Quantum Computing

Linus Ekstrøm

Spring 2021

Abstract

In this document I will collect my thoughts on Quantum Computing.

Contents

1 Short Summaries

1.1 Classical to Quantum

This section will be a short introduction to quantum mechanics for quantum computing.

Classical	Quantum
Phase Space $\mathcal{M}$	Hilbert Space (state space) $\mathcal{H}$
Real valued functions on $\mathcal{M}$	Self-adjoint operators on $\mathcal{H}$
Values of functions on $\mathcal{M}$	Spectra of operators
Hamiltonian H	Self-adjoint operator H on $\mathcal{H}$
Poisson bracket on function S	Operator commutator $[A, B] = AB - BA$

Table 1: Comparison of Classical and Quantum formalism

1.2 Quantum Formalism

We also have the summary table to describe most of the formalism of quantum mechanics (not any quantum field theory here yet i don't think..)

Observables	Self-adjoint operators A on $\mathcal{H}$: $(A \in L(\mathcal{H}), A = A^*)$
Values of an observable A	Elements of the spectrum $\delta(A)$
Pure states	Unit vectors $\|\psi\rangle \in \mathcal{H}$: $(\|\| \|\psi\rangle \|\| = 1)$ considered up to phase factors, or maps $A \mapsto \langle\psi\| A \|\psi\rangle$
Mixed states	Convex combinations of $\langle\psi\| \cdot \|\psi\rangle$, or $\rho : L(\mathcal{H}) \to \mathbb{C}$ such that ρ is linear, positive $(\rho(A) \geq 0$ for $A \geq 0)$, $\rho(I) = 1$
Measurement of A in a state ρ	If $A = \lambda_1 P_1 + \cdots + \lambda_k P_k$ is the spectral decomposition, we get the values $\lambda_1, \cdots, \lambda_k$ with probabilities $\rho(P_1), \cdots, \rho(P_k)$
Hamiltonian (total energy)	$H = H^* \in L(\mathcal{H})$
Equations of motion	Heisenberg picture: $\frac{dA_t}{dt} = \frac{1}{i\hbar}[A_t, H]$ $\big(A_t = e^{-\frac{tH}{i\hbar}} A_0 e^{\frac{tH}{i\hbar}}\big)$. Schrödinger picture: $\frac{d\|\psi_t\rangle}{dt} = \frac{1}{i\hbar} H \|\psi_t\rangle$ $\big(\|\psi_t\rangle = e^{\frac{tH}{i\hbar}} \|\psi_0\rangle \big)$
Composite systems	If we have two quantum mechanical systems with state spaces $\mathcal{H}_1, \mathcal{H}_2$ then the state space of the composite system is $\mathcal{H}_1 \otimes \mathcal{H}_2$

Table 2: Comparison of Classical and Quantum formalism

2 Important Terms

2.1 Quantum Computing

2.1.1 Ancilla bits

In classical computation, any memory bit can be turned on or off at will, requiring no prior knowledge. In the case of classical reversible computing or quantum computing all operations must be reversible an thus operations that would toggle a bit on or off would lose the information about the initial value of that bit, and would result in an irreversible process. In quantum algorithms there is no way to deterministically put bits in a specific state unless there are specific bits whose original state is known in advance. Such bits are called **ancilla bits** is quantum computing.

2.1.2 Quantum Algorithms

There is a variety of Quantum algorithms, with more being developed every year. Below is a figure relative mapping of a couple of algorithms.

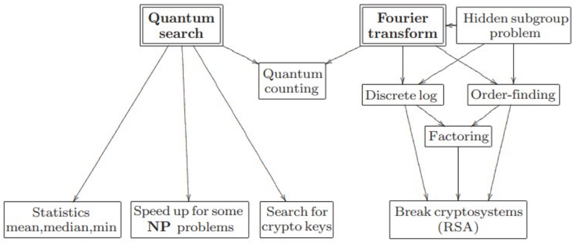


Figure 1: Sourced from [1] p.173, important to realize that this image is from a book written over 20 years ago (as of writing this)

Looking over the Qiskit textbook [2] we see a short list of important algorithms: *Deutsch-Jozsa, Bernstein-Vazirani, Simon's, Quantum Fourier Transform, Quantum Phase Estimation, Shor's, Grover's, Quantum Counting, Quantum Teleportation.*

3 Background Math

3.1 Fourier Series

The following is the explanation of Fourier series given in the book [7] on page 405. Paraphrasing: It is often beneficial to approximate continuous functions $s(x)$, *integrable over the interval* P, as a sum of sine and cosine functions

$$\begin{aligned} s_n &= \frac{a_0}{2} + a_1 \cos t + \cdots + a_n \cos nt \\ &\qquad + b_1 \sin t + \cdots + b_n \sin nt \\ &= \frac{a_0}{2} + \sum_{n=1}^{N} \Big(a_n \cos\big(\frac{2\pi nx}{P}\big) + b_n \sin\big(\frac{2\pi nx}{P}\big)\Big) \end{aligned}$$

where the *Fourier coefficients* a_n and b_n are given by

$$\begin{aligned} a_n &= \frac{2}{P}\int_P s(x)\cos\big(2\pi x\frac{n}{P}\big)dx \\ b_n &= \frac{2}{P}\int_P s(x)\sin\big(2\pi x\frac{n}{P}\big)dx \end{aligned}$$

There are multiple different ways of writing the Fourier series, each with their own interpretations and use cases in a myriad of different fields. We have the so called *amplitude-phase* form

$$s_n = \frac{A_0}{2} + \sum_{n=1}^{N} A_n \cos\Big(\frac{2\pi nx}{P} - \varphi_n\Big)$$

where we have used the definitions $A_n = \sqrt{a_n^2 + b_n^2}$ and $\varphi_n = \arctan 2(b_n, a_n)$ [8]. We also have the, perhaps most commonly seen, *exponential form*

$$s_n = \sum_{n=-N}^{N} c_n e^{i2\pi nx/P}$$

3.2 Fourier Transform

The Fourier transform is a very deep subject, one interpretation is that when you Fourier transform a periodic function you get out a collection of frequencies present in the original function. The Fourier transform is the map

$$F : \big(\text{Functions on } \mathbb{R}\big) \to \big(\text{Functions on } \mathbb{R}\big)$$

$$F(f)(y) = \hat{f}(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)e^{ixy}dx$$

The two most important properties of the Fourier transform are

1. $\frac{d\hat{f}(y)}{dy} = -iy\hat{f}(y)$

2. The Plancherel theorem 5.10, which says F is unitary and thus norm preserving and invertible.

The inverse of the Fourier transform is given by

$$F^{-1}(f)(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(y)e^{-ixy}dy$$

3.2.1 Discrete Fourier

Most the forms written over uses the assumption that $N = \infty$, such that the various sums become integrals. When we use fourier transforms on the computer we will always have an associated error which will decrease as we increase N. [10].

The discrete Fourier transform transforms a sequence of N complex numbers $\{x_n\} := x_0, \cdots, x_{N-1}$ into another sequence of numbers $\{X_n\} := X_0, \cdots, X_{N-1}$ which is defined by

$$\begin{aligned} X_k &= \sum_{n=0}^{N-1} x_n e^{\frac{2\pi i}{N}kn} \\ &= \sum_{n=0}^{N-1} x_n \Big(\cos\big(\frac{2\pi}{N}kn\big) - i \sin\big(\frac{2\pi}{N}kn\big)\Big) \end{aligned}$$

Another way of looking at the discrete Fourier transform is in its matrix form, specifically the unitary transformation

$$\begin{aligned} F_N &= \frac{1}{\sqrt{N}}\begin{bmatrix} \omega_N^{0\cdot 0} & \cdots & \omega_N^{0\cdot(N-1)} \\ \omega_N^{1\cdot 0} & \cdots & \omega_N^{1\cdot(N-1)} \\ \vdots & \ddots & \vdots \\ \omega_N^{(N-1)\cdot 0} & \cdots & \omega_N^{(N-1)\cdot(N-1)} \end{bmatrix} \\ &= \frac{1}{\sqrt{N}}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega_N & \cdots & \omega_N^{1\cdot(N-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_N^{(N-1)} & \cdots & \omega_N^{(N-1)^2} \end{bmatrix} \end{aligned}$$

where $\omega_N = e^{2\pi i/N}$. The inverse transform is then given by $F^{-1} = F^*$. In classical computing there is an efficient algorithm for implementing the discrete Fourier transform. We take advantage of the symmetric nature of the ω's. $\omega_{2N}^{2k} = \omega_N^k$ and $\omega_{2N}^{2k+1} = \omega_{2N}\omega_N^k$. We show by example when $N = 4$, then we have $\omega = i$ and the Fourier transform is

$$\begin{aligned} F_4 &= \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & \omega^3 \\ 1 & \omega^2 & \omega^4 & \omega^6 \\ 1 & \omega^3 & \omega^6 & \omega^9 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & \omega^3 \\ 1 & \omega^2 & 1 & \omega^2 \\ 1 & \omega^3 & \omega^2 & \omega \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega & -1 & -\omega \\ 1 & -1 & 1 & -1 \\ 1 & -\omega & -1 & \omega \end{bmatrix} \end{aligned}$$

Now, we can 'simplify' this by shifting columns so that all the ω's are grouped

$$F_4 = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & \omega & -\omega \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -\omega & \omega \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} F_2 & A_2F_2 \\ F_2 & -A_2F_2 \end{bmatrix}$$

where $A_2 = \begin{bmatrix} 1 & 0 \\ 0 & \omega \end{bmatrix}$. The signs in this above matrix are the same as for the tensor product of Hadamard gates $H \otimes H$. This procedure of grouping the ω's like this allows itself to be generalised to F_{2N}. What we do is we move all columns with even l to the left, and the columns with odd l to the right. This results in the matrix

$$F_{2N} = \frac{1}{\sqrt{2}} \begin{bmatrix} F_N & A_N F_N \\ F_N & -A_N F_N \end{bmatrix}$$
$$A_N = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \omega_2 N & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega_{2N}^{N-1} \end{bmatrix}$$

So if N is a power of two then we get a fast recursive way to compute F_N

3.3 Super Basic Modular Arithmetic

In general, if you are working in $\mod (n)$ (*where n is any whole number*), we write

$$a \equiv b \mod (n)$$

if a and b leave the same remainder when you divide them by n. This is why we do not use equals here, but we say they are congruent. This previous statement is exactly the same as saying we write $a \equiv b \mod (n)$ if n divides $a - b$

3.4 Tensor Product

Granted you can find basis and represent stuff as matrices the tensor product is given by the Kronecker product

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

then the tensor product $A \otimes B$ is

$$A \otimes B = \begin{bmatrix} a_{11} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} & a_{12} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\ a_{21} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} & a_{22} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \end{bmatrix}$$
$$= \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{bmatrix}$$

4 Linear Algebra

4.1 Operators

For any operator C we have

$$\langle\psi| C^* C |\psi\rangle = \big(|\psi\rangle , C^* C |\psi\rangle \big) = \big(C |\psi\rangle , C |\psi\rangle \big) \geq 0$$

meaning that for an any operator C the operator C^*C is positive.

4.1.1 Orthogonal (Real) Matrices

In linear algebra, an orthogonal matrix, or orthonormal matrix, is a real square matrix whose columns and rows are orthonormal vectors.

$$Q^T Q = Q Q^T = I$$

This leads to the equivalent characterization

$$Q^T = Q^{-1}$$

From basic facts about the determinant it follows that the determinant of any orthogonal matrix is ± 1

$$1 = \det(I) = \det\big(Q^T Q\big) = \det\big(Q^T\big)\det(Q) = \big(\det(Q)\big)^2$$

which implies the statement that all orthogonal matrices have determinant ± 1. This is all lifted straight from the wikipedia article on orthogonal matrices, but is probably stated in every linear algebra book.

4.1.2 Normal Operator

A normal operator on a complex Hilbert space is a continuous linear operator $A : \mathcal{H} \to \mathcal{H}$ that commutes with its hermitian adjoint

$$[A, A^*] = AA^* - A^*A = 0$$
$$AA^* = A^*A$$

Normal operators are important because the spectral theorem holds for them.

Normal in Terms of Norms If T is a bounded operator on Hilbert space $\mathcal{H}$. Then T is normal if and only if

$$||T |\psi\rangle || = ||T^* |\psi\rangle ||$$

for all $|\psi\rangle \in \mathcal{H}$

Orthogonal Eigenvectors for Normal Operators Eigenvectors of a normal operator corresponding to distinct eigenvalues are orthogonal. *Proof*: Suppose α, β are distinct eigenvalues of T with corresponding eigenvectors $|\phi\rangle , |\psi\rangle$, then since $T^* f = \bar{\alpha} f$

$$(\beta - \alpha) \langle\phi|\psi\rangle = \langle\beta\phi|\psi\rangle - \langle\phi|\bar{\alpha}\psi\rangle$$
$$= \langle T\phi|\psi\rangle - \langle\phi|T^*\psi\rangle = \langle T\phi|\psi\rangle - \langle T\phi|\psi\rangle = 0$$

since our assumption is that α, β are distinct $\langle\phi|\psi\rangle$ must be zero for the equality to hold.

4.1.3 The Adjoint and Self-Adjoint (Hermitian) Operators

$\mathcal{H}$ is a vector space over $\mathbb{C}$, we denote elements of $\mathcal{H}$ by $|\psi\rangle$. The adjoint of $A : \mathcal{H} \to \mathcal{H}$ is A^*

$$\langle\phi| A^* |\psi\rangle = \overline{\langle\psi| A |\phi\rangle}$$

An operator is called self-adjoint, or *hermitian*, if $A^* = A$. From linear algebra we know that is A is self-adjoint on a finite dimensional Hilbert space $\mathcal{H}$ then

1. All eigenvalues of A are real
2. Eigenvectors corresponding to different eigenvalues are mutually orthogonal
3. There is an orthonormal basis for $\mathcal{H}$ consisting of eigenvectors of A

If we have two self-adjoint operators $A, B \in L(\mathcal{H})_{\text{sa}}$, then we have

$$(AB)^* = B^* A^*$$

4.1.4 Unitary Operators

The mathematical definition of a **unitary matrix** is a matrix A satisfying

$$A^{-1} = A^*$$

where A^* is the **conjugate transpose** of the matrix A. The definition of the conjugate transpose is

$$A^* = \overline{A^T}$$

It follows that unitary matrices are necessarily normal matrices. The evolution of quantum mechanical systems are described by unitary operators. The following facts are all equivalent:

1. U is unitary
2. $||U|\psi\rangle|| = |||\psi\rangle||$ for all $|\psi\rangle \in \mathcal{H}$

3. $\big(\mathrm{U}\,|\phi\rangle, U|\psi\rangle\big) = \langle\phi|\psi\rangle$

Assume $\mathcal{H}$ is a finite dimensional Hilbert space. Then a linear operator U on $\mathcal{H}$ is unitary *if and only if*

$$U = e^{iA}$$

for self-adjoint operator A ($A = A^*$). If we have such an operator U then

$$\det(U) = e^{i\,\mathrm{Tr}\{A\}}$$

and

$$|\det(U)| = 1$$

Which equivalent to saying that the the determinants of unitary matrices lie on the unit circle. We can show this similarly to how we showed that deteminants of orthogonal matrices are ±1

$$\begin{aligned} 1 = \det(I) = \det(U^*U) = \det(U^*)\det(U) \\ \implies \overline{\det(U)}\det(U) = 1 \end{aligned}$$

Thus

$$|\det(U)|^2 = 1 \implies |\det(U)| = 1$$

Which show that the determinants of unitary matrices necessarily lie on the unit circle.

