INTRODUCTION TO QUANTUM MECHANICS

ROHIT TRIPATHI

Abstract. Discussed here is the mathematics of quantum mechanics at an introductory level. Hilbert spaces, the Bra-Ket notation, Fourier Transforms, the Schrödinger Wave Equation, and Quantum Computing are covered.

Introduction

Quantum mechanics is known as the most complete description of the physical world yet. Despite this fact, it is often hard to accept the assumptions and explanations of this theory because it deviates so wildly from everyday experiences. Its notation, terminology, and principles, too, appear alien at first sight. This paper introduces quantum mechanics from a mathematical standpoint, and culminates with an example of its application in quantum computing.

Here we review some of the ideas and entities that populate articles and books on quantum mechanics. These are Hilbert Space, Bras and Kets, Fourier Transforms, and the Schrödinger Wave Equation. We will see how these ideas interweave and describe the theory.

The target audience are readers who would welcome familiarity with the mathematical language of quantum mechanics. It may be treated a starting point for understanding various math-intensive areas of the subject—like quantum computing.

1. Hilbert Space

Hilbert space is the underlying base on which other structures of quantum mechanics are raised. Presented in this section is the sequence of ideas that build the Hilbert Space. An interested reader may refer to [**?**], [**?**] (chapters 4, 5), and [**?**], (chapter 10—for a detailed discussion of the Lebesgue theory).

1.1. **Fields.** A field is a non-empty set **F**, that allows two operations $+$ and $.$, and contains two objects 0 and 1 so that for every x, y, z in **F**, the following are true:
1. $x + y = y + x$ and $x.y = y.x$
2. $(x + y) + z = x + (y + z)$ and $(x.y).z = x.(y.z)$
3. $x + 0 = x$ and $x.1 = x$
4. For all x, there exists a y so that $x + y = 0$
5. For all x when x is not 0, there exists a y so that $x.y = 1$

1.2. **Vector Space.** Let **F** be a field (for example, real numbers $\mathbb{R}$, or complex numbers $\mathbb{C}$). Let elements of **F** be called scalars. A vector space over **F** is a set **V** so that two operations addition and scalar multiplication are allowed over its objects. $\mathbf{x}+\mathbf{y}$ is defined as the sum of $\mathbf{x}$ and $\mathbf{y}$; and c$\mathbf{x}$ is defined as the multiplication of $\mathbf{x}$ by scalar c. For all $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ in **V** the following are true:

1. $\mathbf{x}+\mathbf{y} \in \mathbf{V}$ and c$\mathbf{x} \in \mathbf{V}$ ($\mathbf{V}$ is closed under addition and scalar multiplication).
2. $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$ (vector addition is commutative).
3. $(\mathbf{x}+\mathbf{y})+\mathbf{z}=\mathbf{x}+(\mathbf{y}+\mathbf{z})$ (vector addition is associative).
4. $\exists\ \mathbf{0} \in \mathbf{V}$ so that $\mathbf{x}+\mathbf{0}=\mathbf{0}+\mathbf{x}=\mathbf{x}$ (zero vector $\mathbf{0}$ exists)
5. $\exists\ \mathbf{y} \in \mathbf{V}$ so that $\mathbf{x}+\mathbf{y}=\mathbf{0}$ (negative vector $\mathbf{y}$ exists $\forall \mathbf{x}$).
6. c$(\mathbf{x}+\mathbf{y})$ = c$\mathbf{x}$ + c$\mathbf{y}$ and (c + d)$\mathbf{x}$ = c$\mathbf{x}$ + d$\mathbf{x}$ (distributive law follows).
7. (cd)$\mathbf{x}$=c(d$\mathbf{x}$) (associative law holds).
8. $1\mathbf{x}=\mathbf{x}$, and $0\mathbf{x}=\mathbf{0}$.

1.2.1. *Typical Vector Spaces:* The sets $\mathbb{R}$, $\mathbb{C}$ are vector spaces. Other standard examples are:
1. The set of all $\mathbf{z}$ defined as tuples of n-complex numbers $(z_1,\dots,z_n)$. Vectors $\mathbf{x}=(x_1,\dots,x_n)$ and $\mathbf{y}=(y_1,\dots,y_n)$ are equal if $x_i=y_i$ for all $i=1,2,\dots,n$. Addition is defined on this set as $\mathbf{x}+\mathbf{y}=(x_1+y_1,\dots,x_n+y_n)$. Scalar multiplication $c\mathbf{z}$ is $(cz_1,\dots,cz_n)$. This is called a vector space of n-tuples.
2. The set of all sequences $\mathbf{x}=x_n$, (where $n=1,2,\dots$), such that $x_n=0$ for all $n>N_x$ (where $N_x<\infty$ and $N_x \in \mathbb{N}$ is defined independently for all $\mathbf{x}$). $\mathbf{x}=\mathbf{y}$ if $x_i=y_i$ (for all $i=1,2,3,\dots$). $\mathbf{x}+\mathbf{y}$ is defined as x_n+y_n for all n. If c is a scalar, the $c\mathbf{x}=cx_n$ for all n. This is called the vector space of finitely non-zero sequences.
3. The set of all functions $\mathbf{f}=f(x)$; with addition defined as $\mathbf{f}+\mathbf{g}=(f+g)(x)$, and scalar multiplication defined as $c\mathbf{f}=cf(x)$. Example is the set $\{k.\sin x|k\in\mathbb{R}\}$.

1.2.2. *Treatement of Vectors:* In physics, any quantity with magnitude and direction is known as a vector. Examples are velocity, momentum, etc. In mathematical terms, however, a vector is an object that belongs to a vector space. In the above example, we notice that $\pi.\sin x$ is a vector and belongs to the vector space $\mathbf{V}$.

