

Nielsen and Chuang Solutions

Jacob Watkins

December 2020

Chapter 1

Plan and Progress

0

1.1 Goals of this project

There exist other partial solution manuals to N&C, most of them on github. It appears, taken together, they have covered chapters 2,3,4 and 9 almost completely, with scattered solutions for other chapters. Here, we wish to help fill in the gaps, and perhaps ultimately create the most comprehensive solution manual to date. The strategy here is as follows:

- Create solution manuals for chapters 5-8
- Create solutions for chapters 10-12
- (If motivated) Fill in remaining problems in chapters already covered by others (in chapters 2-4 for example.)
- (If REALLY motivated) Compile together solutions already created, and bring them into a common format, so that we may come closer to a universal solutions manual!

To-do

- Fix QFT circuit, recreate it in Qiskit and set barrier option to True
- Double check circuit shown in problem 5.4

1.2 Progress so far

This progress was last updated February 1st, 2021.

- Chapter 2: Exercises 2.1-2.4 complete. Many exercises remain

- Chapter 5: Exercises 5.1-5.14 complete. Exercises 5.15-5.29 remain, along with 6 practice problems
- Chapter 6: Only first few problems completed
- Appendix 1: All exercises complete!
- Appendix 4: All exercises complete except A4.17 and part 2 of problem 1
- Appendix 5: Almost done. Problem 1 needs to be cleaned up.
- Appendix 6: All exercises complete!

Chapter 2

Introduction to quantum mechanics

2

Exercise 2.1: (Linear dependence: example)

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

Solution: By inspection, one might be able to see that $(2, 1) = (1, -1) + (1, 2)$, hence,

$$(1, -1) + (1, 2) - (2, 1) = 0. \quad (2.1)$$

This demonstrates linear dependence. A more systematic approach would be to cast this problem as a system of linear equations. A desired set of coefficients can be found by solving

$$\begin{pmatrix} 1 & 1 & 2 \\ -2 & 2 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0 \quad (2.2)$$

using standard techniques of linear algebra (such as row reduction).

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: Making the identification

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (2.3)$$

it is easy to see that the matrix representation M of A must be

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (2.4)$$

On the other hand, say we choose the output basis to be $|1\rangle, |0\rangle$ instead of $|0\rangle, |1\rangle$ (note that order matters). In this case, M would take the form of the identity matrix.

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 $|v_i\rangle, |w_j\rangle$, and $|x_k\rangle$ be bases for the vector spaces V, W , and X respectively. Show that the matrix representation for the linear transformation BA is the matrix product of the matrix representations for B and A , with respect to the appropriate bases.

Solution: Let us compute the action of BA on a basis vector $|v_i\rangle$.

$$\begin{aligned}
 BA|v_i\rangle &= B(A|v_i\rangle) \\
 &= B\left(\sum_j A_{ji}|w_j\rangle\right) \\
 &= \sum_j A_{ji}B|w_j\rangle \\
 &= \sum_j \sum_k A_{ji}B_{kj}|x_k\rangle \\
 &= \sum_k \left(\sum_j B_{kj}A_{ji}\right)|x_k\rangle \\
 &= \sum_k (BA)_{ki}|x_k\rangle.
 \end{aligned} \tag{2.5}$$

In the last step, we identified the coefficient as the product of the two matrix representations of A and B . This gives the result.

Exercise 2.4: (Matrix representation for the 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 representation is taken with respect to the same input and output bases. This matrix is known as the *identity matrix*.

Solution: Let $\mathbb{1}$ be the matrix representation of the identity operator I with respect to some basis $\{|v_i\rangle\}$. Then,

$$I|v_i\rangle = |v_i\rangle = \sum_j \delta_{ji}|v_j\rangle, \tag{2.6}$$

where δ_{ji} is the Kronecker delta. The coefficients δ_{ji} are exactly those of the identity matrix $\mathbb{1}$, hence the identity matrix is the matrix representation of I .

Exercise 2.5

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

Solution:

Chapter 3

The Quantum Fourier Transform and its applications

5

Exercise 5.1

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

Solution: It suffices to show that, for any two computational basis states $|j\rangle, |k\rangle$,

$$\langle j | (QFT)^\dagger (QFT) | k \rangle = \langle j | k \rangle = \delta_{ij}. \quad (3.1)$$

To do this, we substitute the definition into the above equation.

$$\begin{aligned} \langle k | (QFT)^\dagger (QFT) | j \rangle &= \left(\frac{1}{\sqrt{N}} \sum_{p=0}^{N-1} e^{-2\pi i k p / N} \langle p | \right) \left(\frac{1}{\sqrt{N}} \sum_{q=0}^{N-1} e^{2\pi i j q / N} | q \rangle \right) \\ &= \frac{1}{N} \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} e^{2\pi i (j q - k p) / N} \langle p | q \rangle \\ &= \frac{1}{N} \sum_{p=0}^{N-1} e^{2\pi i (j - k) p / N} \end{aligned} \quad (3.2)$$

where in the last step we used the orthonormality of the p, q states to eliminate one of the sums. Clearly, if $j = k$, the result is exactly one, as desired. Otherwise, $j - k$ is a nonzero integer, say n , such that $|n| < N$. We will show that in this case the sum above is zero.

The basic idea is that we are taking a sum over phases which are symmetrically distributed around the unit circle, so the result must be zero. To make this argument rigorous, multiply the sum by $e^{2\pi i n / N}$.

$$e^{2\pi i n / N} \sum_{p=0}^{N-1} (e^{2\pi i n / N})^p = \sum_{p=0}^{N-1} (e^{2\pi i n / N})^{p+1} = \sum_{p=1}^N (e^{2\pi i n / N})^p \quad (3.3)$$

In the last equation we simply reindexed. Because of the N -periodicity, $(e^{2\pi i n/N})^N = 1 = (e^{2\pi i n/N})^0$. Hence, we see that the sum is left unchanged by the multiplication. Since $e^{2\pi i n/N} \neq 1$, the sum must in fact be zero. This completes the proof.

Exercise 5.2

Explicitly compute the Fourier transform of the n qubit state $|00\dots 0\rangle$.

Solution: Suppose there are n qubits, so that $N = 2^n$. Using the definition given directly above in the textbook,

$$|00\dots 0\rangle \rightarrow \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} e^0 |k\rangle \quad (3.4)$$

$$= \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} |k\rangle, \quad (3.5)$$

which is simply a uniform superposition over the computational basis states. Evidently, the QFT on the zero state simply acts the same as Hadamards on all the qubits!

We remark that this result is consistent with the interpretation that the Fourier transform decomposes a “signal” into its frequency components. Here, the signal was a sharp spike, which requires a large spread in frequency to construct. Conversely, a uniform superposition without phases is like a constant function signal, which has a frequency of zero.

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). There are 2^n complex numbers we need to compute, which are the output amplitudes of the Fourier transform. If we compute each one using (5.1), each such amplitude involves a sum which contains 2^n terms. Thus, there will be $2^n \times 2^n = 2^{2n}$ summations and therefore at least as many arithmetic operations.