`https://sparse-plex.readthedocs.io/en/latest/book/matrices/unitary_matrices.html`

4.1.5 Spectrum and Spectral Decomposition

Another way to write (and think of) 3). above is: let $\delta(A)$ be the set of different eigenvalues of A

$$\delta(A) = \{\lambda \in \mathbb{C} : A - \lambda I \text{is not invertible}\}$$

we call $\delta(A)$ *the spectrum* of A. Now we let $\delta(A) = \{\lambda_1, \cdots, \lambda_k\}$, for every such λ_i consider the corresponding eigenspace

$$\mathcal{H}_i = \{|\psi\rangle \in \mathcal{H} : A|\psi\rangle = \lambda_i |\psi\rangle\}$$

Then we have

1. $\mathcal{H}_i \perp \mathcal{H}_j$ for $i \neq j$
2. Every $|\phi\rangle \in \mathcal{H}$ can be written as $|\phi\rangle = |\phi_1\rangle + \cdots |\phi_k\rangle$ for $|\phi_i\rangle \in \mathcal{H}_i$

Generally, if both of the above are satisfied for subspace $\mathcal{H}_i \subset \mathcal{H}$ we write $\mathcal{H} = \mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_k$ and say that $\mathcal{H}$ is the *direct sum* of $\mathcal{H}_1, \cdots, \mathcal{H}_k$.

Let $P_i : \mathcal{H} \to \mathcal{H}_i$ be the orthogonal projection:

$$\begin{aligned} P|\psi\rangle = |\psi\rangle \text{if } |\psi\rangle \in \mathcal{H}_i \\ P|\psi\rangle = 0 \text{if } |\psi\rangle\, \mathcal{H}_i \end{aligned}$$

Then 1) is equivalent to saying $P_iP_j = P_jP_i = 0$ for $i \neq j$ and 2) is equivalent

$$\begin{aligned} P_1 + \cdots + P_k = I \\ \lambda_1 P_1 + \cdots + \lambda_k P_k = A \end{aligned}$$

This is called the <u>spectral decomposition</u> of A. Some important facts about the spectral decomposition if we have it like above are

$$\begin{aligned} A^2 &= (\lambda_1)^1 P_1^2 + \cdots + (\lambda_k)^2 P_k^2 \\ &= (\lambda_1)^1 P_1 + \cdots + (\lambda_k)^2 P_k \\ \mu A &= \mu\lambda P_1 + \cdots + \mu\lambda P_k \end{aligned}$$

where in the second line we have used the defining property of projection operators $P^2 = P$. From this it necessarily follows that for any polynomial f we have

$$f(A) = f(\lambda_1)P_1 + \cdots + f(\lambda_k)P_k$$

4.2 Invariant Subspace

A subspace is said to be invariant under a linear operator if its elements are transformed by the linear operator into elements belonging to the subspace itself [32].

4.3 Functional Calculus for Diagonalizable Operators

We start the discussion of functional calculus to be able to better understand entanglement, time evolution of quantum systems. The first remember the Heisenberg Uncertainty relations (5.7) and the orthonormal basis theorem for normal operators (5.8). We also remember that both hermitian and unitary operators are normal operators (4.1.3, 4.1.4).

4.3.1 Using Braket Notation

Let us now consider the normal operator $T \in L(\mathcal{H})$, where $\mathcal{H}$ is a finite dimensional Hilbert space, with a unit eigenvector $|\psi_1\rangle$. (*this is the proof of 5.8*, i.e. that we can find an orthonormal basis of $\mathcal{H}$ consisting of eigenvectors of T.)

$$T|\psi_1\rangle = \lambda_1 |\psi_1\rangle \implies T^*|\psi_1\rangle = \overline{\lambda_1}|\psi_1\rangle$$

Consider now the space

$$\mathcal{H}_1 = |\psi\rangle^{\perp} = \{|\psi\rangle \,|\, \langle\psi|\psi_1\rangle = 0\}$$

We now claim that $T\mathcal{H}_1 \subset \mathcal{H}_1$ and $T^*\mathcal{H}_1 \subset \mathcal{H}_1$. In other words, T and T^* acting on arbitrary vectors in $\mathcal{H}_1$ remain 'within' $\mathcal{H}_1$[1]. This is indeed the case, we take $|\psi\rangle \in \mathcal{H}_1$

$$\begin{aligned}\big(T|\psi\rangle, |\psi_1\rangle\big) &= \big(|\psi\rangle, T^*|\psi\rangle\big) \underset{\text{def}}{=} \big(|\psi\rangle, \overline{\lambda_1}|\psi_1\rangle\big)\\ &= \overline{\lambda_1}\langle\psi|\psi_1\rangle = 0\end{aligned}$$

$$\begin{aligned}\big(T^*|\psi\rangle, |\psi_1\rangle\big) &= \big(|\psi\rangle, T|\psi\rangle\big) \underset{\text{def}}{=} \big(|\psi\rangle, \lambda_1|\psi_1\rangle\big)\\ &= \lambda_1\langle\psi|\psi_1\rangle = 0\end{aligned}$$

Therefore both $T|\psi\rangle \in \mathcal{H}_1$ and $T^*|\psi\rangle \in \mathcal{H}_1$. And we can conclude we have a well-defined $T_1 \in L(\mathcal{H}_1)$ namely

$$T_1 = T|_{\mathcal{H}_1}, \quad T_1^* = T|_{\mathcal{H}_1}$$

In particular T_1 is a yet again a normal operator. So we can again find a unit eigenvector $|\psi\rangle_2 \in \mathcal{H}_1$ for T_1 and continue like this, and since every time we make a new subspace $\mathcal{H}_{++1}$ we require that all vectors in it be orthogonal to the previously found eigenvector we generate an orthogonal basis.

Now since we started out with a finite dimensional Hilbert space $\mathcal{H}$, we will generate this orthogonal basis in a finite number of such steps. Thus we can guarantee that if $T \in L(\mathcal{H})$ is normal, then we can find an orthogonal basis of $\mathcal{H}$ consisting of eigenvectors to T.

4.3.2 Matrix Formulation

In terms of matrices the previous section can be reformulated as:

If $T \in \mathrm{Mat}_n(\mathbb{C})$ is normal then there is a unitary matrix U and $\lambda_1, \cdots, \lambda_n \in \mathbb{C}$ such that

$$UTU^* = \begin{bmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{bmatrix}$$

Indeed, we can find an orthonormal basis $|\psi_1\rangle, \cdots, |\psi_n\rangle \in \mathbb{C}^n$ such that $T|\psi_i\rangle = \lambda_i|\psi_i\rangle$. We define a linear operator U on $\mathbb{C}^n$ by'

$$U|\psi_i\rangle = |i\rangle = (0, \cdots, 0, \underset{i}{1}, 0, \cdots, 0)$$

We have

$$\big(U|\psi_i\rangle, U|\psi_j\rangle\big) = (|i\rangle, |j\rangle) = \delta_{ij} = (|\psi_i\rangle, |\psi_j\rangle)$$

where δ_{ij} is the Kronecker delta function. It follows that

$$\big(U|\psi\rangle, U|\phi\rangle\big) = (|\psi\rangle, |\phi\rangle)$$

and therefore U is unitary by definition. We now have

$$UTU^{-1} = \begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{pmatrix}$$

Remark If we have T as above, then T^* is like it except with $\overline{\lambda_1}, \cdots, \overline{\lambda_n}$ on the diagonal. This implies that if T is a normal operator on a finite dimensional Hilbert space $\mathcal{H}$ then

$$\begin{aligned}T \text{ is self-adjoint } &\Leftrightarrow \delta(T) \subset \mathbb{R}\\ T \text{ is unitary } &\Leftrightarrow \delta(T) \subset \mathbb{T}\end{aligned}$$

where $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, and $\delta(T)$ is the spectra of T.

[1] Not entirely sure if this is the correct interpretation. We are going to a subset..

4.3.3 Functions of Operators

From here on we are going to be dealing with diagonalizable operators on finite dimensional complex vector spaces $\mathcal{H}$. (all we're really saying here is we can find a basis consisting of orthogonal eigenvectors).

Given a function $f : \Omega \to \mathbb{C}, \Omega \subset \mathbb{C}$ how can we make a sensible definition of $f(T)$? First we consider what happens in f is a polynomial function on $\mathbb{C}$. Then we define

$$f(T) = a_0 I + a_1 T + \cdots + a_k T^k$$

Lemma Assume that T is diagonalizable and that f, g are two polynomials such that $f(\lambda) = g(\lambda)$ for $\lambda \in \delta(T)$. Then $f(T) = g(T)$. We can show this by letting noticing if we have $f(z) = a_0 + a_1 z + \cdots + a_k z^k$ then

$$\begin{aligned} f(T) |\psi_i\rangle &= \left(a_0 I + a_1 T + \cdots + a_k T^k\right) |\psi_i\rangle \\ &= a_0 |\psi_i\rangle + a_1 \lambda |\psi_i\rangle + \cdots + a_k \lambda^k |\psi_i\rangle \\ &= f(\lambda_i) |\psi_i\rangle \end{aligned}$$

SHOULD IT BE λ_i in the middle here? Similarly,

$$g(T) |\psi_i\rangle = g(\lambda_i) |\psi_i\rangle$$

So we have now reached some form of conclusion: *If $\mathcal{H}$ is a finite dimensional (complex) vector space and T is a diagonalizable linear operator on $\mathcal{H}$, then for every $f : \delta(T \to \mathbb{C}$ we can define $f(T)$ in either of the following equivalent ways*

1. Find a polynomial function $\tilde{f}(z) = a_0 + a_1 z + \cdots + a_k z^k$ such that $\tilde{f}(\lambda) = f(\lambda)$ for all $\lambda \in \delta(T)$. Then $f(T) = \tilde{f}(T)$

2. Find a basis $|\psi_1\rangle, \cdots, |\psi_n\rangle$ of $\mathcal{H}$ consisting of eigenvectors of T: $T |\psi_i\rangle = \lambda_i |\psi_i\rangle$ then $f(T) |\psi_i\rangle = f(\lambda_i) |\psi_i\rangle$

In terms of matrices 2) above means: *Take $T \in \mathrm{Mat}_n(\mathbb{C})$. If T is diagonalizable, there is an invertible matrix S such that*

$$STS^{-1} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

Then we have

$$f(T) = S^{-1} \begin{pmatrix} f(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & f(\lambda_n) \end{pmatrix} S$$

4.4 Rotation Group and Bloch Sphere

Because of the theorem 5.2 it is of big importance to have a firm grasp of the 1-qubit gates. In a way, understanding these will let you understand quantum computation. In this section we head towards understanding the mathematical group of all 1-qubit gates.

We identify the Hilbert space $\mathcal{H}$ of one qubit with $\mathbb{C}^2$ using the basis $|0\rangle, |1\rangle$. Then the 1-qubit gates are all in (they are, i think thinking about the gates being in U(2) is slightly incorrect. It's more like the gates are the group).

$$U(2) = \text{the } \underline{\text{group of unitary}}\ 2 \times 2 \text{ matrices.}$$

Group means that $U(2)$ is closed under multiplication and inversion. Meaning that if you multiply two elements of $U(2)$ or take their inverse, then the result is also an element of $U(2)$. Since pure states are unit vectors up to phase factors, we are only interested in unitaries up to phase factors. This is because in quantum mechanics any observable is proportional to the amplitude[2]. So, we actually only need to consider the *group*

$$PU(2) = \underline{\text{projective unitary group}} \text{ of } 2 \times 2 \text{ matrices}$$

Now for our last step, which brings us to the group $SU(2)$. For a fact any unitary $U \in U(2)$ can be multiplied by a phase factor to set its determinant to one. Therefore, we really need only consider the so called *special unitary group*

$$\begin{aligned} SU(2) &= \underline{\text{special unitary group}} \text{ of } 2 \times 2 \text{ matrices} \\ &= \{U \in U(2) : det(U) = 1\} \end{aligned}$$

Any unitary $U \in SU(2)$ can be written in the following way

$$U^{iA}, A = A^*, \mathrm{Tr}(A) = 0$$

By diagonalizing U we can show we only have to consider unitaries of the form

$$U = \begin{bmatrix} z & 0 \\ 0 & \overline{z} \end{bmatrix}$$

Then using the polar form $z = e^{it}, t \in \mathbb{R}$. Then we can define our A, which is in $U = e^{iA}$, as follows

$$A = \begin{bmatrix} t & 0 \\ 0 & -t \end{bmatrix}$$

4.4.1 $SU(2)$

Therefore we have to understand the space

$$\{A \in \mathrm{Mat}_2(\mathbb{C}) : A = A^*, \mathrm{Tr}(A) = 0\}$$

lets decode the above statement. We're talking about the space of complex 2×2 matrices, that are self-adjoint where the sum of its eigenvalues are equal to zero.

[2] Good explanation for this is here `https://quantumcomputing.stackexchange.com/questions/5125/what-is-the-difference-between-a-relative-phase-and-a-global-phase-in-particula`

This is (somewhat confusingly) a real vector space, consisting of matrices

$$\begin{bmatrix} a & b \\ \bar{b} & -a \end{bmatrix}, a \in \mathbb{R}, b \in \mathbb{C}$$

This is a three-dimensional real vector space. Real here refers to the fact that the scalars (numbers) of the space are real. It has a basis which is familiar to physicist: the **Pauli matrices**.

$$\begin{aligned} \sigma_X = X &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \sigma_Y = Y &= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\ \sigma_Z = Z &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{aligned}$$

For an arbitrary vector $\vec{a} \in \mathbb{R}^3$ denote

$$\vec{a} \cdot \vec{\sigma} = a_x \sigma_X + a_y \sigma_Y + a_z \sigma_Z$$

Thus, we have a linear isomorphism

$$\begin{aligned} \mathbb{R}^3 &\to \{A \in \mathrm{Mat}_2(\mathbb{C}) : A = A^*, \mathrm{Tr}(A) = 0\} \\ \vec{a} &\mapsto \vec{a} \cdot \vec{\sigma} \end{aligned}$$

The conclusion is

$$\begin{aligned} SU(2) &= \{e^{i(\vec{a}\cdot\vec{\sigma})} | \vec{a} \in \mathbb{R}^3\} \\ &= \{e^{it(\vec{a}\cdot\vec{\sigma})} = (\cos(t))I + i(\sin(t))\vec{a}\cdot\vec{\sigma} : \\ &\quad \vec{a} \in S^2, t \in \mathbb{R}\} \end{aligned}$$

The Pauli matrices are cyclicly anti-commutative 7.1.3 (this is not an accurate term but whatever...).

There's a bunch of stuff about symmetries and adjuncts acting on symmetries and isomorphisms between vectors and symmetries covered in the lecture notes ([14], [15]) which is a bit too heavy for me to include right now.

4.4.2 Bloch Sphere Representation

H is the Hilbert space of one qubit. Every unit vector in $\mathcal{H}$, up to phase factor, can be written as

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

$0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. This defines a point, called the *Bloch vector* or *Bloch coordinates* of $|\psi\rangle$

$$\vec{a} = \big(\sin(\theta)\cos(\phi), \sin(\theta)\sin(\phi), \cos(\theta)\big) \in S^2$$

where S^2 is the three-dimensional sphere.

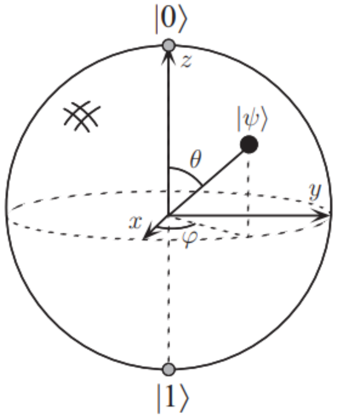


Figure 2: Bloch Sphere representation p.15 Nielsen [1]

In the Bloch coordinates $e^{it\vec{a}\cdot\vec{\sigma}}, \vec{a} \in S^2$ is the rotation about $\vec{a}$, *in the counter-clockwise direction* by $-2t$ degrees.

