Quantum computing and the hidden subgroup problem

Peter Hebden 2274803

March 2024

1 Introduction

This year marks three decades since the first publication of Shor's Algorithm (Shor 1994), and it remains uncontroversially the most celebrated result in the field of quantum computing. It is a story we have all heard before: modern cryptography relies on the intractability of factoring large numbers—in particular semiprimes[1]—and quantum computers are able to subvert this by calculating an endless number of possibilities simultaneously. I would suggest that this popular idea of quantum computing is an awkward one, in that one may argue its truth on a technicality whilst actually painting quite a misrepresentative picture. We will see that in spite of falling short of an infinite gain in computational power, quantum theory provides a fascinating theoretical paradigm of computation—one where the majority of exponential speedups that have been gained over the classical are generalisable to the solution of one particular problem, namely of identifying a "hidden subgroup".

2 A little bit of complexity

Occasionally in this paper we will be dealing with the notion of the *computational complexity* of an algorithm or problem. Many of the problems presented seem trivial to solve—given enough time and paper, anybody could tell you the factors of an arbitrary semiprime, for example. In problems of computation there is a focus given not only to solving a problem, but doing so efficiently with the resources available. Here we will be concerned primarily with the resource of time.

Definition 2.1 (Big O)**.** Let $n \in \mathbf{N}$ repesent the *size* of a problem X, for a reasonable definition of size given the context. Given an algorithm A to solve X, let $T : \mathbf{N} \to \mathbf{R}$ be a function such that $T(n)$ represents the time required by A to solve X. Then for some $g : \mathbf{N} \to \mathbf{R}$, A is said to solve the problem in $O(g(n))$ if there exists $M, n_0 \in \mathbf{N}$ such that

$$n \geq n_0 \implies T(n) \leq Mg(n). \tag{1}$$

As a simple example, let us consider a naïve algorithm to search an arbitrary list of size n for a given element by checking each element in turn. In the worst case the element may not be present in the list, so the number of checks needed in that case would be n. Using the number of checks as a proxy for time, we see that this algorithm is $O(n)$ (noting that this choice of $g(n)$ is not unique, but minimal), and we would say the algorithm's complexity scales linearly with the size of the list.

This formalisation of complexity allows us to compare the efficiency of algorithms. In particular, we are concerned mainly with the distinction between polynomial and non-polynomial time, and an algorithm that is is not $O(n^x)$ for any real x is said to be intractable. Conversely, any algorithm that does run in polynomial time is simply said to be *efficient*.

[1] The product of exactly two prime numbers.

3 Introduction to quantum computing

Quantum computing can be defined as the study of information processing tasks that can be accomplished using quantum mechanical systems (Nielsen and Chuang 2010). Its study was first motivated in the 1970s (ibid.) by the question of finding a model of computation to allow the efficient simulation of any other model of computation, after doubts were cast on the famous Church-Turing thesis, when Solovay and Strassen (1977) were able to efficiently decide primeness by probabilistic means (which a deterministic Turing Machine is unable to simulate).

Richard Feynman (1982) postulated that in order to efficiently simulate quantum mechanical processes, it would be necessary to utilise quantum computers—the implication being that the class of non-polynomial algorithms possible in a quantum setting is different than in a classical setting. In the decades that followed, the field of quantum computing developed, and such speedups were found for certain problems such as factoring. Despite this, to this day it is still unclear whether quantum computing actually allows for a greater set of tractable problems in theory, as nothing has been proven about the general case. We will take a look at some of the foundations of this field before moving on to examining a class of problems where it currently appears to have a massive upper hand.

3.1 Introduction to complex linear algebra

We begin by assuming a general understanding of the linear algebra on real vector spaces, and we will be generalising and extending those ideas to suit the stage upon which quantum mechanics is set—complex vector spaces. We also introduce Dirac notation (or *Bra-Ket* notation), which will be used throughout the paper. Dirac notation is a notational system introduced by Paul Dirac (1939) to ease the writing of the kinds of expressions which come up constantly in quantum mechanics.

It should be understood that whenever we say vector space, we implicitly mean a *finite* vector space.

Definition 3.1. We write $|v\rangle$, pronounced *ket* v, to denote a vector in a complex vector space V.

Definition 3.2. Let V be a complex vector space with inner product $(\cdot, \cdot)$. We write $\langle\phi|$, pronounced *bra* ϕ, to denote a linear form $V \to \mathbf{C}$. Then we write $\langle\phi|v\rangle \in \mathbf{C}$ to denote the action of $\langle\phi|$ on $|v\rangle$.

In particular, for $|u\rangle \in V$ we write $\langle u|$ to denote the **dual** of $|u\rangle$ defined by

$$\langle u|v\rangle = (|u\rangle, |v\rangle). \tag{2}$$

Remark 3.3. When using Dirac notation, it is common to use juxtaposition to denote the composition of linear maps as well as the action of them upon a ket vector. In this light, $\langle u|v\rangle$ may be seen as a shorthand for $\langle u||v\rangle$.

For linear algebra on real vector spaces, an inner product is specified to be a positive-definite bilinear form $\Phi : V \times V \to \mathbf{R}$, where positive-definiteness is defined only when Φ is *symmetric*. For the complex case, in order to maintain positive-definiteness, we generalise the symmetry condition to that of *Hermiticity* and the bilinear form to the *sesquilinear form*, where both are as defined below.

Definition 3.4. A map $\Phi : V \times V \to \mathbf{C}$ is called a **sesquilinear form** if it is linear in the

second argument and conjugate-linear in the first.[2] That is,

$$\Phi\left(|u\rangle, \sum_i a_i|v_i\rangle\right) = \sum_i a_i \Phi(|u\rangle, |v_i\rangle), \tag{3}$$

$$\Phi\left(\sum_i a_i|u_i\rangle, |v\rangle\right) = \sum_i \overline{a_i} \Phi(|u_i\rangle, |v\rangle), \tag{4}$$

where $\overline{z}$ denotes the complex conjugate of z.

