1. Introduction

In this chapter, I will spell out as clearly as possible what are the foundational ideas that underlie the quantum theory, and the theory of gravity according to the theory of general relativity, and how these theories of matter differ from the classical views.

On the quantum theory, we will first discuss how it appeared on the scene in physics in the 1920s, including the ensuing philosophical basis of this theory, and then move on to some of the differing interpretations – of Schrödinger, Bohr/Heisenberg, Born, de Broglie, Bohm and Einstein.

The next section will focus on the history and philosophy of the theories of gravity. This section starts with ideas of Aristotle, of the ancient times, to ideas of Galileo and Newton, of the Renaissance period, and then to Einstein, of the contemporary period. Conceptual differences will be discussed between the classical views of Newton (action at a distance and atomism) and the theory of general relativity (holism, continuity, and forces propagating between interacting bodies at a finite speed).

Comments will then be directed to the possibility of a quantum theory of gravity. There has been a great deal of discussion on this topic in the current literature. Nevertheless, I will argue that because of the incompatible underlying concepts and mathematical expressions of the quantum and general relativity theories, it is not possible, in principle, to unify them into a quantum theory of gravity.

The approach of my research program will then be spelled out, which starts at the outset with the most general expression of the theory of general relativity, as foundational, while discarding the formal expression of the quantum theory. It is found that the formal (Hilbert space) expression of the quantum theory appears as a linear approximation for a generally covariant field theory of the inertia of matter in general relativity, which originates in a unified field theory. The origin of the field unification of the inertial manifestation of matter and its gravitational and electromagnetic force manifestations are shown to follow from fully exploiting the algebraic as well as the geometrical implications of the principle of covariance – the underlying axiom of the theory of general (and special) relativity. The Mach principle, in a generalized form, and gauge covariance also will be shown to play essential roles in the structuring of this unified field theory.

1.1 History and Philosophy of the Quantum Theory

1.1.1 Blackbody Radiation

Near the turn of the 20th century it was discovered by Max Planck that the spectral distribution of blackbody radiation could be fitted to a model whereby the energy in each of the modes of the frequency components of radiation enclosed in a cavity is linearly proportional to its frequency.

The experiment is as follows: a cavity with a small window in its side is maintained in a heat bath at a constant temperature. The walls of the cavity are in thermodynamic equilibrium with the radiation that is emitted and reabsorbed at the same rate by these walls. This is called blackbody radiation. A filter is placed in the window of the cavity that will only transmit a single frequency radiation component. The intensity of this radiation is then measured. The observation is then continued for the visible spectrum of frequencies, by placing the appropriate filters in the window of the cavity. One then plots the intensity of the radiation field in the cavity as a function of frequency. This is the spectral curve for blackbody radiation. Plotting the intensity as a function of frequency ν, it is found to have a characteristic shape with a maximum, and then go to zero at zero frequency and at infinite frequency. Repeating the experiment at different temperatures yields the same general shape for the spectral curve of blackbody radiation.

The blackbody radiation curves were not the expected ones, according to the analysis using classical statistics (Rayleigh–Jeans). Their result depended on the assumption that the energy in each radiation mode in the cavity depends on the square of its frequency. Their prediction was that, as the wavelength $\lambda = c/\nu \to 0$, where c is the speed of light, the intensity of the radiation diverges to infinity. This is the so-called ultraviolet catastrophe. In contrast with this prediction, the empirical result is that as the wavelength decreases, the intensity goes through a maximum and then decreases to approach zero (in the ultraviolet region) as the wavelength approaches zero.

Planck analyzed this problem with the assumption that the energy of a mode of the enclosed radiation field in the cavity is linearly proportional to its frequency, $E_\nu = h\nu = hc/\lambda$. The constant of proportionality h is Planck's constant. It was found to be a universal constant of nature. With this assumption, an analysis by Planck, based on Maxwell–Boltzmann classical statistics for the 'gas' of radiation modes in the cavity (where each mode is distinguishable, i.e., 'tagged'), he correctly predicted the curve that was observed for blackbody radiation.

