Introduction to quantum computation: Qubits and Bell inequalities

Lecture Notes for Quantum Computing

Lecturer: Dorit Aharonov
Scribe: Dan Cohen and Avinatan Hassidim

March 4-9, 2001

1 Preliminaries

1.1 Requirements

- Scribe notes. A couple of students "scribe" the lectures each week: take detailed notes, add missing details, and format them accordingly. Scribe notes are due on Monday of the next week, and will be posted on the course web page (www.cs.huji.ac.il/ doria/qcomp.huji2001.html) hopefully one week later. It is the responsibility of the students to reach the teacher in time (e-mail is doria@cs.huji.ac.il) so that the notes will be ready by the end of the week. Leave two days at least for corrections; Grades will depend on the amount of work I have to put into the notes and into nagging you about them.

- Assignments. There will be six exercises, put on the course web page once every two weeks, to be submitted a week later, on Wednesday, in class.

- Seminar. In the end of the year each student will talk about an article he or she read.

- Web site: **www.cs.huji.ac.il/ doria/qcomp.huji2001.html**

1.2 Prerequisites

I am assuming that you have basic knowledge in linear algebra, and complexity theory. You need to have significant exposure to theoretical computer science, or to know quantum physics and underlying mathematics. It would be difficult to learn all of this material from scratch.

1.3 References

Pointers to several references which you will find useful can be found on the course web page: Preskill's lecture notes, Vazirani's course lecture notes, Nielsen's and Chuang's book, Peres's book and my review about quantum computation.

2 Introduction

What is a computation? A computation is a process in which an input goes through a physical process, and generates an output. There are many beautiful and exotic examples for computations. A famous one is Gaudi's anolog computation. Gaudi was an architect who wanted to build a ceiling of arches for a church. (La Segrada Familia in Barcelona.) He ties ropes in lengths proportional to the lengths of the arches he desired, and tied them one to another. He tied the whole construction to the ceiling, and gravitation immediately put the whole construction in equilibrium. In a mirror on the floor he could see the reflection, which was simply the design of the entire Church. He saw the exact shape he should build the arches. Is that considered a computation? This is a one purpose computation. We will refer to a "computer" as something which can perform a universal computational process.

A breakthrough in computer science was achieved when people understood that if we relax the concept of computation, and require that the answer be correct only with high probability, we can gain a lot in efficiency. An example is when one has access to a database which is one of two possibilities: either for all inputs x $f(x)$ is the same, or for half the inputs $f(x) = 0$ and for half $f(x) = 1$. To give a deterministic answer, one has to query the database more than half its size. If we allow even a tiny error, we already gain a lot since an error ϵ can be achieved after $O(log(1/\epsilon))$ queries, independent of the size of the database. We will allow probabilistic computations, which generate the right answer with a certain probability.

Until the discoveries of the last decade, the class of problems which were considered tractable were those which can be solved by a probabilistic Turing machine in polynomial time. It was believed that probabilistic computation captures the entire notion of computation. This belief is best express in the fundamental Church Turing thesis, which asserts: "A probabilistic Turing machine can simulate efficiently any "reasonable" physical process. " This is of course a thesis, not a theorem, since "reasonable" is not defined mathematically.

However, in the last decade discoveries in a new and exciting area, quantum computation, challenge this very basic belief of theoretical computer scientists. Quantum computation deals with fundamental questions in computer science, quantum physics, and mathematics. The idea is to use computation based on quantum physics effects in order to improve "conventional" computations. I will go through the history of the results in this area, which will go hand in hand with the description of what we will do in the course.

3 Course Outline

- Models of Quantum Computation. Deutsch was the first to define this model: he defined the model of quantum Turing machines in 1985, and the model of quantum circuits in 1989.