1.3. **Basis Vectors.** A basis is a set of vectors that spans the vector space and are linearly independent.The number of vectors in the basis is the dimension of the space. Thus if the basis contains n vectors, the space is an n-dimensional space. The vector space is called *finite dimensional* if the number of its basis vectors is finite, else it is called *infinite dimensional.* An infinite dimensional vector space has infinitely many vectors $x_1,x_2,x_3,\dots$ as a basis.

1.3.1. *Span:* A set of vectors $\mathbf{B}$ spans a vector space $\mathbf{V}$ if every vector v of $\mathbf{V}$ can be written as $v=\sum_i^n c_ie_i$.(Where c_i are scalars—complex numbers in our case—and e_i are vectors from $\mathbf{B}$.)

1.3.2. *Linear Independence:* Let $\mathbf{B}$ be a set of vectors $e_i\in\mathbf{B}$; and let there be scalars $c_i\in\mathbb{C}$, where at least one c_i is non-zero. If $\sum_i^n c_ie_i\neq 0$ then the vectors in $\mathbf{B}$ are linearly independent.

1.4. **Inner Product.** An inner product is a function that takes in as input two vectors from a vector space and produces a scalar as output. A vector space that allows inner products is called an inner-product space. A real vector space with inner-product is called Euclidean space. A complex vector space with inner-product is called unitary space. We represent an inner product between two vectors $\mathbf{x}$ and $\mathbf{y}$ as $\langle\mathbf{x},\mathbf{y}\rangle$.

1.4.1. *Properties of Inner-Product:* The following are some properties that the inner-product function must satisfy [**?**], [**?**]:
1. $\langle \mathbf{f}, \mathbf{g} \rangle = \langle \mathbf{g}, \mathbf{f} \rangle^*$
2. $\langle \mathbf{f}, \alpha_1 \mathbf{g_1} + \alpha_2 \mathbf{g_2} \rangle = \alpha_1 \langle \mathbf{f}, \mathbf{g_1} \rangle + \alpha_2 \langle \mathbf{f}, \mathbf{g_2} \rangle$ where α_1 and α_2 are scalars.
3. $\langle \mathbf{f}, \mathbf{g} \rangle \geqslant 0$ and $\langle \mathbf{f}, \mathbf{g} \rangle = 0 \Rightarrow \mathbf{f} = \mathbf{g} = \mathbf{0}$. Here $0 \in \mathbb{R}$ and $\mathbf{0}$ is the zero vector.

1.4.2. *Norm of a Vector:* The norm of a vector $\mathbf{f}$ is the square-root of $\langle \mathbf{f}, \mathbf{f} \rangle$, and is represented as $\|\mathbf{f}\|$. We thus say that $\sqrt{\langle \mathbf{f}, \mathbf{f} \rangle} = \|\mathbf{f}\|$. We always understand that the norm of any vector is finite. That is $\|\mathbf{f}\| < \infty$. An important property of the inner product is the *Cauchy-Schwarz* inequality: $|\langle \mathbf{f}, \mathbf{g} \rangle| \leqslant \|\mathbf{f}\|.\|\mathbf{g}\|$. The norm follows the triangle inequality: $\|\mathbf{f} \pm \mathbf{g}\| \leqslant \|\mathbf{f}\| + \|\mathbf{g}\|$.

1.4.3. *Examples of Inner Products:* Inner product is defined diferently for n-touples and vectors from function spaces.
1. An inner product of two vectors $\overrightarrow{u}$ and $\overrightarrow{v}$, that are n-touples of the form $\left(u_1, u_2, \ldots, u_n\right)$ and $\left(v_1, v_2, \ldots, v_n\right)$ respectively, is $\sum_i^n u_i^* v_i$. (Such vectors are finite dimensional.)
2. If $\mathbf{x} = x_i$ and $\mathbf{y} = y_i$ belong to a vector space of finitely non-zero sequences, then $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{\infty} x_i^* y_i$.
3. Vectors that are also functions, e.g. $\psi(x), \chi(x)$ and continuous on an interval $a \leqslant x \leqslant b$, have an inner product $\langle \psi, \chi \rangle$ defined as $\int_a^b |\left(\psi^*\right) \chi| dx$.

1.4.4. *Orthogonal Vectors:* Two vectors $\mathbf{x}$ and $\mathbf{y}$ are orthogonal if and only if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. A set of non-zero orthogonal vectors is always linearly independent.

1.4.5. *Orthonormal Vectors:* A set of vectors $\{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}\}$ is orthonormal if $\langle \mathbf{x_i}, \mathbf{x_j} \rangle = \delta_{ij}$. In a finite dimensional vector space, an orthonormal set is *complete* if it is not contained in any larger orthonormal set.

1.5. **Metric Space.** A metric space is a set of points $\mathbf{M}$ that defines a distance between all $\mathbf{f}$, $\mathbf{g}$ in $\mathbf{M}$ as $d(\mathbf{f}, \mathbf{g}) = \|\mathbf{f} - \mathbf{g}\|$. The function $d(\mathbf{x}, \mathbf{y})$ is known as the distance function and it returns a real number [**?**].

