

QUANTUM MECHANICS

SoS-2021

Janaki Ram Puli
200050112
Mentor-Nehal Mittal

June 2021

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 3 |
| 2 | Stern-Gerlach Experiments | 4 |
| 2.1 | Stern-Gerlach Experiment | 4 |
| 2.1.1 | Experiment 1 | 5 |
| 2.1.2 | Experiment 2 | 6 |
| 2.1.3 | Experiment 3 | 6 |
| 2.1.4 | Experiment 4 | 7 |
| 2.2 | Quantum State Vectors | 7 |
| 2.3 | General Quantum Systems | 10 |
| 3 | Operators and Measurement | 11 |
| 3.1 | Operators,Eigen values and Eigen Vectors | 11 |
| 3.2 | Matrix Representation of Operators | 12 |
| 3.3 | Diagonalization of Operators | 13 |
| 3.4 | Hermitian Operators | 14 |
| 3.5 | Projection Operators | 14 |
| 4 | Schrödinger Time evolution | 16 |
| 4.1 | Schrödinger Equation | 16 |
| 5 | Postulates | 18 |
| 6 | Quantum Spookiness | 19 |
| 6.1 | Einstein-Podolsky-Rosen Paradox | 19 |
| 6.2 | Schrödinger Cat Paradox | 20 |
| 7 | Cbits and Qubits | 22 |
| 7.1 | Cbits and their states | 22 |
| 7.2 | Qubits and their states | 22 |
| 7.3 | Bloch Sphere representation | 23 |
| 7.4 | Quantum logic gates | 23 |
| 7.4.1 | Pauli gates(X,Y,Z) | 23 |
| 7.4.2 | Controlled gates | 24 |
| 7.4.3 | Hadamard gate | 24 |
| 7.4.4 | Swap gate | 24 |
| 7.4.5 | Toffoli(CCNOT) gate | 24 |
| 7.4.6 | Fredkin (CSWAP) gate | 25 |
| 7.5 | Quantum Circuit | 26 |
| 7.6 | No-cloning theorem | 26 |
| 7.7 | Teleportation | 26 |
| 7.8 | Entanglement | 27 |
| 8 | Quantum Algorithms | 28 |
| 8.1 | Deutsch's Algorithm | 28 |
| 8.2 | The phase kickback trick | 29 |
| 8.3 | The Deutsch-Josza algorithm | 30 |
| 8.4 | Grover's algorithm | 31 |

Acknowledgement

I would like to express my special gratitude to my mentor, Nehal Mittal, and the Maths And Physics Club IIT Bombay, who gave me this golden opportunity to do this project on the topic of Quantum Mechanics. It helped me to do a lot of research and I came to learn a lot of things related to this topic. I have also improved my command over LaTeX during the course of this project.

1 Introduction

This report provides a summary of everything we covered related to Quantum Mechanics and an introduction to Quantum Computing. We will begin with some experiments which show the big difference between classical and quantum theories. Then we will cover some topics in Quantum Mechanics and then move to Cbits and Qubits and their differences, and some simple Quantum algorithms. Hopefully, throughout this report, one will be able to gain a good knowledge of Quantum Mechanics.

Quantum mechanics is a branch of physics that explores physical world at most fundamental level. At this level particle behave differently from classical world taking more than one state at the same time and interacting with other particles that are very far away. Phenomena like superposition and entanglement take place. Classical physics is still used in much of modern science and technology. However, towards the end of the 19th century, scientists discovered phenomena in both the large (macro) and the small (micro) worlds that classical physics could not explain.

Many aspects of quantum mechanics are counter intuitive and can seem paradoxical because they describe behavior quite different from that seen at larger scales. In the words of quantum physicist Richard Feynman, quantum mechanics deals with nature as She is—absurd. Even though there are many things that are highly confusing about quantum mechanics, the nice thing is that it's relatively easy to apply quantum mechanics to a physical system to figure out how it behaves.

For example, the uncertainty principle of quantum mechanics means that the more closely one pins down one measurement (such as the position of a particle), the less accurate another complementary measurement pertaining to the same particle (such as its speed) must become.

Another example is entanglement, in which a measurement of any two-valued state of a particle (such as light polarized up or down) made on either of two entangled particles that are very far apart causes a subsequent measurement on the other particle to always be the other of the two values (such as polarized in the opposite direction).

There are problems that even the most powerful classical computers are unable to solve because of their scale or complexity. Quantum computers may be uniquely suited to solve some of these problems because of their inherently quantum properties.

2 Stern-Gerlach Experiments

2.1 Stern-Gerlach Experiment

In 1922 Otto Stern and Walther Gerlach performed a experiment in the history of quantum mechanics. In its simplest form, the experiment consisted of an oven that produced a beam of neutral atoms, a region of space with an inhomogeneous magnetic field, and a detector for the atoms, as depicted in Figure 1. Stern and Gerlach used a beam of silver atoms and found that the beam was split into two in its passage through the magnetic field. One beam was deflected upwards and one downwards in relation to the direction of the magnetic field gradient.

We expect such an interaction if the particle possesses a magnetic moment μ . So the potential energy of this interaction will result in a force and deflects the beam in direction of magnetic field gradient.

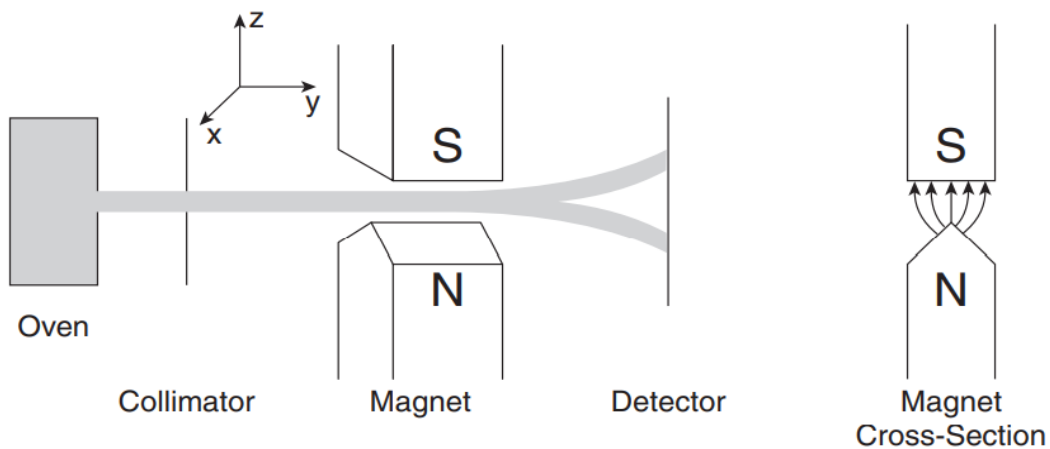


Figure 1: Stern-Gerlach experiment to measure the spin component of neutral particles along the z-axis. The magnet cross section at right shows the inhomogeneous field used in the experiment.

The deflection of the beam in the Stern-Gerlach experiment is thus a measure of the component (or projection) S_z of the spin along the z-axis, which is the orientation of the magnetic field gradient.

If we assume that the 5s electron of each atom has the same magnitude $|S|$ of the intrinsic angular momentum or spin, then classically we would write the z-component as $S_z = |S| \cos \theta$, where θ is the angle between the z-axis and the direction of the spin S . In the thermal environment of the oven, we expect a random distribution of spin directions and hence all possible angles θ . Thus we expect some continuous distribution (the details are not important) of spin components from $S_z = -|S|$ to $S_z = +|S|$, which would yield a continuous spread in deflections of the silver atomic beam. Rather, the experimental result that Stern and Gerlach observed was that there are only two deflections, indicating that there are only two possible values of the z-component of the electron spin. The magnitudes of these deflections are consistent with values of the spin component of

$$S_z = \pm \frac{\hbar}{2},$$

where \hbar is Planck's constant h divided by 2π .

This result of the Stern-Gerlach experiment is evidence of the quantization of the electron's spin angular momentum component along an axis. This quantization is at odds with our classical expectations for this measurement. The factor of $\frac{1}{2}$ in S_z leads us to refer to this as a **spin- $\frac{1}{2}$** system.

In Figure 2, the input and output beams are labeled with a new symbol called a ket. We use the ket $|+\rangle$ as a mathematical representation of the quantum state of the atoms that exit the

upper port corresponding to $S_z = +\hbar/2$. The lower output beam is labeled with the ket $|-\rangle$, which corresponds to $S_z = -\hbar/2$, and the input beam is labeled with the more generic ket $|\psi\rangle$. The kets are representations of the quantum states. They are used in mathematical expressions and they represent all the information that we can know about the state.

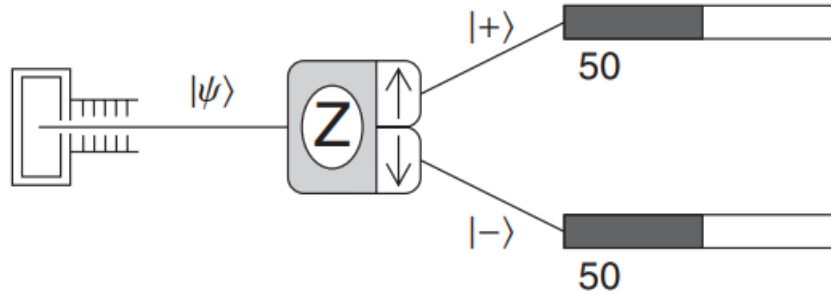


Figure 2: Simplified schematic of the Stern-Gerlach experiment, depicting a source of atoms, a Stern-Gerlach analyzer, and two counters.

Postulate 1

The state of a quantum mechanical system, including all the information you can know about it, is represented mathematically by a normalized ket $|\psi\rangle$.

2.1.1 Experiment 1

The atomic beam coming into the first Stern-Gerlach analyzer is split into two beams at the output, just like the original experiment. Now instead of counting the atoms in the upper output beam, the spin component is measured again by directing those atoms into the second Stern-Gerlach analyzer. The result of this experiment is that no atoms are ever detected coming out of the lower output port of the second Stern-Gerlach analyzer. All atoms that are output from the upper port of the first analyzer also pass through the upper port of the second analyzer. Thus we say that when the first Stern-Gerlach analyzer measures an atom to have a z-component of spin $S_z = +\hbar/2$, then the second analyzer also measures $S_z = +\hbar/2$ for that atom. This result is not surprising, but it sets the stage for results of experiments to follow.

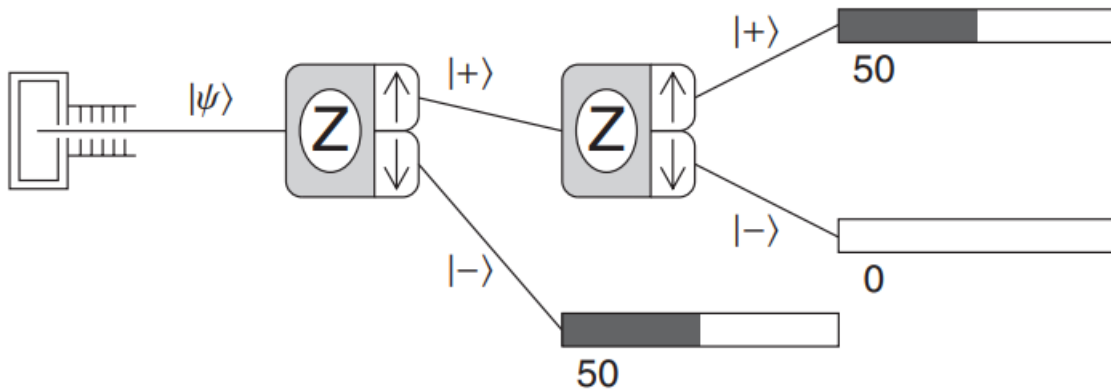


Figure 3: Experiment 1 measures the spin component along the z-axis twice in succession.

Thus our main focus in Experiment 1 is what happens at the second analyzer because we know that any atom entering the second analyzer is represented by the $|+\rangle$ ket prepared by the first analyzer. All the experiments we will describe employ a first analyzer as a state preparation device, though the SPINS program has a feature where the state of the atoms coming from the oven is determined but unknown, and the user can perform experiments to determine the unknown state using only one analyzer in the experiment.

