

# Nielsen and Chuang Solutions

Jacob Watkins

December 2020

## 1 Outline and plan

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!

## 2 To-do

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

## Chapter 5: The Quantum Fourier Transform and its applications

### 5.1

**Problem:** 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  $|j\rangle, |k\rangle$ ,

$$\langle j | (QFT)^\dagger (QFT) | k \rangle = \langle j | k \rangle = \delta_{ij}. \quad (1)$$

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

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

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  $e^{2\pi i n / N}$ .

$$e^{2\pi i n / N} \sum_{p=0}^{N-1} (e^{2\pi i n / N})^p = \sum_{p=0}^{N-1} (e^{2\pi i n / N})^{p+1} = \sum_{p=1}^N (e^{2\pi i n / N})^p \tag{3}$$

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.

## 5.2

**Problem:** 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,

$$|00\dots 0\rangle \rightarrow \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} e^0 |k\rangle \tag{4}$$

$$= \frac{1}{2^{n/2}} \sum_{k=0}^{2^n-1} |k\rangle, \tag{5}$$

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.

## 5.3

**Problem (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

$$\langle k | QFT | j \rangle = \frac{1}{2^{n/2}} (\delta_{k_1 0} + e^{2\pi i 0 \cdot j_n} \delta_{k_1 2}) (\delta_{k_2 0} + e^{2\pi i 0 \cdot j_{n-1} j_n} \delta_{k_2 2}) \dots (\delta_{k_n 0} + e^{2\pi i 0 \cdot j_1 \dots j_n} \delta_{k_n 2}). \quad (6)$$

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.

## 5.4

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

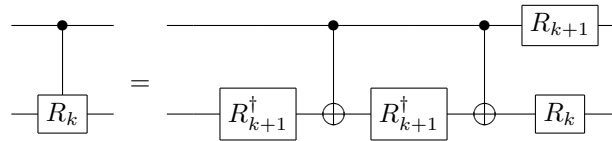
**Solution:** We use the *ABC* construction of Corollary 4.2 to make our controlled  $R_k$  according to Figure 4.6. First, note that

$$R_k = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i / 2^k} \end{pmatrix} = e^{2\pi i / 2^{k+1}} \begin{pmatrix} e^{-2\pi i / 2^{k+1}} & 0 \\ 0 & e^{2\pi i / 2^{k+1}} \end{pmatrix} = e^{i\alpha} R_z(\beta) \quad (7)$$

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,

$$\begin{aligned} A &= R_z(\beta) \\ B &= R_z(-\beta/2) \\ C &= R_z(-\beta/2) \end{aligned} \quad (8)$$

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.

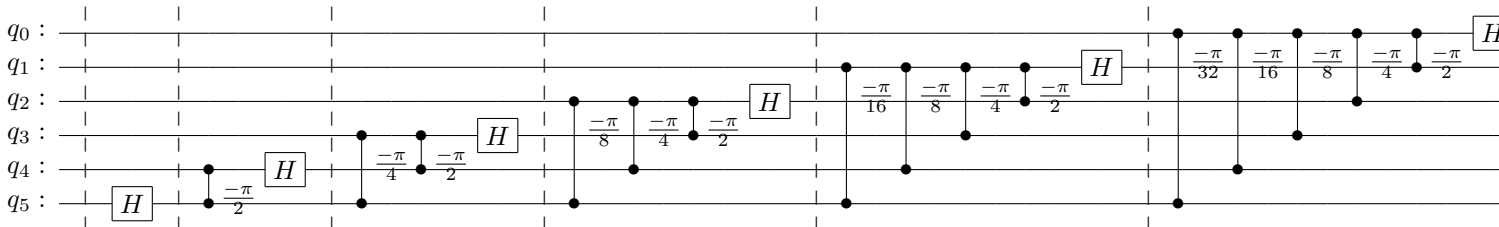


$$\text{Controlled-}R_k = \text{Controlled-}R_{k+1}^\dagger \text{CNOT} \text{Controlled-}R_{k+1}^\dagger \text{CNOT} \text{Controlled-}R_k \quad (9)$$

## 5.5

**Problem:** 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.

## 5.6

**Problem (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

$$E(\mathcal{X}\mathcal{Y}, XY) \leq E(\mathcal{X}, X) + E(\mathcal{Y}, Y). \quad (10)$$