Solution: Let’s now consider a computation based on the factored form of the QFT, Equation (5.4). As before, this involves a computation of 2^n amplitudes, one for each bitstring $k = k_1 k_2 \dots k_n$. Using (5.4) the amplitude a_k corresponding to the state $|k\rangle$ is given by

$$\langle k | QFT | j \rangle = \frac{1}{2^{n/2}} (\delta_{k_1 0} + e^{2\pi i 0 \cdot j_n} \delta_{k_1 2}) (\delta_{k_2 0} + e^{2\pi i 0 \cdot j_{n-1} j_n} \delta_{k_2 2}) \dots (\delta_{k_n 0} + e^{2\pi i 0 \cdot j_1 \dots j_n} \delta_{k_n 2}). \quad (3.6)$$

where $|j\rangle$ is our input state. This involves a multiplication of n terms, hence there are $n \times 2^n$ total multiplications. This is a lower bound for the number of operations.

Exercise 5.4

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

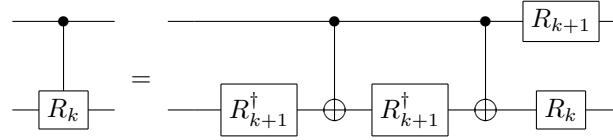
Solution: We use the *ABC* construction of Corollary 4.2 to make our controlled R_k according to Figure 4.6. First, note that

$$R_k = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{pmatrix} = e^{2\pi i/2^{k+1}} \begin{pmatrix} e^{-2\pi i/2^{k+1}} & 0 \\ 0 & e^{2\pi i/2^{k+1}} \end{pmatrix} = e^{i\alpha} R_z(\beta) \quad (3.7)$$

where $\alpha = 2\pi/2^{k+1}$ and $\beta = 2\pi/2^k$. Comparing this to the Euler decomposition formula of Theorem 4.1, we set $\gamma = \delta = 0$. Following through the steps, this implies,

$$\begin{aligned} A &= R_z(\beta) \\ B &= R_z(-\beta/2) \\ C &= R_z(-\beta/2) \end{aligned} \quad (3.8)$$

can be used in the *ABC* construction of R_k . As a final step, primarily one of cosmetics, we notice these gates are related to the R_k through global phases which cancel each other out. Thus, the following circuit implements the controlled- R_k , as is easy to verify.

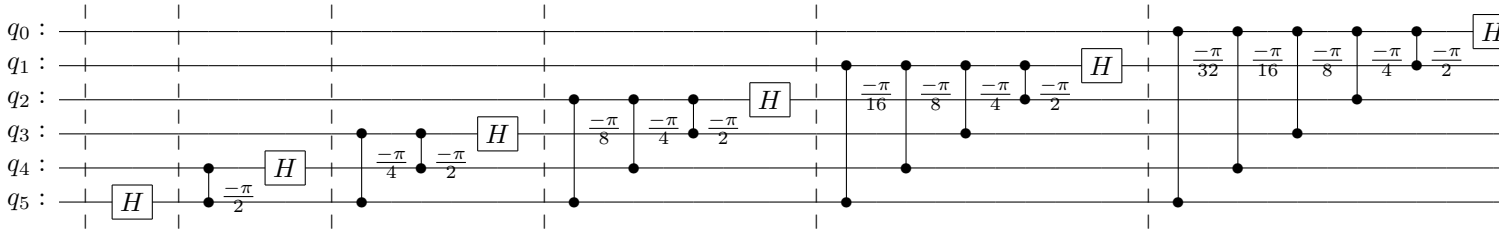


$$\text{Controlled-}R_k = R_{k+1}^\dagger \oplus R_{k+1}^\dagger \oplus R_k \quad (3.9)$$

Exercise 5.5

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

Solution: Here is the circuit for six qubits (generated using qiskit).



Here, the vertical line segments are the R_k gates, where the number alongside indicate angle of phase rotated. Like the inverse to any quantum circuit, can be obtained by reversing the order of the gates and taking the inverse of each gate. Note the Hadamard H is self-inverse.

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 $E(U, V) \equiv \max_{|\psi\rangle} \|(U - V)|\psi\rangle\|$ scales as $\Theta(n^2/p(n))$, and thus polynomial precision in each gate is sufficient to guarantee polynomial accuracy in the output state.

Solution: First we will show a more general result (one which will be further generalized in part 3 of the book, when discussing quantum channels). Let \mathcal{X} and \mathcal{Y} be quantum gates, and let X and Y be gates

which we will think of as approximating \mathcal{X} and \mathcal{Y} respectively. We will show that

$$E(\mathcal{X}\mathcal{Y}, XY) \leq E(\mathcal{X}, X) + E(\mathcal{Y}, Y). \quad (3.10)$$

To proceed, we set things up to make use of our friend: the triangle inequality. For any state $|\psi\rangle$ we have

$$\|(\mathcal{X}\mathcal{Y} - XY)|\psi\rangle\| = \|(\mathcal{X}\mathcal{Y} - X\mathcal{Y} + X\mathcal{Y} - XY)|\psi\rangle\| \quad (3.11)$$

$$= \|(\mathcal{X} - X)\mathcal{Y}|\psi\rangle + X(\mathcal{Y} - Y)|\psi\rangle\| \quad (3.12)$$

$$\leq \|(\mathcal{X} - X)\mathcal{Y}|\psi\rangle\| + \|X(\mathcal{Y} - Y)|\psi\rangle\| \quad (3.13)$$

Since this inequality holds for any $|\psi\rangle$, it certainly holds if we take the max of both sides. Hence,

$$E(\mathcal{X}\mathcal{Y}, XY) \leq \max_{|\psi\rangle} (\|(\mathcal{X} - X)\mathcal{Y}|\psi\rangle\| + \|X(\mathcal{Y} - Y)|\psi\rangle\|) \quad (3.14)$$

Certainly, the maximum of a sum $A + B$ is less than (or equal to) maximizing each individual piece A, B , so we may distribute the max function and maintain the inequality. Let's consider each term on the left hand side. Since \mathcal{Y} is unitary, it is a bijection on the space of valid states. Hence, we can maximize over $|\phi\rangle = \mathcal{Y}|\psi\rangle$ instead. Now consider the rightmost term. Since X is unitary, it preserves norm. Altogether,

$$E(\mathcal{X}\mathcal{Y}, XY) \leq \max_{|\phi\rangle} \|(\mathcal{X} - X)|\phi\rangle\| + \max_{|\psi\rangle} \|(\mathcal{Y} - Y)|\psi\rangle\| \quad (3.15)$$

$$= E(\mathcal{X}, X) + E(\mathcal{Y}, Y). \quad (3.16)$$

This places a bound on the error when composing imperfect gates. Moreover, this bound is tight, since if $X = \mathcal{X}$ and $Y = \mathcal{Y}$ we get strict equality. We can generalize this result, by simple induction, to arbitrary sequences of gates with corresponding approximations. Thus, in our present case, if each controlled R_k has precision Δ ,

$$E(U, V) \leq \frac{n(n+1)}{2} \Delta \in \Theta(n^2 \Delta). \quad (3.17)$$

Here, the use of Θ is appropriate rather than \mathcal{O} since our bound is tight.

