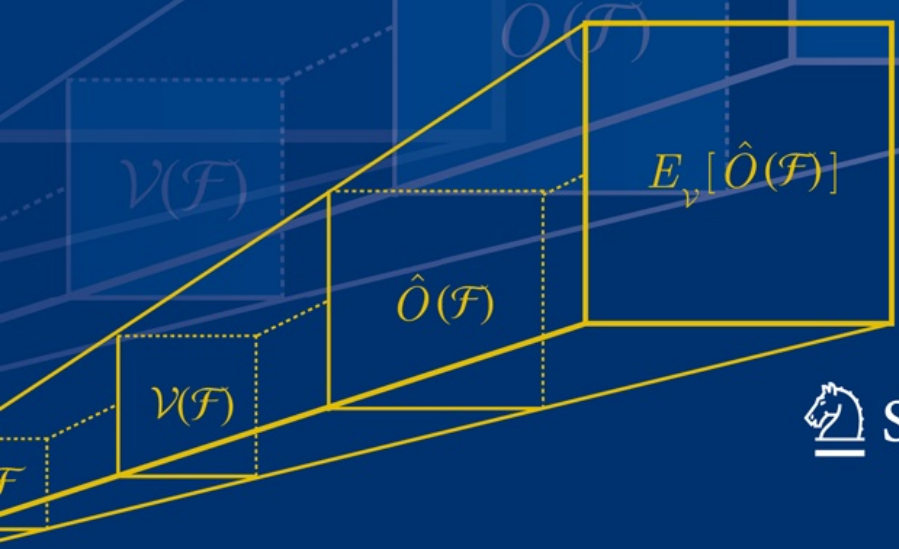
Belal E. Baaquie

The Theoretical Foundations of Quantum Mechanics



Springer

The Theoretical Foundations of Quantum Mechanics

Belal E. Baaquie

The Theoretical Foundations of Quantum Mechanics

Springer

Belal E. Baaquie
Department of Physics
National University of Singapore
Singapore

ISBN 978-1-4614-6223-1 ISBN 978-1-4614-6224-8 (eBook)
DOI 10.1007/978-1-4614-6224-8
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012954422

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Quantum theory introduces a fundamentally new framework for thinking about Nature and entails a radical break with the paradigm of classical physics. In spite of the fact that the *shift of paradigm* from classical to quantum mechanics has been going on for more than a century, a conceptual grasp of quantum mechanics has till today proved elusive. According to leading quantum theorist Richard Feynman, "It is safe to say that no one understands quantum mechanics" [13].

The foundations of quantum mechanics have been studied by many authors, and most of their books have been written for specialists working on the foundations of quantum mechanics and quantum measurement [1, 4, 16]—requiring an advanced knowledge of mathematics and of quantum mechanics [23, 25, 36]. An exception is the book by Isham [19], which is very clearly written and discusses the principles of quantum mechanics for a wider audience.

Given the ubiquitous presence of quantum mechanics in almost all branches of science and of engineering, there is a need for a book on the enigmatic workings of quantum mechanics to be accessible to a wider audience.

This book on the foundations of quantum mechanics is for the nonspecialists and written at a level accessible to undergraduates, both from science and engineering, who have taken an introductory course on quantum mechanics.

The mathematical formalism has been kept to a minimum and requires only a familiarity with calculus and linear algebra. The emphasis in all the topics is on analyzing the concepts and ideas that are expressed in the symbols of quantum mechanics. Linear vector spaces and operators form the mathematical bedrock of quantum mechanics, and a few derivations have been done to clarify these structures.

In this book the Schrödinger equation is never solved; instead, the focus is on the paradoxes and theoretical conundrums of quantum mechanics as well as on the conceptual basis required for addressing these. In particular, this book concentrates on issues such as the inherent (quantum) indeterminateness of Nature and the essential role of quantum measurement in defining a consistent interpretation of quantum mechanics.

The unusual properties of many widely used technologies are due to quantum phenomena. Indeed, most of what goes under the name of high technology is a direct

result of the workings of quantum mechanics, and many modern conveniences that we take for granted today would be impossible without it.[1]

Although quantum mechanics has qualitatively changed our view of Nature, a satisfactory understanding of it is still far from complete, and one can be sure there are a lot of surprises still awaiting us in the future.

The main focus of this book is to address the reasons why quantum mechanics is so enigmatic and extraordinary.

A theoretical framework for quantum mechanics is proposed in an attempt to clarify the underpinnings of quantum mechanics, namely the *transempirical quantum principle*, which states the following: A physical entity has *two forms* of existence, an indeterminate transempirical form when it is *not observed* and a determinate empirical form when it is *observed*. The transempirical and empirical forms have completely different behavior. The empirical form is intuitive and is the (experimentally) observed determinate state of the entity, whereas the indeterminateness of the transempirical form of the entity leads to all the paradoxes of quantum mechanics.

[1]For example, electronic devices, from computers, television, to mobile phones, are all based on semiconductors, and airplanes, ships, and cars all use semiconductors in an essential manner. More complex technologies such as superconductors, scanning electron microscope, magnetic resonance imaging (MRI), and lasers; fabrication of new drugs; modern materials science; and the study of nanoscale phenomenon all draw upon quantum mechanics.

Acknowledgments

I would like to acknowledge and express my heartfelt thanks to many outstanding teachers who inspired me to study quantum mechanics and marvel at its mysteries.

As an undergraduate, my formative views on quantum mechanics were greatly influenced by Khodadad Khan, A.K. Rafiqullah, George Zweig, Gerald "Gerry" Neugebauer, Clifford M. Will, and Jeffrey E. Mandula and by The Feynman Lectures on Physics [24]. As a graduate student, I was a tutor for a course taught by Kurt Gottfried and learned of his views on quantum mechanics; his book on the subject [15] continues to be, in my view, one of the best.

I had the good fortune of conversing with Richard P. Feynman on many occasions, and at times I had the pleasure of even debating with him. His profound observations still ring in my ears.

I had the privilege of doing my Ph.D. thesis under the guidance of Kenneth G. Wilson; his visionary conception of quantum mechanics and of quantum field theory greatly enlightened and inspired me and continues to do so till today.

I thank Kenneth Hong, Thomas Osiopowicz, Setiawan, Pan Tang, Duxin, Kuldip Singh, Rafi Rashid, Oh Choo Hiap, N.D. Hari Das and Cao Yang for helpful discussions. I want to specially thank Dagomir Kaszlikowski and Ravishankar Ramanathan for generously sharing their valuable insights on quantum mechanics.

I owe a special vote of thanks to Frederick H. Willeboordse for a careful reading of the manuscript that clarified many concepts and helped me to make a more coherent presentation of the subtleties of quantum mechanics. I am particularly indebted to Zahur Ahmed for his advice on the book and for his invaluable observations on its draft.

Contents

Synopsis

1

The epoch-making idea of the quantum as a fundamental property of Nature was introduced by Max Planck in 1900. Quantum mechanics is undoubtedly one of the most important and experimentally accurate scientific theory in the history of science.[1] Its range of applications and mathematical depth are unmatched, and quantum mechanics continues to yield novel and unexpected results—in technology as well as in all scientific fields, including physics and mathematics. Paradoxically enough, in spite of all its empirical and mathematical success quantum mechanics—due to its strange and enigmatic conceptual framework—has, until now, defied all attempts to reach a satisfactory understanding of its inner workings.

The human being's five physical senses are based on natural processes that can perceive only a finite range of physical phenomena. In the case of electromagnetic radiation, only a tiny and limited range of its wavelengths are visible to the human eye, with radiation of much longer and much shorter wavelengths being invisible. Since the smallest allowed quantum of energy for light (and for atoms) is truly minuscule when compared to the energies we encounter in daily life, there are only a few physical process, most of them being man-made, where one can directly observe quantum phenomena using one's five senses. When we extend our five senses with experimental devices and instruments, we can probe more deeply into Nature's secrets, and the quantum aspect of Nature becomes more apparent.

Classical mechanics works very well for the kind of objects one encounters in daily life that are moving much slower than the velocity of light. Once objects start to move very fast, we need to modify Newton's equations to Einstein's relativistic equations. On the other hand, for objects that are very small, such as electrons and atoms, quantum mechanics becomes necessary. If one attempts to extend Newton's laws to domains that are far from daily experience, they start to fail and give incorrect results.

[1] Accuracy is defined by the degree to which a theoretical value is close to the measured experimental value. Precision, in contrast, defines the degree to which an experiment, when it is repeated, produces a series of measured values to within a level of precision, namely, to within a certain error.

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8__1, © Springer Science+Business Media New York 2013

One can never expect an understanding of quantum mechanics that is similar in clarity and intelligibility to the one provided by classical mechanics since the connections of the symbols of classical mechanics to the phenomena that it represents are directly based on the perception of Nature by our five sense; "you get what you see." In the case of quantum mechanics, as will become clear as one reads this book, the connections of the symbols of quantum mechanics with observed quantities are more nuanced and opaque than for classical mechanics. One can nevertheless hope that, over time, quantum mechanics will become as intuitively obvious and transparent for future generations as is classical mechanics for the older generations.