Definition 3.5. We say that a sesquilinear form $\Phi : V \times V \to \mathbf{C}$ is **Hermitian** if for all $|u\rangle, |v\rangle \in V$,

$$\Phi(|u\rangle, |v\rangle) = \overline{\Phi(|v\rangle, |u\rangle)}. \tag{5}$$

Notice that a direct consequence of the Hermiticity of Φ is that $\Phi(|v\rangle, |v\rangle)$ is necessarily in $\mathbf{R}$, and so we may use the usual ordering on $\mathbf{R}$ to define positive-definiteness in the usual sense. Then, we may define our complex inner product to be a positive-definite Hermitian form, and most of the notions from real linear algebra—such as those of orthogonality and norms—may be carried over.

Definition 3.6. Let A be a linear operator on a complex vector space V with inner product $(\cdot, \cdot)$. Define the **adjoint** of A, denoted $A^\dagger$, to be the the linear operator such that for all $|u\rangle, |v\rangle \in V$ we have

$$(A|u\rangle, |v\rangle) = \langle u|A^\dagger|v\rangle. \tag{6}$$

Duality theory tells us that there is an isomorphism between the vector space V and the space of linear forms on V (Lang 1968). From this, the existence and uniqueness of such an $A^\dagger$ may be derived.

Definition 3.7. A linear operator U is called **unitary** if U^{-1} exists and is equal to $U^\dagger$. That is, if

$$UU^\dagger = U^\dagger U = I. \tag{7}$$

Corollary 3.8. Let V be a complex vector space with inner product and let $U : V \to V$ be a unitary operator. Then for $|v\rangle \in V$, we have that $\||v\rangle\| = \|U|v\rangle\|$.

Proof.

$$\|U|v\rangle\| = \sqrt{\langle v|U^\dagger U|v\rangle} = \sqrt{\langle v|v\rangle} = \||v\rangle\|. \tag{8}$$

□

Remark 3.9. It should be quite clear from the previous proof that the property of a unitary operator of preserving norms is actually a specific case of a more general property of *preserving inner products*. I have decided to draw special attention to the norm preservation for a reason that should be made clear in the next section.

3.2 Postulates of quantum mechanics

In order to discuss what information processing can be achieved using quantum mechanical systems, it is of course necessary to understand the regime quantum mechanics provides. Foundational to this are the postulates of quantum mechanics, derived experimentally during the early twentieth century. There are many (equivalent) formulations of the postulates, and we will give one that will be familiar to most researchers in quantum computing; namely the one given by Nielsen and Chuang (2010).

[2] In general, the choice of precisely which argument of the sesquilinear form is linear is not important, but it will be convenient for us to assert that the ones we deal with will always have linearity in the second.

Postulate 1 (State space)**.** Any quantum system has associated a complex inner product space, known as its *state space*. The *state vector* is a unit vector (that is, a vector $|\psi\rangle$ such that $\||\psi\rangle\| = 1$) in the state space which describes the system entirely.

Postulate 2 (Evolution)**.** The evolution of a quantum system is unitary. That is, if the state of the same quantum system is described by $|\psi\rangle$ and $|\psi'\rangle$ at times t_1 and t_2 respectively, then there exists a unitary operator U such that

$$|\psi'\rangle = U|\psi\rangle. \tag{9}$$

Postulate 3 (Measurement)**.** The measurement of a quantum system has associated a set of linear operators known as *measurement operators*. Each operator has associated a particular measurement outcome m, and we may denote the operator associated with m by M_m.

Denoting the state vector of the system to be measured by $|\psi\rangle$, the probability $p(m)$ of a given outcome m occurring is given by

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle. \tag{10}$$

Given that outcome m occurred, the state of the system immediately collapses into a new state $|\psi'\rangle$, given by

$$|\psi'\rangle = \frac{M_m|\psi\rangle}{\sqrt{p(m)}}. \tag{11}$$

Naturally, the set of measurement operators must satisfy that the probability of *any* outcome occurring is 1. That is,

$$\sum_m p(m) = \sum_m \langle\psi|M_m^\dagger M_m|\psi\rangle = 1. \tag{12}$$

3.3 Qubits

In classical computation, the smallest unit of information is the *bit*, taking a binary value of 0 or 1. We have just seen from the measurement postulate (Postulate 3) that quantum physics is fundamentally probabilistic, and this is reflected in the bit analogue: the *qubit* (or quantum bit). The state of a qubit can be binary—similarly to a bit—but as it forms a quantum system, we have just seen from Postulate 1 that it may also take a superposition (linear combination) of these two states.

By convention we take $|0\rangle$ and $|1\rangle$ to be an orthonormal basis corresponding with the classical binary values 0 and 1, and as such may write our qubit as the state vector

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{13}$$

for *amplitudes* $\alpha, \beta \in \mathbf{C}$ where $|\alpha|^2 + |\beta|^2 = 1$.

3.4 Measurement in computation

For the rest of the paper we will deal with a specific case of quantum measurement known as *projective measurement*. For this section, the reader should recall the concept of orthogonal projection, in particular that in projecting $|v\rangle \in V$ to a subspace U we have the orthogonal projection of $|v\rangle$, denoted $\text{proj}_U(|v\rangle)$, defined to be the unique vector such that $|v\rangle - \text{proj}_U(|v\rangle)$ is orthogonal to U.

Proposition 3.10. Let V be a complex inner product space over $\mathbf{F}$ and let U be a subspace of V. If $|1\rangle, |2\rangle, \ldots, |\dim V\rangle$ is an orthonormal basis of V such that $|1\rangle, |2\rangle, \ldots, |\dim U\rangle$ is a basis of U, then

$$\mathrm{proj}_U(|v\rangle) = \left(\sum_{i=1}^{\dim U} |i\rangle\langle i|\right)|v\rangle \tag{14}$$

for all $|v\rangle \in V$.

