How to Rise from Ashes

A Comprehensive Guide to Finally Understand Canonical Quantization and Build Quantum Mechanics from Scratch

Karl Oscar Herlin

April 2021

Abstract

In this review essay, we account for the historical development of canonical quantization. By a detailed review of *The Ultraviolet Catastrophe*, *The Photoelectric Effect*, *Compton Scattering*, *Matter Waves* and *Bohr's Atomic Model* we learn about *Old Quantum Theory*. Through the realizations of old quantum theory we account for the development of *Heisenberg-Born-Jordan Matrix Mechanics* and mention in passing the development of *Schrödinger's Quantum Wave Mechanics* to understand how *Canonical Quantization* was developed by P.A.M. Dirac, which we also account for. Through Canonical Quantization we develop the *Schrödinger and Heisenberg equations of motion* from first principles using displacement operators to describe the *Double Slit Experiment* only using results contained in the essay. We conclude by stating that the three main questions posed by this essay,*"How exactly did physicists figure out that classical physics was insufficient?"*, *"Why do we use Hermitian operators in Hilbert space, and how did physicists infer the commutation relations of quantum mechanics?"* and *"How can the equations of motion be developed from first principles?"* have been answered, albeit in realization that no universal scheme for quantizing classical theories exist, leading to recommendations for further study.

Contents

Acknowledgements

First and foremost, I would like to thank Professor Lars-Göran Johansson for our long discussions regarding quantum mechanics and scientific theory, and everything in between. Professor Johansson has been an invaluable asset for me, never accepting a faulty argument and always ready to pull my text appart, as any invested supervisor should do.
Last but not least, I would like to thank my family, Bo, Rose-Marie and Emelie Herlin, as well as my partner Ella Elmgart, for always being there for me at times of stress, excitement and relaxation.

1 Introduction

The goal of the first and second courses on quantum theory is to learn how to *do* quantum mechanics. We are taught how to solve the Schrödinger equation in a number of potentials, how to use perturbation theory and, if we continue with an advanced course, use Dirac notation and present quantum mechanics with the use of Hermitian operators. Often, introductory courses on quantum mechanics starts by presenting a couple of early experiments leading to Planck's radiation law, a corpuscular description of light or perhaps a stationary state description of the Hydrogen atom. As a consequence of these results, we are told that "Therefore, classical physics is insufficient, and a quantum theory of nature is needed.". Proceeding the discussion, sometimes on the very next blackboard, is the Schrödinger equation in one dimension. We are taught how to obtain the time-independent Schrödinger equation, and from there the inexorable quantum train moves on towards greater complexity.

As for many students before me, this led to an enormous confusion. As quantum theory developed before us, questions started to creep up along the way. How exactly did physicists figure out that classical physics was insufficient? Why do we use Hermitian operators in Hilbert space, and how did physicists infer the commutation relations of quantum mechanics? How can the equations of motion be developed from first principles? *How do we build a theory of quantum mechanics from scratch?*

Learning quantum mechanics can feel like you are being the target of the most elaborate April fools joke ever devised, constantly waiting for the professors to eventually crack up and tell you that it was a joke all along. Reluctantly however, your intuition is slowly but steadily crushed under the weight of the formalism, leaving you with stars around your head and, more often than not, smoke coming out of your ears. This essay, in essence, aims to be a treatment for these symptoms. It is directed towards students of physics who have taken at least two courses on quantum mechanics, meaning that they should be comfortable with Dirac notation and some standard mathematical methods used in quantum mechanics.

This essay is built up by four parts. Firstly, a detailed and chronological description of the early experiments of quantum mechanics will be provided. After this, we will take a closer look on how the formal structure of quantum mechanics was developed, ending in presenting the scheme of canonical quantization and why the postulates take the form they do. This discussion is continued by developing some familiar results of quantum mechanics. As a *coup de grâce*, we will end by providing a self contained theoretical description of the double slit experiment, meaning that all concepts used have been introduced and accounted for in the preceding discussion. As is customary, a discussion, conclusion and summary are provided at the end.

2 The Need for a Quantized theory

It is time to enter the head of an early twentieth century physicist. Only this way, without our quantum intuition, can we truly understand the development of quantum mechanics. Our discussion will begin with shedding light on the early developments of quantum physics and end in the Bohr model of the atom, where a classical explanation of nature should be understood to be insufficient.

2.1 The Ultraviolet Catastrophe

In 1859, long before the birth of quantum physics, Gustav Kirchhoff was studying the properties of black body radiation.[1] By creating a small hole piercing into a box containing heated material, he discovered a universal relationship

[1] Gasiorowics, Stephen; *"Quantum Physics"*; john Wiley Sons, Inc., third edition (2003), p. 1.

between the energy emitted per unit area and unit time, $E(\lambda, T)d\lambda$, where λ is the wavelength of the radiation and T the temperature of the radiator, and the energy density of radiation inside the box, $w(\lambda, T)d\lambda$, as

$$w(\lambda, T) = \frac{4E(\lambda, T)}{c}. \tag{1}$$

This seemed to be true for any material being heated, and peaks into the essence of matter itself. Years later, in 1894, Wilhelm Wien deduced that

$$w(\lambda, T) = \frac{f(\lambda T)}{\lambda^5} \tag{2}$$

from the laws of electrodynamics, a derivation that agreed with experiments. Problematically, even though the form of this expression agreed with measurements, derivations of $f(\lambda T)$ could not be obtained from classical physics. Famously, J.W.S. Reyleigh proposed a solution based on the equipartition of energy, as

$$w(\lambda, T) = \frac{8\pi kT}{\lambda^4}, \tag{3}$$

where k is Boltzmann's constant. This solution highlights the central problem that physicists faced during the shift from classical to quantum in the study of thermodynamics. In integrating this expression over all wavelengths to obtain the total energy per unit volume, we see that

$$\int_0^{+\infty} w(\lambda, T)d\lambda = \infty, \tag{4}$$

which obviously is not the case. Because solutions to the structure of $f(\lambda T)$ often faced problems of infinite energy density at wavelengths approaching zero, this problem is called *The Ultraviolet Catastrophe*.

This essential problem was to see its solution in the very first year of the next century in the form of Planck's law of blackbody radiation,

$$w(\lambda, T) = \frac{2\pi hc^2}{\lambda^5}\frac{1}{e^{\frac{hc}{\lambda kT}} - 1}. \tag{5}$$

It is based on the fundamental realization that energy can only be emitted and absorbed in *quanta*, and were to have energy equal to $E = \frac{hc}{\lambda}$, where h is the familiar Planck's constant, whose value were to be determined later in the twentieth century. Planck deduced his formula by theoretically examining the absorption and emission of light inside the Kirchhof box discussed earlier, where we are to imagine its structure as quantized harmonic oscillators. The derivation made by Planck is long and not the focus of this paper, but from other realizations of quantum theory outlined below one can derive his radiation law in a simpler fashion[2]. What is important to us is to realize that his insight

[2]Einstein, Albert; *"On the Quantum Theory of Radiation"*; Phys. Zs. 18 121, (1917).

into the quantization of emission and absorbtion of radiation leads to a theory of physics in line with nature. For example, in integrating over all possible wavelengths, we obtain

$$\begin{aligned} & \int_0^{+\infty} w(\lambda,T)d\lambda \\ &= 2\pi hc^2 \int_0^{+\infty} \frac{1}{\lambda^5}\frac{1}{e^{\frac{hc}{\lambda kT}}-1}d\lambda \\ &= 2\pi hc^2 \left(\frac{kT}{hc}\right)^5 \int_0^{+\infty} \frac{x^5}{e^x-1}\left|\frac{d\lambda}{dx}\right|dx \\ &= 2\pi hc^2 \left(\frac{kT}{hc}\right)^5 \frac{hc}{kT}\int_0^{+\infty} \frac{x^3}{e^x-1}dx, \end{aligned} \tag{6}$$

where

$$\begin{aligned} \int_0^{+\infty} \frac{x^3}{e^x-1}dx &= \int_0^{+\infty} x^3 e^{-x}\frac{1}{1-e^{-x}}dx = \int_0^{+\infty} x^3 e^{-x}\sum_{n=0}^{\infty}(e^{-x})^n dx \\ = \int_0^{+\infty} x^3 \sum_{n=0}^{\infty}(e^{-x})^{n+1}dx &= \int_0^{+\infty} x^3 \sum_{n=1}^{\infty} e^{-nx}dx = \sum_{n=1}^{\infty}\int_0^{+\infty} x^3 e^{-nx}dx \\ = \sum_{n=1}^{\infty}\int_0^{+\infty} \frac{y^3}{n^3}e^{-y}\left|\frac{dx}{dy}\right|dy &= \sum_{n=1}^{\infty}\left(\frac{1}{n}\right)^4\int_0^{+\infty} y^3 e^{-y}dy. \end{aligned} \tag{7}$$

It is clear that

$$\int_0^{+\infty} y^3 e^{-y}dy = 3! = 6 \tag{8}$$

and

$$\sum_{n=1}^{\infty}\frac{1}{n^4} = \frac{\pi^4}{90}, \tag{9}$$

so

$$\int_0^{+\infty} \frac{x^3}{e^x-1}dx = \frac{\pi^4}{90}6 = \frac{\pi^4}{15}. \tag{10}$$

Hence, we are led to believe that

$$\int_0^{+\infty} w(\lambda,T)d\lambda = 2\pi hc^2\left(\frac{kT}{hc}\right)^5\frac{hc}{kT}\frac{\pi^4}{15} = \frac{2\pi^5 k^4}{15h^3c^2}T^4. \tag{11}$$

Of course, this is *Stefan-Boltzmann's law*, which had been confirmed experimentally before Plancks law. It is a significant result that the energy density of radiation only depends on the temperature of the radiator. What is even more significant, though, is the fact that Planck through his theory, which very precise measurements have confirmed[3], solved the ultraviolet catastrophe, and did so by shattering our intuition of a universe estranged to discontinuities.

2.2 The Photoelectric Effect

It is important to understand what Planck's realisation of thermodynamics actually was. As carefully put above he theorized that the absorption and emission of light itself came and went in quanta due to the quantized harmonic oscillators inside the Kirchhoff box, but radiation itself was to follow the familiar classical laws of electrodynamics. That is to say, the idea of particulate radiation had not yet come to fruition. Of course, the idea had long been discussed by physicists[4], but no concrete model had been put forward, particularly after Maxwells equations. That was until Albert Einstein came forward with a model of radiation explaining *The Photoelectric Effect*.

In 1887, Heinrich Hertz discovered the Photoelectric effect almost by accident. His experimental work was dictated towards trying to build a receiver for radio waves. It consisted of a mechanism creating a spark gap, made in such a way to induce a current in a secondary circuit coil by means of electromagnetic radiation. In trying to see the spark in the secondary circuit better, he put it inside a dark box, separated from the primary circuit. To his surprise, the length of the spark gap in the secondary circuit had been reduced by this operation. This intrigued him, and after more experiments concerned with changing the material of the box, blocking different parts of the light spectra, he concluded that UV-radiation and its more energetic counterparts were responsible for the increase in the spark gap length. Thus, it became clear that high energy radiation releases electrons from metal surfaces.

Let us now imagine a simple experiment that we could set up in order to explore the photoelectric effect. We set up a circuit containing a variable power supply with an ammeter and a voltmeter placed appropriately. Closing the circuit is a metal plate connected to the cathode that we illuminate with an external light source, separated from another metal plate that is connected to the anode absorbing the electrons emitted from the other. This way, we create a retarding and stopping potential for the electrons. The type of metal used for the two plates does not change the essential results of the experiment, and can be of your own choosing. Concerning the specifics of the photoelectric effect, some peculiar and troublesome phenomena come to light if we were to run our experiment in the lab[5]:

[3] Xiaomin, Lu; Qishan, Tang; Wendong ,Kang; *"Research and Verification of Blackbody Radiation Law"*; Insight - Energy Science, (2018).

[4] J, Müller; *"Physikens Grunder"* (English, *"The foundations of Physics")*; Stockholm, Zacharias Hæggström, (1854).

[5] Greenstein, George; Zajonc, Arthur G.; , *"The Quantum Challenge: Modern Research on the Foundations of Quantum Mechanics"*; Jones and Bartlett Publishers Inc., second edition (2006), p. 24-37.

1) Electrons are emitted and reach current stability very quickly after the onset of the experiment, in the order of about 10^{-9} seconds.

2) The current decreases with increasing retarding potential V, and eventually stops at the stopping potential V_0.

3) The stopping potential V_0 increases linearly with the frequency ν of radiation, but $V_0 = 0$ at a certain threshold frequency ν_0 specific to the metal used.

4) The current increases linearly with light intensity.

Using our classical model of electrodynamics, some of these results are hard to explain. The idea of a classical photoelectric effect is however not a foreign one. After the photoelectric effect was discovered, but before the emergence of quanta, the existence of electrons was known through the experimental work of J.J. Thompson[6]. Classical electrodynamics allows radiation to be absorbed by electrons in order to accelerate them, but the devil is in the details. In classical electrodynamics, we are to imagine a light wave continuously increasing the energy of the electron until it has left the illumination zone, creating a delay. Of course, this delay should also be dependent on the intensity of the radiation, with an illumination of higher energies decreasing the departure time of the electrons. Of course, a retarding potential in the electron path should slow them down continuously. Thus, observation 2) and 4) can be adequately explained in a classical framework. Concerning observation 1) however, classical electrodynamics demands the delay until current stability is reached to be much larger than that of experiments. The time to reach current stability is also not proportional at all to the intensity of radiation. Concerning observation 3), we have clear evidence that frequency increases the energy of electrons over a certain threshold. This is not explained by our classical model at all. What is needed here is a model that could explain all of these four phenomena at once.

