Quantum computing and entanglement for mathematicians

Nolan R. Wallach

April 22, 2013

These notes are an expanded form of lectures to presented at the C.I.M.E. summer school in representation theory in Venice, June 2004. The sections of this article roughly follow the five lectures given. The first three lectures (sections) are meant to give an introduction to an audience of mathematicians (or mathematics graduate students) to quantum computing. No attempt is given to describe an implementation of a quantum computer (it is still not absolutely clear that any exist). There are also some simplifying assumptions that have been made in these lectures. The short introduction to quantum mechanics in the first section involves an interpretation of measurement that is still being debated which involves the "collapse of the wave function" after a measurement. This interpretation is not absolutely necessary but it simplifies the discussion of quantum error correction. The next two sections give an introduction to quantum algorithms and error correction through examples including fairly complete explanations of Grover's (unordered search) and Shor's (period search and factorization) algorithms and the quantum perfect (five qubit) code. The last two sections present applications of representation and Lie theory to the subject. We have emphasized the applications to entanglement since this is the most mathematical part of recent research in the field and this is also the main area to which the author has made contributions. The material in subsections 5.1 and 5.3 appears in this article for the first time.

1 The basics

In his seminal paper [F1] Richard Feynman introduced the idea of a computer based on quantum mechanics. Of course, all modern digital computers involve transistors that are by their very nature quantum mechanical. However, the quantum mechanics only plays a role in the theory that explains why the transistor switches. The actual switch in the computer is treated as if it were mechanical. In other words as if it were governed by classical mechanics. Feynman had something else in mind. The basic operations of a quantum computer would involve the allowable transformations of quantum mechanics, that is, unitary operators and measurements. The analogue of bit strings for a quantum computer are superpositions of bit strings (we will make this precise later) and the analogue of a computational step (for example the operation not on one bit) is a unitary operator on the Hilbert space of bit strings (say of a fixed length). The reason that Feynman thought that there was a need for such a "computer" is that quantum mechanical phenomena are extremely difficult (if not impossible) to model on a digital computer. The reason why the field of quantum computing has blossomed into one of the most active parts of the sciences is the work of Peter Shor [S1] that showed that on a (hypothetical) quantum computer there are polynomial time algorithms for factorization and discrete logarithms. Since most of the security of the internet is based on the assumption that these two problems are computationally hard (that is, the only known algorithms are superpolynomial in complexity) this work has attracted an immense amount of attention and trepidation. In these lectures we will discuss a model for computation based on this idea and discuss its power, ways in which it differs from standard computation and its limitations. Before we can get started we need to give a crash course in quantum mechanics.

1.1 Basic quantum mechanics

The *states* of a quantum mechanical system are the unit vectors of a Hilbert space, V, over $\mathbb{C}$ ignoring phase. In other words the states are the elements of the projective space of all lines through the origin in V. If $v, w \in V$ then we write $\langle v|w\rangle$ for the inner product of v with w. We will follow the physics convention so the form is conjugate linear in v and linear in w. Following Dirac a vector gives rise to a "*bra*", $\langle v|$ and a "*ket*" $|v\rangle$ the latter is exactly the same as v the former is the linear functional that takes the value $\langle v|w\rangle$ on w. Thus if v is a state then $\langle v|v\rangle = 1$. In these lectures most Hilbert spaces will be finite dimensional. For the moment we will assume that $\dim V < \infty$. An observable is a self adjoint operator, A, on V. Thus A has a spectral decomposition

$$V = \bigoplus_{\lambda \in \mathbb{R}} V_\lambda$$

with $A_{|V_\lambda} = \lambda I$. We can write this as follows. The spaces V_λ are orthogonal relative to the Hilbert space structure. Thus we can define the orthogonal projection $P_\lambda : V \to V_\lambda$. Then we have $A = \sum \lambda P_\lambda$. If v is a state then we set $v_\lambda = P_\lambda v$. A measurement of the state v with respect to an observable A yields a number λ that is an eigenvalue of A with the probability $\|v_\lambda\|^2$. This leads to the following problem. If we do another measurement almost instantaneously we should get a value close to λ. Thus one would expect the probability to be very close to 1 for the state to be in V_λ. In the standard formulation of quantum mechanics this is "explained" by the collapse of the wave function. That is, a measurement of by apparatus corresponding to the observable A has two effects. The first is an eigenvalue, λ of A (the measurement) with probability $\|v_\lambda\|^2$ and the second is that the state has collapsed to

$$\frac{v_\lambda}{\|v_\lambda\|}.$$

This is one of the least intuitive aspects of quantum mechanics. It has been the subject of much philosophical discussion. We will not enter into this debate and will merely take this as an axiom for our system.

If we have a quantum mechanical system then in addition to the Hilbert space V we have a self adjoint operator H the *Hamiltonian* of the system. The evolution of a state in this system is governed by *Schroedinger's equation*

$$\frac{d\phi}{dt} = iH\phi.$$

Thus if we have the initial condition $\phi(0) = v$ then

$$\phi(t) = e^{itH}v.$$

Thus the basic dynamics is the operation of unitary operators. If U is a unitary operator on V then $|Uv\rangle = U|v\rangle$ and $\langle Uv| = \langle v| U^{-1}$. This is the only consistent way to have $\langle Uv|Uv\rangle = \langle v|v\rangle$ for a unitary operator.

Of course, these finite dimensional Hilbert spaces do not exist in isolation. The state of the entire universe, u, is a state in a Hilbert space, U, governed by the Schroedinger equation with Hamiltonian H_U. We we will simplify the situation and think of the finite dimensional space V as a tensor factor of U that is

$$U = V \bigotimes E$$

with E standing for the *environment.* This is not a tremendous assumption since in practice the part of the universe that will have a real effect on V is given by this tensor product. Now, the Hamiltonian H_U will not preserve the tensor product structure. Thus, even though we are attempting to do only operations on states in V the environment will cause the states to change in ways that are beyond the control of the experiment that we might be attempting to do on states in V. Thus if we prepare a state on which we will do a quantum mechanical operation, that is, by applying a unitary transformation or doing a measurement we can only assume that the state will not "morph" into a quite different state for a very short time. This uncontrolled change of the state is called *decoherence* caused by the environment.

The fact that our small Hilbert space V is not completely isolated from the rest of the universe is the reason why it is more natural to use density matrices as the basic states. A *density matrix (operator)* is a self adjoint operator T on V that is positive semi-definite and has trace 1. In this context a state $v \in V$ would then be called a *pure state* and a density matrix a *mixed state.* If v is a pure state then its density matrix is $|v\rangle \langle v|$. We note that this operator is just the projection onto the line corresponding to the pure state v. Thus we can identify the pure states with the mixed states that have rank 1. If T is a mixed state then T transforms under a unitary operator by $T \mapsto kTk^{-1}$ if k is unitary. If we have a pure state u in U then it naturally gives rise to a mixed state on V which is called the *reduced density matrix* and is defined as follows. Let $\{e_i\}$ be an orthonormal basis of E. Then $u = \sum v_i \bigotimes e_i$ with $v_i \in V$. The reduced density matrix is $\sum |v_i\rangle \langle v_i|$. More generally, if T is a mixed state on U then it gives rise to a mixed state $\mathrm{Tr}_2(T)$ on V by the formula

$$\langle w|\mathrm{Tr}_2(T)|v\rangle = \sum_i \langle w \bigotimes e_i|T|v \bigotimes e_i\rangle .$$

This mixed state is the reduced density matrix. One checks easily that since unitary operators don't necessarily preserve the tensor product structure that a unitary transformation of the state, T, will not necessarily entail a unitary transformation of the reduced density matrix. We will mainly deal with pure states in these lectures. However, we should realize that this is a simplification of what nature allows us to see.

1.2 Bits

Although it is not mandatory we will look upon digital computing as the manipulation of bit strings. That is, we will only consider fixed sequences of 0's and 1's. One bit is either 0 or it is 1. Two bits can have one of four values $00, 01, 10, 11$. These four strings can be looked upon as the expansion in base 2 of the integers $0, 1, 2, 3$, they can be looked upon as representatives of the integers mod 4, or they can be considered to be the standard basis of the vector space $\mathbb{Z}_2 \times \mathbb{Z}_2$. In general, an n-bit computer can manipulate bit strings of length n. We will call n the *word length* of our computer. Most personal computers now have word length 32 (soon 64). We will not be getting into the subtleties of computer science in these lectures. Also, we will not worry about the physical characteristics of the machines that are needed to do bit manipulations. A computer also can hold a certain number of words in its *memory.* There are various forms of memory (fast, somewhat fast, less fast, slow) but we will ignore the differences. We will look upon a computer program as a sequence of steps (usually encoded by bit strings of length equal to the word length) which implement a certain set of rules that we will call the *algorithm.* The first step inputs a bit string into memory. Each succeeding step operates on a sequence of words in memory that came from the operation of the preceding step and produces another sequence of words, which may or may not replace some of the words from the previous step and may or may not put words into new memory locations. If properly designed the program will have rules that terminate it and under each of the rules an output of bit strings. That is the actual computation. There are, of course, other ways the program might terminate, for example it runs out of memory, it is terminated by the operating system for attempting to access protected memory locations, or even that it is terminated by the user out of impatience. In these cases there is no (intended) output except possibly an error message.

This is the von Neumann model of computation. The key is that the computer does one step of a program at a time. Most computers can actually do several steps at one time. But this is because the computers are actually several von Neumann computers working simultaneously. For example, a computer might have an adder and a multiplier that can work independently. Or it might have several central processors that communicate with each other and attempt to do program steps simultaneously. These modifications will only lead to a parallelism that its determined by the number of processors and can only lead to a constant speed-up of a computation. For example if we have 10 von Neumann computers searching through a sequence of N elements with the task of finding one with a specified property. For example you have $N-1$ red chips and 1 white one. The program might be set to divide the sequence into 10 subsequences each of size $\frac{N}{10}$ and then each processor is assigned the job of searching through through one part. In the worst case each processor will have to evaluate $\frac{N}{10}$ elements. So we see a speed up of a factor of 10 over using one processor in this simple problem (slightly less since the worst case with one processor is $N-1$).

We will come back to a few more aspects of digital computing as we develop a model for quantum computation.

1.3 Qubits

The simplest description of the basic objects to be manipulated by a quantum computer of word length n are complex superpositions of bit strings of length n. Since a bit string is a sequence of numbers and the coefficients of the superpositions can also be some of these numbers we will use the ket notation for the bit strings as pure states. These superpositions will be called *qubits*. Thus one *qubit* is a an element of the two dimensional vector space over $\mathbb{C}$ with states

$$a\left|0\right\rangle + b\left|1\right\rangle$$

and $|a|^2 + |b|^2 = 1$. We will be dealing with qubits quantum mechanically so we ignore phase (multiples by complex numbers of norm 1). Thus our space of qubits is one dimensional projective space over $\mathbb{C}$. We will think of this qubit as being in state $\left|0\right\rangle$ with probability $|a|^2$ and in state $|b|^2$. Although this is a vast simplification we will take the simplest one step operation on a qubit to be a unitary operator (projective unitary operator to be precise).

Contrasting this with bits we see that on the set of bits $\{0, 1\}$ there are exactly 2 basic reversible operations: the identity map and NOT that interchanges 0 and 1. In the case of qubits we have a 3 dimensional continuum of basic operations that can be done. There is only one caveat. After doing these operations which are difficult to impossible classically we must do a measurement to retrieve a bit. This measurement will yield 0 or 1 with some probability. Thus in a very real sense going to qubits and allowing unitary transformations has not helped at all.

An element of 2 qubit space will be of the form

$$u = a\left|00\right\rangle + b\left|01\right\rangle + c\left|10\right\rangle + d\left|11\right\rangle$$

with $|a|^2 + |b|^2 + |c|^2 + |d|^2 = 1$. We interpret this as u is in state $\left|00\right\rangle$ with probability $|a|^2$, in state $\left|01\right\rangle$ with probability $|b|^2$, etc. Similarly for n qubits. The steps in a quantum computation will be unitary transformations. However, each unitary transformation given in a step will have to be broken up into basic transformations that we can construct with a known and hopefully small cost (time and storage).

A quantum program starts with an n qubit state, u_0, the input, and then does a sequence of unitary transformations T_j on the state so the steps are $u_1 = T_1 u_0, ..., u_m = T_m u_{m-1}$, and a rules for termination and at termination a measurement. The output is the measurement and or the state to which the measured state has collapsed.

References

[F1] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys., 21(1982),

[S1] P. W. Shor, Algorithms for quantum computation, discrete logarithms and factorizing, Proceedings 35th annual symposium on foundations of computer science, IEEE Press, 1994.