Noteworthy 1.1: The experimental accuracy of quantum mechanics

Quantum mechanics is clearly more accurate than classical mechanics, which it supersedes in every way. The special theory of relativity, which describes the structure of empty spacetime, has so far has proven to be experimentally as accurate as quantum mechanics. Einstein's theory of gravity, namely the theory of general relativity is outside the domain of quantum mechanics and we compare their empirical accuracy.

To date, the most accurate test of general relativity is the prediction that a clock slows down by a factor of $1 + \mathrm{U}/\mathrm{c}^2$, where U is the gravitational potential. This prediction has been tested, using the quantum interference of atoms, to an accuracy of 7×10^{-9}, about one part in a hundred million [29].

The experimental value of the electron's magnetic moment is $g\mu_\mathrm{B}$, where $\mu_\mathrm{B} = eh/4\pi m$ is the Bohr magneton and g-factor is the dimensionless constant for electron's magnetic moment; e and m are the charge and mass of the electron and h is Planck's constant. The naive value of $g = 2$, which is given by the Dirac equation for the electron, is corrected by the effects of the electron's interaction with the photon. The most accurate experimental prediction of quantum mechanics is that $g = 2.00231930419922(1 \pm 0.7491312684 \times 10^{-12})$ [7].

This prediction of quantum mechanics—or more accurately of quantum field theory, a formulation of quantum mechanics that incorporates special relativity (and hence the accuracy of special relativity is also being tested)—completely agrees with the experimental result to an accuracy of 10^{-12}, one part in a trillion. As of now, the experimental verification of quantum mechanics is more accurate than general relativity by a factor of more than a thousand, namely, 10^3. This does not mean that Einstein's theory of gravity is not exact, which it may or may not be, but rather that its proven experimental accuracy is less than that of quantum mechanics.

Building on the pioneering work of Max Planck and Niels Bohr, the modern formulation of quantum mechanics rests primarily on the ideas of Max Born, Erwin Schrödinger, Werner Heisenberg, and Paul A.M. Dirac.

In 1926, a quarter century after the Planck's epoch-making quantum hypothesis, Erwin Schrödinger discovered the *dynamical equation* of quantum mechanics. It is worth noting that, for almost a century, Schrödinger's equation has proved to be flawless, successfully facing numerous and precise experimental tests.

The Schrödinger equation replaces Newton's second law of motion and, according to many scientists, is one of the most important cornerstones of science.

In the period of 1926–1929, the founders of quantum mechanics laid down the complete mathematical foundations of quantum mechanics, which continues to hold till today [10].

Einstein's theory of special and general relativity is a logical development of classical physics; relativity reinterprets the meaning of classical concepts such as time, position, mass, velocity, and acceleration. In contrast, quantum mechanics introduces completely new ideas such as indeterminacy and uncertainty, the state vector, operators, path integration, as well as the quantum theory of measurement—concepts that are absent and incomprehensible in the framework of classical physics.

Many books on quantum mechanics follow the historical path by recounting the motivations and reasons that led to the idea of the quantum [15]. A century after the advent of the idea of the quantum, an approach based on the inner logic of quantum mechanics can be now taken.

Most undergraduate textbooks concentrate on the mathematical techniques required for solving the partial differential Schrödinger equation—with questions of interpretation and consistency usually touched upon only in passing. In contrast, this book does not provide any solutions of the Schrödinger equation and, instead, is primarily focused on those fundamental principles and theoretical aspects of quantum mechanics that impinge on its internal workings and clarify its mathematical structure.

In an effort to understand the inner workings of quantum mechanics, the concept of the *trans-empirical quantum principle* is postulated as being inherent in Nature.[2] Using the paradigm of the trans-empirical quantum principle, the book attempts to clarify the world of the quantum by reinterpreting the foundation of quantum mechanics.

The book is organized as follows.

Chapter 2 is a summary of the main ideas of the book. The notion of the quantum entity is reasoned to be inherently and intrinsically *indeterminate* and shown to consist of a pair: the indeterminate quantum mechanical degree of freedom that is the foundation of the quantum entity and the state vector that provides a quantitative description of the quantum entity. Five cardinal principles of quantum mechanics are identified as necessarily arising from the structure of the quantum entity.

Chapter 3 discusses what is *real* and what *exists*, two words that are used synonymously in classical physics but, with appropriate refinements, are shown to be words that have vastly different meanings in quantum mechanics. In order to have a conceptually transparent framework of quantum mechanics, the empirical domain of classical physics is extended to include a new domain termed as the *trans-empirical domain*.

The quantum mechanical degree of freedom is shown to be completely trans-empirical, whereas the state vector straddles two domains—existing in the

[2]The "trans-empirical quantum principle" is stated in Sect. 3.9 and discussed in detail in Chap. 3.

trans-empirical domain when it *is not* experimentally *observed* and having an empirical manifestation when it *is observed*. It is shown that the time evolution of a quantum degree of freedom is via *trans-empirical paths* when the path taken is not experimentally ascertained.

The concept of the *trans-empirical quantum principle* is formulated to define the theoretical framework of quantum mechanics.

Chapter 4 discusses the mathematical framework for describing a quantum entity, namely, the structure and properties of the degree of freedom and the quantum state vector that describes it. The concept of a linear vector space is introduced, and the basic properties of a state vector are stated and analyzed.

In Chap. 5, the concept of a Hermitian operator representing physically observable properties of the quantum state is discussed in some detail. The main properties of operators are stated, and the important examples of a discrete and continuous degree of freedom are discussed.

Chapter 6 discusses the tensor product of vector spaces and operators. This provides the mathematical framework for studying the density matrix, including the pure, mixed, and reduced density matrix. The density matrix provides a criterion for understanding a special class of state vectors, the so-called entangled states.

Chapter 7 shows that the Bell inequality provides a quantitative criterion for differentiating quantum indeterminacy from classical randomness. The BKS theorem further generalizes the Bell inequality to include all quantum states. Quantum probability is defined based on Heisenberg's operator formulation of quantum mechanics.

In Chap. 8, the remarkable properties of quantum superposition are discussed. The Mach-Zehnder interferometer is employed to study the indeterminate paths of a photon and illustrates how quantum interference arises; it is shown that a quantum eraser can partially erase or restore quantum interference.

Chapter 9 discusses how the process of quantum measurement entails the preparation, amplification, entanglement, and collapse of the state vector. The density matrix provides a description of the quantum entity that is mathematically appropriate for describing the process of measurement.

In Chap. 10, the Stern-Gerlach experiment is discussed in detail to illustrate and exemplify the process of quantum measurement.

In Chap. 11 the Feynman path integral is derived by applying the trans-empirical quantum principle to indeterminate paths, and the Dirac-Feynman path integral formulation of quantum mechanics is briefly discussed. Path integral quantization is taken as the starting point of quantum mechanics and is shown to yield the Hamiltonian and its state space.

In Chap. 12 some conclusions are drawn.

The Quantum Entity and Quantum Mechanics 2

A thing, intuitively, seems to be an à priori form of Nature that is most directly experienced by our five senses. A physical entity that can be perceived by sensory perception appears to be one of the most irreducible and primitive notions that underpins our cognition of Nature. A good place to start exploring the quantum realm is to understand the difference in how classical and quantum mechanics conceptualize an *entity*, a *thing*, and an *object*.

The concept of the thing in quantum mechanics soon leads us to a theoretical framework for describing and explaining Nature that goes against our everyday intuition that is based, as it is, on our daily experience.

Buried deep inside the mathematical structure of quantum mechanics are unresolved paradoxes, mysteries, and enigmatic views about Nature. One has to cut through a thick shell of formalism to encounter the theoretical underpinnings of quantum mechanics.

In this chapter, it is shown that the concept of a quantum entity necessarily leads to the cardinal principles of quantum mechanics. The cardinal principles, in turn, will lead us to introduce various theoretical constructs that are necessary for discussing the principles and paradoxes of quantum mechanics.

2.1 What Is a Classical Entity?

The concept of an *object* in classical physics is founded on the idea of an *objective reality*, namely, that a material entity has an *intrinsic* reality and its properties (qualities) are inherent in the entity itself. All the interconnections of classical objects to each other are founded on, and derived from, the objective reality of the classical entity. Since the classical entity exists objectively, its properties do not depend on anything external and, in particular, do not depend on whether it is being observed (measured) or not.