Proof. There exists a_i for $1 \leq i \leq \dim V$ such that $|v\rangle = \sum_{i=1}^{\dim V} a_i|i\rangle$. Then

$$\left(\sum_{i=1}^{\dim U} |i\rangle\langle i|\right)|v\rangle = \left(\sum_{i=1}^{\dim U} |i\rangle\langle i|\right)\left(\sum_{i=1}^{\dim V} a_i|i\rangle\right) = \sum_{\substack{1\leq i\leq \dim U\\ 1\leq j\leq \dim V}} a_j\langle i|j\rangle|i\rangle, \tag{15}$$

but in recalling that $|1\rangle, \ldots, |\dim V\rangle$ is orthonormal, we have

$$\sum_{\substack{1\leq i\leq \dim U\\ 1\leq j\leq \dim V}} a_j\langle i|j\rangle|i\rangle = \sum_{\substack{1\leq i\leq \dim U\\ 1\leq j\leq \dim V}} \delta_{ij}a_j|i\rangle = \sum_{i=1}^{\dim U} a_i|i\rangle. \tag{16}$$

Letting $|u\rangle = \sum_{i=1}^{\dim U} a_i|i\rangle$, it remains to be shown that $|v\rangle - |u\rangle$ is orthogonal to U.

$$|v\rangle - |u\rangle = \sum_{i=1}^{\dim V} a_i|i\rangle - \sum_{i=1}^{\dim U} a_i|i\rangle = \sum_{i=\dim U+1}^{\dim V} a_i|i\rangle, \tag{17}$$

but for all $j \leq \dim U$,

$$\langle j| \sum_{i=\dim U+1}^{\dim V} a_i|i\rangle = \sum_{i=\dim U+1}^{\dim V} a_i\langle j|i\rangle = 0, \tag{18}$$

so we are done. □

The following definition is now justified.

Definition 3.11. The **projector** from V onto a subspace U is defined as the linear transformation $P : V \to U$ given by

$$P := \sum_{i=1}^{\dim U} |i\rangle\langle i|, \tag{19}$$

for any orthonormal basis of U given by $|1\rangle, |2\rangle, \ldots, |\dim U\rangle$, recalling that we can always find such a basis using the Gram-Schmidt orthogonalization process.

A property of any projector P that will be of crucial importance is that $P^2 = P$:

$$P^2 = \left(\sum_i |i\rangle\langle i|\right)\left(\sum_i |i\rangle\langle i|\right) \tag{20}$$

$$= \sum_{ij} |i\rangle\langle i|j\rangle\langle j| \tag{21}$$

$$= \sum_i |i\rangle\langle i|. \tag{22}$$

In particular this means that in U, the projector onto U is the identity. We are almost ready to talk about projective measurement, but not before a couple more definitions from linear algebra that were not included in 3.1.

Definition 3.12. Let A be a linear operator on a complex inner product space. A is said to be **normal** if $AA^\dagger = A^\dagger A$.

Theorem 3.13 (Spectral Decomposition)**.** Let M be a normal linear operator on a complex inner product space V. Then M has a **spectral decomposition**. That is,

$$M = \sum_i \lambda_i |i\rangle\langle i|, \tag{23}$$

where $|i\rangle$ is an orthonormal basis for V, and each $|i\rangle$ an eigenvector of M with eigenvalue λ_i.

Consider a normal linear operator M, with spectral decomposition which we will refer to by the same notation as in the theorem statement (3.13). We may group the terms of the sum by the set of distinct eigenvalues $\{\lambda_m\}$, and then, in denoting by $V(m)$ the set of i such that eigenvector $|i\rangle$ has eigenvalue λ_m, we see

$$M = \sum_m \lambda_m \sum_{i \in V(m)} |i\rangle\langle i|. \tag{24}$$

Notice that each $\sum_{i \in V(m)} |i\rangle\langle i|$ is actually the projector onto the eigenspace of M with eigenvalue λ_m, which we will denote P_m.

We call such an M an *observable*[3], and to take a projective measurement is defined as in Postulate 3 where the measurement operators are given by the $\{P_m\}$ of the observable in question. Notice that we can easily extrapolate (12) holds for $\{P_m\}$ by the following property:

$$\sum_m P_m^\dagger P_m = \sum_m P_m P_m = \sum_m P_m = \sum_m |m\rangle\langle m| = I. \tag{25}$$

The possible outcomes of the measurement correspond to eigenvalues of the observable (Nielsen and Chuang 2010).

Definition 3.14. Given an orthonormal basis $|i\rangle$, we say we **measure in the basis** $|i\rangle$ to mean taking a projective measurement with the implicit observable M as defined by

$$M = \sum_i \lambda_i |i\rangle\langle i|, \tag{26}$$

for an arbitrary choice of distinct eigenvalues λ_i. We say that the outcome of the measurement is i if it corresponds to eigenvalue λ_i.

A useful property of this kind of measurement is that it is easy to quickly see the likelihood of any measurement outcome. In particular, for $|\psi\rangle = \sum_i \alpha_i |i\rangle$, we get the probability of outcome i simply by

$$p(i) = \langle\psi|P_i^\dagger P_i|\psi\rangle = \langle\psi|P_i|\psi\rangle = \langle\psi|i\rangle\langle i|\psi\rangle = \overline{\alpha_i}\alpha_i = |\alpha_i|^2. \tag{27}$$

For example, Consider the state

$$|\psi_1\rangle = \frac{|0\rangle + \sqrt{3}|1\rangle}{2}. \tag{28}$$

The states $|0\rangle$ and $|1\rangle$ have amplitudes $1/2$ and $\sqrt{3}/2$ respectively. As such, when measured in the basis $\{|0\rangle, |1\rangle\}$, the result 1 is three times more likely to occur than 0. Consider now the following, very similar, state

$$|\psi_2\rangle = \frac{|0\rangle + i\sqrt{3}|1\rangle}{2}. \tag{29}$$

Despite having different amplitudes, the probability distribution of outcomes in the basis remains the same. When this occurs—that is, when superposed states differ by a factor of $e^{i\phi}$ for some $\phi \in \mathbf{R}$—we say that they differ in *phase* (ibid.).

[3]Generally, physicists use the term observable to refer to *Hermitian* operators (a subset of normal operators A where $A^\dagger = A$), due to the property that all such operators have real eigenvalues. We will never be interested in the actual eigenvalues and so do not care about this distinction.

3.5 Tensor product

Before we can do anything interesting computationally, we need a way to reason about systems involving multiple qubits. In the single qubit case we saw our system as being a 2-dimensional state space with bases $|0\rangle$ and $|1\rangle$, and it follows naturally that we may see a 2-qubit system as nothing other than a linear combination of basis states $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$—a 4-dimensional state space. In general, then, it is not hard to see that a system involving n qubits is one that has dimension 2^n.

