# Ew Applied Maths

January 27, 2020

## 0 Introduction

### What is information?

We get information when we acquire knowledge. In classical information, the fundamental unit of information is a bit - a Boolean unit, which takes values of either 0 or 1. If we have a question with $2^n$ possible different, answers, then we can label each answer with a string of $n$ bits.

### What is computation?

Computation can be defined as the processing of information - updating a bit string using a prescribed sequence of steps, which we might call a program. Each of these steps can be the actions of a Boolean gate, such as Not, And, Or.

### What is a bit physically?

Physically a bit is given by any 2 *different* physical states of a system that can be reliably distinguished by some physical measurement. For instance, I can write 0 or 1 in this document, which is stored as data on GitHub or your hard drive, and you can conduct the physical measurement of loading this document and reading the following number: 1. This concept is summed up in the saying "There is no information without representation". We see that any computation is then the manipulation of a physical system, and so computers must obey the laws of physics. Modern computers obey classical physics (at least for now - please email me / hololens call me if this is not the case for you), but as Feynmann once said: "Nature isn't classical dammit".

On the atomic and subatomic (and even now molecular) scales, particles are governed by Quantum Mechanics. This has some novel features over classical mechanics, such as the ideas of quantum superposition, entanglement, and quantum measurement. Throughout this course, we will be exploiting these features. They will make things more complicated, but at the same time will give us some advantages in information storage, communication, computation, as well as security/cryptography.

Most of the work for this course was done during the 1980's and 1990's, although the Russian mathematician Alexander Holevo was already working on the ideas of quantum information in the 1970's.

## In what ways is QIC better than its classical counterpart?

A quantum computer cannot compute any computational task which isn't already computable *in principle* by a classical computer. However, some things will require many times the age of the universe to compute even on the fastest of classical supercomputers.

### 0.1 Computational Complexity Theory

We want to have some idea of the "difficulty" of a computational task. We measure this by the amount required of the two computational resources of time and space. For instance, consider the problem of factoring a number with $n$ digits. Then the input size for the computation is $n$, and we can look at how the time/space grow as functions of $n$.

If the time grows polynomially in $n$, we call this a poly-time algorithm. These are computable in practice. If it grows faster than any polynomial, we call it an exponential-time algorithm, which are, whilst theoretically computable, uncomputable in practice for large inputs.

Currently, there is no known poly-time factoring algorithm for a classical computer, but in 1994 Peter Shor came up with a poly-time quantum algorithm for integer factorisation.

### 0.2 Communication & Security Issues

If we use quantum particles as information carriers, we end up with communication possibilities which were the stuff of science fiction previously, such as quantum teleportation, or implement mathematically provably secure communication.

### 0.3 Technological Issue

Moore's law states that, since 1965, devices will miniaturise by a factor of 4 every 3.5 years. Eventually (and this is already happening), we will get close to quantum scales. For instance, modern CPUs are built by manufacturing machines with a resolution of around 10nm, or 100 atoms across. At these scales, quantum mechanics will start to cause problems for classical computation, so why not embrace this new feature? In reality, the technological barrier to actually using quantum bits (qubits) is huge, and the most advanced current quantum computers have about 53 qubits. Moreover, Google has claimed ***quantum supremacy*** in demonstrating a task, known as the "random sampling algorithm", which their quantum computer can complete faster than a classical computer.

### 0.4 Quantum Mechanics

We have 4 key postulates of quantum mechanics:

(QM1) Describes the quantum state of a physical system, $S$

(QM2) Describes the joint state of two composite systems, $S_1S_2$

(QM3) Describes how quantum states evolve throughout time. Here we will be interested in discrete time steps.

(QM4) Describes how quantum states respond to measurement.

The idea behind the mathematics of quantum mechanics, which we'll get to properly in the next section, is that we want to associate to every quantum-mechanical system $S$ a ***complex inner product vector space*** $\mathcal{V}$, or ***Hilbert space***. We will be using ***Dirac's bra-ket notation***.

