$\nu_p\binom{2n}{n} = \sum_{i=1}^{\left\lfloor \frac{\log(2n)}{\log p} \right\rfloor} \left(\left\lfloor \frac{2n}{p^i} \right\rfloor - 2\left\lfloor \frac{n}{p^i} \right\rfloor \right)$$

Now, notice that in this sum, we either have two cases - either $\left\lfloor \frac{2n}{p^i} \right\rfloor = 2\left\lfloor \frac{n}{p^i} \right\rfloor$ or $\left\lfloor \frac{2n}{p^i} \right\rfloor = 2\left\lfloor \frac{n}{p^i} \right\rfloor + 1$ (depending on whether p^i is greater or less than $2 \cdot (n - (p^i)\left\lfloor \frac{n}{p^i} \right\rfloor)$. This means that the terms in the sum have value 1 or 0, and in particular we can thus easily upper bound the sum as:

$$\nu_p\binom{2n}{n} \leq \sum_{i=1}^{\left\lfloor \frac{\log(2n)}{\log p} \right\rfloor} 1 = \left\lfloor \frac{\log(2n)}{\log p} \right\rfloor$$

Now, consider the prime factorization of $\binom{2n}{n} = \prod_{p \leq 2n} p^{\nu_p(\binom{2n}{n})}$ (there are no prime factors greater than $2n$ which enter). Taking the logarithm of this prime factorization, we find:

$$\log \binom{2n}{n} = \sum_{p \leq 2n} \nu_p(\binom{2n}{n}) \log(p)$$

And now applying our upper bound on the exponent:

$$\log \binom{2n}{n} \leq \sum_{p \leq 2n} \left\lfloor \frac{\log(2n)}{\log p} \right\rfloor \log p$$

as claimed.

(3) Combining the inequalities from (1) and (2), we find:

$$n \leq \sum_{p \leq 2n} \left\lfloor \frac{\log(2n)}{\log p} \right\rfloor \log p$$

Further using that $\lfloor x \rfloor y \leq \lfloor xy \rfloor$ for $x, y \geq 1$ and $\lfloor x \rfloor \leq x$, we find:

$$n \leq \sum_{p \leq 2n} \left\lfloor \frac{\log(2n)}{\log p} \log p \right\rfloor = \sum_{p \leq 2n} \lfloor \log(2n) \rfloor \leq \sum_{p \leq 2n} \log(2n) = \pi(2n) \log(2n)$$

And rearranging, we conclude:

$$\pi(2n) \geq \frac{n}{\log(2n)}$$

□

A5 Public key cryptography and the RSA cryptosystem

Exercise A5.1

($\star$) Written examples of the application of RSA tend to be rather opaque. It's better to work through an example yourself. Encode the word 'QUANTUM' (or at least the first letters!), one letter at a time, using $p = 3$ and $q = 11$. Choose appropriate values for e and d, and use a representation of English text involving 5 bits per letter.

Solution

Concepts Involved: RSA Encryption, Modular Arithmetic, Euler φ-function

With $p = 3, q = 11$, we have $n = pq = 33$. We start by picking out e, d - e is a random small odd integer relatively prime to:

$$\varphi(n) = (p-1)(q-1) = (3-1)(11-1) = 20$$

The smallest primes relatively prime to 20 are $7, 11, 13, 17, 19$. A true procedure would randomly pick one of these, but for convenience let's take $e = 7$. Then, we find $d = 3$ as the multiplicative inverse of e mod $\varphi(n)$, as:

$$7 \cdot 3 = 1 + 1 \cdot 20.$$

Using a 5-bit representation of the English alphabet (with the bitstring $\mathbf{b} = b_4b_3b_2b_1b_0$ corresponding to the $\sum_{n=0}^{4} b_n \cdot 2^n$-th letter of the alphabet), we have Q $= 10001 = 17$, U $= 10101 = 21$, A $= 00001 = 1$, N $= 01110 = 14$, T $= 10100 = 20$, U $= 10101 = 21$, M $= 01101 = 13$.
We encode one letter at a time (noting that each letter is represented in $\lfloor \log n \rfloor = \lfloor 5.04 \rfloor = 5$ bits, so the blocks of the message are small enough for the encoding procedure to work), using $E(M) = M^e \pmod{n}$, so:

$$\begin{aligned}
E(Q) &= 17^7 \pmod{33} = 8 \\
E(U) &= 21^7 \pmod{33} = 21 \\
E(A) &= 1^7 \pmod{33} = 1 \\
E(N) &= 14^7 \pmod{33} = 20 \\
E(T) &= 20^7 \pmod{33} = 26 \\
E(U) &= 21^7 \pmod{33} = 21 \\
E(M) &= 13^7 \pmod{33} = 7
\end{aligned}$$