2 Quantum algorithms

In the last lecture we gave simple models for a classical and a quantum computation. In this lecture we will give a very simple example of a quantum algorithm that implements that does something that is impossible to do on a Von Neumann computer. We will next give a more sophisticated example of a quantum algorithm (Grover's algorithm [G1]) that does an unstructured search of N objects of the type described in the last lecture in $\sqrt{N}$ steps. At the end of the lecture we will introduce the quantum (fast) Fourier transform and explain why on a (hypothetical) quantum computer it is exponentially faster than the Fast Fourier transform

2.1 Quantum parallelism

Suppose that we are studying a function, f, on bit strings that takes the values 1 and -1 and assume that it takes only one step on a classical computer to calculate its value given a bit string. For example the function that takes value 1 if the last bit is 0, -1 if it is 1. We will think of bit strings of length n as binary expansions of the numbers $0, 1, ..., 2^n - 1$. Thus our n qubit space, V, has the orthonormal basis $|0\rangle, |1\rangle, ..., |N-1\rangle$ with $N = 2^n - 1$. We can replace f by the unitary operator defined by $T|j\rangle = f(j)|j\rangle$. T operates on a state $v \in V$,

$$v = \sum_{j=0}^{N-1} a_j |j\rangle, \sum_{j=0}^{N-1} |a_j|^2 = 1$$

by

$$Tv = \sum_{j=0}^{N-1} f(j)a_j \left|j\right\rangle .$$

Quantum mechanically this means that we have calculated $f(j)$ with the probability $\left|a_j\right|^2$. In other words the calculation of T on this superposition seems to have calculated all of the values of $f(j)$ simultaneously if all of the $|a_j| > 0$ in one quantum step. In a sense we have, but the rub is that if we do a measurement then all we have after a measurement is $f(j)\left|j\right\rangle$ with probability $|a_j|^2$ and since we ignore phase the value the object we are calculating is lost. Perhaps it would be better to decide that we will operate quantum mechanically and then read of the coordinates classically? I assert that we will still not be able to make direct use of this parallelism. The reason is that we are only interested in very big N. In this situation the set of states, $\sum_{j=0}^{N-1} a_j \left|j\right\rangle$, with $|a_j|^2$ all about the same size have a complement in the sphere of extremely small volume. This implies that most of the states will have probabilities, $|a_j|^2 \sim \frac{1}{N}$. If n is, say, 1000 then all the coordinates will be too small to measure classically. We can see this as follows. We consider the unit sphere in real N dimensional space. Let ω_N be the $O(N)$ invariant volume element on S^{N-1} that is normalized so that

$$\int_{S^{N-1}} \omega_N = 1.$$

We write a state in the form $v = \cos\theta u + \sin\theta \left|N-1\right\rangle$ with $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$. With u an element of the unit sphere in $N-1$ dimensional space. Then we have

$$\omega_N = c_N \cos\theta^{n-2} \omega_{N-1} \bigwedge d\theta$$

and $c_N \sim C\sqrt{N}$ with C independent of N. The set of all v in the sphere with last coordinate a_{N-1}that satisfies $|a_{N-1}|^2 \geq r^2 > \frac{1}{N} + \varepsilon$ with $\varepsilon > 0$ has volume at most

$$C\sqrt{N}(1-r^2)^{\frac{N}{2}-1} = C\sqrt{N}(1-\varepsilon)^{\frac{N}{2}-1}\left(1 - \frac{\frac{1}{1-\varepsilon}}{N}\right)^{\frac{N}{2}-1}$$

which is extremely small for N large.

The upshot is that a quantum algorithm must contain a method of increasing the size of the coefficient of the desired output so that when a measurement is made will have the output with high probability.

2.2 The tensor product structure of n-qubit space

Recall that the standard (sometimes called the *computational*) basis of the space of 2 qubits is $\left|00\right\rangle, \left|01\right\rangle, \left|10\right\rangle, \left|11\right\rangle$. A physicist would also write $\left|0\right\rangle\left|1\right\rangle = \left|01\right\rangle$. We mathematicians would rather think that the multiplication is a tensor product. That is $\left|0\right\rangle, \left|1\right\rangle$ form the standard basis of $\mathbb{C}^2$. Then

$$\left|0\right\rangle \bigotimes \left|0\right\rangle, \left|0\right\rangle \bigotimes \left|1\right\rangle, \left|1\right\rangle \bigotimes \left|0\right\rangle, \left|1\right\rangle \bigotimes \left|1\right\rangle$$

form an orthonormal basis of $\mathbb{C}^2 \bigotimes \mathbb{C}^2$ with the tensor product Hilbert space structure. In other words we identify $|ab\rangle$ with $|a\rangle \bigotimes |b\rangle$. In this form the original bit strings are fully decomposable that is are tensor products of n elements in $\mathbb{C}^2$. We will call an n-qubit state a product state if it is of the form

$$v_1 \bigotimes v_2 \bigotimes \cdots \bigotimes v_n$$

with $\|v_i\|^2 = 1$ for $i = 1, ..., n$. One very important product state is the uniform state ($N = 2^n$):

$$v = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} |j\rangle .$$

To see that it is indeed a product state we set $u = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Then $v = u \bigotimes \cdots \bigotimes u$ (n-fold tensor product). This formula also shows that the uniform state can be constructed in $n = \log_2(N)$ steps. This can be seen by making an apparatus that implements the one qubit unitary transformation (called the *Hadamard transformation*)

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

It has the property that $H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, H|1\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$. We write $H(k)$ for $I \bigotimes \cdots \bigotimes H \bigotimes \cdots I$ with all factors one qubit operations and all factors but one the identity and in the k-th factor the Hadamard transformation. Thus on a quantum computer that can implement a one qubit Hadamard transformation in constant time can construct the uniform state in logarithmic time. We will actually over simplify the model and assume that all one qubit operations can be implemented is one step on a quantum computer, Then

$$u = H(1)H(2) \cdots H(n) |0\rangle .$$

With this in mind we can give our first quantum algorithm. Set up an apparatus that corresponds to an observable, A, with simple spectrum. Here is the algorithm:

Make a uniform state v.
Measure A.
v collapses to $|j\rangle$ with j between 0 and $N-1$ with probability $\frac{1}{N}$.

In other words we are generating truly random numbers. The complexity of this algorithm is n. On a digital computer the best one can do is generate peudo random numbers. The classical algorithms involve multiplication and division. Thus they are slightly more complex. However they do not generate random numbers and no deterministic algorithm can (since the numbers will satisfy the property that they are given by the algorithm).

2.3 Grover's algorithm

We return to unstructured search. We assume that we have a function, f, on n-bit strings that takes the value -1 on exactly one string and 1 on all of the others. We assume that given a bit string the calculation of the value is one step (in computer science f might be called an *oracle*). Here is Grover's algorithm:

Form the uniform state $u = \frac{1}{\sqrt{N}} \sum |j\rangle$. Let T be the unitary transformation defined by $T|j\rangle = f(j)|j\rangle$. Let S be the orthogonal reflection about u. That is

$$S_u(v) = v - 2\langle u|v\rangle u.$$

Then S_u is a unitary operator that in theory can be implemented quantum mechanically with logarithmic complexity (indeed Grover gave a formula for S_u involving the order of n Hadamard transformations). If the number of bits is 2 (we are searching a list of 4 elements) then we observe

$$STu = -|j\rangle$$

with $f(j) = -1$. Thus one quantum operation and one measurement yields the answer. Where as classically in the worst case we would have had to calculate f three times and then printed the answer.

The general algorithm is just an iteration of this step. $u_0 = u$ and $u_{m+1} = STu_m$. A calculation using trigonometry shows that after $[4\pi\sqrt{N}]$ steps the coefficient of $|j\rangle$ with $f(j) = -1$ has absolute value squared .99.. (Here $[x]$ is the maximum of the set of integers less than or equal to x). Thus with almost certainty a measurement at this step in the iteration will yield the answer.

2.4 The quantum Fourier transform

Interpreted as a map of $L^2(\mathbb{Z}/N\mathbb{Z})$ to itself the fast Fourier transform can be interpreted as a unitary operator on this Hilbert space. In general, if G is a finite abelian group of order $|G|$ then we define the Hilbert space $L^2(G)$ to be the space of a complex valued functions on G with inner product

$$\langle f|g\rangle = \sum_{x\in G} \overline{f(x)}g(x).$$

Let $\widehat{G}$ denote the set of unitary characters of G. Then it is standard that the set $\{\frac{1}{\sqrt{|G|}}\chi|\chi \in \widehat{G}\}$ is an orthonormal basis of $L^2(G)$. If $G = \mathbb{Z}/N\mathbb{Z} = \mathbb{Z}_N$ and if we set $\chi_m(n) = e^{\frac{2\pi i n m}{N}}$ for $m = 0, ..., N-1$ then we can define

$$\mathcal{F}(f)(m) = \left\langle \frac{1}{\sqrt{N}}\chi_m|f \right\rangle = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f(n)\chi_m(n)^{-1}$$

and so

$$f(n) = \sum_{m=0}^{N-1} \left\langle \frac{1}{\sqrt{N}}\chi_m|f \right\rangle \frac{1}{\sqrt{N}}\chi_m(n) = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} \mathcal{F}(f)(m)\chi_m(n).$$

As in the case of the fast Fourier transform we will take $N = 2^n$when we estimate its complexity however it makes sense for any N. The standard orthonormal basis of $L^2(\mathbb{Z}_N)$ is the set of delta functions set $\{\delta_m|m = 0, ..., N-1\}$ with

$\delta_m(x) = 1$ if $x = m$ and 0 otherwise. We will identify these delta functions with the computational basis, that is $|m\rangle = \delta_m$. We therefore have

$$\mathcal{F}\,|m\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} \chi_m(j)^{-1} |j\rangle \,.$$

The linear extension to n qubit space is the quantum Fourier transform. The discussion above makes it obvious that this is a unitary operator. What is less obvious is that we can devise a quantum algorithm to implement this operator as (essentially) a tensor product of one qubit operators (which we are assuming are easily implemented on our hypothetical quantum computer). We will conclude this section with the factorization when $N = 2^n$ (due to Shor [S2]) that suggests a fast quantum algorithm

We write if $0 \leq j \leq N-1$ then we write $j = \sum_{i=0}^{n-1} j_i 2^i$ with $j_i \in \{0,1\}$ so that with our convention $|j\rangle = |j_{n-1} j_{n-2} \cdots j_0\rangle$. If $0 \leq m \leq N-1$ then

$$\frac{m}{N} = \sum_{i=1}^{n} m_{n-i} 2^{-i}$$

and since

$$2^{-k} j = \sum_{l=0}^{k-1} j_l 2^{-k+l} + u_{kj}$$

with $u_{kj} \in \mathbb{Z}$. We have

$$e^{2\pi i \frac{m}{N} j} = e^{2\pi i \sum\limits_{k=1}^{n} m_{n-k} \sum\limits_{l=0}^{k-1} j_l 2^{-k+l}} .$$

This leads to the following factorization

$$\mathcal{F}\,|j\rangle = u_n(j) \bigotimes u_{n-1}(j) \bigotimes \cdots \bigotimes u_1(j)$$

with

$$u_k(j) = \frac{|0\rangle + e^{2\pi i \sum\limits_{l=0}^{k-1} j_l 2^{-k+l}} |1\rangle}{\sqrt{2}}.$$

References

[G] L.K.Grover, Quantum mechanics helps in searching for a needle in a haystack, Phys, Rev, Let.,79,1997.

[S2] P.K.Shor, Polynomial time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM J. Comp.,26 1484-1509.

3 Factorization and error correction

In this section we will study the complexity of the quantum Fourier transform and indicate its relationship with Shor's factorization algorithm. We will also discuss the role of error correction in quantum computing and describe a quantum error correcting code.

3.1 The complexity of the quantum Fourier transform

Recall that our simplified model takes a one qubit unitary operator to be one computational step this is a simplification but the one qubit operators that will come into the rest of the discussion of the quantum Fourier transform are provably of constant complexity. We will also be using some two qubit operations which are also each of constant complexity. In addition we assume an implementation of the total flip, τ

$$v_1 \bigotimes v_2 \bigotimes \cdots \bigotimes v_n \mapsto v_n \bigotimes v_{n-1} \bigotimes \cdots \bigotimes v_1$$

One can show that the complexity of this operation is a multiple of n. We will show how to implement the transformation

$$|j\rangle \mapsto u_n(j) \bigotimes u_{n-1}(j) \bigotimes \cdots \bigotimes u_1(j) = \mathcal{F}\,|j\rangle$$

with

$$u_k(j) = \frac{|0\rangle + e^{2\pi i \sum\limits_{l=0}^{k-1} j_l 2^{-k+l}} |1\rangle}{\sqrt{2}}.$$

To describe the steps in the implementation we need the notion of a *controlled* one qubit operation. Let $U \in U(2)$ we define a unitary operator, C_U, on $\mathbb{C}^2 \bigotimes \mathbb{C}^2$ as follows

$$C_U\,|j_1 j_2\rangle = (U\,|j_1\rangle) \bigotimes |j_2\rangle$$

if $j_2 = 1$ and

$$C_U\,|j_1 j_2\rangle = |j_1 j_2\rangle$$

if $j_2 = 0$. We call j_2 the *control bit* if we are operating on n qubits and applying a controlled U operation with the control in the k-th factor and the operation in the l-th factor then we will write $C_U^{l,k}$ (the reader should be warned that this is not standard notation). Thus

$$C_U^{23}\,|0110\rangle = |0\rangle \bigotimes U\,|1\rangle \bigotimes |1\rangle \bigotimes |0\rangle$$

and

$$C_U^{23}\,|0100\rangle = |0100\rangle\,.$$

If U is easily implemented then controlled U is also easily implemented. We define

$$U_k = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{2\pi i}{2^k}} \end{bmatrix}$$

and recall that the Hadamard operator acting on the k-th qubit was denoted $H(k)$ in section 2. We will now describe an operator the implements the quantum Fourier transform. It will be a product $\tau \circ A_n A_{n-1} \cdots A_1 \circ \tau$ with

$$A_1 = C_{U_n}^{1,n} C_{U_{n-1}}^{1,n-1} \cdots C_{U_2}^{1,2} H(1), ...$$
$$A_k = C_{U_k}^{k,n} C_{U_{k-1}}^{k,n-1} \cdots C_{U_2}^{k,k+1} H(k), ..., A_n = H(n).$$

We note that in this expression the operator A_k changes the k-th qubit but doesn't depend on the value of the j-th qubit for $j < k$. We leave it to the reader to expand the product and see that it works. The operator A_k is a product of k operators that we can assume are implemented in constant time. Thus the complexity of the transform is a constant times $\frac{n(n+1)}{2}$. This is exponentially faster than the classical fast Fourier transform which has complexity Nn. It has been pointed out that this algorithm involves high precision in the claculation of the phases $e^{\frac{2\pi i}{2^k}}$ which would cause exponential overhead. This question is adressed in [NC] Exercise 5.6 p. 221. In more detail it is studied by Coppersmith in [Co].