# 1 Mathematical Preliminaries & Bra-Ket Notation

We will be working in an $n$-dimensional complex inner product spaces $\mathcal{V}$. We will denote a column vector by a ***ket***, denoted $|\psi\rangle$. It's conjugate transpose is called a ***bra***, and written $\langle\psi|$. For instance, if $|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix}$, then $\langle\psi| = (\bar{a}, \bar{b})$. We then denote the ***inner product*** $(|\phi\rangle, |\psi\rangle) = \langle\phi|\psi\rangle = (\bar{a}, \bar{b}) \begin{pmatrix} c \\ d \end{pmatrix} = \bar{a}c + \bar{b}d \in \mathbb{C}$.

Recall the axioms for an inner product:

IP1 Positive definite. $\mathbb{R} \ni \langle\psi|\psi\rangle \geq 0$, and $= 0$ if and only if $|\psi\rangle = \mathbf{0}$.

IP2 Linearity in second argument. $\langle\phi|\lambda\psi\rangle = \lambda\langle\phi|\psi\rangle$

IP3 Antilinearity in first argument. $\langle\lambda\phi|\psi\rangle = \bar{\lambda}\langle\phi|\psi\rangle$

IP4 Skew-symmetry. $\overline{\langle\phi|\psi\rangle} = \langle\psi|\phi\rangle$

IP5 Norm. We define $||\psi|| = \sqrt{\langle\psi|\psi\rangle}$

For this course we will take as a convention that $\mathcal{V} = \mathbb{C}^n$, with basis $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, and so on for higher dimensions.

We also have an ***outer product***, which is simply the other way round, so that we have $|\psi\rangle\langle\phi|$ is an $n \times n$ matrix.

Note that this basis is a ***complete orthonormal basis*** for $\mathbb{C}^n$, i.e. we have $\langle i|j\rangle = \delta_{ij}$.

By complete, we mean that $\sum_{i=1}^n |i\rangle\langle j| = I_{n\times n}$. As an exercise, check that $\mathcal{V} = \mathbb{C}^2$ with basis $|0\rangle, |1\rangle$ is complete.

We will also be using the orthonormal basis $\{|+\rangle, |-\rangle\}$, with $|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}$.

Suppose we have spaces $\mathcal{V}, \mathcal{W}$, of dimensions $n, m$, with bases $\{|e_i\rangle\}_{i=1}^m; \{|f_i\rangle\}_{i=1}^n$. Then we define the ***tensor product space*** $\mathcal{V}\otimes\mathcal{W}$ to be the $mn$ dimensional vector space with orthonormal basis $\{|e_i\rangle \otimes |f_j\rangle\}_{i,j}$, where we have:

$$|v\rangle \otimes |w\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a \begin{pmatrix} c \\ d \end{pmatrix} \\ b \begin{pmatrix} c \\ d \end{pmatrix} \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix} \in \mathbb{C}^2 \otimes \mathbb{C}^2.$$

If a vector $|\psi\rangle \in \mathcal{V} \otimes \mathcal{W}$ can be written as $|v\rangle \otimes |w\rangle$, we call it a ***product state***. There are some vectors which cannot be expressed in this way. These are called ***entangled vectors***. Note that we will often omit the $\otimes$, and just write $|\psi\rangle = |v\rangle|w\rangle$.

We can also take $k$-fold ***tensor powers*** of $\mathcal{V}$:

$$\mathcal{V}^{\otimes k} := \mathcal{V} \otimes \mathcal{V} \otimes \ldots \otimes \mathcal{V}$$

$\mathcal{V}^{\otimes k}$ is an $n^k$ dimensional vector space. It has a basis $\{|i_1\rangle|i_2\rangle \ldots |i_k\rangle = |i_1 i_2 \ldots i_k\rangle\}$, where the $|i_j\rangle$ are basis vectors of $\mathcal{V}$. In the case where $\mathcal{V} = \mathbb{C}^2$, we can think of these as bitstrings.