2.1.2 Experiment 2

The second experiment is shown in Figure 4 and is identical to Experiment 1 except that the second Stern-Gerlach analyzer has been rotated by 90° to be aligned with the x-axis. Now the second analyzer measures the spin component along the x-axis rather the z-axis. Atoms input to the second analyzer are still represented by the ket $|+\rangle$. The result of this experiment is that atoms appear at both possible output ports of the second analyzer. Atoms leaving the upper port of the second analyzer have been measured to have $S_z = +\hbar/2$, and atoms leaving the lower port have the $S_z = -\hbar/2$.

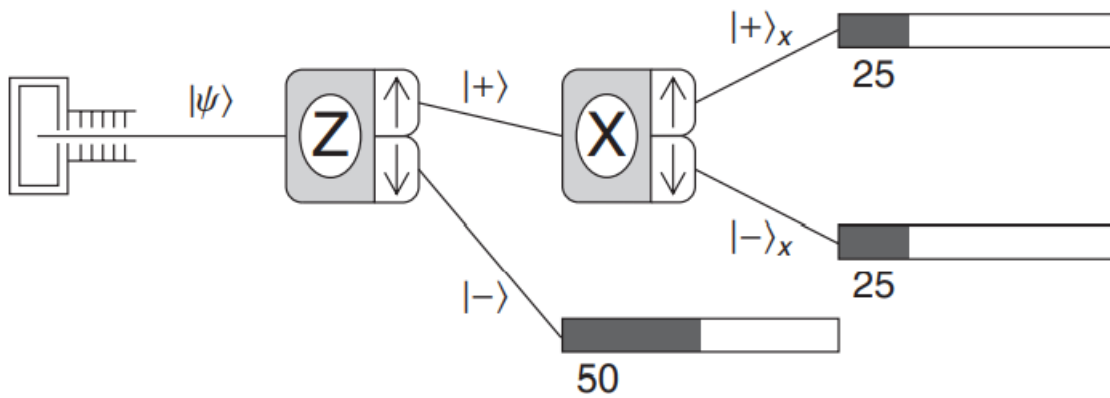


Figure 4: Experiment 2 measures the spin component along the z-axis and then along the x-axis.

A few items are noteworthy about this experiment. First, we notice that there are still only two possible outputs of the second Stern-Gerlach analyzer. The fact that it is aligned along a different axis doesn't affect the fact that we get only two possible results for the case of a spin-1/2 particle. Second, it turns out that the results of this experiment would be unchanged if we used the lower port of the first analyzer. *This probabilistic nature is at the heart of quantum mechanics.*

2.1.3 Experiment 3

Experiment 3, extends Experiment 2 by adding a third Stern-Gerlach analyzer aligned along the z-axis. Atoms entering the third analyzer have been measured by the first Stern-Gerlach analyzer to have spin component up along the z-axis, and by the second analyzer to have spin component up along the x-axis. The third analyzer then measures how many atoms have spin component up or down along the z-axis.

Classically, one would expect that the final measurement would yield the result spin up along the z-axis, because that was measured at the first analyzer. This result demonstrates another key feature of quantum mechanics: a measurement disturbs the system. The second analyzer has disturbed the system such that the spin component along the z-axis does not have a unique value, even though we measured it with the first analyzer.

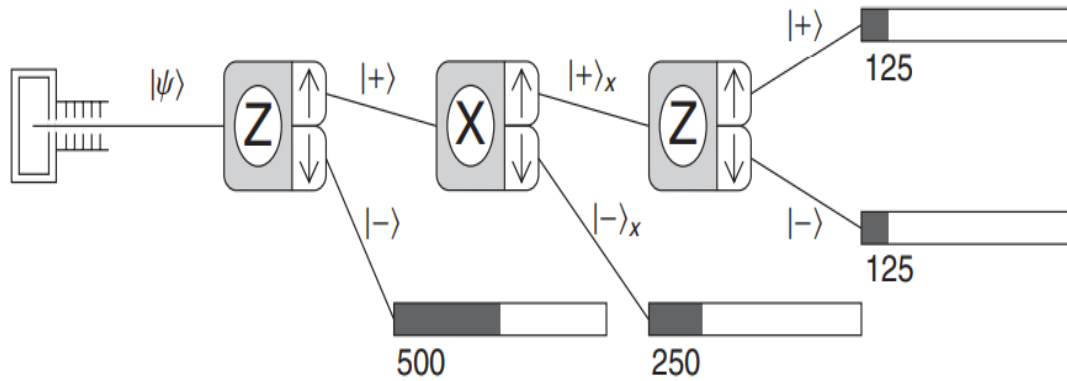


Figure 5: Experiment 3 measures the spin component three times in succession.

2.1.4 Experiment 4

Experiment 4a is identical to Experiment 3. In Experiment 4b, the upper beam of the second analyzer is blocked and the lower beam is sent to the third analyzer. In Experiment 4c, both beams are combined with our new method and sent to the third analyzer. It should be clear from our previous experiments that Experiment 4b has the same results as Experiment 4a. We now ask about the results of Experiment 4c. If we were to use classical probability analysis, then Experiment 4a would indicate that the probability for an atom leaving the first analyzer to take the upper path through the second analyzer and then exit through the upper port of the third analyzer is 25%, where we are now referring to the total probability for those two steps. Likewise, Experiment 4b would indicate that the total probability to take the lower path through the second analyzer and exit through the upper port of the third analyzer is also 25%. Hence the total probability to exit from the upper port of the third analyzer when both paths are available, which is Experiment 4c, would be 50%, and likewise for the exit from the lower port.

However, the quantum mechanical result in Experiment 4c is that all the atoms exit the upper port of the third analyzer and none exits the lower port. The atoms now appear to “remember” that they were initially measured to have spin up along the z-axis. By combining the two beams from the second analyzer, we have avoided the quantum mechanical disturbance that was evident in Experiments 3, 4a, and 4b. The result is now the same as Experiment 1, which means it is as if the second analyzer is not there.

To see how odd this is, look carefully at what happens at the lower port of the third analyzer. In this discussion, we refer to percentages of atoms leaving the first analyzer, because that analyzer is the same in all three experiments. In Experiments 4a and 4b, 50% of the atoms are blocked after the middle analyzer and 25% of the atoms exit the lower port of the third analyzer. In Experiment 4c, 100% of the atoms pass from the second analyzer to the third analyzer, yet fewer atoms come out of the lower port. In fact, no atoms make it through the lower port! So we have a situation where allowing more ways or paths to reach a counter results in fewer counts. It is as if you opened a second window in a room to get more sunlight and the room went dark!

2.2 Quantum State Vectors

Postulate 1 of quantum mechanics stipulates that kets are to be used for a mathematical description of a quantum mechanical system. These kets are abstract entities that obey many of the rules you know about ordinary spatial vectors. Hence they are called **quantum state vectors**. Quantum state vectors are part of a vector space that we call a Hilbert space. The dimensionality of the **Hilbert space** is determined by the physics of the system at hand. In the Stern-Gerlach example, the two possible results for a spin component measurement dictate that the vector space has only two dimensions. That makes this problem mathematically as simple as it can be, which is why we

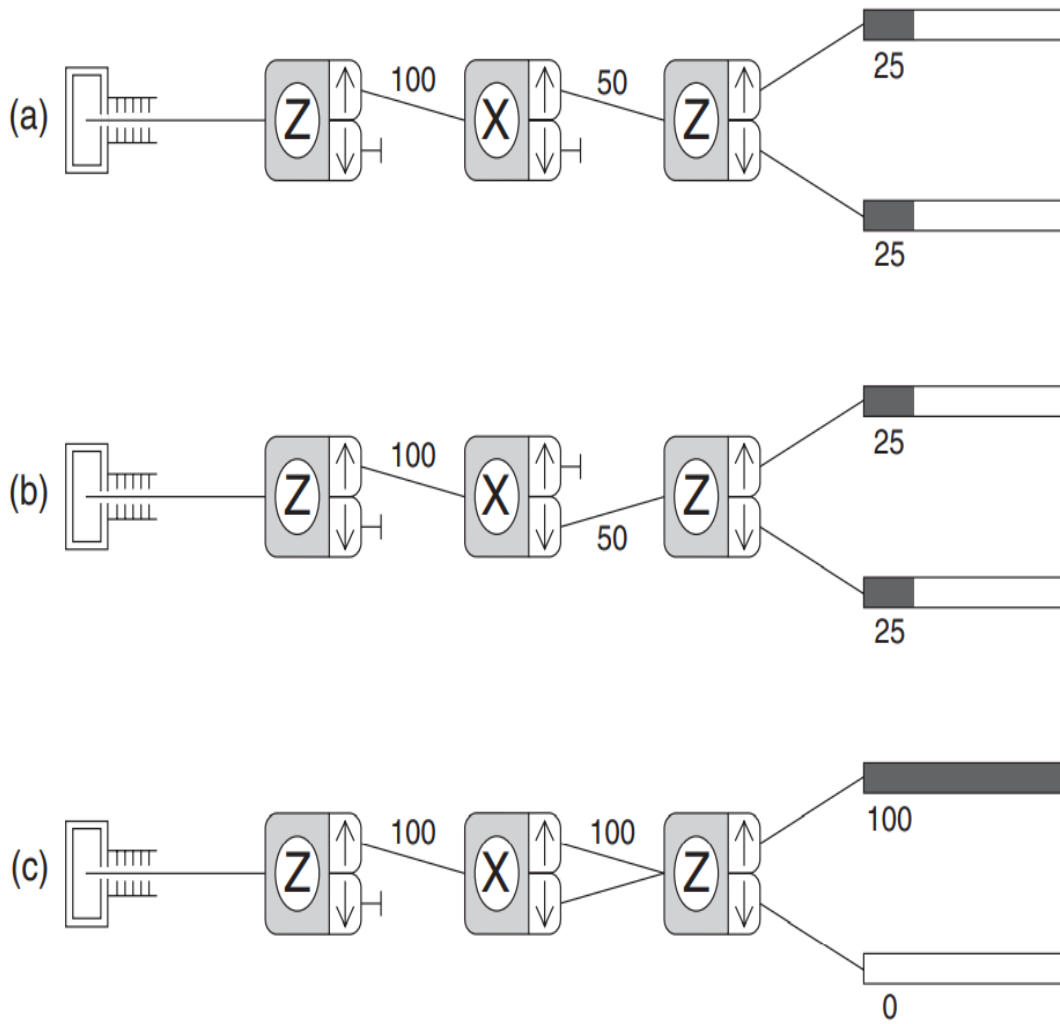


Figure 6: Experiment 4 measures the spin component three times in succession and uses (a and b) one or (c) two beams from the second analyzer.

have chosen to study it. Because the quantum state vectors are abstract, it is hard to say much about what they are, other than how they behave mathematically and how they lead to physical predictions.

Continuing the mathematical analogy between spatial vectors and abstract vectors, we require that these same properties (at least conceptually) apply to quantum mechanical basis vectors. For the S_z measurement, there are only two possible results, corresponding to the states $|+\rangle$ and $|-\rangle$, so these two states comprise a complete set of basis vectors. This basis is known as the S_z **basis**. We focus on this basis for now and refer to other possible basis sets later. The completeness of the basis kets $|\pm\rangle$ implies that a general quantum state vector $|\psi\rangle$ is a linear combination of the two basis kets:

$$|\psi\rangle = a|+\rangle + b|-\rangle,$$

where a and b are complex scalar numbers multiplying each ket. This addition of two kets yields another ket in the same abstract space. The complex scalar can appear either before or after the ket without affecting the mathematical properties of the ket (i.e., $a|+\rangle = |+\rangle a$). It is customary to use the Greek letter ψ (psi) for a general quantum state. You may have seen $\psi(x)$ used before as a quantum mechanical wave function. However, the state vector or ket $|\psi\rangle$ is not a wave function. Kets do not have any spatial dependence as wave functions do.