A more subtle question is to ask how exactly we may combine lower dimensional qubit systems to form larger ones. For example, if we have two systems each with one qubit, it seems reasonable that we should expect to be able to describe the quantum mechanical system of both qubits together. For this, we turn to a final fourth postulate.

Postulate 4 (State combination)**.** Quantum mechanical systems combine in a *bilinear fashion.* That is, if $|\psi\rangle, |\varphi\rangle$ are each state vectors for quantum systems, then denoting their combined state by $|\psi\rangle \otimes |\varphi\rangle$ we see the following properties hold:

1. $\alpha|\psi\rangle \otimes |\varphi\rangle = \alpha(|\psi\rangle \otimes |\varphi\rangle) = |\psi\rangle \otimes \alpha|\varphi\rangle$.
2. $(|\psi_1\rangle + |\psi_2\rangle) \otimes |\varphi\rangle = |\psi_1\rangle \otimes |\varphi\rangle + |\psi_2\rangle \otimes |\varphi\rangle$.
3. $|\psi\rangle \otimes (|\varphi_1\rangle + |\varphi_2\rangle) = |\psi\rangle \otimes |\varphi_1\rangle + |\psi\rangle \otimes |\varphi_2\rangle$.

The distributive laws 2 and 3 illustrate the fact that superposition is maintained in combining states. The first simply indicates that scaling the amplitude of either component system scales the amplitude of the combined system in the same way. This motivates the following definition.

Definition 3.15. Let U and V be a pair of vector spaces with bases $|i\rangle$ and $|j\rangle$ respectively, for $i \in I$ and $j \in J$. The **tensor product** of U and V, denoted $U \otimes V$, is the vector space generated by the basis $|ij\rangle$ for $i \in I$ and $j \in J$.

In addition, there is a bilinear map $U \times V \to U \otimes V$, where for $|u\rangle \in U, |v\rangle \in V$ we write

$$(|u\rangle, |v\rangle) \mapsto |u\rangle \otimes |v\rangle, \tag{30}$$

which we call the **tensor product of elements** $|u\rangle$ and $|v\rangle$. In particular, this map is given by the following:

$$\left(\sum_i u_i|i\rangle\right) \otimes \left(\sum_j v_j|j\rangle\right) = \sum_{ij} u_i v_j |ij\rangle. \tag{31}$$

It is trivial to verify that this map is indeed bilinear. For brevity, we will also use the notation $|u\rangle|v\rangle$ to denote the tensor product $|u\rangle \otimes |v\rangle$.

Corollary 3.16. $\dim U \otimes V = \dim U \dim V$.

It follows that our space of single qubit states $\mathcal{B}$ can be turned into a space of n-qubit states by tensoring with itself n times, written $\mathcal{B}^{\otimes n}$. Note here that the order of applying the tensor product does not matter, as it is associative up to the canonical isomorphism $(|u\rangle \otimes |v\rangle) \otimes |w\rangle \mapsto |u\rangle \otimes (|v\rangle \otimes |w\rangle)$.

Here we will also introduce the notion of the *computational basis*, which is the orthonormal basis for the n-qubit system obtained by tensoring the elements of the standard single qubit basis $\{|0\rangle, |1\rangle\}$ with themselves n times.

Definition 3.17. Let U, V, and the corresponding $U \otimes V$ be vector spaces, with the same notation for their respective bases as in the preceding definition. If A and B represent linear operators on U and V respectively, we may define $A \otimes B$ to be the linear operator such that

$$(A \otimes B)|ij\rangle = A|i\rangle \otimes B|j\rangle. \tag{32}$$

3.5.1 Tensor trouble (EPR pairs & quantum entanglement)

If you have been thinking about the physical implications, you may be wondering if it really is true that we can combine qubit states in the way described above. Let us turn our attention to the following, perfectly well-defined, 2-qubit state:

$$|\beta_{00}\rangle := \frac{|00\rangle + |11\rangle}{\sqrt{2}}. \tag{33}$$

Consider what happens if we measure only the first qubit in the computational basis. Suppose the outcome of the measurement is 0. Then $|0\rangle\langle 0|$ is the relevant projector, and by (11) of Postulate 3 in conjunction with Definition 3.17, we see the combined state collapse into

$$\begin{aligned} |\beta'_{00}\rangle &= \left(\sqrt{2}|0\rangle\langle 0| \otimes I\right)|\beta_{00}\rangle && (34)\\ &= \frac{1}{\sqrt{2}}\left[\left(\sqrt{2}|0\rangle\langle 0||0\rangle\right)|0\rangle + \left(\sqrt{2}|0\rangle\langle 0||1\rangle\right)|1\rangle\right] && (35)\\ &= \langle 0|0\rangle|0\rangle|0\rangle + \langle 0|1\rangle|0\rangle|1\rangle && (36)\\ &= |00\rangle. && (37) \end{aligned}$$

What of the other qubit? It was not measured, yet its state still appears to have collapsed: the measurement of the first qubit seems to have instantly affected the state of the second. At first glance this makes sense, the combined system started with zero amplitude for the state $|01\rangle$. However, recalling that the two qubits may be separated physically—indeed by an arbitrary distance—this could lead to some very troubling implications.[4]

The state $|\beta_{00}\rangle$ is what is known as a *Bell state* or *EPR pair*, and is an example of the famous notion of *quantum entanglement*. When quantum mechanics was still in its infancy, this idea suggested to some (most notably Einstein—the E in *EPR*) that quantum theory may not be a complete picture. Not only does it turn out that this seemingly paradoxical situation agrees with what we observe in reality, but also that we may be able to exploit it in our computation.