With the development of Maxwell's equation and Einstein's theory of gravity, the classical concept of a physical entity was extended to include the *classical field*. A classical field, such as the electromagnetic field, is a physical entity that is spread

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8__2, © Springer Science+Business Media New York 2013

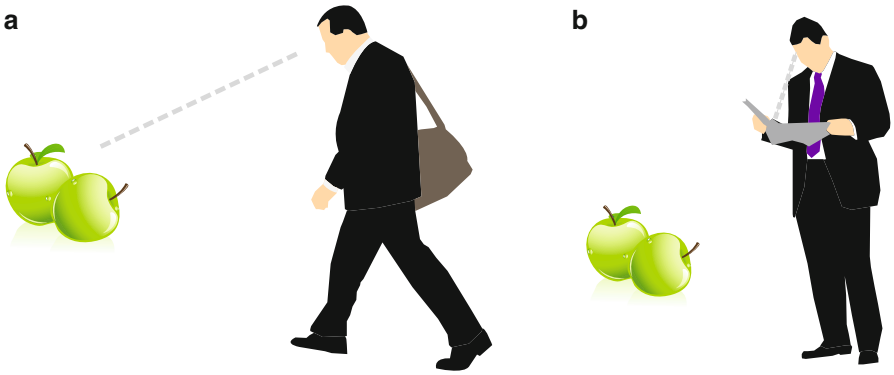


Fig. 2.1 (**a**) A person is looking at the apples. (**b**) The person is not looking at the apples. The classical view is that the existence of the apple is an objective reality independent of the observer (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

over space and propagates in time; the classical field, like a material thing, exists objectively and has intrinsic properties such as having energy and momentum at each spacetime point that it occupies.

According to *classical* physics, an entity is *completely determinate* and exists in an *exactly* defined state; for example, a classical particle has an intrinsic and exact position in space. When it is observed, the classical entity *is what it appears to be*; hence, the classical entity *is completely empirical*, with an observation, in principle, fully and completely describing its state. The Oxford dictionary defines *empirical* as being based on, concerned with, or verifiable by observation or experience rather than theory or pure logic.

We conclude that a classical entity exists objectively and is a determinate quantity.

Consider a person *looking* and *not looking* at the apples, as in Figs. 2.1a,b, respectively. What is the state of the apples for the two cases? Since, in classical physics the apples exist objectively, it follows that even when the person looks away, the apples continue to be in the *same state* in both cases, as in Figs. 2.1a,b.

However, note that if the person is not looking at the apples, then there is *no experimental basis* to claim the apples continue to be in the same state as when the person looked at it. The claim of classical physics that the world exists as an objective reality is an *assumption*.

Dynamics of a Classical Entity

The dynamics (motion, time evolution) of classical dynamical variables (position and velocity for a particle) are determined by Newton's second law of motion; in the modern formulation, it is given by the variation of the system's action S. The

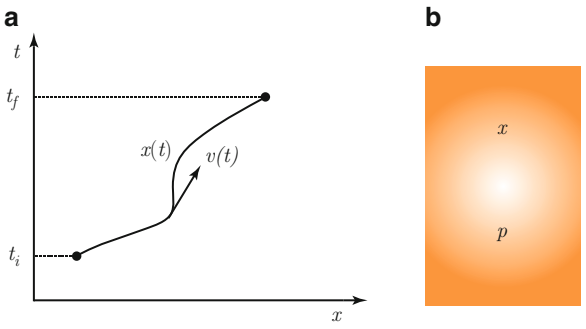


Fig. 2.2 (**a**) The classical trajectory $x(t), v(t) = \mathrm{d}x(t)/\mathrm{d}t$ of a classical particle. (**b**) To specify the classical state of a particle at each moment, its *dynamical variables* (position x and momentum p) have to be specified (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

action is the time integral of the Lagrangian $\mathcal{L}(x, \mathrm{d}x/\mathrm{d}t)$, which is a function of the kinetic and potential energy of a particle's trajectory.

Consider the particle's path for the finite time interval $[t_\mathrm{i}, t_\mathrm{f}]$, as shown in Fig. 2.2a and determined by Newton's law of motion. Note that at every instant t, the particle has a definite position x and momentum $p = mv$, where $v = \mathrm{d}x/\mathrm{d}t$ is its velocity and m its mass. The Lagrangian $\mathcal{L}(x, \mathrm{d}x/\mathrm{d}t)$ for the particle can be computed once the particle's trajectory is specified.

The action S for a particle is given by the following:

$$S = \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}t \mathcal{L}(x, \mathrm{d}x/\mathrm{d}t) \tag{2.1}$$

$$\Rightarrow \delta S = 0 : \text{Equation of motion} \tag{2.2}$$

with the initial and final positions being specified at t_i and t_f, respectively.

The equation of motion given in (2.2) means that if one takes any arbitrary trajectory and computes the value of S, then the numerical value of S will be a minimum (or maximum) only when the trajectory obeys Newton's law and hence will satisfy $\delta S = 0$. We conclude that (2.2) is equivalent to Newton's law.

The state of a classical entity is described by its *dynamical variables*; for the case of a particle, the dynamical variables are x and p, which are fixed for *each instant* of time t. Hence, the *classical state* of the particle is *completely determinate*, with its exact state given by specifying the dynamical variables x and p, as shown in Fig. 2.2b.

Classical Probability

If one views an apple as being composed of a large collection of atoms, then one is hard-pressed to claim that all the atoms that compose the apple continue to be in the apple; it is highly likely that while one was looking away, some atoms have detached themselves from the apple and other atoms from the environment have become attached to it.

The argument that a description of an apple has unavoidable approximations is true for any large classical system, such as the practical impossibility of knowing the precise position and momentum of all the gas atoms in a room. Recall precision was defined in Chap. 1 as the degree to which an experiment, when it is repeated, produces a series of measured values to within a certain error, termed as the level of precision. An approximation is specifying a quantity to a certain well-defined degree of precision.

Furthermore, the emergence of the discipline of (classical) chaos has shown that any practical measurement has only a finite accuracy and introduces the idea of classical randomness in the description of a nonlinear classical system.

Classical randomness, discussed in greater detail in Sect. 7.3, describes phenomenon that lacks predictability, that exhibits the property of chance, and is mathematically modeled by variables having an outcome determined by its probability distribution function.

Nevertheless, classical chaos theory does not change the *ontological* property of a classical system in that it exists in a determinate and intrinsically exact state.

> Ontology: from the Greek term for 'being'; that which 'is,' present participle of the verb 'be'; the term is used for the nature of being, of existence, or of reality.

The ambiguity and imprecision in the knowledge of a large (macroscopic) classical system or a chaotic process is entirely due to our *ignorance* about the exact state of the system and leads to *classical probability theory*.

The *ignorance* of the precise state of a classical object can be modeled and encoded by considering the classical system to be in a *random* state, namely, known to only certain level of *precision*. It is important to note that the intrinsic state of the system is exact; the randomness of classical probability theory is an approximate description of a system that intrinsically exists in an exact state.

The point to note is that an object exhibiting randomness of classical probability is a classical entity that has an *objective existence*, having a specific and precise value *before* it is observed; this point is of cardinal importance in the formulation of classical probability, discussed in Sect. 7.3. The practical inability of providing an *exact* description of a classical system leads one to the field of chaos and complexity; one introduces new concepts drawn from classical probability theory that need to supplement deterministic classical physics for providing a better description of a classical chaotic system.

In fact, it will be shown in Chap. 7 that quantum probability is fundamentally different from classical probability since the concept of quantum uncertainty is essentially different from the idea of classical randomness.

2.2 The Entity in Quantum Mechanics

Our discussion on the "entity" in quantum mechanics does not take the historical route but rather starts from the quantum conception of the entity and then goes on to

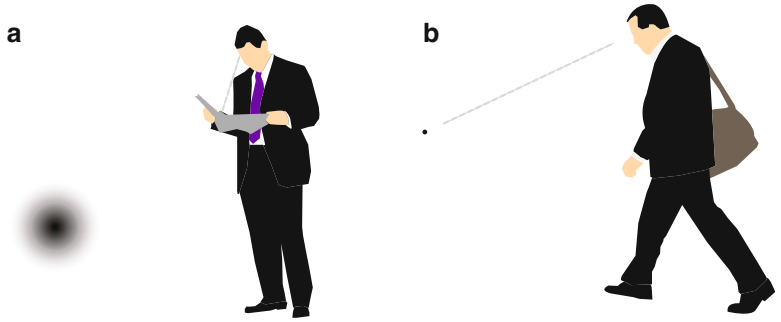


Fig. 2.3 (**a**) The person is *not observing* the atom: The atom is inherently indeterminate; it has the highest likelihood of being observed in the shaded volume. (**b**) A person, by *observing* the atom, puts the atom into a determinate state (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

explain why the mathematical formalism of quantum mechanics logically follows from the need of describing the quantum entity.

Consider a quantum particle in a box and subjected to repeated measurements.[1] When one experimentally measures the position of the particle, one *observes* that it has a definite position x_1; if one then repeats the identical experiment, one observes the particle at *another* position x_2; and a third measurement yields yet *another* position x_3, and so on. Every time one measures the quantum particle's position—prepared in exactly the same way—it is observed at a different position.

There are special quantum states discussed in Sect. 5.3, called eigenstates, with properties, such as energy and angular momentum, that have the same value every time such an eigenstate is observed. The position degree of freedom for particle in a box is not such a property.