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: Suppose there are t qubits in the first register, so the integer j can be expressed in binary as $j_{t-1} \dots j_1 j_0$, with $j_k \in \{0, 1\}$ for every $0 \leq k < t$. By definition, this means $j = j_0 + 2j_1 + \dots + 2^{t-1}j_{t-1}$. The state $|j\rangle$ has tensor product form

$$|j\rangle = \bigotimes_{k=0}^{t-1} |j_k\rangle \equiv |j_{t-1}\rangle \dots |j_0\rangle. \quad (3.18)$$

The action of the controlled- U^{2^k} controlled on the k th qubit is given by

$$|j_k\rangle|u\rangle \longrightarrow |j_k\rangle U^{2^k j_k}|u\rangle \quad (3.19)$$

as can be readily verified. Thus, the full sequence of controlled gates acts as follows.

$$\begin{aligned} |j\rangle|u\rangle &= \bigotimes_{k=0}^{t-1} |j_k\rangle|u\rangle \longrightarrow |j\rangle U^{2^{t-1}j_{t-1}} \dots U^{2^0 j_0}|u\rangle \\ &= |j\rangle U^{2^0 j_0 + \dots + 2^{t-1} j_{t-1}}|u\rangle \\ &= |j\rangle U^j|u\rangle \end{aligned} \quad (3.20)$$

In the last step we reused the definition of j being expressed in binary. This gives the desired result, which, as we see, did not rely on particular knowledge of the state $|u\rangle$.

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 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: If t is chosen as such, then $\tilde{\varphi}_u$ is an n -bit approximation to φ_u with probability $p_{succ} \geq (1 - \epsilon)$. Meanwhile, the probability of measuring $\tilde{\varphi}_u$ on the first register is given by the Born rule: $p_u = |c_u|^2$. These two events are independent, hence, the probability of measuring $\tilde{\varphi}_u$ and having it be an n -bit approximation is

$$p_u p_{succ} \geq |c_u|^2(1 - \epsilon). \quad (3.21)$$

Moreover, any other eigenstate $|v\rangle$ of U such that $\varphi_v \neq \varphi_u$ might still result in an n -bit approximation to φ_u , provided they are sufficiently close (it may even be that $\tilde{\varphi}_v = \tilde{\varphi}_u$). This will only further increase the probability of success. In any case, the right side of (??) remains a lower bound.

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 quantum 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 with Exercise 4.34.

Solution: If the eigenvalues of U are 1 and -1 , the corresponding phases are 0.0 and 0.1 respectively. Because these phases are finite bitstrings, there is no possibility of error and, in fact, we can take $t = n$, which in our case is 1. The inverse-QFT on one qubit is simply the Hadamard, and our circuit reduces to the following.



If a 0 (1) is measured, the final state is known to be in the plus (minus) subspace. The probability of each outcome is simply related to the initial overlap with each subspace via the Born rule.

Exercise 5.10

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

Solution: We proceed by exhaustive calculation.

$$\begin{aligned}
 5^1 \bmod 21 &= 5 \\
 5^2 \bmod 21 &= 4 \\
 5^3 \bmod 21 &= 20 \\
 5^4 \bmod 21 &= 16 \\
 5^5 \bmod 21 &= 17 \\
 5^6 \bmod 21 &= 1
 \end{aligned} \tag{3.23}$$

Hence, 6 is the smallest positive integer r such that $5^r \bmod 21 = 1$. Thus, 6 is the order of 5 modulo 21.

Exercise 5.11

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

Solution: Consider the set $\{x^n \bmod N\}_{n=1}^N$. Because we assume x and N share no common factors, it is not possible for $x^n \bmod N = 0$. Hence, $0 < x^n \bmod N < N$. By the pigeonhole principle, not all the values for $x^n \bmod N$ can be unique. There must exist some $x^i = x^j \bmod N$ for some $1 \leq i, j \leq N$ and $j \neq i$. Assume $j > i$ without loss of generality. Then, we have

$$\begin{aligned}
 x^j - x^i &= 0 \bmod N \\
 x^i(x^{j-i} - 1) &= 0 \bmod N.
 \end{aligned} \tag{3.24}$$

This implies $N | x^i(x^{j-i} - 1)$. But $N \nmid x^i$, again by assumption of no common factors. Hence, we must have $N | (x^{j-i} - 1)$, or $x^{j-i} = 1 \bmod N$. Since r is the smallest integer satisfying this condition, we must have $r \leq j - i < N$. Note this is a strict inequality, unlike what is given in the text.

Exercise 5.12

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

Solution: Since U acts as a map on the computational basis states, it suffices to show this map is injective (one-to-one). If $y \geq N$, then $|y\rangle$ is mapped to itself. On the other hand, if $0 \leq y < N$, then $|y\rangle$ is certainly mapped to some $|z\rangle$ where $z < N$. Hence, U is injective on the subspace where $y \geq N$, and it suffices to focus on the case where $y < N$. To this end, suppose $U|y\rangle = U|z\rangle$ for $y, z < N$. Then,

$$\begin{aligned}
 xy \bmod N &= xz \bmod N \\
 x(y - z) &= 0 \pmod{N}
 \end{aligned} \tag{3.25}$$

This implies N divides $x(y - z)$. However, since N and x are coprime, they share no common factors. This implies $N | (y - z)$. But $|y - z| < N$, so it must be that $y - z = 0$. Hence

$$|y\rangle = |z\rangle \tag{3.26}$$

which proves that U is injective on the basis, as desired.

Exercise 5.13

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

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i s k / r} |u_s\rangle = |x^k \bmod N\rangle \quad (3.27)$$

Solution: Let us crank the wheel: putting in the definition of $|u_s\rangle$ to the left side of (5.44) and regrouping.

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i s k / r} |u_s\rangle = \frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i s k / r} \frac{1}{\sqrt{r}} \sum_{t=0}^{r-1} e^{-2\pi i s t / r} |x^t \bmod N\rangle \quad (3.28)$$

$$= \frac{1}{r} \sum_{s=0}^{r-1} \sum_{t=0}^{r-1} e^{2\pi i s (k-t) / r} |x^t \bmod N\rangle \quad (3.29)$$

$$= \frac{1}{r} \sum_{t=0}^{r-1} |x^t \bmod N\rangle \left(\sum_{s=0}^{r-1} e^{2\pi i s (k-t) / r} \right) \quad (3.30)$$

Making gracious use of the hint, we see the rightmost sum on the last line equals $r \delta_{kt}$. Hence,

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i s k / r} |u_s\rangle = \sum_{t=0}^{r-1} \delta_{tk} |x^t \bmod N\rangle \quad (3.31)$$

$$= |x^k \bmod N\rangle. \quad (3.32)$$

To obtain (5.44) from the text, simply set $k = 0$.

Exercise 5.14

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

$$|\psi\rangle = \sum_{j=0}^{2^t-1} |j\rangle U^j |1\rangle = \sum_{j=0}^{2^t-1} |j\rangle |x^j \bmod N\rangle \quad (3.33)$$

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

$$V |j\rangle |k\rangle = |j\rangle |k + x^j \bmod N\rangle, \quad (3.34)$$

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

Solution: It is obvious that, by setting $k = 0$, we obtain the same output using V as we would if $k = 1$ and we had acted with U . To construct V we can simply use the U^j as before and store the result in an ancillary register whose initial value was one, then *add* that result to the register initialized to 0. The addition step can be carried out bitwise, and only takes $O(L^2)$ gates naively. Hence, the algorithm remains $O(L^3)$ after this addition step.