- Quantum Algorithms. We will study the early quantum algorithms, Deutsch-Jozsa and Simon, and proceed to the celebrated Shor's Algorithm for Factorization, a problem which is believed to be hard for classical computers. The RSA cryptographic system is based on it, and hence the discovery by Shor (in 1994) caused a lot of excitement all over the world. We will discuss Fourier transforms, the hidden subgroup problem and the many open questions in this area.

- Quantum lower bounds: we will study various limits to the advantage of quantum computers over classical ones.

- Quantum noise and Quantum Error Correcting Codes. The existence of such ways of protecting the quantum information from noise is essential to the understanding of quantum computational power- since any physical system will be subjected to noise, the quantum model is meaningful only if it can be made robust to noise. Quantum error corrections came as a surprise to physicists, due to various reasons, eg. quantum information is continuous, and any state can (theoretically at least) be true, which seem to make it hard to detect errors. We would like to make sure we make fair use of our physical resources - for example we can not measure things at absolute precision. Shor(1996) was the first to find a quantum error correcting codes, and a whole theory developed since then.

- Fault tolerant Quantum Computation and threshold - we will show that we can make calculations for noise values smaller than a certain threshold, which means that quantum computation can theoretically be made robust to noise.

- Quantum communication: we will speak about quantum information, quantum entropy and protocols for quantum communication. We will show the exponential separation between the number of communication bits required in the quantum versus the classical models, a result by Ran Raz (1998).

- Quantum cryptography: We will talk about quantum cryptographic protocols which are unconditionally secure against cheaters, like the quantum key distribution protocol and a quantum biased coin flipping protocol. We will also talk about impossibility results in this area, like the impossibility of quantum bit commitment.

- Quantum Complexity Theory: quantum analogues of NP, IP, etc.

- If time permits we will study some of the more recent quantum algorithms, quantum Markov chains, and algebraic quantum algorithms.

- Optional - experimental Realizations and building a quantum, computer

4 So why Quantum physics?

Quantum physics is special because it changed drastically our way of thinking of how things happen in physics and what could happen. What CAN be done has increased(though some things were proved to be impossible) thus leading also the interest in using the new phenomena to improve possible computations and computers.

In order to understand the massive change in understanding of nature that spawned Quantum physics we will describe a simple experiment similar to one conducted in high schools - the two slit experiment.

[The description of this experiment will be added shortly. For now, you can either look at Feynmann's amazing book about quantum physics, or at the description in the first lecture of Vazirani's course in 1997 (not the 2000 course.)]

The lesson from the two slit experiment is that the position of the particle is described by a "wave function", or a linear combination with complex coefficients of all classical possible locations. When measured, though, the detectors either finds the particle there or not. This is the duality between particle and wave behavior. The fact that quantum physics allows superposition of all possible configuration will translate to quantum computers being in a superposition of all possible configurations. In this course we will see how to use these phenomena to design better algorithms.

One has to remember that physics is a model of nature. This means that there is not "right or wrong". We search for a model of a physical system , which agrees with experiments we make and is as simple as possible. Thus , the basics of Quantum physics are a group of postulates one has to accept, and hope they will lead to useful predictions. Some of the properties are very counter intuitive: This only means that our intuition is built from experience which is in the macroscopic scale. in which quantum effects are very rarely seen; Quantum physics is the physics that rules the world in all scales (or so we believe after a century of experiments) but in the macroscopic regime quantum phenomena coincides with classical phenomena, and hence we cannot detect all these peculiar behaviors in which entities behave like waves and so on. Let us now try to turn to study the basic rules of quantum mechanics.

5 A Qubit- One Quantum Bit

The basic entity of quantum computation is a qubit(pronounced "cue-bit"), or a quantum bit. In analogy with classical computation, we want to have as our elementary information unit something which can take two values, 0 and 1. In quantum physics, though, as we have learned from the two slit experiment, the state of the system can be in linear combination of all possible classical states. Thus, the state of a quantum bit, or a “qubit” (pronounced cue-bit) will be a vector in the complex two dimensional vector space. To be able to define a qubit, and in general quantum systems, we now make a detour in which we will define vector spaces, basis states, inner products, and Hilbert spaces.