When it is not observed, the quantum particle does not have any definite position, and, unlike a classical particle, its position does not have an objective existence. What is the form of existence of the particle when one *does not measure* the particle's position? If the entity is large, like a piece of stone, then the classical description in most cases is quite adequate: The observed and unobserved state appear to be the same. However, if the particle is small, like an electron or an atom, all our intuition regarding its behavior fails.

For concreteness, let the quantum entity be an *atom* located in space. When a man directly looks at the atom, as in Fig. 2.3b, he observes a point-like object, but when he does *not* observe the atom, quantum mechanics tells us that the atom no longer has a determinate position, but instead, the atom's position is *indeterminate*; the atom apparently "exists" at many positions simultaneously—with different likelihoods—and the region of greatest likelihood is represented by the shaded portion in Fig. 2.3a; the degree of shading indicates the different likelihood of the

[1]The concept of repeated quantum measurements is discussed in Sect. 9.3.

particle being observed at different points, which in principle can extend to all of space. The word "indeterminate" and the concept of quantum uncertainty are being used synonymously; the term indeterminate needs to be defined more precisely and is addressed in Sect. 3.2.

In other words, unlike a classical particle, a quantum particle *does not* have a precise position before it is observed. This rather unexpected and strange claim of quantum mechanics—that *there is a fundamental difference between the observed and unobserved state of a quantum entity*—is at the foundation of quantum mechanics. This strange claim of quantum mechanics has been shown to be consistent with many experiments designed to test these claims.

The central role of observation*, of* measurement*, is what differentiates the observed from the unobserved state and is the key to quantum mechanics.*

Heisenberg used the term *potentiality* for the indeterminate state of the quantum entity and the term *actuality* for the observed condition. Every act of observation results in the particle making a transition from its state of *potentiality* to one of its possible and *actual* determinate condition [18].

In summary, in quantum mechanics, the entity, the object has *two forms of existence*: when it is *observed*, it is definite and *determinate*, and when it is *not observed*, it is *indeterminate* and uncertain. Fundamental to the two forms of existence of a quantum entity is the *act of observation*, the process of measurement that connects the unobserved with the observed form.

2.3 Describing an Indeterminate Quantum Entity

The *classical* description of an entity starts with the dynamical variables describing the classical state of the entity and completes its description with the equations of motion for the dynamical variables.

Since the quantum entity is intrinsically indeterminate, the classical approach is inadequate. What is the route for describing an indeterminate quantum entity? For concreteness, consider the quantum entity to be a quantum particle. Quantum mechanical indeterminacy requires that the following interrelated issues be addressed:

- The first step for describing a quantum particle is the quantum generalization of the classical dynamical variables. In quantum mechanics, due to indeterminateness, a quantum particle no longer has a classical trajectory; this entails giving up all knowledge of the momentum if one measures the quantum particle's position.
- Since the quantum entity's position is indeterminate, the classical particle's dynamical variables x and p are superseded by the quantum *degree of freedom* $\mathcal{F}$. For a quantum particle moving in one space dimension, the degree of freedom space is given by the real line, namely, $\mathcal{F} = \Re = \{x | x \in [-\infty, +\infty]\}$, and hence, $\mathcal{F}$ *is an entire space*.
- The position of a quantum particle is an indeterminate degree of freedom when it is not being observed. To perform measurements on the particle's degree of freedom, one needs to introduce the concept of *operators* that act on the degree

of freedom, and are discussed in detail in Chap. 5. Suffice for the discussion here is that position projection operators, discussed in Sect. 9.2, can observe all the effects of the position degree of freedom.

Repeated observations by the quantum particle's position operators reveal the range of possible values that the particle's position degree of freedom can take—hence allowing us to theoretically enumerate *all the possible positions* of the particle; in effect, the results of the observations allow us to mathematically reconstruct the degree of freedom space $\mathcal{F}$.

- The quantum degree of freedom is a quantitative entity that numerically describes all the possible allowed values for the quantum entity and constitutes the space $\mathcal{F}$. The degree of freedom is a time-independent quantity, with the space $\mathcal{F}$ being invariant and unchanging over time.
- It is an experimental fact that, when the quantum particle is repeatedly observed using the different position projection operators, the operators acquire different average values, reflecting the properties of the quantum particle's degree of freedom. Repeated observations, besides allowing for the enumeration of the degree of freedom $\mathcal{F}$, also provide the *likelihood* of the particle being found at the different position projection operators.
- A major conceptual leap, following in the footsteps of Max Born, is to *postulate* that the result of repeated *experimental observations* of the state vector yields all the quantitative properties of a quantum entity.
- A quantitative *quantum probabilistic* description of the indeterminate quantum entity is provided by the quantum state vector $\psi(\mathcal{F})$.[2] The state vector is fundamentally statistical in nature, with every outcome being completely unpredictable. The state vector $\psi(\mathcal{F})$ is an element of the *state space* of the degree of freedom, denoted by $\mathcal{V}(\mathcal{F})$.
- The quantum state vector $\psi(\mathcal{F})$ is *postulated* to carry a *complete description* of the quantum entity and is a *superstructure* of the quantum degree of freedom $\mathcal{F}$. Among other things, the state vector $\psi(\mathcal{F})$ determines the likelihood of a particular experimental outcome for the operators observing the degree of freedom.[3] The quantum state can also be represented by the density matrix operator, which is more suitable for analyzing the process of quantum measurements, and is discussed in Chap. 6.
- The dynamics of a quantum entity is determined by the time evolution of its state vector $\psi(\mathcal{F})$, also written as $\psi(t, \mathcal{F})$ to explicitly indicate the dependence on the parameter of time t. The Schrödinger equation determines the dynamics of a quantum entity. It is a first-order partial differential equation in time and yields the time evolution of the state vector, namely, $\partial\psi(t, \mathcal{F})/\partial t$.
- There are two forms of existence of a quantum mechanical entity—the potential (unobserved) and the actual (observed)—that are connected by a process of measurement. Indeterminateness is potential and the determinateness is actual.

[2] Quantum probability is different from classical probability and is discussed in Chap. 7.

[3] The relation of the state to observed quantities is discussed in Sect. 2.4.

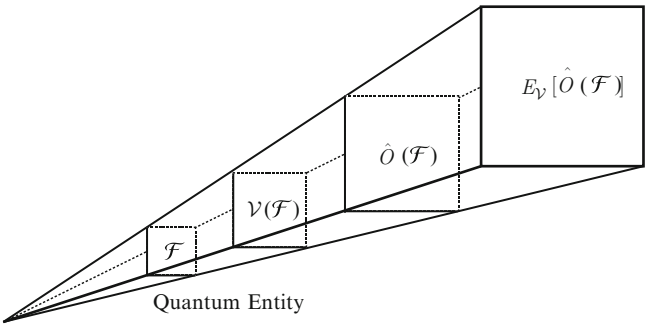


Fig. 2.4 The theoretical superstructure of quantum mechanics; the quantum entity is constituted by the degree of freedom $\mathcal{F}$ and its state vector that is an element of state space $\mathcal{V}(\mathcal{F})$; operators act $\mathcal{O}(\mathcal{F})$ act on the state vector to extract information about the degree of freedom and lead to the final result $E_{\mathcal{V}}[\mathcal{O}(\mathcal{F})]$; only the final result, which is furthest from the quantum entity, is empirically observed (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

According to Werner Heisenberg, all physical properties of the degree of freedom $\mathcal{F}$ are mathematically represented by *operators* $\mathcal{O}(\mathcal{F})$. The process of measurement is mathematically represented by applying operators on the quantum state $\psi(\mathcal{F})$ of the quantum entity.

- Repeating the process of measurement results in the experimental determination of the average, or *expectation value*, of the physical operators and is expressed as $E_{\psi}[\mathcal{O}(\mathcal{F})]$. All physical information about the degree of freedom $\mathcal{F}$ is encoded in the expectation value of operators.

In summary, a quantum entity and its empirically measured properties—shown in Fig. 2.4—are far more elaborately structured than a classical entity. The quantum entity itself, when it is not being observed, consists of its degree of freedom $\mathcal{F}$ together with the state vector $\psi(t, \mathcal{F})$ describing its observable properties.

The observed condition of a classical entity exhausts all its properties—"one sees what one gets." In the case of a quantum entity, there is an entire unobservable superstructure between the empirically observed properties of the quantum entity and the totality of the quantum entity. The hidden quantum superstructure, which is absent for a classical entity, needs an interpretation for making a connection of the quantum entity with its experimentally observed properties and that is provided by the Copenhagen quantum postulate.

2.4 The Copenhagen Quantum Postulate

The Copenhagen interpretation of quantum mechanics was pioneered by Niels Bohr and Werner Heisenberg and is the *standard interpretation* of quantum mechanics that is followed by the majority of practicing physicists—and is the one followed in this book [10, 34]. The discussion in Sects. 2.2 and 2.3 was essentially a summary

and explication of the fundamental tenets of the Copenhagen interpretation of quantum mechanics.