4.4.3 Rotation Operators

We introduce the notation, given $\vec{a} \in S^2$ and $\theta \in \mathbb{R}$,

$$R_{\vec{a}}(\theta) = e^{-\frac{i\theta}{2}\vec{a}\cdot\vec{\sigma}}$$

where we remember that $\vec{\sigma} = \big(\sigma_X, \sigma_Y, \sigma_Z\big)$. This operator $R_{\vec{a}}(\theta)$ we can understand as a rotation about the axis defined by $\vec{a}$ by θ degrees in Bloch coordinates. We note that

$$R_{\vec{a}}(\theta + 2\pi) = -R_{\vec{a}}(\theta)$$

By identifying $\mathcal{H}$ with $\mathbb{C}^2$ using the basis $|0\rangle, |1\rangle$, and identifying V with $\mathbb{R}^3$ using the basis $\sigma_X, \sigma_Y, \sigma_Z$, we can say we have a group isomorphism

$$\begin{aligned} PU(2) &\to SO(3) \\ R_{\vec{a}}(\theta) &\to \text{rotation about } \vec{a} \text{ by } \theta \text{ degrees} \end{aligned}$$

5 Theorems

5.1 Schmidt Decomposition

Suppose $|\psi\rangle$ is a pure state of a composite system $\mathbf{A}, \mathbf{B}$. Then there exists orthonormal states $|i_A\rangle$ for system $\mathbf{A}$, and orthonormal states $|i_B\rangle$ for system $\mathbf{B}$ such that

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

where λ_i are non-negative real numbers satisfying $\sum_i \lambda_i^2 = 1$ known as *Schmidt coefficients*. If the Schmidt decomposition has a Schmidt number of 1 then the above equation can be represented as a product state $\sum_i \lambda_i |i_A\rangle |i_B\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ since only one Schmidt coefficient is nonzero.

From Nielsen p. 109 [1]. If we take the Schmidt decomposition as above then we have reduced density matrices

$$\rho^A = \sum_i \lambda_i^2 |i_A\rangle \langle i_A| = \begin{bmatrix} i_{A_1} i^*_{A_1} & i_{A_1} i^*_{A_2} & \cdots & i_{A_1} i^*_{A_n} \\ i_{A_2} i^*_{A_1} & i_{A_2} i^*_{A_2} & \cdots & i_{A_2} i^*_{A_n} \\ \vdots & \vdots & \ddots & \vdots \\ i_{A_n} i^*_{A_1} & i_{A_n} i^*_{A_2} & \cdots & i_{A_n} i^*_{A_n} \end{bmatrix}$$

$$\rho^B = \sum_i \lambda_i^2 |i_B\rangle \langle i_B|$$

Therefore, if $|\psi\rangle$ has Schmidt number of 1, the reduced density matrices ρ^A, ρ^B have only one non-zero eigenvalue and are pure states.

The Schmidt number is defined as

$$|\eta\rangle = \sum_{k=1}^n |\varphi_k\rangle \otimes |\psi_k\rangle \quad (1)$$

$$m = \dim(\{|\psi_1\rangle, \cdots, |\psi_n\rangle\})$$

is called the Schmidt number of $|\eta\rangle$.

5.2 Gate Decomposition Theorem

Any n-qubit gate decomposes into a product of 1-qubit gates and CNOT gates applied to neighboring pairs of qubits.

5.3 Continued Fraction Theorem 2

If $x = [a_0; a_1; \cdots]$ is a continued fraction presentation of a real number x, then, for all $k \geq 1$ we have

$$|x - \frac{p_k}{q_k}| < \frac{1}{q_k^2}$$

and either

$$|x - \frac{p_k}{q_k}| \leq \frac{1}{2q_k^2} \quad \text{or} \quad |x - \frac{p_{k-1}}{q_{k-1}}| \leq \frac{1}{2q_{k-1}^2}$$

from lecture 21 spring 2021[18]

5.4 Continued Fraction Theorem 3

Assume x is a real number and

$$|x - \frac{p}{q}| \leq \frac{1}{2q^2}$$

for a rational non-integral number p/q. Then p/q is a convergent of the continued fraction expansion of x.

5.5 Chinese Remainder Theorem

Assume $m, n \geq 1$ are relatively prime. Then the map

$$\begin{aligned} \mathbb{Z}_{mn} &\to \mathbb{Z}_m \times \mathbb{Z}_n \\ a &\to (a \mod (m), a \mod (n)) \end{aligned}$$

is a bijection.

5.6 The No-Cloning Theorem

There is no unitary

$$U : \mathcal{H}^{\otimes 2} \mapsto \mathcal{H}^{\otimes 2}$$

Such that

$$U\big(|\phi\rangle \otimes |0\rangle\big) = |\phi\rangle \otimes |\phi\rangle$$

for all unit vectors $|\phi\rangle \in \mathcal{H}$.

5.7 Uncertainty Relations / Heisenbergs Uncertainty Theorem

For any $A = A^*, B = B^* \in L(\mathcal{H})$ and for any state ρ, we have

$$\mathrm{Var}_\rho(A)^{\frac{1}{2}} \mathrm{Var}_\rho(B)^{\frac{1}{2}} \geq \frac{1}{2} |\rho([A, B])|$$

where we define the variation of A with respect to ρ by

$$\begin{aligned} \mathrm{Var}_\rho(A) &= \rho\big((A - \rho(A)I)^2\big) \\ &= \sum_{i=1}^k \big(\lambda_i - \rho(A)\big)^2 \rho(P_i) \end{aligned}$$

where $\rho(A) = \sum_{i=1}^k \lambda_i \rho(P_i)$ from spectral decomposition of A. We can do this because $A = A^*$ means that it is a normal operator, and normal operators can be spectrally decomposed.

$$\mathrm{Var}_\rho(A) = 0$$

if and only if there is i such that $\rho(P_j) = 0$ for $j \neq i$

5.8 Normal Operators Theorem

Assume $\mathcal{H}$ is a finite dimensional Hilbert space, $T \in L(\mathcal{H})$ is normal. Then there is an orthonormal basis in $\mathcal{H}$ consisting of eigenvectors of T. Both unitary and self-adjoint (hermitian) operators are normal operators.

Lemma If $|\psi\rangle$ is an eigenvector of T, $T|\psi\rangle = \lambda |\psi\rangle$ for $\lambda \in \mathbb{C}$. Then it is also an eigenvector for T^* with eigenvalue $\overline{\lambda}$.

5.9 Solovay-Kitav Theorem

There are constants $\epsilon, C_1, C_2 > 0$ such that if $d \geq 1$ and a subset $\mathcal{A} \subset SU(d)$ is closed under inversion such that for any $U \in SU(d)$ there is a $V \in \mathcal{A}$ such that $||U - V|| < \epsilon$, the for any $U \in SU(2)$ and any $\delta > 0$ there are $V_1, \cdots, V_L \in \mathcal{A}$ such that

$$||U - V_1 \cdots V_L|| < \delta, \qquad L \leq C_1 (\log\left(\frac{1}{\delta}\right))^{C_2}$$

5.10 Plancherel Theorem

With F_N being the Fourier transform. F defines a unitary operator $L^2(\mathbb{R}) \to L^2(\mathbb{R})$. SO F preserves the L^2 norm and F is invertible.

6 Quantum Mechanical Formalism

This section is based around the lecture notes in Mat3420: Quantum Computation spring 2021 at the University of Oslo, taught by Sergiy Neshveyev, which are based on Neilsen [1]. Some parts are also based on the Qiskit textbook [2], and sporadically Wikipedia articles.

6.1 Measurement in Quantum Mechanics

Assume we are given a quantum mechanical system with state space $\mathcal{H}$ which we assume to be finite dimensional. Assume $A = A^*$ is an observable (observables in quantum mechanics are indeed defined by self-adjoint / hermitian operators) (4.1.3). Let $\mathrm{spec}(A)$ be the spectral decomposition of A (4.1.5). According to the formalism of quantum mechanics, every unit vector $|\psi\rangle \in \mathcal{H}$ describes a state of the system. Further, if we measure A in this state, then we get values $\delta(A)$ with probabilities

$$\langle\psi| P_1 |\psi\rangle , \cdots , \langle\psi| P_k |\psi\rangle$$

Note that

$$\begin{aligned} \langle\psi| P_i |\psi\rangle &= \big(|\psi\rangle , P_i |\psi\rangle \big) \\ &\underset{P_i=P_i^2}{=} \big(|\psi\rangle , P_i^2 |\psi\rangle \big) \\ &\underset{P_i=P_i^*}{=} \big(P_i |\psi\rangle , P_i |\psi\rangle \big) \\ &= ||P_i |\psi\rangle||^2 \geq 0 \end{aligned} \tag{2}$$

and

$$\sum_{i=1}^{k} \langle\psi| P_i |\psi\rangle \underset{\sum P_i = I}{=} \langle\psi|\psi\rangle = 1$$

Unit vectors in $\mathcal{H}$ do not describe all states of the system. To formulate a more general notion, let us introduce the following notation: $L(\mathcal{H})$ is the set of all linear operators on $\mathcal{H}$. We also have

$$L(\mathcal{H})_{\mathrm{sa}} \subset L(\mathcal{H})$$

the subset of self-adjoint operators (observables). Every $C \in L(\mathcal{H})$ can be written in a unique way for $A, B \in L(\mathcal{H})_{\mathrm{sa}}$.

$$A = \frac{C + C^*}{2}, \qquad B = \frac{C - C^*}{2i}$$

We now consider the mapping

$$\rho : L(\mathcal{H}) \to \mathbb{C}, \qquad \rho(A) = \langle\psi| A |\psi\rangle$$

This map has the following properties

1. ρ is linear, that is, $\rho(\lambda A + B) = \lambda\rho(A) + \mu\rho(B)$ for all $A, B \in L(\mathcal{H})$ and $\lambda, \mu \in \mathbb{C}$
2. ρ is positive, meaning $\rho \geq 0$ if A is positive. An operator A is called positive if $\langle\psi| A |\psi\rangle \geq 0$ for all $|\psi\rangle \in \mathcal{H}$
3. ρ is normalized: $\rho(I) = 1$

Definition of a State A state of the system is any map $\rho : L(\mathcal{H}) \to \mathbb{C}$ with the above three properties. If $\rho(A) = \langle\psi| A |\psi\rangle$ for a unit vector $|\psi\rangle$ then we call ρ a *pure state*. Otherwise ρ is called a *mixed state*.

Fact 1 A state φ is pure if and only if, whenever $\varphi = \lambda\rho_1 + (1-\lambda)\rho_2$ for some states ρ_1, ρ_2 and $0 < \lambda < 1$, we must have $\varphi = \rho_1 = \rho_2$.

Fact 2 Any mixed state can be written as

$$\rho = \alpha_1\varphi_1 + \cdots + \alpha_n\varphi_n$$

for some pure states $\varphi_1, \cdots, \varphi_n$ and numbers $\alpha_1, \cdots, \alpha_n \geq 0$ where the sum of the α's is equal to 1.

6.2 Uncertainty Relations

Consider a quantum mechanical system with a finite dimensional state space $\mathcal{H}$. Take an observable $A = A^*$. Take the spectral decomposition of A. The question is now, in which state do we get a definitive (with probability 1) value of A?

Consider a pure state $\varphi = \langle\psi| \cdot |\psi\rangle$. The probability of getting value λ_i is $\langle\psi| P_i |\psi\rangle = ||P_i |\psi\rangle||^2$ ((2)). The probability of getting λ_i is 1 if and only if

$$P_j |\psi\rangle = 0, \qquad \text{for all } j \neq i$$

This happens if and only if

$$|\psi\rangle \in \mathcal{H}_i = P_i\mathcal{H}$$

that is

$$A|\psi\rangle = \lambda_i |\psi\rangle$$

If we have two observables A, B, then in general they have no common eigenvectors.

For any $A = A^*, B = B^* \in L(\mathcal{H})$ and for any state ρ, we have

$$\mathrm{Var}_\rho(A)^{\frac{1}{2}} \mathrm{Var}_\rho(B)^{\frac{1}{2}} \geq \frac{1}{2}|\rho([A, B])|$$

where we define the variation of A with respect to ρ by

$$\begin{aligned} \mathrm{Var}_\rho(A) &= \rho\big((A - \rho(A)I)^2\big) \\ &= \sum_{i=1}^{k} \big(\lambda_i - \rho(A)\big)^2 \rho(P_i) \end{aligned}$$

where $\rho(A) = \sum_{i=1}^{k} \lambda_i \rho(P_i)$ from spectral decomposition of A. We can do this because $A = A^*$ means that it is a normal operator, and normal operators can be spectrally decomposed.

$$\mathrm{Var}_\rho(A) = 0$$

if and only if there is i such that $\rho(P_j) = 0$ for $j \neq i$

6.3 Time Evolution in Quantum Mechanics

This section follows from the mathematics discussed in the functional calculus section (4.3.3). The most important 'function of an operator' in quantum mechanics will be the exponential function $\exp(T)$. Familiarly, we assume T is a diagonalizable linear operator on $\mathcal{H}$. Take $|\psi\rangle \in \mathcal{H}$ and consider the differential equation for a vector valued function

$$t \to |\psi_t\rangle \in \mathcal{H} : \begin{cases} \frac{d}{dt}|\psi_t\rangle & = T|\psi_t\rangle \\ |\psi_0\rangle & = |\psi\rangle \end{cases} \tag{3}$$

Proposition: the system (3) has a unique solution $|\psi_t\rangle = e^{tT}|\psi\rangle$. e^T can be defined for any linear operator on $\mathcal{H}$ as such

$$e^T = \sum_{n=0}^{\infty} \frac{T^n}{n!}$$

6.3.1 Heisenberg Picture of Evolution

In quantum mechanics the evolution of an observable A is described by the equation

$$\begin{cases} \frac{dA_t}{dt} & = \frac{1}{i\hbar}[A_t, H] \\ A_0 & = A \end{cases} \tag{4}$$

where $[A_t, H]$ is the commutator between the observable and the Hamiltonian of the system and $\hbar$ is Planck's constant. The unique solution of (4) is

$$A_t = e^{-\frac{tH}{i\hbar}} A e^{\frac{tH}{i\hbar}}$$

The evolution in terms of observables is called *the Heisenberg picture* of evolution. Let's show that it is a solution. If T is a diagonalizable operator then

$$\frac{d}{dt}\big(e^{tT}\big) = Te^{tT} = e^{tT}T \tag{5}$$

Further, if A_f and B_f are operator-valued functions then

$$\frac{d}{dt}\big(A_t B_t\big) = \frac{dA_t}{dt} B_t + A_f \frac{dB_t}{dt} \tag{6}$$

This is like the derivative product rule we learn in high school. The proof of both these statements can be found in lecture notes 5 [11]. Now to check if A_t satisfies (4):

$$\begin{aligned} \frac{d}{dt}\Big(e^{-\frac{tH}{i\hbar}} A e^{\frac{tH}{i\hbar}}\Big) &\underset{(6)}{=} \frac{d}{dt}\big(e^{-\frac{tH}{i\hbar}}\big) A e^{\frac{tH}{i\hbar}} \\ &+ e^{-\frac{tH}{i\hbar}} A \frac{d}{dt}\big(e^{\frac{tH}{i\hbar}}\big) \\ &\underset{(5)}{=} -\frac{H}{i\hbar} e^{-\frac{tH}{i\hbar}} A e^{\frac{tH}{i\hbar}} \\ &+ e^{-\frac{tH}{i\hbar}} A e^{\frac{tH}{i\hbar}} \frac{H}{i\hbar} \\ &= \frac{1}{i\hbar}\big[e^{-\frac{tH}{i\hbar}} A e^{\frac{tH}{i\hbar}}, H\big] \end{aligned}$$

which is what we wanted to show

$$\frac{dA_t}{dt} = \frac{1}{i\hbar}[A_t, H]$$

6.3.2 The Schrödinger Picture of Evolution

So how do states evolve in quantum mechanics. We have shown that we know of a general solution to the Heisenberg picture of evolution. Assume we start with a pure state $|\psi\rangle$. Suppose at time t it has evolved to a state $|\psi_t\rangle$. Then for every observable A, we have

$$\begin{aligned} \langle\psi_t| A |\psi\rangle &= \langle\psi| e^{-\frac{tH}{i\hbar}} A e^{\frac{tH}{i\hbar}} |\psi\rangle \\ &= \big(e^{\frac{tH}{i\hbar}}|\psi\rangle, A e^{\frac{tH}{i\hbar}}|\psi\rangle\big) \end{aligned}$$

which means that we can take

$$|\psi_t\rangle = e^{\frac{tH}{i\hbar}}|\psi\rangle$$

In other words

$$\begin{cases} \frac{d}{dt}|\psi_t\rangle & = \frac{H}{i\hbar}|\psi\rangle \\ |\psi_0\rangle = |\psi\rangle \end{cases}$$

And this is what is called the Schrödinger equation and the evolution in terms of states is called *the Schrödinger picture* of evolution. One of the most famous equations in quantum mechanics. Another, maybe more familiar way of representing it is

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|psi(t)\rangle$$

along with the time-independent version

$$H|\psi\rangle = E|\psi\rangle$$

6.4 Composite Systems in Quantum Mechanics

We start of our dive into quantum computing by first describing how we represent quantum mechanical systems composed of other smaller systems.

