1

Description of a Physical System

Paul Dirac and Werner Heisenberg. B464, https://arkiv.dk/en/vis/5940636. Courtesy: Niels Bohr Archive.

The measure of greatness in a scientific idea is the extent to which it stimulates thought and opens up new lines of research.

—P. A. M. Dirac

From our day-to-day experience, we develop our notion of reality. However, our perception of physical properties, such as position and momentum, needs refinement to describe natural phenomena correctly. In this chapter we introduce fundamental principles of the quantum theory that does so with enduring cogency. The laws of nature that govern the functioning of the physical universe cannot be accounted for using classical physics of Newton, Lagrange, and Hamilton. Foundational principles and mathematical structure of quantum theory are introduced in this chapter.

1.1 Quantum versus Classical Physics Theories

In classical physics, the mechanical state of a system is represented by a point in the position–velocity phase space, or equivalently in the position–momentum phase space. The entire theoretical formalism of Newtonian–Lagrangian–Hamiltonian mechanics is based on this Galilean conjecture. The classical hypothesis seems appropriate in a large number of physical situations that concern our day-to-day experiences. With this ansatz, the temporal evolution of a system is described by the trajectory of the point in the phase space. The trajectory is obtained from the equation of motion that governs the time dependence of position and velocity, represented respectively by q and $\dot{q}$ (or alternatively the time dependence of position and momentum, represented respectively by q and p). The alternative equations of motion – Newton's, Lagrange's, and Hamilton's – are equivalent. Their applicability must be reconciled with the upper limit on accuracy with which *conjugate* physical properties of an object, viz. *position* and *momentum*, are simultaneously measured. There is, however, an *inverse* relation between the accuracy of simultaneous measurement of these two properties. The *more* accurately you measure either, the *less* accurately can you measure the other. The act of measuring either of the conjugate properties requires an observation.

Observations resulting in accurate measurements of conjugate variables are, however, not compatible with each other. Heisenberg's principle of uncertainty is the quantitative expression of this law of nature. It is expressed as a rigorous mathematical inequality that is neatly written in a compact form. Nonetheless, we refrain from advancing its mathematical expression too soon. It has no classical analogue. It cannot be written in any terms of what we are familiar with from classical mechanics. *Symbols* for position and momentum that are used to express the uncertainty principle are also found in classical mechanics, but they have a *new meaning* in quantum theory. Heisenberg's principle of uncertainty and also the Schrödinger equation will be introduced in Section 1.4 after the required notation is formally construed in Sections 1.2 and 1.3.

Classical and quantum physics are mathematical models which both aim at describing the state of a physical system and its temporal evolution. There certainly are not *two* laws of nature, one for large objects and the other for small. Many, but not all, macroscopic phenomena are fairly well accounted for by classical physics. If one comes across a claim that classical laws work for macroscopic objects, it must be understood only in the sense that in several macroscopic events, classical laws are very good *approximations* to quantum laws. In the present section, a few salient features of classical physics will be briefly recapitulated. There are, of course, many excellent books on classical mechanics. The formalism and notation used in Reference [1] provides a well-suited platform for topics developed in the present book. Classical physics has two alternative and equivalent formulations. One of these is based on the principle of causality and determinism. Its backbone is the linear stimulus–response hypothesis advanced by Isaac Newton. Newtonian dynamics is based on Galileo's identification of the constancy of momentum of an object as determined only by its *initial* mechanical state. The stimulus (i.e., force) that *changes* the momentum of an object is exactly equal to the *rate* at which the change in momentum occurs. The application of a force on an object therefore imparts to it an acceleration that is *directly* proportional to the force itself. Newtonian formalism is therefore a linear stimulus–response theory.