A similar approach is taken in quantum mechanics. The analog to the complex conjugated vector of classical physics is called a **bra** in the Dirac notation of quantum mechanics. Thus corresponding to a general ket $|\psi\rangle$ there is a bra, or bra vector, which is written as $\langle\psi|$. If a general $|\psi\rangle$ is specified as $|\psi\rangle = a|+\rangle + b|-\rangle$, then the corresponding bra $\langle\psi|$ is defined as

$$\langle\psi| = a^* \langle+| + b^* \langle-|$$

where the basis bras $\langle+|$ and $\langle-|$ correspond to the basis kets $|+\rangle$ and $|-\rangle$, respectively, and the coefficients a and b have been complex conjugated. The scalar product in quantum mechanics is defined as the product of a bra and a ket taken in the proper order—bra first, then ket second:

$$(\langle bra|)(| ket\rangle).$$

When the bra and ket are combined together in this manner, we get a bracket (bra ket) is written in shorthand as

$$\langle bra|ket\rangle$$

Consider the general state vector. Take the inner product of this ket $|\psi\rangle = a|+\rangle + b|-\rangle$ with the bra $\langle+|$ and obtain

$$\begin{aligned}\langle+|\psi\rangle &= \langle+|(a|+\rangle + b|-\rangle) \\ &= \langle+|a|+\rangle + \langle+|b|-\rangle \\ &= a\langle+|+\rangle + b\langle+|-\rangle \\ &= a\end{aligned}\tag{1}$$

using the properties that inner products are distributive and that scalars can be moved freely through bras or kets. Likewise, you can show that $\langle-|\psi\rangle = b$. Hence the coefficients multiplying the basis kets are simply the inner products or projections of the general state $|\psi\rangle$ along each basis ket, albeit in an abstract complex vector space rather than the concrete three-dimensional space of normal vectors.

$$\begin{aligned}\langle\psi|+\rangle &= \langle+|a^*|+\rangle + \langle-|b^*|+\rangle \\ &= a^*\langle+|+\rangle + b^*\langle-|+\rangle \\ &= a^*\end{aligned}\tag{2}$$

Thus, we see that an inner product with the states reversed results in a complex conjugation of the inner product:

$$\langle+|\psi\rangle = \langle\psi|+\rangle^*$$

This important property holds for any inner product. For example, the inner product of two general states is

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$$

Now we come to a new mathematical aspect of quantum vectors that differs from the use of vectors in classical mechanics. The rules of quantum mechanics (postulate 1) require that all state vectors describing a quantum system be normalized, not just the basis kets.

We now have a prescription for predicting the outcomes of the experiments we have been discussing. This basic rule of probabilities is why the rules of quantum mechanics require that all state vectors be properly normalized before they are used in any calculation of probabilities.

Postulate 4(Spin-1/2 system)

The probability of obtaining the value $\pm\hbar/2$ in a measurement of the observable S_z on a system in the state $|\psi\rangle$ is

$$P_{\pm} = |\langle\pm|\psi\rangle|^2,$$

where $|\pm\rangle$ is the basis of ket of S_z corresponding to the result $\pm\hbar/2$.

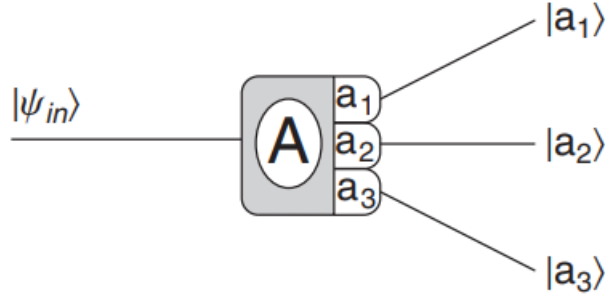


Figure 7: Generic depiction of the quantum mechanical measurement of observable A.

Because the quantum mechanical probability is found by squaring an inner product, we refer to an inner product, $\langle +|\psi \rangle$ for example, as a **probability amplitude** or sometimes just an **amplitude**; much like a classical wave intensity is found by squaring the wave amplitude. Note that the convention is to put the input or initial state on the right and the output or final state on the left: $\langle out|in \rangle$, so one would read from right to left in describing a problem. Because the probability involves the complex square of the amplitude, and $\langle out|in \rangle = \langle in|out \rangle^*$, this convention is not critical for calculating probabilities. Nonetheless, it is the accepted practice and is important in situations where several amplitudes are combined.

2.3 General Quantum Systems

The machinery we have developed for spin-1/2 systems can be generalized to other quantum systems. For example, if an observable A yields quantized measurement results a_n for some finite range of n, then we generalize the schematic depiction of a Stern-Gerlach measurement to a measurement of the observable A, as shown in Figure The observable A labels the measurement device and the possible results a_1, a_2, a_3 , etc. label the output ports. The basis kets corresponding to the results a_n are then $|a_n \rangle$. The mathematical rules about kets in this general case are

$$\langle a_i | a_j \rangle = \delta_{ij} \quad \text{orthonormality} \quad (3)$$

$$|\psi \rangle = \sum_i \langle a_i | \psi \rangle |a_i \rangle \quad \text{completeness} \quad (4)$$

where we use the **Kronecker delta**

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \quad (5)$$

to express the orthonormality condition compactly. In this case, the generalization of postulate 4 says that the probability of a measurement of one of the possible results a_n is

$$P_{a_n} = |\langle a_n | \psi_{in} \rangle|^2$$

3 Operators and Measurement

3.1 Operators, Eigen values and Eigen Vectors

The mathematical theory we developed in Chapter 2 used only quantum state vectors. We said that the state vector represents all the information we can know about the system and we used the state vectors to calculate probabilities. With each observable S_x, S_y , and S_z we associated a pair of kets corresponding to the possible measurement results of that observable. The observables themselves are not yet included in our mathematical theory, but the distinct association between an observable and its measurable kets provides the means to do so.

The role of physical observables in the mathematics of quantum theory is described by the two postulates listed below. Postulate 2 states that physical observables are represented by mathematical operators, in the same sense that physical states are represented by mathematical vectors or kets (postulate 1). An **operator** is a mathematical object that acts or operates on a ket and transforms it into a new ket, for example $A|\psi\rangle = |\phi\rangle$. However, there are special kets that are not changed by the operation of a particular operator, except for a possible multiplicative constant, which we know does not change anything measurable about the state. An example of a ket that is not changed by an operator would be $A|\psi\rangle = a|\psi\rangle$. Such kets are known as **eigenvectors** of the operator A and the multiplicative constants are known as the **eigenvalues** of the operator. These are important because postulate 3 states that the only possible result of a measurement of a physical observable is one of the eigenvalues of the corresponding operator.

Postulate 2

A physical observable is represented mathematically by an operator A that acts on kets.

Postulate 3

The only possible result of a measurement of an observable is one of the eigenvalues a_n of the corresponding operator A .

We now have a mathematical description of that special relationship we saw in Chapter 2 between a physical observable, S_z say, the possible results $\pm\hbar/2$, and the kets $|\pm\rangle$ corresponding to those results. This relationship is known as the **eigenvalue equation** and is depicted in Figure 8 for the case of the spin up state in the z-direction. In the eigenvalue equation, the observable is represented by an operator, the eigenvalue is one of the possible measurement results of the observable, and the eigenvector is the ket corresponding to the chosen eigenvalue of the operator. The eigenvector appears on both sides of the equation because it is unchanged by the operator.

The eigenvalue equations for the S_z operator in a spin-1/2 system are:

$$\begin{aligned} S_z |+\rangle &= +\frac{\hbar}{2} |+\rangle \\ S_z |-\rangle &= -\frac{\hbar}{2} |-\rangle \end{aligned} \tag{6}$$

These equations tell us that $+\hbar/2$ is the eigenvalue of S_z corresponding to the eigenvector $|+\rangle$ and $-\hbar/2$ is the eigenvalue of S_z corresponding to the eigenvector $|-\rangle$. Equation(6) are sufficient to define how the S_z operator acts mathematically on kets. However, it is useful to use matrix notation to represent operators. To determine the matrix representing the operator S_z , assume the most general form for a 2×2 matrix

operator eigenvalue

$$S_z |+\rangle = \frac{\hbar}{2} |+\rangle$$

 eigenvector

Figure 8: Eigenvalue equation for the spin up state.

$$S_z = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (7)$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = +\frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (8)$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (9)$$

Solving equations 8,9 results in:

$$a = +\frac{\hbar}{2} \quad b = 0 \quad (10)$$

$$c = 0 \quad d = -\frac{\hbar}{2} \quad (11)$$

Thus the matrix representation of the operator S_z is

$$S_z = \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (12)$$

Note two important features of this matrix: (1) it is a **diagonal matrix**—it has only diagonal elements—and (2) the diagonal elements are the eigenvalues of the operator, ordered in the same manner as the corresponding eigenvectors. In this example, the basis used for the matrix representation is that formed by the eigenvectors $|\pm\rangle$ of the operator S_z . That the matrix representation of the operator in this case is a diagonal matrix is a necessary and general result of linear algebra that will prove valuable as we study quantum mechanics. In simple terms, we say that an operator is always diagonal in its own basis. This special form of the matrix representing the operator is similar to the special form that the eigenvectors $|\pm\rangle$ take in this same representation—the eigenvectors are unit vectors in their own basis. These ideas cannot be overemphasized, so we repeat them:

An operator is always diagonal in its own basis.

Eigenvectors are unit vectors in their own basis.

3.2 Matrix Representation of Operators

Now consider how matrix representation works in general. Consider a general operator A describing a physical observable (still in the two-dimensional spin-1/2 system), which we represent by the

general matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (13)$$

in the S_z basis. The operation of A on the basis ket $|+\rangle$ yields

$$A|+\rangle = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} \quad (14)$$

The inner product of this new ket $A|+\rangle$ with the ket $|+\rangle$ (converted to a bra following the rules) results in

$$\langle +|A|+\rangle = (1 \ 0) \begin{pmatrix} a \\ c \end{pmatrix} = a. \quad (15)$$

which serves to isolate one of the elements of the matrix. Hence an individual element such as $\langle +|A|+\rangle$ or $\langle +|A|-\rangle$ is generally referred to as a matrix element. This “sandwich” of a bra, an operator, and a ket

$$\langle \text{bra} | \text{OPERATOR} | \text{ket} \rangle$$

plays an important role in many quantum mechanical calculations. Even in cases where the bra and ket are not basis kets, such as in $\langle \psi | A | \phi \rangle$, we still refer to this as a matrix element.

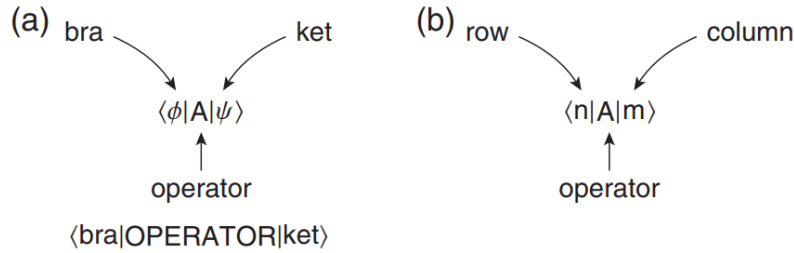


Figure 9: (a) Schematic diagram of a generic matrix element. (b) Schematic diagram of the row and column labeling convention for matrix elements

3.3 Diagonalization of Operators

In the case of the operator S_z above, we used the experimental results and the eigenvalue equations to find the matrix representation of the operator. It is more common to work the other way. That is, one is given the matrix representation of an operator and is asked to find the possible results of a measurement of the corresponding observable. According to the third postulate, the possible results are the eigenvalues of the operator, and the eigenvectors are the quantum states representing them. In the case of a general operator A in a two-state system, the eigenvalue equation is

$$A|a_n\rangle = a_n|a_n\rangle$$

where we have labeled the eigenvalues a_n and we have labeled the eigenvectors with the corresponding eigenvalues. In matrix notation, the eigenvalue equation is

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} c_{n1} \\ c_{n2} \end{pmatrix} = a_n \begin{pmatrix} c_{n1} \\ c_{n2} \end{pmatrix} \quad (16)$$

where c_{n1} and c_{n2} are the unknown coefficients of the eigenvector $|a_n\rangle$ corresponding to the eigenvalue a_n . This matrix equation yields the set of homogeneous equations

$$(A_{11} - a_n)c_{n1} + A_{12}c_{n2} = 0$$

$$A_{21}c_{n1} + (A_{22} - a_n)c_{n2} = 0.$$

The rules of linear algebra dictate that a set of homogeneous equations has solutions for the unknowns c_{n1} and c_{n2} only if the determinant of the coefficients vanishes:

$$\begin{vmatrix} A_{11} - a_n & A_{12} \\ A_{21} & A_{22} - a_n \end{vmatrix} = 0 \quad (17)$$

It is common notation to use the symbol I for the eigenvalues, in which case this equation is

$$\det(A - \lambda I) = 0$$

This procedure of finding the eigenvalues and eigenvectors of a matrix is known as diagonalization of the matrix and is the key step in many quantum mechanics problems. Generally, if we find a new operator, the first thing we do is diagonalize it to find its eigenvalues and eigenvectors. However, we stop short of the mathematical exercise of finding the matrix that transforms the original matrix to its new diagonal form. This would amount to a change of basis from the original basis to a new basis of the eigenvectors we have just found, much like a rotation in three dimensions changes from one coordinate system to another. We don't want to make this change of basis. In the example above, the S_y matrix is not diagonal, whereas the S_z matrix is diagonal, because we are using the S_z basis. It is common practice to use the S_z basis as the default basis, so you can assume that is the case unless you are told otherwise.