In his *anno mirabilis*, Einstein came up with a possible theoretical solution using a particulate theory of light. Tracing back our thoughts to Planck's solution to the ultraviolet catastrophe, Einstein went one step further. Not only did he imagine the harmonic oscillators in matter to be quantized, he also imagined the energy of light to be. We shall hesitate in leaping towards calling his particulate light *particles*, or *photons*, for now, for reasons that will be cleared up later. Einstein's particulate light follows the same formula for energy quantization as Planck's harmonic oscillators, that is, its energy is given by the equation $E = h\nu$. This means, that as a light quanta of energy $E = h\nu$ hits an electron in our experiment above, all of its energy is transferred. Unlike a wave which linearly transfers energy to the electron, a light quanta should transfer the set

[6] Thompson, J.J.; *"Cathode Rays"*; London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, (1897).

amount $h\nu$ instantaneously. However, Einstein imagined there to be a *work function* $W = h\nu_0$ for the metal being illuminated, expected to be different for each metal. The work function gives the energy required to release an electron from the forces holding it inside the metal. If this energy limit is not met, the electron will not absorb the light quanta at all and stay in place. We can imagine this process using the equation

$$E_k = h\nu - W = h(\nu - \nu_0) \tag{12}$$

where E_k is the maximum kinetic energy of the electrons, if indeed $\nu > \nu_0$. Otherwise, we set $E_k = 0$.

Is this model sufficient to explain all four observations above? Indeed it is. As already discussed, light quanta can be absorbed much quicker than a light wave, and hence the small current stabilisation time of observation 1) is no longer an issue. The condition of $\nu > \nu_0$ also explains the threshold frequency, and through equation (12) the linear relationship thereafter. Setting $E_k = eV_0$, where e is the electron charge, equation (12) also gives us the estranged relationship between frequency and stopping potential,

$$V_0 = \frac{h}{e}(\nu - \nu_0), \tag{13}$$

thusly ridding us of the theoretical problems concerning observation 3). As in the classical framework, electrons will be slowed down continuously in this new model, encapsulating observation 2). Intensity, being defined as energy per unit time, corresponds to more light quanta $h\nu$ instead of a higher amplitude of a classical wave. Hence, we should expect to see a linearly increasing current with higher light intensity, even in this model.

Einstein had succeeded. A new theory for the photoelectric effect encapsulating all observations, those which had been thought understood and those which had not, gained him the Nobel prize in 1921. Though the search for photons went on long after his discovery, it had yet again been shown that classical physics was not sufficient to explain all phenomena of nature.

2.3 Compton Scattering

In the above section I have knowingly abstained from calling light quanta *particles*. Einstein established an explanation for the photoelectric effect using particulate energy, but a particle has additional properties such as momentum and a definite position in space. Of course, the theory of quantum particles does not give the same restraint of position in space as classical particles do. On the other hand, *Compton Scattering*, discovered by Arthur H. Compton in 1923-4[7], gives light to the question of whether light quanta carries momentum or not, of

[7] Gasiorowics, Stephen; "*Quantum Physics*"; john Wiley Sons, Inc., third edition (2003), p. 7

which the answer is of course yes. In this section we will go through Compton's experimental work and the theory used to explain his findings.

Compton scattering concerns electron-X-ray scattering, which was a know process long before Compton. It was known that as electromagnetic radiation interacts with a free electron at rest, the electron recoils and the radiation changes the direction of propagation at an angle θ relative to the incident angle. Classically, this is described by the familiar process of Thompson scattering. J.J. Thompson theorised that as radiation interacts with an electron it will start to oscillate in unison with the electric field of the radiation and induce another electromagnetic field, thus changing the final direction. Its intensity I was calculated to be

$$I = 1 + \cos^2 \theta, \tag{14}$$

which had been confirmed experimentally.[8] As we can see, the intensity does not depend on wavelength. A break from this theory was formed by Compton who in 1923 published his results born from an experiment set up as follows, shown in fig. (1).[9] An X-ray tube emits radiation on a thin block of graphite, thus scattering the radiation. A set of slits are placed at θ degrees relative to the incident X-ray to ensure that the radiation measured indeed is scattered from this precise angle, with one slit made from lead to ensure lower experimental noise. Finally the radiation is measured using a Bragg spectrometer giving the intensity as a function of wavelength in our final data. The results of this experiment is given in fig. (2). As can bee seen, for the high energy X-rays, the intensity of the Kα spectral line does not only follow a relationship close to that of Thompson scattering, but at higher θ an additional peak of intensity is seen for wavelengths increasingly further away from the first.

How can this new phenomenon be explained? We can start from the familiar energy-mass equivalence relation,

$$E = \sqrt{(mc^2)^2 + (pc)^2} \tag{15}$$

for a particle with rest mass m and momentum p. The velocity at this momentum is then

$$v = \frac{dE}{dp} = \frac{pc^2}{E} \frac{pc^2}{\sqrt{(mc^2)^2 + (pc)^2}}. \tag{16}$$

For an X-ray, $v = c$ which implies a rest mass of zero, giving us

$$E = pc. \tag{17}$$

[8] Compton, H. Arthur. *A Quantum Theory of the Scattering of X-rays*, Phys. Rev. 21,483; 22,409 (1923).

[9] Compton, H. Arthur. *The Spectrum of Scattered X-rays*, Phys. Rev., vol. 22 No 5 (1923).

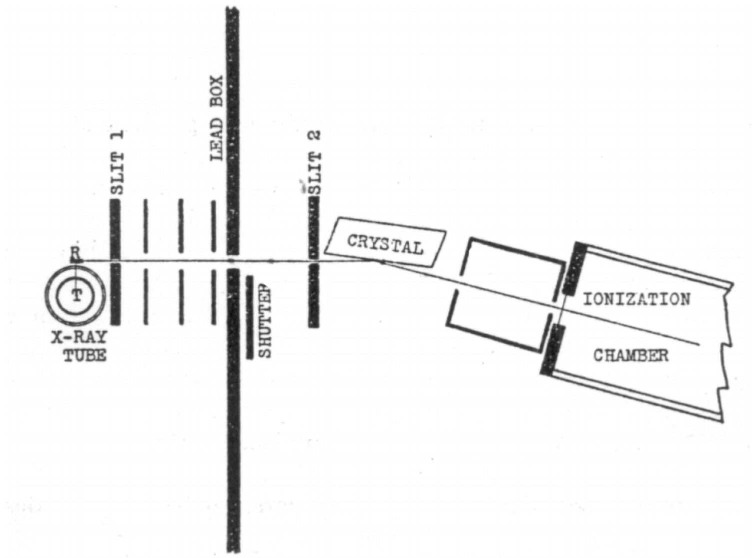

Figure 1: Setup of Compton's Scattering experiment, where "crystal" is the thin block of graphite and "X-ray tube" is the Bragg spectrometer. (Picture copied from: Compton, H. Arthur. *The Spectrum of Scattered X-rays*, Phys. Rev., vol. 22 No 5 (1923))

We can then use Planck's relation $E = h\nu$, giving us the final momentum relation

$$p = \frac{h\nu}{c}. \tag{18}$$

Consider now the momentum conservation corresponding with this scattering process,

$$p = p' + P, \tag{19}$$

where p and p' are the momentum of radiation before and after scattering, respectively, and P is the momentum of the electron after scattering. Consider also the energy conservation of this process,

$$h\nu + m_e c^2 = h\nu' + \sqrt{(m_e c^2)^2 + (Pc)^2} \tag{20}$$

where the same notation is used. We rearrange and square, as

$$m_e^2 c^4 + P^2 c^2 = (h\nu - h\nu' + m_e c^2)^2 = (h\nu - h\nu')^2 + 2m_e c^2 (h\nu - h\nu') + m_e^2 c^4, \tag{21}$$

that is,

$$P^2 c^2 = (h\nu - h\nu')^2 + 2m_e c^2 (h\nu - h\nu'). \tag{22}$$

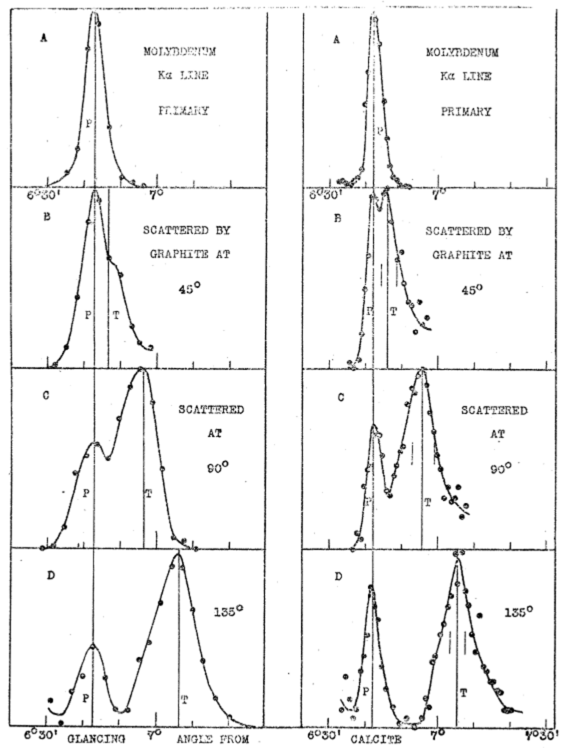

Figure 2: Compton's results. The results show the spreading of the Kα spectral line at two different slit widths (left to right) as well as for different angles θ (up to down). (Picture copied from: Compton, H. Arthur. *The Spectrum of Scattered X-rays*, Phys. Rev., vol. 22 No 5, p. 409 (1923))

Now going back to momentum conservation, we can use four-momentum invariance as found in relativity to find that

$$P^2 = (p - p')^2 = p^2 + p'^2 - 2p \cdot p', \tag{23}$$

$$P^2 = \left(\frac{h\nu}{c}\right)^2 + \left(\frac{h\nu'}{c}\right)^2 - 2\frac{h\nu}{c} \cdot \frac{h\nu'}{c} \cos\theta, \tag{24}$$

rearranged as

$$P^2c^2 = (h\nu - h\nu')^2 + 2(h\nu)(h\nu')(1 - \cos\theta). \tag{25}$$

We now equate our two relationships obtained from momentum and energy equivalence and rearange, as

$$(h\nu - h\nu')^2 + 2m_ec^2(h\nu - h\nu') = (h\nu - h\nu')^2 + 2(h\nu)(h\nu')(1 - \cos\theta) \tag{26}$$

$$\frac{\nu - \nu'}{\nu\nu'} = \frac{h}{m_e c^2}(1 - \cos\theta), \tag{27}$$

Which we rewrite as

$$\lambda' - \lambda = \frac{h}{m_e c}(1 - \cos\theta). \tag{28}$$

This way, Compton had found a new formula for scattering of electromagnetic radiation. In implementing his formula, we are to view radiation as interacting with a carbon nuclei as well as an electron surrounding it, which is assumed to be free due to the high energy of the X-ray. The mass of the entire atom is large ($m_e << m_C$), and so the shift in wavelength is thus very small, and very difficult to detect using the experiments of his day. Contrast this to the electron, where we should observe a shift due to its small mass. This formula found tremendous agreements with his experiments. His insight lied in using the particulate nature of light outside of the photoelectric effect, showing that giving it momentum could describe new experimental findings. His original argument in the referenced paper is more thorough, using the energy and momentum of an electromagnetic wave, but this way we have seen the essence of his argument. Physicists now had experimental backing in assuming that light comes in quanta, with a particulate energy and a given momentum. I have made an effort in our journey to this point not to call it by its more familiar name that it now came to carry; *photons*. Finally, physicists had reason to suspect that light consisted of particles, even though the debate of its essence was far from over.

2.4 Matter waves

In the very same year as Compton's experimental work, Louis De Broglie pondered the "double nature" in the essence of matter.[10] He was particularly intrigued by the way the study of quantum physics had developed a contradictory language concerning a complementarity between particles and waves. De Broglie asked himself how it could be that frequency was used for describing the corpuscular energy of light quanta, $E = h\nu$, a phenomenon classically only used for modeling waves. On the other hand, stationary states for electrons in a spherically symmetric potential were known to exist. The theory was one containing whole numbers as will be discussed in the next section, and this was only known to be present in interference phenomenon and quantized oscillators. Was it then possible for electrons, thought as particles through J.J. Thompson and H.A. Lorentz, to have wave-like properties after all?

De Broglie started with the assumption that a particle indeed has a frequency. He imagined this particle in its inertial system A, with its phase given by

$$\phi_0 = \sin 2\pi\nu_0 t_0 \tag{29}$$

where ν_0 is its frequency and t_0 the intrinsic time of the particle. Consider now an inertial system B, with a velocity v in their common x-axis relative to A.

[10] De Broglie, Louis; *"The wave nature of the electron"*; Nobel Lecture, (1929).

Then, by a Lonertz transformation, we find that

$$t_0 = \frac{t - \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{30}$$

We can then see that the phase observed by observer B shall be given by

$$\phi = \sin 2\pi \frac{\nu_0}{\sqrt{1 - \frac{v^2}{c^2}}}(t - \frac{vx}{c^2}). \tag{31}$$

Thus, we have obtained the frequency that the observer in inertial frame B will observe,

$$\nu = \frac{\nu_0}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{32}$$

and the observed propagation along the x-axis with the phase velocity

$$V = \frac{c^2}{v}. \tag{33}$$

De Broglie's goal was to link up this argumentation, concerning the frequency ν and phase velocity V, within a framework of energy and momenta. He started with imagining the particle in its rest frame A, where we have that $E = m_0c^2$ and $E = h\nu_0$, where m_0 is the rest mass of the particle. Equating these expressions, we find that

$$m_0 = \frac{h\nu_0}{c^2}. \tag{34}$$

He then imagined the momentum of the particle observed in B, as

$$p = \frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{35}$$

This was the final step before executing his *coup de grâce*,

$$p = \frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{\frac{h\nu_0}{c^2} v}{\frac{\nu_0}{\nu}} = h\nu\frac{v}{c^2} = \frac{h\nu}{V} = \frac{h}{\lambda}. \tag{36}$$

Thusly, *the De Broglie wavelength* has been obtained,

$$\lambda = \frac{h}{p}. \tag{37}$$

We note a couple of things with this simple derivation. The only assumption made is that of the phase of the particle. The rest of his argumentation is, and was at his time, unarguably waterproof. We also note how small the De Broglie wavelength ought to be for everyday, as well as quantum objects. As an example, this fundamental equation hints at why we humans do not diffract when running into a pole or interfere when choosing one of two doors, for those of you who have wondered. The key lies in our large momenta and the incredibly small value of Planck's constant. Physicists took interest in his findings, and it was soon proposed that one should be able to test his theory by observing electron diffraction.[11] We will delay the discussion of the famous double slit experiment, which in itself is an excellent demonstration of electron diffraction, see ch. 5.