For an example of an entangled vector, consider $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \in \mathbb{C}^2 \otimes \mathbb{C}^2$. Assume there is $|v\rangle = a|0\rangle + b|1\rangle; |w\rangle = c|0\rangle + d|1\rangle$. Then $|v\rangle \otimes |w\rangle = ac|00\rangle + ad|01\rangle + bc|10\rangle + bd|11\rangle$. But there is no choice of $a, b$ that makes this equal to $|\Phi\rangle$.

If $|\psi_1\rangle = |\alpha_1\rangle|\beta_1\rangle, |\psi_2\rangle = |\alpha_2\rangle|\beta_2\rangle$ are product vectors in $\mathcal{V}\otimes\mathcal{W}$. Then $\langle\psi_1|\psi_2\rangle = \langle\alpha_1|\alpha_2\rangle\langle\beta_1|\beta_2\rangle$. In the case where one of the vectors is not a product vector, we have to write it as a sum of the basis vectors $|e_i\rangle|f_j\rangle$ of $\mathcal{V} \otimes \mathcal{W}$, and then expand that product out.

For example, if $|A\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}|e_i\rangle|f_j\rangle$, and $|B\rangle = \sum\sum b_{ij}|e_i\rangle|f_j\rangle$, then:

$$\begin{aligned}\langle A|B\rangle &= \sum_{i,i'}\sum_{j,j'} a_{ij}^* b_{i'j'}\langle e_i|e_i'\rangle\langle f_j|f_j'\rangle \\ &= \sum_{i,j} a_{ij}^* b_{ij}\end{aligned}$$

# 2 Postulates of Quantum Mechanics

QM1 To any isolated/closed quantum system, one can associate a Hilbert space $\mathcal{V}$, called the ***state space***. The physical states are given by ***unit vectors*** in $\mathcal{V}$. (More precisely, they are given by ***rays***: suppose we have $|\psi\rangle \in \mathcal{V}, |\phi\rangle = e^{\mathrm{i}\theta}|\psi\rangle$ for $\theta \in \mathbb{R}$, where we call $e^{\mathrm{i}\theta}$ a ***global phase factor***. We will see that there is no measurement that is able to distinguish $|\phi\rangle, |\psi\rangle$, so to all intents and purposes they are identical - they refer to the same state. Hence we only distinguish state vectors up to the equivalence relation of differing by a global phase factor, and call the equivalence classes the rays). We will be focusing on the specific case where $\mathcal{V} = \mathbb{C}^2$, with the states $a|0\rangle + b|1\rangle$ for $a, b \in \mathbb{C}$, called ***qubits***.

QM2 Given a composite of two systems $S_1, S_2$ with vector spaces $\mathcal{V}_1, \mathcal{V}_2$, the state space of the composite system $S_1S_2$ is given by $\mathcal{V}_1 \otimes \mathcal{V}_2$. Again, focusing on the qubit case, if we have $n$ cubits then the composite space of $S_1S_2 \dots S_n$ is $\mathcal{V} = (\mathbb{C}^2)^{\otimes n}$, with basis given by the bitstrings discussed above. Note that the number of coefficients required to define a member of $\mathcal{V}$ grows as $2^n$, the number of basis vectors, but the number of coefficients required to define a product state grows only linearly, as such a state can be written as $(a_1|0\rangle + b_1|1\rangle) \otimes \dots \otimes (a_n|0\rangle + b_n|1\rangle)$.