1.5.1. *Properties of Distance Function:* Let M be a set of objects and $d[M, M] \to \mathbb{R}$ be its distance function. Then for all $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ in M the following must be true:
1. $d(\mathbf{x}, \mathbf{y}) \geqslant 0$
2. $d(\mathbf{x}, \mathbf{x}) = 0$
3. If $d(\mathbf{x}, \mathbf{y}) = 0$ then $\mathbf{x} = \mathbf{y}$
4. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$
5. $d(\mathbf{x}, \mathbf{z}) \leqslant d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$

1.5.2. *Examples of metric spaces:* The set of real numbers $\mathbb{R}$ is a metric space. The n-dimensional Euclidean spaces (having vectors that are n-tuples) are metric spaces when distance between two vectors $\overrightarrow{u} = (u_1, \ldots, u_n)$ and $\overrightarrow{v} = (v_1, \ldots, v_n)$ is defined as $\sqrt{\sum_i \left(u_i - v_i\right)^2}$.

1.6. **Completeness of Metric Spaces.** A space in which every Cauchy Sequence converges is called complete [**?**], [**?**].

1.6.1. *Cauchy Sequence.* Let M be a metric space. Let f, and $f_n \in M$ (where $n = 1, 2, 3, \dots$). The sequence $\{f_n\}$ is said to converge to f in M if $\|f_n - f\| \to 0$. $\{f_n\}$ is a Cauchy sequence in M if for every $\epsilon > 0$ there is an integer N such that $n \geqslant N$ implies $\|f_n - f\| \leqslant \epsilon$. All Cauchy sequences converge due to this property. A proof can be found in [**?**].

1.7. **Definition of Hilbert Space.** By now, we have sufficient tools to understand what a Hilbert space is. The well-known description is:*a Hilbert Space is an inner product space, which is complete as a metric space.*

1.7.1. *Examples of Hilbert Spaces:* A common and widely used example of a Hilbert Space is the L_2 space which is defined as follows:
1. It is a space of functions $\mathcal{H} = \{f : [a, b] \to \mathbb{C} | a, b \in \mathbb{R}, a < b, \int_a^b |f|^2 dx < \infty\}$
2. An inner product is defined as $\int_a^b f(x)^* g(x)\, dx$
3. A distance between f, g is defined as $\|f - g\| = \left[\int_a^b \|f - g\|^2 dx\right]^{\frac{1}{2}}$

2. The Bra-Ket Notation

The language of Quantum Mechanics uses a distinct representation of vectors—popularly known as the Bra-Ket notation—and some key ideas regarding it are discussed here [**?**], [**?**].

Let us assume that ψ is a state of a system. (ψ can denote a number of ideas, for example: a state in which the system oscillates about a point.) Then the symbol $|\psi\rangle$ represents the quantum state ψ and is called a Ket.

Kets are abstract entities and are best understod in terms of what they do. Commonly, they are considered to be vectors that represent the state space of their system. Also, if two distinct kets $|\alpha\rangle$ and $|\beta\rangle$ represent two distinct states of the system, then the ket $|\psi\rangle = c_1|\alpha\rangle + c_2|\beta\rangle$ (where $c_1, c_2 \in \mathbb{C}$) is also a possible state of the system. Thus kets form a linear vector space.

Similarly, if there is a ket $|\xi\rangle$ and a complex function $\lambda(\xi)$ then $|\omega\rangle = \int_{\xi_1}^{\xi_2} \lambda(\xi)\, |\xi\rangle d\xi$ is also a vector of the ket space. We then say that $|\omega\rangle$ is a superposition of $|\xi\rangle$.

Any ket $|\alpha\rangle$ has a dual associated with it, called a bra. It is represented as $\langle\alpha|$. Bras and kets are different from each other, due to which operations like $|\alpha\rangle \pm \langle\alpha|$ are not defined. This can be explained by the following description of finite dimensional bras and kets:

Let $|\alpha\rangle = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \dots \\ c_n \end{pmatrix}$, $c_i \in \mathbb{C}$. Then, the associated bra $\langle\alpha|$ is $(c_1^* \;\; c_2^* \;\; c_3^* \;\; \dots \;\; c_n^*)$.

(Where c_i^* is a complex conjugate of c_i.)

We see that by the rules of matrix addition and subtraction, $|\alpha\rangle \pm \langle\alpha|$ are invalid.

If $|\psi\rangle = c_1|\alpha\rangle + c_2|\beta\rangle$ then the dual of $|\psi\rangle$ is $\langle\psi| = c_1^*\langle\alpha| + c_2^*\langle\beta|$

An inner product of $|\alpha\rangle$ and $\langle\beta|$ is defined as $\langle\beta|\alpha\rangle = \langle\alpha|\beta\rangle^*$. In finite dimensional spaces, the outcome is obvious. In function spaces, the inner product is defined as in Hilbert spaces, that is $\langle\psi|\chi\rangle = \int_a^b \psi^* \chi$. As is the case with inner products, a complex number is the result.

An *outer product* is represented as $|\alpha\rangle\langle\beta|$. This is a very special alignment as it assists the development of a variety of operations. Let us say $\mathbf{X} = |\alpha\rangle\langle\beta|$. If we now pick another ket $|\psi\rangle$ and multiply, then $\mathbf{X}|\psi\rangle = c|\alpha\rangle$ where $c \in \mathbb{C}$ and $c = \langle\beta|\psi\rangle$. This behavior makes $\mathbf{X}$ an 'operator' in Hilbert space. (Operators are instructions to do something on a function. For example, $\frac{\partial}{\partial x}$ is an operator, that when given a function $f(x,y)$ returns $\frac{\partial f(x,y)}{\partial x}$.) The dual of $\mathbf{X}|\psi\rangle$ is $\langle\psi|\mathbf{X}^\dagger$, where $\mathbf{X}^\dagger$ is the Hermitian conjugate[1] of $\mathbf{X}$. If $\mathbf{X}$=$\mathbf{X}^\dagger$, then $\mathbf{X}$ is called a Hermitian operator.