2.5 Bohr's Atomic Model

Along with the discovery of the electron by J.J. Thompson and of radiation by Henri Becquerel, new experimental methods had been developed to pierce into the structure of the atom, a task once only considered solvable by philosophy. During the turn of the century it was believed that electrons were embedded into a hypothesised *cloud* of unknown positively charged material. This model is called Thompson's plum pudding model, and was an attempt to reconcile the new findings of electrons with the neutral charge of the atom. Soon however, Thompson's model fell. By firing alpha-particles towards a thin gold foil, Ernest Rutherford and his students were able to find that Thompson's model of the atom was unable to explain the erratic reflections observed. The alpha-particles sometimes passed right through the golden foil, while at other times it flew off in a completely different direction.[12] This was not how to expect a uniform positively charged cloud to interact with massive positively charged particles, and the electrons were far to light to deflect the alpha particles in this way. Thus, Rutherford theorised that there should exist a concentrated positive charge inside the atoms for the alpha particles to scatter against. He further proposed that the electrons existed in circular or elliptical orbits around this concentrated positive charge, thus keeping the atom overall neutral. Therefore, this model came to be called *The planetary model* of the atom.[13]

Rutherford and his colleagues all knew what fundamental problems the planetary model of the atom carried with it. Consider an electron in circular orbit around a spherically symmetric coulomb potential of the same charge as the electron,

$$\frac{mv^2}{r} = m\omega^2 r = \frac{e^2}{4\pi\epsilon_0 r^2}. \tag{38}$$

We can use this expression to derive Kepler's third law, as

$$\omega^2 r^3 = \frac{e^2}{4\pi\epsilon_0 m}. \tag{39}$$

---

[11] Gasiorowics, Stephen; *"Quantum Physics"*; john Wiley Sons, Inc., third edition (2003), p. 10

[12] Rutherford, Ernest; *"The Scattering of and Particles by Matter and the Structure of the Atom"*; Philosophical Magazine Series 6, vol. 21 (1911).

[13] Gasiorowics, Stephen; *"Quantum Physics"*; john Wiley Sons, Inc., third edition (2003), p. 15

The total energy of this system is given by

$$E = \frac{mv^2}{2} - \frac{e^2}{4\pi\epsilon_0 r} = \frac{1}{2}\frac{e^2}{4\pi\epsilon_0 r} - \frac{e^2}{4\pi\epsilon_0 r} = -\frac{1}{2}\frac{e^2}{4\pi\epsilon_0 r}. \tag{40}$$

Consider now the Larmor formula

$$P = \frac{dE}{dt} = \frac{2}{3}\frac{q^2}{c^3}\left|\frac{d^2x}{dt^2}\right|^2, \tag{41}$$

obtained from classical electrodynamics and known during Rutherford's experiments, which gives us the energy loss per unit time of an accelerating particle with charge distribution q. In our system, it is given by

$$\begin{aligned} P &= \frac{2}{3}\frac{e^2}{4\pi\epsilon_0}\frac{1}{c^3}\left(\frac{v^2}{r}\right)^2 \\ &= \frac{2}{3}\frac{e^2}{4\pi\epsilon_0}\frac{1}{c^3}(\omega^2 r)^2 \\ &= \frac{2}{3}\frac{e^2}{4\pi\epsilon_0 r^4}\frac{1}{c^3}(\omega^2 r^3)^2 \\ &= \frac{2}{3}\frac{e^2}{4\pi\epsilon_0 r^4}\frac{1}{c^3}\left(\frac{e^2}{4\pi\epsilon_0 m}\right)^2 \\ &= \frac{2}{3}\left(\frac{e^2}{4\pi\epsilon_0}\right)^3\frac{1}{m^2c^3r^4}. \end{aligned} \tag{42}$$

However, if we decide to calculate the energy loss per unit time from eq. (40), we obtain

$$\frac{dE}{dt} = \frac{1}{2}\frac{e^2}{4\pi\epsilon_0}\frac{1}{r^2}\frac{dr}{dt}. \tag{43}$$

We combine these two expressions as

$$3r^2\frac{dr}{dt} = \frac{4}{m^2c^3}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2. \tag{44}$$

In integrating both sides with respect to time, we are left with

$$r^3 = \frac{4}{m^2c^3}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2 t \tag{45}$$

leading us to the final expression,

$$t = \frac{m^2c^3}{4}\left(\frac{4\pi\epsilon_0}{e^2}\right)^2 r^3. \tag{46}$$

From the assumption of energy loss through the Larmor equation, which was well known in the early twentieth century, this expression gives us the time t for a charge in circular orbit of radius r around a spherically symmetric potential to collapse. Plugging in fitting numbers for an electron, with r being in the order of 10^{-10} meters, a back of the envelope calculation gives us a value of 10^{-10} seconds. This is of course disastrous for the planetary model of the atom, and a confusing theoretical result. Indeed, the planetary model did explain the scattering data of Rutherford's experiments, and it was an immediate consequence of it when considering classical electrodynamics. Classical electrodynamics predicted the atom to collapse extremely quickly, and it was clear that the planetary model could not be the final, correct model.

The classical model also predicted that the frequency of radiation emitted from a particle in periodic motion is equal to that motion. It was expected that a continuous flow of radiation with increasing frequency would be measured, before the signal dies at the moment of collapse. Experiments carried out by Anders Ångström and collegues on the spectra of hydrogen was summerized by Johann Balmer in 1885, as

$$\frac{1}{\lambda} = R\Big(\frac{1}{n_1^2} - \frac{1}{n_2^2}\Big) \tag{47}$$

where R is Rydbergs constant and n_1 and n_2 are integers. This formula, called *the Rydberg formula*, showed good agreement with Ångströms experiments, and stood in direct conflict with classical electrodynamics.

The conflict between theory and experiment hinted at new physics. In 1913, Niels Bohr formulated three postulates, built upon a quantum mechanical account, to try and swing away at the problems outlined above:

1) Electrons exist in discrete *stationary states* of discrete energy around a spherically symmetric potential. In changing an electron from one stationary state to another, a discrete amount of energy must be emitted or absorbed, equal to the difference in energy between the two states.

2) The frequency of the radiation emitted or absorbed during a transition between two states of energy E_1 and E_2, $E_1 > E_2$, is given by

$$h\nu = E_1 - E_2 \tag{48}$$

3) Electrons in a spherically symmetric potential have discrete angular momenta, given by

$$L_n = n\hbar \tag{49}$$

where $n \in \mathbb{N}$ is the occupied stationary state.

Through these postulates, Bohr had *quantized* the atom. Instead of orbits with a continuum of allowed energies and angular momenta, we now have a system of a set amount of discrete allowed values. Bohr had come up with the first two postulates in consultation with equation (47) and Planck's formula. Concerning his third postulate, the stationary states correspond to a set of allowed circular orbits in the planetary model of the atom. He related the kinetic energy of the electron by the frequency of motion, as

$$\frac{m_e v^2}{2} = \frac{1}{2} n h \nu \tag{50}$$

so that

$$m_e v r = n \frac{h}{2\pi} \tag{51}$$

or

$$L_n = n\hbar. \tag{52}$$

Now, going back to the system of an electron in a stationary state around a spherically symmetric potential, we should again have

$$\frac{Ze^2}{4\pi\epsilon_0 r^2} = \frac{m_e v^2}{r}. \tag{53}$$

where we have added Z corresponding to the strength of the potential. In nature, this relates to the number of protons in the nucleus, where we have assumed the nuclear mass to be infinite and compressed to a point. Let us consider this system in light of Bohr's postulates. Combining this with eq. (51), we obtain

$$v = \frac{2\pi Z e^2}{4\pi\epsilon_0} \frac{1}{hn} \tag{54}$$

which, when reinserted into eq. (51) gives

$$r = \frac{1}{4\pi^2} \left(\frac{Ze^2}{4\pi\epsilon_0} \right)^{-1} \frac{h^2 n^2}{m_e}. \tag{55}$$

Thus, the total energy of the system is given by

$$E = \frac{m_e v^2}{2} - \frac{e^2}{4\pi\epsilon_0 r} = -\frac{Z^2 e^4 m_e}{8h^2\epsilon_0^2} \frac{1}{n^2} \tag{56}$$

where we have simply inserted v and r and simplified. Using Planck's relationship,

$$E = h\nu = \frac{hc}{\lambda} \tag{57}$$

we see that

$$\frac{1}{\lambda} = -\frac{Z^2 e^4 m_e}{8h^3\epsilon_0^2 c} \frac{1}{n^2} \tag{58}$$

which by our second postulate leads directly to eq. (47), where

$$R = \frac{Z^2 e^4 m_e}{8h^3\epsilon_0^2 c}. \tag{59}$$

Thus, Bohr's postulates has lead us to Rydbergs formula, and a theoretical value for the Rudberg constant which showed good agreement with experiments. He had done so by assuming that electrons occupy *stationary states*, a break from orbits in classical electrodynamics.

With this powerful result in hand, it was time for Bohr to await experimental results showing a clear connection between electron transitions between his postulated stationary states and the wavelengths put forth in the Rydberg formula. He did not have to wait long, as experiments performed by James Franck and Gustav Hertz which were published just one year later in 1914 showed strong agreement with his prediction.[14][15]. In Franck's and Hertz' experiment, a cathode and an anode, both negatively charged, were placed in plane parallel relative to each other inside of a glass tube only containing mercury gas (see fig. 3). at a distance of $z = d$ from the cathode placed at $z = 0$, a positively charged grid was placed creating a uniform electric field $\vec{E} = \frac{U}{d}$ between cathode and grid. Between the grid and the anode, we call the oppositely directed electric field $\vec{E}_{retard}$. Now, as electrons were ejected from the source S in the schematic, they were accelerated towards the grid, and on their way, they collided with mercury atoms inside the tube. As they reached the grid, the electrons start moving through $\vec{E}_{retard}$, slowing them down. Franck and Hertz wanted to measure the current I_A created by electrons reaching the anode and find its relation to the amount of voltage applied between the cathode and the grid, U. Their result is just as astonishing as it is classic, and is shown below (fig. 4).

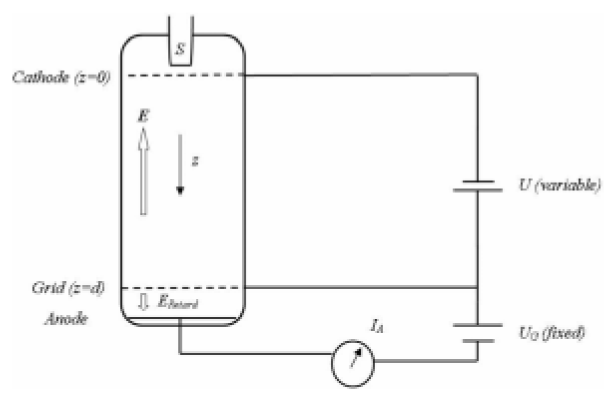

Figure 3: A Schematic of the Frank-Hertz experiment. (Picture Copied from: Robson, Robert E.; White Ronald D.; Hildebrandt, Malte; *"One hundred years of the Franck-Hertz experiment"*; Eur. Phys. J. D 68: 188, (2014).)

[14]Frank, James; Hertz, Gustav; *"Über Zusammenstöße zwischen Elektronen und Molekülen des Quecksilberdampfes und die Ionisierungsspannung desselben" (english: "On the collisions between electrons and molecules of mercury vapor and the ionization potential of the same")*; Verhandlungen der Deutschen Physikalischen Gesellschaft (1914); Translated version in: Boorse, A. Henry; Motz, Lloyd; *"The World of the Atom"*; Basic Books Inc, Publishers (1966).

[15]Robson, Robert E.; White, Ronald D.; Hildebrandt, Malte; *"One hundred years of the Franck-Hertz experiment"*; Eur. Phys. J. D 68: 188, (2014).

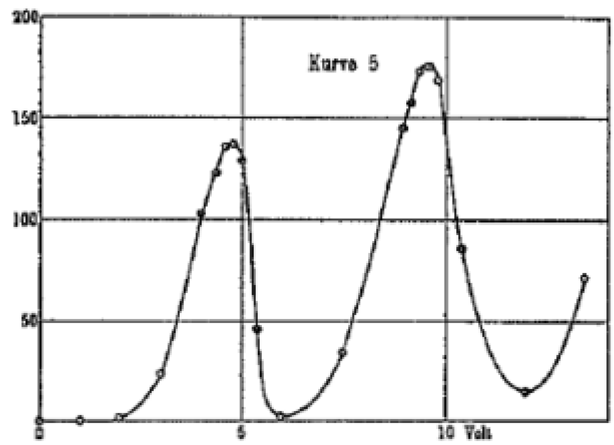

Figure 4: Results from the Frank-Hertz experiment, where the y-axis is the current I_A and the x-axis is the voltage U. (Picture Copied from: Robson, Robert E.; White Ronald D.; Hildebrandt, Malte; *"One hundred years of the Franck-Hertz experiment"*; Eur. Phys. J. D 68: 188, (2014).)

In classical electrodynamics, it would be expected that the relationship between I_A and U would be linear. This was not observed, instead Franck and Hertz obtained distinct peaks of current at voltage differences of about $\Delta U = 4.9$ V. First interpreted as an ionization potential for mercury atoms, subsequent experiments made them theorize that they had found electron excitation levels of energy $E = e\Delta U = 4.9$ eV. Furthermore, during their experiments it was noticed that excited atoms returning to their ground state emitted radiation of wavelength $\lambda = 253.6$ nm. Using Planck's radiation formula $E = \frac{hc}{\lambda}$, they noticed that

$$h = \frac{E\lambda}{c} = \frac{4.9 \text{ eV} \times \ 253.6 \text{ nm}}{c} = 6.59 \times 10^{-34} \text{ Js}, \tag{60}$$

extraordinarily close to the real value of $h = 6.62607015 \times 10^{-34}$ Js.

So what happened? As electrons accelerated towards the grid, inelastic and elastic collisions with gaseous mercury atoms took place, de-accelerating the electrons. In a saw toothed fashion, as electrons slowly accelerated in the electric field, they transferred this kinetic energy to instantaneously excite the mercury atoms. Of course, due to the excitation level spacing of $\Delta E = 4.9$ eV and the mean free path only allowing the electrons to be accelerated to a maximum kinetic energy, if this energy is sufficiently high, close to a multiple of 4.9 eV, the electrons are fast enough to peer through $\vec{E}_{retard}$, closing the circuit. If their energy recently has been transferred, the electrons wouldn't have been able to. Alas, Bohr had strong experimental evidence to support his stationary state model of the atom.