3.6 Quantum computation

With our understanding of qubit systems, along with evolution and measurement in quantum mechanics, we can finally examine what information processing tasks can actually be achieved. To start with a simple example, let us consider a XOR gate from classical computing—that is, a function on two bits $b_1, b_2 \in \mathbf{B} := \{0, 1\}$ computing addition modulo 2, which we denote $b_1 \oplus b_2$. We can attempt to define our quantum version analogously as below:

$$|b_1\rangle|b_2\rangle \mapsto |b_1 \oplus b_2\rangle, \tag{38}$$

for $|b_1\rangle|b_2\rangle$ in the computational basis with two qubits. There are a couple of problems here. The first is that our system goes from $\mathcal{B}^{\otimes 2}$ to $\mathcal{B}$, which violates Postulate 1. We can attempt to fix this issue by using only the second qubit for storing the output:

$$|b_1\rangle|b_2\rangle \mapsto |0\rangle|b_1 \oplus b_2\rangle. \tag{39}$$

Could *this* function be computed using quantum mechanics? The answer is, again, no. Postulate 2 stipulates that quantum mechanical systems must evolve according to unitary transformations, and it is quite easy to see that there is no unitary operator which could compute our function. In particular, all unitary operators are invertible (by definition), and our function as defined above is not even surjective.

[4]The relevant search term is "EPR paradox", for those interested.

Proposition 3.18. Let V be an inner product space with orthonormal basis $|i\rangle$. Any linear operator $A : V \to V$ such that A restricted to $|i\rangle$ is a permutation is unitary.

Proof. Let $|v\rangle = \sum_i a_i|i\rangle \in V$. Then

$$A|v\rangle = \sum_i a_i|\phi(i)\rangle \tag{40}$$

where $\phi : \{i\} \to \{i\}$ is a permutation. In particular,

$$\begin{aligned}\langle v|A^\dagger A|v\rangle &= \sum_{ij} a_i\overline{a_j}\langle\phi(j)|\phi(i)\rangle &(41)\\ &= \sum_{ij} a_i\overline{a_j}\delta_{ij} &(42)\\ &= \sum_i |a_i|^2 &(43)\\ &= \langle v|v\rangle, &(44)\end{aligned}$$

but $\langle v|A^\dagger A|v\rangle = \langle v|v\rangle \iff A^\dagger A = I$. Further, elementary algebra shows that $A^\dagger A = I \iff AA^\dagger = I$, and hence that A is unitary. □

A solution to our XOR problem is the map $|b_1\rangle|b_2\rangle \mapsto |b_1\rangle|b_1 \oplus b_2\rangle$. It can be verified quickly that this is indeed a permutation:

$$\begin{aligned}|00\rangle &\mapsto |00\rangle &(45)\\ |01\rangle &\mapsto |01\rangle &(46)\\ |10\rangle &\mapsto |11\rangle &(47)\\ |11\rangle &\mapsto |10\rangle. &(48)\end{aligned}$$

Fortunately, it is not so finicky to construct a unitary operator for a general function f. The above notion generalises such that it can be shown that for $|x\rangle|y\rangle$ in the computational basis, the map $|x\rangle|y\rangle \mapsto |x\rangle|y \oplus f(x)\rangle$ is unitary.[5][6]

Definition 3.19. Let $|i\rangle$ be the computational basis for $n+m$ qubits. For a function $f : \mathbf{B}^n \to \mathbf{B}^m$, write $U_f : \mathcal{B}^{\otimes n+m} \to \mathcal{B}^{\otimes n+m}$ to be the unitary operator defined by

$$\begin{aligned}U_f|\psi\rangle &= U_f \sum_i \alpha_i|i\rangle &(49)\\ &= U_f \sum_i \alpha_i|x_i\rangle|y_i\rangle &(50)\\ &:= \sum_i \alpha_i|x_i\rangle|y_i \oplus f(x_i)\rangle, &(51)\end{aligned}$$

where $|x_i\rangle \in \mathcal{B}^{\otimes n}$ is the "first" n qubits in $|i\rangle$ and $|y_i\rangle \in \mathcal{B}^{\otimes m}$ is the remaining m. Often, if the specific binary representation is not important or is obvious,[7] we will use this notation even when the underlying domain and codomain of the function f are not of the form $\mathbf{B}^n$.

Thus, if $|y_i\rangle$ is prepared to be $|0\rangle^{\otimes m}$, we can extract $f(x_i)$ by examining the last m qubits of $U_f|\psi\rangle$. Or can we? Those qubits will be in superposition—and if measuring a qubit in superposition were so straightforward then we would not have had an entire grueling section on it.

[5] If $y = a_1, a_2, \ldots, a_n \in \mathbf{B}^n$, and $f(x) = b_1, b_2, \ldots, b_n \in \mathbf{B}^n$, then $y \oplus f(x) := a_1 \oplus b_1, \ldots, a_n \oplus b_n$.

[6] Indeed, the map $x \mapsto x \oplus y$ is actually self-inverse, from which the suffient permutation property results.

[7] And, in particular, when the necessary domain extension to a set of size 2^N for some N is obvious.

3.6.1 So close, and yet so far

To illustrate this point, let us imagine we have a large semiprime $s = p_1 \cdot p_2$ which we wish to factor. A naïve way one may try to implement this is the following: take a function $f : \mathbf{N} \to \mathbf{B}$ such that $n \mapsto 1$ if and only if n divides s non-trivially. Then applying U_f to the state vector

$$|\psi\rangle = \frac{1}{\sqrt{2^N}} \sum_n |n\rangle|0\rangle, \tag{52}$$

where N is the number of binary digits required to encode the natural numbers less than or equal to $\sqrt{s}$, yields

$$U_f|\psi\rangle = \frac{1}{\sqrt{2^N}} \sum_n |n\rangle|f(n)\rangle \tag{53}$$

$$= \frac{1}{\sqrt{2^N}} \sum_n x_n, \tag{54}$$

where

$$x_n = \begin{cases} |n\rangle|1\rangle, & \text{if } n = p_1 \text{ or } n = p_2, \\ |n\rangle|0\rangle, & \text{otherwise.} \end{cases} \tag{55}$$

It seems like the work has been done, we have checked all potential factors of s in parallel! However, measuring in the computational basis yields us p_1 or p_2 only with probability $1/2^{N-1}$, with the vast majority of attempts resulting in a random non-factor of s. Even (very generously) assuming the action of U_f—as well as the time to set up the initial state $|\psi\rangle$—to be $O(1)$, the implication is that to get p_1 or p_2 with an arbitrary constant probability, the method of applying U_f and measuring the result is $O(2^N)$—in other words, intractable.