The Copenhagen interpretation is not universally accepted by the physics community, with many alternative explanations being proposed for understanding quantum mechanics, and which are summarized in Sect. 12.2. Instead of entering this debate, this book endeavors to clearly represent the theoretical assumptions that are implicit in the Copenhagen interpretation—assumptions that are generally quite opaque due to the mathematical formalism of quantum mechanics.

The Copenhagen interpretation can be summarized by the following postulate:

> The quantum entity consists of its degree of freedom $\mathcal{F}$ and its state vector $\psi(t,\mathcal{F})$. The foundation of the quantum entity is its degree of freedom, which takes a range of values and constitutes a space $\mathcal{F}$. The quantum degree of freedom is completely described by the quantum state $\psi(t,\mathcal{F})$, a complex-valued function of the degree of freedom that is an element of state space $\mathcal{V}(\mathcal{F})$.
> All physically observable quantities are obtained by applying Hermitian operators $\mathcal{O}(\mathcal{F})$ on the state $\psi(t,\mathcal{F})$.
> The quantum entity is an *inseparable pair*, namely, the degree of freedom and its state vector.
> Experimental observations collapse the quantum state and repeated observations yield $E_\psi[\mathcal{O}(\mathcal{F})]$, which is the expectation value of the operator $\mathcal{O}(\mathcal{F})$ for the state $\psi(t,\mathcal{F})$.
> The Schrödinger equation determines the time dependence of the state vector, namely of $\psi(t,\mathcal{F})$, but does not determine the process of measurement.

It needs to be emphasized that the state vector $\psi(t,\mathcal{F})$ provides only statistical information about the quantum entity; the result of any particular experiment is impossible to predict.

The organization of the theoretical superstructure of quantum mechanics is shown in Fig. 2.4.

The quantum state $\psi(t,\mathcal{F})$ is a complex number that describes the degree of freedom and is more fundamental than the observed probabilities, which are always real positive numbers. The scheme of assigning expectation values to operators, such as $E_\psi[\mathcal{O}(\mathcal{F})]$, leads to a generalization of classical probability to quantum probability and is discussed in detail in Chap. 7.

To give a concrete realization of the Copenhagen quantum postulate, consider a quantum particle moving in one dimension; the degree of freedom is the real line, namely, $\mathcal{F}=\Re=\{x|x\in(-\infty,+\infty)\}$ with state $\psi(t,\Re)$. Consider the position operator $\mathcal{O}(x)$;[4] a measurement projects the state to a point $x\in\Re$ and collapses the quantum state to yield

$$P(t,x)\equiv E_\psi[\mathcal{O}(x)]=|\psi(t,x)|^2\ ;\ P(t,x)>0\ ;\ \int_{-\infty}^{+\infty}\mathrm{d}xP(t,x)=1 \tag{2.3}$$

Note from (2.3) that $P(t,x)$ obeys all the requirements to be interpreted as a probability distribution. A complete description of a quantum system requires

[4]The position projection operator $\mathcal{O}(x)=|x\rangle\langle x|$ and is discussed in Sect. 9.2.

specifying the probability $P(t,x)$ for all the possible states of the quantum system. For a quantum particle in space, its possible quantum states are the different positions $x \in [-\infty, +\infty]$. The position of the quantum particle is indeterminate, and $P(t,x) = |\psi(t,x)|^2$ is the probability of the state vector collapsing at time t and at $\mathcal{O}(x)$—the projection operator for position x.

The moment that the state $\psi(t,\Re)$ is observed at *specific* projection operator $\mathcal{O}(x)$, the state $\psi(t,\Re)$ instantaneously becomes zero everywhere else. The transition from $\psi(t,\Re)$ to $|\psi(t,x)|^2$ is an expression of the collapse of the quantum state. It needs to be emphasized that no classical wave undergoes a collapse on being observed; the collapse of the state $\psi(t,\Re)$ is a purely quantum phenomenon.

The pioneers of quantum mechanics termed it as "wave mechanics" since the Newtonian description of the particle by its trajectory $x(t)$ was replaced by the state $\psi(t,\Re)$ that looked like a classical wave spread over all of space $\Re$. Hence, the term "wave function" was used for denoting $\psi(t,\Re)$.

The state $\psi(t,\mathcal{F})$ of a quantum particle is *not* a classical wave; rather, the only thing it has in common with a classical wave is that it is sometimes spread over space. However, there are quantum states that are *not spread over space*. For example, the up and down spin states of a quantum particle exist at a single point; such quantum states are described by a state that has no dependence on space and hence is *not* spread over space.

In the text, the terms *state, quantum state, state function*, or *state vector* are henceforth used for $\psi(t,\mathcal{F})$, as these are more precise terms than the term wave function.

The result given in (2.3) is an expression of the great discovery of quantum theory, namely, that *behind what is directly observed*—the outcome of experiments from which one can compute the probabilities $P(t,x) = |\psi(t,x)|^2$—there lies an *unobservable world of the probability amplitude* that is fully described by the quantum state $\psi(t,\mathcal{F})$.

2.5 Five Pillars of Quantum Mechanics

The description and dynamics of a quantum entity given in Sect. 2.3 can be summarized as follows. Quantum mechanics is built on five main conceptual pillars that are given below.

- The quantum degree of freedom space $\mathcal{F}$
- The quantum state vector $\psi(\mathcal{F})$
- Time evolution of $\psi(\mathcal{F})$ given by the Schrödinger equation
- Operators $\mathcal{O}(\mathcal{F})$
- The process of measurement, with repeated observations yielding the expectation value of the operators, namely, $E_{\mathcal{V}}[\mathcal{O}(\mathcal{F})]$

The five pillars of quantum mechanics are illustrated in Fig. 2.5. Each pillar of quantum mechanics is briefly summarized in the following sections.

Degree of freedom	State space	Dynamics	Operators	Observation
$\mathcal{F}$	$\mathcal{V}(\mathcal{F})$	$\frac{\partial \psi(t,\mathcal{F})}{\partial t}$	$\hat{O}(\mathcal{F})$	$E_{\mathcal{V}}[\hat{O}(\mathcal{F})]$

Fig. 2.5 The five cardinal pillars of quantum mechanics (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

2.6 Degree of Freedom Space $\mathcal{F}$

Recall in classical mechanics, a system is described by its dynamical variables, and its time dependence is given by Newton's equations of motion. In quantum mechanics, the description of a quantum entity starts with the generalization of the classical dynamical variables and, following Dirac [10], is called the quantum *degree of freedom*.

The degree of freedom is the root and ground on which the quantum entity is anchored. The degree of freedom embodies the *qualities* and *properties* of a quantum entity. A single quantum entity, such as the electron, can simultaneously have many degrees of freedom, such as spin, position, and angular momentum that all, taken together, describe the quantum entity. The symbol $\mathcal{F}$ is taken to represent all the degrees of freedom of a quantum entity.

The indeterminacy of the quantum degree of freedom is studied in detail in Chap. 7. A remarkable conclusion of the study—validated by experiments—is that a quantum degree of freedom *does not* have any precise value before it is observed; the degree of freedom is *inherently indeterminate*. One interpretation of the degree of freedom being intrinsically indeterminate is that it *simultaneously* has a *range of possible values*; the collection of all possible values of the degree of freedom constitutes a *space* that is denoted by $\mathcal{F}$; the space $\mathcal{F}$ is *time independent*.

The entire edifice of quantum mechanics is built on the degree of freedom and, in particular, on the space $\mathcal{F}$.

2.7 State Space $\mathcal{V}(\mathcal{F})$

In the quantum mechanical framework, a quantum degree of freedom is inherently indeterminate and, metaphorically speaking, simultaneously has a range of possible values that constitutes the space $\mathcal{F}$.

An experimental device is constructed to examine and explore the properties of a degree of freedom and is built to correspond to the properties of the degree of freedom. Consider a quantum entity that has spin ℓ and for which the degree of freedom consists of $2\ell + 1$ discrete points. A device built for observing a spin ℓ system needs to have $2\ell + 1$ possible positions, one for each of the possible values of the degree of freedom.

The experiment needs to be repeated many times due to the indeterminacy of the quantum degree of freedom, as discussed in Sect. 9.3. The outcome of each particular experiment is *completely uncertain* and *indeterminate*, with the degree of freedom inducing the device to take any *one* of its (the device's) many possible values.[5] However, the cumulative result of repeated experiments shows a pattern—for example, with the device pointer having some positions being more likely to be observed than others.