3. Fourier Transforms

This section is adapted from [**?**], [**?**], and from [**?**]. For more details on the topic, these are excellent sources to start with.

3.1. **Fourier Series.** Fourier series are a way to represent periodic functions as superpositions of other periodic functions. A trigonometric series of the form:

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

is known as a Fourier series. The numbers a_n, b_n are known as the coefficients of the series.

If $f(x)$ is a function periodic on the interval $[-\pi, +\pi]$, and we want to express it as a Fourier series, our first task is to evaluate the coeffecients. We start as:

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

and evaluate

$$\int_{-\pi}^{+\pi} f(x)dx = \int_{-\pi}^{+\pi} \left[\frac{a_0}{2} + \sum_{i=1}^{\infty} (a_n \cos nx + b_n \sin nx)\right] dx$$

to obtain

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{+\pi} f(x)dx \tag{3.1}$$

Also, upon evaluating the integral of $f(x)\sin nx$ and $f(x)\cos nx$ we find:

$$a_n = \frac{1}{\pi}\int_{-\pi}^{+\pi} f(x)\cos nx dx \tag{3.2}$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{+\pi} f(x)\sin nx dx \tag{3.3}$$

This method can be extended to functions periodic over arbitrary intervals. It may be noticed that the functions $\sin nx$ and $\cos nx$ behave as orthonormal vectors in the vector space $\{k_1 \sin nx + k_2 \cos nx | k_1, k_2 \in \mathbb{R}, n \in \mathbb{N}\}$.

---

[1] $\langle\psi|\mathbf{X}|\psi\rangle = \langle\psi|\mathbf{X}\psi\rangle = \langle\mathbf{X}^\dagger\psi|\psi\rangle$

3.2. **Fourier Transform.** A Fourier transform of a function $f(x)$, if it exists is:

$$\boxed{F(u) \equiv \mathscr{F}[f] = \left(\frac{\alpha}{2\pi}\right)^{1/2} \int_{-\infty}^{+\infty} e^{-i\alpha ux} f(x)\, dx} \tag{3.4}$$

Here α is a constant defined according to situation. In quantum mechanics, we use $\alpha = 1/\hbar$. Some authors also prefer to do away with the scaling factor $\frac{\alpha}{2\pi}$. We can obtain $f(x)$ from $F(u)$ by the inverse Fourier transform:

$$\boxed{f(x) = \mathscr{F}^{\dagger}[F] = \left(\frac{\alpha}{2\pi}\right)^{1/2} \int_{-\infty}^{+\infty} e^{i\alpha ux} F(u)\, du} \tag{3.5}$$

Existence of Fourier Transform:
Condition 1. If the function $f(x)$ is continuous and integrable in the following way:

$$\int_{-\infty}^{+\infty} |f(x)|dt < \infty \tag{3.6}$$

then its Fourier transform $F(u)$ exists, and its inverse Fourier transform is also possible. Known as the *Dirichlet condition*, this is sufficient for the existence of a function's FT, but not necessary.
Condition 2. When $f(x)$ is of the form $\beta(x)\sin(2\pi kx + \phi)$ where k and ϕ are arbitrary constants, and $\beta(x+\delta) < \beta(x)$, then if for $0 < \lambda < |x|$, the function $f(x)$ is integrable as above, then the Fourier transform of $f(x)$ exists, and it also satisfies the inverse FT.
Condition 3. If $f(x)$ is an impulse function like Dirac's δ-function $\delta(x)$ defined as:

$$\int_{-\infty}^{+\infty} \delta(x - x_0)\, g(x)\, dx = g(x_0) \tag{3.7}$$

In the special case when $g(x) = 1$ always, we have $\int_{-\infty}^{+\infty} \delta(x - x_0)\, dt = 1$.
Let $h(x) = K\delta(x)$. Then

$$H(u) = \int_{-\infty}^{+\infty} K\delta(x) e^{-i(2\pi ux)} dx = Ke^0 = K. \tag{3.8}$$

The inverse Fourier transform is given as

$$h(x) = \int_{-\infty}^{+\infty} [K]\, e^{i(2\pi ux)} du \tag{3.9}$$

3.2.1. *Discrete Fourier Transform* [**?**]. The discrete fourier transform changes a vector $\mathbf{x} = (x_0, \ldots, x_n)$ to the vector $\mathbf{y} = (y_0, \ldots, y_n)$ in the following way:

$$y_k \longrightarrow \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{i(2\pi jk)/N} \tag{3.10}$$

3.2.2. *Quantum Fourier Transform* [**?**]. If a complex vector space $\mathcal{H}$ (for example, a Hilbert space) has for its basis the kets $|0\rangle, \ldots, |N-1\rangle$, then the Quantum Fourier Transform (or QFT) is defined as an operator that has the following action on the basis vectors:

$$|j\rangle \longrightarrow \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{i(2\pi jk)/N} |k\rangle \tag{3.11}$$

4. The Wave Equation

When studying quantum-mechanical systems, we are interested in obtaining the wavefunction of the particle(s). This is a function with complex coefficients and it depends on position and time. We can thus write it as $\Psi(\overrightarrow{r}, t)$.
To obtain the wave equation for systems, we solve a partial differential equation called the Wave Equation. The *a priori* conditions to be satisfied by a wave equation are [**?**]:
1. The equation must be linear and homogeneous. This ensures that the wave has properties of superposition. If Ψ_1 and Ψ_2 are two solutions, then $\Psi = \lambda_1\Psi_1 + \lambda_2\Psi_2$ is also a solution to the wave equation. Thus, two waves can create a new wave upon superposition.
2. It must be a differential equation of the first order with respect to time. Due to this, knowledge of Ψ at one instant entirely defines its later evolution.

