A CRASH COURSE IN BASIC QUANTUM THEORY PRELIMINARIES

Abstract. We study basic preliminaries for quantum theory as a preliminary of QFT(quantum field theory), mainly based on "Quantum Theory for Mathematicians", written by Brain C.Hall. Readers are assumed to have little knowledge of physics(because I did!), but basic knowledge of functional analysis is necessary.

1. Introduction

1.1. **Motivation.** Quantum mechanics is a very strange theory. It was not invented because anyone thought this is the way the world should behave, but because various experiments showed that this is the way the world does behave, like or not. Very roughly speaking, 'quantized' can be understood as 'discretized', and the main difference between classical mechanics and quantum mechanics comes from the 'randomness from discretization'.

Example 1.1 (Polarization)**.** *Consider the light passing through a filter. Suppose that light has an energy of E, and the angle between light and filter is $\theta \in [0, \pi/2]$ (for example, $\theta = 0$ means that light is aligned with filter). Then the amount of energy passing through the filter is measured by $E\cos\theta$.*

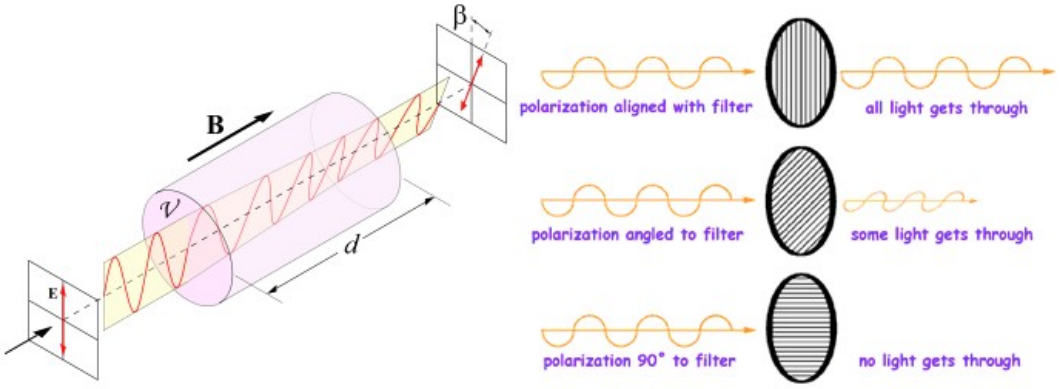

Figure 1. Polarization (Reference: https://www.askamathematician.com/2011/06/q-what-is-a-measurement-in-quantum-mechanics/)

On the other hand, from the theory of quantum mechanics, the energy of light is discretized. That is, energy E must be of the form

$$E = \hbar k \quad (k = 1, 2, 3, \cdots), \quad \text{for some positive constant } \hbar. \tag{1.1}$$

Then, here is a question: "What if we shoot a light with unit energy $\hbar$?" We cannot expect to obtain a light of energy $\hbar\cos\theta$ unless $\theta = 0$, because this will contradicts the assumption

Date: September 8, 2021.

(1.1). *The answer to the question is rather unexpected: we will get a light of energy $\hbar$ in probability $\cos^2\theta$, which means that light passes through the filter randomly!*

As an example 1.1 suggests, phenomena in quantum mechanics should be interpreted randomly(statistically) as determining the probabilities for observations of the system. Under this interpretation, the state of the system is not directly observable(recall that we could not determine if the light is the one that would pass through the filter until we observe the result), but described by the wave function ψ. In particular, if ψ is properly normalized, then the quantity $|\psi|^2$ is the probability distribution for the position of the particle. Even ψ itself is spread out over a large region in space, any measurement of the position of the particle will show that the particle is located at a single point. Such phenomena describe wave-particle duality, which will be explained right away.

There are two key ingredients to quantum theory, both of which arose from experiments. The first one is wave-particle duality, in which objects are observed to have both wave-like and particle-like behavior. The second one is its probabilistic behavior. Those two aspects of quantum theory come together in the wave function ψ, which is complex-valued[1]. The wave function is a function of a variable $x \in \mathbb{R}^d$, which we interpret as describing the possible values of the position of a particle, and it evolves in time according to a wave-like equation, the Schrödinger equation. The wave function and its time-evolution account for the wave aspect of quantum theory. The particle aspect of theory comes from the interpretation of the wave function. *Whenever we attempt to measure the position of a single electron, we always find the electron at a single point. The wave function merely tells us the probability distribution for its position.*

1.2. **Remind about Operators and their Adjoints.** In quantum mechanics, physical quantities such as position, momentum, and energy are represented by operators on a certain Hilbert space $\boldsymbol{H}$. These operators are (usually) unbounded operators, reflecting that in classical mechanics, these quantities are unbounded functions on the classical phase space. In this subsection, we look briefly at some technical issues related to unbounded operators and their adjoints.