3 Canonical Quantization: Its History and Postulates

It is perhaps surprising that *there exist no proof of quantum theory*[16], meaning that there is no clear and coherent way to move from a classical to a quantum formalism purely using mathematical reasoning or the like. We must look to nature and use a fair amount of guess-work as to how a quantum theory of nature might look like. In fact, this *should* not be surprising at all. The history of physics moves in the "wrong" direction, with the correspondence principle demanding quantum theory to collapse into classical physics when quantum numbers are large, and not some other way around. Our intuition is in the realm of classical physics, and that of nature seems to be in line with quantum theory. Nonetheless, there are reasons for the educated guesses physicists made for quantizing theories in the past, of course. In this chapter, we will discuss the postulates of quantization and their origins to later discuss their implications.

Let us summarize what we found above through the early experiments of quantum mechanics. In essence, we have found that quantizing certain parts of classical theories shows promising results in explaining natural phenomena. It is clear that classical physics is unable to give an account of these processes. The story told, though hold together by a red thread, however lacks any formal structure[17]. As *Old Quantum Theory* developed, it was clear that the theories obtained were essentially different, sometimes contradictory fragments of an incomplete theory, demanding answers on such matters as the self-contradiction of wave-particle duality and the existence of quantized *and* unquantized dynamics existing simultaneously in one model. A first step came in the summer of 1925 when Werner Heisenberg realized that non-commuting operators might explain the scattered data of old quantum theory. Together with Max Born and Pascual Jordan, this method was further developed and ended in a joint paper[18] following two other influential papers, one written by Heisenberg[19] and one written by Born and Jordan[20]. Heisenberg's approach was not to solve the question of electron dynamics in a spherically symmetric potential, a sought after goal

[16] Nakahara, Mikio. *"Geometry, Topology and Physics"*, second edition, p. 9. Taylor and Francis group (2003).

[17] Von Neumann, John; *"Mathematical Foundations of Quantum Mechanics"*; Princeton University Press, (1955), p. 3-17.

[18] Born, Max; Heisenberg, Werner; Jordan, Pascual; *"Zur Quantenmechanik II"*; Zeitschrift für Physik 35 (1926). English translation: *"On Quantum Mechanics II"*; Translated version in: van der Waerden, B. L. (editor); *"Sources of Quantum Mechanics"*; Dover Publications, (1968).

[19] Heisenberg, Werner; *"Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen"*; Zeitschrift für Physik 33 (1925). English translation: *"Quantum-Theoretical Re-interpretation of Kinematic and Mechanical Relations"*; Translated version in: van der Waerden, B. L. (editor); *"Sources of Quantum Mechanics"*; Dover Publications, (1968).

[20] Born, Max; Jordan, Pascual; *"Zur Quantenmechanik"*; Zeitschrift für Physik 34 (1925). English translation: *"On Quantum Mechanics"*; Translated version in: van der Waerden, B. L. (editor); *"Sources of Quantum Mechanics"*; Dover Publications, (1968).

in old quantum theory, but only focus on observable quantities, in this case hydrogen spectral lines[21]. Guided by the correspondence principle, Heisenberg first recalled that[22]

$$X(t) = \sum_{n=-\infty}^{\infty} e^{\frac{2\pi i n t}{T}} X_n \tag{61}$$

is the Fourier description of a classical particle in circular motion while interacting weakly with an electromagnetic field, which also happens to be a description of its emitted frequency, where $X(t)$ is the particle position, and T its period, under the condition

$$X_n = X_{-n}^* \tag{62}$$

to keep $X(t)$ real at all times. For large orbits n and m, where $n - m$ is very small, then, we should expect that

$$f = \frac{E_n - E_m}{h} \approx \frac{n - m}{T} \tag{63}$$

where T is assumed to be the period of both orbits n and m and f is the frequency of emitted radiation. Of course, if n and m are small or if $n - m$ is large, this approximation would not be sufficient. As the frequencies from a quantum description is approximately the same as that of a classical description under these conditions, Heisenberg assumed that there should be something that could be described by oscillation in the quantum description. He introduced the new quantity X_{nm}, where for n and m large and $n - m$ small, X_{nm} is the coefficient X_{n-m}, and only this coefficient in equation (61). This way, he could develop equations of motion for one term in the Fourier description dependent on the difference of frequency, obtaining radiation of one frequency as is required from the experimental data on spectral lines. Of course, he demanded that this description is to hold below the classical limit. Realising that X_{n-m} has the opposite frequency of X_{m-n}, the new condition to keep these objects real in a Fourier description was

$$X_{nm} = X_{mn}^*. \tag{64}$$

Heisenberg further realised that two of these Fourier coefficients representing position X_{nm} one the one hand and momentum P_{nm} on the other could be added together as

$$(XP)_{nm} = \sum_{k=0}^{\infty} X_{nk} P_{km}, \tag{65}$$

where $(XP)_{nm}$ oscillates with the same frequency as X_{nm} and P_{nm}.

[21] Fedaka, William A. ; Prentis, Jeffrey J.; *"The 1925 Born and Jordan paper "On quantum mechanics""*;American Journal of Physics 77, 128 (2009).

[22] Mehra, Jagdish; Rechenberg, Helmut; *"The historical development of quantum theory: Vol. 3, The formulation of matrix mechanics and its modifications 1925-1926"*; Springer-Vlg, cop. (1982).

It is particularly interesting to note that Heisenberg, when developing this scheme, was not familiar with matrix theory. Born pointed out to Heisenberg that equation (65) is indeed the law of matrix multiplication and that Heisenberg's condition that $X_{nm} = X^*_{mn}$ is nothing more than the demand of only considering Hermitian matrices. Heisenberg and Born, together with Jordan, then developed the theory of *Matrix Mechanics*, stating that we should find a set of Hermitian matrices

$$X(t)_{nm} = e^{\frac{2\pi i f(nm)t}{T}} X_{nm}, \tag{66}$$

$$P(t)_{nm} = e^{\frac{2\pi i f(nm)t}{T}} P_{nm}, \tag{67}$$

where $X(t)_{nm}$ and $P(t)_{nm}$ can be interpreted as corresponding to one term in (61), where $f(nm)$ is the frequency difference $n-m$. In their system, diagonal elements represented stationary states, as $X(t)_{nn} = X_{nn}$, while $n \neq m$ represented transition amplitudes between states. Particularly, we note that

$$X(t)_{nm} X(t)^*_{mn} = \left|X(t)_{nm}\right|^2, \tag{68}$$

giving the probability of a transition.

The Heisenberg-Born-Jordan formulation gave some familiar results of quantum mechanics. For example, they introduced the Heisenberg equation of motion and defined the commutator between two Hermitian matrices. Concerning the commutator, Heisenberg, Born and Jordan set out to remove any artificial dependence on integers such as n and m in fundamental relations of quantum mechanics, as was the case in old quantum theory we saw above. The idea was to set up relations not explicitly dependent on n or m, and from them be able to find their relations. We shall see that the commutation relations they found,

$$[X_n, X_m] = [P_n, P_m] = 0, \tag{69}$$

$$[X_n, P_m] = i\frac{h}{2\pi}\delta_{nm}, \tag{70}$$

are very intimately connected with *Poisson brackets* of classical Hamiltonian mechanics, and so we should not elucidate their points any further, other than commenting that they first appeared in their papers.

It is worth to remember that when the Heisenberg-Born-Jordan papers were published, the physics community had hardly been exposed to matrices, let alone Hermitian matrices. This abstract formulation initially got the cold shoulder, as physicists were mostly still interested in finding a new quantum theory of wave mechanics, somewhat familiar to classical wave mechanics. This would come in 1926 by a series of papers published by *Erwin Schrödinger*[23]. Inspired by the results of De Broglie, Schrödinger wanted to find a three dimensional

[23] Galler, Anna; Canfield, Jeremy; Freericks, James K.; *"Schrödinger's original quantum-mechanical solution for hydrogen"*; arXiv:2007.14798, (2020).

wave equation explaining the dynamics of electrons, which would naturally exist if electrons were represented by waves. His results culminated in the famous equation that bears his name,

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = \hat{H}\psi(\vec{r},t), \tag{71}$$

whose time-independent form familiarly can be used to find stationary states,

$$\hat{H}\phi(\vec{r}) = E\phi(\vec{r}). \tag{72}$$

We will spend some time below developing this famous equation through first principles, but what is of importance here is to realise that two theories of quantum mechanics, from entirely different starting points, was proposed very close to each other in time, one based on matrix mechanics and one on wave mechanics, both capable of reproducing in a systematic way some famous results of quantum mechanics, for example the stationary states of the hydrogen atom. They even gave the same predictions in some aspects not present in old quantum theory. However, as was first shown by Schrödinger considering function spaces of discrete sequences of matrix theory and wave functions, and later by *Paul A. M. Dirac* and Jordan by introducing the *Dirac-delta distribution*, functions on the mathematical spaces of these two theories are equivalent[24], which indeed proves that they will always give equivalent results of measurements. Hence, it was possible to use the languages of both formulations in a joint fashion.

Parallel to the development of this formalism, Paul A. M. Dirac introduced the "Method for classical analogy" in his 1926 doctoral thesis, today more known as *The Method of Canonical Quantization*. The correspondence principle had been a valuable guiding light for developing matrix and quantum wave mechanics, but it was Dirac who noticed the close relationship between the Poisson bracket of canonical variables in Hamiltonian mechanics and the commutator relations presented above. Through his method, Dirac had developed the first scheme of quantization that in a standardized way could give the quantum counterpart of classical systems, giving physicists somewhere to start when building new quantum systems. The method of canonical quantization is very well detailed in his seminal work of 1930[25], a work that we will use heavily in the further.

Without further ado, let us introduce the postulates of canonical quantization, seeing them as the results of the preceding discussion. When considering an isolated system in classical Hamiltonian mechanics, canonical quantization gives us four rules for quantizing the given system:

Postulate 3.1: Any quantum system has a corresponding Hilbert space $\mathscr{H}$, where

[24]Von Neumann, John; *"Mathematical Foundations of Quantum Mechanics"*; Princeton University Press, (1955), p. 17-33.

[25]Dirac, Paul A. M.; *"The Principles of Quantum Mechanics"*; Oxford University Press, Third edition (1947) (first edition published (1930)).

the state of a quantum system is described by a state vector $|\psi\rangle \in \mathscr{H}$*,* $|\psi\rangle \sim c\,|\psi\rangle$*,* $c \in \mathbb{C} \setminus \{0\}$*.*[26]

Postulate 3.2: We replace a classical quantity A, represented by a function, with a Hermitian operator $\hat{A}$ *acting on* $|\psi\rangle \in \mathscr{H}$*. Measurements of the physical quantity A result in one of the eigenvalues* a_n *of* $\hat{A}$*, according to* $\hat{A}\,|\psi_n\rangle = a_n\,|\psi_n\rangle$ *.*

Postulate 3.3: The Poisson bracket in classical Hamiltonian mechanics should be replaced by the commutator,

$$i\hbar\{A, B\} = [\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \tag{73}$$

with the fundamental commutation relations

$$[\hat{q}_i, \hat{q}_j] = [\hat{p}_i, \hat{p}_j] = 0, \tag{74}$$

$$[\hat{q}_i, \hat{p}_j] = i\hbar\delta_{ij}. \tag{75}$$

Postulate 3.4: After many measurements of the physical quantity A of a set of equivalent states $|\psi\rangle \in \mathscr{H}$ *at time t, the expectation value is given by*

$$\langle A\rangle_t = \frac{\langle\psi|\,\hat{A}(t)\,|\psi\rangle}{\langle\psi|\psi\rangle} \tag{76}$$

Postulate 3.5: A state $|\psi\rangle = \sum_n c_n\,|\psi_n\rangle$ *and its displaced state* $|\psi d\rangle = \sum_n c_n\,|\psi_n d\rangle$ *are related as* $|\psi d\rangle = \hat{D}\,|\psi\rangle$*, where* $\hat{D}$ *is a unitary operator.*

Some comments on these postulates are in order. It is clear from postulate 3.1 that we define *quantum states* $|\psi\rangle \in \mathscr{H}$. Thus, the Hilbert space defines all the possible states a quantum system can possess, while $|\psi\rangle \in \mathscr{H}$ represents the state that the quantum system is in at a particular moment in time. A classical analogue to this system would be to regard $\mathscr{H}$ as the phase space, while $|\psi\rangle \in \mathscr{H}$ is a point in that phase space. Correspondingly, specifying a point in Hilbert space allows us to solve the equations of motion uniquely, i.e the Schrödinger equation.

Why are quantum states defined on a complex Hilbert space? A complex Hilbert space is, as may be familiar, a complex inner product space such that the norm

$$\big||x|\big| = \sqrt{\langle x|x\rangle} \tag{77}$$

turns the space into a complete metric space. Looking at our postulates, we note that they naturally do include complex inner product spaces, for example

[26]Nakahara, Mikio. *"Geometry, Topology and Physics"*, second edition, Taylor and Francis group (2003), p. 10.

in defining $\hat{A}\,|\psi\rangle = a_n\,|\psi\rangle$ or $\langle A\rangle_t$, which is of course absolutely necessary. Even more pressing is the fact that Hermitian operators always have orthogonal eigenvectors, an undefined concept without an inner product. The demand of completeness might be rather more tricky, but an elegant example could help us understand this demand. Consider the vector space of all polynomials with the naturally chosen orthonormal basis $(x^0, x^1, x^2, ...)$. We could easily check that in general a finite number of elements in this space has the properties of being a complete vector space. Now, consider the Taylor series

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x. \tag{78}$$

Evidently, any finite linear combination of polynomials is itself a polynomial, but this no longer holds true for infinite series of polynomials. This Taylor series is an example of a *Cauchy sequence*, and a complete metric space has the property that all Cauchy sequences with points from a space is itself an element of that space. Because e^x is not a polynomial, the vector space of all polynomials is *not* a complete metric space. Consider now the general solution to the time in-dependent Schrödinger equation,

$$|\psi\rangle = \sum_{n=1}^{\infty} c_n\,|\psi_n\rangle\,. \tag{79}$$

This is indeed also a Cauchy sequence. Given that these infinite series often appear in quantum mechanics, and the fact that linearity of the Schrödinger equation implies that any linear combination of solutions is also itself a solution, we need infinite series in quantum mechanics to be Cauchy sequences in order for our postulates to be properly defined. Hence, $|\psi\rangle \in \mathscr{H}$.