## 2.1 Observables

An ***observable*** $A$ is a linear hermitian (or self-adjoint) operator that acts on the state space $\mathcal{V}$ of $S$. Recall that linearity means that $A(a|v\rangle + b|w\rangle) = aA|v\rangle + bA|w\rangle$, and in this case we can represent $A$ by some matrix with respect to a basis of $\mathcal{V}$. We have for each $A$ a *unique* linear operator $A^\dagger$, the ***adjoint*** of $A$, with the property that $\langle v|Au\rangle = \langle A^\dagger v|u\rangle$. From 1B linear algebra, we know $A^\dagger$ has matrix given by the conjugate transpose of the matrix representing $A$ (when taken with the same basis of course). Hence to say that $A$ is ***self-adjoint*** is to say that $A = A^\dagger$.

An ***eigenvector*** of $A$ is $|phi\rangle$ with the property that $A|phi\rangle = \lambda|\phi\rangle$ for $\lambda$ in the base field of $\mathcal{V}$. We call $\lambda$ an ***eigenvalue*** of $A$. Some important examples of these linear maps will be the

Pauli operators, given by:

$$I = \sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$
$$\sigma_y = \begin{pmatrix} 0 & -\mathfrak{i} \\ \mathfrak{i} & 0 \end{pmatrix}$$
$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Note that $\sigma_i^2 = I$, and $\sigma_x \sigma_y = \mathfrak{i}\sigma_z$.

Moreover, $\sigma_x|0\rangle = |1\rangle, \sigma_x|1\rangle = |0\rangle$. We call $\sigma_x$ a ***bit flip*** - it flips the value of a bit. The eigenvectors of $\sigma_x$ are $|+\rangle, |-\rangle$.

$\sigma_z|0\rangle = |0\rangle, \sigma_z|1\rangle = -|1\rangle$, so we call $\sigma_z$ a ***phase flip***.

$\sigma_y|0\rangle = \mathfrak{i}|1\rangle, \sigma_y|1\rangle = -\mathfrak{i}|0\rangle$, and we call $\sigma_y$ a ***combined flip***.

| | $\|0\rangle$ | $\|1\rangle$ |
|---|---|---|
| $\sigma_x$ | $\|1\rangle$ | $\|0\rangle$ |
| $\sigma_y$ | $\mathfrak{i}\|1\rangle$ | $-\mathfrak{i}\|0\rangle$ |
| $\sigma_z$ | $\|0\rangle$ | $-\|1\rangle$ |

## 2.2 Dirac Notation For Linear Operators/Maps

A simple example of a linear operation on $\mathcal{V}$: let $M$ be the matrix given by $|v\rangle\langle w|$ for $|v\rangle, |w\rangle \in \mathbb{C}^2$. Then $M|x\rangle = |v\rangle\langle w||v\rangle = \langle w|x\rangle|v\rangle$, so this is a rank 1 linear operation with image span $|v\rangle$. The kernel of $M$, $\ker M = \{|a\rangle : |a\rangle \perp |\omega\rangle\}$, since if $|a\rangle$ is perpendicular to $|w\rangle$ then $\langle w|a\rangle = 0$.

In general, we can represent any linear operation $A : \mathbb{C}^2 \to \mathbb{C}^2$ by a $2 \times 2$ matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a|0\rangle\langle 0| + b|0\rangle\langle 1| + c|1\rangle\langle 0| + d|1\rangle\langle 1|$. In general, if $A : \mathcal{V} \to \mathcal{V}$ where $\dim \mathcal{V} = n$, we can write $A = \sum_{i,j}^n a_{ij}|e_i\rangle\langle e_j|$, and so we have $A|\psi\rangle = \sum_{i,j,k} a_{ij}\psi_k|e_i\rangle\langle e_j|e_k\rangle = \sum_{i,j} a_{ij}\psi_j|e_i\rangle$.

We define the ***trace*** of a linear map to be the sum along the leading diagonal of a matrix representing the linear map, i.e. $\operatorname{tr} A = \sum_i \langle e_i|A|e_i\rangle$. Note that $\langle w|v\rangle = \operatorname{tr}|v\rangle\langle w|$.