Chapter 4

Quantum Search Algorithms

6

Exercise 6.1

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

Solution: For $|x\rangle = |0\rangle$,

$$(2|0\rangle\langle 0| - I)|0\rangle = 2|0\rangle - |0\rangle = |0\rangle \quad (4.1)$$

Meanwhile, for $|x\rangle \neq |0\rangle$,

$$(2|0\rangle\langle 0| - I)|x\rangle = 2|0\rangle\langle 0|x\rangle - |x\rangle = -|x\rangle \quad (4.2)$$

Altogether,

$$(2|0\rangle\langle 0| - I)|x\rangle = (-1)^{\delta_{0x}}|x\rangle \quad (4.3)$$

Exercise 6.2

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

$$\sum_k (-\alpha_k + 2\langle\alpha\rangle) |k\rangle \quad (4.4)$$

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

Solution: By linearity,

$$\begin{aligned} (2|\psi\rangle\langle\psi| - I) \sum_k \alpha_k |k\rangle &= \sum_k 2\alpha_k (|\psi\rangle\langle\psi|) |k\rangle - \sum_k \alpha_k |k\rangle \\ &= 2|\psi\rangle \sum_k \alpha_k \langle\psi|k\rangle - \sum_k \alpha_k |k\rangle \end{aligned}$$

Because $|\psi\rangle$ is uniform superposition over the computational basis states, for all k we have $\langle\psi|k\rangle = 1/\sqrt{N}$. Hence,

$$(2|\psi\rangle\langle\psi| - I) \sum_k \alpha_k |k\rangle = \frac{2}{\sqrt{N}} \left(\sum_k \alpha_k \right) |\psi\rangle - \sum_k \alpha_k |k\rangle \quad (4.5)$$

$$= 2\sqrt{N}\langle\alpha\rangle |\psi\rangle - \sum_k \alpha_k |k\rangle \quad (4.6)$$

Finally, we expand out the definition of $|\psi\rangle$ and cancel the factors of \sqrt{N} . This gives our result.

$$2\sqrt{N}\langle\alpha\rangle |\psi\rangle - \sum_k \alpha_k |k\rangle = \sum_k (2\langle\alpha\rangle - \alpha_k) |k\rangle \quad (4.7)$$

Exercise 6.3

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

$$G = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (4.8)$$

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

$$\sin \theta = \frac{2\sqrt{M(N-M)}}{N} \quad (4.9)$$

Solution: As discussed in the text, both the oracle O and the reflection $2|\psi\rangle\langle\psi| - I$ leave the subspace $V = \text{span}(|\alpha\rangle, |\beta\rangle)$ invariant. Hence, so does the product G . Therefore, we will from here on speak of G only in terms of its action on the 2-dimensional subspace $V = \text{span}(|\alpha\rangle, |\beta\rangle)$, and consider the matrix representation in the orthonormal basis $\{|\alpha\rangle, |\beta\rangle\}$. This representation is unitary (since G itself is), and in fact it is orthogonal, since both O and $(2|\psi\rangle\langle\psi| - I)$ have real matrix elements. More specifically, G is *special* orthogonal, meaning it has determinant one, because it is the product of two reflections. All of this implies G is a proper rotation in the plane, and any such matrix may be parametrized as equation (??) for *some* angle θ . It remains to show θ satisfies relation (??). To do this, we simply compute the matrix element $\langle\beta|G|\alpha\rangle$.

$$\begin{aligned} \langle\beta|G|\alpha\rangle &= \langle\beta|(2|\psi\rangle\langle\psi| - I)O|\alpha\rangle \\ &= \langle\beta|(2|\psi\rangle\langle\psi| - I)|\alpha\rangle \\ &= \langle\beta|(2|\psi\rangle\langle\psi|\alpha\rangle - |\alpha\rangle) \\ &= 2\langle\beta|\psi\rangle\langle\psi|\alpha\rangle, \end{aligned} \quad (4.10)$$

where, along the way, we used the orthogonality of $|\alpha\rangle, |\beta\rangle$ and the fact that $O|\alpha\rangle = |\alpha\rangle$. Finally, using the expression given in the text for $|\psi\rangle$ expanded in the $|\alpha\rangle, |\beta\rangle$ basis, we arrive at our result.

$$\langle\beta|G|\alpha\rangle = \sin \theta = \frac{2\sqrt{M(N-M)}}{N} \quad (4.11)$$

Note that, in fact, we did not require the assumption that $M \leq N/2$ in our derivation.

Appendix 1: Notes on basic probability theory

Exercise A1.1

Prove Bayes' rule.

Solution: From the definition of conditional probability, we have

$$p(x, y) = p(y|x)p(x) = p(x|y)p(y) \quad (4.12)$$

Rearranging the last of these equations gives the desired result.

Exercise A1.2

Prove the law of total probability.

Solution: We start with the notion that, in the joint probability distribution for (X, Y) , one sums over all outcomes of X to get a probability distribution on Y alone.

$$p(y) = \sum_x p(x, y) \quad (4.13)$$

We arrive at our result by noting that, from the definition of conditional probability, $p(x, y) = p(y|x)p(x)$.

Exercise A1.3

Prove there exists a value of $x \geq \mathbf{E}(X)$ such that $p(x) > 0$.

Solution: Suppose, for sake of contradiction, that every value x of X with nonzero probability has the property $x < \mathbf{E}(X)$. Intuitively, we'd expect that the expectation value would have to be less than $\mathbf{E}(X)$. Indeed, using the inequality in the definition of expectation value,

$$\mathbf{E}(X) = \sum_{x \in X} xp(x) < \mathbf{E}(X) \sum_{x \in X} p(x) = \mathbf{E}(X) \quad (4.14)$$

Hence, $\mathbf{E}(X) < \mathbf{E}(X)$, a clear contradiction. We conclude our premise was false, hence there does exist a value of $x \in X$ such that $x \geq \mathbf{E}(X)$ and $p(x) > 0$.

Exercise A1.4

Prove that $\mathbf{E}(X)$ is linear in X .

Solution: The following computation gives us the result.

$$\begin{aligned}
 \mathbf{E}(aX + bY) &= \sum_{(x,y) \in (X,Y)} (ax + by)p(x, y) \\
 &= \sum_{x \in X} \sum_{y \in Y} axp(x, y) + byp(x, y) \\
 &= \sum_{x \in X} ax \sum_{y \in Y} p(x, y) + \sum_{y \in Y} by \sum_{x \in X} p(x, y) \\
 &= a \sum_{x \in X} xp(x) + b \sum_{y \in Y} yp(y) \\
 &= a\mathbf{E}(X) + b\mathbf{E}(Y).
 \end{aligned} \tag{4.15}$$

Here, a, b are constants. Along the way, we used $p(x) = \sum_y p(x, y)$ and the definition of expectation value.

Exercise A1.5

Prove that for independent random variables X and Y , $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$.

Solution: Recall that, for independent random variables, the joint probability distribution breaks into a product of individual probabilities. This yields the following computation.

$$\begin{aligned}
 \mathbf{E}(XY) &= \sum_x \sum_y xy p(x, y) \\
 &= \sum_x \sum_y xy p(x)p(y) \\
 &= \sum_x p(x) \sum_y p(y) \\
 &= \mathbf{E}(X)\mathbf{E}(Y)
 \end{aligned} \tag{4.16}$$

Exercise A1.6

Prove Chebyshev's inequality.