The equivalence relation in 3.1 should be understood as the assumption where superposing a state upon itself should result in the same state.[27] Suppose that we superpose a state upon itself, as

$$c_1\,|\psi\rangle + c_2\,|\psi\rangle = (c_1 + c_2)\,|\psi\rangle = c_3\,|\psi\rangle \tag{80}$$

where $c_1, c_2 \in \mathbb{C}$. As c_3 is any complex number, by our assumption it has to correspond to the original state $|\psi\rangle$, if $c_3 \neq 0$. If the condition $c_3 = 0$ is met, we can physically say that some kind of interference phenomenon has occurred, removing the state entirely. It is thus necessary to only consider the punctured complex plane.

One could represent a classical system within a Hilbert space. A classical state corresponds to systems where all Hermitian operators commute with each other and where the state is an eigenstate of all Hermitian operators. Put more clearly,

[27] Dirac, Paul A. M.; *"The Principles of Quantum Mechanics"*; Oxford University Press, Third edition (1947) (first edition published (1930)), p. 17.

we can *know everything* there is to know about a classical system. When moving from a system described by classical Hamiltonian mechanics to a system described by quantum mechanics, though, we need to make use of postulate 3.2. The reason for introducing this postulates goes back to Heisenberg's, Born's and Jordan's realization that in order to describe hydrogen spectral lines, a two-indexed mathematical object that is invariant under Hermitian conjugation is required: Hermitian operators. Their action on a particular quantum state within a Hilbert space also guarantees real eigenvalues, which we interpret as the results of a measurement. The question of how to find these Hermitian operators is in essence the project of quantum mechanics. To find a suitable Hermitian operator, we need to study quantum systems by making standardized measurements that are contextually suitable to measure a particular observable, obtain these observables and interpret them as eigenvalues of a particular operator and work ourselves backward to find an operator that predicts them.

We know that the Poisson bracket is essential in Hamiltonian mechanics as a means to derive the equations of motion. Now that we have changed from the formalism of continuous functions to Hermitian operators, we need a new Poisson bracket to integrate this into the new theory. In essence, this was Dirac's realization as stated above, which built on top of the Heisenberg-Born-Jordan formulation and gave us a clear scheme towards quantizing classical theories. As is well known, the Poisson bracket in Hamiltonian mechanics is defined as

$$\{A, B\} = \sum_k \left(\frac{\partial A}{\partial q_k} \frac{\partial B}{\partial p_k} - \frac{\partial A}{\partial p_k} \frac{\partial B}{\partial q_k} \right) \tag{81}$$

where $A(q,p)$ and $B(q,p)$ (and later $C(q,p)$ and $D(q,p)$) are functions defined on the phase space of a Hamiltonian. As can easily be shown, the Poisson bracket enjoys the following properties,[28]

$$\{A, B\} = -\{B, A\}, \tag{82}$$

$$\{A, c\} = 0, \quad c \in \mathbb{R}, \tag{83}$$

$$\{A + B, C\} = \{A, C\} + \{B, C\}, \tag{84}$$

$$\{A, B + C\} = \{A, B\} + \{A, C\}, \tag{85}$$

$$\{AB, C\} = \{A, C\}B + A\{B, C\}, \tag{86}$$

$$\{A, BC\} = \{A, B\}C + B\{A, C\}, \tag{87}$$

$$\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0. \tag{88}$$

When determining the relationship between the commutator and the Poisson bracket, which is given as equation (73) in postulate 3.3, Dirac realized that both follow these properties, with the extra condition that $\hat{A}\hat{B} \neq \hat{B}\hat{A}$ for the

[28]Dirac, Paul A. M.; *"The Principles of Quantum Mechanics"*; Oxford University Press, Third edition (1947) (first edition published (1930)), p. 84-89.

quantum mechanical relations. Assuming this condition on the Poisson bracket properties above, Dirac realized that

$$\begin{aligned} &\{AB, CD\} \\ &= \{A, CD\}B + A\{B, CD\} \\ &= \big(\{A, C\}D + C\{A, D\}\big)B + A\big(\{B, C\}D + C\{B, D\}\big) \\ &= \{A, C\}DB + C\{A, D\}B + A\{B, C\}D + AC\{B, D\} \end{aligned} \tag{89}$$

and

$$\begin{aligned} &\{AB, CD\} \\ &= \{AB, C\}D + C\{AB, D\} \\ &= \big(\{A, C\}B + A\{B, C\}\big)D + C\big(\{A, D\}B + A\{B, D\}\big) \\ &= \{A, C\}BD + A\{B, C\}D + C\{A, D\}B + CA\{B, D\}, \end{aligned} \tag{90}$$

where we have used property (86) and (87) in two different orders. Equating these results, we obtain

$$\begin{aligned} \{A, C\}DB + AC\{B, D\} &= \{A, C\}BD + CA\{B, D\}, \\ \{A, C\}(BD - DB) &= (AC - CA)\{B, D\}. \end{aligned} \tag{91}$$

The kicker is to realize that this expression holds for A and C independently of B and D. Thus, we must have that

$$BD - DB = i\hbar\{B, D\} \tag{92}$$

$$AC - CA = i\hbar\{A, C\}, \tag{93}$$

where $\hbar$ should be independent of the Hermitian operators, and needs to commute with $BD - DB$ and $AC - CA$. To ensure that the bracket remains real as in classical mechanics, we have inserted a factor i.[29] This requires $\hbar \in \mathbb{R}$. Seeing that the left hand side of this equation is indeed equal to the commutator postulated in the Heisenberg-Born-Jordan formulation if the functions on phase space are turned into Hermitian operators, we are led towards guessing the relationship

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} = i\hbar\{A, B\} \tag{94}$$

between the commutator and the Poisson bracket, where $\hbar$ is to be seen as a universal constant with the dimensions of action. Of course, this constant is $\hbar = \frac{h}{2\pi}$, where h is Planck's constant. This, in essence, is how Dirac was led to introduce equation (73).

Through this postulate of canonical quantization, we see that it is indeed similar

[29] Dirac, Paul A. M.; *"The Principles of Quantum Mechanics"*; Oxford University Press, Third edition (1947) (first edition published (1930)), p. 28.

to the Poisson bracket of classical mechanics. What we can do at this stage is to guess what the quantum Poisson bracket might look like for quantum systems, based on analogies of classical mechanics. With the classical fundamental commutation relations

$$[p_i, p_j] = [q_i, q_j] = 0 \tag{95}$$

$$[q_i, p_j] = \delta_{ij} \tag{96}$$

at hand, where q_i and p_i are the canonical coordinates and momenta in phase space, respectively, we now make the educated guess that

$$[\hat{p}_i, \hat{p}_j] = [\hat{q}_i, \hat{q}_j] = 0 \tag{97}$$

$$[\hat{q}_i, \hat{p}_j] = i\hbar\delta_{ij}, \tag{98}$$

for our fundamental commutation relations of our bracket in quantum mechanics. In this case, it seems to correspond with nature, and we have successfully found one such relation. Thus, we have found a procedure to find quantum phenomena which has a classical analogue and are described in terms of canonical coordinates and momenta. Famously, it should be stated that the quantum commutation relations transforms into the classical Poisson bracket if we let $\hbar \longrightarrow 0$. This is the second way to formulate the correspondence principle, giving the same results as considering systems with large quantum numbers. It should however be noted that this does not work for *every* quantum system, as not all quantum systems has a classical analogue. In these cases, for example concerning spin states, canonical coordinates and momenta doesn't exist, even though we can still give meaning to the commutation relations above.

Postulate 3.4 needs a shorter introduction. This is because it relates to how to interpret experimental results given the formalism that we have already established. For discussing postulate 3.4, let us assume that $\left\| |\psi\rangle \right\|^2 = \langle\psi|\psi\rangle = 1$ for simplicity. Let us start with a physical quantity A which has a set of discrete eigenvalues, as

$$\hat{A}(t)\,|n\rangle = a_n\,|n\rangle\,, \qquad \langle n|n\rangle = 1. \tag{99}$$

Considering an arbitrary state

$$|\psi\rangle = \sum_n \psi_n\,|n\rangle\,, \qquad \psi_n = \langle n|\psi\rangle \tag{100}$$

we find that

$$\langle A\rangle_t = \langle\psi|\,\hat{A}(t)\,|\psi\rangle = \sum_{n,m} \psi_m^*\psi_n\,\langle m|\,\hat{A}(t)\,|n\rangle = \sum_n a_n|\psi_n|^2. \tag{101}$$

The $|\psi_n|^2$ factor is called the *probability*. This is realized by the fact that a measurement of A in the state $|n\rangle$ always yields a_n, and thus the probability that the outcome of a measurement is indeed a_n, or the probability of the state $|\psi\rangle$ being in $|n\rangle$, is given by

$$|\psi_n|^2 = \left|\langle n|\psi\rangle\right|^2. \tag{102}$$

This argument is similar for physical quantities on a continuous spectrum. Then, we have that

$$|\psi\rangle = \int da\psi(a)\,|a\rangle\,, \tag{103}$$

where the completeness relation is given by

$$\int da\,|a\rangle\,\langle a| = I.^{30} \tag{104}$$

From this, it is easy to see that

$$\int da'\,|a'\rangle\,\langle a'|a\rangle = \int da'\,\langle a'|a\rangle\,|a'\rangle = |a\rangle \tag{105}$$

must be normalized as $\langle a'|a\rangle = \delta(a'-a)$, where we have made use of the Dirac-delta function. It is then clear that we should have

$$\langle\psi|\psi\rangle = \int dada'\psi^*(a)\psi(a')\,\langle a|a'\rangle = \int da\big|\psi(a)\big|^2. \tag{106}$$

It then directly follows that

$$\langle\psi|\,\hat{A}\,|\psi\rangle = \int da\;a\big|\psi(a)\big|^2. \tag{107}$$

By a similar argument as above, the probability of measuring the physical quantity A with the outcome of $(a, a+da)$ is $\big|\psi(a)\big|^2 da$. Then, the probability density is given by

$$\rho(a) = \left|\langle a|\psi\rangle\right|^2. \tag{108}$$

Also, it is clear from our assumption of $\langle\psi|\psi\rangle = 1$ that

$$1 = \langle\psi|\psi\rangle = \sum_{n,m}\psi_m^*\psi_n\,\langle m|n\rangle = \sum_n|\psi_n|^2 = \int da\big|\psi(a)\big|^2 \tag{109}$$

directly follows. That is, the sum of all probability amplitudes for a given system is equal to 1.

Some technical comments on these postulates are in order. Firstly, the reader might be used to seeing the *equations of motion* as a postulate in some texts. We will see shortly that this is not necessary, as we can develop them from first principles using the above postulates, particularly postulate 3.5. As this development is central for our discussion, and also quite long, we have devoted

[30] Remember that $\langle a|\psi\rangle = \int da\psi(a)\,\langle a|a\rangle = \psi(a)$. Then, $|\psi\rangle = \int da\,\langle a|\psi\rangle\,|a\rangle = \int da\,|a\rangle\,\langle a|\psi\rangle \implies \int da\,|a\rangle\,\langle a| = I$.

an entire chapter to developing them. Secondly, it is often taken as a postulate that quantum states may be expanded as

$$|\psi\rangle = \sum_n \psi_n |n\rangle, \qquad \psi_n = \langle n|\psi\rangle, \tag{110}$$

often refereed to as the *expansion postulate*.[31] This is redundant as we have already postulated that state vectors exists in Hilbert space and that operators acting on state vectors are Hermitian. Of course, we are allowed to write any state vector as a linear combination of any set of orthogonal basis vectors in its Hilbert space, and the eigenstates of linear Hermitian operators always form such a basis. The "expansion postulate" is only a mathematical fact about linear Hermitian operators, already included by the postulate stating that we should use them representing physical quantities.

4 The Schödinger and Heisenberg equations of motion

Missing from the discussion now under our belt is the question of *dynamics*. It has been clear, for example from postulate 3.4, that quantum states or operators evolve over time, but the question of how has not been elucidated. In this section, we will enlight the nature of displacement operators, as well as develop the dynamical equations of quantum mechanics.

4.1 Displacement Operators

Let us create a scheme for the displacement of quantum states or Hermitian operators[32]. This could for example be a displacement δx in space from the original position x to $x + \delta x$, but as of now we will treat displacement operators in general. We shall take it as given that the superposition between states remain invariant under displacement if the system remains undisturbed, as is stated in postulate 3.5. Otherwise, the displaced state would not keep its essential form corresponding to the un-displaced state. That is, if we have the un-displaced state

$$|\psi\rangle = \sum_n c_n |\psi_n\rangle, \tag{111}$$

we demand that the displaced state is obtained in a way such that states correspond with the same linear equation between them, as

$$|\psi d\rangle = \sum_n c_n |\psi_n d\rangle. \tag{112}$$

[31] Jaffe, R. L.; *"SUPPLEMENTARY NOTES ON DIRAC NOTATION, QUANTUM STATES, ETC. "*; Physics 8.05, (2007).

[32] Dirac, Paul A. M.; *"The Principles of Quantum Mechanics"*; Oxford University Press, Third edition (1947) (first edition published (1930)), p. 99-103.