3.2 The Shor period search algorithm

Shor introduced this transform in order to give a generalization of an algorithm of Deutch (cf. [NC]). Shor's algorithm finds the period of a function with an unknown period with complexity a power of the number of bits involved. His reason for doing this was that he had a method of reducing factorization to period search (see the next section). An exposition of a period finding algorithm can be found in [NC] where they reduce the problem to what they call phase approximation. We will give a description of (essentially) Shor's original argument [S2].

We assume that n is a positive integer and $0 < x < n$ is relatively prime to n. We follow [S2] to calculate the order of x in the ring $\mathbb{Z}_n$. That is the smallest $r > 0$ such that $x^r \equiv 1 \bmod n$. We choose m with

$$n^2 \leq 2^m < 2n^2.$$

Set $q = 2^m$. Let $\mathbb{C}\mathbb{Z}_n$ be the group algebra of $\mathbb{Z}_n$ with orthonormal basis $|j \bmod n\rangle$, $j = 0, ..., n-1$ which will be taken as the computational basis. We consider the tensor product Hilbert space $\left(\otimes^m \mathbb{C}^2\right) \otimes \mathbb{C}\mathbb{Z}_n$ and the state

$$\frac{1}{\sqrt{q}} \sum_{a=0}^{q-1} |a\rangle \otimes |1\rangle$$

(see the discussion of Grover's algorithm for how this state might be constructed). We assume that the U^k operation on $\mathbb{C}\mathbb{Z}_n$ given by $U^k\ |j \bmod n\rangle = \left|x^k j \bmod n\right\rangle$has been implimented. Using the first factor (register) as a control we apply $I \otimes U^j$ to the superposition and have

$$\frac{1}{\sqrt{q}} \sum_{a=0}^{q-1} |a\rangle \otimes |x^a \bmod n\rangle \, (*).$$

Shor gives a detailed exposition of how this state is not hard to impliment that is basically the classical way of taking high powers in the ring $\mathbb{Z}_n$. In any event we begin with the above state and explain method.

First we take the quantum Fourier transorm in the first factor so

$$|a\rangle \longmapsto \frac{1}{\sqrt{q}}\sum_{c=0}^{q-1} e^{2\pi i\frac{ac}{q}}\left|c\right\rangle .$$

We now have

$$\frac{1}{q}\sum_{a,c=0}^{q-1} e^{2\pi i\frac{ac}{q}}\left|c\right\rangle \otimes \left|x^{a} \operatorname{mod} n\right\rangle .$$

This we rewrite

$$\frac{1}{q}\sum_{k=0}^{r}\sum_{\substack{a,c=0\\ a\equiv k \operatorname{mod} r}}^{q-1} e^{2\pi i\frac{ac}{q}}\left|c\right\rangle \otimes \left|x^{k} \operatorname{mod} n\right\rangle .$$

Now for each term with the second factor $\left|x^{k} \operatorname{mod} n\right\rangle$ .We have $a = br + k$ and the b vary between 0 and $\left\lfloor \frac{q-k-1}{r}\right\rfloor$ so we have

$$\frac{1}{q}\sum_{k=0}^{r}\sum_{c=0}^{q-1}\sum_{b=0}^{\left\lfloor \frac{q-k-1}{r}\right\rfloor} e^{2\pi i\frac{c(br+k)}{q}}\left|c\right\rangle \otimes \left|x^{k} \operatorname{mod} n\right\rangle .(**)$$

Now comes the surprise. If we measure the first register (tensor factor) then with probability at least $\frac{A}{\log\log r}$(this is the estimate that Shor gets) for some constant $A > 0$ the value of c to which the first register collapses will satisfy

$$\left|\frac{d}{r}-\frac{c}{q}\right| < \frac{1}{2q}$$

for some d relatively prime to r. Before we prove this result we will show how Shor uses it to complete the argument. We have assumed that $q \geq n^2$. Since $n > r$ we see that

$$\left|\frac{d}{r}-\frac{c}{q}\right| < \frac{1}{2r^2}.$$

This implies that $\frac{d}{r}$ appears in the sequence of continued fraction approximations (usaully called a convergent) to $\frac{c}{q}$. The continued fraction approximations can be calculated efficiently on a classical computer (we will see in the appendix to this subsection that the complexity is the same as the Euclidean algorithm for computing the greatest common divisor of c and q). This implies that we can compute r classically if we know c as above. We also note that since $r < n$ if we repeat the method a multiple of $\log n$ times we have found with probability close to 1 an appropriate c. Thus we have found r with probability $1-\varepsilon$ at the cost of $C(\varepsilon)\ \alpha \times \beta \times \gamma \times \delta$ with

$\alpha =$ the cost of the initial state (*).

$\beta =$ the cost of a quantum Fourier transform of order 2^m with $m = \lfloor 2\log n\rfloor$.

$\gamma =$ the cost of a partial fraction expansion of a fraction with denominator 2^m.

$\delta =$ a constant times $\log n$ (the number of tests necessary).

We are left with the probability of the appropriate measurement of c in (**). Set $s_k = \left\lfloor \frac{q-k-1}{r} \right\rfloor$. The measurement of c has probability

$$\frac{1}{q^2}\sum_{k=0}^{r}\left| e^{2\pi i \frac{ck}{q}} \sum_{b=0}^{s_k} e^{2\pi i \frac{brc}{q}} \right|^2 = \frac{1}{q^2}\sum_{k=0}^{r}\left|\sum_{b=0}^{s_k} e^{2\pi i b \frac{rc}{q}}\right|^2 .$$

We note that the inner sum is

$$\sum_{b=0}^{s_k} e^{2\pi i b\frac{rc}{q}} = \frac{1-e^{2\pi i (s_k+1)\frac{rc}{q}}}{1-e^{2\pi i \frac{rc}{q}}} = \frac{e^{\pi i (s_k+1)\frac{rc}{q}}}{e^{\pi i \frac{rc}{q}}} \frac{\sin(\pi(s_k+1)\frac{rc}{q}}{\sin(\pi\frac{rc}{q})}.$$

Hence the probability is

$$\frac{1}{q^2}\sum_{k=0}^{r}\left(\frac{\sin(\pi(s_k+1)\frac{rc}{q}}{\sin(\pi\frac{rc}{q})}\right)^2 (***)$$

Let $-\frac{1}{2}q \le [x]_q < \frac{1}{2}q$ be such that $x \equiv [x]_q \bmod q$. We only consider values of c such that

$$|\, [rc]_q \,| \le \frac{1}{2}r.$$

Under this condition

$$\left|(s_k+1)\frac{[rc]_q}{q}\right| < \frac{1}{2}\frac{r}{q}(\frac{q-k-1}{r}+1) < \frac{1}{2}(1+\frac{1}{\sqrt{q}}).$$

So if $m > 6$ then

$$\left|(s_k+1)\frac{[rc]_q}{q}\right| < 0.51.$$

We now observe that if $0 < t < 0.51\pi$ then

$$\frac{\sin t}{t} > \frac{0.49}{\pi}.$$

To see this observe that $\frac{d}{dt}\frac{\sin t}{t} \le 0$ for $0 < t < \pi$. This implies that

$$\left|\frac{\sin(\pi(s_k+1)\frac{rc}{q}}{\sin(\pi\frac{rc}{q})}\right| \ge \left|\frac{\sin(\pi(s_k+1)\frac{rc}{q}}{\pi\frac{rc}{q}}\right| = (s_k+1)\frac{0.51}{\pi}.$$

Now

$$s_k + 1 \ge \frac{q-k-1}{r} = \frac{q}{r} - \frac{k+1}{r} \ge \frac{q}{r} - 1.$$

We now estimate the the sum (***)

$$\frac{1}{q^2}\sum_{k=0}^{r}\left(\frac{\sin(\pi(s_k+1)\frac{rc}{q}}{\sin(\pi\frac{rc}{q})}\right)^2 \ge \frac{r}{q^2}\left(\frac{0.51}{\pi}\right)^2\left(\frac{q}{r}-1\right)^2 =$$

$$r\left(\frac{0.49}{\pi}\right)^2\left(\frac{1}{r}-\frac{1}{q}\right)^2$$

thus if n is large we can absorb the $\frac{1}{q}$ in the constant and have the estimate

$$\left(\frac{0.48}{\pi}\right)^2\frac{1}{r}.$$

We also note that if $|\left[rc\right]_q| < \frac{1}{2}r$ then there exists d such that

$$|rc-dq| < \frac{r}{2}$$

and dividing both sides by rq we have

$$\left|\frac{c}{q}-\frac{d}{r}\right| < \frac{1}{2q}.$$

We note that this argument is reversable. So it is equivalent to

$$|\left[rc\right]_q| \leq \frac{1}{2}r.$$

Now if $0 < d < r$ then we can find $0 < c < r$ such that $\left|\frac{c}{q}-\frac{d}{r}\right| < \frac{1}{2q}$. Let ϕ be the number of $0 < d < r$ such that d and r are relatively prime (Euler's Totient function). Then we have shown that the probability of measureing c such that there exists d relatively prime to r with $\left|\frac{c}{q}-\frac{d}{r}\right| < \frac{1}{2q}$ is at least a constant times $\frac{\phi(r)}{r}$. Now there are various lower bounds to $\frac{\phi(r)}{r}$ we will explain the one that Shor uses in the second appendix to this subsection. The first will prove the needed property of continued fractions.

3.2.1 Appendix1. Continued fractions

In this appendix we prove the characterization of a convergent of the continued fraction that Shor uses. The reason we give details is that the characterization is usually given for irrational numbers. Our argument involves a first step that allows the more standard method to work for rational numbers. We also begin the dsicussion with the observation that for rational numbers continued fractions are calculated using the Euclidean algorithm. If $q_{0,}q_1, q, ...$ are positive real numbers then we define the symbol $[q_0, q_1, q_2, ..., q_m]$ recursively by

$$[q_0] = q_0$$

and

$$[q_0,q_1, ..., q_{m+1}] = q_0 + \frac{1}{[q_1, ..., q_{m+1}]}.$$

Thus

$$[q_0, q_1, q_2, q_3, q_4] = q_0 + \frac{1}{q_1 + \frac{1}{q_2 + \frac{1}{q_3 + \frac{1}{q_4}}}}.$$

Suppose that $0 < a < b$ are integers we consider the Euclidean algorithm to calculate $\gcd(a, b)$.

$$b = aq_1 + r_1, 0 \le r_1 < a, r_1 \in \mathbb{Z}_{\ge}0.$$

If $r_1 = 0$ then $a = \gcd(a, b)$ ortherwise $r_1 > 0$ so

$$a = r_1 q_2 + r_2, 0 \le r_2 < r_1, q_2 \in \mathbb{Z}_{>0}, r_2 \in \mathbb{Z}_{\ge 0}.$$

If $r_2 = 0$ then $r_1 = \gcd(a, b)$ and otherwise

$$r_1 = r_2 q_3 + r_3.0 \le r_3 < r_2, q_3 \in \mathbb{Z}_{>0}, r_3 \in \mathbb{Z}_{\ge 0}.$$

Euclid's argorithm eventually stops at

$$r_{k-1} = r_k q_k$$

and according to the algrithm $r_k = \gcd(a, b)$. The amazing part of this algorithm is that

$$[0, q_1, ..., q_k] = \frac{a}{b}$$

in lowest terms.

For example $a = 25, b = 90$ then the algorithm yields:

$$q_1 = 3, r_1 = 15; q_2 = 1, r_2 = 10; q_3 = 1, r_3 = 5; q_4 = 2; r_4 = 0.$$

So $\gcd(25, 90) = 5$ and a direct calculation shows that

$$[0, 3, 1, 1, 2] = \frac{5}{18}.$$

Thus the Euclidean algorithm gives an algorithm that calculates for $0 < a < b$ integers, $\frac{a}{b}$ in lowest terms in the form

$$[0, q_1, q_2, ..., q_k]$$

with $q_k \in \mathbb{Z}_{>0}$. This is the continued fractions decomposition of $\frac{a}{b}$. One can check that

$$[q_0, q_1, q_2, ..., q_k, 1] = [q_0, q_1, q_2, ..., q_k + 1].$$

It turns out that this is the only ambiguity in expressing a rational number as a continued fraction. Thus we will use the outcome of the Euclidean algorithm as "the" continued fraction expansion. For $1 < r \le k$ terms

$$[0, q_1, ..., q_r]$$

are called the convergents of the continued fraction expansion of $\frac{a}{b}$.