Solution: Our solution is taken from Kliesch and Roth, 2021. We start by proving a more fundamental result: *Markov's inequality*. Let Y be a nonnegative random variable, and $t > 0$. Then

$$p(Y \geq t) \leq \frac{\mathbf{E}(Y)}{t}. \tag{4.17}$$

To show this, let Ω be the set of outcomes over which our random variable Y is defined. Consider the indicator function $\mathbf{1}_A$ for some $A \subset \Omega$, defined as follows.

$$\mathbf{1}_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases} \tag{4.18}$$

Take the particular choice $A = \{\omega \in \Omega | Y(\omega) \geq t\}$, one can observe that

$$t\mathbf{1}_A(\omega') \leq Y(\omega') \quad (4.19)$$

for any $\omega' \in \Omega$. Taking the expectation value of our results gives us Markov's inequality.

To obtain Chebyshev's inequality, let $Y = |X - \mathbf{E}(X)|^2$ for some probability distribution X , and let $\lambda^2 = t/\Delta(X)^2$. Note that $\mathbf{E}(Y) = \Delta(X)^2$. Making these substitutions for t and Y gives

$$p(|X - \mathbf{E}(X)|^2 \geq \lambda^2 \Delta(X)^2) \leq \frac{1}{\lambda^2} \quad (4.20)$$

$$p(|X - \mathbf{E}(X)| \geq \lambda \Delta(X)) \leq \frac{1}{\lambda^2} \quad (4.21)$$

This proves our result.

Appendix 4: Number theory

Exercise A4.1 (Transitivity)

Show that if $a|b$ and $b|c$, then $a|c$.

Solution: The premises imply, by definition, that there exist integers j, k such that $b = aj$ and $c = bk$. Substituting the first of these two equations into the second, we see $c = ajk = al$, where $l = jk \in \mathbb{Z}$. Hence, there exists an integer, namely l , such that $c = al$, proving $a|c$.

Exercise A4.2

Show that if $d|a$ and $d|b$ then d also divides a linear combination of a and b , $ax + by$, where x and y are integers.

Solution: From the definition of $d|a$ and $d|b$, there exist integers j, k such that $a = dj$ and $b = dk$. Hence,

$$ax + by = djx + dky = d(jx + ky). \quad (4.22)$$

Define $m = jx + ky \in \mathbb{Z}$. Then we see $ax + by = dm$. From the definition of dividing, we have $d|(ax + by)$.

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: Suppose that $a|b$. Then there exists some $k \in \mathbb{Z}$ such that $b = ak$. By the hypothesis that a and b are positive, it must be that $k > 0$. Hence, $k - 1 \geq 0$. This, of course, implies

$$b(k - 1) \geq 0 \quad (4.23)$$

since, again, b is nonnegative (positive, in fact). The result we desire comes from basic manipulations of inequalities.

$$\begin{aligned} b(k - 1) \geq 0 &\implies bk \geq b \\ &\implies a \geq b \end{aligned} \quad (4.24)$$

As an immediate corollary, if $a|b$ and $b|a$ (both being positive integers), we have $a \geq b$ and $b \geq a$. Of course, this implies $a = b$. Note that the assumption of positivity was crucial for the proof to hold, and indeed, it is easy to see how it can be broken if negative numbers are included.

Exercise A4.4

Find the prime factorizations of 697 and 36 300.

Solution: I do not pretend to solve these in any mechanical fashion. Looking online, I notice 697 is a product of 41 and 17. Each of these numbers are themselves prime, so the prime factorization is

$$697 = 41^1 17^1 \quad (4.25)$$

Unlike the first, the second number is easier to do in your head. We can pull out two factors of 10 and easily prime factor those. Meanwhile, 363 is divisible by 3, and then if you cared to memorize some perfect squares, $121 = 11^2$. Altogether.

$$36300 = 2^2 3^1 5^2 11^2. \quad (4.26)$$

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: For any integer $a \in [1, p - 1]$, a is coprime with p . Hence, a has an inverse modulo p . In the case of p^2 , the only integer which is not coprime with p^2 in the range $[0, p^2 - 1]$ is p itself. Every other integer in the range has a multiplicative inverse.

Exercise A4.6

Find the multiplicative inverse of 17 modulo 24.

Solution: We seek an positive integer $n < 24$ such that $17 * n = 1 \pmod{24}$. Without yet an efficient method, we can perform an exhaustive check by hand or with a computer. It turns out the answer is $n = 17$ itself.

Exercise A4.7

Find the multiplicative inverse of $n + 1$ modulo n^2 , where n is any integer greater than 1.

Solution: The answer, which might be reasonably guessed (or not). Is $n - 1$.

$$(n + 1)(n - 1) = n^2 - 1 = 1 \pmod{n^2}. \quad (4.27)$$

Exercise A4.8 (Uniqueness of the inverse)

Suppose b and b' are multiplicative inverses of a , modulo n . Prove that $b = b' \pmod{n}$.

Solution: If b and b' are both inverses of a , then $ab = ab' \pmod{n}$. This implies

$$a(b - b') = 0 \pmod{n}. \quad (4.28)$$

From this, we conclude $n|a(b-b')$. But we also know, by Corollary A4.4, that n and a are coprime. Hence, $n|(b-b')$, so $b = b' \pmod{n}$.

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: If the prime factorization of a and b are known, simply find the largest set (counting multiplicity) of shared prime factors.

The prime factorization of 6825 and 1430 are $3^1 5^1 15^1 419^1$ and $2^1 5^1 11^1 13^1$ respectively. The only shared prime factor is 5, hence this is also the gcd.

Exercise A4.10

What is $\varphi(187)$?

Solution: The prime factorization of 187 is 11×17 . Hence,

$$\varphi(187) = \varphi(17 \times 11) = \varphi(17)\varphi(11) = 16 \times 10 = 160 \quad (4.29)$$

Exercise A4.11

Problem: Prove that

$$n = \sum_{d|n} \varphi(d) \quad (4.30)$$

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: Follow the advice of the hint, suppose $n = p^\alpha$ where p is prime. The divisors of n are p^j , where $0 \leq j \leq \alpha$. Hence, starting from the right hand side,

$$\begin{aligned} \sum_{d|n} \varphi(d) &= \sum_{j=0}^{\alpha} \varphi(p^j) = 1 + \sum_{j=1}^{\alpha} p^{j-1}(p-1) \\ &= 1 + (p-1) \sum_{j=1}^{\alpha} p^{j-1} \\ &= 1 + \sum_{j=0}^{\alpha} p^j - \sum_{j=0}^{\alpha} p^{j-1} \\ &= p^\alpha, \end{aligned} \quad (4.31)$$

where, in the last step, all but p^α cancel from subtractions. This proves the result when n is a power of a prime.

To generalize the argument, we use the fundamental theorem of arithmetic, which says any $n \in \mathbb{Z}$ has a prime factorization.