Thus, whenever the first equation holds, so does the other. Therefore, we have created a relationship between the states such that the displaced states are linear functions of the un-displaced states, leading us to deduce the existence of a linear operator of the form

$$|\psi d\rangle = \hat{D}\,|\psi\rangle\,, \tag{113}$$

depending only on the displacement performed. Let us investigate this linear operator further under our assumptions. Introducing a state $|\phi\rangle$ with the same constraints as above, it is clear that

$$\langle\phi d|\psi d\rangle = k = \langle\phi|\psi\rangle\,,\ k\in\mathbb{R} \tag{114}$$

where we have used the invariance under displacement of states. It is therefore true that

$$\langle\phi d|\psi d\rangle = \langle\phi|\,\hat{D}^\dagger\hat{D}\,|\psi\rangle = \langle\phi|\psi\rangle \tag{115}$$

or

$$\hat{D}^\dagger\hat{D} = 1. \tag{116}$$

As a second realisation, we know that for any Hermitian operator $\hat{A}$, we have that

$$\hat{A}\,|\psi\rangle = a_n\,|\psi\rangle\,, \tag{117}$$

which is given by postulate 3.2. Then, for the corresponding displaced state, $|\psi d\rangle$, we must have that

$$\hat{A}_d\,|\psi d\rangle = a_n\,|\psi d\rangle \tag{118}$$

where $\hat{A}_d$ is the displaced operator. Thus, we should have that

$$\hat{A}_d\,|\psi d\rangle = a_n\hat{D}\,|\psi\rangle = \hat{D}a_n\,|\psi\rangle = \hat{D}\hat{A}\,|\psi\rangle = \hat{D}\hat{A}\hat{D}^{-1}\hat{D}\,|\psi\rangle = \hat{D}\hat{A}\hat{D}^{-1}\,|\psi d\rangle \tag{119}$$

from which we can deduce that

$$\hat{A}_d = \hat{D}\hat{A}\hat{D}^{-1} \tag{120}$$

by realising that this relationship is true for any quantum state $|\psi d\rangle$.

Let us return to the infinitesimal displacements mentioned above. That is, we wish to show the behaviour of a quantum state $|\psi d\rangle$ as $\delta x\to 0$, under our position displacement example above. Under the assumption of physical continuity of quantum states under its dynamics, we should then expect $|\psi d\rangle$ to tend to $|\psi\rangle$, allowing us to expect that the limit

$$\lim_{\delta x\to 0}\frac{|\psi d\rangle - |\psi\rangle}{\delta x} = \lim_{\delta x\to 0}\frac{\hat{D}-1}{\delta x}\,|\psi\rangle \tag{121}$$

exists. Thusly, through this limit, we have defined a new linear operator for an infinitesimal position displacement of quantum states. Let us define it as *the displacement operator*,

$$\left(\frac{d}{dx}\right) = \hat{d}_x \equiv \lim_{\delta x \to 0} \frac{\hat{D} - 1}{\delta x}. \tag{122}$$

We take notice of the fact that the application of a phase factor $e^{i\theta}$, $\theta \in \mathbb{R}$ on $\hat{D}$ does not change the expectation value of any dynamical operator, as can be seen from equation (116) and postulate 3.4. Of course, this is the essence of *normalization*, and we wish for this phase factor to tend to unity as $\delta x \to 0$. We use this fact and see that

$$\lim_{\delta x \to 0} \frac{\hat{D} e^{i\theta} - 1}{\delta x} = \lim_{\delta x \to 0} \frac{\hat{D} - 1 + i\theta}{\delta x} = \hat{d}_x + i \lim_{\delta x \to 0} \frac{\theta}{\delta x} \tag{123}$$

where the remaining limit is a real number. Thusly, we realise that $\hat{d}_x$ can be added by an arbitrary imaginary number by this normalization condition.

Using this new linear operator, it is true that for small δx,

$$\hat{D} = 1 + \delta x d_x. \tag{124}$$

Using the fact that $\hat{D}^\dagger \hat{D} = 1$, we see that

$$\left(1 + \delta x \hat{d}_x^\dagger\right)\left(1 + \delta x \hat{d}_x\right) = 1 + \delta x (\hat{d}_x^\dagger + \hat{d}_x) = 1 \tag{125}$$

Where we have discarded δx^2, because of its insignificance in size. Thus, we see that

$$\delta x (\hat{d}_x^\dagger + \hat{d}_x) = 0. \tag{126}$$

This statement is only true for arbitrary δx for anti-Hermitian operators, and therefore we can conclude that $\hat{d}_x$ is an anti-Hermitian operator. Now, using the fact that $\hat{A}_d = \hat{D}\hat{A}\hat{D}^{-1}$, we obtain

$$\hat{A}_d = (1 + \delta x \hat{d}_x)\hat{A}(1 - \delta x \hat{d}_x) = \hat{A} + \delta x (\hat{d}_x \hat{A} - \hat{A}\hat{d}_x) \tag{127}$$

where we have again discarded δx^2. Thusly, under the same argumentation of infinitesimal displacement of quantum states, we can show that for displaced operators,

$$\lim_{\delta x \to 0} \frac{\hat{A}_d - \hat{A}}{\delta x} = \lim_{\delta x \to 0} \frac{\hat{A} + \delta x (\hat{d}_x \hat{A} - \hat{A}\hat{d}_x) - \hat{A}}{\delta x} = \hat{d}_x \hat{A} - \hat{A}\hat{d}_x, \tag{128}$$

Which is the generalized quantum condition for displacement operators. We can use these results to describe a system given by the Cartesian spatial coordinates, (x, y, z) and their respective canonical momenta conjugates (p_x, p_y, p_z). Suppose we introduce the position operators $\hat{x}$ and $\hat{x}_d = \hat{x} + \delta x$, corresponding to the original and displaced states of a quantum system. Then, using equation (126), we have that

$$\hat{x}_d = \hat{x} + \delta x = \hat{x} + \delta x (\hat{d}_x \hat{x} - \hat{x}\hat{d}_x), \tag{129}$$

leading us to

$$\hat{d}_x\hat{x} - \hat{x}\hat{d}_x = 1, \tag{130}$$

as the specific quantum condition relationship between the distance displacement operator and the position operator. Comparing this result with the fundamental commutation relations in postulate 3.3, we see that if we write this equation as

$$\hat{x}(-i\hbar\hat{d}_x) - (-i\hbar\hat{d}_x)\hat{x} = i\hbar, \tag{131}$$

$-i\hbar\hat{d}_x$ satisfies the same commutation relations as $\hat{p}_x$ in $[\hat{q}_x, \hat{p}_x] = i\hbar$. It follows that all other position and momentum operators commute with the distance displacement operator under the fundamental commutation relations. Considering their difference, $\hat{p}_x - (-i\hbar\hat{d}_x) = \hat{p}_x + i\hbar\hat{d}_x$, it is clear that it commutes with all operators introduced. For example,

$$[\hat{p}_x + i\hbar\hat{d}_x, \hat{x}] = [\hat{p}_x, \hat{x}] + i\hbar[\hat{d}_x, \hat{x}] = -i\hbar + i\hbar = 0 \tag{132}$$

Where all else is more trivial. It follows that $\hat{p}_x + i\hbar\hat{d}_x$ is a number, cf. the discussion above. $\hat{p}_x$ is necessarily real by assumption, and the factor of i in $-i\hbar\hat{d}_x$ ensures that the anti-Hermitian operator $\hat{d}_x$ turns real. Hence, $\hat{p}_x + i\hbar\hat{d}_x$ is a real number, say $\hat{p}_x + i\hbar\hat{d}_x = k$. Before, we found out that $\hat{d}_x$ can be added by an arbitrary imaginary number. Thusly, through $i\hbar(\hat{d}_x + i\alpha_x) = i\hbar\hat{d}_x - \hbar\alpha_x$, where α_x is the limit $\lim_{\delta x \to 0} \frac{\theta}{\delta x}$, we can assume that $-\hbar\alpha_x = k$, to obtain

$$\hat{p}_x + i\hbar\hat{d}_x = 0 \tag{133}$$

leading us to the fundamental relation

$$\hat{p}_x = -i\hbar\hat{d}_x. \tag{134}$$

Which, of course, has similar relationships as

$$\hat{p}_y = -i\hbar\hat{d}_y \tag{135}$$

$$\hat{p}_z = -i\hbar\hat{d}_z \tag{136}$$

from arguments that transfer trivially from above. This is *the momentum operator in position representation*, which are used diligently in quantum mechanics. We can now write the momentum operator as

$$\hat{p}_x |\psi\rangle = -i\hbar\frac{\partial}{\partial x} |\psi\rangle \tag{137}$$

where we take the partial derivative because of possible time dependence of the quantum state. This result, that you have undoubtedly seen before, comes from the arguments above concerning displacement operators. With this result we end this discussion, and it is time to develop additional familiar relationships.

4.2 The Dynamics of States

The example of displacement above was that of displacement in position. However, an example just as valid is that of displacement in time[33]. Considering

[33]Dirac, Paul A. M.; *"The Principles of Quantum Mechanics"*; Oxford University Press, Third edition (1947) (first edition published (1930)), p. 108-111.

the assumption of invariance of superposition of states under displacement as above, as stated in postulate 3.5, we now specify the invariant equations

$$\left|\psi(t_0)\right\rangle = \sum_n c_n \left|\psi_n(t_0)\right\rangle, \tag{138}$$

$$\left|\psi(t)\right\rangle = \sum_n c_n \left|\psi_n(t)\right\rangle, \tag{139}$$

where the quantum state is displaced from time t_0 to time t. It is then clear that we can define a linear operator

$$\left|\psi(t)\right\rangle = \hat{T}(t, t_0) \left|\psi(t_0)\right\rangle \tag{140}$$

where we now have specified the dependence of displacement. Under the same argumentation as above, we can deduce its unitarity and its affect on displacement of operators,

$$\hat{T}^\dagger \hat{T} = 1 \tag{141}$$

$$\hat{A}_t = \hat{T}\hat{A}\hat{T}^{-1}. \tag{142}$$

In introducing infinitesimal displacements in time, instead of the above example of $\delta x \to 0$ we introduce $t \to t_0$ and enlight the existence of

$$\lim_{t\to t_0} \frac{\left|\psi(t)\right\rangle - \left|\psi(t_0)\right\rangle}{t - t_0} = \frac{d}{dt_0} \left|\psi(t_0)\right\rangle, \tag{143}$$

which is just the derivative of $|\psi t_0\rangle$ with respect to t_0. As an operator, it is given by

$$\frac{d}{dt_0} \left|\psi(t_0)\right\rangle = \lim_{t\to t_0} \frac{\hat{T} - 1}{t - t_0} \left|\psi(t_0)\right\rangle. \tag{144}$$

By the same arguments as for the position displacement operator, this new operator is anti-Hermitian and invariant under addition of an arbitrary imaginary number. Multiplying the above equation with $i\hbar$, and setting

$$i\hbar \lim_{t\to t_0} \frac{\hat{T} - 1}{t - t_0} = \hat{H}(t_0) \tag{145}$$

where $\hat{H}(t_0)$ is the *Hamiltonian*, we obtain the new expression

$$i\hbar \frac{d}{dt_0} \left|\psi(t_0)\right\rangle = \hat{H}(t_0) \left|\psi(t_0)\right\rangle \tag{146}$$

or, for a general time t,

$$i\hbar \frac{d}{dt} \left|\psi(t)\right\rangle = \hat{H}(t) \left|\psi(t)\right\rangle. \tag{147}$$

This is *the Schrödinger equation of motion*, or just the *Schrödinger equation.* Thus, we have an equation describing the dynamics of undisturbed states between two time intervals. We will keep building upon the Schrödinger equation, particularly on the question of why the Hamiltonian is introduced in this way, but first we need to discuss *the Heisenberg equation of motion.*

4.3 The Dynamics of Operators

As we have found the equation of which quantum states propagates through time, it is time for a very simple observation proceeding our continued discussion[34]. We have already seen that operators such as $\hat{D}$ and $\hat{T}$ can displace a quantum state, for example as

$$\left|\psi(t_0)\right\rangle \to \hat{T}(t_0,t)\left|\psi(t_0)\right\rangle = \left|\psi(t)\right\rangle. \tag{148}$$

However, we saw that for these linear operators, it is true that

$$\left\langle\phi(t)\middle|\psi(t)\right\rangle = \langle\phi|\,\hat{T}^{\dagger}\hat{T}\,|\psi\rangle = \left\langle\phi(t_0)\middle|\psi(t_0)\right\rangle, \tag{149}$$

leaving the inner product of quantum states unchanged under displacement operators. Using this fact, we infer how a system $\left\langle\phi(t_0)\right|\hat{A}\left|\psi(t_0)\right\rangle$ must change as

$$\left\langle\phi(t_0)\right|\hat{A}\left|\psi(t_0)\right\rangle \to \left\langle\phi(t_0)\right|\hat{T}^{\dagger}\hat{A}\hat{T}\left|\psi(t_0)\right\rangle = \left\langle\phi(t_0)\right|\hat{A}_t\left|\psi(t_0)\right\rangle \tag{150}$$

The *simple* observation, however, is the physical fact we obtain from the associative law of multiplication, as

$$(\left\langle\phi(t_0)\right|\hat{T}^{\dagger})\hat{A}(\hat{T}\left|\psi(t_0)\right\rangle) = \left\langle\phi(t_0)\right|(\hat{T}^{\dagger}\hat{A}\hat{T})\left|\psi(t_0)\right\rangle = \left\langle\phi(t_0)\right|\hat{A}_t\left|\psi(t_0)\right\rangle. \tag{151}$$

The key insight is this: either we treat the displacement operators as acting on our quantum states, leaving the Hermitian operator $\hat{A}$ independent of time as is shown on the left hand side of this equation, or we let the displacement operators act on our Hermitian operator $\hat{A}$, leaving the quantum states independent of time. This is of course the *Schrödinger picture* and the *Heisenberg picture*, respectively. The Schrödinger equation fixes the dynamics of states under the Schrödinger picture, and we now wish to find the dynamical equation of operators under the Heisenberg picture.