5.1 vector spaces

A vector space V is a group of items(called vectors).It is defined together with a field F like the real numbers(R) or the complex numbers(C). the items in the field are called scalars. In Quantum Physics , C is almost always used.

The vector space has to satisfy the a number of properties :

- Closure under sum and multiplication by scalars for all $v \in V, u \in V$, $(v + u) \in V$ and also $\lambda v \in V$
- Existence of a zero , i.e. a vector $0 \in V$ such as for every $v \in V$ $(v + 0) = v$
- commutative $v \in V, u \in V$, u+v=v+u
- multiplication by a scalar
- associative

5.2 Basis

A Basis for a vector space V is a group of vectors $\{v_1...v_n\} \subseteq V$ which satisfies two conditions :

- It spawns V : for every $u \in V$ one can find $\{\alpha_1...\alpha_n\} \subseteq F$ so $\sum \alpha_i v_i = u$
- linear independence : if $\sum \alpha_i v_i = 0$ for $\{\alpha_1...\alpha_n\}$ then $\forall i \alpha_i = 0$

An Orthonormal base is a base $\{v_1...v_n\} \subseteq V$ which satisfies

- $\forall i \ |v_i| = 1$
- $\forall i, j \ (v_i, v_j) = 0$ if $i \neq j$

A number of facts are known :

- For every vector space V a base can be found , furthermore , an Orthonormal Base always exists.
- Every two basis of V have the same number of vectors. Thus we can define the "dimension" of a vector space as the number of vectors in its base.

When a Basis is agreed upon one can represent a vector $u = \sum \alpha_i v_i$ by those α_i, writing them in vectorial notation :

$$\left(\alpha_0..\alpha_n \right)$$

5.3 inner products

An inner product is a function defined given a vector space V, $f : V \times V \to \mathcal{C}$. it is popular to write (u,v) instead f(u,v). we shall discuss inner products defined on a vector space V whose field is C. f satisfies that for any $v, u, w \in V \alpha, \beta \in C$:

- symmetry : $(u, v) = (v, u)^\star$ (* is the complex conjugate)
- linearity in second argument: $(w, \alpha u + \beta v, w) = \alpha(w, u) + \beta(w, v)$
- positive : $(u, u) \geq 0$ (u,u)=0 iff u=0

By using it's linearity one can verify that if vectors v,u are represented in an orthonormal basis by vectors

$$\left(\alpha_0..\alpha_n \right)$$

$$\left(\beta_0..\beta_n \right)$$

then $(v, u) = \sum \alpha_i \beta_i^\star$

5.4 Norm spaces

A norm is a function defined given a vector space V, $f : V \to R^+$. it is popular to write $|v|$ instead of f(v). f satisfies that for any $u, v \in V \alpha \in C$:

- positivity : $|v| \geq 0$ and $|v| = 0$ iff v is the zero vector
- linearity : $|\alpha v| =| \alpha | \, |v|$
- the triangle inequality : $|\alpha + \beta| \leq |\alpha| + |\beta|$

It is easy to prove that if an inner product is defined on V then one can define a norm on V which is defined : $|v| = \sqrt{(v, v)}$ This is called the induced norm(induced by the inner product)

A "normalized" vector v is a vector whose norm $|v|$ is 1.

5.5 Hilbert space

A Hilbert space is a Vector space V , on which an inner product is defined and which is complete under the induced norm (a complete vector space is one in which every Cauchy sequence converges.) For finite dimensional spaces, the restriction regarding completeness is trivially fulfilled. Since in this course we will only be interested in finite dimensional vector spaces, a Hilbert space for us is a vector space accompanied with an inner product. The vector spaces which we are interested in are $\mathcal{C}^d$ for some finite d, with the inner product defined above.