Fix $q_1, q_2, ..., q_m, ...$ We define $n_0 = 0, d_0 = 1, n_1 = 1, d_1 = q_1$ then we assert that if

$$[0, q_1, q_2, ..., q_r] = \frac{n_r}{d_r}$$

then

$$[0, q_1, q_2, ..., q_r, q_{r+1}] = \frac{q_{r+1}n_r + n_{r-1}}{q_{r+1}d_r + d_{r-1}}$$

for $r \geq 2$. The numbers $\frac{n_r}{d_r}$ are called the convergents of $[q_0, q_1, q_2, ...]$. In fact we have the double recurrence

$$\begin{array}{c} n_{r+1} = q_{r+1}n_r + n_{r-1} \\ d_{r+1} = q_{r+1}d_r + d_{r-1} \end{array}.$$

Using this it is easily seen that

$$n_{r+1}d_r - n_r d_{r+1} = (-1)^r.$$

One can also check that the even convergents are increasing and the odd convergents are decreasing. Furthermore every odd convergent is larger than every even one. This implies in particular that if $\alpha = \frac{a}{b}$ and the convergents are calculated as above then

$$(d_k\alpha - n_k)(d_{k+1}\alpha - n_{k+1}) < 0.$$

Lemma 1 *If* $\gcd(x, N) = 1$ *and* p, q *are natural numbers and*

$$\left|\frac{p}{q} - \frac{x}{N}\right| < \frac{1}{2q^2}$$

then $\frac{p}{q} = \frac{n_k}{d_k}$ *a convergent of the continued fraction decomposition of* $\alpha = \frac{x}{N}$.

Proof. Suppose that $q \geq N$. Then since

$$\left|\frac{pN - xq}{qN}\right| < \frac{1}{2q^2}$$

we have

$$|pN - xq| < \frac{N}{2q} \leq \frac{1}{2}.$$

Since the left hand side of this equation is an integer we see that $\frac{p}{q} = \frac{q}{N}$. So we may assume that $q < N$. Now the standard argument applies which we recall. Let the convergents of $\frac{x}{N}$ be given as $\frac{n_l}{d_l}$. We note that $x = n_r, N = d_r$ for some r. We note that $0 < d_1 < d_2 < ... < d_r$. So there exists $k \leq r - 1$ such that

$$d_k \leq q < d_{k+1}.$$

We assume that $\frac{p}{q} \neq \frac{n_k}{d_k}$ and derive a contradiction.

We assert that $q < d_{k+1}$ implies

$$|q\alpha - p| \geq |d_k\alpha - n_k|. \qquad (*)$$

Assuming this we complete the proof. We have

$$d_k\left|\alpha - \frac{n_k}{d_k}\right| \leq q\left|\alpha - \frac{p}{q}\right| < \frac{1}{2q}.$$

Hence

$$\left|\alpha - \frac{n_k}{d_k}\right| < \frac{1}{2qd_k}.$$

$$\left|\frac{p}{q} - \frac{n_k}{d_k}\right| = \frac{|d_k p - n_k q|}{qd_k} \geq \frac{1}{qd_k}.$$

Thus

$$\frac{1}{qd_k} \leq \left|\frac{p}{q} - \frac{n_k}{d_k}\right| = \left|\frac{p}{q} - \alpha + \alpha - \frac{n_k}{d_k}\right| \leq$$

$$\left|\frac{p}{q} - \alpha\right| + \left|\alpha - \frac{n_k}{d_k}\right| < \frac{1}{2q^2} + \frac{1}{2d_k q}.$$

This implies that

$$\frac{1}{2d_k q} < \frac{1}{2q^2}$$

which says that $d_k > q$ which is a contradiction.

We are left with the proof of $(*)$. We note that

$$\det\begin{bmatrix} n_k & n_{k+1} \\ d_k & d_{k+1} \end{bmatrix} = (-1)^{k+1}.$$

So the inverse to $\begin{bmatrix} n_k & n_{k+1} \\ d_k & d_{k+1} \end{bmatrix}$ is $(-1)^{k+1}\begin{bmatrix} d_{k+1} & -n_{k+1} \\ -d_k & n_k \end{bmatrix}$. Thus if

$$\begin{gathered} n_k u + n_{k+1} v = p, \\ d_k u + d_{k+1} v = q \end{gathered}$$

then

$$\begin{gathered} (-1)^{k+1}(d_{k+1} p - n_{k+1} q) = u, \\ (-1)^{k+1}(-d_k p + n_k q) = v. \end{gathered}$$

Assert that we may assume $uv \neq 0$. If $u = 0$ then $d_{k+1}p = n_{k+1}q$.This implies d_{k+1} divides q since $\gcd(n_{k+1}, d_{k+1}) = 1$.But this contradicts the assumption $q < d_{k+1}$. If $v = 0$ we have $n_k u = p$ and $d_k u = q$ so

$$|q\alpha - p| = |uq_k\alpha - up_k| = |u|\,|d_k\alpha - n_k| \geq |d_k\alpha - n_k|\,.$$

Thus we may assume $uv \neq 0$. We assert that $uv < 0$. Indeed, if $uv > 0$ then, since $q = d_k u + d_{k+1} v$, we would have $q > d_{k+1}$ which is contrary to our hypthesis. We have already observed that

$$(d_k\alpha - n_k)(d_{k+1}\alpha - n_{k+1}) < 0.$$

Now $q\alpha - p = (d_k u + d_{k+1} v)\alpha - (n_k u + n_{k+1} v) = u(d_k\alpha - n_k) + v(d_{k+1}\alpha - n_{k+1})$. This is a sum of terms that are of the same sign so

$$q\alpha - p = \pm\left(|u(d_k\alpha - n_k)| + |v(d_{k+1}\alpha - n_{k+1})|\right)$$

Thus

$$|q\alpha - p| = |u|\,|(d_k\alpha - n_k)| + |v|\,|(d_{k+1}\alpha - n_{k+1})| \geq |(d_k\alpha - n_k)|$$

since $u \in \mathbb{Z}$. ■

3.2.2 Appendix 2: Euler's Totient function

If n is a positive integer then we set $\phi(n)$ equal to the number of integers $0 < r < n$ with $\gcd(r,n) = 1$. Clearly if n is prime then $\phi(n) = n - 1$ and if $n = p^k$ with p a prime $\phi(n) = p^{k-1}(p-1) = n(1 - \frac{1}{p})$. One can also show that if $\gcd(m,n) = 1$ then $\phi(mn) = \phi(m)\phi(n)$. So

$$\phi(n) = n\prod_{p|n}\left(1 - \frac{1}{p}\right).$$

Proposition 2 *There exists a constant $C > 0$ such that $\phi(n) \geq C\frac{n}{\log\log(n)}$ if (say) $n > 3$.*

The argument basically the method of an exercise in [R]. Since the exercise is only scetched out in [R] we give it a bit more detail here. We note that

$$\log\left(\frac{\phi(n)}{n}\right) = \sum_{p|n}\log(1 - \frac{1}{p}).$$

Since $\log(1-t) = -\sum_{m=1}^{\infty}\frac{t^m}{m}$ for $|t| < 1$we have $\log(1-t) + t = -t^2\sum_{m=1}^{\infty}\frac{t^m}{m+1}$ so if $0 < t < 1$

$$\log(1-t) = -t - t^2\sum_{m=0}^{\infty}\frac{t^m}{m+2}$$

This implies that if $0 < t \leq \frac{1}{2}$ then $\log(1-t) = -t + O(t^2)$ thus

$$\sum_{p|n}\log(1 - \frac{1}{p}) > -\sum_{p|n}\frac{1}{p} + C_1$$

with C_1 a constant. We now break up the sum as follows

$$\sum_{p|n}\frac{1}{p} = \sum_{\substack{p|n \\ p \leq \log n}}\frac{1}{p} + \sum_{\substack{p|n \\ p > \log n}}\frac{1}{p}$$

we note that the second sum is bounded by a fixed positive constant C_2. Theorem 2.3 in chapter 12 od [R] implies that

$$\sum_{p\leq x}\frac{1}{p} < \log\log x + C_3$$

with C_3 a positive constant. Thus

$$\sum_{p|n}\frac{1}{p} \leq \sum_{\substack{p|n \\ p\leq \log n}} \frac{1}{p} + C_2 <$$

$$\sum_{p\leq \log n}\frac{1}{p} + C_2 < \log\log\log n + C_2 + C_3.$$

So

$$\sum_{p|n}\log(1-\frac{1}{p}) > -\sum_{p|n}\frac{1}{p} + C_1 > -\log\log\log n + C_1 - C_2 - C_3$$

yielding

$$\frac{\phi(n)}{n} > \frac{C_5}{\log\log n}.$$

3.3 Reduction of factorization to period search

We will now describe the method Shor uses to reduce the problem of factorization to period search for which he had devised a fast quantum algorithm. Consider an integer N. for which we want to find a nontrivial factor. We may assume that it is odd and composite. Chose a number $1 < y < N-1$ randomly. If the greatest common divisor (gcd) of N and y is not one then we are done. We can therefore assume that $\gcd(y,N) = 1$. Hence y is invertible as an element of $\mathbb{Z}_N$ (under multiplication). Consider $f(m) = y^m \bmod N$. Then since the group of invertible elements of the ring $\mathbb{Z}_N$ is a finite group the function will have a minimum period. We can thus use Shor's algorithm to find the period, T. If T is even we assert that $y^{\frac{T}{2}}+1$ and N have a common factor larger than 1. We can thus use the Euclidean algorithm (which is easy classically) to find a factor of N. Before we demonstrate that this works consider $N = 55$ and $y = 3$. Then the smallest period is $T = 20$, , so $f(20) = 1 = f(0)$.

$$\gcd(3^{10}+1, 55) = 5.$$

We will now prove the assertion about the greatest common divisor. We first note that

$$(y^{\frac{T}{2}}+1)^2 = y^T + 2y^{\frac{T}{2}} + 1.$$

But $y^T = 1 + m\cdot N$ by the definition of T. Thus $(y^{\frac{T}{2}}+1)^2 \equiv 2(y^{\frac{T}{2}}+1) \bmod N$. Hence

$$(y^{\frac{T}{2}}+1)^2 - 2(y^{\frac{T}{2}}+1)$$

is evenly divisible by N. We therefore see that

$$\left((y^{\frac{T}{2}}+1)-2\right)\left(y^{\frac{T}{2}}+1\right)=\left(y^{\frac{T}{2}}-1\right)\left(y^{\frac{T}{2}}+1\right)$$

is evenly divisible by N. Hence, if $y^{\frac{T}{2}}+1$ and N have no common factor then $y^{\frac{T}{2}}-1$ is evenly divisible by N. This would imply that $\frac{T}{2}$ (which is smaller than T) satisfies

$$f(x+\frac{T}{2})=f(x).$$

This contradicts the choice of T as the minimal period. This is still not enough to get a non-trivial divisor of N. We must still show that y can be chosen so that N doesn't divide $y^{\frac{T}{2}}+1$ and that we can choose y so that T is even. Neither can be done with certainty. What can be proved is that if N is not a pure prime power then the probability of choosing $1<y<N-1$ such that $\gcd(y,N)=1$,y has even period and N doesn't divide $y^{\frac{T}{2}}+1$ is at least $\frac{3}{4}$. The proof of this is will take us too far afield a good reference is [NC]. We note that classically the test whether a number, N, is a pure power of a number $a>2$ and if so to calculate the number is polynomial in the number of bits of N. The upshot is that a quantum computer will factor a number with very large probability (if the algorithm is done say 10 times then the probability of success would be 0.999999 in polynomial time).

3.4 Error correction

So far we have ignored several of the difficulties that we had indicated in section 1 having to do with two problems that are caused by the environment. The first is that we can only really look at mixed states since we cannot compute the actual action of the environment and the second is the decoherence caused by the dynamics of the total system. We will assume that our quantum computations are divided into steps that take so little time that our initial pure states remain close enough to being pure states that we can ignore the first difficulty. For the second we will look at the decoherence over this small period as a small error. For most of the systems that are proposed the most likely error is a one qubit error. Thus as in classical error correction we will show how to set up a quantum error correcting code that corrects a one qubit error. The standard procedure is to encode a qubit as an element of a two dimensional subspace of a higher qubit space.

That is we take V to be the space of n qubits and we take u_0 and u_1 orthonormal in V and assign

$$a\left|0\right\rangle+b\left|1\right\rangle\mapsto au_0+bu_1.$$

The right hand side will be called the *encoded* qubit. The question is what is the most likely error if we transmit the encoded qubit? The generally accepted answer is that it would be a transformation of the form

$$E=I\bigotimes\cdots\bigotimes A\bigotimes\cdots\bigotimes I$$

with all factors the identity except for an A in the k-th factor and this A is a fairly arbitrary linear map on 1-qubit space that is close to the identity. The problem is to fix the error which means change $E\left(au_0 + bu_1\right)$ to $au_0 + bu_1$ without knowing which qubit has an error, what the error is and not collapsing the wave function of the unknown qubit. Classically one can transmit one bit in terms of 3 bits. $0 \mapsto 000, 1 \mapsto 111$. The most likely error is a NOT in one bit. To fix such an error one reads the sum of the entries of this possibly erroneous output and if it is at most 1 then change it to 000 if it is at least 2 change it to 111. This will correct exactly one NOT in any position. Quantum mechanically we must correct a continuum of possible errors. This seems to be impossible and if it were impossible then quantum computation looked impossible also since decoherence would set in before we could do any useful computation. As usual, Shor [S3] found a method. We will describe a later development that yielded a quantum analog of a perfect code (such as the three bit classical error correction scheme described above).