3.4 Hermitian Operators

So far we have defined how operators act upon kets. For example, an operator A acts on a ket $|\psi\rangle$ to produce a new ket $|\phi\rangle = A|\psi\rangle$. The operator acts on the ket from the left; if the operator is on the right of the ket, the result is not defined, which is clear if you try to use matrix representation. Similarly, an operator acting on a bra must be on the right side of the bra

$$\langle\epsilon| = \langle\psi| A$$

and the result is another bra. However, the bra $\langle\epsilon| = \langle\psi| A$ is not the bra $\langle\phi|$ that corresponds to the ket $|\phi\rangle = A|\psi\rangle$. Rather than bra $\langle\phi|$ is found by defining a new operator A^\dagger that obeys

$$\langle\phi| = \langle\psi| A^\dagger$$

This new operator A^\dagger is called the **Hermitian adjoint** of the operator A . Hermitian adjoint A^\dagger is found by transposing and complex conjugating the matrix representing A .

This is consistent with the definition of Hermitian adjoint used in matrix algebra. An operator A is said to be Hermitian if it is equal to its Hermitian adjoint A^\dagger . In quantum mechanics, all operators that correspond to physical observables are Hermitian. The Hermiticity of physical observables is important in light of two features of Hermitian matrices:

- Hermitian matrices have real eigenvalues, which ensures that results of measurements are always real.
- The eigenvectors of a Hermitian matrix comprise a complete set of basis states, which ensures that we can use the eigenvectors of any observable as a valid basis.

3.5 Projection Operators

$$\begin{aligned} |+\rangle\langle+| + |-\rangle\langle-| &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \quad (18)$$

Now consider the individual operators $|+\rangle\langle+|$ and $|-\rangle\langle-|$. These operators are called **projection operators**. In terms of these new operators the completeness relation can also be written as

$$P_+ + P_- = 1$$

When a projection operator for a particular eigenstate acts on a state $|\psi\rangle$, it produces a new ket that is aligned along the eigenstate and has a magnitude equal to the amplitude (including the phase) for the state $|\psi\rangle$ to be in that eigenstate.

Because the projection operator produces the probability amplitude, we expect that it must be intimately tied to measurement in quantum mechanics.

The projection postulate is at the heart of quantum measurement. This effect is often referred to as the **collapse (or reduction or projection)** of the quantum state vector. The projection postulate clearly states that quantum measurements cannot be made without disturbing the system (except in the case where the input state is the same as the output state), in sharp contrast to classical measurements. The collapse of the quantum state makes quantum mechanics irreversible, again in contrast to classical mechanics.

Postulate 5

After a measurement of A that yields the result a_n , the quantum system is in a new state that is the normalized projection of the original system ket onto the ket (or kets) corresponding to the result of the measurement:

$$|\psi'\rangle = \frac{P_n |\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}}$$

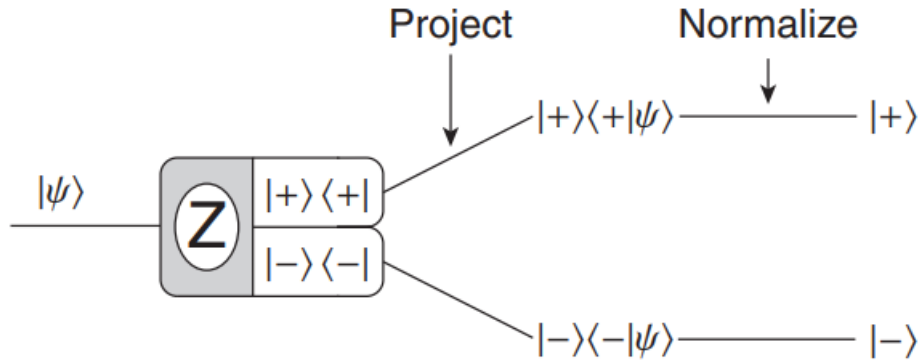


Figure 10: Schematic diagram of the role of the projection operator in a Stern-Gerlach spin measurement.

We do not really know what is going on in the measurement process, so we cannot explain the mechanism of the collapse of the quantum state vector. This lack of understanding makes some people uncomfortable with this aspect of quantum mechanics and has been the source of much controversy surrounding quantum mechanics. Trying to better understand the measurement process in quantum mechanics is an ongoing research problem. However, despite our lack of understanding, the theory for predicting the results of experiments has been proven with very high accuracy.

4 Schrödinger Time evolution

In this we marks our final step in developing the mathematical basis of a quantum theory. The key missing aspect is the ability to predict the future. Physics theories are judged on their predictive power. Classical mechanics relies on Newton's second law $F = ma$ to predict the future of a particle's motion. The ability to predict the quantum future started with Erwin Schrödinger and bears his name.

4.1 Schrödinger Equation

The sixth postulate of quantum mechanics says that the time evolution of a quantum system is governed by the differential equation

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

where the operator H corresponds to the total energy of the system and is called the **Hamiltonian** operator of the system because it is derived from the classical Hamiltonian. This equation is known as the **Schrödinger equation**.

Postulate 6

The time evolution of a quantum system is determined by the Hamiltonian or total energy operator $H(t)$ through the Schrödinger equation

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

The Hamiltonian is a new operator, but we can use the same ideas we developed to understand its basic properties. The Hamiltonian H is an observable, so it is a Hermitian operator. The eigenvalues of the Hamiltonian are the allowed energies of the quantum system, and the eigenstates of H are the energy eigenstates of the system. If we label the allowed energies as E_n , then the energy eigenvalue equation is

$$H |E_n\rangle = E_n |E_n\rangle$$

The eigenvectors of the Hamiltonian form a complete basis because the Hamiltonian is an observable, and therefore a Hermitian operator. Because H is the only operator appearing in the Schrödinger equation, it would seem reasonable (and will prove invaluable) to consider the energy eigenstates as the basis of choice for expanding general state vectors:

$$|\psi(t)\rangle = \sum_n c_n(t) |E_n\rangle$$

The basis of eigenvectors of the Hamiltonian is also orthonormal, so

$$\langle E_k | E_n \rangle = \delta_{kn}$$

We refer to this basis as the energy basis. For now, we assume that the Hamiltonian is time independent. Thus if a general state $|\psi\rangle$ is to be time dependent, as the Schrödinger equation implies, then the time dependence must reside in the expansion coefficients $c_n(t)$. Substitute general state into the Schrödinger equation

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

and using eigenvalue equation to obtain

$$i\hbar \sum_n \frac{dc_n(t)}{dt} |E_n\rangle = \sum_n c_n(t) E_n |E_n\rangle$$

Each side of this equation is a sum over all the energy states of the system. To simplify this equation, we isolate single terms in these two sums by taking the inner product of the ket on each side with one particular ket $|E_k\rangle$ (this ket can have any label k , but must not have the label n that is already used in the summation). The orthonormality condition $\langle E_k|E_n\rangle = \delta_{kn}$ then collapse the sums:

$$\begin{aligned}
 \langle E_k|i\hbar \sum_n \frac{dc_n(t)}{dt}|E_n\rangle &= \langle E_k|\sum_n c_n(t)E_n|E_n\rangle \\
 i\hbar \sum_n \frac{dc_n(t)}{dt} \langle E_k|E_n\rangle &= \sum_n c_n(t)E_n \langle E_k|E_n\rangle \\
 i\hbar \sum_n \frac{dc_n(t)}{dt} \delta_{kn} &= \sum_n c_n(t)E_n \delta_{kn} \\
 i\hbar \frac{dc_k(t)}{dt} &= c_k(t)E_k
 \end{aligned} \tag{19}$$

We are left with a single differential equation for each of the possible energy states of the systems $k = 1, 2, 3, \dots$. This first-order differential equation can be rewritten as

$$\frac{dc_k(t)}{dt} = -i\frac{E_k}{\hbar}c_k(t).$$

A solution for this is a complex exponential

$$c_k(t) = c_k(0)e^{-iE_k t/\hbar}$$

Here we have denoted the initial condition as $c_k(0)$, but we denote it simply as c_k hereafter. Each coefficient in the energy basis expansion of the state obeys the same form of the time dependence in above equation, but with a different exponent due to the different energies. The time-dependent solution for the full state vector is summarized by saying that if the initial state of the system at time $t = 0$ is

$$|\psi(0)\rangle = \sum_n c_n |E_n\rangle$$

then the time evolution of this state under the action of the time-independent Hamiltonian H is

$$\boxed{|\psi(t)\rangle = \sum_n c_n e^{-iE_n t/\hbar} |E_n\rangle}$$

So the time dependence of the original state vector is found by multiplying each energy eigenstate coefficient by its own phase factor $e^{-iE_n t/\hbar}$ that depends on the energy of that eigenstate. Note that the factor E/\hbar is an angular frequency, so that the time dependence is of the form $e^{-i\omega t}$, a form commonly found in many areas of physics. It is important to remember that one must use the energy eigenstates for the expansion in order to use the simple phase factor multiplication to account for the Schrödinger time evolution of the state. This key role of the energy basis accounts for the importance of the Hamiltonian operator and for the common practice of finding the energy eigenstates to use as the preferred basis.

5 Postulates

We have introduced all postulates of quantum mechanics. The postulates of quantum mechanics dictate how to treat a quantum mechanical system mathematically and how to interpret the mathematics to learn about the physical system in question. These postulates cannot be proven, but they have been successfully tested by many experiments, and so we accept them as an accurate way to describe quantum mechanical systems. New results could force us to reevaluate these postulates at some later time. All six postulates are listed below to give an idea where we are headed and a framework into which you can place the new concepts as we confront them.

Postulates of Quantum Mechanics

1. The state of a quantum mechanical system, including all the information you can know about it, is represented mathematically by a normalized ket $|\psi\rangle$.
2. A physical observable is represented mathematically by an operator A that acts on kets.
3. The only possible result of a measurement of an observable is one of the eigenvalues of the corresponding operator A .
4. The probability of obtaining the eigenvalue a_n in a measurement of the observable A on the system in the state $|\psi\rangle$ is

$$P_{a_n} = |\langle a_n | \psi \rangle|^2,$$

where $|a_n\rangle$ is the normalized eigenvector of A corresponding to the eigenvalue a_n .

5. After a measurement of A that yields the result a_n , the quantum system is in a new state that is the normalized projection of the original system ket onto the ket (or kets) corresponding to the result of the measurement:

$$|\psi'\rangle = \frac{P_n |\psi\rangle}{\sqrt{\langle \psi | P_n | \psi \rangle}}$$

6. The time evolution of a quantum system is determined by the Hamiltonian or total energy operator $H(t)$ through the Schrödinger equation

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

.