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

$$\|(\mathcal{X}\mathcal{Y} - XY)|\psi\rangle\| = \|(\mathcal{X}\mathcal{Y} - X\mathcal{Y} + X\mathcal{Y} - XY)|\psi\rangle\| \quad (11)$$

$$= \|(\mathcal{X} - X)\mathcal{Y}|\psi\rangle + X(\mathcal{Y} - Y)|\psi\rangle\| \quad (12)$$

$$\leq \|(\mathcal{X} - X)\mathcal{Y}|\psi\rangle\| + \|X(\mathcal{Y} - Y)|\psi\rangle\| \quad (13)$$

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

$$E(\mathcal{X}\mathcal{Y}, XY) \leq \max_{|\psi\rangle} (\|(\mathcal{X} - X)\mathcal{Y}|\psi\rangle\| + \|X(\mathcal{Y} - Y)|\psi\rangle\|) \quad (14)$$

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,

$$E(\mathcal{X}\mathcal{Y}, XY) \leq \max_{|\phi\rangle} \|(\mathcal{X} - X)|\phi\rangle\| + \max_{|\psi\rangle} \|(\mathcal{Y} - Y)|\psi\rangle\| \quad (15)$$

$$= E(\mathcal{X}, X) + E(\mathcal{Y}, Y). \quad (16)$$

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  $\Delta$ ,

$$E(U, V) \leq \frac{n(n+1)}{2} \Delta \in \Theta(n^2 \Delta). \quad (17)$$

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

## 5.7

**Problem:** 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

$$|j\rangle = \bigotimes_{k=0}^{t-1} |j_k\rangle \equiv |j_{t-1}\rangle \dots |j_0\rangle. \quad (18)$$

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

$$|j_k\rangle |u\rangle \longrightarrow |j_k\rangle U^{2^k j_k} |u\rangle \quad (19)$$

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

$$\begin{aligned} |j\rangle |u\rangle &= \bigotimes_{k=0}^{t-1} |j_k\rangle |u\rangle \longrightarrow |j\rangle U^{2^{t-1}j_{t-1}} \dots U^{2^0 j_0} |u\rangle \\ &= |j\rangle U^{2^0 j_0 + \dots + 2^{t-1} j_{t-1}} |u\rangle \\ &= |j\rangle U^j |u\rangle \end{aligned} \quad (20)$$

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  $|u\rangle$ .

## 5.8

**Problem:** 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  $\varphi_u$  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  $\varphi_u$  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

$$p_u p_{succ} \geq |c_u|^2 (1 - \epsilon). \quad (21)$$

Moreover, any other eigenstate  $|v\rangle$  of  $U$  such that  $\varphi_v \neq \varphi_u$  might still result in an  $n$ -bit approximation to  $\varphi_u$ , 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 (21) remains a lower bound.

## 5.9

**Problem:** Let  $U$  be a unitary transform with eigenvalues  $\pm 1$ , 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.

## 5.10

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

**Solution:** We proceed by exhaustive calculation.

$$\begin{aligned}
 5^1 \bmod 21 &= 5 \\
 5^2 \bmod 21 &= 4 \\
 5^3 \bmod 21 &= 20 \\
 5^4 \bmod 21 &= 16 \\
 5^5 \bmod 21 &= 17 \\
 5^6 \bmod 21 &= 1
 \end{aligned}
 \tag{23}$$

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

## 5.11

**Problem:** 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

$$\begin{aligned}
 x^j - x^i &= 0 \bmod N \\
 x^i(x^{j-i} - 1) &= 0 \bmod N.
 \end{aligned}
 \tag{24}$$

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.

## 5.12

**Problem:** 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,

$$\begin{aligned}
 xy \bmod N &= xz \bmod N \\
 x(y - z) &= 0 \pmod{N}
 \end{aligned}
 \tag{25}$$

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

$$|y\rangle = |z\rangle \quad (26)$$

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

### 5.13

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

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i s k / r} |u_s\rangle = |x^k \bmod N\rangle \quad (27)$$

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

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i s k / r} |u_s\rangle = \frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i s k / r} \frac{1}{\sqrt{r}} \sum_{t=0}^{r-1} e^{-2\pi i s t / r} |x^t \bmod N\rangle \quad (28)$$

$$= \frac{1}{r} \sum_{s=0}^{r-1} \sum_{t=0}^{r-1} e^{2\pi i s (k-t) / r} |x^t \bmod N\rangle \quad (29)$$

$$= \frac{1}{r} \sum_{t=0}^{r-1} |x^t \bmod N\rangle \left( \sum_{s=0}^{r-1} e^{2\pi i s (k-t) / r} \right) \quad (30)$$