It is clear, then, that thanks to quantum measurement theory, any exponential increase in information processing will not be gained for free. The field of quantum algorithms is a deep one, but for our purpose of leading onto the hidden subgroup problem we will be focusing on a particularly successful strategy; namely of exploiting the same *quantum parallelism* (Deutsch 1985) afforded by superposed entangled states as above by transforming the final state to expose certain periodicities.

4 The Hidden Subgroup Problem

The hidden subgroup problem is a group theoretic abstraction, and so in this section we will give a very brief reminder to some elementary notions of group theory so as to build a group theoretic idea of periodicity; one which we will later exploit using some elementary concepts from the theory of representations in the general solution for the problem on finite abelian groups as presented by (R. Jozsa 2001; R. Jozsa 1998; Ekert and Richard Jozsa 1998).

4.1 A group theoretic notion of periodicity

The notion of a group is a powerful abstraction lying at the heart of countless structures in algebra, from rings and fields to symmetries of geometric objects. A group is nothing other than a set G coupled with a binary operation $* : G \times G \to G$, satisfying the following axioms:

1. (**Existence of identity**). $\exists e \in G : \forall g \in G, g * e = e * g = g$.
2. (**Existence of inverses**). $\forall g \in G$, $\exists g^{-1} \in G : g * g^{-1} = g^{-1} * g = e$.

3. (**Associativity**). $\forall g_1, g_2, g_3 \in G$, $g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3$.

From now on, we will be concerned only with those groups which are *abelian*, satisfying a fourth axiom of commutitivity. As such we will often use the additive notation $+$, with the inverse of g expressed by $-g$.

Definition 4.1. Let G be a group with $H \leq G$ (that is, H is a subgroup of G). We define the **cosets** of H in G to be all sets of the form

$$g + H := \{g + h : h \in H\}, \tag{56}$$

for g in G.

A property of cosets that will be important is that they split the group elements into a partition of equal-sized subsets. Indeed, it may be easily shown that for all $g_1, g_2 \in G$, we have $g_1 \in g_0 + H \iff g_0 + H = g_1 + H$. Further, the map $h \mapsto g + h$ is a bijection from H onto $g + H$, and so in particular $|g + H| = |H|$ for all $g \in G$.

It is this property that allows us to use cosets to generalise the idea of periodicity to abelian groups. If a function f is constant on cosets of $H \leq G$, we see that necessarily

$$f(g + h) = f(g) \tag{57}$$

for all $g \in G, h \in H$, and we may think of f as being h-periodic.

4.2 Complex representations of groups

Representation theory is a branch of mathematics that deals with representing groups by linear maps. In particular, for a vector space V we consider the set of all invertible linear maps $\phi : V \to V$ along with ordinary composition. Together, this forms a group, called the *general linear group* of V, denoted $GL(V)$.

Definition 4.2. For a group G, a **representation** of G in V is a map

$$\chi : G \to GL(V), \tag{58}$$

where χ is a group homomorphism. That is,

$$\chi(g_1 + g_2) = \chi(g_1)\chi(g_2), \tag{59}$$

for all $g_1, g_2 \in G$.

Consider the one dimensional vector space $\mathbf{C}$, and the corresponding group $GL(\mathbf{C}) \cong \mathbf{C}^\times$, where $\mathbf{C}^\times$ denotes the multiplicative group of complex numbers without 0. It may not be surprising that representations in $\mathbf{C}$ will be useful to us, being as the complex numbers is our field of interest for quantum state spaces. What do these representations look like?

In particular, we may ask for representations of the additive group $G = \mathbf{Z}_n$ in $\mathbf{C}$. That is, homomorphisms $\chi : \mathbf{Z}_n \to \mathbf{C}^\times$. For χ to be a group homomorphism, the identity must be preserved, and hence we have

$$1 = \chi(0) = \chi(\underbrace{a + \cdots + a}_{n \text{ times}}) = \underbrace{\chi(a) \cdots \chi(a)}_{n \text{ times}} \tag{60}$$

for all $a \in \mathbf{Z}_n$. The implication here is that by mapping each group element a to an nth root of unity, we may hope to find a representation. Looking at the generators of $\mathbf{Z}_n$, we may obtain a homomorphism by choosing an arbitrary root ω to be that which is mapped to by 1. For example, the map

$$a \mapsto e^{2\pi i \frac{a}{n}} \tag{61}$$

is a representation of $\mathbf{Z}_n$ in $\mathbf{C}$. This result generalises easily to show that *any* finite abelian group G has $|G|$ distinct representations $\chi : G \to \mathbf{C}^\times$ such that each $\chi(g)$ is a $|G|^{\text{th}}$ root of unity. We will make use of this fact shortly.

4.3 The hidden subgroup problem

Let G be any finite abelian group, and H a subgroup of G. Given a quantum machine to evaluate a function $f : G \to X$, such that f is constant on cosets of H (and distinct on different cosets), the hidden subgroup problem asks for us to determine H.

It is not immediately clear that quantum computing should give us an advantage here. However, the claim has already been made that we will be able to make ground using the previously mentioned notion of quantum parallelism, so let us follow that route. Forgetting about systems of qubits for the time being, suppose we have a state space $V \otimes W$ with $\dim V = |G|$, $\dim W = |X|$. Then, we may construct an orthonormal basis for V of the form $\{|g\rangle : g \in G\}$, and similarly for W, and after reaching an analogous position to (53) we have the state

$$|\psi\rangle = \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g\rangle |f(g)\rangle. \tag{62}$$

Of course, measuring $|\psi\rangle$ now would not be a great strategy, resulting in $|g_0\rangle|f(g_0)\rangle$ for some uniformly random choice of $g_0 \in G$. Our only hope for finding out information about H this way is to try it successively, learning something whenever we fortuitously find some g_1, g_2 with $f(g_1) = f(g_2)$. This could be exceedingly unlikely, and regardless is no better than could be done classically. The first insight is that we may instead try to exploit the entanglement in the state.