5.6 Bras and kets

Quantum physics describes the state of a physical system with d possible states by a vector, called the "state" of the system, $v \in \mathcal{C}^d$. This vector is often denoted using the Dirac notation. When one wants to talk about a column vector in $\mathcal{C}^d$ one writes instead of "v" : $|v\rangle$, which is called a "ket" vector. We can denote the basis states of an orthonormal basis for $\mathcal{C}^d$ by $|0\rangle, |1\rangle, ...|d-1\rangle$. Then $|v\rangle = \sum_i v_i |i\rangle$. This is just an equivalent notation to the vector

$$\begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{d-1} \end{pmatrix}$$

Remember, this is just a notation, which once one gets used to, becomes extremely intuitive. It turns out to be a very convenient notation, as we will see later. The transposed complex conjugate of v is a raw vector, $(v_0^*, ...v_{d-1}^*)$, which is denoted by $\langle v|$, and is called a "bra" vector. Note that the inner product between two vectors, $(|\alpha\rangle, |\beta\rangle)$ is actually equal to the product (as matrices): $\langle\alpha| \cdot |\beta\rangle$. For brevity, we will omit one of the two lines between the two vectors and denote: $(|\alpha\rangle, |\beta\rangle) = \langle\alpha|\beta\rangle$. This is the reason for the weird names "bra" and "ket": In the inner product, they form together "brackets".

5.7 A qubit

We define a qubit to be the two dimensional Hilbert space $\mathcal{C}^2$. In Physics it is called a two-state system. The orthonormal basis, $\{|0\rangle, |1\rangle\}$ represents the "classical" states of the system, but in general the state of a qubit is a "superposition" of these two classical states, i.e. any combination $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$.

We will require that any quantum state describing a physical system is normalized, i.e. it has norm one.

$$\sum_i |c_i|^2 = 1.$$

This requirement is due to the fact that, as we will soon see in the definition of measurements, the norm squared of projections on subspaces of the state have a probabilistic meaning. The normalization requirement simply means that the sum of probabilities of all possible measurement outcomes should add up to 1.

6 Measurements

We now know how a quantum state looks like. But how do we observe it? The definition of a measurement in Quantum Physics is very different from what we are familiar with from everyday life. A measurement is defined by an Hermitian matrix (O is hermitian if $O = O^\dagger$). An Hermitian matrix can be diagonalized by an orthonormal set of eigenvectors with real eigenvalues. This means that there is a basis $\{|v_1\rangle ... |v_n\rangle\} \subseteq V$ for V such that $O|v_i\rangle = v_i|v_i\rangle$ with v_i the eigenvalue associated with $|v_i\rangle$. let's assume the system is in state $|\psi\rangle$ and $\sum \alpha_i|v_i\rangle = |\psi\rangle$ then the basic postulate of Quantum Physics is that the result of a measurement of an observable O can be only $v_1...v_n$. The decision which result will be is probabilistic and the probability for result v_i is $|\alpha_i|^2$. Furthermore , if the result is v_i then after the measurement is over the system will be in the projection of $|\psi\rangle$ on the subspace spanned by $|v_i\rangle$ (This is what is known as the "collapse of the wave function." Thus measuring can change the state of the system! What is most important is that if we measure one value in basis A and then a second basis B which is not orthogonal to A then we will be projected into a state that is not orthogonal to the states of A. Thus if we measure again A there is a chance we will get a different result. This phenomena leads to the famous idea that measuring some quantities can affect other- The uncertainty principle.

we can deduce some important facts from this postulate which agree with intuition :

- if the system is already in a state which is an eigenvector of O , i.e. $|v_i\rangle$ then the chance for the measurement to be v_i is 1(because $\alpha_i = 1$)
- the total probability of all possibilities is 1. that is because we use normalized states so $1 = |v|^2 = \sum \alpha_i^\star \alpha_i = \sum |\alpha_i|^2 = \sum prob(measuring v_i)$

Note that the measurement operation does NOT correspond to applying the linear matrix O on $|psi\rangle$, but the basis in which we measure, and the possible classical results, are determined by O.