Firstly, some comments on the Schrödinger equation are in order. We have that

$$i\hbar\frac{d}{dt}\left|\psi(t)\right\rangle = \hat{H}(t)\left|\psi(t)\right\rangle \tag{152}$$

for quantum states of general time parameter t. However, as

$$\left|\psi(t)\right\rangle = \hat{T}\left|\psi(t_0)\right\rangle \tag{153}$$

we can write the Schrödinger equation as

$$i\hbar\frac{d}{dt}\hat{T}\left|\psi(t_0)\right\rangle = \hat{H}(t)\hat{T}\left|\psi(t_0)\right\rangle. \tag{154}$$

This is true for quantum states in general, and thus we have that

[34] Dirac, Paul A. M.; *"The Principles of Quantum Mechanics"*; Oxford University Press, Third edition (1947) (first edition published (1930)), p. 111-118.

$$i\hbar\frac{d\hat{T}}{dt} = \hat{H}(t)\hat{T}. \tag{155}$$

Now, consider Hermitian operators in the Heisenberg picture which changes as

$$\hat{A}_t = \hat{T}^\dagger \hat{A}\hat{T}, \tag{156}$$

or

$$\hat{T}\hat{A}_t = \hat{A}\hat{T}. \tag{157}$$

Differentiating with respect to t gives us

$$\frac{d\hat{T}}{dt}\hat{A}_t + \hat{T}\frac{d\hat{A}_t}{dt} = \hat{A}\frac{d\hat{T}}{dt}. \tag{158}$$

Using equation (155), we can rewrite this as

$$\hat{H}(t)\hat{T}\hat{A}_t + i\hbar\hat{T}\frac{d\hat{A}_t}{dt} = \hat{A}\hat{H}(t)\hat{T}, \tag{159}$$

or equivalently

$$\begin{aligned} & i\hbar\frac{d\hat{A}_t}{dt} \\ &= \hat{T}^{-1}\hat{A}\hat{H}(t)\hat{T} - \hat{T}^{-1}\hat{H}(t)\hat{T}\hat{A}_t \\ &= \hat{T}^{-1}\hat{A}\hat{T}\hat{T}^{-1}\hat{H}(t)\hat{T} - \hat{T}^{-1}\hat{H}(t)\hat{T}\hat{A}_t \\ &= \hat{A}_t\hat{H}(t)_t - \hat{H}(t)_t\hat{A}_t. \end{aligned} \tag{160}$$

Using postulate 3.3, we can directly deduce that

$$\frac{d\hat{A}_t}{dt} = [\hat{A}_t, \hat{H}(t)_t], \tag{161}$$

and we have obtained our desired result. This is *the Heisenberg equation of motion.*

Let us compare the Heisenberg equation of motion thus obtained by the Hamilton equations of motion,

$$\frac{dA}{dt} = [A, H] \tag{162}$$

where $A(q, p)$ is any physical quantity defined on the phase space of the canonical coordinates, H is the classical Hamiltonian and the bracket is the Poisson bracket. It is of course strikingly similar to the Heisenberg equation of motion. From classical analogy, we then assume that the Hamiltonian found in quantum theory should define the total energy of the system, just as it does in classical mechanics. This assumption is further strengthened by the fact that

$$\hat{p}_x = -i\hbar \lim_{\delta x\to 0}\frac{\hat{D}-1}{\delta x} = -i\hbar d_x \tag{163}$$

and

$$\hat{H}(t_0) = i\hbar \lim_{t\to t_0} \frac{\hat{T}-1}{t-t_0} = i\hbar d_t, \tag{164}$$

due to *Noether's theorem*[35]. It states that every *continuous symmetry* in a closed system corresponds to a conservation law of nature. For example, if a system inherits translation invariance in space we find that linear momentum is conserved in that system, and if a system inherits translation invariance in time we find that energy is conserved in that system. Indeed, our linear translation operators above enjoys such continuous symmetries in space and time, a fact we will not prove but state nonetheless. Hence, it is reasonable to assume that $\hat{H}(t)$ is given by the total energy of the system, in harmony with Noether's theorem.

4.4 Discussion of the Equations of Motion

It is now time to finalize the discussion on the Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}\left|\psi(t)\right\rangle = \hat{H}(t)\left|\psi(t)\right\rangle. \tag{165}$$

Using postulate 3.4, we introduce a representation which we assume without loss of generality to be a complete set of commuting operators α acting on our quantum state $\left|\psi(t)\right\rangle$, giving us

$$\left\langle\alpha\middle|\psi(t)\right\rangle \equiv \psi_\alpha(t). \tag{166}$$

Thus, making this bra act on the Schrödinger equation, we obtain

$$i\hbar\frac{\partial}{\partial t}\psi_\alpha(t) = \hat{H}(t)\psi_\alpha(t). \tag{167}$$

This is the *Schrödinger wave equation*, which you have seen during your first courses on quantum mechanics. As is commonly done, we write the quantum state in position representation, as

$$\left\langle x\middle|\psi(t)\right\rangle \equiv \psi_x(t) \equiv \psi(x,t). \tag{168}$$

Using our discussion of the Hamiltonian as the total energy of a system, we again see by classical analogy that

$$H(t) = \frac{p^2}{2m} + V(t). \tag{169}$$

It is our goal to translate this classical Hamiltonian to that of a quantum framework. We know from ch. 4.1 that the momentum operator in position representation is given as

$$\hat{p} = -i\hbar\frac{\partial}{\partial x}. \tag{170}$$

[35] Rubens, de Melo Marinho Jr; *"Noether's theorem in classical mechanics revisited"*; arXiv:physics/0608264, (2006).

Inserting this into our Hamiltonian, we obtain

$$\hat{H}(t) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(t) \tag{171}$$

which we insert into the Schrödinger wave equation as

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x,t) + V(t)\psi(x,t). \tag{172}$$

This equation gives us the dynamics of the wave function on the x-axis, which translates trivially to other axes and combinations of them. We will use this equation in analysing one of the most famous experimental results of quantum mechanics.

5 To Marvel at Progress: A self Contained Guide to the Double Slit Experiment

As any enthusiast of physics can tell you, the double slit experiment condensates some of the more peculiar aspects of the nature of quanta. Nevertheless, it is often presented as a thought experiment or in a manner not up to date with current experiments. To combat this problem, let us represent the experimental work of *A. Tonomura et al.*,[36] who in 1989 performed a realization of this famous thought experiment.

The experimental setup looks something like this. Electrons from a source pass through a set of electrostatic lenses in order to create an electron beam, see fig. (5). Directly in front of the beam is a positively charged wire placed perpendicularly, and on either side of the wire are two grounded plates. At the end of the beam path is a detector screen, producing a cascade of photons once an electron is detected. This way, we can see that the classic thought experiment has been recreated. The positively charged wire together with the grounded plates creates two potential "tunnels", or slits, for an electron to go through, and the task at hand is to fire electrons, one by one, towards the detector. Fig. (6) shows the results from five experiments, each with a difference in the total number of electrons fired. The first experiment gives us a seemingly random pattern. In the experiments with the longest time exposure, however, an *interference pattern* emerges.

[36]Tonomura, A.; Endo, J.; Matsuda, T.; et al.; *"Demonstration of single-electron buildup of an interference pattern"*; American Journal of Physics 57, 117 (1989).

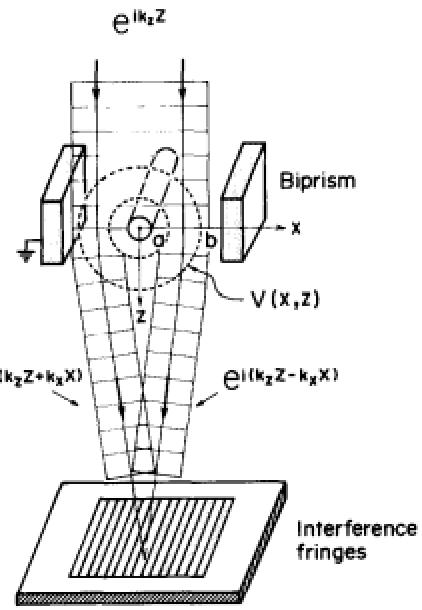

Figure 5: A. Tonomura et al. experimental setup. (Picture copied from: Tonomura, A.; Endo, J.; Matsuda, T.; et al.; *"Demonstration of single-electron buildup of an interference pattern"*; American Journal of Physics 57, 117 (1989).)

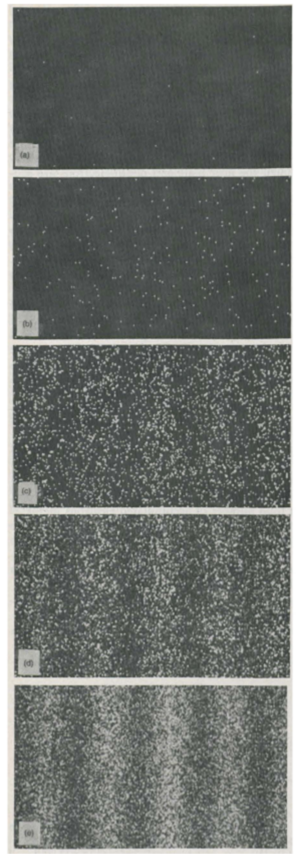

Figure 6: Results from A. Tonomura et al. experiments showing electron interference buildup. (Picture copied from: Tonomura, A.; Endo, J.; Matsuda, T.; et al.; *"Demonstration of single-electron buildup of an interference pattern"*; American Journal of Physics 57, 117 (1989).)

This result can be thought as a strange one. The electron is a particle, and how could it produce an interference pattern, characteristic of waves? Famously, we can try to obtain more information about how the electrons behave by trying to measure its path and answer the question, "which slit did the electrons pass through?". We make a potential blockade in the left slit such that electrons only could pass through the right slit. A. Tonomura et al. did not perform this experiment, but there exist no disagreement among physicists as to what would happen. We would see that the interference pattern would have completely disappeared and been replaced by one strong band, directly in front the right slit. This new experiment could be performed with the same number of electrons, the same number of times, with the same amount of detections. The implication seems clear: during the experiment of A. Tonomura et al., the electrons could not have been passing through the right slit at all, as blocking the left slit completely changed the outcome of the experiment. They must all have been passing through the left potential slit. Secure in our conviction, we now make a potential blockade in the right slit to only let electrons pass the left slit. Shockingly, we see how we now have created a strong band directly above the left slit. When performing our experiment, it seems as though the electron is affected by whether there exists one or two slits to pass through, regardless of our convictions of which of these slits it passes through. We can try and claim that the electrons somehow affected each other on the way to the detector in the experiment of A. Tonomura et. al., but after more experiments we see that it makes no difference at all how far the electrons are from each other, neither in space, time or both. Logically, there seems to be only one solution: *One electron passes through both slits at once.*

Another famous way of performing this thought experiment that we might try after the above observations is to place a detector measuring passing electrons in front of the right or the left slit, instead of blocking it. This way, we allow the electrons to pass through either slit, as both remain open. The results are just as staggering. Even when both slits are open we do not obtain an interference pattern but instead two strong bands of detection. Remove the detector, and the interference pattern is back again. Not only have we found that electrons seem to be affected by the possibility of whether one or two slits remain open, but it seems as though the process of measuring it on its path to the final detector has changed the outcome of the experiment. Try to find what path the electron is going, and the interference pattern always dissapears.

How can we explain this phenomena with the quantum physics developed in this essay? Starting with the Schrödinger equation (172) in one spatial dimension[37], we have that

$$-\frac{\hbar^2}{2m_e}\frac{\partial^2}{\partial x^2}\psi(x,t)+V(x,t)\psi(x,t)=i\hbar\frac{\partial}{\partial t}\psi(x,t). \tag{173}$$

In the first part of Tonomura et al. experiment, the electrons are traveling through free space. In fact, the potential created by the positively charged wire is only used for the purpose of recreating the double slit experiment, and if diffraction caused by this fact is treated later, we do not have to consider any potential in the Schrödinger equation. Therefore, we set $V(x,t)=0$,

$$-\frac{\hbar^2}{2m_e}\frac{\partial^2}{\partial x^2}\psi(x,t)=i\hbar\frac{\partial}{\partial t}\psi(x,t). \tag{174}$$

It is reasonable to assume that the solution of this equation is in the form of a traveling wave, due to its similarity with the wave equation. We therefore guess that

[37] Greenstein, George; Zajonc, Arthur G.; *"The Quantum Challenge: Modern Research on the Foundations of Quantum Mechanics"*; Jonas and Bartlett Publishers Inc., second edition (2006), p. 13-18.

$$\psi(x,t) = Ae^{i(kx-\omega t)} \tag{175}$$

where A is the real amplitude and k and ω are constants. We directly see that

$$-\frac{\hbar^2}{2m_e}\frac{\partial^2}{\partial x^2}\psi(x,t) = +\frac{k^2\hbar^2}{2m_e}\psi(x,t) \tag{176}$$

$$i\hbar\frac{\partial}{\partial t}\psi(x,t) = +\omega\hbar\psi(x,t), \tag{177}$$

and therefore

$$\frac{k^2\hbar^2}{2m_e}\psi(x,t) = \omega\hbar\psi(x,t). \tag{178}$$

Then, $\psi(x,t) = Ae^{i(kx-\omega t)}$ will be a solution to the Schrödinger equation under the constraint

$$k = \sqrt{\frac{2m_e\omega}{\hbar}} \tag{179}$$

which is of course familiar, and it seems as though we are correct in considering the wave function of having the behavior of a traveling wave. Let us rephrase this equation with more intuitive quantities, wavelength λ and period P. We begin to consider the wave function at $t = 0$, as $\psi(x,0) = Ae^{ikx}$, where at $x = 0$ we must have $\psi(0,0) = A$. One wavelength is then of course the quantity $x = \lambda$ that returns A. From Eulers identity, we find that

$$e^{2\pi i} = 1 \tag{180}$$

which equals one revolution. Hence, we must have that $k\lambda = 2\pi$, which gives us

$$k = \frac{2\pi}{\lambda}. \tag{181}$$

Similarly, from the fact that $\psi(0,t) = Ae^{-i\omega t}$, we use the same argumentation to obtain

$$\omega = \frac{2\pi}{P}. \tag{182}$$

Thus, we can rewrite our initial guess as

$$\psi(x,t) = Ae^{2\pi i\left(\frac{x}{\lambda}-\frac{t}{P}\right)}. \tag{183}$$

Now that we have a correct and intuitive solution to the Schrödinger equation in free space, we can start to discuss interference phenomena. As a plane wave approaches the double slit in the experiment of A. Tonomura et al., we should expect diffraction at each slit. The entire system should then be described as as the sum of these two diffracted waves,

$$\psi(x,t) = \psi_L(x_L,t) + \psi_R(x_R,t) \tag{184}$$

where L and R represents the wave passing through the Left and Right slit, respectively, and

$$\psi_L(x_L,t) = A_L e^{2\pi i\left(\frac{x_L}{\lambda}-\frac{t}{P}\right)} \tag{185}$$

$$\psi_R(x_R,t) = A_R e^{2\pi i\left(\frac{x_R}{\lambda}-\frac{t}{P}\right)}. \tag{186}$$

We can factor out the time parameter, as it is the same for both waves, as

$$\psi(x,t) = \left(A_L e^{2\pi i\frac{x_L}{\lambda}} + A_R e^{2\pi i\frac{x_R}{\lambda}}\right) e^{-2\pi i\frac{t}{P}} = \alpha(x_L,x_R)e^{-2\pi i\frac{t}{P}}. \tag{187}$$