It will be argued in the next chapter that the alternative formalism of classical physics based on the *principle of variation* is in some sense more powerful than Newton's method, although the two approaches produce equivalent results. In it, the notion of force is not used. Instead, methods of variational calculus are used. The principle of variation stipulates that a mechanical system is described by its *Lagrangian*, $L(q,\dot{q})$. The essential premise of this formulation is therefore the same as that of Newton's method, since the Lagrangian is given in terms of the generalized position q and the generalized velocity $\dot{q}$. Its simplest form is

$$L(q,\dot{q}) = f_1(\dot{q}^2) + f_2(q), \tag{1.1}$$

where f_1 and f_2 are suitable functions respectively of velocity and position. The dependence on the square of the velocity rather than on its first power is prompted by the isotropy of space. The equation of motion of the system is then determined on the basis of an ansatz that the definite integral over time from the initial time t_i to the final time t_f, i.e.,

$$S = \int_{t_i}^{t_f} L(q,\dot{q})\,dt, \tag{1.2}$$

called *action*, is an extremum. The choice

$$L(q,\dot{q}) = T(\dot{q}^2) - V(q) = \frac{1}{2}m\dot{q}^2 - V(q), \tag{1.3}$$

with the requirement that action is stationary, i.e.,

$$\delta S = 0, \tag{1.4}$$

leads one [1] to the equation of motion as the necessary and sufficient condition that describes motion:

$$-\frac{\partial V}{\partial q} = \frac{\partial L}{\partial q} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}. \tag{1.5}$$

In Eq. 1.3, T is the kinetic energy of the object, V its potential energy, and m its mass. That Eq. 1.3 is the right choice for the Lagrangian is seen on recognizing that it reproduces Newton's second law, since

$$\text{force} = -\frac{\partial V}{\partial q} \tag{1.6a}$$

$$\text{and } \frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{dp}{dt}, \tag{1.6b}$$

where p is the generalized momentum *defined* as the partial derivative of the Lagrangian with respect to $\dot{q}$.

Equation 1.5 is the Lagrange's equation of motion.

Instead of representing the state of a system by $(q,\dot{q})$ or $L(q,\dot{q})$, one may of course represent it by (q,p), or for that matter by a function $H(q,p)$ of the same.

The necessary and sufficient conditions for action to be an extremum turn out to be given by two first-order equations of motion,

$$\dot{p} = -\frac{\partial H(q,p)}{\partial q} \tag{1.7a}$$

and

$$\dot{q} = \frac{\partial H(q,p)}{\partial p}, \tag{1.7b}$$

$$\text{with } H(q,p) = T + V = \frac{p^2}{2m} + V(q). \tag{1.8}$$

Equation 1.7a, b are known as the Hamilton equations of motion, and H (Eq. 1.8) as the Hamiltonian.

Funquest: What constants of integration are required to solve (i) Newton's, (ii) Lagrange's, and (iii) Hamilton's equations of motion? How are these constants to be obtained?

Time-reversal symmetry of the equations of motion (Newton's, Lagrange's, or Hamilton's) ensures that we can not only predict from its solution the mechanical state of the system anytime in the future but also determine what it would have been at any time in the past. We advance a caution here that time-reversal symmetry has a very different connotation in quantum mechanics. We shall discuss it later.

Funquest: Show that (i) Newton's equation of motion, (ii) Lagrange's equation of motion, and (iii) Hamilton's equation of motion are symmetric under time-reversal $t \to -t$.

Classical mechanics is a study of the pair $(q, \dot{q})$ or equivalently of (q, p). The classical equations of motion require the initial conditions $(q_i, \dot{q}_i)$ or (q_i, p_i) at the initial time t_i. Lack of knowledge of either of the two physical properties in each of these pairs makes it impossible to determine the mechanical state of a system at an arbitrary time using the equations of motion.

A *pair of measurements* is necessary to determine $\left(q_i(t=0), \dot{q}_i(t=0)\right)$ or $\left(q_i(t=0), p_i(t=0)\right)$. Likewise, the correctness of the solution at an arbitrary time t to the equation of motion can also be verified only by a *measurement* of the *pair*, $\left(q(t), \dot{q}(t)\right)$, or $\left(q(t), p(t)\right)$.