How does one describe the statistical regularities of the indeterminate and uncertain outcomes of an experiment carried out on a degree of freedom? As mentioned in Sect. 2.6, the subject of quantum probability arose from the need to describe quantum indeterminacy. A complex-valued state vector, also called the state function and denoted by ψ, is introduced to describe the observable properties of the degree of freedom. The quantum state ψ maps the space $\mathcal{F}$ to the complex numbers $\mathbb{C}$; in particular, for the special case of coordinates $x \in \Re = \mathcal{F}$, ψ is a complex function of x. Hence

$$\begin{aligned} \psi &: \mathcal{F} \rightarrow \mathbb{C} \\ x \in \Re &\Rightarrow \psi(x) \in \mathbb{C} \end{aligned}$$

The state vector is an element of an infinite-dimensional linear vector state space. For a consistent probabilistic interpretation of quantum mechanics, discussed in Chap. 4, it is necessary that the norm of ψ be unity, namely,

$$|\psi|^2 = 1$$

The state vector hence is an element of a time-independent normed linear vector space, namely, Hilbert space $\mathcal{V}$, which is the subject matter of Chap. 4. In symbols

$$\psi \in \mathcal{V}(\mathcal{F})$$

2.8 Operators $\mathcal{O}(\mathcal{F})$

The connection of the quantum degree of freedom with its observable and measurable properties is indirect and roundabout and is always, of necessity, *mediated* by the process of measurement. A consistent interpretation of quantum mechanics requires that the measurement process plays a central role in the theoretical framework of quantum mechanics.

In classical mechanics, observation and measurement of the physical properties plays no role in the definition of the classical system. For instance, a classical particle is fully specified by its position and velocity at time t and denoted

[5] It is always assumed, unless stated otherwise, that a quantum state is not an eigenstate.

by $x(t), v(t)$; it is immaterial whether a measurement is performed to ascertain the position and velocity of the classical particle; in other words, as mentioned earlier, the position and velocity of the classical particle $x(t), v(t)$ exist *objectively*, regardless of whether its position or velocity is measured or not.

In contrast to classical mechanics, in quantum mechanics, the degree of freedom $\mathcal{F}$, *in principle*, can *never* be directly observed. All the observable physical properties of a degree of freedom are the result of a process of measurement carried out on the state vector ψ. Operators, discussed in Chap. 5, are mathematical objects that represent physical properties of the degree of freedom $\mathcal{F}$ and act on the state vector; the action of an operators on the state vector is a mathematical representation of the process measuring these physical properties. Following Dirac [10], operators that represent physical quantities are also called observables.

The degree of freedom $\mathcal{F}$ and its measurable properties—represented by the operators $\mathcal{O}_i$—are *separated* by the quantum state vector $\psi(t,\mathcal{F})$; see Fig. 3.3. An experiment can only measure the *effects* of the degree of freedom—via the state vector $\psi(t,\mathcal{F})$—*on the operators* $\mathcal{O}_i$. Furthermore, each experimental device is designed and tailor-made to measure a specific physical property of the degree of freedom, represented by an operator $\mathcal{O}_i$.

2.9 The Schrödinger Equation for State $\psi(t,\mathcal{F})$

The discussion so far has been kinematical, in other words, focused on the framework for *describing* a quantum system. One of the fundamental goals of physics is to obtain the dynamical equations that *predict* the future state of a system. This requirement in quantum mechanics is met by the Schrödinger equation that determines the *future time evolution* of the state function $\psi(t,\mathcal{F})$, where t parametrizes time. The Schrödinger equation is time reversible.

To write down the Schrödinger equation, one first needs to specify the quantum generalization of energy. The *Hamiltonian* operator H represents the energy of a quantum entity; H determines the form and numerical range of the possible allowed energies of a given quantum entity.

To exist, all physical entities must have energy; hence, it is reasonable that the Hamiltonian H should enter the Schrödinger equation. Furthermore, energy is the quantity that is conjugate to time, similar to position being conjugate to momentum, and one would consequently expect that H should play a central role in the state vector's time evolution. However, in the final analysis, there is no derivation of the Schrödinger equation from any underlying principle, and one has to simply postulate it to be true.

The celebrated Schrödinger equation is given by

$$-\frac{\hbar}{i}\frac{\partial\psi(t,\mathcal{F})}{\partial t} = H\psi(t,\mathcal{F}) \tag{2.4}$$

where $\hbar = h/2\pi$, with h being Planck's constant. Consider a quantum particle in one dimension; the degree of freedom is given by $\mathcal{F} = \Re$; for the coordinate $x \in \Re$, the Schrödinger equation is given by

$$-\frac{\hbar}{i}\frac{\partial \psi(t,x)}{\partial t} = H\psi(t,x) \tag{2.5}$$

The Hamiltonian, for potential $V(x)$, is given by

$$H = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \tag{2.6}$$

The Schrödinger equation given in (2.4) is a *linear equation* for the state function ψ; if ψ_1 and ψ_2 are solutions of the Schrödinger equation, then their linear combination

$$\psi = \alpha\psi_1 + \beta\psi_2 \tag{2.7}$$

is also a solution, where α and β are complex numbers. The *quantum superposition* of state vectors given in (2.7) is of far-reaching consequence, and its implications are discussed in Sect. 3.7.

Quantum mechanics introduces a great complication in the description of Nature by replacing the dynamical variables x and p of classical mechanics, which consist of only six real numbers for every instant of time, by an entire space $\mathcal{F}$ of the indeterminate degree of freedom; a description of the quantum entity requires, in addition, a state vector that is a *function* of the space $\mathcal{F}$. According to Dirac, the *great complication* introduced by quantum indeterminacy is "offset" by the *great simplification* due to the linearity of the Schrödinger equation [10].

The mathematical reason that state vector ψ is an element of a normed *linear vector space* is due to the *linearity* of the Schrödinger equation. The fact that ψ is an element of a linear vector space leads to many nonclassical and unexpected phenomena such as the existence of entangled states and the quantum superposition principle—to be discussed later in Chaps. 6 and 8, respectively.

2.10 Indeterminate Quantum Paths

The time evolution of physical entities is fundamental to our understanding of Nature. For a classical entity evolving in time, its *trajectory* exists objectively, regardless of whether it is observed or not and shown in Fig. 2.2a, with both its position $x(t)$ and velocity $v(t)$ having exact values for each instant of time t.

We need to determine the mode of existence of quantum indeterminacy for the case of the time evolution of a quantum particle.

Consider a quantum particle with degree of freedom $x \in \Re = \mathcal{F}$. Suppose that the particle is observed at time t_i, with the position operator finding the particle at point x_i and a second observation is at time t_f, with the position operator finding the

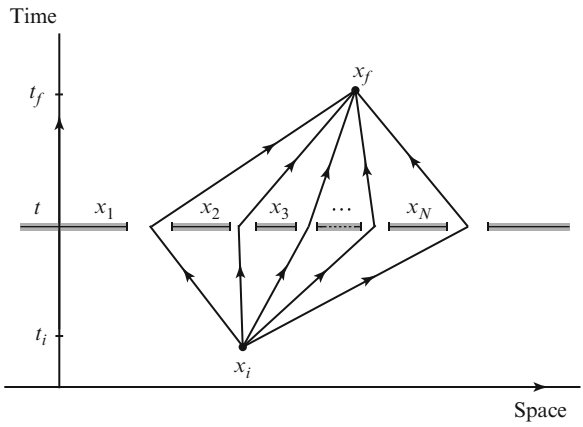


Fig. 2.6 A quantum particle is observed at first at initial position x_i at time t_i and a second time at final position x_f at time t_f. The quantum particle's path being indeterminate means that the single particle simultaneously *exists in all the allowed paths* (published with permission of © Belal E. aaquie 2012. All Rights Reserved)

particle at point x_f. To simplify the discussion, suppose there are N-slits between the initial and final positions, located at positions $x_1, x_2, \ldots, x_N$ and shown in Fig. 2.6.

There are two cases for the quantum particle making a transition from x_i, t_i to x_f, t_f, namely, when the path taken at intermediate time t is *observed* and *not observed*. The two cases are studied in detail in Chaps. 8 and 11, and the two-slit case is discussed in Sect. 3.7. For the case when the paths taken at intermediate time t is observed, one simply obtains the classical result and is discussed further in Chap. 8.

What is the description of the quantum particle making a transition from x_i, t_i to x_f, t_f when it is *not observed* at intermediate time t? The following is a summary of the conclusions:

- The quantum indeterminacy of the degree of freedom leads to the conclusion that the *path* of the quantum particle is *indeterminate*.
- The indeterminacy of the path is realized by the quantum particle by existing in all possible paths simultaneously, or metaphorically speaking, the single quantum particle simultaneously "takes" all possible paths.
- For the case of N-slits between the initial and final positions shown in Fig. 2.6, the quantum particle simultaneously exists in all the N-paths.

The term *probability amplitude* is used for describing the indeterminate paths of a quantum system. The *probability amplitude* is a complex number, and each *determinate path* is assigned a probability amplitude.

Since no observation was made to determine which path was taken, all the paths are indistinguishable, and hence, the particle's path is indeterminate, with the particle *simultaneously existing in all the N-paths*, as shown in Fig. 2.6. The probability amplitude for the quantum particle that has an *indeterminate* path is obtained by combining the probability amplitudes for the different *determinate* paths using the quantum superposition principle; this procedure is discussed for the two-slit case in Sect. 3.7 and for the general case in Chap. 8.