So the encrypted message is $(8, 21, 1, 20, 26, 21, 7)$. □

Exercise A5.2

Show that d is also an inverse of e modulo r, and thus $d = d' \pmod{r}$

Solution

Concepts Involved: Divisibility, Modular Arithmetic

Since d is the inverse of e modulo $\varphi(n)$, there exists integer k such that $ed = 1 + k\varphi(n)$. Since r divides $\varphi(n)$, there exists integer k' such that $\varphi(n) = k'r$. Substituting the second equation into the first, we obtain $ed = 1 + kk'r$, and since kk' is an integer, this means d is an inverse of e modulo r.
Since d' is also an inverse of e modulo r, there exists integer k'' such that $ed' = 1 + k''r$. Substituting this equation from the previous, we obtain:

$$e(d - d') = (kk' - k'')r \implies d - d' = \frac{kk' - k''}{e}r$$

Since e, r are coprime, it must be the case that $\frac{kk'-k''}{e}$ is an integer, and hence $d = d' \pmod{r}$. □

Problem A5.1

($\star\star$) Write a computer program for performing encryption and decryption using the RSA algorithm. Find a pair 20 bit prime numbers and use them to encrypt a 40 bit message.

Solution

Concepts Involved: RSA Encryption, Modular Arithmetic, Euler-φ function

The Python program RSA.py can be found in the Github Repository. The program generates an RSA public and private key using the procedure given in the appendix (though this was not asked for explicitly, it is certainly the more involved aspect of the problem). A given binary message can then be encoded/decoded block by block.
As an example, we chose to encode the 20-bit message "MBQC" using the same 5-bit representation of the English alphabet as in Ex. A5.1. One run of the program yielded:

$$\begin{aligned} p &= 786161472611 \\ &= 10110111000010101110000000010100011000011 \end{aligned}$$

$$\begin{aligned} q &= 905514821797 \\ &= 11010010110101001110001110101000101001 01 \end{aligned}$$

wherein the message:

$$[M, B, Q, C] = [01101, 00010, 10001, 00011]$$

was encoded block-by-block to:

$$\begin{aligned} E(01101) &= 1461920290375446110677 \\ &= 000000001001111010000000011100000110110100000101100010001100001001000011101010 1 \end{aligned}$$

$$
\begin{aligned}
E(00010) &= 524288 \\
&= 00000000000000000000000000000000000000000000000000000000000010000000000000000000
\end{aligned}
$$

$$
\begin{aligned}
E(10001) &= 23907243568515132484715 3 \\
&= 01100101010000000100100100010000111101110011000001000100010000100111100001100 01
\end{aligned}
$$

$$
\begin{aligned}
E(00011) &= 1162261467 \\
&= 00000000000000000000000000000000000000000000000010001010100011010110011111011011
\end{aligned}
$$

□

A6 Proof of Lieb's theorem

Exercise A6.1: $\leq$ is preserved under conjugation

If $A \leq B$, show that $XAX^\dagger \leq XBX^\dagger$ for all matrices X.

Solution

Concepts Involved: Positive Operators

If $A \leq B$, then $\langle\psi| (B-A) |\psi\rangle \geq$ for all ψ. It then follows that:

$$\langle\psi| (XBX^\dagger - XAX^\dagger) |\psi\rangle = \langle\psi| X(B-A)X^\dagger |\psi\rangle = (\langle\psi| X)(B-A)(X^\dagger |\psi\rangle) \geq 0$$

for all matrices X. Hence $XAX^\dagger \leq XBX^\dagger$. □

Exercise A6.2

Prove that $A \geq 0$ if and only if A is a positive operator.

Solution

Concepts Involved: Positive Operators

If A is a positive operator, then $0 \leq \langle\psi| A |\psi\rangle = \langle\psi| (A-0) |\psi\rangle$ for all $|\psi\rangle$ so $A \geq 0$. Suppose instead $A \geq 0$. Then $(A-0) = A$ is positive. □

Exercise A6.3: $\leq$ is a partial order

Show that the relation $\leq$ is a partial order on operators - that is, it is transitive ($A \leq B$ and $B \leq C$ implies $A \leq C$), asymmetric ($A \leq B$ and $B \leq A$ implies $A = B$), and reflexive ($A \leq A$).