If we take a measurement in only the second register[8], we again see a result $f(g_0)$ for some uniformly chosen $g_0 \in G$. The result of this measurement is not important and may be discarded. What is important, however, is that only the values in the first register with nonzero amplitudes in the combined state will remain, and that is exactly those $g \in G$ which are in the same coset as g_0. In other words, the state after measurement will be given by[9]

$$|\psi'\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |g_0 + h\rangle. \tag{63}$$

The second insight is a subtle change of perspective on this state in (63). Consider a set of linear *shifting operators* (R. Jozsa 2001) $U(g) : V \to V$ for $g \in G$ given by

$$U(g_1)|g_2\rangle = |g_1 + g_2\rangle \tag{64}$$

for all $g_1, g_2 \in G$. Then by the group axioms, it is not hard to see that these operators are each permutations on $|g\rangle$ and hence unitary by Proposition 3.18. Furthermore, we may rewrite (63) as

$$|\psi'\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} U(g_0)|h\rangle. \tag{65}$$

The final and most crucial idea is to ask whether there exists some basis $|\chi_i\rangle$ which is *invariant under these shifting operators*, in the sense that

$$U(g)|\chi_i\rangle = e^{j\phi(g,i)}|\chi_i\rangle, \tag{66}$$

where j here is the imaginary unit, for some arbitrary $\phi : G \times I \to \mathbf{R}$. If there were, then in that basis we would see that the state in (65) and the state

$$\frac{1}{\sqrt{|H|}} \sum_{h \in H} |h\rangle \tag{67}$$

[8]Writing $|\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle$, we may use the terminology i*th register* to refer to the system represented by $|\psi_i\rangle$.

[9]Note that we have not violated Postulate 1 or 2 by changing the state space. In the physical system, the second register still exists. We omit it because it is no longer of interest.

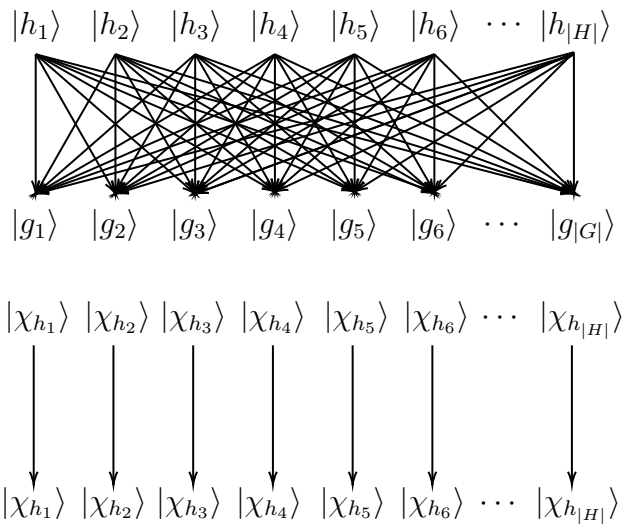

Figure 1: The action of $U(g_0)$ on the superposed states of (67) with respect to bases $|g\rangle$ and $|\chi_g\rangle$, for $h_i \in H$, subject to differing phase factors. Without knowing g_0, $|h_i\rangle$ maps to an effectively random state whilst χ_{h_i} maintains information about H.

would have the same distribution of outcomes when measured in the basis $|g\rangle$. Hence, if we were able to transform the basis of the quantum state from $|\chi_i\rangle$ to $|g\rangle$, then we would be able to gain some kind of information about H with *every* measurement! (R. Jozsa 2001) It turns out, of course, that there is such a basis. The provocatively unified notation should indicate that this is where we will be finally applying the result from our foray into representation theory.

We established that for any finite abelian group G, there exist $|G|$ distinct representations $\chi : G \to \mathbf{C}^\times$. Then, we may index each of these by a unique element $g \in G$ to obtain $|G|$ representations χ_g, and consider, for each $k \in G$, the state (ibid.)

$$|\chi_k\rangle := \frac{1}{\sqrt{|G|}} \sum_{g \in G} \chi_k(g)|g\rangle. \tag{68}$$

We will now show that $|\chi_g\rangle$ is suitable to be our alternate basis. The shift invariance is satisfied:

Proof. For all $k, g_0 \in G$,

$$U(g_0)|\chi_k\rangle = \frac{U(g_0)}{\sqrt{|G|}} \sum_g \chi_k(g)|g\rangle \tag{69}$$

$$= \frac{1}{\sqrt{|G|}} \sum_g \chi_k(g)|g + g_0\rangle, \tag{70}$$

but $g \mapsto g + g_0$ is a permutation, so

$$\frac{1}{\sqrt{|G|}} \sum_g \chi_k(g)|g + g_0\rangle = \frac{1}{\sqrt{|G|}} \sum_{g^*} \chi_k(g^* - g_0)|g^*\rangle \tag{71}$$

where $g^* = g + g_0$. Then, recalling that χ_k is a homomorphism,

$$\frac{1}{\sqrt{|G|}} \sum_{g^*} \chi_k(g^* - g_0)|g^*\rangle = \frac{1}{\sqrt{|G|}} \sum_g \chi_k(g - g_0)|g\rangle \tag{72}$$

$$= \frac{\chi_k(-g_0)}{\sqrt{|G|}} \sum_g \chi_k(g)|g\rangle \tag{73}$$

$$= \chi_k(-g_0)|\chi_k\rangle, \tag{74}$$

but $\chi_k(-g_0)$ is a $|G|^{\text{th}}$ root of unity, and so we have the desired relation. □

Further, they are linearly independent. Indeed, they are in fact orthonormal.

Proof.

$$\langle \chi_i | \chi_j \rangle = \frac{1}{|G|} \sum_{kg} \overline{\chi_i(k)} \chi_j(g) \langle k | g \rangle \tag{75}$$

$$= \frac{1}{|G|} \sum_{g} \overline{\chi_i(g)} \chi_j(g), \tag{76}$$

where an application of Shur's Lemma (see (R. Jozsa 1998)) tells us that (76) is equal to δ_{ij}. □

It is this orthonormality which actually guarantees that the linear map $|\chi_g\rangle \mapsto |g\rangle$ is unitary, and so can be applied to the state. The proof of this is extremely similar to that of Proposition 3.18, and so we will not include it; indeed, 3.18 may be thought of as a specific case of this more general fact—that linear maps which are bijections between orthonormal bases are unitary.