Let probability amplitude ϕ_n be assigned to the determinate path going through slit at x_n with $n = 1, 2, \ldots, N$, as shown in Fig. 2.6, and let ϕ be the probability amplitude for a particle that is observed at position x_i at time t_i and then observed at position x_f at later time t_f. The probability amplitude ϕ for the transition is obtained by *superposing* the probability amplitudes for all *indistinguishable determinate paths* and yields

$$\phi(x_\mathrm{f}, t_\mathrm{f}|x_\mathrm{i}, t_\mathrm{i}) = \sum_{n=1}^{N} \phi_n : \text{indistinguishable paths} \tag{2.8}$$

Once the probability amplitude is determined, its modulus squared, namely, $|\phi|^2$, yields the *probability* for the process in question. For the N-slit case,

$$|\phi(x_\mathrm{f}, t_\mathrm{f}|x_\mathrm{i}, t_\mathrm{i})|^2 = P(x_\mathrm{f}, t_\mathrm{f}|x_\mathrm{i}, t_\mathrm{i}) \;\; ; \;\; \int \mathrm{d}x_\mathrm{f} P(x_\mathrm{f}, t_\mathrm{f}|x_\mathrm{i}, t_\mathrm{i}) = 1$$

where $P(x_\mathrm{f}, t_\mathrm{f}|x_\mathrm{i}, t_\mathrm{i})$ is the conditional probability that a particle, observed at position x_i at time t_i, will be observed at position x_f at later time t_f.

Quantum mechanics can be formulated entirely in terms of indeterminate paths, a formulation that is independent of the framework of the state vector and the Schrödinger equation; this approach, known as the Dirac–Feynman formulation, is discussed in Chap. 11.

2.11 The Process of Measurement

Ignore for the moment details of what constitutes an experimental device. What is clear from numerous experiments is that the experimental *readings* obtained by observing a quantum entity (by a measuring device) cannot be explained by deterministic classical physics and, in fact, require quantum mechanics for an appropriate explanation.

Consider a degree of freedom $\mathcal{F}$; the existence of a range of possible values of the degree of freedom is encoded in its state vector $\psi(\mathcal{F})$. Let physical operators $\mathcal{O}(\mathcal{F})$ represent the observables of the quantum degree of freedom. The degree of freedom cannot be directly observed; instead, what can be measured is the effect of the degree of freedom on the operators that is mediated by the state vector $\psi(\mathcal{F})$.

The *preparation* of a quantum state yields the quantum state $\psi(\mathcal{F})$, which is then subjected to repeated measurements.

A concrete example of how the quantum state of a "quantum particle in a box" is prepared is discussed in Sect. 9.3, which we briefly review.

Electrons are obtained by heating a metal. The electrons ejected by the metal are fairly well localized in space and can be treated as semiclassical particles; the electrons are focused towards a cavity using electric and magnetic fields, as shown in Fig. 9.7a. Once an electron is inside the cavity, a mechanism has to be provided to ensure that the electron does not hit the wall of the cavity. For a special cavity,

called a Penning trap, electric and magnetic fields are used to confine the movement of the electron to a finite region of space, as shown in Fig. 9.7b.

Once the electron is inside the cavity, no further measurement is performed; the energy of the electron is chosen so that it remains confined inside the "box"; the electron is described by the quantum state $\psi(\mathcal{F})$ for a quantum particle in a box.

Operators $\mathcal{O}(\mathcal{F})$, discussed in Chap. 5, are the mathematical basis of measurement theory. The experimental device is *designed* to measure the properties of the operator $\mathcal{O}(\mathcal{F})$. Measurement theory requires knowledge of special quantum states, namely, the eigenstates χ_n of the operator $\mathcal{O}(\mathcal{F})$, which are defined as follows[6]:

$$\mathcal{O}(\mathcal{F})\chi_n = \lambda_n \chi_n$$

Because the process of measurement ascertains the properties of the degree of freedom by subjecting it to the experimental device. The measurement is mathematically represented by applying the operator $\mathcal{O}(\mathcal{F})$ on the state of the system $\psi(\mathcal{F})$ and projecting it to one of the eigenstates of $\mathcal{O}(\mathcal{F})$, namely,

$$\psi(\mathcal{F}) \;\rightarrow\; \text{Measurement} \;\rightarrow\; \chi_n \;:\; \text{collapse of state } \psi(\mathcal{F})$$

Applying $\mathcal{O}(\mathcal{F})$ on the state vector collapses it to one $\mathcal{O}(\mathcal{F})$'s eigenstates.

The projection of the state vector ψ to one of the eigenstates χ_n of the operator $\mathcal{O}(\mathcal{F})$ is *discontinuous* and *instantaneous*; it is termed as the *collapse of the state vector* ψ. The result of a measurement has to be *postulated* to lead to the collapse of the state vector and is a feature of quantum mechanics that *is not governed* by the Schrödinger equation.

Unlike classical mechanics, where the same initial condition yields the same final outcome, in quantum mechanics the same initial condition leads to a wide range of possible final states. The result of identical quantum experiments is inherently uncertain.[7] For example, radioactive atoms, even though identically prepared, decay randomly in time precisely according to the probabilistic predictions of quantum mechanics.

After many repeated observations performed on state $\psi(\mathcal{F})$, all of which in principle are identical to each other, the experiment yields the average value of the physical operator $\mathcal{O}(\mathcal{F})$, namely,

$$\mathcal{O} \;\rightarrow\; \text{measurements on } \psi(\mathcal{F}) \;\rightarrow\; E_\psi[\mathcal{O}(\mathcal{F})]$$

Because the process of measurement *cannot* be modeled by the Schrödinger equation, this has long been a point of contention among physicists since many physicists hold that the fundamental equations of quantum mechanics should determine both the evolution of the quantum state and the collapse of the state

[6] Eigenstates are discussed in (5.5).

[7] Except, as mentioned earlier, for eigenstates.

caused by the process of measurement. As of now, there has been no resolution to this conundrum.

2.12 Summary: Quantum Entity

In light of the superstructure of quantum mechanics, what is a quantum entity? A careful study of what is an entity, a thing, and an object leads to the remarkable conclusion that the quantum entity is intrinsically indeterminate and its description requires a framework that departs from our classical conception of Nature.

The quantum entity's foundation is its degree of freedom $\mathcal{F}$, and quantum indeterminacy is due to the intrinsic indeterminacy of the degree of freedom. A landmark step, taken by Max Born, was to postulate that quantum indeterminacy can be described by a state vector $\psi(\mathcal{F})$ that obeys the laws of (quantum) probability. The state vector is inseparable from the degree of freedom and encodes all the information that can be obtained from the indeterminate degree of freedom and is illustrated in Fig. 2.7.

It should be noted that the state vector $\psi(\mathcal{F})$ does not "surround" the degree of freedom in physical space; rather, Fig. 2.7 illustrates the fact that all observations carried out on the degree of freedom always encounter the state vector and no observation can ever come into direct "contact" with the degree of freedom itself. All "contact" of the measuring device with the degree of freedom is mediated by the state vector.

In summary, the following is a definition of the quantum entity:

> A quantum entity is constituted by a pair, namely, the degree(s) of freedom $\mathcal{F}$ and the state vector $\psi(\mathcal{F})$ that encodes all of its properties. This inseparable pair, namely, the degree of freedom and the state vector, embodies the condition in which the quantum entity exists.

According to the Copenhagen quantum postulate, observations carried out on the degree of freedom collapse the quantum state to a definite state, namely, an eigenstate of a physical operator; the outcome of every experiment (except for eigenstates) is uncertain. Repeated observations yield quantities that describe the observable and quantifiable properties of the quantum entity.

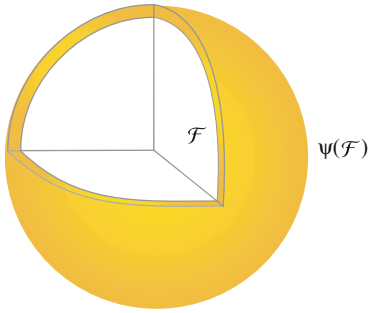


Fig. 2.7 A quantum entity is constituted by its degrees of freedom $\mathcal{F}$ and the state vector $\psi(\mathcal{F})$ that permanently encompasses and envelopes its degrees of freedom (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Heisenberg pioneered a framework for describing quantum mechanics based on the algebra of operators; quantum indeterminacy and the concepts of quantum probability are defined in terms of these operators, discussed in Chap. 7.

A third and independent formulation of quantum mechanics is based on the Dirac–Feynman framework of the degree of freedom having *indeterminate paths* (indeterminate time evolution) and is expressed in terms of the Feynman path integral, discussed in Chap. 11.