Solution

Concepts Involved: Positive Operators

- (Transitivity) If $A \leq B$ and $B \leq C$ then $\langle\psi| (B-A) |\psi\rangle \geq 0$ and $\langle\psi| (C-B) |\psi\rangle \geq 0$, and hence:

  $$\langle\psi| A |\psi\rangle \leq \langle\psi| B |\psi\rangle \leq \langle\psi| C |\psi\rangle$$

  and so $\langle\psi| (C-A) |\psi\rangle \geq 0$ and $A \leq C$.

- (Asymmetry) If $A \leq B$ and $B \leq A$ then $\langle\psi| (B-A) |\psi\rangle \geq 0$ and $\langle\psi| (A-B) |\psi\rangle \geq 0$ for all $|\psi\rangle$. But this implies that $\langle\psi| B |\psi\rangle \geq \langle\psi| A |\psi\rangle$ and $\langle\psi| B |\psi\rangle \leq \langle\psi| A |\psi\rangle$ and hence $\langle\psi| A |\psi\rangle = \langle\psi| B |\psi\rangle$ for all $|\psi\rangle$, and hence $A = B$.

- (Reflexivity) $A - A = 0$ is positive (as $\langle\psi| 0 |\psi\rangle = 0 \geq 0$ for all $|\psi\rangle$ so $A \leq A$.

□

Exercise A6.4

($\star$) Suppose A has eigenvalues λ_i. Define λ to be the maximum of the set $|\lambda_i|$. Prove that:

1. $\|A\| \geq \lambda$
2. When A is Hermitian, $\|A\| = \lambda$.
3. When

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

$\|A\| = 3/2 > 1 = \lambda$.

Solution

Concepts Involved: Eigenvalues, Operator Norms

1. Let $|v\rangle$ be the unit eigenvector with maximal magnitude eigenvalue λ_i, i.e. such that $|\lambda_i| = \lambda$. Then:

$$\|A\| = \max_{\langle u|u\rangle=1} \left|\langle u| A |u\rangle\right| \geq \left|\langle v| A |v\rangle\right| = |\lambda_i| = \lambda$$

2. By Ex. 2.21 (Spectral decomposition), there exists an orthonormal basis $\left\{|i\rangle\right\}_i$ such that A is diagonal. Therein:

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

with λ_i real. Let us write any normalized $|u\rangle$ in this basis such that $|u\rangle = \sum_i c_i |i\rangle$ with $\sum_i |c_i|^2 = 1$. Then:

$$\begin{aligned}
\|A\| &= \max_{\langle u|u\rangle=1} \left|\langle u| A |u\rangle\right| \\
&= \max_{\{c_i\}} \left| \left(\sum_i c_i^* \langle i|\right) \left(\sum_j \lambda_j |j\rangle\langle j|\right) \left(\sum_k c_k |k\rangle\right) \right| \\
&= \max_{\{c_i\}} \left| \sum_i |c_i|^2 \lambda_i \right| \\
&\leq \max_{\{c_i\}} \left| \sum_i |c_i|^2 \right| \lambda \\
&= \lambda
\end{aligned}$$

where in the third equality we use the orthonormality of the basis and in the fourth equality we use that λ is the magnitude of the largest eigenvector. Thus $\|A\| \leq \lambda$ for Hermitian A, and combining

this with the general previous result that $\|A\| \leq \lambda$ we conclude $\|A\| = \lambda$.

3. Note that A is not Hermitian, and hence we do not expect the equality of the previous part to hold. First, solving for the eigenvalues we have:

$$0 = \det(A - I\lambda_i) = \det\left(\begin{bmatrix} 1-\lambda_i & 0 \\ 1 & 1-\lambda_i \end{bmatrix}\right) = (1-\lambda_i)^2 - 0 = (1-\lambda_i)^2$$

Hence A has eigenvalues $\lambda_1 = \lambda_2 = 1$ and so $\lambda = \max\left\{|\lambda_i|\right\} = 1$.
To solve for the operator norm, we observe that in the $\left\{|+\rangle, |-\rangle\right\}$ basis A can be represented as:

$$A \cong \begin{matrix} \\ |+\rangle \\ |-\rangle \end{matrix}\begin{matrix} \begin{matrix} \langle +| & \langle -| \end{matrix} \\ \begin{bmatrix} 3/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \end{matrix}$$

wherein it becomes clear that $|+\rangle$ would maximize $\langle u| A |u\rangle$, and so $\|A\| = 3/2$.

□

Exercise A6.5: AB and BA have the same eigenvalues

Prove that AB and BA have the same eigenvalues. (Hint: For invertible A, show that $\det(\lambda I - AB) = \det(\lambda I - BA)$, and thus the eigenvalues of AB and BA are the same. By continuity this holds even when A is not invertible).