A fundamental question therefore arises: Is simultaneous *accurate* measurement of both members of the pair $\left(q(t), p(t)\right)$ possible at all, at any instant t? To address this question, we consider the measurement of position and momentum of an electron (Fig. 1.1). The experiment involves shining light of wavelength λ on the electron to observe it through the lens of a microscope. This apparatus is called as the *Heisenberg microscope*. An analysis of this experiment would lead us to the quantum theory, which supersedes classical physics. Our discussion of this experiment requires acquaintance with two brilliant advances that were made in the early days of quantum physics. Historical accounts of concurrent progress in quantum theory and atomic physics are as fascinating and intriguing as romantic. Exciting accounts of these advances are available [2, 3, 4]. We restrict ourselves only to two rather significant milestones used in Eq. 1.9a,b that would be employed in the analysis of the experiment with the Heisenberg microscope. The first is Planck's hypothesis, made in December 1899, about the corpuscular nature of energy in an electromagnetic field, succinctly stated as $E = \mathrm{h}\nu$, ν being the frequency of an electromagnetic wave. Planck described his hypothesis as an *act of desperation* and struggled hard for many years later to negate his own theory. Planck's postulate however gained robust support in Einstein's Nobel Prize–winning explanation (in 1905) of the 1887 experiments, conducted by Hertz and Lenard, in which they discovered the photoelectric effect. The corpuscular nature of the quantum of energy in the electromagnetic field was advanced by Einstein in this work. It was firmly established afterward, in 1924, in the *statistical* description of the electromagnetic field by S. N. Bose. The constant h is commonly named after Planck, but considering the necessary and decisive contributions made by Einstein and Bose that established its significance, it is more appropriately called as *Planck–Einstein–Bose* (PEB) constant. The second significant milestone we shall employ in Eq. 1.9a is *de Broglie's hypothesis* of *wave–particle dualism*. It associates a wave having a wavelength

$$\lambda = \frac{\mathrm{h}}{p} \tag{1.9a}$$

with every particle, p being the particle's momentum. Using these two landmark advances, we see that a photon of energy

$$E = h\nu = \frac{h}{\lambda}c = pc, \tag{1.9b}$$

carries a momentum p. This is well known from the observation of comet tails, which are always pointed *away* from sunlight. The momentum carried by an electromagnetic field is well described by the Poynting vector (Chapter 13 of Reference [1]). It is taken full advantage of in the laser cooling of atoms.

We will now use (a) the Planck–Einstein–Bose quantum of energy of a photon and (b) the de Broglie wave–particle duality hypothesis, to discuss the Heisenberg microscope experiment to determine the position of an electron. Our limited objective in deliberating on this experiment is to demonstrate that the assumption we make in classical physics that the state of a system is represented by a point in the position–momentum phase space cannot withstand scrutiny by an experiment. It would thus expose the inadequacy of classical mechanics and prepare us to look for an alternative theory, which turns out to be the quantum theory.

An electron (assumed to be initially at rest) would be seen through a microscope's lens (Fig. 1.1). This requires a photon to be scattered by the electron along OB, after maximal momentum transfer to the electron, or along OF, after minimal momentum transfer to the electron, or of course at some intermediate angle determined by the conservation of momentum between the photon and the electron.

For minimum momentum transfer, we have

$$p = \frac{h}{\lambda} = \frac{h}{\lambda'}\sin\theta + m_e v_x', \tag{1.10a}$$

and for maximum momentum transfer, we have

$$p = \frac{h}{\lambda} = \frac{h}{\lambda''}(-\sin\theta) + m_e v_x''. \tag{1.10b}$$

In Eq. 1.10a and Eq. 1.10b m_e stands for the electron's mass and v_x' and v_x'' are its x-components of the velocity.