Another important fact that we would like to stress here has to do with the amount of information encoded in one qubit. One could believe that it is possible to encode infinite information in one Qubit because c_0 and c_1 are real number so they could be represented with infinite accuracy in the state of a qubit. Although a single (very complex) distribution of a single variable (or quantum bit) could theoretically encode the whole bible (for example) measuring it (or taking a single test of the distribution) will give us 1 or 0 - which is at most one bit worth of information.

The definition we have defined above is a **complete** measurement of the system, i.e. it projects the state to subspaces of dimension 1 (The eigenvectors of O.) We will define later a slightly more general notion, in which we allow to measure only partial information about the system, projecting the state to eigenspaces of dimensionality more than 1.

One of the most useful measurements which we will be using all the time is the one in which O is the one qubit measurement in the standard basis, represented by the matrix

$$O = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

which projects the state of a qubit in the direction of one of the basic states $|0\rangle$ or $|1\rangle$.

7 Two Qubits

When combining two quantum systems, one described by A and the other by the Hilbert space B, the Hilbert space that describes the new system is the **tensor product** of A and B.

Given two vector spaces U and V with dimensions n and m we define their tensor multiplication as a vector space $W = V \otimes U$ with dimension mn, associated with a mapping, or a function, from pairs of vectors $v \in V, u \in U$ to the vector space $W = V \otimes U$. This function is denoted by the sign $\otimes$, so

$$v, u \longmapsto v \otimes u \in W.$$

The function $\otimes$ is required to be linear, and associative.

We define the inner product in the space W to be induced from the inner product of the spaces V and U: $(|\alpha\rangle \otimes |\beta\rangle, |\alpha'\rangle \otimes |\beta'\rangle) = (|\alpha\rangle, |\alpha'\rangle)(|\beta\rangle, |\beta'\rangle)$. (We have to extend this definition to all vectors, not only tensor products, and we do that using linearity. It is easy to check that this indeed satisfies the requirements of an inner product.) It follows that if $|i\rangle, |j\rangle$ are orthonormal basis states for V, U respectively, all pairs $(|i\rangle, |j\rangle)$ are mapped to orthogonal vectors, then the states $\{|i\rangle \otimes |j\rangle\}$ is an orthonormal basis for W.

Due to linearity, it suffices to know what the pairs of basis states are mapped to, in order to determine the mapping on the entire space. So we can view $W = V \otimes U$ as the vector space spanned by the orthonormal states $(\{|i\rangle \otimes |j\rangle\})$ for two bases of V and W. For simplicity and brevity, we will use the short notation $|i,j\rangle \equiv |i\rangle \otimes |j\rangle$. The tensor product of two vectors $\alpha = \sum \alpha_i |i\rangle \in U$ and $\beta = \sum \beta_j |j\rangle \in V$ is by linearity $\sum \alpha_i \beta_j |i,j\rangle \in W$.

Note that the axiom saying that the combination of two quantum systems is described by their tensor product fits our intuition that a quantum system should be able to be in a superposition of all classical configuration of the underlying classical system; Indeed, two classical systems, where the first can be in configurations i and the second in configurations j can be together in all pairs of configurations i, j. This is why the dimensionality of a system combined of two quantum systems is the **product** of the dimensions, (rather than the sum).

A very important observation is that in the tensor product space, not every vector is a tensor product; A general vector is a linear combination of tensor product vectors. This means that the state of two qubits cannot be described as the state of the first qubit and the state of the second qubit; there are correlations between them.

Here is an example of how this works for two qubits. We will build a simple isomorphism between $C^2 \otimes C^2$ and C^4 in the following way: $(|i\rangle, |j\rangle) \mapsto |i, j\rangle$; for example $(|0\rangle, |1\rangle) \mapsto |01\rangle$. We can now create a superposition of two states of two particle by a single vector. We will assume that the first coordinate describes the first qubit, and that the second describes the second one.