From now on, $\boldsymbol{H}$ will represent a Hilbert space over $\mathbb{C}$, always assumed to be separable. We follow the convention in the physics literature that the inner product be linear in the second factor:

$$\langle \phi, \lambda\psi\rangle = \lambda\langle\phi,\psi\rangle, \quad \lambda\langle\phi,\psi\rangle = \bar{\lambda}\langle\phi,\psi\rangle, \quad \text{for all } \lambda\in\mathbb{C},\ \phi,\psi\in\boldsymbol{H}.$$

First, assume that $A:\boldsymbol{H}\to\boldsymbol{H}$ is a bounded operator, which means

$$\|A\psi\| \le C\|\psi\| \quad \text{for all } \psi\in\boldsymbol{H},$$

for some positive constant C, where $\|\cdot\|$ is a norm of $\boldsymbol{H}$. Then for each $\phi\in\boldsymbol{H}$, the map $\psi\mapsto\langle\phi, A\psi\rangle$ is a bounded linear functional on $\boldsymbol{H}$. Therefore, from Riesz's Theorem[2], there

[1]This is because a complex number suits perfectly for describing a wave. For $z = re^{i\theta}$, we can regard r and θ as magnitude and frequency of a wave, respectively.

[2]If $\zeta:\boldsymbol{H}\to\mathbb{C}$ is a bounded linear functional, then there exists a unique $\chi\in\boldsymbol{H}$ such that $\zeta(\psi)=\langle\chi,\psi\rangle$ for all $\psi\in\boldsymbol{H}$. Furthermore, the operator norm of ζ as a linear functional is equal to the norm of χ as an elemnt of $\boldsymbol{H}$.

exists the unique bounded operator A^*, called the adjoint of A, such that

$$\langle \phi, A\psi \rangle = \langle A^*\phi, \psi \rangle, \quad \text{for all } \phi, \psi \in \boldsymbol{H}.$$

A bounded operator A is said to be self-adjoint if $A^* = A$. For various reasons, which will be declared a bit later, we want the operators of quantum mechanics operators to be self-adjoint. However, then on is confronted with a serious technical difficulty: the operators are not bounded.

Why the boundedness matter? If A is a linear operator defined on all of $\boldsymbol{H}$ satisfying $\langle \phi, A\psi \rangle = \langle A\phi, \psi \rangle$ for all $\phi, \psi \in \boldsymbol{H}$, then A is automatically bounded. In other words, an unbounded self-adjoint operator cannot be defined on the entire Hilbert space. Thus, to deal with the unbounded operators of quantum mechanics, we must deal with operators that are defined only on a subspace of the relevant Hilbert space, called the domain of the operator.

Definition 1.1 (Unbounded operator and its domain)**.** *An unbounded operator A on $\boldsymbol{H}$ is a linear operator from a dense subspace $\mathcal{D}(A) \subset \boldsymbol{H}$ into $\boldsymbol{H}$.*

In defining the adjoint of an unbounded operator, we immediately encounter a difficulty: for a given $\phi \in \boldsymbol{H}$, the linear functional $\psi \mapsto \langle \phi, A\psi \rangle$ may not be bounded, in which case we cannot use the Riesz theorem to define $A^*\phi$. Hence adjoint of A will be defined not on all of $\boldsymbol{H}$ but only on some subspace thereof.

Definition 1.2 (Adjoint of an unbounded operator)**.** *For and unbounded operator A on $\boldsymbol{H}$, the adjoint of A is denoted A^* and $\tilde{\phi} := A^*\phi$ is the unique vector such that*

$$\langle \tilde{\phi}, \psi \rangle = \langle \phi, A\psi \rangle, \quad \textit{for all } \psi \in \mathcal{D}(A),$$

where $\mathcal{D}(A^*)$, domain of A^*, is defined by the collection of ϕ which makes

$$\psi \mapsto \langle \phi, A\psi \rangle$$

into 'bounded' linear functional defined on $\mathcal{D}(A)$.