6.4.1 Definition of Tensors / Tensor Product

Given finite dimensional, *real or complex*, vector spaces V_1 and V_2 with bases $|\psi_1\rangle, \cdots, |\psi_n\rangle$ and $|\phi_1\rangle, \cdots, |\phi_m\rangle$ respectively the tensor product $V_1 \otimes V_2$ is defined as the vector space with basis $|\psi_i\rangle \otimes |\phi_j\rangle$ for $1 \leq i \leq n$ and $1 \leq j \leq m$.

It is often convenient to think of $V_1 \otimes V_2$ in a basis-independent way: *For every $|\psi\rangle \in V_1$ and $|\phi\rangle \in V_2$ we define $|\psi\rangle \otimes |\phi\rangle \in V_1 \otimes V_2$ as follows*

$$\begin{aligned} |\psi\rangle &= a_1|\psi_1\rangle + \cdots + a_n|\psi_n\rangle \\ |\phi\rangle &= b_1|\phi_1\rangle + \cdots + b_m|\phi_m\rangle \\ |\psi\rangle \otimes |\phi\rangle &= \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j |\psi_i\rangle \otimes |\phi_j\rangle \end{aligned}$$

The elements of $V_1 \otimes V_2$ are called tensors while the elements of the form $|\psi\rangle \otimes |\phi\rangle$ are called elementary tensors.

This next part is technical, but I wanted to leave it in.

In a more basis-independent way, $V_1 \otimes V_2$ can be described as a vector space together with a map

$$\begin{gathered} V_1 \times V_2 \to V_1 \otimes V_2 \\ (|\psi\rangle , |\phi\rangle) \mapsto |\psi\rangle \otimes |\phi\rangle \end{gathered}$$

And this map satisfies the following properties.

1. The map is *bilinear*
$$\begin{aligned} \big(a_1 |\psi_1\rangle + a_2 |\psi_2\rangle\big) \otimes |\phi\rangle &= a_1 |\psi_1\rangle \otimes |\phi\rangle \\ &\quad + a_2 |\psi_2\rangle \otimes |\phi\rangle , \\ |\psi\rangle \otimes \big(b_1 |\phi_1\rangle + b_2 |\phi_2\rangle\big) &= b_1 |\psi\rangle \otimes |\phi_1\rangle \\ &\quad + b_2 |\psi\rangle \otimes |\phi_2\rangle \end{aligned}$$

2. $V_1 \otimes V_2$ is spanned by $|\psi\rangle \otimes |\phi\rangle$

3. If V is any vector space and $B : V_1 \times V_2 \mapsto V$ is any bilinear map, then there is a unique linear map $\overline{B} : V_1 \otimes V_2 \mapsto V$ such that
$$B(|\psi\rangle , |\phi\rangle) = \overline{B}(|\psi\rangle \otimes |\phi\rangle)$$

I think the best way to understand this is that for any two vector spaces V_1, V_2 we can have a bilinear map B and when we use this map on the pair $|\psi\rangle , |\phi\rangle$ then we end up in a space V. And then we are guaranteed to have a unique linear map from the elementary tensor $|\psi\rangle \otimes |\phi\rangle$ which ends up the same 'place' in V.

Proposition: Suppose we have a complex vector space V. Then we have a linear isomorphism

$$\begin{gathered} V \times \cdots \times V \to V \otimes \mathbb{C}^n \\ (|\psi_1\rangle , \cdots , |\psi_n\rangle) \mapsto \sum_{i=1}^{n} |\psi_i\rangle \otimes |i\rangle \end{gathered}$$

6.4.2 Observables of Composite Systems

We think of observables for a composite quantum mechanical system with state spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ as 'living in' the space

$$L(\mathcal{H}_1)_{\text{sa}} \underset{\mathbb{R}}{\otimes} L(\mathcal{H}_2)_{\text{sa}} \tag{7}$$

There is some more technical details in the lecture notes showing that $\mathcal{H}_1 \otimes \mathcal{H}_2$ is indeed the underlying Hilbert space to (7) [12], but from here on when we write $\mathcal{H}_1 \otimes \mathcal{H}_2$ we take it to mean the underlying Hilbert space to (7).

Proposition: If $\mathcal{H}_1, \mathcal{H}_2$ are finite dimensional Hilbert spaces, then we have a linear isomorphism

$$\begin{aligned} &L(\mathcal{H}_1) \otimes L(\mathcal{H}_2) \to L(\mathcal{H}_1 \otimes \mathcal{H}_2) \\ &S \otimes T \mapsto S \otimes T \end{aligned}$$

for $S \in \mathcal{H}_1, T \in \mathcal{H}_2$. Therefore there is no need to distinguish between $L(\mathcal{H}_1) \otimes L(\mathcal{H}_2)$ and $L(\mathcal{H}_1 \otimes \mathcal{H}_2)$.

We have reached our conclusion: *In quantum mechanics, given two systems with state spaces $\mathcal{H}_1, \mathcal{H}_2$ the Hilbert space of the composite system is $\mathcal{H}_1 \otimes \in$, and the observables of the system live in $L(\mathcal{H}_1)_{\text{sa}} \otimes L(\mathcal{H}_2)_{\text{sa}}$.*

6.5 Entanglement

One of the most 'mysterious' parts of quantum mechanics. Now we will look at what it means in terms of the mathematics of the previous section. Suppose we have two quantum mechanical systems with finite dimensional state spaces $\mathcal{H}_1, \mathcal{H}_2$ and composite state space $\mathcal{H}_1 \otimes \mathcal{H}_2$. If we let $\varphi_1 \in L(\mathcal{H}_1)$ and $\varphi_2 \in L(\mathcal{H}_2)$, then there is a unique state $\varphi_1 \otimes \varphi_2$ on $L(\mathcal{H}_1) \otimes L(\mathcal{H}_2) = L(\mathcal{H}_1 \otimes \mathcal{H}_2)$ such that

$$\big(\varphi_1 \otimes \varphi_2\big)\big(S \otimes T\big) = \varphi_1(S)\varphi(T)$$

Definition A state φ on $L(\mathcal{H}_\infty \otimes \mathcal{H}_2)$ is called *separable* if it is a convex combination of states of the form $\varphi_1 \otimes \varphi_2$, that is

$$\varphi = \lambda_1 \varphi_1 \otimes \psi_1 + \cdots + \lambda_k \varphi_k \otimes \psi_k$$

where

$$\begin{gathered} \lambda_1, \cdots , \lambda_k \geq 0 \\ \lambda_1 + + \lambda_k = 1 \end{gathered}$$

Otherwise, it is called *entangled.* Now we let $|\eta\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ be a unit vector. Then $\langle\eta| \cdot |\eta\rangle$ is separable *if and only if*

$$|\eta\rangle = |\varphi\rangle \otimes |\psi\rangle$$

for unit vectors $|\varphi\rangle \in \mathcal{H}_1, |\psi\rangle \in \mathcal{H}_2$.

6.5.1 Numerical Measure of Entanglement

This section will lead to the Schmidt decomposition and Schmidt number.

We first ask ourselves when is a tensor $|\eta\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ elementary? Choose a basis $|\varphi_1\rangle , \cdots , |\varphi_n\rangle \in \mathcal{H}_1$. Then we have

$$|\eta\rangle = \sum_{k=1}^{n} |\varphi_k\rangle \otimes |\psi_k\rangle$$

for uniquely defined $|\psi_k\rangle \in \mathcal{H}_2$. They are uniquely defined because we have a linear isomorphism, (1 to 1 mapping) $\mathcal{H}_2^n \to \mathcal{H}_1 \otimes \mathcal{H}_2$. Next we show by assumption what happens if $|\eta\rangle$ is an elementary tensor. If it is elementary then

$$|\eta\rangle = |\varphi\rangle \otimes |\psi\rangle$$

and we can write $|\varphi\rangle = \sum_{k=1}^{n} a_k |\varphi_k\rangle$. Then we have

$$|\eta\rangle = \big(\sum_{k=1}^{n} a_k |\varphi_k\rangle\big) \otimes |\psi\rangle = \sum_{k=1}^{n} |\varphi_k\rangle \otimes a_k |\psi\rangle$$

and hence

$$|\psi_k\rangle = a_k |\psi\rangle$$

for $k = 1, \cdots, n$. We could repeat the same argument if we had a basis in $\mathcal{H}_2$ to start with instead, but the conclusion would still be the same:

If $|\eta\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ and $|\varphi_1\rangle, \cdots, |\varphi_n\rangle$ is a basis in $\mathcal{H}_1$, and

$$|\eta\rangle = \sum_{k=1}^{n} |\varphi_k\rangle \otimes |\psi_k\rangle \tag{8}$$

then $|\eta\rangle$ is an elementary tensor if and only if $|\psi_1\rangle, \cdots, |\psi_n\rangle$ lie in a one-dimensional subspace of $\mathcal{H}_2$. Said another way: pure states can be represented by a ray in Hilbert space. The above argument leads to a numerical measure of entanglement: the *Schmidt number*. If we take $|\eta\rangle$ as (8) then the number

$$m = \dim(\{|\psi_1\rangle, \cdots, |\psi_n\rangle\})$$

is called the Schmidt number of $|\eta\rangle$. This number **does not** depend on the choice of basis.

One can rephrase the above conclusion in terms of the Schmidt number: $|\eta\rangle$ is an elementary tensor (separable state) *if and only if* the Schmidt number $m = 1$.

Another rephrasing is: If $|\eta\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ has Schmidt number m, then we can choose linearly independent vectors

$$\begin{aligned} &|\theta_1\rangle, \cdots, |\theta_m\rangle \in \mathcal{H}_1 \\ &|\omega_1\rangle, \cdots, |\omega_m\rangle \in \mathcal{H}_2 \quad \text{s.t.} \\ &|\eta\rangle = \sum_{i=1}^{m} |\theta_i\rangle \otimes |\omega_i\rangle \end{aligned}$$

7 Quantum Computing and Algorithms

In this section we start to deal with the actual subject of quantum computation, and look into various algorithms that are able to to things that we cannot do classically. Exponential speedups and whatnot.

7.1 Introduction to Quantum Computing

7.1.1 One-Qubit

We define a qubit as any quantum mechanical system with a two-dimensional state space $\mathcal{H}$, and a distinguished orthonormal basis $|0\rangle, |1\rangle$ in $\mathcal{H}$. In reality the state space is infinite-dimensional, but we ignore some degrees of freedom of the system in order to consider only a two-dimensional part of it.)

The distinguished basis $|0\rangle, |1\rangle$ is called the computational basis. By measuring the state of the qubit we mean measuring the observable $|1\rangle\langle 1| =$ *the orthogonal projection onto* $\mathbb{C}|1\rangle \subset \mathcal{H}$. Thus, is the qubit is in the state

$$a|0\rangle + b|1\rangle, \qquad (a, b \in \mathbb{C}, |a|^2 + |b|^2 = 1)$$

then the result of a measurement is 0 or 1 with probabilities $|a|^2$ or $|b|^2$ respectively.

7.1.2 N-Qubits

Next we consider composing a system with n qubits we have a state space, as discussed in (6.4),

$$\mathcal{H}^{\otimes n} = \underbrace{\mathcal{H} \otimes \cdots \otimes \mathcal{H}}_{n}$$

This space has a distinguished basis consisting of vectors

$$|i_1\rangle \otimes |i_2\rangle \otimes \cdots \otimes |i_n\rangle \underset{\text{denote}}{=} |i_1 \cdots i_n\rangle$$

for all $i_1, \cdots, i_n = 0, 1$. This basis is again called the *computational basis.* By measuring the state of n qubits we mean a measurement of the states of all qubits. I.e. we measure the observables

$$I \otimes \cdots \otimes \underset{k}{\big(|1\rangle\langle 1|\big)} \otimes I \otimes \cdots \otimes I$$

Thereby, if we take $|\eta\rangle$ as

$$|\eta\rangle = \sum_{i_1, \cdots, i_n = 0}^{1} a_{i_1 \cdots i_n} |i_n \cdots i_n\rangle$$

then the result of the measurement is a bit-string $i_1 \cdots i_n$ with probability $|a_{i_1 \cdots i_n}|$

7.1.3 Quantum Gates

In a parallel to classical logic gates we also have quantum gates. The main difference is that quantum gates are represented by unitary matrices (4.1.4) and as such they must necessarily be invertible. In addition, the quantum gates can be thought of as time evolving (6.3) our qubits by a Hamiltonian defined by the quantum gate. We refer to a series of such quantum gates as a *quantum circuit.* Since all unitary matrices are invertible, quantum computing is closely linked with the classical field of reversible computing (ex: XOR-gate is irreversible). There are tricks to make arbitrary Boolean circuits into quantum circuits which we discuss in (7.6). Some common quantum gates are:

The Pauli gates

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

$$\sigma_Y = Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

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

The Hadamard gate

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

The controlled X-gate or CNOT

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

The CNOT gate is an example of a two qubit gate. All single qubit gates can be 'controlled' in the same way as the CNOT-gate. What the CNOT gate does is that depending on the input to the 'first' qubit, in terms of the computational basis states $|0\rangle, |1\rangle$, it either applies the X-gate to the second qubit or not. Controlled gates introduce entanglement into our quantum circuit.

The SWAP-gate "swaps" two qubits

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

Making Controlled-U Gates To make controlled U-gates for an arbitrary unitary U, you simply add whatever gate you wish to control into a matrix as such

$$U = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}$$

$$C_{U_{1\to 2}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & u_{11} & u_{12} \\ 0 & 0 & u_{21} & u_{22} \end{bmatrix}$$

If you want to switch which qubit is the control-qubit [35] then you transform the above C_U as such

$$C_{U_{2\to 1}} = \text{SWAP}C_U\text{SWAP}$$

Gate Identities Here is a list of some commonly used gate identities of the Pauli gates

$$\begin{aligned} \sigma_X\sigma_Y = -\sigma_Y\sigma_Z = iZ \\ \sigma_Y\sigma_Z = -\sigma_Z\sigma_Y = iX \\ \sigma_Z\sigma_X = -\sigma_Z\sigma_X = iY \\ \sigma_X\sigma_Y\sigma_Z = iI \end{aligned}$$

Gate identities with the Hadamard gate

$$\begin{aligned} H\sigma_X H = \sigma_Z \\ H\sigma_Y H = -\sigma_Y \\ H\sigma_Z H = \sigma_X \end{aligned}$$

A good source `https://people.math.gatech.edu/~jeanbel/4782/Year09/problem207.pdf`

7.1.4 Universal Gate Sets

In quantum computations it is neither realistic nor needed to realize every unitary gate precisely. It is enough to be able to approximate every gate with an arbitrary good precision.

Definition: A collection $\mathcal{A}$ of quantum gates on $\mathcal{H}^{\otimes n}$ closed under inversion is called a *universal gate set*, or a complete basis, if it generates a dense subgroup of $PU(\mathcal{H}^{\otimes n})$. That is, if U is a unitary on $\mathcal{H}^{\otimes}$, then for all ($\forall$) $\epsilon > 0$ we can find $\phi \in \mathbb{R}$ and unitaries $U_1, \cdots, U_m \in \mathcal{A}$ such that

$$||U - e^{i\phi}U_1 \cdots U_m|| < \epsilon$$

For every $n \geq 2$ the one-qubit gates $H, T = \sqrt[4]{Z}, T^{-1} = T^3$ and the 2-qubit CNOT gates acting on neighboring qubits form a universal gate set on $\mathcal{H}^{\otimes}$.

There is nothing special about H and T. It can be shown that 2 any randomly chosen unitary 2×2 matrices generate a dense subgroup of $PU(2)$ with probability 1. We divide the proof for H and T into several lemmas.