Solution

Concepts Involved:

We recall that that the determinant is multiplicative, i.e. $\det(AB) = \det(A)\det(B)$, and as a consequence that $\det\left(A^{-1}\right) = \frac{1}{\det(A)}$ for invertible A. Then:

$$\begin{aligned}
\det(\lambda I - AB) &= 1 \cdot \det(\lambda I - AB) \cdot 1 \\
&= \det(I)\det(\lambda I - AB)\det(I) \\
&= \det\left(AA^{-1}\right)\det(\lambda I - AB)\det\left(AA^{-1}\right) \\
&= \det(A)\det\left(A^{-1}\right)\det(\lambda I - AB)\det(A)\det\left(A^{-1}\right) \\
&= \det(A)\det\left(\lambda A^{-1}A - A^{-1}ABA\right)\det\left(A^{-1}\right) \\
&= \det(A)\det(\lambda I - BA)\det\left(A^{-1}\right) \\
&= \det(A)\det(\lambda I - BA)\frac{1}{\det(A)} \\
&= \det(\lambda I - BA)
\end{aligned}$$

Thus $\det(\lambda I - AB) = \det(\lambda I - BA)$ for invertible A. Since the eigenvalues are the roots of the charac-

teristic polynomial given by the determinant, this tells us that AB, BA have the same eigenvalues for all invertible A, and by continuity all A, B. □

Exercise A6.6

Suppose A and B are such that AB is Hermitian. Using the previous two observations, show that $\|AB\| \leq \|BA\|$.

Solution

Concepts Involved:

By Ex. A6.5 AB and BA have the same eigenvalues, so $\lambda = \max\left\{\left|\lambda_{i,AB}\right|\right\} = \max\left\{\left|\lambda_{i,BA}\right|\right\}$. By part 1 of A6.4 we then have that $\|BA\| \geq \lambda$ and by part 2 we have that $\|AB\| = \lambda$ (as AB is Hermitian). Combining these relations $\|AB\| \leq \|BA\|$. □

Exercise A6.7

Suppose A is positive. Show that $\|A\| \leq 1$ if and only if $A \leq I$.

Solution

Concepts Involved:

Since $A \geq 0$, by the spectral theorem

$$A = \sum_i \lambda_i \, |v_i\rangle\langle v_i|, \qquad \lambda_i \geq 0,$$

and for positive (Hermitian) A,

$$\|A\| = \sup_{\|x\|=1} \langle x, Ax\rangle = \lambda_{\max}(A).$$

($\Rightarrow$) If $\|A\| \leq 1$, then $\lambda_{\max}(A) \leq 1$, hence $\lambda_i \leq 1$ for all i, and

$$\langle x, (I - A)x\rangle = \sum_i (1 - \lambda_i) \, |\langle v_i|x\rangle|^2 \geq 0,$$

so $I - A \geq 0$, i.e. $A \leq I$.
($\Leftarrow$) If $A \leq I$ (equivalently $I - A \geq 0$), then for any unit vector x,

$$\langle x, Ax\rangle \leq \langle x, Ix\rangle = 1.$$

Thus, $\|A\| = \sup_{\|x\|=1}\langle x, Ax\rangle \leq 1$.
Overall we have, $\|A\| \leq 1 \iff A \leq I$. □

Exercise A6.8

Let A be a positive matrix. Define a superoperator (linear operator on matrices) by the same equation $\mathcal{A}(X) \equiv AX$. Show that $\mathcal{A}$ is positive with respect to the Hilbert-Schmidt inner product. That is, for all X, $\mathrm{tr}\big(X^\dagger \mathcal{A}(X)\big) \geq 0$. Similarly, show that the superoperator defined by $\mathcal{A}(X) \equiv XA$ is positive with respect to the Hilbert-Schmidt inner product on matrices.

Solution

Concepts Involved:

For any X, it must be that $X^\dagger X$ and $XX^\dagger$ are positive (Ex. 2.25). Further, any positive operator A satisfies $\mathrm{Tr}(A) \geq 0$ as its eigenvalues are all positive and so $\mathrm{Tr}(A) = \sum_i \lambda_i \geq 0$. We then observe that for any X:

$$\mathrm{Tr}\Big(X^\dagger \mathcal{A}(X)\Big) = \mathrm{Tr}\Big(X^\dagger AX\Big) = \mathrm{Tr}\Big(XX^\dagger A\Big) \geq 0$$

and also for $\mathcal{A}(X) \equiv XA$:

$$\mathrm{Tr}\Big(X^\dagger \mathcal{A}(X)\Big) = \mathrm{Tr}\Big(X^\dagger XA\Big) \geq 0$$

where we use that $XX^\dagger A$ and $X^\dagger XA$ are positive, being products of two positive operators. □