7.1 Measuring One Qubit

One can also be interested in the results of measuring the state of just one qubit in a system of more qubits. How to define such a process? Technically speaking, such a measurement is represented by a measurement of one qubit, O tensor product with identity on the rest of the qubits, $O \otimes I$. Alternatively, we generalize the definition of the measurement from before. Instead of projecting on eigenvectors, we allow projections on orthogonal subspaces. To measure one qubit in a basis $|\alpha\rangle, |\alpha^{\perp}\rangle$ means to project the state on one of the two subspace, the first spanned by all vectors of the form $|\alpha\rangle \otimes |i\rangle$ for all i basis states for the rest of the qubits, and the second subspace is spanned by all vectors of the form $|\alpha^{\perp}\rangle \otimes |i\rangle$ for all i basis states for the rest of the qubits.

An example is a measurement of the first qubit in the standard basis, in a two qubit state:

$$\begin{pmatrix} \alpha_{00} \\ \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{pmatrix}$$

If the result of the first measurement is 0, the state will be projected to

$$\begin{pmatrix} \lambda\alpha_{00} \\ \lambda\alpha_{01} \\ 0 \\ 0 \end{pmatrix}$$

for $\lambda = \frac{1}{\sqrt{\alpha_{00}^2+\alpha_{01}^2}}$ which is used to normalize the vector. If the result is 1, the state will collapse to the vector

$$\begin{pmatrix} 0 \\ 0 \\ \lambda\alpha_{10} \\ \lambda\alpha_{11} \end{pmatrix}$$

for $\lambda = \frac{1}{\sqrt{\alpha_{10}^2+\alpha_{11}^2}}$. This rule, is very reminiscent of the rule for measuring one bit in a probability distribution over two bits; we simply condition on what we know, and renormalize the remaining probability distribution.

8 EPR

8.1 Entanglement

Entanglement is THE effect that contributes to many of the "bizzare" quantum phenomena that can be used for example in quantum communication. An entangled state $|\psi\rangle$ is a state in the resulting space C^4 which cannot be represented as a tensor multiplication of a state $|\phi_1\rangle$ of the first particle and a state $|\phi_2\rangle$ of the second. In entangled states the states of the qubits are correlated. These correlations seem to be much stronger than classical correlations between random bits, as we will see shortly.

8.2 EPR pair

Definition 1 *The state in* C^4 $|\psi^-\rangle \equiv \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ *is called an EPR (Einstein Podulsky Rosen) pair or the Bell state.*

Measuring the first bit of $|\Psi^-\rangle$ in the standard basis yields a 0 with probability 1/2, and 1 with probability 1/2. Likewise, measuring the second bit of $|\Psi^-\rangle$ yields the same outcomes with the same probabilities. Measuring one bit of this state yields a perfectly random outcome. Yet the important phenomenon is that measuring one bit determines the other. Suppose , for example That the first bit is measured to be 0 then as was shown the state collapses from

$$\begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \\ 0 \end{pmatrix}$$

to

$$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

so it is certain that measuring the second bit will be 1.

It is important to notice that this property is invariant to change of basis. This means that the strong correlations between measurements of the two qubits exist not only in the standard basis, but in any basis.

Theorem 1 *Consider an arbitrary orthonormal basis for one qubit,* $v, v^\perp$ *where* $v = \alpha|0\rangle + \beta|1\rangle$. *Then* $v^\perp = \bar{\alpha}|1\rangle + \bar{\beta}|0\rangle$, *where* $\bar{\alpha}, \bar{\beta}$ *are complex conjugates of* α, β. *Claim:* $|\psi\rangle = \frac{1}{\sqrt{2}}\left(|vv^\perp\rangle - |v^\perp v\rangle\right)$.