Accordingly, the least momentum gained by the scattered electron is

$$p_{\text{least}} = \frac{h}{\lambda} - \frac{h}{\lambda'}\sin\theta = m_e v_x', \tag{1.11a}$$

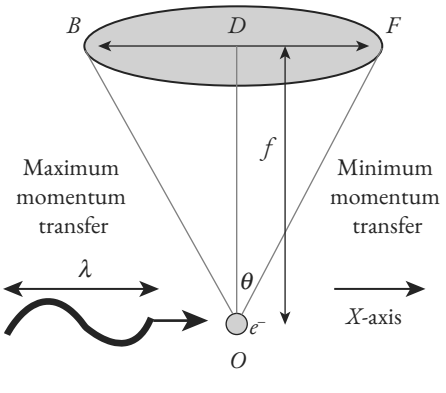

Fig. 1.1 The *Heisenberg microscope* is an apparatus we employ to discuss a thought experiment that demonstrates that accuracy in the measurement of an object's position is inversely related to that in a concurrent measurement of its momentum.

and the most momentum gained by it is

$$p_{\text{most}} = \frac{\text{h}}{\lambda} + \frac{\text{h}}{\lambda''}\sin\theta = m_e v''_x. \tag{1.11b}$$

The electron gains momentum from the probing photon, and the change in momentum along the x-axis is therefore

$$\Delta p_{\text{electron}} = \frac{\text{h}}{\lambda} \mp \frac{\text{h}}{\lambda}\sin\theta \approx \frac{\text{h}}{\lambda}(1 \mp \theta). \tag{1.12a}$$

This range would be minimal for small angles, and hence the momentum uncertainty is

$$\Delta p \approx \frac{2\text{h}}{\lambda}\theta. \tag{1.12b}$$

In order to minimize uncertainty in momentum, λ must therefore be large. Any attempt to minimize this uncertainty therefore comes at a price, since it would require a photon having a large wavelength, but a photon with larger wavelength would result in a greater uncertainty in position. Now, the resolving power (RP) of a microscope is

$$RP = \frac{1}{D} \simeq \frac{\sin\theta}{\lambda}, \tag{1.13a}$$

$$\text{or } \sin\theta \approx \frac{\lambda}{D}, \tag{1.13b}$$

$$\text{i.e., } \theta \approx \frac{\lambda}{D}. \tag{1.13c}$$

The location of the electron can at best be placed between $-f\tan\theta$ and $+f\tan\theta$, where

$$\tan\theta = \frac{\frac{1}{2}D}{f}, \tag{1.14}$$

f being the focal length of the microscope's lens (Fig. 1.1).

$$\text{Hence, } \frac{f}{D} = \frac{1}{2\tan\theta} \approx \frac{1}{2\theta}. \tag{1.15}$$

The uncertainty in locating the position of the electron therefore is

$$\Delta x \approx 2f\tan\theta \approx 2f\theta \approx 2f\frac{\lambda}{D} \simeq \frac{\lambda}{\theta}. \tag{1.16}$$

From Eqs. 1.12 and 1.16, we see that the product of the uncertainty in position with that in momentum is

$$\Delta p \Delta x \approx \frac{2\text{h}}{\lambda}\theta \times \frac{\lambda}{\theta} \simeq 2\text{h}, \tag{1.17}$$

which is of the order of the PEB constant. It is small, but not zero! The position uncertainty increases with the wavelength of the photon used, and that in momentum decreases! There is thus an inverse relationship between the two. This discussion on the Heisenberg microscope is mostly qualitative, but it brings out an essential limitation of classical physics. It exposes its intrinsic inconsistency, since the formalisms of Newton's, Lagrange's, and Hamilton's mechanics *all* depend on the *simultaneous* knowledge of position and momentum.

Physicists ran also into many other limitations of classical physics. These included the spectral intensity distribution of the blackbody radiation that was alluded to in Planck's hypothesis, analysis of atomic spectra, specific heat of solids, and so on. Readers should consult other sources for details on these topics. The way forward would be as romantic as challenging. A new formalism had to be developed which must admit the impossibility of simultaneous accurate measurement of position and momentum of a particle. These physical properties could not be represented merely by two numbers; a number (having dimension L) for the position of a particle, and another (having dimension MLT^{-1}) for momentum. Instead, it had to be stipulated that each particle would be represented by a wave–particle duality. The new theory that was developed using this scheme was called *wave mechanics*. It was developed in the 1920s by de Broglie and Schrödinger. An alternative theory that provided a viable replacement of classical physics was called *matrix mechanics*. It was developed by Heisenberg, Born, and Jordan around the same time. In matrix mechanics, position and momentum were represented by non-commuting matrices, and not merely by single numbers. Wave mechanics and matrix mechanics were raised on revolutionary ideas that would bring about an upheaval which impacted not just physics but nearly every single aspect of human life.