Lemma 1 Assume $\vec{v}, \vec{w} \in S^2 \subset \mathbb{R}^3$ are mutually orthogonal. Then every $U \in SU(2)$ can be written as

$$U = R_{\vec{v}}(\alpha)R_{\vec{w}}(\beta)R_{\vec{v}}(\gamma) =$$

That is, every operator in $SU(2)$ can be written as a product of three rotations about two axes $\vec{v}, \vec{w}$ by the angles α, β, γ. This makes intuitive sense. Full proof in the lecture notes [16]. The rotation operators were defined in 4.4.3.

Lemma 2 Take $\vec{v} \in S^2 \subset \mathbb{R}^3$. Assume $\theta \in \mathbb{R}$ such that $\frac{\theta}{2\pi}$ is irriational. Then for any $\phi \in \mathbb{R}$ and $\epsilon > 0$ there is $n \in \mathbb{N}$ such that

$$||R_{\vec{n}}(\phi) - R_{\vec{v}}(\theta)^n|| < \epsilon$$

This is saying that by repeated rotations by an angle θ we can get arbitrarily close to another rotation by an angle ϕ. Another way to think about this is by thinking of the unit circle S^1. What it is saying is we can rotate around the unit circle to any other number. (still a bit unclearly explained but the gist is there.)

7.1.5 Quantum Circuits

Running a quantum computer means that we have n-qubits represented by $\mathcal{H}^{\otimes n}, \dim(\mathcal{H}) = 2$. We initialize the quantum computer, usually in the state $|0\cdots 0\rangle = |0\rangle \otimes \cdots \otimes |0\rangle$. Then we apply a series of gates, unitary operators on $\mathcal{H}^{\otimes n}$, which perform some calculation. We can represent this sequence of gates graphically. For each qubit we have a wire, which we read from left to right interpreting this as 'time steps' in our calculation.

Adding later some diagrams probably, dont need to do it before exam tbh, bunch of diagrams below...

7.2 Quantum Fourier Transform

For this section we consider an n-qubit system. We identify every bit string $j_1 \cdots j_n$ with the integer it represents in binary

$$2^{n-1}j_1 + 2^{n-2}j_2 + + \cdots + 2^0 j_n$$

therefore the computational basis in $\mathcal{H}^{\otimes n}$ is $|0\rangle, \cdots, |N-1\rangle$, where we take $N = 2^n$. From our discussion on the discrete Fourier we know that having N be divisible by two is beneficial. In reality, the quantum Fourier transform is nothing more than the discrete Fourier transform in the mentioned computational basis. Now we tackle the question of how to implement the quantum Fourier transform. For reasons similar to the argumentation regarding the discrete Fourier transform 3.2.1, it is more convenient to work with a slightly changed gate. We denote F'_N to mean the composition of F_N and the operator that reverses the order of the qubits $|j_1 \cdots j_n\rangle \to |j_n \cdots j_1\rangle$.

Lemma: We have

$$\begin{aligned} F'_{2N} &= \frac{1}{\sqrt{2}}\begin{bmatrix} F'_N & F'_N \\ F'_N A_N & -F'_N A_N \end{bmatrix} \\ &= \begin{bmatrix} F'_N & 0 \\ 0 & F'_N \end{bmatrix}\begin{bmatrix} I_N & 0 \\ 0 & A_N \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} I_N & I_N \\ I_N & -I_N \end{bmatrix} \end{aligned}$$

This last equality is proved in the lecture notes [17], and where A_N is the same matrix we talked about in the discrete Fourier section 3.2.1. We now observe that the last matrix is exactly

$$\frac{1}{\sqrt{2}}\begin{bmatrix} I_N & I_N \\ I_N & -I_N \end{bmatrix} = H \otimes \underbrace{I \otimes \cdots \otimes I}_{n}$$

Therefore we can write this whole lemma as the circuit

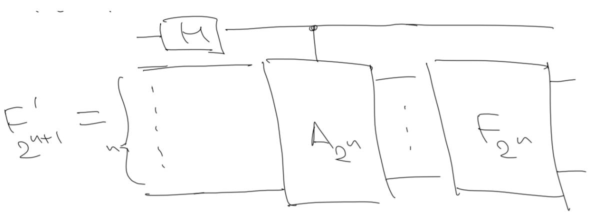


Figure 3: Lemma circuit from lecture 18 [17]

Next we deal with the controlled A_{2^n} gate. It is the operator

$$A_{2^n}\,|j_1 \cdots j_n\rangle = e^{\frac{2\pi i}{2^{n+1}}(2^{n-1}j_1+\cdots+j_n)}\,|j_1 \cdots j_n\rangle$$

What it essentially is, is a composition of two-qubit controlled rotations by factors of

$$e^{\frac{2\pi i}{2^{n+1}}2^{n-k}j_k}$$

This can be realized with the controlled z-rotation gate

$$R(\theta) = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{i\theta}{2}} \end{bmatrix}$$

Thus we are able to write the circuit for F'_{2N} as

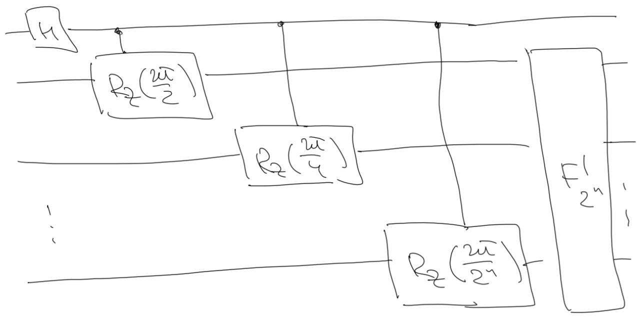


Figure 4: The F'_{2N} gate circuit. Image from lecture 18 [17]

Recursively, this gives a decomposition of $F_{2^{n+1}}$ into $n+1$ Hadamard gates and $\frac{n(n+1)}{2}$ controlled rotation gates.

7.3 Shor's Period Finding Algorithm

Let us summarize Shor's period finding algorithm! *Implementation paper* [25]

We have a function

$$f : \mathbb{Z}_+ \to \{0, 1, \cdots, 2^m - 1\}$$

such that $f(x) = f(y)$ **if and only if**

$$x \equiv y \mod r \quad \text{i.e.} \quad r|(x-y)$$

where we understand the last piece of notation as: $x - y$ *is divisible by* r. We call r the period of the function f.

We operate on a quantum computer with $n \times m$ qubits, where n is such that $N = 2^n > r^2$. We assume that we have a gate V_f, *a quantum black box*, computing f in the sense that

$$V_f\big(\,|x\rangle \otimes |0^m\rangle\,\big) = |x\rangle \otimes |f(x)\rangle$$

The algorithm is as follows:

1) We run the following circuit twice:

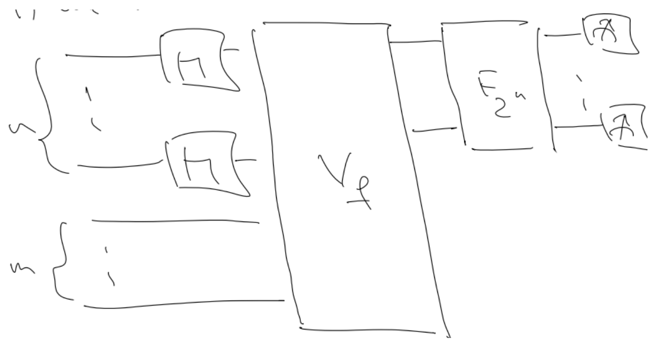


Figure 5: Drawing from mat3420 lecture 22 spring 2021

and from this we get two numbers k and l:

$$(0 \leq k, l \leq N-1)$$

2) We compute all the convergents of the continued fraction expansions of k/N and l/N and we call these p_i/q_i and r_j/s_j respectively [26].

3) Compute the least common multiple, $\text{lcm}(q_i, s_j)$, for all pairs of i, j. [27], [28].

4) Test the numbers $\text{lcm}(q_i, s_j)$ to see if they are multiples of r. That is, we check whether

$$f(0) = f\big(\text{lcm}(q_i, s_j)\big)$$

If this is the case, we return the smallest such $\mathrm{lcm(q_i, s_j)}$ as the answer of our algorithm.

7.3.1 Euclid's Algorithm

I have linked to the wiki page for this algorithm, but since we covered it in lectures I also felt like writing down the explanation given there.

We deal with the Euclidean algoerithm for computing $lcm(q, s)$, which is very close to the algorithm producing the continued fraction expansions of $\frac{q}{s}$.

We have

$$lcm(q, s) = \frac{qs}{gcd(q, s)}$$

where *gcd* is the greatest common divisor of q, s. The algorithm goes as follows:

Assume $q > s$. Then divide q by s with remainder

$$q = as + q_1, \quad 0 \leq q_1 < s$$

If $q_1 = 0$ then we are done and $gcd(q, s) = s$. Else we continue with

$$gcd(q, s) = gcd(s, q_1)$$

Again, we divide with remained, only this time we do

$$s = bq_1 + s_1, \quad 0 \leq s_1 < q_1$$

Again if $s_1 = 0$ we are done and $gcd(s, q_1) = q_1 = gcd(q, s)$. Else we continue as earlier only this time we have

$$gcd(q, s) = gcd(s, q_1) = gcd(q_1, s_1)$$

And the algorithm continues exactly like this until we have found the greatest common divisor of q and s. The algorithm is simply just a 'cycling' through dividing with remained, checking if the remainder is zero, and then dividing again with remainder if not. *We may conclude that to find $gcd(q, s)$ we need not use more than*

$$2 \log_2(\max\{q, s\}) + 1$$

divisions.

7.3.2 Efficiency of Shor's Period Finding Algorithm

1) In the circuit, apart from the black box V_f, we have the F_{2^n}-gate whose composition is $\mathcal{O}(n^2)$ 1-and 2-qubit gates. Amongst the 2-qubit gates we require controlled rotation gates which need to be approximated, for implementational simplicity from a universal gate set. The Solovay-Kitaev algorithm can be used to approximate these controlled rotations by $\sim \mathcal{O}(n^6)$ quantum gates.
2) The number of arithmetic operations needed to calculate the convergents of the continued fractions is of order $\mathcal{O}(n^2)$
3) Further, the number of arithmetic operations needed to find all $lcm(q_i, s_j)$ is $\mathcal{O}(n^3)$, because the number of such pairs is at most $\mathcal{O}(n^2)$.
4) Following the above statement the maximum number of calls to f to check if $f(0) = f(lcm(q_i, s_j))$ is of order $\mathcal{O}(n^2)$ also because of the number of such pairs.

To summarize, we need the following numbers of operations for every run of the period-finding algorithm:

- 2 calls of V_f
- $\mathcal{O}(n^6)$ gates from a universal gate set
- $\mathcal{O}(n^3)$ arithmetic operations
- $\mathcal{O}(n^2)$ calls of f

7.3.3 Finalizing Shor's Period Finding Algorithm

To reiterate we have the setup:
We are given a natural number n, n is odd, n is not a power of a prime number. Let m be the number of digits in the binary form of n

$$2^{m-1} \leq n < 2^m$$

Let $2 \leq x \leq n-1$ be a natural number relatively prime to n. We consider the function

$$\begin{aligned} f &: \mathbb{Z}_+ \to \mathbb{Z}_n^* \subset \mathbb{Z}_n \\ f(k) &= x^k \mod (n) \end{aligned}$$

The goal is to find the period r of this function. We have $r \leq n-1$, $r^2 < n^2 < 2^{2m}$. To apply Shor's period finding algorithm it is enough to construct a gate

$$\begin{aligned} V_f &: \mathcal{H}^{\otimes 3m} \to \mathcal{H}^{\otimes 3m} \\ &|k\rangle \otimes |0\rangle \to |k\rangle \otimes |f'(k)\rangle \end{aligned}$$

In fact we end up needed $(3m + L_m)$ qubits instead of the $3m$ above and we end up with the map

$$|k\rangle \otimes |1\rangle \otimes |0\rangle \to |k\rangle \otimes \left|x^k \mod (n)\right\rangle \otimes |0\rangle$$

7.3.4 Discussion on the V_f Quantum Gate

This is probably one of the most difficult and complex subjects I have learned in my life so far. This whole discussion of Shor's period finding and factoring algorithm in fact! Nonetheless here we are :)

For this next part I feel the background reasoning is pretty opaque.

1) Consider the function

$$\begin{aligned} F_a &: \{0,1\}^m \to \{0,1\}^m \\ F_a(b) &= \begin{cases} (a+b) \mod (n) & \text{if } 0 \leq b \leq n \\ b & \text{if } n \leq b \leq 2^m \end{cases} \\ &\text{if } 0 \leq a \leq n: \\ F_a(b) &= \begin{cases} a+b & \text{if } 0 \leq b < n-a \\ a+b-n & \text{if } n-a \leq b < n \\ b & \text{if } n \leq b \leq 2^m \end{cases} \end{aligned}$$

where a is an integer. The above function is reversible, and we note that $F_a = F_{a \mod (n)}$ and $F_a^{-1} = F_{-a}$. This last fact will be especially important when combining our derivation with discussion in ((9)) (<u>second conclusion</u>). From this conclusion we know it is possible to construct F_a as a quantum circuit operating on $m + L'_m$ qubits.

$$|b\rangle \otimes |0\rangle \to |F_a(b)\rangle \otimes |0\rangle$$

From now we denote this quantum circuit for F_a by the name $\text{ADD}_n(a)$. So we now have a circuit which computes addition of two numbers a and b mod n.

2) We wish to compose this circuit in a way such that we get a new 'bigger' circuit which computes multiplication of two numbers a and b mod n. We can pretty much do this by phrasing binary multiplication in terms of a sequence of additions using our now defined $\mathrm{ADD}_n(a)$ circuit. We observe that with b in binary form $b = 2^{m-1}b_1 + \cdots + 2^1 b_{m-1} + 2^0 b_m$ then

$$ab = (2^{m-1}a)b_1 + \cdots + (2^1 a)b_{m-1} + 2^0 ab_m$$

Consider the quantum circuit

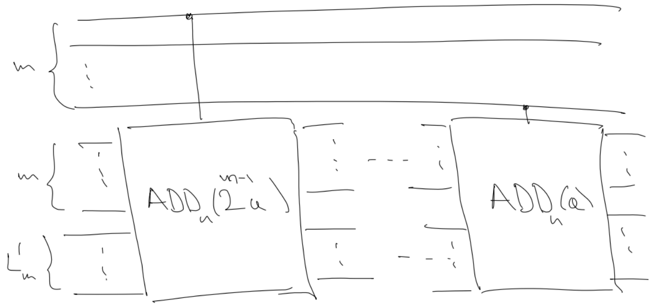


Figure 6: Drawing from lecture 27 [21]

It maps

$|b\rangle \otimes |c\rangle \otimes |0\rangle \to |b\rangle \otimes |(ab + c) \mod (n)\rangle \otimes |0\rangle$
Particularly :
$|b\rangle \otimes |0\rangle \otimes |0\rangle \to |b\rangle \otimes |ab \mod (n)\rangle \otimes |0\rangle$

Now if we have an a' such that

$$aa' \equiv 1 \mod (n) \quad \text{then } a'ab \equiv b \mod (n)$$

And so we have that the inverse of the map above is

$$b \to a'b \mod (n)$$

This allows us to use the trick from ((9)), which allows us to construct a circuit

$$|b\rangle \otimes |0\rangle \otimes |0\rangle \to |ab \mod (n)\rangle \otimes |0\rangle \otimes |0\rangle$$

for all $0 \leq b < n$. This above circuit we now reference as U_a. *It is important to remember this U_a circuit is in turn m $\mathrm{ADD}_n(a)$ circuits after one another like in figure (6)*, and a few details were glossed over here, they are further explained in the lecture notes [21].