We will describe a special class of error correcting codes that are known as *orthogonal codes* (or non-degenerate codes). In fact, Shor's original example was not an orthogonal code, but we feel that these codes are easier for mathematicians to understand. We will need some additional notation.

If $X, Y \in M_2(\mathbb{C})$ then we define $\langle X|Y\rangle = \frac{1}{2}\mathrm{tr}(X^*Y)$ ($X^* = X^\dagger$ to a physicist) is the Hermitian adjoint of X). Given $j = 1, ..., n$ we define $F_j : M_2(\mathbb{C}) \to \mathrm{End}\left(\bigotimes^n \mathbb{C}^2\right)$ by

$$F_j(A) = I \bigotimes \cdots \bigotimes A \bigotimes \cdots \bigotimes I$$

where all of the factors on the right hand side are I except for the k-th term which is A. We say that an isometry, $T : \mathbb{C}^2 \to \bigotimes^n \mathbb{C}^2$ defines an *orthogonal code space* if it has the following properties:

1. The maps $T_j : M_2(\mathbb{C}) \bigotimes \mathbb{C}^2 \to \bigotimes^n \mathbb{C}^2$ given by $T_j(X \bigotimes v) = F_j(X)T(v)$ are isometries (onto their images) for $i = 1, ..., n$.

2. If $V = \{X \in M_2(\mathbb{C}) | tr(X) = 0\}$. Then the sum

$$Z = T(\mathbb{C}^2) \bigoplus \bigoplus_{1 \leq j \leq n} T_j(V \bigotimes \mathbb{C}^2)$$

is an orthogonal direct sum.

We will now show how to correct a one qubit error if we have an orthogonal code. Let X_1, X_2, X_3 be an orthonormal basis of V consisting of invertible elements. For example we could choose the Pauli matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

We write $\bigotimes^n \mathbb{C}^2 = Z \bigoplus Z'$ with Z' the orthogonal complement to Z. Let A be an observable that acts by distinct scalars as indicated $\lambda_0 I$ on $T(\mathbb{C}^2), \lambda_{ij} I$ on $T_j(X_i \bigotimes \mathbb{C}^2), 1 \leq j \leq n, 1 \leq i \leq 3$ and μI on $Z^{'}$. If we start with $T(v)$ and it has incurred an error and we have w rather than $T(v)$ then we do a measurement of A on w. If the measurement is μ then with high probability the error wasn't a one qubit error. Otherwise we assume a one qubit error then there is j such that $w = T_j(X \bigotimes v)$. $X = aI + bX_1 + cX_2 + dX_2$. Thus with probability 1 the eigenvalue will be one of $\lambda_0, \lambda_{1j}, \lambda_{2j}, \lambda_{3j}$. If it is λ_0 then w will have collapsed to v. If it is λ_{ij} then if w collapses to z then $F_j(X_i^{-1})z = T(v)$ we have thus corrected the error.

Obviously, to use this idea we must have a way of finding T. We note first of all that $\dim Z \leq 2^n$ and $\dim Z = 6n + 2$. If $6n + 2 \leq 2^n$ then $n \geq 5$ and if $n = 5$ then $2^5 = 6 \cdot 5 + 2$. Thus the smallest n that we could use would be $n = 5$. We will now give conditions on a map T that are equivalent to having an orthogonal code. If $w \in \bigotimes^n \mathbb{C}^2$ then $w = \sum w_j \left| j \right\rangle$. Let $0 \leq p < q < n$ be two bit positions. Then we form a $4 \times 2^{n-2}$ matrix as follows. The $i = i_0 + i_1 2$, $j = j_0 + j_1 2 + ... + j_{n-3} 2^{n-3}$ entry is w_k where

$$k = j_0 + ... j_{p-1} 2^{p-1} + i_0 2^p + j_p 2^{p+1} + ... + j_{q-2} 2^{q-1} + i_1 2^q + j_{q-1} 2^{q+1} + ... + j_{n-3} 2^{n-1}.$$

If $p = 0, q = 1$ this is just $k = i_0 + i_1 2 + 2^2(j_0 + j_1 2 + ... + j_{n-3} 2^{n-3})$. Let $W(p, q, w)$ denote this matrix. We have

Theorem 3 *$T : \mathbb{C}^2 \to \bigotimes^n \mathbb{C}^2$ defines an orthogonal code if and only if*

$$W(p, q, T \left| i \right\rangle) W(p, q, T \left| j \right\rangle)^* = \frac{1}{4} \delta_{i,j} I.$$

For all p, q and $i, j \in \{0, 1\}$.

This can be written as a system of quadratic equations. If we put them into Mathematica for $n = 5$ the first solution is given as follows: We define

$$\left\langle j_1 j_2 j_3 j_4 j_5 \right\rangle = \left| j_1 j_2 j_3 j_4 j_5 \right\rangle + \left| j_5 j_1 j_2 j_3 j_4 \right\rangle + \left| j_4 j_5 j_1 j_2 j_3 \right\rangle + \left| j_3 j_4 j_5 j_1 j_2 \right\rangle + \left| j_2 j_3 j_4 j_5 j_1 \right\rangle .$$

Then set

$$T \left| 0 \right\rangle = \frac{1}{4} \left(\left| 0000 \right\rangle + \left\langle 11000 \right\rangle - \left\langle 10100 \right\rangle - \left\langle 11110 \right\rangle \right)$$

and

$$T \left| 1 \right\rangle = \frac{1}{4} \left(\left| 11111 \right\rangle + \left\langle 00111 \right\rangle - \left\langle 01011 \right\rangle - \left\langle 00001 \right\rangle \right).$$

Because of the symmetry it is easy to check that the condition of the theorem is satisfied. This code was originally found by other methods (c.f. [KL]).

References

[Co] D. Coppersmith, An approximate Fourier Transform Useful in Quantum Factoring,arXiv:quant-ph/0201067v1.

[NC] Michael Nielson and Isaac Chang, Quantum Computation and Quantum Information, Cambridge University Press, Cambridge, 2000.

[S3] P. Shor, A scheme for reducing decoherence in quantum computer memory, Phys. Rev. A,52,1995.

[KL] E. Kroll and R. Laflamme, A theory of quantum error correcting codes, Phys. Rev. A, 50:900-911,1997.

[R] H. E. Rose, A Course in Number Theory, Second Edition, Oxford Science Publications,Oxford, 1994.

4 Entanglement

As we have seen the only non-trivial reversible one bit operation is NOT which interchanges 0 and 1. We have made the simplifying assumption that all one qubit unitary operators are easily implementable on a quantum computer. The operation NOT gives rise to the unitary operator in one qubit with matrix

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

relative to the computational basis $|0\rangle, |1\rangle$. We will say that a transformation of n bits that is given by applying either NOT or the identity to each bit is a *classical local transformation.* A *quantum local transformation* in n qubits is a unitary operator of the form

$$A_1 \bigotimes A_2 \bigotimes \cdots \bigotimes A_n$$

where $A_i \in U(2)$. There is a major distinction between the classical and the quantum cases. The classical local transformations act transitively on the set of all n bit bit strings. Whereas the quantum local transformations act transitively only in the case when $n = 1$. For example there is no local transformation that takes the state

$$\frac{|00\rangle + |11\rangle}{\sqrt{2}}$$

to $|00\rangle$ (see the next section for a proof). We will call a state that is not a product state (not in the orbit of $|00\rangle$ under local transformations) an *entangled* state. The two code words of the five bit error correcting code are entangled. Furthermore, entanglement explains some of the apparent paradoxes that appeared in the early thought experiments of quantum mechanics. It is also basic to quantum teleportation (a subject that we will not be covering in these sections). In this section we will study the orbit structure of the local transformations on the pure states and in particular functions that help to separate these orbits: the measures of entanglement. We will emphasize methods that allow one to determine if two states are related by a local transformation and to determine the extent of the entanglement of a state.

4.1 Measures of entanglement

We will first look at the example of an entangled state: $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$. One way that one can see that it is entangled is by observing that if we act on $\mathbb{C}^2 \bigotimes \mathbb{C}^2$ by $G = SL(2,\mathbb{C})\times SL(2,\mathbb{C})$ by the tensor product action. Then G leaves invariant a symmetric form, the tensor product of the symplectic forms on each of the $\mathbb{C}^2$ factors that are $SL(2,\mathbb{C})$ invariant. This form, $(\ , \ )$, is given by

$$(|00\rangle, |11\rangle) = (|11\rangle, |00\rangle = 1,$$
$$(|01\rangle, |10\rangle) = (|10\rangle, |01\rangle) = -1$$

and all the other products are 0. We note that $(\frac{|00\rangle+|11\rangle}{\sqrt{2}},\frac{|00\rangle+|11\rangle}{\sqrt{2}})=1$ and $(|00\rangle,|00\rangle)=0$ so there can't be a local transformation taking one to another since the function $\phi(u)=|(u,u)|$ is invariant under local transformations. It is an example of a *measure of entanglement*. Indeed, one can prove that a state,u, in 2 qubits is entangled if and only if $\phi(u)>0$. Another property enjoyed by this function is that for all pure 2 qubit states $\phi(u)\leq 1$ and $\phi(u)=1$ if and only if u is in the orbit of $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$ under local transformations. To prove the upper bound we consider $u=a\,|00\rangle+b\,|01\rangle+c\,|10\rangle+d\,|11\rangle$. Then $\phi(u)=2(ad-bc)$. Since $2|a||d|\leq|a|^2+|d|^2$ we have $|\phi(u)|\leq|a|^2+|b|^2+|c|^2+|d|^2=1$. Although it is not hard to prove the assertion about the orbit directly we will use a result of Kempf and Ness [KN] which is useful in other contexts. For those of you who are unfamiliar with semisimple Lie groups take G to be the product of n copies of $SL(2,\mathbb{C})$ and K to be n copies of $SU(2)$.

Theorem 4 *Let G be a semisimple Lie group over $\mathbb{C}$ and let K be a maximal compact subgroup of G. Let (π,V) be a finite dimensional holomorphic representation of G with the K-invariant Hilbert space structure $\langle\ |\ \rangle$. If $v\in V$ and $m=\inf\{\langle\pi(g)v|\pi(g)v\rangle\,|g\in G\}$. Then if $u\in\pi(G)v$ and $\langle u|u\rangle=m$ then $\pi(K)u=\{w\in\pi(G)v|\,\langle w|w\rangle=m\}$. Furthermore the infimum is actually attained if and only if the orbit $\pi(G)v$ is closed.*

In words this says that the elements of minimal norm in a G orbit form a single K-orbit.

We will now give an idea of the proof. We note that $Lie(G)=Lie(K)+iLie(K)$. We therefore have

$$G=K\exp(iLie(K)).$$

If $X\in iLie(K)$ then $d\pi(X)^*=d\pi(X)$. Thus

$$\begin{aligned}&\frac{d^2}{dt^2}\langle\pi(\exp tX)v|\pi(\exp tX)v\rangle\\&=4\,\langle d\pi(X)\pi(\exp tX)v|d\pi(X)\pi(\exp tX)v\rangle\geq 0\end{aligned}$$

With equality if and only if $d\pi(X)v=0$. Everything follows from this.

We will now show how the Kempf-Ness result applies to our situation for 2 qubits. We first note that relative to $G=SL(2,\mathbb{C})\times SL(2,\mathbb{C})$ the space $V=\mathbb{C}^2\bigotimes\mathbb{C}^2$ has the following orbit structure. For each $\lambda\in\mathbb{C}-\{0\}$ the set $M_\lambda=\{w\in V|(v,v)=\lambda\}$ is a single orbit. The other orbits are $\pi(G)\,|00\rangle$ and $\{0\}$. The union of the latter two is M_0. We set $u_0=\frac{|00\rangle+|11\rangle}{\sqrt{2}}$. We note that we have $M_\lambda=z\pi(G)u_0$ with $z^2=\lambda$. We therefore see that the elements in the unit sphere that maximize ϕ are contained in the set of elements the form $w=e^{i\theta}\frac{\pi(g)u_0}{\|\pi(g)u_0\|}$, $g\in G$. For such a w we have $\phi(w)=\frac{1}{\|\pi(g)u_0\|^2}$. Thus maximizing ϕ on the unit sphere means (up to phase) minimizing the norm on $\pi(G)u_0$. The Kempf-Ness theorem implies that this subset of $\pi(G)u_0$ is $\pi(K)u_0$. This completes the proof of the assertion.

We note that the group of local transformations on $\bigotimes^n\mathbb{C}^2$ is the image of $S^1 \times SU(2)^n$ with S^1 the circle group acting by scaler multiplication and $SU(2)^n$ acting by the tensor product action (i.e. by local transformations). We will therefore concentrate on invariants for $SU(2)^n$. We also note that if we consider $\phi(u)^2$ rather than $\phi(u)$ then it is a polynomial function on $\mathbb{C}^2 \bigotimes \mathbb{C}^2$ as a real vector space. We will only consider measures of entanglement that are polynomials invariant under $K = SU(2)^n$ on $\bigotimes^n\mathbb{C}^2$ as a real vector space. We will use the term measure of entanglement for such a polynomial. We will denote the algebra of such polynomials by $\mathcal{P}_{\mathbb{R}}(\bigotimes^n\mathbb{C}^2)^K$. These are exactly what we need to separate the K-orbits.