1.2 Hilbert Space Description of a Physical System

The experiment in Fig. 1.1 requires us to concede that representing the mechanical state of a system by a point in the phase space by its position and momentum is intrinsically flawed. This scheme must be replaced by a new theory. How about representing the state of a system by a vector in the Hilbert space? One may well ask why one should represent the state of a system by a vector. Well, what if it turns out to be useful in determining physical properties of the system? Quantum physics is the theory that answers this question magnificently! Quantum physics is the only theory that not only accounts for physical properties, but all of its predictions have come out to be correct! From the properties of fundamental particles to those of atoms, molecules, and condensed matter, it explains everything; nothing else does! Quantum theory may shock and confuse us, but it is a successful theory of the physical world. It is cast in a mathematical framework that must be learned with patience and rigor. At some point on this journey, we figure out *how* it works, if not *why* it works. Questions about the laws of nature reduce to *what the laws are* and *how they account for natural phenomena*, and not *why are the laws what they are*. The success of quantum theory is a compelling reason to admit that the theory is good, *extremely good*. It is a mathematical framework, and what makes it immensely appropriate to describe the laws of nature is "a wonderful gift that we neither understand nor deserve," as Wigner would say [5]. We therefore proceed to learn the mathematical formulation of quantum physics and hope that after putting in sufficient effort, we will discover how physical properties of a system can be extracted from this theory. More than a century of scintillating success of quantum theory is a strong argument to expect this hope to be very well placed.

Funquest: If, after patiently and thoroughly studying the mathematical framework described in the next few chapters, you discover that quantum theory accounts for most of the observable properties of the physical universe, what arguments would you have *against* the ansatz that the physical state of a system is described by a vector in the Hilbert space?

Intellectual giants like Bohr, Einstein, Feynman, and many others have all conceded, in telling ways, that quantum theory is shocking and mindboggling. Feynman went as far as saying that "nobody understands quantum mechanics." Quantum theory's success, and our quest to determine the laws of nature, is however a sufficient reason to explore, and hopefully master, its ansatz and methods. At a fundamental level, the ansatz of quantum theory has two components. That the state of the system is represented by a vector is the first of these. The other is how physical dynamical variables, such as position q, momentum p, angular momentum $\vec{j}$, energy E, etc., are represented. In quantum theory, these are represented by the operators $q_{op}, p_{op}, \vec{j}_{op}$, and the Hamiltonian H_{op}, respectively. These operators operate on the vectors (operands) in the Hilbert space. However, most often, the subscript "op" on the operators is omitted; the context would tell us if the symbol we are referring to corresponds to their quantum connotations. There, of course, are many physical dynamical variables of interest, over and above the four mentioned. As in classical mechanics, we would be interested in the results of measurements of dynamical variables, each of which would be represented by an appropriate operator.

We must interpret the result of a measurement in terms of the *operator algebra* that we have barely begun to develop. Nowhere have we referred to the size or mass of the physical system under investigation. We must therefore expect the new formalism to be applicable to all objects, whether microscopic or large. Quantum theory is based on a set of ansatz, a set of mathematical rules, which initially appear to be abstract and unrelated to the questions about physical properties of matter and energy. Yet it would turn out that it is just these mathematical structures that provide the *best* description of the laws of nature, albeit only *after* the mathematical framework is developed further.