Proof:

$|\psi\rangle$ can be written in terms of the new basis as $|vv^\perp\rangle - |v^\perp v\rangle$:

$$\begin{aligned}
|vv^\perp\rangle - |v^\perp v\rangle &= (\alpha|0\rangle + \beta|1\rangle)\left(\bar{\alpha}|1\rangle - \bar{\beta}|0\rangle\right) \\
&\qquad - \left(\bar{\alpha}|1\rangle - \bar{\beta}|0\rangle\right)(\alpha|0\rangle + \beta|1\rangle) \\
&= \left(\alpha\bar{\alpha} + \bar{\beta}\beta\right)|01\rangle + \left(-\alpha\bar{\beta} + \bar{\beta}\alpha\right)|00\rangle \\
&\qquad + (\beta\bar{\alpha} - \bar{\alpha}\beta)|11\rangle + \left(-\beta\bar{\beta} - \bar{\alpha}\alpha\right)|10\rangle \\
&= \left(\alpha\bar{\alpha} + \beta\bar{\beta}\right)|01\rangle - \left(\alpha\bar{\alpha} + \beta\bar{\beta}\right)|10\rangle \\
&= |01\rangle - |10\rangle \quad \square
\end{aligned}$$

8.3 The EPR paradox

The famous EPR paradox was described in 1935 by Einstein, Podolski and Rosen. This paradox was a gedanken(imaginary) experiment in which they contradicted the two physical assumptions they believed to be true : realism and locality.

Let's begin with understanding these two assumptions :

- realism is the assumption that if one can predict the result of an experiment with probability 1 then this result must describe a physical property, a real property. For example, suppose Alice and Bob each receive a colored ball. They know that one of the balls is black and one is white.Furthermore Alice knows that her ball is white so she can predict with certainty that Bob's ball is black. The assumption of realism implies that the black color of Bob's ball is a property of the ball so if we keep measuring it's color it will remain the same.

- locality is the assumption that information cannot be passed faster then light. Thus nothing that is done in one location cannot affect a distant location immediately.

Now, let us assume that Alice and Bob share and EPR state $|\Psi^-\rangle$. By measuring it's qubit in the standard basis, Alice can predict with certainty Bob's result which will be measured in the standard basis too.Thus the value measured by Bob in the standard basis is a real property(because the information could not pass from Alice as assumed in the locality assumption). The same will happen if they both measure in a basis rotated by 45 degrees. Thus the values measured by Bob in both bases are real properties. Yet as was explained before values measured in non-orthogonal bases cannot be both real properties because measuring one can changes the result of measuring the other.

Thus we see that these "natural" assumptions contradict Quantum theory. Much later it was verified experimentally that this contradiction was not a proof of the incompleteness of Quantum theory but a proof that these assumptions are not true in nature as we know it.

8.4 Bell's Inequality

Almost three decades after the EPR paradox was first presented, in 1964, Bell showed that these properties of EPR states were not merely fodder for a philosophical discussion, but had verifiable consequences.

Bell's inequality states: There does not exist any local hidden variable theory (A theory which satisfies both locality and realism) consistent with these outcomes of quantum physics.

Consider the following communication protocol in the classical world: A and B are two parties who share a common random string. They receive bits X_A, X_B, and try to output bits a, b respectively, such that$X_A \wedge X_B = a \oplus b$.

classically, since they have no knowledge they will both return 0 (or 1, it doesn't matter) because they know that for two random bits X_A, X_B there is a probability of 3/4 that $X_A \wedge X_B = 0$. Since Alice and Bob can not transfer information, that is the best they can do. assume not, that is assume that a certain choice they have made can grant them better chances. Notice that if Alice and Bob use a deterministic method, The probability of success is an integer multiplication of 1/4, because there are 4 cases of X_A, X_B. Since the probability of success can not be 1 (because that would be equivalent to knowing the other's bit) then 3/4 is the best success probability we can expect.