Planck's discovery of the linear relation $E_\nu = h\nu$ between the energy and the frequency of a radiation mode in the cavity led to the idea of the 'quantization' of electromagnetic radiation, and the idea of a particle of light, called a photon. This theory is referred to as the old quantum theory.

It is important to note at this stage of the discussion that the 'photon' is not the same sort of particle as the 'electron'. The electron has inertial mass. It can thus be slowed down or speeded up by an external electric force. But, as we will see later on in our discussion of relativity theory, the 'photon' can only propagate at a constant speed c, the speed of light. Since it has no inertial mass, it cannot be slowed down or speeded up from the constant speed c. All that can happen to a photon (after it as been created with energy E_γ and speed c) is that it can be annihilated when it is absorbed by matter, which in turn elevates the energy of the matter by $E_i - E_j = E_\gamma = h\nu$ units of energy.

1.1.2 Photoelectric Effect

A beam of monochromatic light (single frequency) impinges on a piece of metal that in turn is in an electrical circuit, with a battery, a resistor and an ammeter. The ammeter reads the current in the circuit, $I = V/R$, where V is the voltage across the resistance R. At first, the current is constant. When the frequency of the light is increased to some threshold value ν_0, the current reading increases. Continuously increasing the light frequency from that threshold value increases the current reading linearly. This is the photoelectric effect.

The voltage change ΔV across the resistor corresponding to the frequency change $\Delta\nu$, multiplied by the electron charge e is

$$\Delta E = e\Delta V = eR\Delta I = h\Delta\nu \; .$$

Thus, Planck's constant h appears again in the photoelectric effect, the same constant that appeared in the fit to the blackbody radiation spectral curve.

There is much more to say about the photoelectric effect, but this is the essence of it. It was a further substantiation of the quantization of monochromatic light in terms of its energy $h\nu$. In spite of these results from blackbody radiation, analyzed by Planck, and the photoelectric effect, analyzed by Einstein, neither Einstein nor Planck accepted the concept of the 'photon' as an elementary particle of light. In a letter to Einstein (6 July, 1907), Planck said: "For I do not seek the meaning of the quantum of action (light quantum) in the vacuum, but at its site of absorption and emission." In a letter to Laub (4 November, 1910), Einstein said: "I am very hopeful that I will solve the radiation problem and I will do it without light quanta."[1]

In the theory developed by Einstein, the photon plays the role of a virtual field that propagates as a signal between interacting electrically charged matter, to affect their mutual interaction. But the 'photon' is not a thing on its own, in this view. Similarly, the view of Planck was that what is 'quantized' is a gas of radiation in the cavity – a set of modes of radiation that couple charged matter of the walls of the cavity and the matter that constitutes the measuring apparatus that 'looks into' the cavity to measure the frequencies of these modes.

1.1.3 Compton Effect

Another important experiment that ushered in the 'old quantum theory' was the Compton effect. Here, a photon scatters from an electron, thereby changing its energy (and therefore its frequency) to a lower value, while the electron increases its energy by the same amount. It was found that the scattering angle of the electron was dependent on the same constant h determined in the blackbody radiation spectrum and the photoelectric effect, thereby verifying the assumption of the quantization of light.

1.1.4 Atomic Spectra and the Bohr Atom

A seminal observation in the early days of the 'old quantum theory' related to the measurement of the emission spectrum of excited atoms. The Ritz combination rule said that the frequencies of any of the various lines of the emission spectrum of hydrogen added as follows:

$$\nu_1 = \nu_2 + \nu_3 = \nu_4 + \nu_5 + \nu_6 = \dots ,$$

---

[1] Both quotations are from Anna Beck, *The Collected Papers of Albert Einstein*, Vol. 5: The Swiss Years Writings, 1902–1914, English translation, Princeton, 1994.