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_m^{\alpha_m}. \quad (4.32)$$

Since all terms are powers of prime, they are coprime with each other, and we may use the multiplicative property of φ .

$$\begin{aligned} \varphi(n) &= \prod_{j=1}^m \varphi(p_j^{\alpha_j}) \\ &= \prod_{j=1}^m \sum_{k_j=0}^{\alpha_j} \varphi(p^{k_j}) \end{aligned} \quad (4.33)$$

In the second step we used the first result derived above for powers of primes. By repeated use of the distributive property, the sum and product in the second line of (??) can be reversed, and cast as a sum over m variables.

$$\begin{aligned} \prod_{j=1}^m \sum_{k_j=0}^{\alpha_j} \varphi(p^{k_j}) &= \sum_{k_1=0}^{\alpha_1} \sum_{k_2=0}^{\alpha_2} \dots \sum_{k_m=0}^{\alpha_m} \prod_{j=1}^m \varphi(p^{k_j}) \\ &= \sum_{k_1=0}^{\alpha_1} \sum_{k_2=0}^{\alpha_2} \dots \sum_{k_m=0}^{\alpha_m} \varphi(p^{k_1} p^{k_2} \dots p^{k_m}) \end{aligned} \quad (4.34)$$

A careful examination of this last equation reveals it is nothing more than a sum over all possible divisors d of n , expressed via the prime factorization. Hence,

$$\varphi(n) = \sum_{d|n} \varphi(d) \quad (4.35)$$

as desired.

Exercise A4.12

Verify that \mathbf{Z}_n^* forms a group of size $\varphi(n)$ under the operation of multiplication modulo n .

Solution: That \mathbf{Z}_n^* is a set of size $\varphi(n)$ follows directly from the definition of φ . Let $a, b \in \mathbf{Z}_n^*$, with inverses a^{-1}, b^{-1} . Then, the product ab has inverse $a^{-1}b^{-1}$, hence is in \mathbf{Z}_n^* (note the order doesn't matter since multiplication is commutative). Thus, the set is closed under the binary operator. Moreover, multiplication modulo n is associative. Finally, it is easy to see that $1 \in \mathbf{Z}_n^*$ (being its own inverse) and it acts as the identity operator. Of course inverses exist, by definition, therefore we have shown that \mathbf{Z}_n^* satisfies the properties of a group under multiplication modulo n .

Exercise A4.13

Let a be an arbitrary element of \mathbf{Z}_n^* . Show that $S \equiv \{1, a, a^2, \dots\}$ forms a subgroup of \mathbf{Z}_n^* , and that the size of S is the least value of r such that $a^r = 1 \pmod{n}$.

Solution: For any finite group G , if I take a single element $g \in G$ and generate a subset $S \subset G$ by repeatedly multiplying g by itself, the result will be a subgroup (when I include the induced binary operation).

More generally, I can have multiple generators g_1, g_2, \dots, g_m and the result will still be a subgroup. Note this does not hold for infinite groups such as \mathbb{Z} , unless we allow negative exponents.

If r is the smallest positive integer satisfying $a^r = 1 \pmod{n}$, it follows that each a^i is unique for $i = 0, 1, \dots, r-1$. Otherwise, $a^i = a^j$ for some $i, j < r$, which implies $a^{j-i} = 1$. This contradicts the assertion that r is the *least* such value. Hence, S has at least r values. In fact, it cannot have more than r unique values, since for any $k > r$ we have

$$k = qr + i \quad (4.36)$$

for some $q \in \mathbb{Z}^+$ and $i < r$. But this will give the same power of a as i does.

$$a^k = a^{qr+i} = (a^r)^q a^i = 1^q a^i = a^i \quad (4.37)$$

Here all powers are taken modulo n . Thus, S has r elements.

Exercise A4.14

Suppose g is a generator for \mathbf{Z}_n^* . Show that g must have order $\varphi(n)$.

Solution: If g generates \mathbf{Z}_n^* , then every $a \in \mathbf{Z}_n^*$ must be some power of g . Hence, \mathbf{Z}_n^* is cyclic. By the results from the previous exercise, the size of \mathbf{Z}_n^* , which is $\varphi(n)$ must equal the order of the generator g .

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 alternative proof of Theorem A4.9, that is, show that $a^{\varphi(n)} = 1 \pmod{n}$ for any $a \in \mathbf{Z}_n^*$.

Solution: Consider the subgroup $A \subset G$ generated by a . Then the size of A is the order of a , say, r . By Lagrange's theorem, r must divide $\varphi(n)$, the size of \mathbf{Z}_n^* . That is, $\varphi(n) = kr$ for some $k \in \mathbb{Z}^+$. Given this,

$$a^{\varphi(n)} = a^{kr} = (a^r)^k = 1^k = 1 \quad (4.38)$$

where all values are taken modulo n . This proves Euler's generalization of the little theorem.

Exercise A4.16

Use Theorem A4.9 to show that the order of x modulo N must divide $\varphi(N)$.

Solution: This follows directly from Lagrange's theorem (see the previous cluster of exercises). We've already shown that the size of a cyclic subgroup of \mathbf{Z}_n^* is the order of a generating element. Lagrange's theorem says this order r must divide the size of the larger group \mathbf{Z}_n^* , which is $\varphi(n)$.

Exercise A4.17 (Reduction of order-finding to factoring)

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:

Exercise A4.18

Find the continued fraction expansion for $x = 19/17$ and $x = 77/65$

Solution: In both cases we apply the repeated fraction algorithm. The case $x = 19/17$ is only a few steps.

$$\frac{19}{17} = 1 + \frac{2}{17} = 1 + \frac{1}{\frac{17}{2}} = 1 + \frac{1}{8 + \frac{1}{2}} \quad (4.39)$$

For the case $x = 77/65$, we have to work a little harder. Here are the intermediate steps.

$$\begin{aligned} 77/65 &= 1 + 12/65 \\ 65/12 &= 5 + 5/12 \\ 12/5 &= 2 + 2/5 \\ 5/2 &= 2 + 1/2. \end{aligned} \quad (4.40)$$

Hence the result is

$$\frac{77}{65} = 1 + \frac{1}{5 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2}}}} \quad (4.41)$$

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: As the hint suggests, we proceed by induction on n . In the case $n = 1$, using the definitions provided in the text,

$$q_1 p_0 - p_1 q_0 = a_1 a_0 - (1 + a_0 a_1)1 = -1 \quad (4.42)$$

as desired. By inductive hypothesis, assume the statement holds for $n = m$. Then, using the recursive definition for p and q ,

$$q_{m+1} p_m - p_{m+1} q_m = (a_{m+1} q_m + q_{m-1}) p_m - (a_{m+1} p_m + p_{m-1}) q_m \quad (4.43)$$

$$= \cancel{a_{m+1} q_m p_m} + q_{m-1} p_m - \cancel{a_{m+1} p_m q_m} - p_{m-1} q_m \quad (4.44)$$

$$= -(q_m p_{m-1} - p_m q_{m-1}) \quad (4.45)$$

$$= (-1)^m + 1, \quad (4.46)$$

where in the last step we invoked the inductive hypothesis. Hence, the statement also holds for $n = m + 1$. By induction, the statement holds for all $n \geq 1$.

Note that the result may be reexpressed as

$$(-1)^n (q_n p_{n-1} - p_n q_{n-1}) = 1. \quad (4.47)$$

By Theorem A4.2, we must have $\gcd(p_n, q_n) = 1$.

Problem 4.1 (Prime number estimate)

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 $\lim_{n \rightarrow \infty} \pi(n) \log(n)/n = 1$ and thus $\pi(n) \approx n/\log(n)$. This problem gives a poor man's version of the prime number theorem which gives a pretty good lower bound on the distribution of prime numbers.

(1) Prove that $n \leq \log \binom{2n}{n}$.