In 1926 Erwin Schrödinger introduced the wave equation for non-relativistic quantum-mechanical systems which is widely used.

4.1. **The Schrödinger Equation.** The complete Schrödinger Wave Equation is compactly written as:

$$\boxed{i\hbar\frac{\partial\Psi}{\partial t} = \mathbf{H}|\Psi\rangle} \tag{4.1}$$

Just as classical mechanics uses Newton's Second Law to study the behavior of large objects under the action of forces, quantum mechanics uses the Schrödinger equation to identify the behavior of particles under the influence of potential. Our problem is to identify the wave-function of a particle, denoted by Ψ; which is a function of the radius vector $\overrightarrow{r}$ and time t.

H is called the Hamiltonian (operator) of the system. It is written as $\frac{-i\hbar^2}{2m}\Delta + V$ and returns the total energy of the system. V is the potential energy function. Δ is known as the Laplacian operator, and $\Delta = \nabla.\nabla$ where ∇ is the gradient operator. Given a function $f(x, y, z)$, ∇ returns the vector (f_x, f_y, f_z). That is

$$\nabla f = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}).$$

Thus,

$$\Delta f = \nabla.\nabla f = (\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}).$$

The Schrödinger equation can now be written as:

$$\boxed{i\hbar\frac{\partial\Psi}{\partial t} = \left[\frac{-\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + V\right]\Psi} \tag{4.2}$$

4.1.1. *Variations of the Schrödinger Equation:* The wave equation in one dimension is written as:

$$\boxed{i\hbar\frac{\partial\psi(x,t)}{\partial t} = \frac{-\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V\psi(x,t)} \tag{4.3}$$

The Schrödinger equation in polar spherical coordinates (r, θ, ϕ) is a very important form due to its use in studying atoms. To obtain this variation, we first calculate the Laplacian in spherical coordinates and then apply it to the original equation. The Laplacian turns out to be

$$\frac{\partial^2}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{1}{r^2 \sin\theta^2}\frac{\partial^2}{\partial \phi^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{\cot\theta}{r^2}\frac{\partial}{\partial \theta}.$$

Thus, the complete equation in polar spherical cordinates is written as:

$$\boxed{i\hbar\frac{\partial \Psi}{\partial t} = \frac{-i\hbar^2}{2m}\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{1}{r^2 \sin\theta^2}\frac{\partial^2}{\partial \phi^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{\cot\theta}{r^2}\frac{\partial}{\partial \theta}\right)\Psi + V\Psi} \tag{4.4}$$

4.1.2. *About the Wave Function:* One of the assumptions of quantum mechanics is that a particle is spread out in space, instead of being at one definite point. Using the wave function we can calculate the probability of finding it between two given locations.
For example, in one dimension the probability of finding a particle between x and $x + dx$ is $p(x) = |\Psi|^2 dx$ and the probability of finding a particle between points a and b is:

$$\boxed{p = \int_a^b |\Psi|^2 dx} \tag{4.5}$$

The above is known as the Born Interpretation of the wave function.

4.1.3. *Normalization:* If a particle exists, *we can find it somewhere.* Though we don't know where exactly the particle is, we are sure that it must be *somewhere.* This means that:

$$\boxed{\int_{-\infty}^{+\infty} |\Psi\left(\overrightarrow{\mathbf{r}}, t\right)|^2 d\overrightarrow{\mathbf{r}} = 1} \tag{4.6}$$

This process is called normalizing the wave function [**?**]. Since the wavefunction is linear, therefore if $\Psi\left(\overrightarrow{\mathbf{r}}, t\right)$ is a solution, so is $A\Psi\left(\overrightarrow{\mathbf{r}}, t\right)$. Under the light of this requirement, if some results to equation 5.6 are $infinity$ or are 0, we must reject the corresponding wavefunction.

4.1.4. *The Wavefunction and Hilbert Space:* The solutions of the wave function for a particular system constitute a Hilbert space.

4.2. **Relativistic Quantum Mechanics: Klein-Gordon Equation.** The equation below is known as the Klein-Gordon equation and plays a vital role in relativistic quantum mechanics [**?**]. A discussion on this is beyond the scope of this paper.

$$\left[\Box + \left(\frac{mc}{\hbar}\right)^2\right]\psi = 0 \tag{4.7}$$

Here, $\Box$ is the Dalembertian operator: $\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta$. It turns out that solutions to the above equation cannot be considered wave functions, because the equation does not satisfy condition 2 above.

5. Solutions to the Schrödinger Equation

The wavefunction for a system is found by solving the partial differential equation with the appropriate V. The resulting solution is normalized as in 5.1.3, and is accepted if the normalization conditions are satisfied. As an example, we discuss a simple quantum mechanical model below.

5.1. The Infinite Square-Well Potential.

Assume a particle free to move along the x-axis, subjected to the potential conditions

$$V(x) = \begin{cases} +\infty, & |x| > a \\ 0, & |x| < a \end{cases}$$

Using the one-dimensional wave equation 5.3, we notice that infinite potential outside the interval $(-a, a)$ renders the wave equation meaningless. Thus, $\psi(x, t)$ represents a particle only when $-a < x < a$, in which case $V(x) = 0$. The wavefunction must also disappear at the boundry $x = \pm a$ due to discontinuity.