Note that if $\mathcal{D}(A)$ is dense and $\psi \mapsto \langle \phi, A\psi \rangle$ is bounded, it has a unique bounded extension to whole $\boldsymbol{H}$ and then Riesz's theorem guarantee the uniqueness of $\tilde{\phi}$.

Now we are ready to define self-adjointness and some related notions for unbounded operators.

Definition 1.3 (Self-adjoint operator)**.** *Let A be an unbounded operator on $\boldsymbol{H}$.*

(1) A is symmetric if

$$\langle \phi, A\psi \rangle = \langle A\phi, \psi \rangle, \quad \textit{for all } \phi, \psi \in \mathcal{D}(A).$$

(2) A symmetric operator A is self-adjoint if

$$\mathcal{D}(A^*) = \mathcal{D}(A) \quad \textit{and} \quad A^* = A \textit{ on } \mathcal{D}(A).$$

(3) A symmetric operator A is essentially self-adjoint if it has a unique self-adjoint extension.[3]

[3]Equivalently, A is essentially self-adjoint if the closure in $\boldsymbol{H} \times \boldsymbol{H}$ of the graph of A is the graph of a self-adjoint operator.

Remark 1.1. *Suppose that A is a symmetric operator. Then for $\phi, \psi \in \mathcal{D}(A)$, we have*

$$\left|\left\langle \phi, A\frac{\psi}{\|\psi\|} \right\rangle\right| = \left|\left\langle A\phi, \frac{\psi}{\|\psi\|} \right\rangle\right| \leq A\|\phi\|,$$

and hence $\psi \mapsto \langle \phi, A\psi\rangle$ is a bounded linear functional. Therefore, $\phi \in \mathcal{D}(A^*)$ holds. Since ϕ can be chosen arbitrary in $\mathcal{D}(A)$, we have $\mathcal{D}(A) \subset \mathcal{D}(A^*)$. Because $A^* = A$ in $\mathcal{D}(A)$, we have:

(1) A^* is an extension of A.
(2) If $\mathcal{D}(A) = \mathcal{D}(A^*)$, then A is self-adjoint.

That is to say, A is self-adjoint operator if A^* and A are the same operator with the same domain. The reason a symmetric operator may fail to be self-adjoint is that $\mathcal{D}(A^*)$ may be strictly larger than $\mathcal{D}(A)$.

Although the condition of being symmetric is certainly easier to understand and to verify than the condition of being self-adjoint, the self-adjointness is the "right" condition: the spectral theorem, which is essential to much of quantum mechanics, applies only to self-adjoint operators but not to merely symmetric. If A is essentially self-adjoint, then we simply extend it into self-adjoint operator and apply the spectral theorem. We close this subsection by citing some additional reason that A should be self-adjoint(in fact, symmetric).

Proposition 1.1. *Suppose A is a symmetric operator on $\boldsymbol{H}$.*

(1) For all $\psi \in \mathcal{D}(A)$, the qunatity $\langle \psi, A\psi\rangle$ is real. More generally, if $\psi, A\psi, \cdots, A^{m-1}\psi$ all belong to $\mathcal{D}(A)$, then $\langle \psi, A^m\psi\rangle$ is real.
(2) Suppose λ is an eigenvector for A, meaning that $A\psi = \lambda\psi$ for some $0 \neq \psi \in \mathcal{D}(A)$. Then $\lambda \in \mathbb{R}$.

Proof. (1) Since A is symmetric, we have

$$\langle \psi, A^m\psi\rangle = \langle A\psi, A^{m-1}\psi\rangle = \cdots = \langle A^m\psi, \psi\rangle = \overline{\langle \psi, A^m\psi\rangle}.$$

(2) Since $\psi \neq 0$, the following computation

$$\lambda\langle \psi, \psi\rangle = \langle \psi, A\psi\rangle = \langle A\psi, \psi\rangle = \bar{\lambda}\langle \psi, \psi\rangle$$

implies $\lambda = \bar{\lambda}$.