We therefore begin with an abstract formalism of the quantum theory and rely on your motivation to patiently learn the rules of the game, and begin to have fun playing it. You will discover as you go along how beautifully the quantum theory connects to physical measurements and to *tangible* properties of the physical universe. The *Hilbert* space is the mathematical space in which the physical state vectors reside. It was named by John von Neumann after his Guru, David Hilbert. A vector that represents the physical state of a system at time t, if its state is known at an initial time t_0, is denoted as $\left|\gamma, t_0 : t\right\rangle$. The label γ must appropriately *designate* the quantum state. It is actually a *set* of one or more physical properties of the system that are *measurable* by an observer and that characterize the physical properties of the system. Temporal evolution of the system is then described by its derivative with respect to time, i.e., by $\frac{\partial}{\partial t}\left|\gamma, t_0 : t\right\rangle$. In Section 1.4, we shall see that it is given by the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t}\left|\gamma, t_0 : t\right\rangle = H\left|\gamma, t_0 : t\right\rangle, \tag{1.18}$$

where H is the Hamiltonian operator, also to be introduced in Section 1.4. The symbol $|\ \rangle$ represents a vector, called a "ket." We shall soon define a mathematical object that will be represented by $\langle\ |\ \rangle$, which looks like a *bracket*. Hence the terminology *ket* for $|\ \rangle$ and *bra* for $\langle\ |$. This terminology was introduced by Dirac, and the notation we have used is named

after him. $\langle\ |$ is also a vector in a space that is called Dual Conjugate (D.C.) Hilbert vector space, defined through a one-to-one correspondence with the ket Hilbert space:

$$\forall\ |\ \rangle \overset{\substack{\text{one}\\ \text{to}\\ \text{one}}}{\underset{\text{D.C.}}{\leftrightarrow}} \langle\ |. \tag{1.19}$$

The space of ket vectors $|\ \rangle$ is referred to as the *direct* space and that of bra vectors $\langle\ |$ as the *dual conjugate* (D.C.) space. The set of measurable properties represented by γ cannot include both members of the pair (q,p); in Section 1.1 we have already found that position and momentum are not simultaneously measurable. The label γ can therefore only be a set of physical properties whose measurements are *compatible* with each other. A physicist is interested in whatever *most* that can be learned about the target of her/his curiosity. Hence, we are interested in determining the ***C**omplete **S**et of **C**ompatible **O**bservables* that would constitute the set γ. Dirac abbreviated this as CSCO, which would also stand for ***C**omplete **S**et of **C**ommuting **O**perators*. By the end of this chapter, this dual meaning of CSCO would become clear.

Vectors in the Hilbert space can be added to each other, weighted by *complex* numbers α_i to get new vectors,

$$\left|x_{\text{resultant}}\right\rangle = \alpha_1\left|x_1\right\rangle + \alpha_2\left|x_2\right\rangle. \tag{1.20a}$$

The corresponding, one-to-one, relation in the D.C. space is

$$\left\langle x_{\text{resultant}}\right| = \alpha_1^*\left\langle x_1\right| + \alpha_2^*\left\langle x_2\right|. \tag{1.20b}$$

The asterisk employed here denotes complex conjugation. We shall work with an N-dimensional linear vector space, spanned by N linearly independent base vectors $\left\{\left|x_1\right\rangle,\left|x_2\right\rangle,..,\left|x_N\right\rangle\right\}$. This basis is identified by the criterion that for *complex* numbers $\alpha_i, i = 1,2,..,N$,

$$\alpha_1\left|x_1\right\rangle + \alpha_2\left|x_2\right\rangle + .. + \alpha_N\left|x_N\right\rangle = |0\rangle, \tag{1.21a}$$

if, and only if, *each* $\alpha_i = 0$.

Since the correspondence between the direct space and the D.C. space is one-to-one, one must expect a linearly independent basis $\left\{\left\langle x_1\right|,\left\langle x_2\right|,..,\left\langle x_N\right|\right\}$ for the D.C. space such that

$$\alpha_1^*\left\langle x_1\right| + \alpha_2^*\left\langle x_2\right| + .. + \alpha_N^*\left\langle x_N\right| = 0, \tag{1.21b}$$

if and only if *each* $\alpha_i^* = 0$.

Observe that complex conjugate coefficients are employed in Eq. 1.20b. An arbitrary vector in the Hilbert space can always be written as a linear superposition of a complete set of base vectors. We now define products of vectors in the direct and the D.C. spaces. Three types of products can be defined:

i. an inner product, represented by a bra-ket $\langle\ |\ \rangle$, which is a scalar,
ii. an outer product, represented by a ket-bra $|\ \rangle\langle\ |$, which is an operator, and
iii. a direct product, also called as tensor product, $|\ \rangle|\ \rangle$, of two ket vectors, which belong to two *disjoint* vector spaces. We shall use these in Chapter 4.