Theorem 5 *If $u, v \in \bigotimes^n\mathbb{C}^2$ then $u \in \pi(K)v$ if and only if $f(u) = f(v)$ for all $f \in \mathcal{P}_{\mathbb{R}}(\bigotimes^n\mathbb{C}^2)^K$.*

We also note that if we look at the action of the circle group by multiplication on $\bigotimes^n\mathbb{C}^2$ we can define a $\mathbb{Z}$-grading on $\mathcal{P}_{\mathbb{R}}(\bigotimes^n\mathbb{C}^2)^K$ by $f \in \mathcal{P}^j_{\mathbb{R}}(\bigotimes^n\mathbb{C}^2)^K$ if $f \in \mathcal{P}_{\mathbb{R}}(\bigotimes^n\mathbb{C}^2)^K$ and $f(zu) = z^j f(u)$ for all $z \in S^1$ and $u \in \bigotimes^n\mathbb{C}^2$.

Theorem 6 *If $u, v \in \bigotimes^n\mathbb{C}^2$ then $u \in S^1\pi(K)v$ if and only if $f(u) = f(v)$ for all $f \in \mathcal{P}^0_{\mathbb{R}}(\bigotimes^n\mathbb{C}^2)^K$.*

Both of these theorems are consequences of the following result.

Theorem 7 *Let U be a compact Lie group. Let (ρ, W) be a finite dimensional representation of U on a real Hilbert space. Let $\mathcal{P}(W)^U$ be the algebra of all complex valued polynomials on W that are invariant under U. If $u, v \in W$ then $u \in \rho(U)v$ if and only if $f(u) = f(v)$ for all $f \in \mathcal{P}(W)^U$.*

Proof. The necessity is obvious. Since $v \mapsto \|v\|^2$ is in $\mathcal{P}(W)^U$ we will prove that if $\|v\| = \|u\| = r > 0$ and $f(u) = f(v)$ for all $f \in \mathcal{P}(W)^U$ then $u \in \rho(U)v$. The Stone-Weierstrauss theorem implies that the restriction of $\mathcal{P}(W)$ to the sphere of radius r,S_r,is uniformly dense in the space of continuous functions on S_r. Suppose that $\rho(U)v \cap \rho(U)u$ is empty then Uryson's Lemma implies there is a continuous function φ on S_r such that $\varphi_{|\rho(U)v} \equiv 1$ and $\varphi_{|_{\rho(U)u}} \equiv 0$. The uniform density implies that there exists an $f \in \mathcal{P}(W)$ such that $|f(x) - \varphi(x)| < \frac{1}{4}$ for all $x \in S_r$. Let du denote normalized invariant measure on U. We define $\overline{f}(x) = \int\limits_U f(\rho(z)x)dz$. Then $\overline{f} \in \mathcal{P}(W)^U$. We have

$$|\overline{f}(v) - 1| = \left|\int\limits_U f(\rho(z)v)dz - 1\right| \leq \int\limits_U |f(\rho(z)v)dz - \varphi(\rho(z)v)|dz \leq \frac{1}{4}$$

hence $|\overline{f}(v)| > \frac{3}{4}$ similarly $|\overline{f}(u)| < \frac{1}{4}$. This proves the theorem. ■

4.2 Three qubits

These results make it reasonable to assert that the orbit of $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$under local transformations consists of the most entangled two qubit states. In the case of 3 qubits there is a similar result. First the ring of invariant (complex polynomials) on $\mathbb{C}^2 \bigotimes \mathbb{C}^2 \bigotimes \mathbb{C}^2$ under the tensor product action of

$$G = SL(2,\mathbb{C}) \times SL(2,\mathbb{C}) \times SL(2,\mathbb{C})$$

is generated by one element, f, of degree 4 (here we will be stating several results without proof in this case the details can be found in [GrW]). We can define it as follows: If $v \in \mathbb{C}^2 \bigotimes \mathbb{C}^2 \bigotimes \mathbb{C}^2$ the we can write it as

$$v = |0\rangle \bigotimes v_0 + |1\rangle \bigotimes v_1$$

with $v_0, v_1 \in \mathbb{C}^2 \bigotimes \mathbb{C}^2$. If we use the symmetric form defined above we have

$$f(v) = \det \begin{bmatrix} (v_0, v_0) & (v_0, v_1) \\ (v_1, v_0) & (v_1, v_1) \end{bmatrix}.$$

As in the case of two qubits most of the orbits under G are described by the values of f. Here we set $M_\lambda = \{v \in \mathbb{C}^2 \bigotimes \mathbb{C}^2 \bigotimes \mathbb{C}^2 | f(v) = \lambda\}$. Then if $\lambda \neq 0$ we have M_λ consists of a single orbit. If $\lambda = 0$ then there are 6 orbits in M_0, We note that $f(\frac{|000\rangle + |111\rangle}{\sqrt{2}}) = \frac{1}{4}$. Thus $M_\lambda = zG\left(\frac{|000\rangle + |111\rangle}{\sqrt{2}}\right)$ with $z^4 = 4\lambda$. We will now describe the orbits in M_0. First there is the open orbit in this quartic given as the orbit of

$$w_0 = \frac{|001\rangle + |010\rangle + |100\rangle}{\sqrt{3}}.$$

If we remove this orbit from M_0 then there are three open orbits in what remains. They are the orbits of $\frac{|000\rangle + |011\rangle}{\sqrt{2}}$, $\frac{|000\rangle + |101\rangle}{\sqrt{2}}$ and $\frac{|000\rangle + |110\rangle}{\sqrt{2}}$. If in addition these are removed then what we have left is the union of 0 and the product states (which form a single orbit).

One can show by an argument similar to that in two qubits that if u is a state then $|f(u)| \leq \frac{1}{4}$ and if $u_o = \frac{|000\rangle + |111\rangle}{\sqrt{2}}$ then $f(u_o) = \frac{1}{4}$. Since the set where f is non-zero is exactly the set of all elemets $\mathbb{C}^\times G\left(\frac{|000\rangle + |111\rangle}{\sqrt{2}}\right)$ we see that if u is a state with $|f(u)| \neq 0$ then $u = \frac{gu_o}{\|gu_o\|}$ with $g \in G$. Thus $f(u) = f\left(\frac{gu_o}{\|gu_o\|}\right) = \frac{1}{\|gu_o\|^4} f(gu_o) = \frac{1}{4\|gu_o\|^4}$. Thus the set of states with $|f(u)| = \frac{1}{4}$ are exactly the elements that minimize the value of $\|gu_o\|^4$ for $g \in G$. Thus Theorem 2 implies:

Proposition 8 *If $K = S^1 SU(2) \times SU(2) \times SU(2)$ then*

$$\{v \in \mathbb{C}^2 \bigotimes \mathbb{C}^2 \bigotimes \mathbb{C}^2 || f(v)| = \frac{1}{4}, \|v\| = 1\} = K\left(\frac{|000\rangle + |111\rangle}{\sqrt{2}}\right).$$

Thus one value of one invariant is enough to determine if a state can be gotten from $\frac{|000\rangle + |111\rangle}{\sqrt{2}}$ by local transformations. For example

$$v = \frac{|111\rangle + |001\rangle + |010\rangle + |100\rangle}{2}$$

has the property that $f(v) = \frac{1}{4}$. So it can be obtained by a local transformation from $\frac{|000\rangle + |111\rangle}{\sqrt{2}}$.

So far we have been analyzing only one polynomial measure of entanglement. There is the natural problem of determining a generating set for these measures. To do this it is useful to reduce the problem to a problem involving complex algebraic groups and complex polynomials. The basic idea is that if G is a simply connected semi-simple Lie group over $\mathbb{C}$ then G is a linear algebraic group. If K is a maximal compact subgroup of G and if (ρ, V) is a finite dimensional unitary representation of K then ρ extends to a regular representation of G on V. The real polynomials on V are the complex polynomials in both the bra and the ket vectors. The ket vectors give a copy of V as a complex vector space whereas the ket vectors give a copy of the complex dual representation of V. This implies that the algebra $\mathcal{P}_R(V)^K$ is naturally isomorphic with $\mathcal{P}(V \bigoplus V^*)^G$. In the case when we are dealing with qubits the representation of $G = SL(2)^n$ on $\bigotimes^n \mathbb{C}^2$ is self dual. We are thus looking at the problem of determining the invariants of G acting on two copies of $\bigotimes^n \mathbb{C}^2$ by the diagonal action. We analyze this problem for two and three qubits.

4.3 Measures of entanglement for two and three qubits

We first look at 2 qubits and continue the discussion begun in the previous subsection. As we have observed $G = SL(2,\mathbb{C}) \times SL(2,\mathbb{C})$ leaves invariant a symmetric bilinear form on $\mathbb{C}^2 \bigotimes \mathbb{C}^2$. A dimension count shows that the image of G on $\mathbb{C}^2 \bigotimes \mathbb{C}^2$ is the full orthogonal group for this form. Thus the action on $\mathbb{C}^2 \bigotimes \mathbb{C}^2$ can be interpreted as the action of $SO(4,\mathbb{C})$ on $\mathbb{C}^4$. We are thus looking at the invariants of $SO(4,\mathbb{C})$ on two copies of $\mathbb{C}^4$. Classical invariant theory implies that the algebra of invariants is generated by the three polynomials $\alpha(v \bigoplus w) = (v,v), \beta(v \bigoplus w) = (v,w)$ and $\gamma(v \bigoplus w) = (w,w)$, This implies

Lemma 9 *The algebra of measures of entanglement in* 2 *qubits is the set of polynomials in* $(v,v), \langle v|v\rangle$ *and* $\overline{(v,v)}$.

Thus in this case we were using the only "interesting" measure, since we are only considering states which are assumed to satisfy $\langle v|v\rangle = 1$.

The situation is different for three qubits. We will describe a set of generators in this case that was determined in [MW1] our method is a modification which is an outgrowth of joint work with H. Kraft. As above we look upon $\mathbb{C}^2 \bigotimes \mathbb{C}^2 \bigotimes \mathbb{C}^2$ as $\mathbb{C}^2 \bigotimes \mathbb{C}^4$ and $G = SL(2,\mathbb{C}) \times SL(2,\mathbb{C}) \times SL(2,\mathbb{C})$ acting as $SL(2,\mathbb{C}) \times SO(4,\mathbb{C})$. For the moment we will ignore the $SL(2,\mathbb{C})$ factor and look at $I \bigotimes SO(4,\mathbb{C})$ acting on two copies of $\mathbb{C}^2 \bigotimes \mathbb{C}^4$. If we consider only the action of $SO(4,\mathbb{C})$ then we are looking at its action on 4 copies of $\mathbb{C}^4$. We look at this as $SO(4,\mathbb{C})$ acting on $X \in M_4(\mathbb{C})$ under right multiplication by the transpose of the matrix. Then the invariants for $SO(4,\mathbb{C})$ are generated by the matrix entries of XX^T (the upper T stands for transpose) and $\det(X)$ (for these results and others stated without proof in this subsection please see

[GW]). We now look at the action of the remaining $SL(2,\mathbb{C})$. The $SL(2,\mathbb{C})$ is acting on the left on the matrix via multiplication by the block diagonal matrix

$$h = \left[\begin{array}{cc} g & 0 \\ 0 & g \end{array}\right].$$

Thus the $SL(2,\mathbb{C})$ factor is acting on the generators of the $SO(4,\mathbb{C})$ invariants trivially on $\gamma = \det X$ (an invariant under the full G of degree 4) and via hXX^Th^T with h as above, We write XX^T in block form

$$\left[\begin{array}{cc} A & B \\ B^T & C \end{array}\right]$$

then the $SL(2,\mathbb{C})$ is acting on the components via $A \mapsto gAg^T, B \mapsto gBg^T, C \mapsto gCg^T$. We note that A and C are symmetric and completely general and B is an arbitrary 2×2 matrix which we can write as

$$a\left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right] + Z$$

with Z a general two by two symmetric matrix. The coefficient a defines an invariant for G of degree 2 on the qubits which we will call α. The rest of the action is by three copies of the action of $SL(2,\mathbb{C})$ on the symmetric 2×2 matrices. Using the trace form we see that this is just the action of $SO(3,\mathbb{C})$ on three copies of $\mathbb{C}^3$. Again we look upon this as the action of $SO(3,\mathbb{C})$ on $Y = M_3(\mathbb{C})$ via left multiplication. The invariants in this case are generated by $\beta = \det Y$ (an invariant of degree 6 on the qubits) and the matrix coefficients of Y^TY which yield 6 invariants of degree 4. The upshot is the invariants are generated by an invariant of degree 2 (α), an invariant of degree 4 (γ), an invariant of degree 6 (β) and 6 invariants of degree 4 (the matrix coefficients of Y^TY), $\mu_1, ..., \mu_6$. We note that the invariants γ and β have the property that their squares are invariant under $O(3)\times O(4)$. Thus γ^2 and β^2 are in the algebra generated by α and $\mu_1, ..., \mu_6$. We can also see from the invariant theory of $SO(3)$ that the functions $\alpha, \mu_1, ..., \mu_6$ are algebraically independent. We therefore see that the full ring of invariants is $\mathbb{C}[\alpha, \mu_1, .., \mu_6] \bigoplus \mathbb{C}[\alpha, \mu_1, .., \mu_6]\beta \bigoplus \mathbb{C}[\alpha, \mu_1, .., \mu_6]\gamma \bigoplus \mathbb{C}[\alpha, \mu_1, .., \mu_6]\beta\gamma$.