In the quantum mechanical analogue of this protocol, A and B share the EPR pair $|\Psi^-\rangle$. As before, they receive bits X_A, X_B, and try to output bits a, b respectively, such that $a, b \in 0, 1, \quad X_A \wedge X_B = a \oplus b$.

If the odd behavior of Ψ^- can be explained using some hidden variable theory, then the best they can do in both scenarios should be the same, i.e. the two scenarios are equivalent.

We will show that the quantum mechanical system can do better. Specifically, if A and B share an EPR pair, we will show that the probability $Prob(X_A \wedge X_B = a \oplus b)$ is greater than 3/4. Thus it is not possible to explain the EPR paradox using a hidden variable model, i.e. a probabilistic model which is local, i.e what happens in Alice's side is unknown to Bob, and cannot effect him, and vice versa.

To begin, notice that if a measurement in the standard basis yields $|0\rangle$ with probability 1, then if a state is rotated by θ, measurement yields $|0\rangle$ with probability $\cos^2(\theta)$.

We can setup the following protocol:

- if $X_A = 0$, then A measures in the standard basis, and outputs the result.
- if $X_A = 1$, then A measures in a basis rotated by $\pi/8$, and outputs the result.

- if $X_B = 0$, then B measures in the standard basis, and outputs the complement of the result.
- if $X_B = 1$, then B measures in a basis rotated by $-\pi/8$, and outputs the complement of the result.

Let us analyze the case of $X_A = 0, X_B = 1$. If A looks at it's bit first then it will measure her part of the Bell state in the standard basis and it will collapse. For example if Alice will measure 1 then the Bell state will collapse into $|10\rangle$. Then Bob will measure the second bit of this state in a Basis rotated by $-\pi/8$. A mistake will happen if B will output 0 because then $1 = a \oplus b \neq X_A \wedge X_B = 0$ and this will happen if Bob will measure 1. The state of the second bit $|0\rangle$ in the standard so the chance of measuring 0 in a basis rotated by $-\pi/8$ is as was explained $\cos^2(-\pi/8)$ so the chance of measuring 1 is $1 - cos^2(-\pi/8) = sin^2(\pi/8)$. so the chance of mistake in this case is $sin^2(\pi/8)$. The explanation for other cases and for $X_A = X_B = 1$ is similar.

If $X_A = X_B = 0$ then they both measure in the same base and because they measure a bell state they will measure opposite values, thus outputing the same value and then $0 = a \oplus b = X_A \wedge X_B = 0$ with certainty.

Now we calculate $Prob(a \oplus b \neq x_A \wedge x_B)$

$$Prob(a \oplus b \neq X_A \wedge X_B) = \sum_{X_A, X_B} Pr(a \oplus b \neq X_A \wedge X_B | X_A, X_B)$$

$$\begin{aligned}
Pr(a \oplus b \neq X_A \wedge X_B | X_A = 0, X_B = 0) &= 1/4 \times 0 = 0 \\
Pr(a \oplus b \neq X_A \wedge X_B | X_A = 0, X_B = 1) &= 1/4 \times \sin^2(\pi/8) \\
Pr(a \oplus b \neq X_A \wedge X_B | X_A = 1, X_B = 0) &= 1/4 \times \sin^2(\pi/8) \\
Pr(a \oplus b \neq X_A \wedge X_B | X_A = 1, X_B = 1) &= 1/4 \times 1/2
\end{aligned}$$

$$\begin{aligned}
Proba \oplus b \neq X_A \wedge X_B &= \frac{\big(2\sin^2(\pi/8) + 1/2\big)}{4} \\
&= \frac{(1 - \cos(2 * \pi/8) + 1/2)}{4} \\
&= \frac{(3/2 - \cos(\pi/4))}{4}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\left(3/2 - \sqrt{2}/2\right)}{4} \\
&\approx 1.6/8 = .2
\end{aligned}
$$

The error probability is smaller than any classical protocol can achieve, proving that the correlations between qubits are stronger than any possible correlations between classical bits.