3) If we take $0 \leq k < 2^{2m}$, where k is in the binary representation $k = 2^{2m-1}k_1 + \cdots + 2^1 k_{2m-1} + 2^0 k_{2m}$ then

$$x^k b = x^{2^{2m-1}k_1} \cdot \ldots \cdot x^{2^1 k_{2m-1}} x^{2^0} \cdot b$$

SHOULD THERE BE A k_{2m} IN THE ABOVE LAST TERM? Maybe not.. but i feel it would fit with pattern. Therefore the circuit

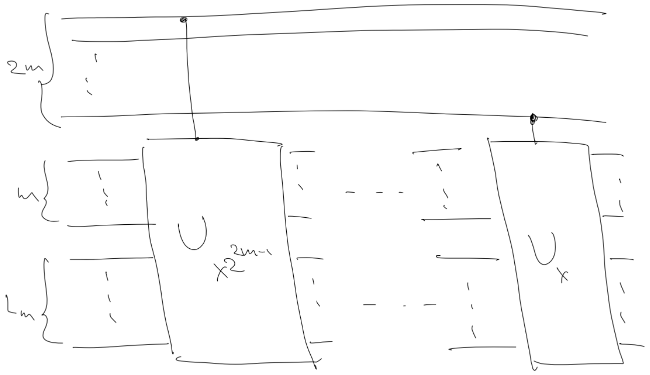


Figure 7: Final circuit for the implementation of V_f in Shor's period finding algorithm. This figure is from lecture 27 [21].

maps

$$|k\rangle \otimes |b\rangle \otimes |0\rangle \to |k\rangle \otimes \left|x^k b \mod (n)\right\rangle \otimes |0\rangle$$

In particular, when we take $b = 1$, it maps

$$|k\rangle \otimes |1\rangle \otimes |0\rangle \to |k\rangle \otimes \left|x^k \mod (n)\right\rangle \otimes |0\rangle$$

Which is exactly what we set out to obtain! This last circuit operates on $3m + L_m$ qubits, and uses $\mathcal{O}(m^3)$ simple gates. This is a very good improvement over classical circuits for which all known algorithms require a number of operations which is exponential in m!

7.4 RSA Algorithm

One of the most used cryptosystems is the so-called RSA (Rivest - Shamir - Adleman) algorithm. The idea is as follows:

Suppose we want to encrypt and send a message. We represent the message as a bit-string, which we can think of as a number m. (*m for message, this could for example be the bit values of a text file.*) First, we choose a number $n > m$, it can be shown that for all m relatively prime (*co-prime*) to n there is a number $k \geq 2$ such that

$$m^k \equiv 1 \mod (n)$$

The smallest such number k is denoted by $\lambda(n)$, where λ is *Carmichael's totient function.* Now take any number $e \geq 2$ relatively prime to $\lambda(n)$, *the smallest of the k's*, then we can find numbers d, a such that

$$ed = 1 + a\lambda(n)$$

then we have

$$m^{ed} = m^{1+a\lambda(n)} = m\big(m^{\lambda(n)}\big)^a \equiv m \mod (n)$$

In RSA one takes $n = pq$ for different primes p and q. To encrypt a message m we take

$$x = m^e \mod (n)$$

that is x is the remainder of the division of m^e by n. Then to decrypt the message x we take

$$x^d \mod (n)$$

The numbers e and n are kept open, often $e = 2^{16} + 1 = 65537$ is used, while the numbers $d, p, q, \lambda(n)$ are kept secret. Therefore anyone can encrypt the message, but to decrypt one needs to know d. We call the number e the *public key* and the number d the *private key.*

For $n = pq$, we have $\lambda(n) = lcm(p-1, q-1)$. For the most naive way of breaking RSA one would try numbers one by one up to $\sqrt{n}$ to see if we get divisors. This requires $\mathcal{O}(e^{\frac{1}{2}\log(n)})$ operations. When key sizes are typically 2,048 to 4,096bits, meaning a decimal number with somewhere around 200 digits then this already takes longer than there will be protons in the universe to do. The best classical algorithm we have for '*breaking*' RSA is the *randomized number field sieve.* This algorithm requires around $\mathcal{O}(e^{2(\log(n)^{\frac{1}{3}}})$ operations, still way to slow!

7.5 Shor's Factoring Algorithm

In section (7.4) we discussed one of the most common crypto-systems, and the naive / best case scenario for 'breaking' it using classical computing. In this section we discuss Shor's factoring algorithm, in which Shor's period finding algorithm discussed in (7.3) is the most essential sub-routine.

We first introduce some helpful notation, first we have integers $\mathbb{Z}_n = \{0, 1, \cdots, n-1\}$. Next we define, $\mathbb{Z}_n^* \subset \mathbb{Z}_n$ the subset of numbers relatively prime to n. Now, take $m \in \mathbb{Z}_n^*, m \neq 1$ and consider the function

$$f : \mathbb{Z}_+ \to \mathbb{Z}_n^*, \quad f(k) = m^k \mod (n)$$

Shor's factoring algorithm for finding a divisor of n

1) Take a random number $m \in \mathbb{Z}_n, m \geq 2$
2) Compute *gcd*(m, n) to check if we get a non-trivial divisor of n, if so we are done. If not we have $m \in \mathbb{Z}_n^*$ and we continue
3) Use Shor's period-finding algorithm to find, r of m, the period / order of the function $f(k) = m^k \mod (n)$ **SEE EXERCISE 7 (NOT ADDED YET)**
4) If r is even, compute the numbers

$$gcd(m^{\frac{r}{2}} - 1, n) \quad \text{and} \quad gcd(m^{\frac{r}{2}} + 1, n)$$

if one of these numbers is different from $1, n$, we have a non-trivial divisor of n!

Now onto our conclusion of Shor's factoring algorithm. Provided n is odd and in the prime factorization $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ for $k \geq 2$, the probability of failure for the factoring part of the algorithm, *by which we mean the proportion of* $1 \leq m \leq n-1$ *for which the algorithm does not return a non-trivial factor of* n, is

$$\text{prob} \leq \frac{1}{2^{k-1}} \leq \frac{1}{2}$$

So by applying the algorithm a few times we have probability of success arbitrarily close to 1.

7.5.1 Orders of Elements in $\mathbb{Z}_n^*$

We denote by $\text{ord}_n(a)$ the order of $a \in \mathbb{Z}_n^*$. By this we mean the smallest number ≥ 1 such that

$$a^{\text{ord}_n(a)} \equiv 1 \mod (n)$$

It can be shown that

$$a^k \equiv a^l \mod (n)$$

if and only if

$$\text{ord}_n(a) | (k - l)$$

in other words if and only if $(k-l)$ is divisible by $\text{ord}_n(a)$

7.6 From Boolean to Quantum Circuits

A classical algorithm computing the function f has a k-bit number as input, and a l-bit number as output.

$$f : \{0,1\}^k \to \{0,1\}^l$$

The definition of what we mean by a Boolean circuit is

- The computation is done in n steps
- The result r_i at the i-th step is obtained by applying a logic gate g_i to some bits x and the previous results $r_1, \cdots, r_{i-1}$
- $f(x)$ is obtained by taking some bits of $r_1, \cdots, r_n$

A presentation of f in the above form is called a Boolean or logic circuit. To execute such a circuit we need to operate on $(k + L + l)$ bits, where the k bits represent the input bits, the L bits represent the ancilla or intermediary bits, and the l bits represent the output bits. Our question is given a Boolean circuit can we construct a quantum circuit? The answer is yes, with the caveat that quantum gates have the requirement that they be reversible. But there is indeed a trick to produce a reversible gate out of any arbitrary gate, the reversible gate will often be a bit more complicated though. What we do is for any logic gate $g(x)$ we define

$$g_\oplus(x, y) : \{0,1\}^{k_g + l_g} \to \{0,1\}^{k_g + l_g}$$
$$g_\oplus(x, y) = (x, g(x) \oplus y)$$

Now given an arbitrary classical Boolean circuit on $(k + L + l)$ bits computing a function f as described above, we can use the reversible trick to make the whole circuit reversible. We now refer to the reversible version of our logic gates as

$$f_1, \cdots f_n : \{0,1\}^{k+L+l} \to \{0,1\}^{k+L+l}$$

meaning that if in the original irreversible formulation we at the i-th step have a logic gate g which acts on input bits $s_1, \cdots, s_a$ with output bits $t_1, \cdots, t_b$ then f_i acts on the exact same bits but with the logic gate $g_\oplus$.

The composition of all these f's: $f_n \circ \cdots \circ f_1$ has the property (quite expected):

$$\big(f_n \circ \cdots \circ f_1\big)(x, 0, 0) = (x, G(x), f(x))$$

where x is the input, $G(x)$ is what we end up with in the ancilla bits, and $f(x)$ is the output.

Notice this function is reversible as we keep the input 'stored'. An example of a classical irreversible function is the XOR function, whose reversible counterpart is the CNOT function [29].

As every part of our circuit/algorithm/function f_i is reversible we can now define a quantum circuit on $(k + L + l)$ qubits consisting of n unitary quantum logic gates U_i

$$U_i \left|x, y, z\right\rangle = \left|f_i(x, y, z)\right\rangle$$

It would be nice if we were done here, but due to our circuit now being implemented with quantum gates, obeying quantum mechanics brings some troubles. Mainly, for *entangled states* the ancilla qubits are going to affect measurements of our output qubits.

7.6.1 Quantum Garbage Removal

To deal with this, we need quantum "garbage removal"; *we need to get rid of* $G(x)$. It turns out this can be done by using l more ancilla bits via the following trick. We redefine our f_i's from earlier as $\tilde{f}_i : \{0,1\}^{k+L+2l} \to \{0,1\}^{k+L+2l}$

$$\begin{aligned}\tilde{f}_i(x, a, y, z) &= \big(f_i(x, a, y), z\big)\\ &k, L, l, l\\ \tilde{f}_{n+1}(x, a, y, z) &= \big(x, a, y, y \oplus z\big)\end{aligned}$$

then we have the compositions

$$\begin{aligned}\big(\tilde{f}_n \circ \cdots \circ \tilde{f}_1\big)(x, 0, 0, z) &= \big(x, G(x), f(x), z\big)\\ \big(\tilde{f}_{n+1} \circ \tilde{f}_n \circ \cdots \circ \tilde{f}_1\big)(x, 0, 0, z) &= \big(x, G(x), f(x), f(x) \oplus z\big)\end{aligned}$$

from this it follows that

$$\begin{aligned}\big(\tilde{f}_1^{-1} \circ \cdots \circ \tilde{f}_n^{-1} \circ \tilde{f}_{n+1} \circ \tilde{f}_n \circ \cdots \circ \tilde{f}_1\big)(x, 0, 0, z)&\\ = \big(x, 0, 0, f(x) \oplus z\big)&\end{aligned}$$

for all x, z. When $z = 0$ we get the result $(x, 0, 0, f(x))$, so we have successfully avoided garbage in the ancilla bits. The cost of this modification is not detrimentally high: for every $g_\oplus$ used to define our f_i's we also need to use one $g_\oplus^{-1} = g_\oplus$ and l CNOT gates to express $\tilde{f}_{n+1}$.

And so we have arrived at our **first conclusion**: *given a Boolean circuit with k input bits, L ancilla bits, and l output bits computing a function f using n logic gates g, we can construct a quantum circuit that operates on $(k + L + l)$ qubits using $2n$ unitary quantum gates $g_\oplus$ and l CNOT gates*

$$\left|x\right\rangle \otimes \left|0\right\rangle \otimes \left|0\right\rangle \otimes \left|z\right\rangle \to \left|x\right\rangle \otimes \left|0\right\rangle \otimes \left|0\right\rangle \otimes \left|f(x) \oplus z\right\rangle$$

Sometimes, *as with the discussion of V_f in Shor's period finding algorithm*, we want to get rid of the input state $\left|x\right\rangle$ and have a map of the form

$$\left|x\right\rangle \otimes \left|0\right\rangle \to \left|f(x)\right\rangle \otimes \left|0\right\rangle$$

To do this we need a few extra steps, a more detailed description can be found in lecture 26 [20]. The conclusion drawn there is the following.

Second conclusion: *We have a reversible function f and Boolean circuits computing both f and f^{-1}, both using L ancilla bits and n logic gates g. Then we can construct a quantum circuit on $(k + L + 2k = 3k + L)$ qubits, which uses at most $4n$ unitary quantum gates defined by $g_\oplus$, $2k$ CNOT gates and k SWAP gates such that*

$$\begin{array}{c}\left|x\right\rangle \otimes \left|0\right\rangle \to \left|f(x)\right\rangle \otimes \left|0\right\rangle\\ \uparrow\\ \mathcal{H}^{\otimes(2k+L)}\end{array} \tag{9}$$

Qiskit reference [30].

7.7 Grover's Search Algorithm

In this section we discuss the Grover search algorithm which offers a quadratic speed up in unstructured search problems. The setup is as follows. Most of this section is based on the relevant chapter in [1], Lecture 28 [22] and the Qiskit textbook [31].

Suppose we have 2^n boxes, which lets us enumerate them by bit-strings of length n, and we know that in one of these boxes there is an item we wish to find. We call this x^*. Classically, one cannot *on average* do better than spending $N/2$ steps opening these boxes, with the worst case situation being opening all N. On a quantum computer it is possible to find the item x^* using $\sqrt{N}$ steps using Grover's amplitude amplification trick. In addition to offering a quadratic speed up Grover's algorithm does not at all depend on the internal structure of our 'list' and as such it offers an immediate quadratic speed up for many classical algorithms.

Consider the function

$$f : \{0,1\}^n \to \{0,1\}$$

$$f(x) = \begin{cases} 0 & \text{if } x \neq x^* \\ 1 & \text{if } x = x^* \end{cases}$$

Suppose we operate on n qubits and have an implementation of the above function f as a quantum gate

$$V_f \left|x\right\rangle = (-1)^{f(x)} \left|x\right\rangle$$

Such an implementation is commonly referred to as an oracle. Grover's algorithm solves oracles that add a negative phase to the solution states.

Because the list is unstructured and we have no idea where x^* is, we start of our algorithm with a uniform superposition

$$|Q\rangle = \sum_x |x\rangle$$

Now the point of the algorithm is to increase the amplitude corresponding to the box we want x^*, which will simultaneously shrink the amplitudes for all the other boxes (states). The way we do this is with what we referred to as *Grover's amplitude amplification trick* earlier. This has a nice geometric interpretation in terms of two reflections, i.e. a rotation.