Let $\psi(x, t)$ be written as $\psi(x) f(t)$. We can now write 5.3 as

$$i\hbar\psi(x)\frac{\partial f(t)}{\partial t} = \frac{-\hbar^2}{2m} f(t)\frac{\partial^2\psi(x)}{\partial x^2}$$

which when divided by $\psi(x)f(t)$ gives

$$i\hbar\frac{1}{f(t)}\frac{\partial f(t)}{\partial t} = \frac{-\hbar^2}{2m}\frac{1}{\psi(x)}\frac{\partial^2\psi(x)}{\partial x^2}$$

Since t and x denote different variables, the above must be a constant. We call it E. Changing $\partial f/\partial t$ to df/dt and $\partial^2\psi/\partial x^2$ to $d\psi/dx$ we get:

$$\frac{d^2\psi(x)}{dx^2} = -E\frac{2m}{\hbar^2}\psi(x) \tag{5.1}$$

Solving 6.1 for ψ gives

$$\psi(x) = A\sin kx + B\cos kx$$

where

$$A, B \in \mathbb{R},\ \ k = \left(E\frac{2m}{\hbar^2}\right)^{\frac{1}{2}}.$$

We can now solve the simultaneous equations for $x = \pm a$

$$\begin{aligned} -A\sin ka + B\cos ka &= 0 \\ A\sin ka + B\cos ka &= 0 \end{aligned}$$

to obtain two possible cases

i. $B = 0$, and $A\sin ka = 0$

ii. $A = 0$, and $B\cos ka = 0$

In the first case,

$$ka = n\pi,\ \ E = \frac{n^2\pi^2\hbar^2}{2a^2m}\ (\text{where } n = 1, 2, 3, \dots)\ . \tag{5.2}$$

E for a particular n is denoted by E_n. We write $\psi(x) = A\sin kx$, and then normalize the wavefunction as follows:

$$\int_{-a}^{+a} |\psi(x)|^2 dx = \int_{-a}^{+a} A^2 \sin^2\frac{n\pi x}{a} dx = 1$$

which gives

$$A = \sqrt{\frac{1}{a}} \text{ and } \psi(x) = \sqrt{\frac{1}{a}} \sin kx.$$

$f(t)$ turns out to be $e^{-iE_nt/\hbar}$ where n is the same as in equation 6.2. Finally,

$$\psi(x,t) = e^{-iE_nt/\hbar}\sqrt{\frac{1}{a}} \sin \frac{n\pi}{a} x \quad (n = 1, 2, 3, \dots) \tag{5.3}$$

Solution is similar for the second case. More on this method and its consequences may be found in [**?**].

6. Postulates of Quantum Mechanics

The following are from [**?**]. For additional discussion see [**?**], 5.11.

6.1. **State-Space.** Every isolated physical system has a Hilbert space associated with it, which is called the state-space. The system is completely described by a unit vector in this space, which is called the state vector of the system.

6.2. **Evolution.** The change of state of an isolated quantum system is called its evolution. As mentioned previously, the evolution of a closed non-relativistic quantum system is described by the Schrödinger equation $i\hbar\frac{\partial|\psi\rangle}{\partial t} = \mathbf{H}|\psi\rangle$, where $|\psi\rangle$ is the state vector. The generalization of this postulate says that the state of a system at time t_1 is related to its state at time t_2 by a unitary[2] operator U which depends only on t_1 and t_2. If $|\psi_1\rangle$ and $|\psi_2\rangle$ are the state at time t_1 and t_2 respectively, then:

$$|\psi_2\rangle = U|\psi_1\rangle \tag{6.1}$$

6.3. **Measurement.** Measurement changes the state of a quantum system. The measurements are described by a set of measurement operators referred to as $\{M_m\}$ where m denotes the possible outcome of the measurement.

Given a state $|\psi\rangle$, the probability that m occurs is $p(m) = \langle\psi|{M_m}^\dagger M_m|\psi\rangle$. After measurement, the new state is given by:

$$\boxed{\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|{M_m}^\dagger M_m|\psi\rangle}}} \tag{6.2}$$

The probabilites sum to 1, so $\sum p(m) = 1$, and $\sum_m {M_m}^\dagger M_m = I$. Here I is the identity matrix.

6.4. **Composite Systems.** If there are n-physical components of a physical system, then its state space is the tensor product of the state spaces of its individual parts. This is written as

$$|\psi_1\rangle \otimes \cdots \otimes |\psi_n\rangle$$

where $|\psi_i\rangle$ is the state vector of the i^{th} system.

[2] has the property $U^\dagger U = I$

6.4.1. *Tensor Product.* A tensor product of two vector spaces $\mathcal{H}^n$ and $\mathcal{H}^m$ produces a new vector space $\mathcal{H}^{mn}$. Example of tensor product of two finite dimensional vectors:

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \otimes \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 1\times 4 \\ 1\times 5 \\ 2\times 4 \\ 2\times 5 \\ 3\times 4 \\ 3\times 5 \end{bmatrix}$$

If ψ_i and χ_j form a basis set for $\mathcal{H}^n$ and $\mathcal{H}^m$ respectively, then the basis for $\mathcal{H}^n \otimes \mathcal{H}^m$ is $\psi_i \otimes \chi_j$ where $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$.

7. Quantum Computing[?]

Let $|\psi_0\rangle$ and $|\psi_1\rangle$ form an orthonormal basis for a system's state space. Let α and β be two complex numbers. Then the ket $|\psi\rangle = \alpha|\psi_0\rangle + \beta|\psi_1\rangle$ is called qubit. In common literature, the states $|\psi_0\rangle$ and $|\psi_1\rangle$ are denoted by $|0\rangle$ and $|1\rangle$ respectively. (A qubit can be understood as a generalization of a classical bit of the form $|b\rangle = x|0\rangle + y|1\rangle$ where $x, y \in \{0, 1\}$ and $x \neq y$. A classical bit is therefore either in the state 1 or in the state 0, but not both simultaneously.)