Solution to (1): Note this is equivalent proving $2^n \geq \binom{2n}{n}$ (note the logarithm is base two). By definition,

$$\binom{2n}{n} = \frac{2n!}{n!n!} = \prod_{i=1}^n \frac{n+i}{i}. \quad (4.48)$$

Moreover, for each i in the product, $(n+i)/i = 1 + n/i \geq 2$. Hence,

$$\prod_{i=1}^n \frac{n+i}{i} \geq \prod_{i=1}^n 2 = 2^n. \quad (4.49)$$

This proves the result.

(2) Show that

$$\log \binom{2n}{n} \leq \sum_{p \leq 2n} \left\lfloor \frac{\log(2n)}{\log p} \right\rfloor \log p \quad (4.50)$$

where the sum is over all primes p less than or equal to $2n$.

Solution to (2): This one is hard! I could rewrite the problem as showing

$$\binom{2n}{n} \leq \prod_{p < 2n} e^{\log p \left\lfloor \frac{\log(2n)}{\log p} \right\rfloor} \quad (4.51)$$

(3) Use the previous two results to show that

$$\pi(2n) \geq \frac{n}{\log(2n)} \quad (4.52)$$

Solution to (3): From the previous two parts, we have

$$n \leq \sum_{p \leq 2n} \left\lfloor \frac{\log(2n)}{\log p} \right\rfloor \log p. \quad (4.53)$$

Moreover, for any two positive real numbers x and y ,

$$\lfloor x \rfloor y \leq \lfloor xy \rfloor. \quad (4.54)$$

Hence,

$$\begin{aligned} \sum_{p \leq 2n} \left\lfloor \frac{\log(2n)}{\log p} \right\rfloor \log p &\leq \sum_{p \leq 2n} \left\lfloor \frac{\log(2n)}{\log p} \log p \right\rfloor \\ &\leq \sum_{p \leq 2n} \lfloor \log(2n) \rfloor \\ &\leq \log(2n) \pi(2n) \end{aligned} \quad (4.55)$$

From this, we have $n \leq \log(2n) \pi(2n)$, and we have our result from rearranging.

Appendix 5: Public key cryptography and the RSA cryptosystem

Exercise A5.1

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 few 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: There is choice in how we represent the letters, but a natural one is to label from 1 to 26. In this representation, we have

$$\begin{aligned} Q &= 17 = 10001 \\ U &= 21 = 10101 \\ A &= 01 = 00001 \\ N &= 14 = 01110 \\ T &= 20 = 10100 \\ M &= 13 = 01101. \end{aligned} \tag{4.56}$$

In our encoding, the message is 35 bits in length, and given by

$$S = 10001101010000101110101001010101101 \tag{4.57}$$

$$= 18959782573 \tag{4.58}$$

where we converted to decimal in the last step. Next we choose an odd number e relatively prime to $\phi(n) = (p-1)(q-1) = 20$. We will choose $e = 9$. To compute the multiplicative inverse modulo 20, d , we employ Euler's algorithm. Following the steps outlined in appendix 4,

$$\begin{aligned} 20 &= 2 \times 9 + 2 \\ 9 &= 4 \times 2 + 1 \\ 2 &= 2 \times 1. \end{aligned} \tag{4.59}$$

Now we back substitute to find coefficients x, y such that $1 = 9x + 20y$.

$$\begin{aligned} 1 &= 9 - 4 \times 2 \\ &= 9 - 4 \times (20 - 2 \times 9) \\ &= 9 - 4 \times 20 + 8 \times 9 \\ &= 9 \times 9 - 4 \times 20. \end{aligned} \tag{4.60}$$

Reading off the coefficient, we can readily see that 9 is its own inverse modulo 20. Hence $d = e = 9$.

Alas, with such a small n we can only encode in 5 bit chunks. We'll therefore simply encode each letter separately. We have

$$E(Q) = 17^9 \pmod{33} = 02 = 00010 \quad (4.61)$$

$$E(U) = 21^9 \pmod{33} = 21 = 10101 \quad (4.62)$$

$$E(A) = 01^9 \pmod{33} = 01 = 00001 \quad (4.63)$$

$$E(N) = 14^9 \pmod{33} = 26 = 00010 \quad (4.64)$$

$$E(T) = 20^9 \pmod{33} = 05 = 00101 \quad (4.65)$$

$$E(M) = 13^9 \pmod{33} = 28 = 11100 \quad (4.66)$$

$$(4.67)$$

You can readily check, as expected, that taking the encoded message to the power of 9 (in 5 bit chunks) gets you back to the original message.

Exercise A5.2

Show that d is also an inverse of e modulo r , and thus $d = d' \pmod{r}$.

Solution: We will prove a somewhat more general result, namely if $ab = 1 \pmod{n}$ and $d|n$, then $ab = 1 \pmod{d}$. Indeed, the first statement implies $ab = qn + 1$ for some $q \in \mathbb{Z}$. On the other hand, since $d|n$, there is an integer k such that $n = dk$. Using these relations, we have $ab = q(dk) + 1 = (qk)d + 1$. Thus, $ab = 1 \pmod{d}$.

This solves the exercise when we recognize that $de = 1 \pmod{\phi(n)}$ and $r|\phi(n)$. A result from the previous appendix shows that the two inverses d and d' are equivalent modulo r .

Problem 5.1:

Write a computer program for performing encryption and decryption using the RSA algorithm. Find a pair of 20 bit prime numbers and use them to encrypt a 40 bit message.

Solution: I will first write pseudocode, then give an actual implementation in a common language such as python. Here is some pseudocode for the two major subroutines employed: RandomPrime and InverseMod.

Algorithm 1: RSA algorithm for public key cryptography

1 RSA algorithm (L, M) ;

Input : An integer L specifying bit length of primes, and a $2L$ -bit message M .

Output: A public key $P = (e, n)$ and private key $M = (d, n)$.

2 $p = \text{RandomPrime}(L)$;

3 $q = \text{RandomPrime}(L)$;

4 $n = pq$;

5 $\varphi = (p - 1)(q - 1)$;

6 $d = \text{InverseMod}(e, \varphi)$;

7 $P = (e, n)$;

8 $S = (d, n)$;

9 return P, S

Algorithm 2: Algorithm for producing random prime p of given length.

```
1 RandomPrime ( $L$ );  
   Input : An integer  $L$  specifying the bit length of the desired prime  
   Output: A random prime  $p$  of that length  
2  $p = \text{RandomInt}(L)$ ;  
3 do  
4    $p = \text{RandomInt}(L)$ ;  
5 while not prime( $p$ );
```

Appendix 6: 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: We will first prove that positivity is preserved under conjugation with X . That is, if P is positive, so is XPX^\dagger . Taking the adjoint shows that XPX^\dagger is hermitian.

$$(XPX^\dagger)^\dagger = (X^\dagger)^\dagger P^\dagger X^\dagger = XPX^\dagger \quad (4.68)$$

To prove positive-semidefiniteness, suppose λ is a (real) eigenvalue of XPX^\dagger , so that there is a normalized vector v such that.

$$XPX^\dagger v = \lambda v \quad (4.69)$$

Taking the inner product of both sides of this equation with v itself,

$$\begin{aligned} \langle v, XPX^\dagger v \rangle &= \langle v, \lambda v \rangle \\ \langle X^\dagger v, PX^\dagger v \rangle &= \lambda \langle v, v \rangle \\ \langle u, Pu \rangle &= \lambda, \end{aligned} \quad (4.70)$$

where $u = X^\dagger v$. Because P is positive semidefinite, we see that $\lambda \geq 0$. Hence, every eigenvalue of XPX^\dagger is nonnegative. This proves our result.