A very useful class of operators is the ***projection operators***, which are ***idempotent*** (i.e. $\pi^2 = \pi$) hermitian operators. For example, $\Pi_v = |v\rangle\langle v|$, a rank 1 projection operator, satisfies $\Pi_v^2 = \Pi_v, \Pi_v^\dagger = \Pi_v$. This projection map returns only the component of the vector along the vector onto which it is a projection. We can generalise this to higher dimensional subspaces: if $\mathcal{E} \subseteq \mathcal{V}$, where $\mathcal{E} = \operatorname{span}\{e_1, \ldots, e_d\}$, then we can define the projection onto $\mathcal{E} = \sum_{i=1}^d \Pi_{e_i}$.

## 2.3 Tensor Products of Linear Operations

If $a, B$ act on $\mathbb{C}^2$, then given $|i\rangle|j\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$, we can define the tensor product of the linear operations $(A \otimes B)|i\rangle|j\rangle$ to be the $4 \times 4$ linear map $(A \otimes B)_{ij} = A_{\lceil i/2 \rceil \lceil j/2 \rceil} B_{i\%2, j\%2}$, i.e. $a_{ij}B$ for each $i, j$.

QM3 We can describe the time evolution of a ***closed*** quantum system - we can think about how the states evolve by writing them as say $|\psi(t_1)\rangle, |\psi(t_2)\rangle$ for some time points $t_2 > t_1$. The third postulate says that this evolution is given by a ***unitary operator***, i.e. a linear map $U$ for which $UU^\dagger = I = U^\dagger U$, so that $|\psi(t_2)\rangle = U(t_1, t_2)|\psi(t_2)\rangle$.

**Proposition 2.1.** *Let $U : \mathcal{V} \to \mathcal{V}$ be a linear operator. The following are equivalent:*

1. *$U$ is unitary*
2. *$U$ maps complete orthonormal bases to complete orthonormal bases*
3. *$U$ preserves inner products, so $\langle Uv|Uw\rangle = \langle v|w\rangle$*
4. *The columns of $U$ form an orthonormal basis of $\mathcal{V}$*
5. *The rows of $U$ form an orthonormal basis of $\mathcal{V}$*

*Proof.* Done in lectures but was a mess. Check this for yourself. (If you can be bothered). □

In Quantum Mechanics, one measures an ***observable***, say $A$. An important representation of $A$ is as ***spectral projection***: we can write $A = \sum_n a_n P_n$, where $a_n$ are the (real) eigenvalues of $A$, and $P_n$ are the spectral projection operators. In the case that $a_n$ is a ***non-degenerate*** eigenvalue (i.e. has only one $|\phi_n\rangle$ with $A|\phi_n\rangle = a_n|\phi_n\rangle$), then $P_n = |\phi_n\rangle\langle\phi_n|$. If $a_n$ is degenerate, say with multiplicity $m > 1$, then $P_n = \sum_{i=1}^m |\phi_n^{(i)}\rangle\langle\phi_n^{(i)}|$, where the $|phi_n^{(i)}\rangle$ are the eigenvectors with $a_n$ as their eigenvalue.

The outcome of a measurement is one of the eigenvalues, $a_k$. If the ***initial state*** of the system to be measured is $|\phi\rangle \in \mathcal{V}$, then $\mathbb{P}(\text{outcome is } a_k) = p(a_k)$, where $p(a_k) = \langle\phi|P_k|\phi\rangle$. If the outcome is $a_k$, then after the measurement, the state collapses to $\frac{P_k|\phi\rangle}{\sqrt{p(a_k)}}$. This is called the ***Born Rule.*** Note that this measurement is completely specified by the projection operations.

We can also measure ***relative to a basis***: given a basis $\mathcal{B} = \{|e_i\rangle\}_{i=1}^n$, and $|\phi\rangle \in \mathcal{V}$. Then the probability we observe outcome $j \in [n]$ is $\langle\phi|e_j\rangle\langle e_j|\phi\rangle = |a_j|^2$, and the state collapses to $\frac{a_j|e_j\rangle}{\sqrt{|a_j|^2}} = |e_j\rangle$.