If $\mathcal{H}$ is the vector space of all such qubits, then the basis for this two dimensional space can be represented in vector form as:

$$|0\rangle = |\psi_0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad |1\rangle = |\psi_1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

We can measure a qubit using the operators M_0 and M_1 which measure $|0\rangle$, and $|1\rangle$ respectively. $M_0 = |0\rangle\langle 0|$, and $M_1 = |1\rangle\langle 1|$. Applying M_0 to $|\psi\rangle$ we get

$$M_0|\psi\rangle = M_0\alpha|0\rangle + M_0\beta|1\rangle$$

which is

$$|0\rangle\langle 0|\alpha|0\rangle + |0\rangle\langle 0|\beta|1\rangle = \alpha|0\rangle + 0$$

The probability that we measure $|0\rangle$ is

$$p(0) = \langle\psi|M_0^\dagger M_0|\psi\rangle = \alpha^*\alpha = |\alpha|^2 \text{ (As } M_0^\dagger = M_0.\text{)}$$

As the above example shows, we cannot determine by measurment what state a qubit is in. When a qubit is measured, it shows properties of being in state $|0\rangle$ with a probability of $|\alpha|^2$, and appears to be in state $|1\rangle$ with a probability of $|\beta|^2$. Since there are no other states in consideration, $|\alpha|^2 + |\beta|^2 = 1$ to satisfy the normalization condition.

7.1. **Square Well and Qubits.** Let $|\psi_1\rangle$ and $|\psi_2\rangle$ correspond to $\psi(x,t)$ in equation 6.3 with $n = 1$ and $n = 2$ respectively. Then $|\psi\rangle = a|\psi_1\rangle + b|\psi_2\rangle$ may be treated as a qubit (from [**?**] page 280).

7.2. **Quantum Parallelism.** The properties of qubits allow us to generate a superposition of 'inputs'. Such a state represents many inputs simultaneously. An operation U on this state can be understood as being applied simultaneously to all the inputs that create the superposition. The outcome of this operation resembles the superposition of all outcomes we get upon performing U on the constituent states. However this process is not so useful, because individual results cannot be measured separately.

7.3. **Quantum Gates.** In general, quantum gates are unitary operations on qubits. From 7.1 we know

$$UU^\dagger = I.$$

Thus

$$|\psi_2\rangle = U|\psi_1\rangle \;\Rightarrow\; U^\dagger|\psi_2\rangle = |\psi_1\rangle$$

which means quantum gates can be reversible. One such example is the NOT gate:

$$\text{NOT} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \text{ and } \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^\dagger = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The NOT gate swaps the states of $|0\rangle$ and $|1\rangle$. Applying it to the qubit $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ is equivalent to:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \times \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \beta \\ \alpha \end{bmatrix} \equiv \beta|0\rangle + \alpha|1\rangle$$

Other well-known gates are as follows:

Hadamard (**H**)	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
Pauli-x (σ_x)	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Pauli-y (σ_y)	$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$
Pauli-z (σ_z)	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Phase (**S**)	$\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$
Toffoli (**T**)	$\begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$

As the gates above can act on one qubit only, they are called *single-qubit* gates.

7.4. **Multiple Qubits.** Two qubits belonging to spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ will belong to the space $\mathcal{H}_2 \otimes \mathcal{H}_2$ which is a four-dimensional Hilbert space. The four basis vectors for this new combined space can be written as:

$$|00\rangle,\ |01\rangle,\ |10\rangle, \text{ and } |11\rangle$$

8. Quantum Algorithms

Let $\mathcal{A}$ be an algorithm that fails with probability ϵ (where $0 < \epsilon < 1$). If we run it k times, then the probability that it fails each time is ϵ^k. Hence the probability of its sucess is $1 - \epsilon^k$ (more on this in [**?**]). If we are allowed to run $\mathcal{A}$ for arbitrarily large number of times, we can bring the probability of sucess very close to 1.

Quantum computing hinges on the peculiar way quantum-mechanical systems evolve. Using quantum parallelism we can perform several operations at once on a set of inputs, but measurement of the result is inherently probabilistic.

8.1. **Factoring of numbers[?].** Consider the function $f_N(a) = x^a(mod\, N)$, where x is coprime to N. According to a result from Number Theory, f is a periodic function, with a period r. That means, $f_N(a) = f_N(a+r)$. For example, putting $x = 8$ and $N = 15$ gives $r = 4$. Once r is known, a factor of N can be found as $gcd(x^{r/2} \pm 1, N)$. This happens when r is even. The problem of finding factors of N can be reduced to finding the order r of N given a particular x.

Peter Shor's celebrated algorithm provides an efficient way to compute order r of N for a randomly chosen x.

Steps to Shor's Algorithm[**?**]*:*

We denote a discrete fourier transform on a $q-dimensional$ vector as DFT_q. This transform is the engine of Shor's algorithm. Also we denote the prime factors of q by $q_i^{e_i}$. A q is defined as *smooth* if $q_i^{e_i} \leqslant A\,(\log q)^c$. DFT_q can be performed *efficiently* if q is smooth.