References

[GrW] Benedict H. Gross and Nolan R. Wallach, On quaternionic discrete series and their continuations, J. Reine Angew. Math. 481 (1996),73-123.

[GW] Roe Goodman and Nolan R. Wallach, Representations and invariants of the classical groups, Cambridge University Press, Cambridge, 1998.

[MW1] David Meyer and Nolan Wallach, Invariants for multiple qubits: the case of 3 qubits. Mathematics of quantum computation,77–97, Comput. Math. Ser., Chapman & Hall/CRC, Boca Raton, FL, 2002.

[KN] George Kempf and Linda Ness, The length of vectors in representation spaces. Algebraic geometry (Proc. Summer Meeting, Univ. Copenhagen, Copenhagen, 1978), pp. 233–243, Lecture Notes in Math., 732, Springer, Berlin, 1979.

5 Four and more qubits

In the cases of 2 and 3 qubits it is fairly clear what the maximally entangled states should be or at least there are just a few candidates for that honor. We will see that there is an immense variety of states that are highly entangled in the case of 4 qubits. This and the calculation of Hilbert series for measures of entanglement (see subsection 5.2) indicate that the search for all measures of entanglement or the complete description of the orbit structure for arbitrary numbers of qubits will be so hard and complicated as to become useless. However the case of 4 qubits gives some indications of how to find more invariants. Also, methods similar to the Kempf-Ness theorem can be used to prove uniqueness theorems (for example the theorem of Rains [R] that implies that the 5 bit error correcting code we discussed earlier is unique up to local transformations).

As it turns out the orbit structure under

$$G = SL(2,\mathbb{C})\times SL(2,\mathbb{C})\times SL(2,\mathbb{C})\times SL(2,\mathbb{C}) \text{ on } \mathbb{C}^2 \bigotimes \mathbb{C}^2 \bigotimes \mathbb{C}^2 \bigotimes \mathbb{C}^2$$

can be determined using the results of Kostant and Rallis [KR]. Since it fits in their theory in case of the symmetric pair $(SO(4,4), SO(4)\times SO(4))$. We will now describe the outgrowth of this theory purely in terms of qubits.

5.1 Four qubits

We are therefore analyzing the action of $G = SL(2)\times SL(2)\times SL(2)\times SL(2)$ on the space $V = \mathbb{C}^2 \bigotimes \mathbb{C}^2 \bigotimes \mathbb{C}^2 \bigotimes \mathbb{C}^2$ via the tensor product action

$$(g_1, g_2, g_3, g_4)(v_1 \bigotimes v_2 \bigotimes v_3 \bigotimes v_4) = g_1v_1 \bigotimes g_2v_2 \bigotimes g_3v_3 \bigotimes g_4v_4$$

We first note that if $H = SL(2)\times SL(2)$ and if $W = \mathbb{C}^2 \bigotimes \mathbb{C}^2$ and if we have H act on W by the tensor product action then there is a H-invariant non-degenerate symmetric bilinear form, $(..., ...)$, on W given as follows

$$(v \bigotimes w, x \bigotimes y) = \omega(v,x)\omega(w,y).$$

Here $\omega((x_1,y_1),(x_2,y_2)) = x_1y_2 - x_2y_1$. This form allows us to define a linear map, T, of V onto $End(W)$ in the following way

$$T(v_1 \bigotimes v_2 \bigotimes v_3 \bigotimes v_4)(w_1 \bigotimes w_2) = \omega(v_3, w_1)\omega(v_4, w_2)v_1 \bigotimes v_2.$$

We look upon G as $H \times H$. Thus if $g = (h_1, h_2)$ then

$$T(gv)(w) = h_1T(v)(h_2^{-1}w).$$

If $A \in End(W)$ then we define $A^\#$ by $(Aw_1, w_2) = (w_1, A^\# w_2)$. We note that if $h \in H$ then $h^\# = h^{-1}$. This implies that

$$\begin{aligned} T(gv)T(gv)^\# &= h_1T(v)h_2^{-1}(h_1T(v)h_2^{-1})^\# \\ &= h_1T(v)h_2^{-1}h_2T(v)^\# h_1^{-1} = h_1T(v)T(v)^\# h_1^{-1}. \end{aligned}$$

We therefore have invariants $f_{2j}(v) = tr((T(v)T(v)^{\#})^j)$, $j = 1, 2, ...$ and $g_4(v) = \det(T(v))$.

Theorem 10 *The ring of invariants under the action of G on V is generated by the algebraically independent elements f_2, f_4, g_4, f_6.*

The following discussion gives a sketch of a proof.

We will use qubit notation for elements of V. Thus V has a basis consisting of elements $|i_0 i_1 i_2 i_3\rangle$ with $i_j = 0, 1$. We set

$$\begin{aligned} v_1 &= \frac{1}{2}(|0000\rangle + |1111\rangle + |0011\rangle + |1100\rangle), \\ v_2 &= \frac{1}{2}(|0000\rangle + |1111\rangle - |0011\rangle - |1100\rangle), \\ v_3 &= \frac{1}{2}(|1010\rangle + |0101\rangle + |0110\rangle + |1001\rangle), \\ v_4 &= \frac{1}{2}(|1010\rangle + |0101\rangle - |0110\rangle - |1001\rangle). \end{aligned}$$

These states can be described in terms of the Bell states for 2 qubits. Let $u_\pm = \frac{|00\rangle \pm |11\rangle}{\sqrt{2}}$ and $v_\pm = \frac{|01\rangle \pm |10\rangle}{\sqrt{2}}$ then

$$v_1 = u_+ \bigotimes u_+, v_2 = u_- \bigotimes u_-, v_3 = v_+ \bigotimes v_+, v_4 = v_- \bigotimes v_-.$$

We note that if $v = x_1 v_1 + x_2 v_2 + x_3 v_3 + x_4 v_4$ then

$$T(v) = \begin{bmatrix} \frac{x_1 - x_2}{2} & 0 & 0 & \frac{x_1 + x_2}{2} \\ 0 & \frac{x_4 - x_3}{2} & -\frac{x_3 + x_4}{2} & 0 \\ 0 & -\frac{x_3 + x_4}{2} & \frac{x_4 - x_3}{2} & 0 \\ \frac{x_1 + x_2}{2} & 0 & 0 & \frac{x_1 + x_2}{2} \end{bmatrix}.$$

Hence

$$f_{2j}(v) = \sum x_i^{2j}$$

and

$$g_4(v) = x_1 x_2 x_3 x_4.$$

We note that this implies that the functions f_2, f_4, g_4, f_6 are algebraically independent. Set $\mathfrak{a} = \{v = x_1 v_1 + x_2 v_2 + x_3 v_3 + x_4 v_4 | x_j \in \mathbb{C}\}$ and $\mathfrak{a}' = \{v = x_1 v_1 + x_2 v_2 + x_3 v_3 + x_4 v_4 | x_i \neq \pm x_j$ for $i \neq j\}$. One can check that the map $G \times \mathfrak{a}' \to V$ given by $g, v \longmapsto gv$ is regular. Furthermore, if $x \in \mathfrak{a}'$ then the set of $g \in G$ such that $gx = x$ is finite. Since $\dim G = 12$ and $\dim \mathfrak{a} = 4$ we see that if f is a G invariant polynomial then f is completely determined by its restriction to $\mathfrak{a}$ (since $G\mathfrak{g}'$ has interior). We also note that if $N = \{g \in G | g\mathfrak{a} = \mathfrak{a}\}$ then the group $W = N_{|\mathfrak{a}}$ is the subgroup of the group generated by the linear maps given by the permutations of v_1, v_2, v_3, v_4 and those that involve an even number of sign changes. For example,

$$\left(\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)$$

corresponds to $v_1 \to v_3, v_2 \to v_4, v_3 \to v_1, v_4 \to v_2$,

$$\left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \right)$$

corresponds to $v_1 \to v_1, v_2 \to -v_3, v_3 \to -v_2, v_4 \to v_4$. Thus W is the subgroup of the group of signed permutations with an even number of sign changes. One can check directly that every invariant under W is a polynomial in $(f_2)_{|\mathfrak{a}}$, $(f_4)_{|\mathfrak{a}}$, $(g_4)_{|\mathfrak{a}}$,$(f_6)_{|\mathfrak{a}}$ This completes the sketch of the proof of the theorem.

Remark 11 *This result is an explicit form of the Chevalley restriction theorem for the group $SO(4,4)$.*

We will now relate the space $\mathfrak{a}$ to the orbit structure. For this we need another construct. If $v, w \in \mathbb{C}^2$ then we write vw for the product of v, w in $S^2(\mathbb{C}^2)$. We set

$$[u_1 \bigotimes u_2 \bigotimes u_3 \bigotimes u_4, w_1 \bigotimes w_2 \bigotimes w_3 \bigotimes w_4]_i = \left(\prod_{j \neq i} \omega(u_j, w_j) \right) u_i w_i, i = 1, 2, 3, 4.$$

We say that $v, w \in V$ commute if $[v, w]_i = 0$ for $i = 1, 2, 3, 4$. We note that $[v_i, v_j]_k = 0$ for $i, j, k = 1, 2, 3, 4$. We also observe that if $v, w \in V$ and $g = (g_1, ..., g_4) \in G$ then $[gv, gw]_i = g_i[v, w]_i$ with the latter given by the action of $SL(2)$ on $S^2(\mathbb{C}^2)$. If $v \in V$ we will say that v is nilpotent if $T(v)T(v)^{\#}$ is nilpotent (that is, some power of $T(v)T(v)^{\#}$ is 0). This is the same as saying that $f_{2j}(v) = 0$ for all $j = 1, 2, ...$. Hilbert's criterion for this condition is

Theorem 12 *v is nilpotent if and only if there is a rational homomorphism, ϕ, of the group $\mathbb{C}^\times = \{z \in \mathbb{C} | z \neq 0\}$ into G such that $\lim_{z \to 0} \phi(z)v = 0$. We note that the action of G stabilizes the set of nilpotent elements.*

If $v \in V$ set $G_v = \{g \in G | gv = v\}$. We can now state the basic result on the orbit structure of G on V. We will call an element of $G\mathfrak{a}$ semi-simple. Then the Jordan decomposition of [KR] implies

Theorem 13 *An element $v \in V$ is semi-simple if and only if Gv is closed. Let v be an element of V then $v = s + n$ with s semi-simple and n nilpotent such that $[s, n]_i = 0$ for $i = 1, 2, 3, 4$. If s, s' are semi-simple and n, n' are nilpotent and commute with s, s' respectively then $s + n = s' + n'$ if an only if $s = s'$ and $n = n'$. If $g \in G$, $v \in \mathfrak{a}$ and $gv \in \mathfrak{a}$ then there exists $w \in W$ such that $wv = gv$. If $s \in \mathfrak{a}$ and $n, n' \in V$ are nilpotent and commute with s then if there exists $g \in G$ such that $g(s + n) = s + n'$ then there exists $h \in G_s$ such that $hn = n'$. Finally, if $s \in \mathfrak{a}$ and if $\mathcal{N}_s = \{v \in V | v$ is nilpotent and commutes with $s\}$ then $\mathcal{N}_s$ consists of a finite number of G_s orbits.*

We will next give a quantitative version of this theorem. We will first establish a bit more terminology.

We will say that a nilpotent element, n, is regular if setting $U = T(n)T(n)^{\#}$, $R = T(n)^{\#}$ then $R, RU + UR, RU^2 + U^2R, RU^3 + U^3R$ are linearly independent operators. A family of such examples is

$$a\,|0011\rangle + b\,|0100\rangle + c\,|1001\rangle + d\,|1010\rangle$$

with $abcd \neq 0$. It is easily seen that all of the regular elements of the above form are in the G orbit of the element with a, b, c, d all equal to 1. Let us call this element n_o. It turns out that there are 4 distinct regular nilpotent orbits. There are 20 distinct nilpotent orbits. The general theory also allows us to determine the general orbits. The number of different "types" of orbits is 90. The term "type" will become clear in the course of the discussion below leading to an explanation of the quantitative statement.

For each $i = 1, 2, 3, 4$ we define $\varepsilon_i \in V^*$ by $\varepsilon_i(v_j) = \delta_{ij}$. Let $\Phi = \{\pm(\varepsilon_i + \varepsilon_j) | 1 \leq i < j \leq 4\} \cup \{\varepsilon_i - \varepsilon_j | 1 \leq i \neq j \leq 4\}$. Set $\Delta = \{\alpha_1 = \varepsilon_1 - \varepsilon_2, \alpha_2 = \varepsilon_2 - \varepsilon_3, \alpha_3 = \varepsilon_3 - \varepsilon_4, \alpha_4 = \varepsilon_3 + \varepsilon_4\}$. If $s \in \mathfrak{a}$ then we define $\Phi_s = \{\alpha \in \Phi | \alpha(s) = 0\}$. One can show that if $s \in \mathfrak{a}$ then there exists $w \in W$ such that $\Phi_{ws} = \Phi \cap span_{\mathbb{Z}}(\Delta \cap \Phi_{ws})$. The main theorem implies that we need only look at elements s satisfying

$$\Phi_s = \Phi \cap span_{\mathbb{Z}}(\Delta \cap \Phi_s).$$

Here are the possibilities with $|\Delta \cap \Phi_s| \leq 1$.