6 Quantum Spookiness

As we have seen in the previous chapters, many aspects of quantum mechanics run counter to our physical intuition, which is formed from our experience living in the classical world. The probabilistic nature of quantum mechanics does not agree with the certainty of the classical world—we have no doubt that the sun will rise tomorrow. Moreover, the disturbance of a quantum mechanical system through the action of measurement makes us part of the system, rather than an independent observer. These issues and others make us wonder what is really going on in the quantum world. As quantum mechanics was being developed in the early twentieth century, many of the world's greatest physicists debated the true meaning of quantum mechanics.

6.1 Einstein-Podolsky-Rosen Paradox

Albert Einstein was never comfortable with quantum mechanics. He is famously quoted as saying Gott würfelt nicht or God does not play dice, to express his displeasure with the probabilistic nature of quantum mechanics. But his opposition to quantum mechanics ran deeper than that. He felt that properties of physical objects have an objective reality independent of their measurement, much as Erwin felt that his socks were black or white, or long or short, independent of his pulling them out of the drawer. In quantum mechanics, we cannot say that a particle whose spin is measured to be up had that property before the measurement. It may well have been in a superposition state. Moreover, we can only know one spin component of a particle, because measurement of one component disturbs our knowledge of the other components. Because of these apparent deficiencies, Einstein believed that quantum mechanics was an *incomplete description of reality*.

In 1935, Einstein, Boris Podolsky, and Nathan Rosen published a paper presenting a gedanken experiment designed to expose the shortcomings of quantum mechanics. The **EPR Paradox** (Einstein-Podolsky-Rosen) tries to paint quantum mechanics into a corner and expose the **absurd** behavior of the theory. The essence of the argument is that if you believe that measurements on two widely separated particles cannot influence each other, then the quantum mechanics of an ingeniously prepared two-particle system leads you to conclude that the physical properties of each particle are really there—they are **elements of reality** in the authors' words.

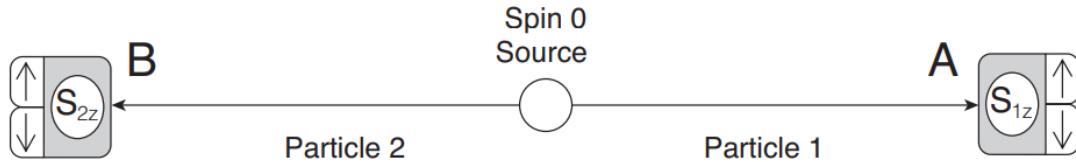


Figure 11: Einstein-Podolsky-Rosogen *gedanken* experiment.

The experimental situation is depicted in figure. An unstable particle with spin 0 decays into two spin-1/2 particles, which by conservation of angular momentum must have opposite spin components and by conservation of linear momentum must travel in opposite directions. For example, a neutral pi meson decays into an electron and a positron. Observers A and B are on opposite sides of the decaying particle and each has a Stern-Gerlach apparatus to measure the spin component of the particle headed in its direction. Whenever one observer measures spin up along a given direction, then the other observer measures spin down along that same direction. The quantum state of this two-particle system is

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|1\rangle_1 |-\rangle_2 - |-\rangle_1 |+\rangle_2) \quad (20)$$

As shown in Figure 11 observer A measures the spin component of particle 1 and observer B measures the spin component of particle 2. The probability that observer A measures particle 1 to be spin up is 50% and the probability for spin down is 50%. The 50-50 split is the same for observer

B. For a large ensemble of decays, each observer records a random sequence of spin up and spin down results, with a 50/50 ratio. But, because of the correlation between the spin components of the two particles, if observer A measures spin up (i.e., $S_{1z} = +\hbar/2$), then we can predict with 100% certainty that the result of observer B's measurement will be spin down ($S_{2z} = -\hbar/2$). The result is that even though each observer records a random sequence of ups and downs, the two sets of results are perfectly anticorrelated. The state ψ in Equation 20 that produces this strange mixture of random and correlated measurement results is known as an **entangled state**. The spins of the two particles are entangled with each other and produce this perfect correlation between the measurements of observer A and observer B.

Imagine that the two observers are separated by a large distance, with observer B slightly farther from the decay source than observer A. Once observer A has made the measurement $S_{1z} = +\hbar/2$, we know that the measurement by observer B in the next instant will be spin down $S_{2z} = -\hbar/2$. We conclude that the state ψ in Equation 20 instantaneously collapses onto the state $|0+91\ 0-92\rangle$, and the measurement by observer A has somehow determined the measurement result of observer B. Einstein referred to this as spooky action at a distance. The result that observer B records is still random, it is just that its randomness is perfectly anticorrelated with observer A's random result.

The EPR argument contends that because we can predict a measurement result with 100% certainty (e.g., $S_z = -\hbar/2$), then that result must be a real property of the particle - it must be an element of reality. Because the particles are widely separated, this element of reality must be independent of what observer A does, and hence, must have existed all along. The independence of the elements of reality of the two particles is called **Einstein's locality principle**, and is a fundamental assumption of the EPR argument.

6.2 Schrödinger Cat Paradox

The Schrödinger cat paradox is an experiment designed by Schrödinger to illustrate some of the problems of quantum measurement, particularly in the extension of quantum mechanics to classical systems. The apparatus of Schrödinger's experiment consists of a radioactive nucleus, a Geiger counter, a hammer, a bottle of cyanide gas, a cat, and a box, as shown in Figure 12. The nucleus has a 50% probability of decaying in one hour. The components are assembled such that when the nucleus decays, it triggers the Geiger counter, which causes the hammer to break the bottle and release the poisonous gas, killing the cat. Thus, after one hour there is a 50% probability that the cat is dead.

After the one hour, the nucleus is in an equal superposition of undecayed and decayed states:

$$|\psi_{nucleus}\rangle = \frac{1}{\sqrt{2}} (|\psi_{undecayed}\rangle + |\psi_{decayed}\rangle) \quad (21)$$

The apparatus is designed such that there is a one-to-one correspondence between the undecayed nuclear state and the live-cat state and a one-to-one correspondence between the decayed nuclear state and the dead-cat state. Though the cat is macroscopic, it is made up of microscopic particles and so should be describable by a quantum state, albeit a complicated one. Thus, we expect that the quantum state of the cat after one hour is

$$|\psi_{cat}\rangle = \frac{1}{\sqrt{2}} (|\psi_{alive}\rangle + |\psi_{dead}\rangle) \quad (22)$$

Both quantum calculations and classical reasoning would predict 50-50 probabilities of observing an alive or a dead cat when we open the box. However, quantum mechanics would lead us to believe that the cat was neither dead nor alive before we opened the box, but rather was in a superposition of states, and the quantum state collapses to the alive state $|\psi_{alive}\rangle$ or dead state $|\psi_{dead}\rangle$ only when we open the box and make the measurement by observing the cat. But our classical experiences clearly run counter to this. We would say that the cat really was dead or alive, we just did not know it yet. (Imagine that the cat is wearing a cyanide sensitive watch—the time will tell us when the cat was killed, if it is dead!)

The **Copenhagen interpretation** of quantum mechanics championed by Bohr and Heisenberg maintains that there is a boundary between the classical and quantum worlds. We describe

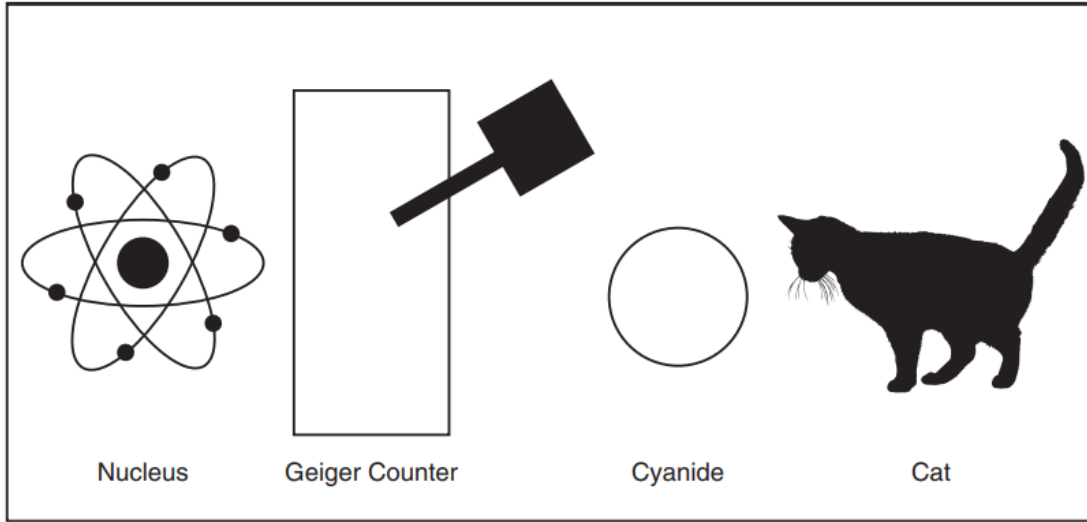


Figure 12: Schrödinger cat experiment.

microscopic systems (the nucleus) with quantum states and macroscopic systems (the cat, or even the Geiger counter) with classical rules. The measurement apparatus causes the quantum state to collapse and to produce the single classical or meter result. The actual mechanism for the collapse of the wave function is not specified in the Copenhagen interpretation, and where to draw the line between the classical and the quantum world is not clear. Others have argued that the human consciousness is responsible for collapsing the wave function, while some have argued that there is no collapse, just bifurcation into alternate, independent universes. Many of these different points of view are untestable experimentally and thus raise more metaphysical than physical questions.

These debates about the interpretation of quantum mechanics arise when we use words, which are based on our classical experiences, to describe the quantum world. The mathematics of quantum mechanics is clear and allows us to calculate precisely. No one is disagreeing about the probability that the cat will live or die. The disagreement is all about “what it really means!” To steer us toward the clear mathematics, Richard Feynman admonished us to “Shut up and calculate!” Two physicists who disagree on the words they use to describe a quantum mechanical experiment generally agree on the mathematical description of the results.

Recent advances in experimental techniques have allowed experiments to probe the boundary between the classical and quantum worlds and address the quantum measurement issues raised by the Schrödinger cat paradox. The coupling between the microscopic nucleus and the macroscopic cat is representative of a quantum measurement whereby a classical meter (the cat) provides a clear and unambiguous measurement of the state of the quantum system (the nucleus). In this case, the two possible states of the nucleus (undecayed or decayed) are measured by the two possible positions on the meter (cat alive or cat dead). The quantum mechanical description of this complete system is the entangled state

$$|\psi_{system}\rangle = \frac{1}{\sqrt{2}} (|\psi_{undecayed}\rangle |\psi_{alive}\rangle + |\psi_{decayed}\rangle |\psi_{dead}\rangle) \quad (23)$$

7 Cbits and Qubits

7.1 Cbits and their states

A classical computer operates on strings of zeroes and ones, such as 1110011001010, converting them into other such strings. Each position in such a string is called a *bit*, and it contains either a 0 or a 1. To represent such collection of bits the computer must contain a corresponding collection of physical systems, each of which can exist in two unambiguously distinguishable physical states, associated with the value (0 or 1) of the abstract bit that the physical system represents. Such a physical system could be, for example, a switch that could be open (0) or shut (1), or a magnet whose magnetization could be oriented in two different directions, up (0) or down (1).

We shall represent the state of each Cbit as a kind of box, depicted by the symbol $|\rangle$, into which we place the value, 0 or 1, represented by that state. Thus the two distinguishable states of a Cbit are represented by the symbols $|0\rangle$ and $|1\rangle$.