Now we use the born rule to obtain the probability of finding the particle at a point on the screen during a measurement. We find that

$$\begin{aligned} &\left|\psi(x,t)\right|^2 \\ &= \psi^*(x,t)\psi(x,t) \\ &= \alpha^*(x_L,x_R)e^{2\pi i\frac{t}{P}}\alpha(x_L,x_R)e^{-2\pi i\frac{t}{P}} \\ &= \alpha^*(x_L,x_R)\alpha(x_L,x_R) = \left|\alpha(x_L,x_R)\right|^2 \end{aligned} \tag{188}$$

After which we then see that

$$\begin{aligned} &\left|\alpha(x_L,x_R)\right|^2 \\ &= \left(A_L e^{-2\pi i\frac{x_L}{\lambda}} + A_R e^{-2\pi i\frac{x_R}{\lambda}}\right)\left(A_L e^{2\pi i\frac{x_L}{\lambda}} + A_R e^{2\pi i\frac{x_R}{\lambda}}\right) \\ &= A_L^2 + A_R^2 + A_L A_R\big(e^{-\frac{2\pi i}{\lambda}(x_L-x_R)} + e^{\frac{2\pi i}{\lambda}(x_L-x_R)}\big) \\ &= A_L^2 + A_R^2 + 2A_L A_R\cos\left[\frac{2\pi}{\lambda}(x_L-x_R)\right], \end{aligned} \tag{189}$$

Where we have used the fact that

$$\cos\theta = \frac{e^{-i\theta}+e^{i\theta}}{2}. \tag{190}$$

Hence, the final result is

$$\left|\psi(x,t)\right|^2 = A_L^2 + A_R^2 + 2A_L A_R\cos\left[\frac{2\pi}{\lambda}(x_L-x_R)\right]. \tag{191}$$

This result can be interpreted as the frequency of which electrons hit the position $x_L - x_R$ on the screen. As can be see from the last term, A_L and A_R "interact" with each other and oscillate through the cosine factor. This is the so called *interference term*, and should be present when modeling all kinds of interference experiments. As it happens this equation predicts the structure observed in fig. (6), and we have hence found a quantum mechanical account for the double slit interference phenomena. If we were to block electrons from traveling through the left slit, we should follow the argumentation above using $\psi(x,t) = \psi_R(x_R,t)$ instead of $\psi(x,t) = \psi_L(x_L,t) + \psi_R(x_R,t)$, and similarly for blocking the right slit. No doubt, we then obtain A_R^2 or A_L^2, respectively, which correctly models the pattern obtained during this type of experiment. In explaining the last version of our experiment, where we make use of detectors instead of potential barriers, what we obtain is A_R^2 *and* A_L^2. Even though both wave function from diffraction are present, observation does not conclude in an interference pattern, and we should therefore not expect any interference terms in our theory. As stringently put by Gasiorowicz, there exists a simple rule: *"If the paths are not determined, add the wave function and square; if the paths are determined, square the wave function and add."*[38]. That is,

$$\left|\psi(x,t)\right|^2 = \left|\psi_L(x_L,t) + \psi_R(x_R,t)\right|^2 = A_L^2 + A_R^2 + 2A_L A_R \cos\left[\frac{2\pi}{\lambda}(x_L - x_R)\right] \tag{192}$$

$$\left|\psi(x,t)\right|^2 = \left|\psi_L(x_L,t)\right|^2 + \left|\psi_R(x_R,t)\right|^2 = A_L^2 + A_R^2 \tag{193}$$

for the undetermined electron path and the determined electron path, respectively. With this rule at hand, we have been able to theoretically explain all aspects of the double slit experiment brought up here from the theory of quantum physics built in this essay.

6 Discussion

So, *How exactly did physicists figure out that classical physics was insufficient? Why do we use Hermitian operators in Hilbert space, and how did physicists infer the commutation relations of quantum mechanics? How can the equations of motion be derived from first principles?* It is clear that these questions all have been answered in this essay. Through the use of quanta, we saw how physicists of the early twentieth century could explain a large set of experimental data. In particular, we note how physicists never actually believed in the planetary model of the atom, as is often claimed, and that it was directly realised that classical electrodynamics could not explain what was observed. To solve this problem, Bohr introduced his theory of stationary states, which neither Bohr himself nor physicists at the time understood through the classical motion of particles.

Before introducing the postulates of canonical quantization we saw how Heisenberg, Born and Jordan developed matrix mechanics almost simultaneously as Schrödinger developed quantum wave mechanics. Dirac further discovered that the commutator presented by Born and Jordan was closely linked to the Poisson bracket in Hamiltonian mechanics, leading him to develop canonical quantization. Through this history, we understood why we make use of Hermitian opera-

[38] Gasiorowics, Stephen; *"Quantum Physics"*; john Wiley Sons, Inc., third edition (2003), p. 30

tors in Hilbert space and why the commutator is so central for quantization. Using this new found understanding we developed the Heisenberg and Schrödinger equations of motion through first principles, and used the Schrödinger equation to account for the results of the double slit experiment of electrons.

It is of great importance to realize what it is that we have done, and not done, in this essay. This work, excluding the discussion of the double slit experiment, concerns the history of quantum mechanics from 1900 to around 1926-27, and therefore the development of *first quantization* as opposed to the *second quantization*. First quantization only considers semi-classical models, most notably through quantizing particles within a classically described potential, as is done when for example obtaining the stationary states of the hydrogen atom in early courses of quantum mechanics. As was first developed by Dirac, second quantization involves the quantization of the potential itself by the use of *field operators*. It is evident from the chronological presentation provided in this paper that a natural continuation of it would be the development of canonical quantization of quantum field theories, specifically quantum electrodynamics. There, we would see how a Lorentz invariant field description of quanta gives rise to quantized particles containing a natural language for describing their respective interactions.

As hinted above, it would be hearty to claim that we have actually *built* quantum mechanics from scratch. As detailed in the *Groenewold Van Hove* theorem[39], an invertible linear map

$$\mathscr{M} : f(q,p) \mapsto \hat{Q}_f \tag{194}$$

of a function $f(q,p)$ on phase space to a Hermitian operator $\hat{Q}_f$ such that

$$\hat{Q}_{\{f,g\}} = \frac{1}{i\hbar}[\hat{Q}_f, \hat{Q}_g], \tag{195}$$

where $g(q,p)$ is another function on phase space, does *not* exist in general[40]. In switching the equal sign to a mapping arrow, equation (195) would be a condensed way of stating postulate 3.3, meaning that we not even in principle can describe all functions on phase space with Hermitian operators, and vise versa. An above example was spin quantum numbers, which is indeed not described by canonical coordinates. To describe spin, we need to add an additional postulate above describing spin statistics[41]. Hence, the above postulates of canonical quantization cannot account for spin statistics, and we could further build on quantum mechanics.

What is easy to overlook is the fact stated outright in chapter 3: *there exists no proof of quantum theory.* To *actually* build quantum mechanics from

[39] Todorov, Ivan; *"'Quantization is a mystery'"*; arXiv:1206.3116, (2012).

[40] By "invertible linear map" it is meant that there exists a linear operator D such that $Df(q,p) = \hat{Q}_f$, and that D^{-1} exists.

[41] In non-relativistic quantum mechanics, which was considered above.

scratch, we must move beyond canonical quantization using other quantization methods such as *Deformation* or *Geometric* quantization. It must be said, however, that even through these methods there is no universally valid scheme to quantize classical theories satisfactorily. For the interested reader, consider as an example the quantization of gravity. General relativity *can* be quantized if represented as an effective field theory, but it is hopelessly unrenormalizable at the high energy limit[42]. In fact, this is one of the core reasons for why learning about the development of canonical quantization *matters*; quantization is an unfinished project, and in order to understand what must be done in the future we must understand the ideas that has been used in the past. Also, and in some ways because, there is no magical recipe for developing a satisfactory quantum theory, a vast set of experiments and models under our belt today has been left out here in the interest of time. Nature is messy; without studying all fringe cases and their experimental context, one cannot claim to have built quantum mechanics from scratch, only parts of it.

What we *have* done, however, is realizing how we can obtain one such schematic method for quantization, and indeed see how it can account for experimental results. Chapter 4 stopped at a peculiar place, namely where you probably started at your very first course of quantum mechanics. Hopefully, this essay has given you the tools to understand the large jump that we spoke of in the introduction, helping you grasp the though process behind some of the earlier results of quantum mechanics.

7 Conclusions and Summary

During chapter two, we accounted for some famous early experiments of old quantum theory. In chapter 3 we saw how quantum mechanics was developed through matrix and wave quantum mechanics as a direct consequence of old quantum theory, leading to canonical quantization and the modern formalism we use today. We developed the Heisenberg and Schrödinger equations through the use of displacement operators from first principles to describe the double slit experiment in a self contained manner, as presented in chapter 4. We discussed the limitations of canonical quantization, realizing that there is no universal scheme for quantization, but concluded that the questions posed in the introduction of this essay has indeed been answered.

8 Bibliography

[1][7][11][13][37][38] Gasiorowics, Stephen; *"Quantum Physics"*; john Wiley Sons, Inc., third edition (2003), p. 1, p. 7, p. 10, p. 15, p. 30.

[42] Donoghue, John F.; *"The effective field theory treatment of quantum gravity"*; arXiv:1209.3511, (2012).

[2] Einstein, Albert; *"On the Quantum Theory of Radiation"*; Phys. Zs. 18 121, (1917).

[3] Xiaomin, Lu; Qishan, Tang; Wendong ,Kang; *"Research and Verification of Blackbody Radiation Law"*; Insight - Energy Science, (2018).

[4] Müller, J.; *"Physikens Grunder"* (English, *"The foundations of Physics")*; Stockholm, Zacharias Hæggström, (1854).

[5][37] Greenstein, George; Zajonc, Arthur G.; , *"The Quantum Challenge: Modern Research on the Foundations of Quantum Mechanics"*; Jones and Bartlett Publishers Inc., second edition (2006), p. 24-37.

[6] Thompson, Joseph J.; *"Cathode Rays"*; London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, (1897).

[8] Compton, H. Arthur. *A Quantum Theory of the Scattering of X-rays*, Phys. Rev. 21,483; 22,409 (1923).

[9] Compton, H. Arthur. *The Spectrum of Scattered X-rays*, Phys. Rev., vol. 22 No 5 (1923).

[10] De Broglie, Louis; *"The wave nature of the electron"*; Nobel Lecture, (1929).

[12] Rutherford, Ernest; *"The Scattering of and Particles by Matter and the Structure of the Atom"*; Philosophical Magazine Series 6, vol. 21 (1911).

[14] Frank, James; Hertz, Gustav; *"Über Zusammenstößе zwischen Elektronen und Molekülen des Quecksilberdampfes und die Ionisierungsspannung desselben" (english: "On the collisions between electrons and molecules of mercury vapor and the ionization potential of the same")*; Verhandlungen der Deutschen Physikalischen Gesellschaft (1914); Translated version in: Boorse, A. Henry; Motz, Lloyd; *"The World of the Atom"*; Basic Books Inc, Publishers (1966).

[15] Robson, Robert E.; White Ronald D.; Hildebrandt, Malte; *"One hundred years of the Franck-Hertz experiment"*; Eur. Phys. J. D 68: 188, (2014).

[16][26] Nakahara, Mikio. *"Geometry, Topology and Physics"*, second edition, Taylor and Francis group (2003), p. 9, p. 10.

[17][24] Von Neumann, John; *"Mathematical Foundations of Quantum Mechanics"*; Princeton University Press, (1955), p. 3-17, p. 17-33.

[18] Born, Max; Heisenberg, Werner; Jordan, Pascual; *"Zur Quantenmechanik II"*; Zeitschrift für Physik 35 (1926). English translation: *"On Quantum Mechanics II"*; Translated version in: van der Waerden, B. L. (editor); *"Sources*

of Quantum Mechanics"; Dover Publications, (1968).

[19] Heisenberg, Werner; *"Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen"*; Zeitschrift für Physik 33 (1925). English translation: *"Quantum-Theoretical Re-interpretation of Kinematic and Mechanical Relations"*; Translated version in: van der Waerden, B. L. (editor); *"Sources of Quantum Mechanics"*; Dover Publications, (1968).

[20] Born, Max; Jordan, Pascual; *"Zur Quantenmechanik"*; Zeitschrift für Physik 34 (1925). English translation: *"On Quantum Mechanics"*; Translated version in: van der Waerden, B. L. (editor); *"Sources of Quantum Mechanics"*; Dover Publications, (1968).

[21] Fedaka, William A. ; Prentis, Jeffrey J.; *"The 1925 Born and Jordan paper "On quantum mechanics""*;American Journal of Physics 77, 128 (2009).

[22] Mehra, Jagdish; Rechenberg, Helmut; *"The historical development of quantum theory: Vol. 3, The formulation of matrix mechanics and its modifications 1925-1926"*; Springer-Vlg, cop. (1982).

[23] Galler, Anna; Canfield, Jeremy; Freericks, James K.; *"Schrödinger's original quantum-mechanical solution for hydrogen"*; arXiv:2007.14798, (2020).

[25][27][28][29][32][33][34] Dirac, Paul A. M.; *"The Principles of Quantum Mechanics"*; Oxford University Press, Third edition (1947) (first edition published (1930)), p. 17, p. 84-89, p. 28, p. 99-103, p. 108-111, p. 111-118.

[31] Jaffe, R. L.; *"SUPPLEMENTARY NOTES ON DIRAC NOTATION, QUANTUM STATES, ETC. "*; Physics 8.05, (2007).

[35] Rubens, de Melo Marinho Jr; *"Noether's theorem in classical mechanics revisited"*; arXiv:physics/0608264, (2006).

[36] Tonomura, Akira.; Endo, Junji; Matsuda, Takashi; Kawasaki, Takeo; *"Demonstration of single-electron buildup of an interference pattern"*; American Journal of Physics 57, 117 (1989).

[39] Todorov, Ivan; *"'Quantization is a mystery'"*; arXiv:1206.3116, (2012).

[42] Donoghue, John F.; *"The effective field theory treatment of quantum gravity"*; arXiv:1209.3511, (2012).