□

Physically, $\langle \psi, A\psi\rangle$ represents the expectation value for measurements of A in the state ψ, whereas the *eigenvalue λ represents one of the possible values for this measurement.* On physical grounds, we want both of these numbers to be real. If A is self-adjoint, then the spectral theorem will give a canonical way of associating to each $\psi \in \boldsymbol{H}$ a probability measure on the real line that encodes the probabilities for measurements of A in the state ψ.

2. Physical Quantities and their Operators

2.1. **Position Operator.** Let us consider a single particle moving on the real line. The wave function for such a particle is a map $\psi : \mathbb{R} \to \mathbb{C}$. Although this map will evolve in time, let us think for now that the time is fixed. *The function $|\psi(x)|^2$ is supposed to be the*

probability density for the position of the particle. This means that the probability that the position of the particle belongs to some Borel set $E \subset \mathbb{R}$ is

$$\int_E |\psi(x)|^2 dx.$$

For this prescription to make sense, ψ should be normalized so that

$$\int_{\mathbb{R}} |\psi(x)|^2 dx = 1, \quad \text{i.e. } \psi \text{ should be a unit vector in } \boldsymbol{H} = L^2(\mathbb{R}),$$

and expectation value of position and any moment(= expectation value of some power of the position) is

$$\boldsymbol{E}(x) = \int_{\mathbb{R}} x|\psi(x)|^2 dx, \quad \boldsymbol{E}(x^m) = \int_{\mathbb{R}} x^m |\psi(x)|^2 dx, \tag{2.1}$$

provided that the integral is absolutely convergent.

A key idea in quantum theory is to express expectation values of various quantities in terms of operators and the inner product on the relevant Hilbert space, in this case, $L^2(\mathbb{R})$. In the case of position, we may introduce that position operator X defined by

$$(X\psi)(x) := x\psi(x).$$

Then the expectation values of position and moment in (2.1) may now be written as

$$\boldsymbol{E}_\psi(X) = \langle \psi, X\psi \rangle, \quad \boldsymbol{E}_\psi(X^m) = \langle \psi, X^m \psi \rangle. \tag{2.2}$$

Remark 2.1.

(1) *At this point, it is not clear that we have gained anything by writing moments in terms of an operator and the inner product (2.2) instead of in terms of the integral (2.1). The operator description will, however, motivate a parallel description of moments for the momentum, energy, and so on.*
(2) *For a given $\psi \in \boldsymbol{H} = L^2(\mathbb{R})$, $X\psi$ might fail to be in $\boldsymbol{H}$. This failure reflects that X is an unbounded operator. Even if $X\psi \in \boldsymbol{H}$, $X^m\psi$ might fail to be in $\boldsymbol{H}$ for some m.*

2.2. **Momentum Operator.** Although the wave function ψ directly encodes the probabilities for the position of the particle through $|\psi(x)|^2$, it is not as clear *how information about the particle's momentum is encoded.* A crucial idea in quantum mechanics is the *de Broglie hypothesis*, which proposes a particular relationship between the frequency of oscillation of the wave function and its momentum.

Proposition 2.1 (de Broglie hypothesis)**.** *If the wave function of a particle has spatial frequency k, then the momentum p of the particle is*

$$p = \hbar k,$$

where $\hbar$ is Planck's constasnt.[4]

[4]Spatial frequency means the frequency in space variable x. On the contrary, temporal frequency means frequency in time variable t and this will be considered in time-evolution of wave function i.e. Schrödinger Equation.

Let us first consider $\boldsymbol{H} = L^2([0, 2\pi])$ and assume that ψ us 2π-periodic in x. *Proposition 2.1 is supposed to mean that a wave function of the form $\psi = e^{ikx}$ represents a particle with momentum $p = \hbar k$.* Hence, it makes sense to say that the normalized wave function $\psi = e^{ikx}/\sqrt{2\pi}$ represents a definite(=nonrandom) momentum $p = \hbar k$ for any integer k. In other word, a measurement of the particle's momentum should(=with probability 1) give the value $\hbar k$.