where ν_j is the frequency of any line in the observed spectrum. Niels Bohr explained this rule with his model of the atom, whereby the orbital electrons are in 'quantized orbits', with respective energies $E_1, E_2, \dots, E_6$, relative to the positively charged nucleus of the atom. The idea is the following. When the orbital electron drops in energy from the j th state to the i th state, it loses energy $E_j - E_i$, which then transforms into the energy of a created photon $E_\gamma = h\nu_{ij}$. It then follows that for a series of de-excitations between the different states of the atom, the energy losses (its emissions) are:

$$\begin{aligned} E_\mathrm{f} - E_\mathrm{i} &= (E_\mathrm{f} - E_1) + (E_1 - E_\mathrm{i}) \\ &= (E_\mathrm{f} - E_3) + (E_3 - E_5) + (E_5 - E_\mathrm{i}) \ , \end{aligned}$$

and so on. Dividing this equation by Planck's constant h, and using the relation

$$\frac{E_j - E_k}{h} = \nu_{ki} \ ,$$

we have the empirically verified Ritz combination rule for the frequencies in the emission spectrum:

$$\nu_\mathrm{fi} = \nu_\mathrm{f1} + \nu_\mathrm{1i} = \nu_\mathrm{f3} + \nu_{35} + \nu_\mathrm{5i} \ ,$$

and so on.

Questions that arise in regard to the Bohr model are as follows:

- What is the physical cause for the de-excitation of the atomic electron from the higher to the lower energy levels?
- What is the physical cause for the creation of a photon when the atom de-excites? Of course, the latter process conserves energy in that the energy lost by the atomic electron is given to the created photon. But it does not explain how the photon was created from a vacuum (i.e., from no photons).
- When the electron loses energy by dropping from one energy level to a lower one, but before it reaches the lower energy level, it would have lost energy while the photon is not created until the electron does reach the lower energy level. Thus it seems that energy is conserved only when one sees the electron in one energy state or another, but not when it is in transition between energy levels. How is this explained in the context of an energy-conserving system? That is, does the law of conservation of energy apply only when one is looking at the atom in one state or another, but not when the electron is in transition between the states of the atom? The idea of the

> Copenhagen school is that the conservation laws apply only when a macro-observer is observing the atom in one state or another, but not when the atomic system is in transition between its states. The physical properties of a microsystem of matter that is measured by a macro-observer then defines the micromatter. This is a crucial aspect in the theory of matter according to the new quantum view that was to emerge with the Copenhagen school.

Summing up, the experiments of the 'old quantum theory' at the turn of the 20th century, on blackbody radiation, the Compton effect, the photoelectric effect, atomic spectra, were all explained in terms of the 'photon', and the new universal constant, Planck's constant h, that entailed the quantum of radiation $h\nu$. This evolved into the new quantum theory, called quantum mechanics, in the 1920s. This development will now be outlined in an historical context.

1.1.5 The Seminal Experiment: Electron Diffraction

The seminal experiment that led to the ideas of the Copenhagen School was the set of observations in 1927, by C.J. Davisson and L.N. Germer in the US and by G.P. Thomson, in the UK, that electrons can scatter from a crystal lattice as though they are continuous waves. These experiments were preceded three years earlier by the theoretical speculation of Louis de Broglie on the possibility that an electron (or any other elementary particle with mass) has a wave nature. According to de Broglie's hypothesis, its particle-like nature – its momentum p – relates to its wave-like nature – its wavelength λ – according to the reciprocal relation $p = h/\lambda$.

The experiments on electron diffraction verified de Broglie's speculation about the 'matter wave'. A question then arises. If the electron is indeed a massive particle, then one should expect that its scattering from a crystal lattice would reveal, on an absorbing screen, a geometrical mapping of the atoms of the crystal lattice. Instead, what was seen was that, when the lattice spacing is the same order of magnitude as the de Broglie wavelength of the electron $\lambda = h/p$, the distribution of scattered electrons on the absorbing screen was not unlike the pattern of X-rays that have been scattered by a crystal lattice. That is, there are regions of constructive interference, where the electrons bunch together, and there are regions of destructive interference where no electrons are seen to land. Of course, one expects scattered X-rays to land this way on the absorbing screen (as had been seen earlier by Bragg) because we know at the outset that X-rays are electromagnetic waves, whose wavelength is the order of magnitude of the lattice spacing of the crystal. It is the expected diffraction pattern for the scattering of waves from a crystal lattice.