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

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i s k / r} |u_s\rangle = \sum_{t=0}^{r-1} \delta_{tk} |x^t \bmod N\rangle \quad (31)$$

$$= |x^k \bmod N\rangle. \quad (32)$$

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

### 5.14

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

$$|\psi\rangle = \sum_{j=0}^{2^t-1} |j\rangle U^j |1\rangle = \sum_{j=0}^{2^t-1} |j\rangle |x^j \bmod N\rangle \quad (33)$$

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

$$V |j\rangle |k\rangle = |j\rangle |k + x^j \bmod N\rangle, \quad (34)$$

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 6: Quantum Search Algorithms

### 6.1

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

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

$$(2|0\rangle\langle 0| - I)|0\rangle = 2|0\rangle - |0\rangle = |0\rangle \quad (35)$$

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

$$(2|0\rangle\langle 0| - I)|x\rangle = 2|0\rangle\langle 0|x\rangle - |x\rangle = -|x\rangle \quad (36)$$

Altogether,

$$(2|0\rangle\langle 0| - I)|x\rangle = (-1)^{\delta_{0x}}|x\rangle \quad (37)$$

### 6.2

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

$$\sum_k (-\alpha_k + 2\langle\alpha\rangle) |k\rangle \quad (38)$$

where  $\langle\alpha\rangle \equiv \sum_k \alpha_k / N$  is the mean value of the  $\alpha_k$ . 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,

$$(2|\psi\rangle\langle\psi| - I) \sum_k \alpha_k |k\rangle = \frac{2}{\sqrt{N}} \left( \sum_k \alpha_k \right) |\psi\rangle - \sum_k \alpha_k |k\rangle \quad (39)$$

$$= 2\sqrt{N}\langle\alpha\rangle |\psi\rangle - \sum_k \alpha_k |k\rangle \quad (40)$$

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

$$2\sqrt{N}\langle\alpha\rangle |\psi\rangle - \sum_k \alpha_k |k\rangle = \sum_k (2\langle\alpha\rangle - \alpha_k) |k\rangle \quad (41)$$

### 6.3

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

$$G = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad (42)$$



where  $\theta$  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

$$\sin \theta = \frac{2\sqrt{M(N-M)}}{N} \quad (43)$$

**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 (42) for *some* angle  $\theta$ . It remains to show  $\theta$  satisfies relation (43). To do this, we simply compute the matrix element  $\langle\beta|G|\alpha\rangle$ .

$$\begin{aligned} \langle\beta|G|\alpha\rangle &= \langle\beta|(2|\psi\rangle\langle\psi| - I)O|\alpha\rangle \\ &= \langle\beta|(2|\psi\rangle\langle\psi| - I)|\alpha\rangle \\ &= \langle\beta|(2|\psi\rangle\langle\psi|\alpha\rangle - |\alpha\rangle) \\ &= 2\langle\beta|\psi\rangle\langle\psi|\alpha\rangle, \end{aligned} \quad (44)$$

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.

$$\langle\beta|G|\alpha\rangle = \sin \theta = \frac{2\sqrt{M(N-M)}}{N} \quad (45)$$

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

### A1.1

**Problem:** Prove Bayes' rule.

**Solution:** From the definition of conditional probability, we have

$$p(x, y) = p(y|x)p(x) = p(x|y)p(y) \quad (46)$$

Rearranging the last of these equations gives the desired result.

### A1.2

**Problem:** 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.

$$p(y) = \sum_x p(x, y) \quad (47)$$

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

### A1.3

**Problem:** 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,

$$\mathbf{E}(X) = \sum_{x \in X} xp(x) < \mathbf{E}(X) \sum_{x \in X} p(x) = \mathbf{E}(X) \quad (48)$$

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  $p(x) > 0$ .