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

In the case of two Cbits the vector space is four-dimensional, with an orthonormal basis

$$|00\rangle, |01\rangle, |10\rangle, |11\rangle.$$

7.2 Qubits and their states

In quantum computing, a Qubit or Quantum bit is the basic unit of quantum information - the quantum version of the classic binary bit. A quantum state is represented as a ray in an abstract linear vector space known as the Hilbert's space. The only vectors with any classical meaning in the whole two-dimensional vector space are the two orthonormal vectors $|0\rangle$ and $|1\rangle$, since those are the only two states a Cbit can have. The state $|\psi\rangle$ associated with a Qubit can be any unit vector in the two-dimensional vector space spanned by $|0\rangle$ and $|1\rangle$ over the complex numbers. The general state of a Qubit is

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

where α and β are two complex numbers constrained only by the requirement that $|\psi\rangle$, like $|0\rangle$ and $|1\rangle$, should be a unit vector in the complex vector space. The state $|\psi\rangle$ is said to be *superposition* of the states $|0\rangle$ and $|1\rangle$ with *amplitudes* α and β .

Suppose we have two qubits. If these were two classical bits, then there would be four possible states, 00, 01, 10, and 11. Correspondingly, a two qubit system has four *computational basis states* denoted by $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. A pair of qubits can also exist in superpositions of these four states, so the quantum state of two qubits involves associating a complex coefficient - sometimes called an amplitude - with each computational basis state, such that the state vector describing the two qubits is

$$|\psi\rangle = \alpha_{00} |00\rangle + \alpha_{01} |01\rangle + \alpha_{10} |10\rangle + \alpha_{11} |11\rangle$$

We can write the computational basis as direct or tensor product of two single qubits as shown: $|00\rangle = |0\rangle \otimes |0\rangle$. Each of the two qubits can be measured separately. There are some special two qubit states which cannot be written as a product of two single qubits known as the Bell states. An example is:

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

If we measure one of the qubits, the state of the other qubit is determined without performing a measurement. These are called entangled states.

7.3 Bloch Sphere representation

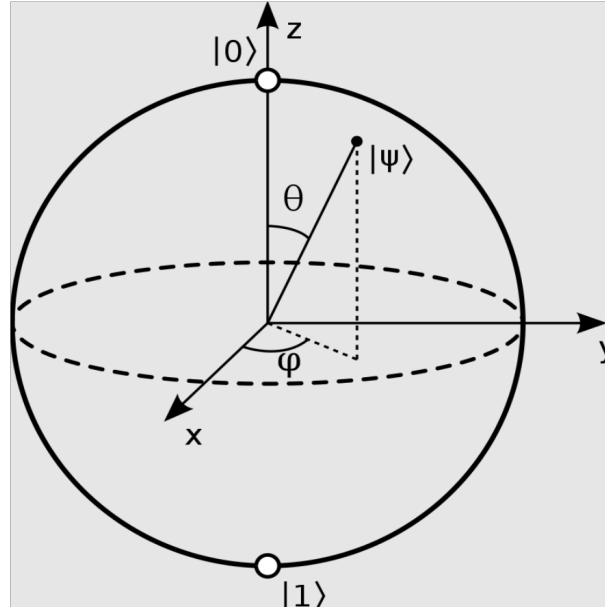


Figure 13: Bloch sphere

The Bloch Sphere is a sphere with a radius of one and a point on its surface represents the state of a qubit. The basis as $|0\rangle$ and $|1\rangle$ and their linear combinations $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ describe the state of a single qubit. But because the coefficients α and β are not just real numbers, but can be imaginary or even complex, visualizing a qubit requires a special tool called the Bloch Sphere. We also know from quantum mechanics that total probability of the system has to be one. Given this constraint, we can write $|\psi\rangle$ using the following representation:

$$|\psi\rangle = \cos \theta/2 |0\rangle + e^{i\phi} \sin \theta/2$$

where $0 \leq \theta \leq \pi$ and $0 \leq \phi < 2\pi$.

7.4 Quantum logic gates

In quantum computing and specifically the quantum circuit model of computation, a quantum logic gate (or simply quantum gate) is a basic quantum circuit operating on a small number of qubits. They are the building blocks of quantum circuits, like classical logic gates are for conventional digital circuit. Unlike many classical logic gates, quantum logic gates are reversible. Quantum gates are unitary operators, and are described as unitary matrices relative to some basis.

7.4.1 Pauli gates(X,Y,Z)

The Pauli gates(X,Y,Z) are the three Pauli matrices ($\sigma_x, \sigma_y, \sigma_z$) and act on a single qubit. The Pauli X,Y and Z equate, respectively, to a rotation around the x,y and z axes of the Bloch sphere by π radians.

The Pauli-X gate is the quantum equivalent of the NOT gate for classical computers with respect to the standard basis $|0\rangle, |1\rangle$ which distinguishes the z-axis on the Bloch sphere. It is sometimes called a bit-flip as it maps $|0\rangle$ to $|1\rangle$ and $|1\rangle$ to $|0\rangle$. Similarly, the Pauli-Y maps $|0\rangle$ to $i|1\rangle$ and $|1\rangle$ to $-i|0\rangle$. Pauli Z leaves the basis state $|0\rangle$ unchanged and maps $|1\rangle$ to $-|1\rangle$. Due to this nature, it is sometimes called phase-flip.

These matrices are usually represented as

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

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

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

The Pauli matrices are involutory, meaning that the square of a Pauli matrix is the identity matrix.

$$I^2 = X^2 = Y^2 = Z^2 = -iXYZ = I$$

7.4.2 Controlled gates

Controlled gates act on 2 or more qubits, where one or more qubits act as a control for some operation. For example, the controlled NOT gate (or CNOT or CX) acts on 2 qubits, and performs the NOT operation on the second qubit only when the first qubit is $|1\rangle$, and otherwise leaves it unchanged. With respect to the basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$, it is represented by the matrix:

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

7.4.3 Hadamard gate

The Hadamard gate acts on a single qubit. It maps the basis state $|0\rangle$ to $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $|1\rangle$ to $\frac{|0\rangle-|1\rangle}{\sqrt{2}}$, which means that a measurement will have equal probabilities to result in 1 or 0. It represents rotation of π about $(\hat{x} + \hat{z})/\sqrt{2}$ at the Bloch sphere. It is represented by the Hadamard matrix:

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

7.4.4 Swap gate

The swap gate swaps two qubits. With respect to the basis $|00\rangle, |01\rangle, |10\rangle, |11\rangle$, it is represented by the matrix:

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

7.4.5 Toffoli(CCNOT) gate

The Toffoli gate, also called CCNOT gate is a 3-bit gate, which is universal for classical computation but not for quantum computation. The quantum Toffoli gate is the same gate, defined for 3 qubits. If we limit ourselves to only accepting input qubits that are $|0\rangle$ and $|1\rangle$, then if the first two bits are in the state $|11\rangle$ it applies a Pauli-X (or NOT) on the third bit, else it does nothing. It is an example of a controlled gate. Since it is the quantum analog of a classical gate, it is completely specified by its truth table. It is represented in matrix form as:

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

7.4.6 Fredkin (CSWAP) gate

The Fredkin gate (also CSWAP), is a 3-bit gate that performs a controlled swap. It is universal for classical computation. It has the useful property that the numbers of 0s and 1s are conserved throughout, which in the billiard ball model means the same number of balls are output as input. It is represented in matrix form as:

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

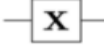

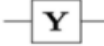
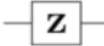
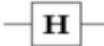
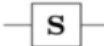
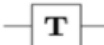
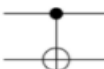
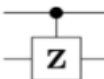



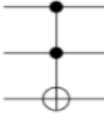
| Operator | Gate(s) | Matrix |
|-----------------------------------|---|--|
| Pauli-X (X) |   | $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ |
| Pauli-Y (Y) |  | $\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$ |
| Pauli-Z (Z) |  | $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ |
| Hadamard (H) |  | $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ |
| Phase (S, P) |  | $\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$ |
| $\pi/8$ (T) |  | $\begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$ |
| Controlled Not (CNOT, CX) |  | $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ |
| Controlled Z (CZ) |   | $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$ |
| SWAP |   | $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ |
| Toffoli (CCNOT, CCX, TOFF) |  | $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$ |

Figure 14: Common quantum logic gates by name (including abbreviation), circuit form(s) and the corresponding unitary matrices.

7.5 Quantum Circuit

A quantum circuit contains logic gates connected by straight lines, which don't represent physical wires but indicates the direction of logic flow with time, earlier time being to the left. Classical computer circuits consist of wires and logic gates. The wires are used to carry information around the circuit, while the logic gates perform manipulations of the information, converting it from one form to another. Quantum Circuits are similar with few important points. Inputs to the circuits are qubits, as are the outputs. Unlike classical circuits, it do not allow looping. Looping in the circuit or Fan-inns are not permitted. (Fan-out being a copying circuit is illegal and Fan-in being it's inverse is ruled out by reversibility.)

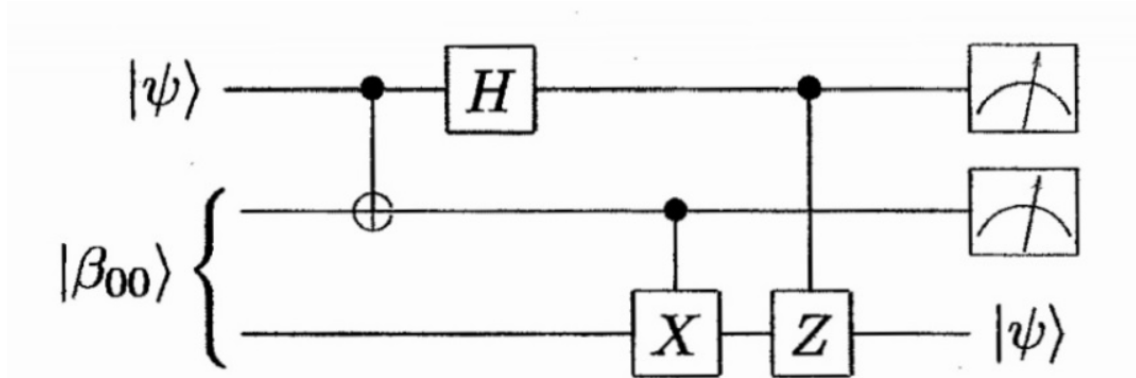


Figure 15: Example of a quantum circuit

7.6 No-cloning theorem

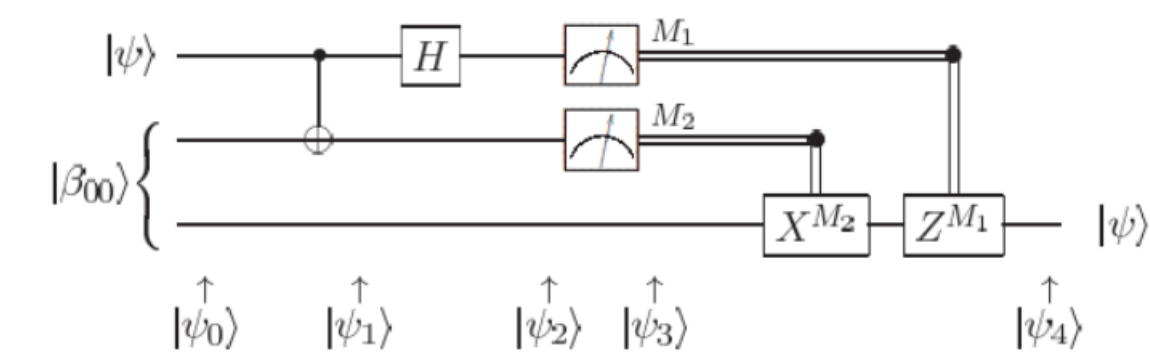
The theorem states that it is impossible to create an exact copy of an unknown quantum state. In other words, we cannot find an unitary matrix U such that

$$|\psi\rangle \otimes |s\rangle \xrightarrow{U} U(|\psi\rangle \otimes |s\rangle) = |\psi\rangle \otimes |\psi\rangle.$$

where $|\psi\rangle$ is the state to be copied and $|s\rangle$ is some input state of target qubit. It can be shown that for $|\psi\rangle$ to be copied, $|\psi\rangle$ must be either $|0\rangle$ or $|1\rangle$. An arbitrary superposition cannot be copied. It is impossible to create an independent and identical copy of an arbitrary unknown quantum state. Alice and Bob have an entangled pair of qubits $|\beta_{00}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$. Now Alice keeps one of the qubit while Bob takes the other qubit to some place separated physically.