Exercise A6.2

Prove that $A \geq 0$ if and only if A is a positive operator.

Solution: If $A \geq 0$, then $A - 0$ is positive semidefinite, hence so is A . Conversely, if A is positive, so is $A - 0$, and thus $A \geq 0$.

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: Let's start by proving transitivity. If $A \leq B$ and $B \leq C$, then $B - A$ and $C - B$ are positive matrices. Hence so is their sum, $C - A$. This implies $A \leq C$ by definition.

To prove asymmetry, suppose $A \leq B$ and $B \leq A$. Let λ be an eigenvalue of $A - B$. It is then clear that $B - A$ must have eigenvalue $-\lambda$. By assumption of positive semidefiniteness of $A - B$ and $B - A$ we must have

$$\lambda \leq 0 \quad \lambda \geq 0. \quad (4.71)$$

Hence, $\lambda = 0$. Thus every eigenvalue of $A - B$ is zero, so $A - B = 0$. This proves asymmetry.

Finally, we note that $A - A = 0$ is positive semidefinite. Thus, $A \leq A$, proving the reflexive property.

Exercise A6.4

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

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad (4.72)$$

$$\|A\| = 3/2 > 1 = \lambda$$

Solution: (1) Since A is a matrix, its set of eigenvalues is finite. Hence, there exists an eigenvalue λ_m such that $|\lambda_m| = \lambda$. Let $|u_m\rangle$ be the corresponding eigenvector, normalized. Then,

$$|\langle u_m | A | u_m \rangle| = |\lambda_m \langle u_m | u_m \rangle| = \lambda \quad (4.73)$$

Since $\|A\|$ is the maximum over all such inner products, it is certainly at least as big as the value set by $|u\rangle = |u_m\rangle$. Hence, $\|A\| \geq \lambda$.

(2) Using part (1), it suffices to show that $\|A\| \leq \lambda$ for hermitian A . If $|u\rangle$ is a normalized state, it can be expressed as a linear combination in an orthonormal basis defined by the eigenstates of A .

$$|u\rangle = \sum_i c_i |\lambda_i\rangle \quad (4.74)$$

Here, $|\lambda_i\rangle$ is an eigenstate of A with eigenvalue λ_i . Computing the inner product as in the definition of $\|A\|$,

$$|\langle u | A | u \rangle| = \left| \sum_i \lambda_i |c_i|^2 \right| \leq \sum_i |\lambda_i| |c_i|^2 \leq \lambda \sum_i |c_i|^2 = \lambda \quad (4.75)$$

Along the way, we used the triangle inequality, the fact that $|\lambda_i| \leq \lambda$, and the normalization of $|u\rangle$. Since λ is an upper bound for every $|u\rangle$, it is also an upper bound for the maximum, which is precisely $\|A\|$. Thus, $\|A\| \leq \lambda$, which combined with the previous result gives $\|A\| = \lambda$.

(3) A has a single eigenvalue $\lambda = 1$ with eigenvector $|\lambda\rangle = (0, 1)^T$. Hence, $\lambda = 1$. On the other hand, for some normalized $(a, b) \in \mathbb{C}^2$,

$$(a^* \quad b^*) \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = |a|^2 + |b|^2 + ab^* \quad (4.76)$$

$$= 1 + ab^* \quad (4.77)$$

If $a = b = 1/\sqrt{2}$, then this value is $3/2$. On the other hand, for more arbitrary complex values of a and b , the triangle inequality puts $3/2$ as an upper bound on the magnitude.

$$|1 + ab^*| \leq 1 + |a||b| \leq 3/2. \quad (4.78)$$

Hence, $\|A\| = 3/2$, and we have our result.

Exercise 6.5: (AB and BA have the same eigenvalues)

Prove that AB and BA have the same eigenvalues. (*Hint:* For invertible A , show that $\det(xI - AB) = \det(xI - BA)$, and thus the eigenvalues of AB and BA are the same. By continuity this holds even when A is not invertible.

Solution: As the hint suggests, first suppose A is invertible. Because the eigenvalues of AB and BA are the zeros of the characteristic polynomial, showing these two polynomials are the same amounts to proving the statement. And these polynomials are precisely the determinants shown in the hint.

We will use the property that the determinant is indifferent to the permutation of matrices in a product.

$$\det(xI - AB) = \det(I(xI - AB)) \quad (4.79)$$

$$= \det(A^{-1}A(xI - AB)) \quad (4.80)$$

$$= \det(A^{-1}(xI - AB)A) \quad (4.81)$$

$$= \det(xI - BA) \quad (4.82)$$

This proves our result when A is invertible. If A is singular, then there exists an $\epsilon > 0$ such that

$$A' = A + \epsilon I \quad (4.83)$$

is invertible. Then the theorem carries over as before for A' , and to get the result for A we take $\epsilon \rightarrow 0$. This is valid since the determinant is only a polynomial in ϵ .

Exercise 6.6

Suppose A and B are such that AB is Hermitian. Using the previous two observations show that $\|AB\| \leq \|BA\|$.

Solution: Since AB is Hermitian, then $\|AB\| = |\lambda|$ for some eigenvalue of AB . By the previous exercise, λ is also an eigenvalue of BA , and by that same exercise we have $\lambda \leq \|BA\|$. Thus, $\|AB\| \leq \|BA\|$.

Exercise 6.7

Suppose A is positive. Show that $\|A\| \leq 1$ if and only if $A \leq I$.

Solution: (\implies) Suppose $\|A\| \leq 1$. Then every eigenvalue λ of A is such that $\lambda \in [0, 1]$. This implies the eigenvalues of $I - A$, which are given by $1 - \lambda$ are also in this range. In particular, $I - A$ is positive, so $A \leq I$.

(\impliedby) Suppose $A \leq I$, so that $I - A$ is positive. As above, the eigenvalues of $I - A$ are $1 - \lambda$, where λ is an eigenvalue of A . Since both A and $I - A$ are positive, we have

$$1 - \lambda \geq 0 \quad \lambda \geq 0. \quad (4.84)$$

This implies $\lambda \leq 1$ for each λ , so we have $\|A\| \leq 1$.

Exercise 6.8

Let A be a positive matrix. Define a superoperator (linear operator on matrices) by the 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 , $\text{tr}(X^\dagger \mathcal{A}(X)) \geq 0$. 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: Suppose A is positive. Then for any matrix X , both XAX^\dagger and $X^\dagger AX$ are positive. Moreover, the trace of any positive matrix is itself positive. Therefore, the desired result comes from the simple fact that the condition for positivity of \mathcal{A} amounts to taking traces of the above matrices. For the first definition of \mathcal{A} ,

$$\text{tr}(X^\dagger \mathcal{A}(X)) = \text{tr}(X^\dagger AX). \quad (4.85)$$

For the second definition,

$$\text{tr}(X^\dagger \mathcal{A}(X)) = \text{tr}(X^\dagger XA) = \text{tr}(XAX^\dagger) \quad (4.86)$$