### A1.4

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

**Solution:** The following computation gives us the result.

$$\begin{aligned} \mathbf{E}(aX + bY) &= \sum_{(x,y) \in (X,Y)} (ax + by)p(x, y) \\ &= \sum_{x \in X} \sum_{y \in Y} axp(x, y) + byp(x, y) \\ &= \sum_{x \in X} ax \sum_{y \in Y} p(x, y) + \sum_{y \in Y} by \sum_{x \in X} p(x, y) \\ &= a \sum_{x \in X} xp(x) + b \sum_{y \in Y} yp(y) \\ &= a\mathbf{E}(X) + b\mathbf{E}(Y). \end{aligned} \quad (49)$$

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

### A1.5

**Problem:** 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.

$$\begin{aligned} \mathbf{E}(XY) &= \sum_x \sum_y xy p(x, y) \\ &= \sum_x \sum_y xy p(x)p(y) \\ &= \sum_x p(x) \sum_y p(y) \\ &= \mathbf{E}(X)\mathbf{E}(Y) \end{aligned} \quad (50)$$

## Appendix 4: Number theory

### A4.1

**Problem: (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  $a|c$ .

### A4.2

**Problem:** 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,

$$ax + by = djx + dky = d(jx + ky). \quad (51)$$

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

### A4.3

**Problem:** 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

$$b(k - 1) \geq 0 \quad (52)$$

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

$$\begin{aligned} b(k - 1) \geq 0 &\implies bk \geq b \\ &\implies a \geq b \end{aligned} \quad (53)$$

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.

### A4.4

**Problem:** 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

$$697 = 41^1 17^1 \quad (54)$$

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.

$$36300 = 2^2 3^1 5^2 11^2. \quad (55)$$

## A4.5

**Problem:** 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.

## A4.6

**Problem:** 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.

## A4.7

**Problem:** 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  $n - 1$ .

$$(n + 1)(n - 1) = n^2 - 1 = 1 \pmod{n^2}. \quad (56)$$

## A4.8

**Problem: (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

$$a(b - b') = 0 \pmod{n}. \quad (57)$$

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  $b = b' \pmod{n}$ .

## A4.9

**Problem:** 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 6285 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.

#### A4.10

**Problem:** What is  $\varphi(187)$ ?

**Solution:** The prime factorization of 187 is  $11 \times 17$ . Hence,

$$\varphi(187) = \varphi(17 \times 11) = \varphi(17)\varphi(11) = 16 \times 10 = 160 \quad (58)$$

#### A4.11

**Problem:** Prove that

$$n = \sum_{d|n} \varphi(d) \quad (59)$$

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  $\varphi$  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,

$$\begin{aligned} \sum_{d|n} \varphi(d) &= \sum_{j=0}^{\alpha} \varphi(p^j) = 1 + \sum_{j=1}^{\alpha} p^{j-1}(p-1) \\ &= 1 + (p-1) \sum_{j=1}^{\alpha} p^{j-1} \\ &= 1 + \sum_{j=0}^{\alpha} p^j - \sum_{j=0}^{\alpha} p^{j-1} \\ &= p^\alpha, \end{aligned} \quad (60)$$

where, in the last step, all but  $p^\alpha$  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.

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_m^{\alpha_m}. \quad (61)$$

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

$$\begin{aligned} \varphi(n) &= \prod_{j=1}^m \varphi(p_j^{\alpha_j}) \\ &= \prod_{j=1}^m \sum_{k_j=0}^{\alpha_j} \varphi(p^{k_j}) \end{aligned} \quad (62)$$

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.

$$\begin{aligned} \prod_{j=1}^m \sum_{k_j=0}^{\alpha_j} \varphi(p^{k_j}) &= \sum_{k_1=0}^{\alpha_1} \sum_{k_2=0}^{\alpha_2} \cdots \sum_{k_m=0}^{\alpha_m} \prod_{j=1}^m \varphi(p^{k_j}) \\ &= \sum_{k_1=0}^{\alpha_1} \sum_{k_2=0}^{\alpha_2} \cdots \sum_{k_m=0}^{\alpha_m} \varphi(p^{k_1} p^{k_2} \cdots p^{k_m}) \end{aligned} \quad (63)$$

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,

$$\varphi(n) = \sum_{d|n} \varphi(d) \quad (64)$$

as desired.

## A4.12

**Problem:** 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  $\varphi$ . 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  $n$ .

## A4.13

**Problem:** 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

$$k = qr + i \quad (65)$$

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

$$a^k = a^{qr+i} = (a^r)^q a^i = 1^q a^i = a^i \quad (66)$$

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

#### A4.14

**Problem:** 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  $g$ .

#### A4.15

**Problem:** *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,

$$a^{\varphi(n)} = a^{kr} = (a^r)^k = 1^k = 1 \tag{67}$$

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

#### A4.16

**Problem:** 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  $\varphi(n)$ .

#### A4.17

**Problem:**

**Solution:**

#### A4.18

**Problem:**

**Solution:**

#### A4.19

**Problem:**

**Solution:**