We consider the vectors $|x^*\rangle$ and $|Q\rangle$, basically what we want to do is rotate $|Q\rangle$ onto $|x^*\rangle$. Consider the operator

$$D = 2\,|Q\rangle\,\langle Q| - I$$

1) We now think of the two-dimensional subspace W of $\mathcal{H}^{\otimes n}$ spanned by $|Q\rangle$, our starting state, and $|x^*\rangle$ our goal state. This space W is invariant under DV_f 4.2, which means that acting with DV_f on any vector in $\mathrm{span}\{|\mathrm{Q}\rangle\,, |\mathrm{x}^*\rangle\}$ returns another vector in W. Said compactly

$$DV_f W \subset W \implies DV_f W = W$$

because DV_f is invertible.
2) Let $|y\rangle \in W$ be the unit vector such that $|y\rangle \perp |x^*\rangle$ and the angle between $|y\rangle$ and $|Q\rangle$ is $\theta/2$ with

$$\sin\left(\frac{\theta}{2}\right) = \langle x^*|Q\rangle = \frac{1}{\sqrt{N}}$$

and we have

$$(DV_f)|_W = D|_W V_f|_W$$

where

$$V_f|_W \text{is the reflection through } \mathbb{R}\,|y\rangle$$
$$D|_w \text{is the reflection through } \mathbb{R}\,|Q\rangle$$

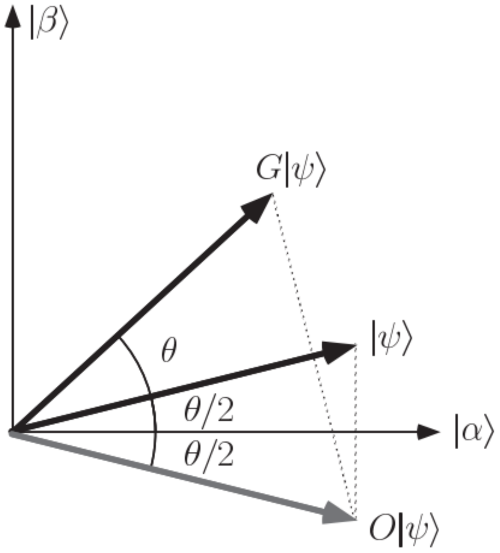


Figure 8: Image from p.253 in Nielsen [1]. Should make my own using the names in this text in the future. Translating the names we have $|Q\rangle \leftrightarrow |\psi\rangle$, $|y\rangle \leftrightarrow |\alpha\rangle$, $|x^*\rangle \leftrightarrow |\beta\rangle$, $DV_f \leftrightarrow G\,|\psi\rangle$. *Do note as they mention in [1] that the length of $|\alpha\rangle\,, |\beta\rangle$ is not correct to make the picture a bit clearer. All vectors are indeed unit vectors.*

Now, let k be such that

$$\frac{\pi}{2k} \leq \theta < \frac{\pi}{2(k-1)}, \qquad \text{so } k \approx \frac{\pi\sqrt{N}}{4}$$

Then $(DV_f)^k\,|Q\rangle$ will be close to $|x^*\rangle$ and with measurement we get the state, or 'box', $|x^*\rangle$ with probability close to 1! That is Grover's quantum search algorithm with quantum circuit

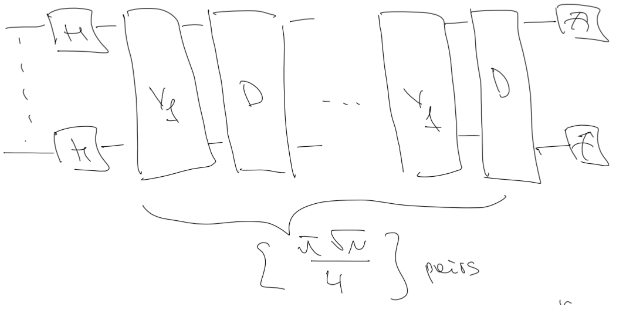


Figure 9: Grover search circuit from lecture 28 [22]

7.8 Quantum Error Correction

In quantum computing there are more possibilities for errors than in classical computing. As such quantum error correction is essential to study to understand quantum computing and its challenges. Error types which are exclusive to quantum computing include: *decoherence, coherent errors, currupt input, and leakage.* (These might be explained in the future but for now I'm skipping the explanation.) First we consider how to mitigate bit-flip errors.

7.8.1 Bit Flip Errors

Classically, one would send the same information multiple times, however the no-cloning theorem 5.6 prevents copying of information in quantum mechanics. Turns out there is a 'way around' this limitation of not being able to send the same information multiple times. The main idea is to encode the state of k-qubits in a state of n-qubits $(n > k)$. In other words, we consider the *isometry*

$$C : \mathcal{H}^{\otimes k} \to \mathcal{H}^{\otimes n}$$

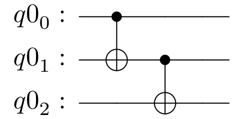
Isometry means that C *preserves the scalar product*

$$\big(C\,|\phi\rangle\,, C\,|\psi\rangle\,\big) = \big(\,|\phi\rangle\,, |\psi\rangle\,\big)$$
$$||C\,|\phi\rangle\,|| = ||\,|\phi\rangle\,||$$

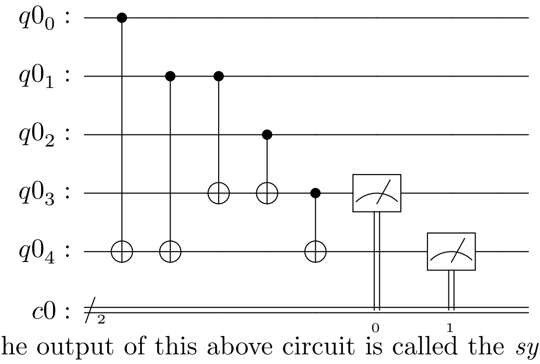
Now to make it a bit concrete we consider the 3-qubit repetition code C_3, we define

$$C_3 : \mathcal{H} \to \mathcal{H}^{\otimes 3}$$
$$C_3\,|0\rangle = |000\rangle$$
$$C_3\,|1\rangle = |111\rangle$$

By identifying $\mathcal{H}$ with $\mathcal{H} \otimes \mathbb{C}\,|00\rangle \subset \mathcal{H}^{\otimes 3}$, we can extend the map $|\phi\rangle \otimes |00\rangle \to C_3\,|\phi\rangle$ to a unitary map on $\mathcal{H}^{\otimes 3}$ as such:



checking for input states $|0\rangle$ and $|1\rangle$ does indeed produce $|000\rangle$ and $|111\rangle$ respectively. Now we extend this circuit to work on 3-qubits, we need two ancilla qubits



The output of this above circuit is called the *syndrome*, and as such the ancilla qubits are often called the *syndrome-bits*. The syndrome, in this case, is a two bit string which will tell us where the error occurred. In the above diagram $q0_0, q0_1, q0_2$ are the *code bits* across which the logical state is encoded, and $q0_3, q0_4$ are the syndrome bits which we measure. So now we can ask what do we get as output in our syndrome bits when we apply this circuit to a state in $\mathcal{H}^{\otimes 3}$?

Input	Syndrome	Decimal
$\lvert 000\rangle$	00	0
$\lvert 001\rangle$	11	3
$\lvert 010\rangle$	10	2
$\lvert 011\rangle$	01	1
$\lvert 100\rangle$	01	1
$\lvert 101\rangle$	10	2
$\lvert 110\rangle$	11	3
$\lvert 111\rangle$	00	0

Table 3: Decimal column is the syndrome value in decimal. We label $|abc\rangle$, $a : 1, b : 2, c : 3$.

So the syndrome tells us in which bit in the bit string of $|\phi\rangle$ the bit flip error has occured, and we can correct this by applying an X-gate to that corresponding qubit. In fact we are able to perform this error correction directly in the circuit as follows

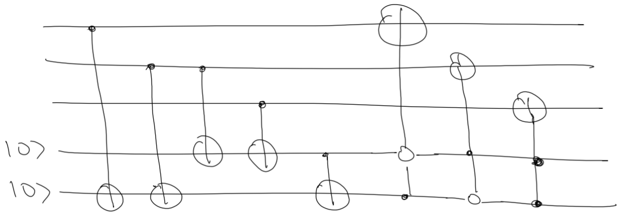
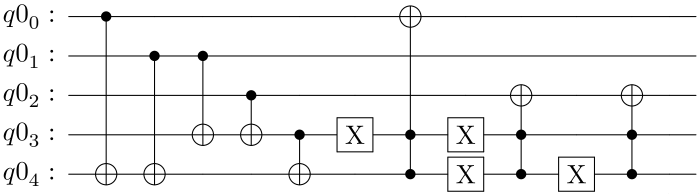


Figure 10: Full 3-qubit repetition code to protect against bit-flip errors. Image taken from lecture 29 [23]

where the white circle in the double control gates means that the control for that qubit is inverse (so it activates when the control = 0). From Nielsen [1] p.184 we see that such an inverse control can be realized by 'sandwiching' a CNOT by two X-gates.. The first 'part' of the circuit performs the bit-flip detection and the last 'part' (with the Toffoli-gates) performs the automatic correction. The above circuit implemented in qiskit would look like



(10)

Thus, if we apply the above circuit to $X_i C_3\,|\phi\rangle$, where X_i is the X-gate applied to the i-th qubit, then we get a state on the form $C_3\,|\phi\rangle \otimes |\psi\rangle$ for some $|\psi\rangle \in \mathcal{H}^{\otimes 2}$. We see that the 3-qubit repetition code is enough to protect against bit-flip errors in one qubit.

7.8.2 Phase Flip Errors

Phase flip errors have no classical counter-part, however the discussion on this type of error mirrors the previous discussion on bit flip errors where we considered flipping by a 'imagined' gate X_i previously we not consider a flip by the gate Z_i-gate. The formulation of our flipping errors as gates provides a fortunate short hand when dealing with phase flips errors. We use the fact that

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

and so we are able to use the previous code C_3 (**??**), and modify it using the above relation to protect against single-qubit phase flip errors. Basically, we wrap the whole of (**??**) in a '*hadamard sandwich.*

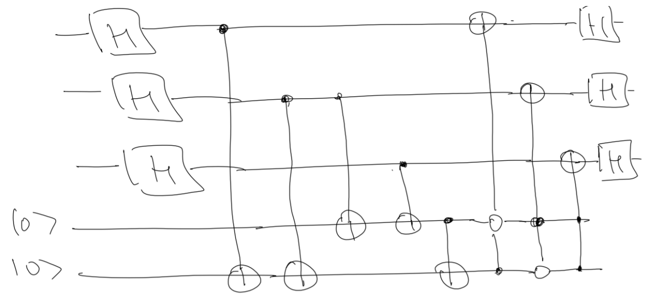


Figure 11: 3-qubit phase flip correction circuit. Image from lecture 29 [23].

7.8.3 Shor's 9 Qubit Code

In the two previous sections we saw that we can protect against single-qubit bit- and phase-flip errors. Now to see how we can combine the concepts to protect against both types of error simultaneously. Shor showed that this is possible if we concatenate the two codes ((10), 11). Our results is a 9-qubit code called *Shor's 9-qubit code. Note, this is for correcting both phase and bit-flip for 1-qubit using 8 ancilla qubits.*

$$C_9 : \mathcal{H} \to \mathcal{H}^{\otimes 9}$$

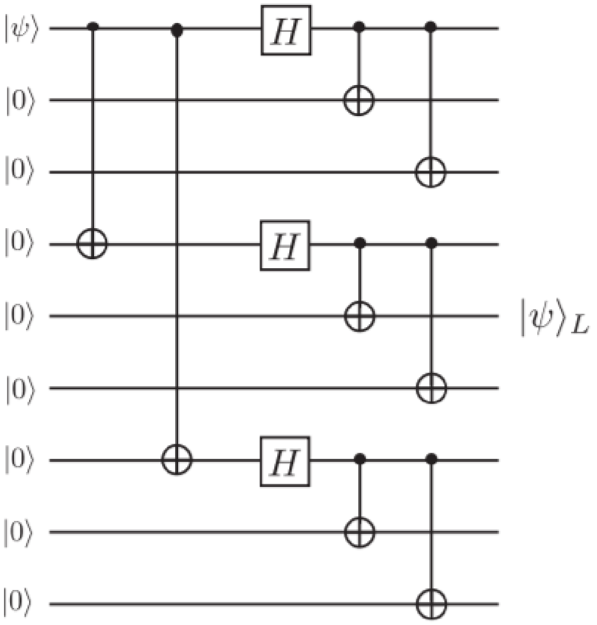


Figure 12: Circuit for encoding a single qubit to use with Shor's 9 qubit code C_9. Image is from the excellent review paper [33] by Devitt, Munro, Nemoto.

So then C_9 is defined by

$$C_9\,|0\rangle = \frac{1}{2\sqrt{2}}\big(\,|000\rangle + |111\rangle\,\big) \otimes \big(\,|000\rangle + |111\rangle\,\big) \\ \otimes \big(\,|000\rangle + |111\rangle\,\big)$$

$$C_9\,|1\rangle = \frac{1}{2\sqrt{2}}\big(\,|000\rangle - |111\rangle\,\big) \otimes \big(\,|000\rangle - |111\rangle\,\big) \\ \otimes \big(\,|000\rangle - |111\rangle\,\big)$$

these above are referred to as the basis states of the code [33]. The full Shor 9-qubit code is described as follows

1) First we apply the error-correction code that corrects 1-qubit bit flips (10) to each of the three 3-qubit blocks $\{1,2,3\}, \{4,5,6\}, \{7,8,9\}$.
2) Next we want to apply the error-correction code that corrects one-qubit phase flips to the 3 qubits we entangled in the first stage: $\{1,4,7\}$. However, we can't do this directly because each of these have been encoded into a 3-qubit state by (12). The following circuit detects in which of these three 3-qubit blocks a phase flip error has occurred

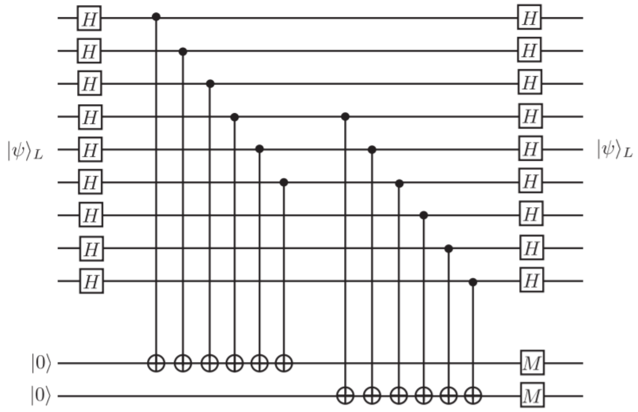


Figure 13: Circuit to perform the Z-error correction for the 9-qubit code. Image from [33].

The output of the above circuit gives a syndrome that specifies which of the three blocks had a phase-flip error. After than we can apply a Z-gate to any one of three qubits in that block.

Theorem: Shor 9 Code The above error-correcting procedure correct any one-qubit error. In other words if we apply this code to

$$U_i C_9 |\phi\rangle$$

for some one-qubit gate U with $1 \leq i \leq 9$, then we get

$$C_9 |\phi\rangle \otimes |\psi\rangle$$

where ψ is a state of the ancilla qubits. The proof of this is evident if we consider the gate-set (universal) $U \in \{I, X, Z, XZ\}$, so since the above circuit corrects errors induced by all these gates, and these gates form a universal gate set, then the above circuit is able to correct any combination of them (happening to one qubit). If we don't want the full 9-code as output from the circuit we can achieve the discussed 1-qubit error correction by a smaller circuit

References

[1] Quantum Computation and Quantum Information, Isaac Chuang and Michael Nielsen

[2] *Qiskit textbook from IBM*, https://qiskit.org/textbook/preface.html, Abraham Asfaw, Luciano Bello, Yael Ben-Haim, Mehdi Bozzo-Rey, Sergey Bravyi, Nicholas Bronn, Lauren Capelluto, Almudena Carrera Vazquez, Jack Ceroni, Richard Chen, Albert Frisch, Jay Gambetta, Shelly Garion, Leron Gil, Salvador De La Puente Gonzalez, Francis Harkins, Takashi Imamichi, Hwajung Kang, Amir h. Karamlou, Robert Loredo, David McKay, Antonio Mezzacapo, Zlatko Minev, Ramis Movassagh, Giacomo Nannicni, Paul Nation, Anna Phan, Marco Pistoia, Arthur Rattew, Joachim Schaefer, Javad Shabani, John Smolin, John Stenger, Kristan Temme, Madeleine Tod, Stephen Wood, James Wootton.

[3] Postulates of Quantum Mechanics, http://vergil.chemistry.gatech.edu/notes/quantrev/node20.html

[4] The Character of Physical Law, R. Feynman 1964, Cornell University, https://youtu.be/kEx-gRfuhhk?t=15047

[5] Qiaochu Yuan, Math StackExchange 13 Aug 2011, https://math.stackexchange.com/questions/57148/matrices-which-are-both-unitary-and-hermitian#:~:text=Unitary%20matrices%20are%20precisely%20the,are%20on%20the%20unit%20circle.&text=So%20unitary%20Hermitian%20matrices%20are,corresponding%20eigenvalues%20are%20%C2%B11.

[6] Unitary and Hermitian matrices http://www.bumatematikozelders.com/altsayfa/matrix_theory/unitary_and_hermitian_matrices.pdf

[7] Linear Algebra and its Applications, Fifth Edition, David C. Lay, Steven R. Lay, Judi J. McDonald, Pearson Publishing.