$\Delta \cap \Phi_s = \emptyset, s = x_1v_1 + x_2v_2 + x_3v_3 + x_4v_4, x_i \neq \pm x_j$ for all $i \neq j$.

$\Delta \cap \Phi_s = \{\alpha_1\}, s = x_1(v_1 + v_2) + x_3v_3 + x_4v_4, x_i \neq \pm x_j$ for all $i \neq j$.

$\Delta \cap \Phi_s = \{\alpha_2\}, s = x_1v_1 + x_2(v_2 + v_3) + x_4v_4, x_i \neq \pm x_j$ for all $i \neq j$.

$\Delta \cap \Phi_s = \{\alpha_3\}, s = x_1v_1 + x_2v_2 + x_3(v_3 + v_4), x_i \neq \pm x_j$ for all $i \neq j$.

$\Delta \cap \Phi_s = \{\alpha_4\}, s = x_1v_1 + x_2v_2 + x_3(v_3 - v_4), x_i \neq \pm x_j$ for all $i \neq j$.

We note that the permutation (123) maps the set $\{s | s = x_1(v_1 + v_2) + x_3v_3 + x_4v_4, x_i \neq \pm x_j\}$ for all $i \neq j$ bijectively onto the set $\{s | s = x_1v_1 + x_2(v_2 + v_3) + x_4v_4, x_i \neq \pm x_j$ for all $i \neq j\}$. Similarly, there is a permutation that maps the set indicated by $\Delta \cap \Phi_s = \{\alpha_2\}$ onto the set indicated by $\Delta \cap \Phi_s = \{\alpha_3\}$. Finally, the sign change $v_1 \to v_1, v_2 \to -v_2, v_3 \to v_3, v_4 \to -v_4$ takes the set indicated by $\Delta \cap \Phi_s = \{\alpha_3\}$ onto the set indicated by $\Delta \cap \Phi_s = \{\alpha_4\}$. Thus by the basic theorem we need only consider the first two in our list. For $|\Delta \cap \Phi_s| \geq 2$ we will only list the cases up to the action of signed permutations involving an even number of sign changes. Here are all of the examples

1. $\Delta \cap \Phi_s = \emptyset, s = x_1v_1 + x_2v_2 + x_3v_3 + x_4v_4, x_i \neq \pm x_j$ for all $i \neq j$.
2. $\Delta \cap \Phi_s = \{\alpha_1\}, s = x_1(v_1 + v_2) + x_3v_3 + x_4v_4, x_i \neq \pm x_j$ for all $i \neq j$.
3. $\Delta \cap \Phi_s = \{\alpha_1, \alpha_2\}, s = x_1(v_1 + v_2 + v_3) + x_4v_4, x_1 \neq \pm x_4$.
4. $\Delta \cap \Phi_s = \{\alpha_1, \alpha_3\}, s = x_1(v_1 + v_2) + x_3(v_3 + v_4), x_1 \neq \pm x_3$.
5. $\Delta \cap \Phi_s = \{\alpha_1, \alpha_4\}, s = x_1(v_1 + v_2) + x_3(v_3 - v_4), x_1 \neq \pm x_3$.
6. $\Delta \cap \Phi_s = \{\alpha_1, \alpha_2, \alpha_3\}, s = x_1(v_1 + v_2 + v_3 + v_4), x_1 \neq 0$.
7. $\Delta \cap \Phi_s = \{\alpha_1, \alpha_2, \alpha_4\}, s = x_1(v_1 + v_2 + v_3 - v_4), x_1 \neq 0$.
8. $\Delta \cap \Phi_s = \{\alpha_2, \alpha_3, \alpha_4\}, s = x_1v_1, x_1 \neq 0$.

9. $\Delta\cap\Phi_s=\{\alpha_1,\alpha_3,\alpha_4\},s=x_1(v_1+v_2),x_1\neq 0$.
10. $\Delta\cap\Phi_s=\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\},s=0$.

We now count the number of G_s orbits in $\mathcal{N}_s$ in each of the 10 cases above. Case 1. Yields 1 since $\mathcal{N}_s=\{0\}$. 2. Yields 2. 3. Yields 3. 4. and 5. Yield 8. 6.,7.,8.Yield 7. 9. Yields 27. 10. Yields 20. The total is our promised 90.

Here are some examples. The extremes in part 10. of the list involving the non-zero orbits are the 4 regular nilpotent orbits and the orbit of product states

$$\{u_1\bigotimes u_2\bigotimes u_3\bigotimes u_4|u_i\in\mathbb{C}^2-\{0\}\}.$$

We now look at the so called WHZ state. This is (up to normalization) $s=|0000\rangle+|1111\rangle=v_1+v_2$. It appears in case 9. Thus there are 26 additional orbits with s-component the WHZ state. Here is how you find them. We note that

$$G_s=\left\{\left(\left[\begin{array}{cc}a_1&0\\0&a_1^{-1}\end{array}\right],\left[\begin{array}{cc}a_2&0\\0&a_2^{-1}\end{array}\right],\left[\begin{array}{cc}a_3&0\\0&a_3^{-1}\end{array}\right],\left[\begin{array}{cc}a_4&0\\0&a_4^{-1}\end{array}\right]\right)|a_1a_2.a_3a_4=1\right\}$$

The space of all elements $v\in V$ such that $[s,v]_i=0$ for all i is spanned by s and

$$\{|0,0,1,1\rangle,|0,1,0,1\rangle,|1,0,0,1\rangle,|1,0,1,0\rangle,|1,1,0,0\rangle\}.$$

Let $S_1=\{|0,0,1,1\rangle,|1,1,0,0\rangle\}$, $S_2=\{|0,1,0,1\rangle,|1,0,1,0\rangle\}$, $S_3=\{|1,0,0,1\rangle,|0,1,1,0\rangle\}$. Then the orbits corresponding to s are the orbits through $s+\sum_{j\in J}n_j$ where J is a subset of $\{1,2,3\}$ and $n_j\in S_j$. There are 27 such orbits. The orbits with minimal stability groups are the ones corresponding to $|J|=3$. There are 8 of them.

We note that for 2 and 3 qubits the state $\frac{|00...\rangle+|11...\rangle}{\sqrt{2}}$ was arguably the most entangled. In the case of 4 qubits this state is just $\frac{v_1+v_2}{\sqrt{2}}$ and so it is not even in $\mathfrak{a}'$.

This discussion indicates that the measures of entanglement for 4 qubits will form a complicated algebra. One useful invariant of such an algebra is the Hilbert series.

5.2 Some Hilbert series of measures of entanglement

If V is a real vector space then we set $\mathcal{P}^j(V)$ equal to the complex vector space of all polynomials on V that are homogeneous of degree j. If W is a complex vector space that $\mathcal{P}^j_{\mathbb{R}}(W)=\mathcal{P}^j(V)$ where V is W as a real vector space. We say that a subalgebra, A, of $\mathcal{P}^j_{\mathbb{R}}(W)$ is homogeneous if it is the direct sum of $A^j=\mathcal{P}^j_{\mathbb{R}}(W)\cap A$. If A is a homogeneous subalgebra of $\mathcal{P}_{\mathbb{R}}(W)$ then the formal power series

$$h_A(q)=\sum_{j\geq 0}q^j\dim A^j$$

is called the *Hilbert series* of A.

In the results we have described for 2 and 3 qubits imply that

$$h_{\mathcal{P}_{\mathbb{R}}(\mathbb{C}^2 \bigotimes \mathbb{C}^2)^K} = \frac{1}{(1-q^2)^3}$$

and

$$h_{\mathcal{P}_{\mathbb{R}}(\mathbb{C}^2 \bigotimes \mathbb{C}^2 \bigotimes \mathbb{C}^2)^K} = \frac{(1+q^4)(1+q^6)}{(1-q^2)(1-q^4)^6}.$$

As we predicted the case of 4 qubits is much more complicated. Here is the series (see [W])

Numerator: $1+3q^4+20q^6+76q^8+219q^{10}+654q^{12}+1539q^{14}+3119q^{16}+$
$5660q^{18}+9157q^{20}+12876q^{22}+16177q^{24}+18275q^{26}+18275q^{28}+$
$16177q^{30}+12876q^{32}+9157q^{34}+5660q^{36}+3119q^{38}+1539q^{40}+$
$654q^{42}+219q^{44}+76q^{46}+20q^{48}+3q^{50}+q^{54}$
Denominator: $(1-q^2)^3(1-q^4)^{11}(1-q^6)^6$.

5.3 A measure of entanglement for n qubits

In this subsection we will describe a specific measure of entanglement introduced in [M-W2] that has been used experimentally as test of entanglement. We will give a formula for it in terms of representation theory and show how it can be slightly modified to be an entanglement monotone.

Let $V = \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$ n-fold product. We look upon $V \otimes V$ as $\left(\mathbb{C}^2 \otimes \mathbb{C}^2\right) \otimes \cdots \otimes \left(\mathbb{C}^2 \otimes \mathbb{C}^2\right)$. Let

$$S : \mathbb{C}^2 \otimes \mathbb{C}^2 \to S^2(\mathbb{C}^2)$$

and

$$A : \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \bigwedge \mathbb{C}^2$$

be the canonical orthogonal projections. If $F \subset \{1, ..., n\}$ then we define p_F to be the product

$$R_1 \otimes \cdots \otimes R_n$$

with $R_i = A$ if $i \in F$ and $R_i = S$ otherwise. Then if $v \in V$ we have

$$v \otimes v = \sum_{|F| \text{ even}} p_F(v \otimes v).$$

The p_F are orthogonal projections so we have in particular

$$\|v\|^4 = \sum_{|F| \text{ even}} \|p_F(v \otimes v)\|^2 .$$

We set

$$\Upsilon(v) = \|v\|^4 - \|p_{\emptyset}(v \otimes v)\|^2 .$$

The following result is not completely obvious. We will sketch a reduction to the same assertion for another measure of entanglement.

Theorem 14 *A state $v \in V$ is a product state if and only if $\Upsilon(v) = 0$.*

This measure of entanglement is related to one denoted Q in [MW-2] (they are the same for 2 and 3 qubits) and which was defined as follows. If $0 \leq j < N = 2^n$ and $j = \sum_{m=0}^{n-1} j_m 2^m$ then if $0 \leq i < n$ define $t_i(j) = \sum_{0 \leq m < i} j_m 2^m + \sum_{i<m<n} j_m 2^{m-1}$. If $v = \sum_{0 \leq j < N} v_j \left|j\right\rangle$ then we set $v_{i,0} = \sum_{j_i=0} v_j \left|t_i(j)\right\rangle$ and $v_{i,1} = \sum_{j_i=1} v_j \left|t_i(j)\right\rangle$. Thus if

$$v = \frac{|111\rangle + |001\rangle + |010\rangle + |100\rangle}{2}$$

then

$$v_{2,0} = \frac{|01\rangle + |10\rangle}{2}, v_{2,1} = \frac{|00\rangle + |11\rangle}{2}.$$

We set $Q(v) = \sum_{i=0}^{n-1} \left\| v_{i,0} \bigwedge v_{i,1} \right\|^2$. Here in $W \bigotimes W$, $u \bigwedge w = \frac{u \bigotimes w - w \bigotimes u}{2}$ and we use the tensor product inner product. We note that $Q\left(\frac{|000\rangle + |111\rangle}{\sqrt{2}}\right) = \frac{3}{8}$ and $Q\left(\frac{|001\rangle + |010\rangle + |100\rangle}{\sqrt{3}}\right) = \frac{1}{3}$. One can show that if $v \in \bigotimes^n \mathbb{C}^2$ then

$$Q(v) = \sum_{k=1}^{\frac{n}{2}} k \sum_{|F|=2k} \left\| p_F(v \bigotimes v) \right\|^2.$$

We note that in [MW-2] we proved that the (relatively easy) result that the Theorem above is true for Q replacing Υ. Since $Q(v) = 0$ if and only if $\left\| p_F(v \bigotimes v) \right\|^2 = 0$ for all $|F| > 0$ and Υ has the same property. Hence $\Upsilon(v) = 0$ if and only if $Q(v) = 0$.

In a forthcoming article we will prove that Υ is an *entanglement monotone*. This assentially means that quantum operations (such as measurements and local transformations) cannot increase its value. This condition is sometimes included in the definition of a measure of entanglement.

References

[KR] B. Kostant and S. Rallis, Orbits and Lie group representations associated to symmetric spaces, Amer. J. Math.,93(1971),753-809.

[MW2] David Meyer and Nolan Wallach, Global entanglement in multiparticle systems. Quantum information theory. J. Math. Phys. 43 (2002), no. 9, 4273–4278.

[R] Erik Rains, Quantum codes of minimum distance two. IEEE Trans. Inform. Theory 45 (1999), no. 1, 266–271.

[W] Nolan R. Wallach, The Hilbert series of measures of entanglement for 4 qubits, Acta Appl. Math. 86 (2005), no.1-2,203-220.