But why should electrons diffract in this way if indeed they are discrete particles of matter? One answer is that it may be an illusion that electrons are truly point particles; rather, they may be, most fundamentally, (matter) waves at the outset. This is a difficult explanation because there are experimental circumstances wherein electrons appear to be point particles of matter, such as the cathode ray experiment of J.J. Thomson.

With these two empirical facts in mind, Niels Bohr and the Copenhagen school proposed a resolution to extend Einstein's idea of wave–particle dualism from the (massless) photons of electromagnetic radiation to material particles with finite mass, such as electrons. The idea was then that, when one does an experiment to view the electron as a discrete particle, it is such a particle at that time. But when an experiment is carried out to view the electron as a wave, it is a wave at that time. Both statements – that the electron can be a particle and that it can be a wave – are taken to be true, so long as the observations of these states of the electron are not seen by the macro-observer simultaneously.

This assertion fits in with the epistemological stand of logical positivism. It is a philosophical approach, first proposed by Ernst Mach and the Vienna Circle, around 1900. It is the idea that in principle the only meaningful statements about nature must be verifiable with the human senses or their measuring instruments. This is called 'the principle of verifiability'. It was claimed by Bohr and Heisenberg to be a natural epistemological basis of knowledge about the natural world. In his initial paper on quantum mechanics, W. Heisenberg said: "The present paper seeks to establish a basis of theoretical quantum mechanics founded exclusively upon relationships between quantities which in principle are observable."[2] This is the view of logical positivism, based on the principle of verifiability. Bertrand Russell gave a well-known refutation of the principle of verifiability: "This principle is not, in itself, verifiable by the human senses or instruments. Thus, if it is true, it must be false. Therefore it is false".

According to Bohr's and Heisenberg's positivistic view, then, the observer must be involved in the definition of what an electron is, rather than an outside experimenter who probes the nature of the ob-

---

[2] W. Heisenberg, Zeits f. Physik **33**, 879 (1925).

jective thing, called ‘electron’, independent of himself (or herself). The Copenhagen view then defines the elements of matter in an irreducibly subjective manner.

An epistemological view different than this one is the idea of realism, wherein the elements of matter are what they are, independent of who or what may be probing their properties. In this approach, what we see in experimentation must then be rationally interpreted to arrive at assertions about what the nature of this matter is, independent of ourselves as observers and our mode of measurement.

Niels Bohr and Werner Heisenberg led the philosophical view of positivism in physics. Most of the physics community has followed this approach from the 1920s until this time. But there were some very notable physicists in the 20th century who took the stand of realism in physics, such as Einstein, Planck, Schrödinger, de Broglie and Bohm. The former positivistic philosophy is that of the quantum theory. The latter realist philosophy is that of the theory of general relativity. It is the realist philosophy that I believe is the one where the truth lies in science, and will flourish in 21st century physics.

Not too long after the experimental discoveries of the wave nature of matter, in the 1920s, Erwin Schrödinger discovered the equation whose solutions are the matter waves. This is the so-called Schrödinger wave equation. His formalism correctly predicted the energy levels of atoms and the transitions between its states. Around the same time Werner Heisenberg discovered an equation in terms of matrices of numbers, representing the discrete observables such as energy values, in particular states of an atom, as well as predicting correctly the transitions between the states of the system. This formalism is called Heisenberg’s matrix mechanics. Thus both the continuous wave theory of Schrödinger and the discrete matrix theory of Heisenberg made identical predictions for the atomic states of matter.

Not long afterward, it was shown by C. Lanczos, and independently by Schrödinger, that the Schrödinger representation of the quantum theory and the Heisenberg representation could be mathematically transformed into one another. Lanczos showed this equivalence by transforming the differential equation of Schrödinger into an integral equation. From there, he was able to demonstrate the Heisenberg form of matrix mechanics. Thus the Schrödinger equation and the Heisenberg equation are equivalent formalisms; this is the reason that they gave identical predictions, though this was not obvious to the physicists at first glance! They called these mathematical expressions, which correctly gave predictions of the physical properties of atomic matter, quantum mechanics.