4.4 Some loose ends

We saw in the previous section how the hidden subgroup problem may be solved more efficiently by a quantum machine, but a few questions still remain. Primarily, what kind of information is actually gained after measuring the transformed basis? Clearly, it is not necessarily elements of H itself. Without going out of scope with regards to the underlying group theory, this is best seen by means of example; perhaps the most natural of which is that of period finding in $\mathbf{Z}_n$. That is, the hidden subgroup problem where $G = \mathbf{Z}_n$ and H is the subgroup generated by some $r \in \mathbf{Z}_n$.

After applying the transformation, we will have a system where the superposed states are of the form (R. Jozsa 2001; Ekert and Richard Jozsa 1996)

$$\left| j \frac{n}{r} \right\rangle \tag{77}$$

for $j = 0, \ldots, r-1$. Measuring will result in a value which is an integer multiple of n/r, and the majority of the time cancellations will be possible to determine r. In fact, the above example is actually the crux of Shor's algorithm. Given a number to factor, N, and some $a < N$ coprime to N, the problem reduces (R. Jozsa 2001) to determining the periodicity of the function

$$f(x) = a^x \mod N. \tag{78}$$

The other question pertains to the quite concerning modification made at the beginning of section 4.3; namely of *"forgetting about systems of qubits"*. Our entire computational system is founded on qubits!

If it happens that our group of interest is of size 2^N for some natural N, then there is no problem. Of course, however, the sparsity of such sizes grows exponentially. What can be done in most practical cases is that the group is "truncated" to a suitable set of size 2^N—for example[10], in Shor's algorithm we have the periodic function (78) on $\mathbf{Z}$, but the computation is actually done on some $\mathbf{Z}_{2^N}$ with addition modulo 2^N. Clearly, it is very unlikely that r should divide 2^N exactly, but for large enough N the "slightly spoilt" periodicity does not ruin the effectiveness of the algorithm (Ekert and Richard Jozsa 1996; R. Jozsa 2001).

[10]The N given in this example is distinct from the number the algorithm aims to factor, which was previously denoted also by N.

5 Conclusion

We have seen how the properties of quantum mechanics may be used to construct an alternative computational paradigm, and further seen an example of a class of problems where it appears very advantageous to do so. As hinted in the introduction, we saw that such advantageous gains were not acquired automatically. Let us reflect, then, on where exactly they *were* gained.

In a classical setting, scaling the number of bits in the system exponentially scales the operations that must be performed to act on each possible state. In the quantum paradigm, it is the tensor product acting on spaces of unitary operators which dictates the way operations can be performed. When we add one qubit to the system, the dimension of the corresponding operator space is doubled, and appears that this is where the exponential increase comes from. The properties of quantum measurement seek to thwart this gain, but as we have seen in an explicit example, it is sometimes still possible to overcome that barrier by further manipulations on the state. The (effectively equivalent) idea that entanglement is what gives rise to the increased computation is explored in more detail in (Ekert and Richard Jozsa 1998).

Bibliography and References

Deutsch, David (1985). "Quantum theory, the church–turing principle and the universal quantum computer". In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 400 (1818), pp. 97–117. DOI: 10.1098/rspa.1985.0070.

Dirac, P.A.M. (1939). "A new notation for quantum mechanics". In: *Mathematical Proceedings of the Cambridge Philosophical Society*.

Ekert, Artur and Richard Jozsa (July 1996). "Quantum computation and Shor's factoring algorithm". In: *Rev. Mod. Phys.* 68 (3), pp. 733–753. DOI: 10.1103/RevModPhys.68.733. URL: https://link.aps.org/doi/10.1103/RevModPhys.68.733.

— (Aug. 1998). "Quantum algorithms: entanglement–enhanced information processing". In: *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 356.1743. Ed. by A. Ekert, R. Jozsa, and R. Penrose, pp. 1769–1782. ISSN: 1471-2962. DOI: 10.1098/rsta.1998.0248. URL: http://dx.doi.org/10.1098/rsta.1998.0248.

Feynman, Richard P. (June 1982). "Simulating Physics with Computers". In: *International Journal of Theoretical Physics* 21.6-7, pp. 467–488. DOI: 10.1007/BF02650179.

Jozsa, R. (1998). "Quantum algorithms and the Fourier transform". In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 454.1969, pp. 323–337. ISSN: 1471-2946. DOI: 10.1098/rspa.1998.0163. URL: http://dx.doi.org/10.1098/rspa.1998.0163.

— (2001). "Quantum factoring, discrete logarithms, and the hidden subgroup problem". In: *Computing in Science & Engineering* 3.2, pp. 34–43. ISSN: 1521-9615. DOI: 10.1109/5992.909000. URL: http://dx.doi.org/10.1109/5992.909000.

Kitaev, A. Yu. (1995). *Quantum measurements and the Abelian Stabilizer Problem*. arXiv: quant-ph/9511026 [quant-ph].

Kitaev, A. Yu., A. H. Shen, and M. N. Vyalyi (1999). *Classical and Quantum Computation*. American Mathematical Society.

Lang, S. (1968). *Linear Algebra*. Addison-Wesley series in mathematics. Addison-Wesley.

— (1984). *Algebra*. Menlo Park, California: Addison-Wesley. ISBN: 0-201-05487-6.

Lomont, Chris (2004). "The Hidden Subgroup Problem - Review and Open Problems". In: arXiv: quant-ph/0411037 [quant-ph].

Nielsen, Michael A. and Isaac L. Chuang (2010). *Quantum Computation and Quantum Information*. Cambridge University Press.

Shor, P.W. (1994). “Algorithms for quantum computation: discrete logarithms and factoring”. In: *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pp. 124–134. DOI: 10.1109/SFCS.1994.365700.

Solovay, R. and V. Strassen (1977). “A Fast Monte-Carlo Test for Primality”. In: *SIAM Journal on Computing* 6.1, pp. 84–85. DOI: 10.1137/0206006.