7.7 Teleportation

Teleportation is the technique of sending quantum state from one point in space to another. Let us assume Alice has a quantum state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ which she needs to send to Bob. Alice don't know the details about the state. The circuit depicting this is:



The first two lines are the qubits with Alice while the last line is the qubit with Bob. The starting state is

$$|\psi_0\rangle = |\psi\rangle |\beta_{00}\rangle$$

Alice now applies a CNOT gate on her qubits to get

$$|\psi_1\rangle = \frac{1}{\sqrt{2}} [\alpha |0\rangle (|00\rangle + |11\rangle) + \beta |1\rangle (|10\rangle + |01\rangle)]$$

Alice then applies a Hadamard gate on the first qubit to get

$$|\psi_2\rangle = \frac{1}{2} [|00\rangle (\alpha |0\rangle + \beta |1\rangle) + |01\rangle (\alpha |1\rangle + \beta |0\rangle) + |10\rangle (\alpha |0\rangle - \beta |1\rangle) + |11\rangle (\alpha |1\rangle - \beta |0\rangle)]$$

Now, Alice measures the two qubits she has. By knowing the outcome of the measurement, Bob can easily convert the qubit he has to the original qubit to be teleported, by applying X M2 times and then Z M1 times. In effect, we have transported a qubit using two classical bits of information which were the measurements done by Alice.

7.8 Entanglement

Recall the two principles of quantum computing:

- A physical system in a definite state can still behave randomly.
- Two systems that are too far apart to influence each other can nevertheless behave in ways that, though individually random, are somehow strongly correlated.

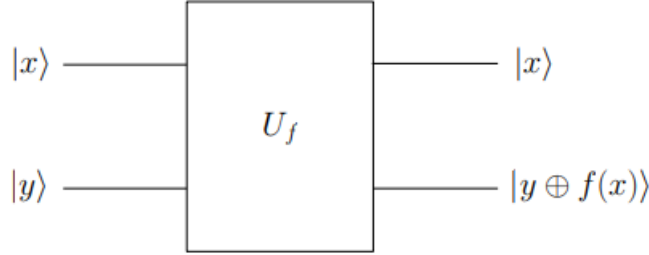
The core idea behind the second principle is entanglement. Upon reading the principle, one might be inclined to think that entanglement is simply strong correlation between two entities – but entanglement goes well beyond mere perfect (classical) correlation. If you and I read the same paper, we will have learned the same information. If a third person comes along and reads the same paper, they also will have learned this information. All three people in this case are perfectly correlated, and they will remain correlated even if they are separated from each other.

Quantum entanglement is a bit more subtle. In the quantum world, you and I could read the same quantum paper, and yet we will not learn what information is actually contained in the paper until we get together and share our information. However, when we are together, we find that we can unlock more information from the paper than we initially thought possible. Thus, quantum entanglement goes much further than perfect correlation.

Entanglement is a primary feature of quantum mechanics lacking in classical mechanics.

8 Quantum Algorithms

The first thing we'll need to understand is how to model functions in quantum circuit model. All quantum operations must be unitary and hence reversible. In general, however, given the output $f(x)$ of a function, it is not always possible to invert f to obtain the input x . In other words, we have to compute $f(x)$ in such a way as to guarantee that the computation can be undone. This is achieved via the following setup:



Here, U_f is a unitary operator mapping $|x\rangle |y\rangle \leftarrow |x\rangle |x \oplus y\rangle$ for any $x, y \in \{0, 1\}$.

A suitably programmed quantum computer should act on a number x to produce another number $f(x)$ for some specified function f . Each integer is represented in the quantum computer by the corresponding computational-basis state of k Qubits. Some basics:

- State of quantum register : linear combination of states.
- Quantum Parallelism : Computation of a function for each of the states in the input register.
- Oracle : A black box computation analogous to a classical function or subroutine.
- Measurement to extract required result.

8.1 Deutsch's Algorithm

Let both input and output registers each contain only one Qubit, so we are exploring functions f that take a single bit into a single bit. There are two rather different ways to think about such functions.

- Function is constant $f(0) = f(1) = 0$ or $f(0) = f(1) = 1$.
- Function is balanced either $f(0) = 0, f(1) = 1$ or $f(0) = 1, f(1) = 0$.

The circuit for Deutsch's algorithm is given as follows:

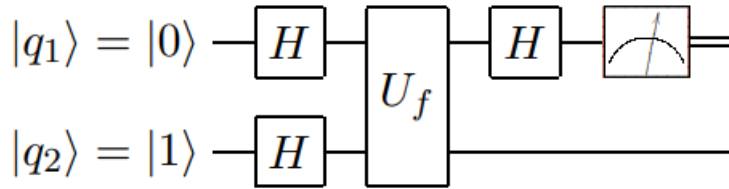


Figure 16: Quantum circuit for the Deutsch's algorithm

Divide the computation into 4 stages denoted by the quantum state in that stage: At the start of the circuit ($|\psi_1\rangle$), after the first Hadamards are applied ($|\psi_2\rangle$), after U_f is applied ($|\psi_3\rangle$), and after the last Hadamard is applied ($|\psi_4\rangle$). It is clear that

$$|\psi_1\rangle = |0\rangle |1\rangle,$$

$$|\psi_2\rangle = |+\rangle |-\rangle = \frac{1}{2}(|0\rangle |0\rangle - |0\rangle |1\rangle + |1\rangle |0\rangle - |1\rangle |1\rangle).$$

After the oracle U_f is applied, we have state

$$|\psi_3\rangle = \frac{1}{2}(|0\rangle|f(0)\rangle - |0\rangle|1 \oplus f(0)\rangle + |1\rangle|f(1)\rangle - |1\rangle|1 \oplus f(1)\rangle).$$

Before we apply the final Hadamard, it will be easier to break our analysis down into two cases: When f is constant and when f is balanced.

Case 1: Constant function. By definition, if f is constant, then $f(0) = f(1)$. Therefore, we can simplify $|\psi_3\rangle$ to

$$\begin{aligned} |\psi_3\rangle &= \frac{1}{2}(|0\rangle|f(0)\rangle - |0\rangle|1 \oplus f(0)\rangle + |1\rangle|f(0)\rangle - |1\rangle|1 \oplus f(0)\rangle) \\ &= \frac{1}{2}((|0\rangle + |1\rangle) \otimes |f(0)\rangle - (|0\rangle + |1\rangle) \otimes |1 \oplus f(0)\rangle) \\ &= \frac{1}{2}(|0\rangle + |1\rangle) \otimes (|f(0)\rangle - |1 \oplus f(0)\rangle) \\ &= \frac{1}{\sqrt{2}}|+\rangle \otimes (|f(0)\rangle - |1 \oplus f(0)\rangle). \end{aligned}$$

Thus, qubit 1 is now in state $|+\rangle$. We conclude that

$$|\psi_4\rangle = \frac{1}{\sqrt{2}}|0\rangle \otimes (|f(0)\rangle - |1 \oplus f(0)\rangle)$$

i.e., qubit 1 is exactly in state $|0\rangle$. Thus, measuring qubit 1 in the standard basis now yields outcome 0 with certainty.

Case 2: Balanced function. By definition, if f is balanced, then $f(0) \neq f(1)$. Since f is a binary function, this means $f(0) \oplus 1 = f(1)$ and equivalently $f(1) \oplus 1 = f(0)$. Therefore, we can simplify $|\psi_3\rangle$ to

$$\begin{aligned} |\psi_3\rangle &= \frac{1}{2}(|0\rangle|f(0)\rangle - |0\rangle|f(1)\rangle + |1\rangle|f(1)\rangle - |1\rangle|f(0)\rangle) \\ &= \frac{1}{2}((|0\rangle - |1\rangle) \otimes |f(0)\rangle - (|0\rangle - |1\rangle) \otimes |f(1)\rangle) \\ &= \frac{1}{2}(|0\rangle - |1\rangle) \otimes (|f(0)\rangle - |f(1)\rangle) \\ &= \frac{1}{\sqrt{2}}|-\rangle \otimes (|f(0)\rangle - |f(1)\rangle). \end{aligned}$$

Thus, qubit 1 is now in state $|-\rangle$. We conclude that

$$|\psi_4\rangle = \frac{1}{\sqrt{2}}|1\rangle \otimes (|f(0)\rangle - |f(1)\rangle)$$

i.e., qubit 1 is exactly in state $|1\rangle$. Thus, measuring qubit 1 in the standard basis now yields outcome 1 with certainty.

Conclusion. If f is constant, the algorithm outputs 0, and if f is balanced, the algorithm outputs 1. Thus, the algorithm decides whether f is constant or balanced, using just a single query!

8.2 The phase kickback trick

We've analyzed Deutsch's algorithm using a brute force calculation, but there's a more intuitive view which will be used repeatedly in later algorithms, and which simplifies our calculation here greatly. This view is in terms of the phase kickback trick, which Deutsch's algorithm uses. To explain the trick, consider for any $x \in \{0, 1\}$ what happens if we run U_f on input $|x\rangle|-\rangle$:

$$\begin{aligned} |\psi\rangle &= U_f |x\rangle|-\rangle = \frac{1}{\sqrt{2}}(U_f |x\rangle|0\rangle - U_f |x\rangle|1\rangle) = \frac{1}{\sqrt{2}}(|x\rangle|f(x)\rangle - |x\rangle|1 \oplus f(x)\rangle) \\ &= |x\rangle \otimes \frac{1}{\sqrt{2}}(|f(x)\rangle - |1 \oplus f(x)\rangle) \end{aligned}$$

Now, there are two possibilities: Either $f(x) = 0$, or $f(x) = 1$. If $f(x) = 0$, the equation above simplifies to

$$|\psi\rangle = |x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |x\rangle |-\rangle,$$

i.e. the input state is unchanged by the action of U_f . If, on the other hand, $f(x) = 1$, we instead have

$$|\psi\rangle = |x\rangle \otimes \frac{1}{\sqrt{2}}(|1\rangle - |0\rangle) = -|x\rangle |-\rangle,$$

i.e. a 1 phase factor is produced. We can summarize both these cases in a single equation:

$$U_f |x\rangle |-\rangle = (-1)^{f(x)} |x\rangle |-\rangle.$$

8.3 The Deutsch-Josza algorithm

Deutsch's algorithm works in the simple case where f acts on a single input qubit. Developing the n -bit generalization of Deutsch's algorithm, known as the Deutsch-Josza algorithm. Specifically, imagine now we have an n -bit function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ which is promised to be constant or balanced, and we wish to determine which is the case. Here, constant means $f(x)$ is the same for all $x \in \{0, 1\}^n$, and balanced means $f(x) = 0$ for precisely half the $x \in \{0, 1\}^n$ and $f(x) = 1$ for the remaining inputs.

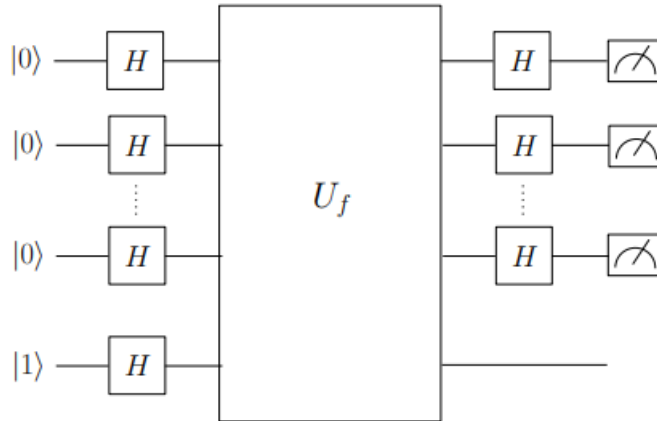


Figure 17: Quantum circuit for the Deutsch-Josza algorithm.