1.1.6 Interpretations of Quantum Mechanics

Soon after the discovery of quantum mechanics, Max Born found that one could express the formalism in terms of a probability calculus. Thus, Born, Bohr and Heisenberg interpreted Schrödinger's matter waves as waves of probability. This was to be the (complex number) amplitude whose absolute square is the probability that the particle of matter whose properties are being measured is at a particular point of space. The probability was then tied to quantum mechanics as a theory of measurement – made by a macro-observer on micro-matter. This view was then in line with the positivistic philosophy, whereby the elements of matter are defined subjectively in terms of the measurements of their properties, expressed with a probability calculus. These ideas will be discussed in more detail in Chap. 4. Proponents of these ideas are said to belong to the 'Copenhagen school'.

While the majority of physicists have accepted the truth of the ideas of the Copenhagen school, there have been other interpretations of the empirically successful equations of quantum mechanics. I will now briefly describe some of these.

Schrödinger's View. E. Schrödinger did not accept the probability interpretation of his wave equation at the outset. He thought of his wave solutions – the matter waves – as real waves, just like ocean waves or waves of electromagnetic radiation. His idea, after seeing the results of the electron diffraction experiments, was to complete the Maxwell formulation of electromagnetic field theory by properly expressing the real-number-valued source terms (on their right-hand side) in terms of complex functions that are the de Broglie matter waves ψ – the solutions of Schrödinger's wave equation. He found that these source terms might be factorized into a product of ψ – a complex function – and its complex conjugate $\overline{\psi}$. Thus the charge density becomes $\rho = e\overline{\psi}\psi$ and the current density becomes $\boldsymbol{j} \propto \mathrm{Re}(\overline{\psi}\nabla\psi)$, where the overline denotes (henceforth) the complex conjugate, e is the electron charge and ∇ is the gradient operator.

Thus Schrödinger believed that the wave nature of electrons (and any other electrically charged elementary matter) is implicit in the real number-valued source terms of Maxwell's field equations for electromagnetism. It is not 'unfolded', that is, available to be observed, until experiments such as electron diffraction are carried out.

De Broglie, Bohm and the Hidden Variable View. Not long after the successes of quantum mechanics, Louis de Broglie suggested his 'double solution' interpretation of the quantum formalism. This was an approach, like Schrödinger's, that attempts to restore determinism to physics. His idea was that, in addition to the probability calculus of quantum mechanics, there must be variables that relate to the objective electron, independent of anyone's observation of it. He saw the Schrödinger wave equation as a subjective part of the theory of the electron, relating only to the measurements of its properties. But he believed that, buried inside of the probability wave there must be a singular function ζ representing the real electron, independent of any measurements on its properties. Thus, to complete the description of the electron, there must be another mathematical equation in ζ depending on the space and time variables. It is this equation that entails the actual dynamics of the real electron.

De Broglie argued that this added function must be a point singularity, coupled to the Schrödinger wave ψ. The latter complex function, in turn, must influence the function ζ, because, he felt, the probability wave must in some way guide the point singularity wave of the electron. He then concluded that both of these functions must influence each other, and that the equation in ζ must be nonlinear, while the equation in ψ must remain linear, since it is to represent a solution for a probability calculus. The added function ζ is then not directly observed – it is a hidden variable that relates to the deterministic electron. This view is called de Broglie's double-solution interpretation of quantum mechanics. The analysis has yet to be further analyzed and taken to completion.

In the 1950s, in trying also to restore determinism to physics, David Bohm took a different view of hidden variables to de Broglie. What he did was to add to the independent variables, additional hidden parameters that the matter field depends on:

$$\psi(\boldsymbol{r}, t) \longrightarrow \psi(\boldsymbol{r}, t, \Lambda) ,$$

where $(\boldsymbol{r}, t)$ are the ordinary space and time independent variables and $\{\Lambda\}$ are the additional hidden parameters. Then, the change of Λ with time is to denote the actual velocity of the particle's trajectory. The matter wave is viewed as the dependent variable $\psi\big(\boldsymbol{r}, \Lambda(t)\big)$. The equation in the matter field must then also entail a dependence on the changes of the hidden parameter Λ, to restore determinism in the description of the particle. What appears in this regard is an extra 'potential' in the equation for the matter field.