[8] Arctan2 definition https://en.wikipedia.org/wiki/Atan2

[9] Visualizing multiple entangled qubits https://algassert.com/post/1716

[10] Weisstein, Eric W. "Discrete Fourier Transform." From MathWorld–A Wolfram Web Resource. https://mathworld.wolfram.com/DiscreteFourierTransform.html

[11] Lecture 5 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture05.pdf

[12] Lecture 6 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture06.pdf

[13] Lecture 7 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture07.pdf

[14] Lecture 8 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture08.pdf

[15] Lecture 9 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture09.pdf

[16] Lecture 11 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture11.pdf

[17] Lecture 18 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture18.pdf

[18] Lecture 21 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture21.pdf

[19] Lecture 22 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture22.pdf

[20] Lecture 26 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture26.pdf

[21] Lecture 27 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture27.pdf

[22] Lecture 28 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture28.pdf

[23] Lecture 29 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture29.pdf

[24] Lecture 30 https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/lecture30.pdf

[25] https://arxiv.org/pdf/quant-ph/0205095.pdf

[26] Wikipedia article on convergents of continued fractions https://en.wikipedia.org/wiki/Continued_fraction#Infinite_continued_fractions_and_convergents

[27] Wikipedia article on least common multiple https://en.wikipedia.org/wiki/Least_common_multiple

[28] Wikipedia article on Euclid's algorithm for fast computation of greatest common divisor. https://en.wikipedia.org/wiki/Euclidean_algorithm

[29] Wikipedia article on reversible computing https://en.wikipedia.org/wiki/Reversible_computing#Logical_reversibility

[30] Article on quantum garbage removal, Qiskit Textbook https://qiskit.org/textbook/ch-gates/oracles.html

[31] Article on Grover search, Qiskit textbook https://qiskit.org/textbook/ch-algorithms/grover.html

[32] Article on invariant subspaces https://www.statlect.com/matrix-algebra/invariant-subspace

[33] Quantum Error Correction for Beginners, S. Devitt, W. Munro, K. Nemoto, https://arxiv.org/pdf/0905.2794.pdf

[34] Linear Maps on Hilbert Spaces, springer https://link.springer.com/content/pdf/10.1007%2F978-3-030-33143-6_10.pdf

[35] Quantum Gates Chapter https://www.asc.ohio-state.edu/perry.6/p5501_sp17/articles/quantum_gates.pdf

A MAT3420 Exam Preparation

A.1 Exam 2020

A.1.1 Qubit states 1)

Which of the following are possible qubit states?

$$\frac{1}{2}|0\rangle + \frac{1}{2}|1\rangle$$
$$\frac{3}{5}|0\rangle + \frac{4}{5}|1\rangle$$
$$\frac{\sqrt{3}}{2}|0\rangle + \frac{1}{2}|1\rangle$$

Answer: It is widely accepted that all quantum states should have the property that the square of the amplitude should add up to one, in other words, it makes sense that if you add up all the possibilities you should be left with 100% likelihood that whatever you are looking at will be in one of the possible states [3]. Thus, for the three examples the first is an impossible state in quantum mechanics while the two remaining ones are possible. $(\frac{1}{2})^2 + (\frac{1}{2})^2 \neq 1$, $(\frac{3}{5})^2 + (\frac{4}{5})^2 = 1$, $(\frac{\sqrt{3}}{2})^2 + (\frac{1}{2})^2$.

The word amplitude, or probability amplitude, stems from the famous result of the double slit experiment. Richard Feynman has an excellent description of the weirdness and concept in his lecture "The Character of Physical Law" [4] (highly recommended watch). The reason why we still use the word amplitude is that one part of the behavior of 'particles' fits with the classical formalism for the motion of waves through a double slit. A subject where the concept of an amplitude is well understood.

A.1.2 Unitary Operators 2)

Suppose we have m input/output qubits and n ancilla qubits. Consider a quantum circuit consisting of one unitary gate $\mathbf{U}$ operating on n ancilla qubits. Prove that we won't see any effect of $\mathbf{U}$.

Answer: A *unitary operator* is an operator which preserves the vector norm when it operates on a vector. In other words, it makes sure that the amplitude previously discussed stays at 1. Another way to say this is: *Unitary matrices are precisely the matrices admitting a complete set of orthonormal eigenvectors such that the corresponding eigenvalues are on the unit circle* [5].

A *pure state* of the entire system is represented by a unit vector of the following form

$$\sum_{x=0}^{2^m-1} |x\rangle \otimes v_x$$

where v_x are vectors in the state space of the ancilla qubits, with $\sum_x ||v_x||^2 = 1$. The specified circuit transforms this into

$$\sum_{x=0}^{2^m-1} |x\rangle \otimes \mathbf{U}v_x$$

The probability of the outcome x is therefore given by $||\mathbf{U}v_x||^2 = ||v_x||^2 = 1$ which is independent of $\mathbf{U}$.

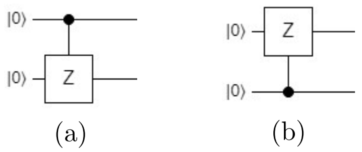
A.1.3 Modular Inverse, Shor's Algorithm 3)

One of the classical subroutines in Shor's algorithm computes the *modular inverse* of a number. Explain this classical algorithm and consider the following example:

Find the modular inverse of 16mod21, that is, find a number n such that $16n = 1\text{mod}21$

A.1.4 Circuit Equivalence 4)

Prove the following equality of quantum circuits

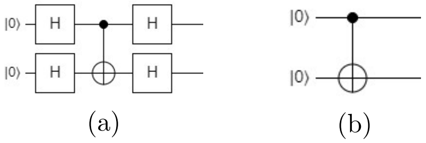


(a) (b)

Figure 15: Caption

A.1.5 Circuit Equivalence 5)

Prove the following equality of quantum circuits



(a) (b)

Figure 16: Caption

A.1.6 6)

Assume we have two quantum circuits $\mathbf{U}$ and $\mathbf{U}$' with m input/output qubits and n ancilla qubits, both computing a function f, but $\mathbf{U}$ does it without leaving 'garbage' in the ancilla qubits, while $\mathbf{U}$' possibly not. In other words

$$\mathbf{U}\big(\,|x\rangle \otimes |0\rangle\,\big) = |f(x)\rangle \otimes |0\rangle$$
$$\mathbf{U}'\big(\,|x\rangle \otimes |0\rangle\,\big) = |f(x)\rangle \otimes |g(x)\rangle$$

What are the necessary conditions on $g(x)$ for not seeing any difference between $\mathbf{U}$ and $\mathbf{U}$' for any *mixed* input state?

A.1.7 Schmidt Decomposition and Number. 7)

Consider two quantum systems **A**, **B** with finite dimensional state spaces $\mathbf{H_A}, \mathbf{H_B}$

Let $\eta \in \{\mathbf{H_A} \otimes \mathbf{H_B}\}$ be a pure state, *unit vector*, of the composite system. It can be shown that there exists an orthonormal system of vectors $e_1, ..., e_n \in \mathbf{H_A}$ and $f_1, ..., f_n \in \mathbf{H_B}$ and numbers $\lambda_k > 0$ such that

$$\eta = \sum_{k=1}^{n} \lambda_k (e_k \otimes f_k)$$

This is called a *Schmidt decomposition.*

Can you find n without knowing the decomposition? Show that n depends only on η. It is called the *Schmidt number* of η and can be considered a measure of entanglement of η.

`http://www.fmt.if.usp.br/~gtlandi/04---reduced-dm-2.pdf`

A.2 Hand-out Exercises

1. `https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/exercise1.pdf`
2. `https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/exercise2.pdf`
3. `https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/exercise3.pdf`
4. `https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/exercise4.pdf`
5. `https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/exercise5.pdf`
6. `https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/exercise6.pdf`
7. `https://www.uio.no/studier/emner/matnat/math/MAT3420/v21/filer/exercise7.pdf`

A.2.1 Exercise 1

In the following exercises $\mathcal{H}$ denotes a finite dimensional Hilbert space.

1) Show that if $|\psi\rangle$ and $|\phi\rangle$ are eigenvectors of a normal operator T on $\mathcal{H}$, so that $T|\psi\rangle = \mu|\psi\rangle$ and $T|\phi\rangle = \lambda|\phi\rangle$, and $\lambda \neq \mu$, then $\langle\phi|\psi\rangle = 0$
To show this we write

$$\begin{aligned}(\lambda - \mu)\langle\phi|\psi\rangle &= \langle\lambda\phi|\psi\rangle - \langle\phi|\overline{\mu}\psi\rangle \\ &= \langle T\phi|\psi\rangle - \langle\phi|T^*\psi\rangle \\ &= \langle T\phi|\psi\rangle - \langle\phi|\psi\rangle = 0\end{aligned}$$

Because by definition $\lambda \neq \mu$ then $\langle\phi|\psi\rangle$ must be zero.

2) Assume T is a normal operator on $\mathcal{H}$ with spectrum $\delta(T) = \{\lambda_1, \cdots, \lambda_k\}$. For a fixed index i, consider the function

$$\chi_i(\lambda_j) = \begin{cases} 1 \text{ if j = i} \\ 0 \text{ if j} \neq i \end{cases}$$

Show that $X_i(T)$ is the orthogonal projection

$$\mathcal{H} \to \mathcal{H}_i = \{|\psi\rangle : T|\psi\rangle = \lambda_i|\psi\rangle\}$$

To show this we start with

$$\begin{aligned}\chi_i(T)|\psi\rangle &= \chi_i(\lambda_1 P_1 + \cdots + \lambda_k P_k)|\psi\rangle \\ &= \sum_j \chi_i(\lambda_j) P_j |\psi\rangle \\ &= 1P_{j=i}|\psi\rangle = \lambda_i|\psi\rangle\end{aligned}$$

which is the orthogonal projection we were asked for.

3) Show that if T is a normal operator on $\mathcal{H}$ and S is a linear operator commuting with T, so that $ST = TS$, then we have $Sf(T) = f(T)S$ for every function $f : \delta(T) \to \mathbb{C}$
To show this we start with $f(T) = f(\lambda_1)P_1 + \cdots + f(\lambda_k)P_k$

$$\begin{aligned}Sf(t) = S\big(f(\lambda_1)P_1 + \cdots + f(\lambda_k)P_k\big) \\ = \big(f(\lambda_1)P_1 + \cdots + f(\lambda_k)P_k\big)S = f(T)S \\ Sf(\lambda_1)P_1 + \cdots + Sf(\lambda_k)P_k \\ = f(\lambda_1)P_1S + \cdots + f(\lambda_k)P_kS\end{aligned}$$

Not sure if it's correct to say here that we can say order doesn't matter here because S and T commute and therefore S and the projector commute. The are both linear operators, so the order shouldn't matter and the equality holds. (but not all linear transformations commute?)

4) We know that every normal operator T on $\mathcal{H}$ is diagonalizable in an orthonormal basis. Prove the following extension of this result: *If we have a (possibly infinite) family of pairwise commuting normal operators $T_j(j \in J)$ on $\mathcal{H}$, then they are simultaneously diagonalizable in an orthonormal basis.*
We assume there is a non-scalar operator T_{j0} and we let λ be an eigenvalue of T_{j0}. Now we consider the space

$$\mathcal{H}' = \{|\phi\rangle \in \mathcal{H} : T_{j0}|\phi\rangle = \lambda|\phi\rangle\}$$

Then $\mathcal{H}' \neq \mathcal{H}$ and $\mathcal{H}' \neq \mathcal{H}$, so we can conclude that $1 \leq \dim(\mathcal{H}') \leq n - 1$. We have $T_j\mathcal{H}' \subset \mathcal{H}'$ and $T_j^*\mathcal{H}' \subset \mathcal{H}'$. If we take $|\psi\rangle \in \mathcal{H}'$ then we have

$$T_{j0}T_j|\psi\rangle = T_jT_{j0}|\psi\rangle = \lambda T_j|\psi\rangle$$

ergo T_j is in the space $\mathcal{H}'$. Because $T_j \left|\psi\right\rangle$ is an eigenvector for T_{j0}, which is how we defined $\mathcal{H}'$. From normality we can conclude that this is also the case for $T_j^* \left|\psi\right\rangle$.

Now we consider the whole family

$$\left(T_j|_{\mathcal{H}'}\right)_{j\in J}$$

of normal operators on $\mathcal{H}'$. These operators pairwise commute. By the induction hypothesis we conclude that $\mathcal{H}'$ has a non-zero vector $\left|\psi_1\right\rangle$ that is an eigenvector for T_j for all values of j. Thus we can use the same procedure as in the proof for orthonormal basis for one such operator using induction on the dimension of $\mathcal{H}$ like (4.3.1) to find an orthonormal basis consisting of eigenvectors of T_j for all j.

A.2.2 Exercise 2

1) a) Show that if $ST = TS$ and they're normal on a finite dimensional Hilbert space then

$$e^{S+T} = e^S e^T$$

We have the definition of e^A for any A

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

Writing both sides out and expanding

$$\sum \frac{(S+T)^n}{n!} = c_0 S^n T^0 + \cdots + c_n S^0 T^n$$

where i have absorbed the reciprocal factorial into the coefficients. Now we look at the right hand side of the equation expanded we have

$$\begin{aligned}\sum \frac{S^n}{n!} \sum \frac{T^n}{n!} = c_0' S^0 T^0 + \cdots + c_k' S^0 T^n \\ + c_{k+1}' S^1 T^0 + \cdots + \cdots + c_n' S^n T^n\end{aligned}$$

Cant seem to complete this argument fully...

Doing it using same basis. Since S and T are normal commuting operators we have that they are both diagonalizable in the same orthonormal basis: $\left|\phi_1\right\rangle, \cdots, \left|\phi_n\right\rangle$ such that

$$\begin{aligned}S\left|\phi_i\right\rangle = a_i \left|\phi_i\right\rangle, T\left|\phi_i\right\rangle &= b_i \left|\phi_i\right\rangle \\ e^S e^T \left|\phi_i\right\rangle = e^S e^{b_i} \left|\phi_i\right\rangle &= e^{a_i} e^{b_i} \left|\phi_i\right\rangle \\ &= e^{a_i+b_i} \left|\phi_i\right\rangle = e^{S+T} \left|\phi_i\right\rangle\end{aligned}$$

Hence $e^S e^T = e^{S+T}$.

b) Show the converse if $S = S^*$ and $T = T^*$ with $e^{S+T} = e^S e^T$ then $ST = TS$.

$$e^S e^T = (e^S e^T)^* \underset{(AB)^*=B^*A^*}{=} (e^T)^* (e^S)^* = e^T e^S$$

Hence there is an orthonormal basis in which e^S and e^T are diagonal. Thus, we can take $S = log(e^S), T = log(e^T)$, and they are diagonal in the same basis. It follows that $ST = TS$.

2) Given $T \in L(\mathcal{H})$, the operator $(TT)^{1/2}$ is denoted by $|T|$. Explain why $|T|$ is a well-defined positive operator. Show next that if T is invertible, then $|T|$ is invertible as well and $U = T|T|^1$ is unitary. *The decomposition $T = U|T|$ is called the polar decomposition. It is a close relative of the singular value decomposition discussed in the linear algebra course.*

Every positive operator is self-adjoint is a fact. We check if this is the case for $|T|$

$$|T| = (T^*T)^{\frac{1}{2}} = \left((T^*T)^{\frac{1}{2}}\right)^* = \left(T^*T^{*^*}\right)^{\frac{1}{2}} = (T^*T)^{\frac{1}{2}}$$

We conclude that since $|T|$ is self-adjoint then it must be positive. If T is invertible, then must also be invertible T^*. Then we can conclude that since $|T|$ is just a product of invertible matrices, that it itself must also be invertible.
To show that $T|T|^{-1}$ is unitary we check using the definition of unitarity $U^*U = UU^* = I$)

$$\begin{aligned}\left(T(T^*T)^{\frac{1}{2}}\right)^{-1} &= \left(T(T^*T)^{\frac{1}{2}}\right)^* \\ T^{-1}(T^*T)^{\frac{1}{2}} &= \left((T^*T)^{-\frac{1}{2}}\right)^* T^* \\ T^{-1}T^{*^{\frac{1}{2}}} T^{\frac{1}{2}} &= \left(T^{*^{\frac{1}{2}}} T^{-\frac{1}{2}}\right)^* T^* \\ T^{*^{\frac{1}{2}}} T^{-\frac{1}{2}} &= T^{*^{-\frac{1}{2}}} T^{*^{*^{\frac{1}{2}}}} T^* \\ &= T^{*^{-\frac{1}{2}} T - \frac{1}{2}} T^* \\ &= T^{*^{\frac{1}{2}}} T^{-\frac{1}{2}}\end{aligned}$$

and we have shown unitarity.