For the time being, we shall mostly work with the scalar (i.e., "inner") product and with the outer product. Without even explicitly stating it, we have already employed *addition* of vectors in the Hilbert space. For now we will work with a finite dimensional Hilbert space, i.e., one for which N is finite. When required, especially when we work with continuum/scattering states, we shall introduce techniques to extend the mathematical machinery to address an infinite dimensional Hilbert space.

The Hilbert space is a *metric* space. An example of a metric space is the familiar n-dimensional Euclidean space, $R^{(n)}$. It is an ordered set of n-tuples $\left(x_1, x_2, x_3, ..., x_n\right)$ such that a *measure* of distance between two points in the space can be defined by employing a suitable criterion. In $R^{(3)}$, the distance between two Cartesian points $\left(x_1, x_2, x_3\right)$ and $\left(x_1', x_2', x_3'\right)$ is

$$d = d\left\{\vec{r}, \vec{r}'\right\} = d\left\{\left(x_1, x_2, x_3\right), \left(x_1', x_2', x_3'\right)\right\} = \sqrt{\left\{\sum_{i=1}^{3}\left(x_i - x_i'\right)^2\right\}}, \tag{1.22a}$$

$$\text{i.e., } d = \sqrt{\left(\vec{r} - \vec{r}'\right)\bullet\left(\vec{r} - \vec{r}'\right)}, \tag{1.22b}$$

which is, of course, the Pythagoras theorem. The notion of such a measure is readily extended to vector spaces of higher dimensions, such as the non-Euclidean *flat* space–time continuum of the Special Theory of Relativity, and even the *curved* space–time of the General Theory of Relativity (Chapters 2, 13, 14 of Reference [1]). The function d that can be appropriately defined is called a "metric" on the vector space. We see that the metric is defined with respect to the scalar product. The generalization of this idea to an n-dimensional linear Hilbert space leads us to the *inner* product of two vectors, $|a\rangle$ and $|b\rangle$, denoted by $\langle b|a\rangle$. With reference to any three vectors $|a\rangle, |b\rangle, |c\rangle$ and arbitrary complex numbers λ, μ, κ, the inner product has the following properties:

i. $\langle b|a\rangle = \langle a|b\rangle^*$, (1.23a)

ii. $\langle a|(\lambda b + \mu c)\rangle = \langle a|(\lambda b)\rangle + \langle a|(\mu c)\rangle = \lambda\langle a|b\rangle + \mu\langle a|c\rangle$, (1.23b)

and

iii. $\langle(\lambda a + \mu b)|\kappa c\rangle = \langle(\lambda a)|\kappa c\rangle + \langle(\mu b)|\kappa c\rangle = \lambda^* \kappa\langle a|c\rangle + \mu^* \kappa\langle b|c\rangle$. (1.23c)

The inner product of a ket vector with itself, i.e., with its dual conjugate bra vector, is $\langle a|a\rangle$, and is defined to be essentially a non-negative quantity, i.e.,

$$\langle a|a\rangle \geq 0. \tag{1.24}$$

Equation 1.24 makes it possible to define a *real norm* of a vector as

$$\|a\| = \sqrt{\langle a|a\rangle}. \tag{1.25}$$

A vector whose norm is equal to unity is said to be *normalized*. Two vectors $|a\rangle, |b\rangle$ are said to be *orthogonal* if their inner product is zero. There is a significant parallel between the algebra of the linear vector Hilbert space and that of $R^{(3)}$. The difference however is not only in its possibly different, finite or infinite, dimensionality but also in the use of complex conjugation employed above.

An arbitrary vector in an N-dimensional Hilbert space is expressible as a linear combination of a complete set of *linearly independent* base vectors $\left\{|e_1\rangle, |e_2\rangle, ..., |e_N\rangle\right\}$, not necessarily