In this more general setting, note that we define the oracle U_f implementing f analogously to before: $U_f |x\rangle |y\rangle = |x\rangle |y \oplus f(x)\rangle$, where now x is an n -bit string. As before, each wire denotes a single qubit. The first n qubits are initialized to $|0\rangle$; these are the input qubits. The final, i.e. $(n + 1)$ st, qubit is initialized to $|1\rangle$. Observe that the algorithm is the straightforward generalization of Deutsch's algorithm to the setting of n input qubits. We claim that using a single query to U_f , the Deutsch-Josza algorithm can determine if f is constant or balanced. Let us now see why this is so.

As before, we divide the computation into 4 stages denoted by the quantum state in that stage: At the start of the circuit ($|\psi_1\rangle$), after the first Hadamards are applied ($|\psi_2\rangle$), after U_f is applied ($|\psi_3\rangle$), and after the last Hadamard is applied ($|\psi_4\rangle$). It is clear that

$$\begin{aligned} |\psi_1\rangle &= |0\rangle \dots |0\rangle |1\rangle = |0\rangle^{\otimes n} |1\rangle \\ |\psi_2\rangle &= |+\rangle \dots |+\rangle |-\rangle = |+\rangle^{\otimes n} |1\rangle \end{aligned}$$

Since we have defined the action of U_f in terms of the standard basis, i.e. $U_f |x\rangle |y\rangle = |x\rangle |y \oplus f(x)\rangle$, in order to understand how U_f applies to $|\psi_2\rangle$, we first need to rewrite $|+\rangle^{\otimes n}$ in terms of the

standard basis. For this, note that

$$|+\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}}(|0\rangle + |1\rangle) \otimes \cdots \otimes (|0\rangle + |1\rangle) = \frac{1}{2^{n/2}} \sum_{x \in \{0,1\}^n} |x\rangle,$$

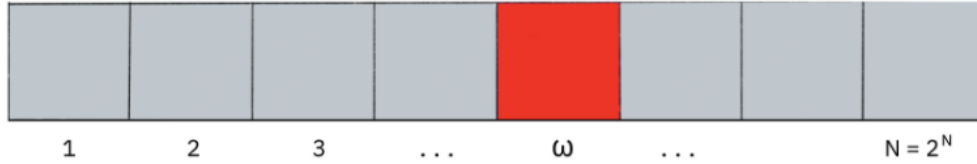
where the last equality holds since expanding the tensor products out yields 2^n terms in the sum, each of which corresponds to a unique string $x \in \{0,1\}^n$.

8.4 Grover's algorithm

Quantum computer has over a classical computer is its superior speed searching databases. Grover's algorithm demonstrates this capability. This algorithm can speed up an unstructured search problem quadratically, but its uses extend beyond that; it can serve as a general trick or subroutine to obtain quadratic run time improvements for a variety of other algorithms. This is called the *amplitude amplification trick*.

Unstructured search

Suppose you are given a large list of N items. Among these items is one item with a unique property that we wish to locate. We will call this one the winner, w . Think of each item in the list as a box of a particular color. Say all items in the list are gray except the winner w , which is red.



To find the red box – the marked item – using classical computation, one would have to check on average $N/2$ of these boxes, and in the worst case, all N of them. On a quantum computer, however, we can find the marked item in roughly \sqrt{N} steps with Grover's amplitude amplification trick. A quadratic speedup is indeed a substantial time-saver for finding marked items in long lists. Additionally, the algorithm does not use the list's internal structure, which makes it generic; this is why it immediately provides a quadratic quantum speed-up for many classical problems.

The Oracle

How will the list items be provided to the quantum computer? For the examples in this topic, our 'database' is comprised of all the possible computational basis states our qubits can be in. For example, if we have 3 qubits, our list is the states $|000\rangle, |001\rangle, \dots, |111\rangle$ (i.e the states $|0\rangle \rightarrow |7\rangle$).

Grover's algorithm solves oracles that add a negative phase to the solution states. That is, for any state $|x\rangle$ in the computational basis:

$$U_w |x\rangle = \begin{cases} |x\rangle & \text{if } x \neq w \\ -|x\rangle & \text{if } x = w \end{cases}$$

This oracle will be a diagonal matrix, where the entry that correspond to the marked item will have a negative phase. For example, if we have three qubits and $w = 101$, our oracle will have the matrix:

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

What makes Grover's algorithm so powerful is how easy it is to convert a problem to an oracle of this form. There are many computational problems in which it's difficult to find a solution, but relatively easy to verify a solution. For these problems, we can create a function f that takes a

proposed solution x and returns $f(x) = 0$ if x is not a solution ($x \neq w$), and $f(x) = 1$ for a valid solution ($x = w$). Our oracle can then be described as:

$$U_w |x\rangle = (-1)^{f(x)} |x\rangle$$

and the oracle's matrix will be a diagonal matrix of the form:

$$U_w = \begin{bmatrix} (-1)^{f(0)} & 0 & \dots & 0 \\ 0 & (-1)^{f(1)} & \dots & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & (-1)^{f(2^n-1)} \end{bmatrix}$$

Amplitude amplification

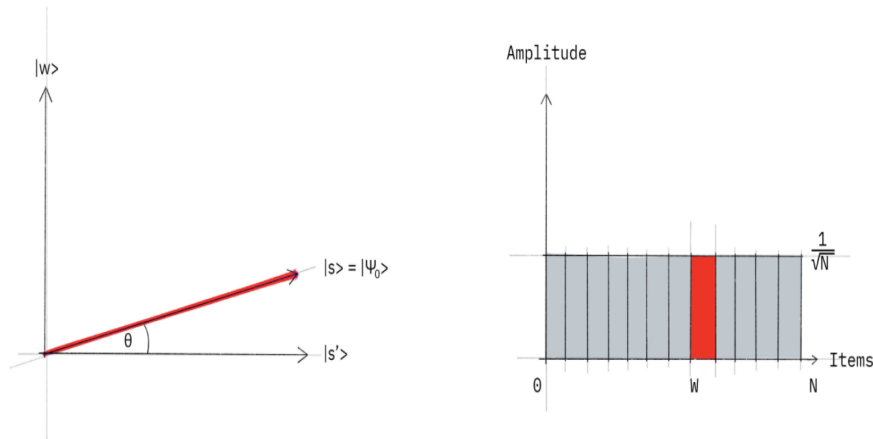
So how does the algorithm work? Before looking at the list of items, we have no idea where the marked item is. Therefore, any guess of its location is as good as any other, which can be expressed in terms of a quantum state called a uniform superposition:

$$|s\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle$$

If at this point we were to measure in the standard basis $|x\rangle$, this superposition would collapse to any one of the basis states with the same probability of $\frac{1}{N} = \frac{1}{2^n}$. Our chances of guessing the right value is therefore 1 in 2^n , as could be expected. Hence, on average we would need to try about $N = 2^n$ times to guess the correct item. Enter the amplitude amplification procedure, which is how a quantum computer significantly enhances this probability. This procedure stretches out (amplifies) the amplitude of the marked item, which shrinks the other items' amplitudes, so that measuring the final state will return the right item with near certainty.

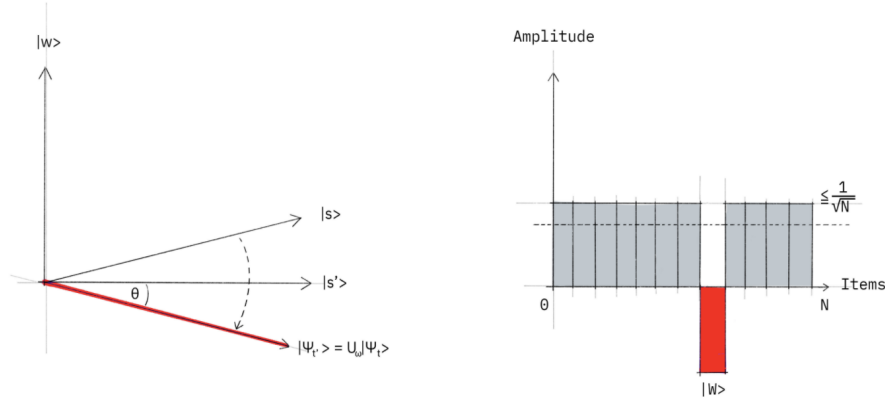
This algorithm has a nice geometrical interpretation in terms of two reflections, which generate a rotation in a two-dimensional plane. The only two special states we need to consider are the winner $|w\rangle$ and the uniform superposition $|s\rangle$. These two vectors span a two-dimensional plane in the vector space. They are not quite perpendicular because $|w\rangle$ occurs in the superposition with amplitude $N^{-1/2}$ as well. We can, however, introduce an additional state $|s'\rangle$ that is in the span of these two vectors, is perpendicular to $|w\rangle$, and is obtained from $|s\rangle$ by removing $|w\rangle$ and rescaling.

Step 1 The amplitude amplification procedure starts out in the uniform superposition $|s\rangle$. (The uniform superposition is easily constructed from $|s\rangle = H^{\otimes n} |0\rangle^n$).



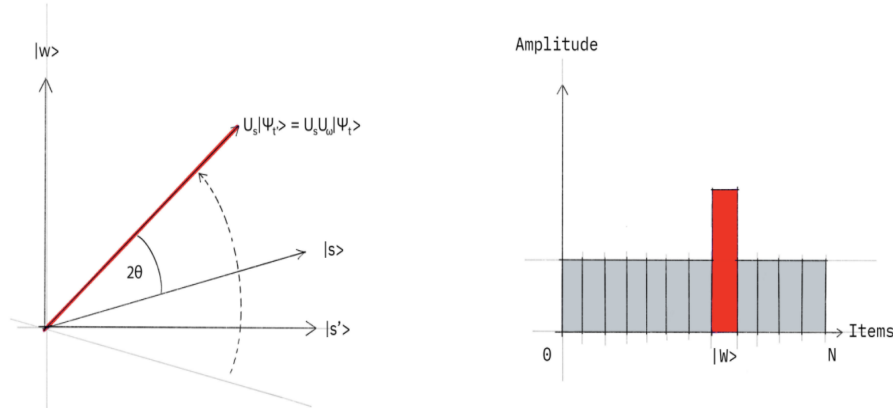
The left graphic corresponds to the two-dimensional plane spanned by perpendicular vectors $|w\rangle$ and $|s'\rangle$, which allows us to express the initial state as $|s\rangle = \sin\theta |w\rangle + \cos\theta |s'\rangle$, where $\theta = \arcsin \langle s|w\rangle = \arcsin \frac{1}{\sqrt{N}}$. The right graphic is a bar graph of the amplitudes of the state $|s\rangle$.

Step2 We apply the oracle reflection U_f to the state $|s\rangle$.



Geometrically this corresponds to a reflection of the state $|s\rangle$ about $|s'\rangle$. This transformation means that the amplitude in front of the state becomes negative, which in turn means that the average amplitude (indicated by a dashed line) has been lowered.

Step3 We now apply an additional reflection U_s about the state $|s\rangle$: $U_s = 2|s\rangle\langle s| - 1$. This transformation maps the state to $U_s U_f |s\rangle$ and completes the transformation.



Two reflections always correspond to a rotation. The transformation $U_s U_f$ rotates the initial state $|s\rangle$ closer toward the winner $|w\rangle$. The action of the reflection U_s in the amplitude bar diagram can be understood as a reflection about the average amplitude. Since the average amplitude has been lowered by the first reflection, this transformation boosts the negative amplitude of $|w\rangle$ to roughly three times its original value, while it decreases the other amplitudes. We then go to Step 2 to repeat the application. This procedure will be repeated several times to focus in on the winner.

After steps, the state will have transformed to $|\psi_t\rangle$, where $|\psi_t\rangle = (U_s U_f)^t |s\rangle$. How many times do we need to apply the rotation? It turns out that roughly \sqrt{N} rotations suffice. This becomes clear when looking at the amplitudes of the state. We can see that the amplitude of $|w\rangle$ grows linearly with the number of applications $\approx tN^{-1/2}$. However, since we are dealing with amplitudes and not probabilities, the vector space's dimension enters as a square root. Therefore it is the amplitude, and not just the probability, that is being amplified in this procedure.

In the case that there are multiple solutions, M , it can be shown that roughly $\sqrt{\frac{N}{M}}$ rotations will suffice.