On the other hand, since $\{e^{ikx}/\sqrt{2\pi}\}_{k\in\mathbb{Z}}$ forms an orthonormal basis of $L^2([0, 2\pi])$, unit vector ψ is represented as

$$\psi(x) = \sum_{k=-\infty}^{\infty} a_k \frac{e^{ikx}}{\sqrt{2\pi}}, \quad \text{where} \quad \sum_{k=-\infty}^{\infty} |a_k|^2 = \|\psi\|^2_{L^2([0,2\pi])} = 1, \tag{2.3}$$

where the sum of first equation is convergent in $L^2([0, 2\pi])$. For a particle with wave function ψ in (2.3), the momentum of the particle is no longer definite. Rather, we are supposed to think that a measurement of the particle's momentum will yield one of the values $\hbar k$ ($k \in \mathbb{Z}$), with the probability of getting a particular value $\hbar k$ being $|a_k|^2$. Hence the expectation values for the momentum and its higher moments should be

$$\boldsymbol{E}(p) = \sum_{k=-\infty}^{\infty} \hbar k |a_k|^2, \quad \boldsymbol{E}(p^m) = \sum_{k=-\infty}^{\infty} (\hbar k)^m |a_k|^2. \tag{2.4}$$

We would like to encode (2.4) in a momentum operator P, which should be consistent with (2.1): $\boldsymbol{E}(p^m) = \langle \psi, P^m \psi \rangle$. This is achieved if P satisfies

$$Pe^{ikx} = \hbar k e^{ikx} = -i\hbar \frac{d}{dx} e^{ikx} \text{ for each } k, \quad \text{i.e.} \quad P = -i\hbar \frac{d}{dx}. \tag{2.5}$$

Turning to $\boldsymbol{H} = L^2(\mathbb{R})$, now it is natural to postulate the momentum operator P on $\mathbb{R}$ also should be given by the one satisfying (2.5). Although e^{ikx} is not in $\boldsymbol{H}$, the Fourier transform allows us to build up $\psi \in \boldsymbol{H}$ as a "superposition" of functions of the form e^{ikx}. This means that, by analogy to $(2.3)_1$, we have

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ikx} \hat{\psi}(k) dk,$$

where $\hat{\psi}(k)$ is the Fourier transform of ψ, defined by

$$\hat{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-ikx} \psi(x) dx.$$

Since Fourier transform is a unitary map of $\boldsymbol{H}$ onto $\boldsymbol{H}$ from Plancherel theorem, unit vector $\psi \in \boldsymbol{H}$ satisfies

$$\int_{\mathbb{R}} |\psi(x)|^2 dx = \int_{\mathbb{R}} \left|\hat{\psi}(k)\right|^2 dk = 1.$$

In light of $(2.3)_2$, *it is natural to think that $\left|\hat{\psi}(k)\right|^2$ is essentially the probability density for the (scaled) momentum $k = p/\hbar$ of the particle.*

Proposition 2.2. *Define the momentum operator P by*

$$P = -i\hbar \frac{d}{dx}.$$

Then for all sufficiently nice(for example, Schwartz function) unit vectors ψ in $L^2(\mathbb{R})$, we have

$$\langle \psi, P^m \psi \rangle = \int_{\mathbb{R}} (hk)^m \left| \hat{\psi}(k) \right|^2 dk, \tag{2.6}$$

for all positive integers m. The quantity in (2.6) is interpreted as the expectation value of the mth power of the momentum, denoted by $\boldsymbol{E}_\psi(P^m)$.[5]

Proof. Suppose that ψ is in, say, the Schwartz space. Then we have

$$\widehat{\left(\frac{d^m}{dx^m} \psi \right)} = (ik)^m \hat{\psi}(k), \quad \text{so that} \quad \widehat{P^m \psi} = (hk)^m \hat{\psi}(k).$$

Meanwhile, since the Fourier transform is unitary, we have

$$\langle \psi, P^m \psi \rangle = \left\langle \hat{\psi}(k), (\hbar k)^m \hat{\psi}(k) dk \right\rangle = \int_{\mathbb{R}} (hk)^m \left| \hat{\psi}(k) \right|^2 dk.$$

□

[5]Schwartz space on $\mathbb{R}^d$, say $\mathcal{S}(\mathbb{R}^d)$, is a space of smooth function defined on $\mathbb{R}^d$, whose any derivative decays faster than arbitrary polynomial. Since $\mathcal{S}(\mathbb{R}^d)$ is closed under differentiation and Fourier transform maps $\mathcal{S}(\mathbb{R}^d)$ into $\mathcal{S}(\mathbb{R}^d)$, Schwartz space might be regarded as a sort of safe zone, free from any danger which might occur differentiation when dealing with a Fourier transform.