1. Given N, find a smooth q between N and N^2. Choose constants A, c independent of q.
2. Choose random $x(modN)$ and begin with a state labelled as $|N, x, q, 0\rangle$
3. Apply DFT_q to the last slot to obtain:

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

4. Compute $x^a\,(modN)$ for all a and store the result in a new slot giving:

$$\frac{1}{\sqrt{q}}\sum_{a=0}^{q-1}|N, x, q, a, x^a(modN)\rangle$$

This can be done efficiently due to quantum parallelism

5. Perform a measurement on the last slot, and measure $y(modN)$ where $y = x^l$. As the function $x(modN)$ is periodic with period r, this measurement also selects all a, where $a = l, l+r, l+2r, \ldots, Ar$ where A is the largest integer less than $(q-1)/r$.
6. The post measurement state is

$$|\psi_l\rangle = \frac{1}{\sqrt{A+1}}\sum_{j=0}^{A}|jr+l\rangle$$

Considering a situation when $A = (q/r-1)$ the above becomes

$$|\phi_{in}\rangle = \sqrt{\frac{r}{q}}\sum_{j=0}^{\frac{q}{r}-1}|jr+l\rangle$$

7. Let $M = \frac{q}{r}$ and perform DFT_q on the $|\phi_{in}\rangle$ to obtain

$$\begin{aligned}|\phi_{out}\rangle &= \sqrt{\frac{r}{q}}\sum_{c=0}^{q-1}e^{\frac{2\pi ilc}{q}}\sum_{j=0}^{M-1}e^{\frac{2\pi ijc}{M}}|c\rangle \\ &= \sum_{c=0}^{q-1}\alpha_c|c\rangle\end{aligned}$$

$$\Rightarrow \sum_{j=0}^{M-1}e^{\frac{2\pi ijc}{M}} = \begin{cases} M, & \text{if } c \text{ is a multiple of } M \\ 0, & \text{otherwise}\end{cases}$$

The amplitudes of $|c\rangle$ are

$$\alpha_c = \begin{cases} e^{\frac{2\pi i l c}{q}}, & \text{if } c \text{ is a multiple of } M \\ 0, & \text{otherwise} \end{cases}$$

So the DFT_q of a state with a period r is a state with period $M = q/r$.

$$|\phi_{out}\rangle = \frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} e^{\frac{2\pi i l c}{q}} |j\frac{q}{r}\rangle$$

Now a measurement will yield $\frac{\lambda q}{r}$, where $\lambda = 0, 1, 2, \ldots, r-1$. After our measurement of the label, we see a value of $|c\rangle$ satisfying $\frac{c}{q} = \frac{\lambda}{r}$ where c, q are known, and we can determine r.

9. Acknowledgement

I am grateful to Dr. Kurt Fisher, under whose guidence I completed this independent study project. Many thanks also to Dr. Raymond Grinnell and Dr. Herbert Bernstein for providing insights into the mathematics of the subject.

Appendix A. Used Symbols and Terms

The following is the list of symbols used in this paper and their meaning:

δ_{ij} : The Kronecker-delta. $\delta_{ij} = 1$ if $i = j$, else it is 0.
i : The imaginary number $\sqrt{-1}$.
ψ^* : The complex conjugate of ψ.
$\hbar$: Planck's constant, to be determined experimentally.
$\dagger$: Hermitian conjugate,
also denotes the inverse Fourier Transform

Appendix B. Roadmap

Important terms and definitions are presented in subsections (example: 2.3 **Basis Vectors**). New ideas mentioned in these are described subsequently. Vectors are in bold (example: $\mathbf{x}$) or with an arrow over (example: $\overrightarrow{u}$). Sets (and operators) are in bold caps (example: $\mathbf{F}$). It is assumed that a reader follows this paper from the first section to the last, and accordingly builds up from information in the previous sections. Used symbols and notation are borrowed from cites sources.

References

[1] Barenco A., Bennett C. H., Cleve R., DiVincenzo D. P., Margolus N., Shor P., Sleator T., Smolin J., Weinfurter H. *Elementary Gates for Quantum Computation* Physical Review A, March 22, 1995
[2] Berberian, Sterling K. *Introduction to Hilbert Space*, Oxford University Press, New York, 1961
[3] Brigham, E. Oran, *The Fast Fourier Transform and its Applications*, Prentice Hall, NJ, 1988
[4] Byron F. W. Jr., Fuller R. W. *Mathematics of Classical and Quantum Physics: Volume 1*, Addison-Wesley Publishing Company, Philippines, 1969
[5] Ekert, A., Jozsa R. *Nist Workshop on Quantum Computing and Communication, Supplementary Materials for Participants containing notes on Shor's Efficient Algorithm for factoring on quantum computer*, Gaithersburg, Maryland 1994.
[6] Gerlach, U. H. *http://www.math.ohio-state.edu/~gerlach/math/BVtypset/*, 2003
[7] Griffiths, D. J. *Introduction to Quantum Mechanics*, Prentice Hall, New Jersey, 1994

[8] Messiah A. *Quantum Mechanics: Volume 1*, North-Holland Publishing Company, Amsterdam, 1961
[9] Nielsen, M. A. and Chuang I. L. *Quantum Computation and Quantum Information*, Cambridge University Press, 2000
[10] Patel, Ketan N., Markov, Igor L., Hayes, John P. *Efficient Syhthesis of Linear Revirsible Circuits*, arXiv:quant-ph/0302002
[11] Rudin, W. *Principles of Mathematical Analysis*, McGraw-Hill, New York, 1953
[12] Shor, Peter W. *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*, Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, IEEE Computer Society Press, arXiv:quant-ph/9508027
[13] Weinstein Y.S., Pravia M.A., Fortunato E.M., Lloyd S., Cory D.G. *Implementation of the Quantum Fourier Transform*, Physical Review Letters 86, 1889-1891, 2001
[14] Zygelman, Bernard *http://www.physics.unlv.edu/ bernard/phy721_99/tex_notes/node6.html*, 2001

Dowling College, Oakdale, NY 11769
E-mail address: trip7294@dowling.edu