For example, if $\mathcal{V} = \mathbb{C}^2$, and $\mathcal{B} = \{|0\rangle, |1\rangle\}$, then we measure $|\phi\rangle = a|0\rangle + b|1\rangle$ relative to $\mathcal{B}$. The outcomes are 0 or 1. $\mathbb{P}(0) = \langle\phi|P_0|\phi\rangle = \langle\phi|0\rangle\langle 0|\phi\rangle = |a|^2$, and if we get 0 then the state collapses to $|0\rangle$.

## 2.4 Incomplete Projective Measurements

In the previous case, where we have a complete orthonormal basis, we decomposed the space up into dimension 1 subspaces, i.e. the spans of the eigenvectors, and then used rank 1 projection operators onto these subspaces. We can instead, more generally, decompose $\mathcal{V}$ into mutually orthogonal subspace of dimension $\geq 1$, so that $\mathcal{V} = \bigoplus_{i=1}^d \mathcal{E}_i$, where $\sum_{i=1}^d \dim \mathcal{E}_i = \dim \mathcal{V} = n$. Then if $\Pi_i$ is the projection operator onto $\mathcal{E}_i$, we have $\Pi_j \Pi_k = \delta_{jk}\Pi_j$.

The incomplete measurement of $|\psi\rangle \in \mathcal{V}$ relative to the orthogonal decomposition $\{\mathcal{E}_1, \ldots, \mathcal{E}_d\}$ returns $i$ with probability $p(i) = \langle\psi|\Pi_i|\psi\rangle$. If the outcome of the measurement is $i$, then the post measurement state is $\frac{\Pi_i|\psi\rangle}{\sqrt{p(i)}}$.

Given an incomplete measurement with decomposition $\{\mathcal{E}_i, \ldots, \mathcal{E}_d\}$ of $\mathcal{V}$, we can choose a basis $\{|e_j^{(i)}\rangle\}_{j=1}^{d_i}$ of each $\mathcal{E}_i$, so that together they form an orthonormal basis of the whole space

$\mathcal{V}$. Then the measurement will return $k$ (corresponding to the subspace $\mathcal{E}_k$) with probability $\sum_{i=1}^{d_k}\left(\langle\psi|P_i^{(k)}|\psi\rangle\right) = \langle\psi|\left(\sum_{i=1}^{d_k} P_i^{(k)}\right)|\psi\rangle = \langle\psi|\Pi_k|\psi\rangle$.

For example, if we want to measure the ***parity*** of a 2-bit string $b_1b_2$ (i.e. $b_1$ XOR $b_2$), we might have some state given by $|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle$. Then the space of 0 parity, $\mathcal{E}_0 = \text{span}\{|00\rangle, |11\rangle\}$, and likewise we have $\mathcal{E}_1 = \text{span}\{|01\rangle, |11\rangle\}$. Then $\mathbb{P}(0) = \langle\psi|\Pi_0|\psi\rangle = \langle\psi|\Big(|00\rangle\langle00| + |11\rangle\langle11|\Big)|\psi\rangle = \langle\psi|00\rangle\langle00|\psi\rangle + \langle\psi|11\rangle\langle11|\psi\rangle = |a|^2 + |d|^2$, and the post measurement state is $\frac{a|00\rangle + d|11\rangle}{\sqrt{|a|^2+|d|^2}}$.

We can also do this on the vector level: $\mathbb{P}(0) = \mathbb{P}(00) + \mathbb{P}(11) = \langle\psi|00\rangle\langle00|\psi\rangle + \langle\psi|11\rangle\langle11|\psi\rangle = |a|^2 + |d|^2$. It is not a coincidence that we get the same answer, but of course since this isn't a pure course we won't formalise this into a nice little lemma or proposition, and instead we'll just assert it's obvious and move on.