Quantum indeterminacy—formulated by Born in terms of the state vector—was later realized to have a more transparent representation in Heisenberg's operator formulation. Physical quantities are represented by operators, and every experimental device is custom designed for measuring the properties of the operator. The quantum state is represented by the density matrix operator, discussed in Chap. 6. The operators extract all the information obtainable about the quantum entity by acting on its quantum state, which encodes all the physically observable properties of the degree of freedom, and in effect of the quantum entity as well.

In the ultimate analysis all that one can finally measure in an experiment is the effect of the quantum entity on the readings of our experimental devices, which in many cases consists of discrete changes ('clicks') in the readings of the device's counters.

All of our understanding of a physical entity is based on our theoretical and mathematical concepts, which in turn are constructed from, and have to be explain, a plethora of experimental data. In the case of quantum mechanics, the mathematical construction has led us to infer the existence of the quantum entity. The theoretical constructions of quantum mechanics are far from being arbitrary and ambiguous; to the contrary, given the maze of links from the quantum entity to its empirical properties, it is highly unlikely that there are any major gaps or redundancies in the theoretical superstructure of quantum mechanics.

Quantum Mechanics: Empirical and Trans-empirical

3

The founders and leading proponents of quantum mechanics were well aware of the paradoxical and opaque workings of quantum mechanics that do not conform to our everyday intuition, an intuition derived from the world of macroscopic objects. Niels Bohr had the following to say about quantum mechanics: "Those who are not shocked when they first come across quantum mechanics cannot possibly have understood it" (Quoted by Heisenberg in [18]). Richard Feynman made the following observation: "I think it is safe to say that no one understands quantum mechanics. Do not keep saying to yourself, if you can possibly avoid it, 'But how can it be like that?' because you will get 'down the drain' into a blind alley from which nobody has yet escaped. Nobody knows how it can be like that" [13].

The mysteries and paradoxes of quantum mechanics arise due to the following two reasons:

- The intrinsic indeterminacy and uncertainty exhibited by a quantum entity is completely absent in our everyday life.
- The linearity of the Schrödinger equation that determines the dynamics of the quantum state.

The mystery of quantum mechanics is not just only about indeterminism but rather also about the *manner* in which this indeterminism is realized.

Dirac has a famous statement that a "picture", namely, a metaphor, in a scientific theory need not be a classical construct but rather should be a "way of looking at the fundamental laws which makes their self-consistency obvious" [10]. In this spirit of Dirac, a new concept, namely, the *trans-empirical*, is proposed for inclusion into the lexicon and conceptual framework of quantum mechanics. The term trans-empirical is introduced for providing a "picture," a metaphor, for the equations of quantum mechanics; the hope is that the trans-empirical concept will bring greater clarity to the inner workings of quantum mechanics by making it theoretically more transparent as well as accessible to deeper mathematical analysis.

The following are the topics covered in this chapter:

- Definition of the concept of trans-empirical.
- The reexamination of the various domains of the quantum entity's superstructure.

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8_3, © Springer Science+Business Media New York 2013

- The study of the two-slit experiment as an exemplar for studying the relation of the trans-empirical domain to quantum superposition.
- The trans-empirical quantum principle is stated and essentially entails enhancing the concept of what "exists" in Nature.

3.1 *Real* Versus *Exist*

According to Dirac, quantum mechanics is "built up from physical concepts .. which cannot even be explained in words at all" [10]. Referring to quantum mechanics, Niels Bohr had the following to say about what is "real":"We are suspended in language in such a way that we cannot say what is up and what is down. The word 'reality' is also a word, a word which we must learn to use correctly" [30].

Due to its opaque and unfamiliar workings, words that are employed in discussing, understanding, and explaining the concepts of quantum mechanics have come to be used in a very specific manner, and one needs to carefully define these words. Two central concepts that we need to discuss are what is "real" and what "exists"; these two terms have many implications in quantum mechanics, and hence it is necessary to clarify their meaning. Whether something is "real" or "exists" comes down to how these words are defined; the dictionary meaning of the word "exist" is extended so that it can be used to discuss many concepts that arise in quantum mechanics.

The Oxford dictionary defines *real* as *actually existing as a thing or occurring in fact; not imagined or supposed* and defines *exist* to *have objective reality or being*. Hence, the words "to exist" and "to be real" seem to have the same meaning, with "to exist" meaning "to be real" and "to be real" meaning "to exist"; in fact, these two words are normally used synonymously.

An entity that is real is defined as *existing . . . in fact* and points to the entity's empirical, objective, and factual existence. Quantum mechanics has adopted this definition of "real" and also provides a new dimension to the concept of what is "real"; according to Heisenberg, "Reality is in the observations, not in the electron" [18]: What is *defined* by Heisenberg to be real is not the electron itself but rather the *result* of observations on the electron. The results of observations are taken to be real, and Heisenberg is silent on the ontological status of the electron itself.

For Heisenberg, similar to its dictionary meaning, the word "real" stands for an *empirical reality*, a reality that has an objective and factual existence. Instead of debating whether this is the most appropriate use of the word "real," for the sake of clarity, Heisenberg's definition of the word "real" is adopted in this book. Henceforth, the word "real" refers only to the result of observations, or in short, to what is empirical.

As discussed in Sect. 2.2 and elaborated in detail in Chap. 7 on Quantum Indeterminacy, a quantum particle, say, an electron, is inherently indeterminate; what this means is that, before its position operator is measured, the particle's "position" does not have an objective reality; instead, only on being observed does the particle yield a determinate and real value to the observation. In other words, Heisenberg's

statement "Reality is . . . not in the electron" means the following: that in quantum mechanics, what is *real* and *existing in fact* is only the result of an observation (of the electron) and that the electron itself, as such, does not have an objective reality.

The quantum entity (the degree of freedom and its inseparable quantum state) is inherently indeterminate and, unlike a classical entity, does not have an objective reality, but it certainly has some *mode of existence* since the quantum degree of freedom continues to "exist" when it is not observed. How should we refer to the "being" of a quantum entity when it is not being observed, when it is *intrinsically indeterminate*?

It is most logical to use the word "exist" for entities that are empirical as well as trans-empirical. What is at issue is the concept of an *objective existence* of an entity and which has crept into the dictionary definition of "exist". There are many entities such as mathematics, language, and so on that don't have an empirical and objective existence, but nevertheless do exist. To describe such entities as well as the being of an indeterminate entity, in this book, the meaning of "exist" is no longer taken to entail having objective reality—which is now reserved for something that is "real"—and the definition of *exist* is *generalized* and *extended* to describe any entity that *is*, namely, *has being*, and not necessarily has an objective being or an objective existence or an objective reality.

The word "exist" can now be used for describing the being of entities that *don't have* an objective existence and which includes entities that are inherently indeterminate. With the modified definition of the word "exist", the *unobservable* and indeterminate quantum degree of freedom is said to *exist*, namely, it has a "nonobjective" mode of existence.

In summary, the point of view and ontological stand adopted in this book is that the unobservable and indeterminate quantum degree of freedom *exists*; in contrast, the empirically observed properties of the degree of freedom are *real*.

3.2 Empirical, Trans-empirical, and Indeterminate

The rather inexplicable and unexpected feature of quantum mechanics—that the degree of freedom $\mathcal{F}$, *in principle*, cannot be empirically observed—is explored in this book. In particular, to explain and understand the unusual properties of the degree of freedom, the term *trans-empirical* is introduced for describing $\mathcal{F}$.

The Oxford dictionary defines *empirical* as being *based on, concerned with, or verifiable by observation or experience rather than theory or pure logic*. The empirical realm is accessible to direct experimental observations.

The term trans-empirical is defined as a domain of existence that lies beyond, that transcends, the empirical realm. The trans-empirical domain, in particular, is *inaccessible to direct observation, experiential and experimental knowledge,* and *only accessible to theory or pure logic*.

The term *trans-empirical* has been defined in a manner so that it can be used to denote the virtual and *unobservable* form of existence of a quantum entity.[1] The adjective trans-empirical is sometimes used for qualifying existence to emphasize the difference in the meaning of trans-empirical *existence* from the empirically *real*.

An entity that is determinate is in a definite state, and this definiteness is a property intrinsic to the entity; hence, a determinate entity is taken to be synonymous to being an empirical entity; the quality of determinateness is intrinsic to the entity and hence exists regardless of whether it is observed or not.

The Cambridge Dictionary defines *indeterminate* as *not measured, counted or clearly described*; this definition is far too literal and narrow a definition since it implies the objective existence of the entity that is indeterminate; this definition of indeterminate will *not* be adopted in this book.

The term "indeterminate" has been used in the foregoing sections as being synonymous to quantum uncertainty—based on an intuitive understanding of this word as being the "opposite" of being determinate, definite, and factual. The concept of "indeterminateness" needs to be defined more precisely so that it can be related to quantitative questions such as, in an experiment, what can and cannot be termed as being an "indeterminate" entity.

A classical entity is intrinsically determinate since, at each instant, it has definite and exact values for its dynamical variables. In contrast, a quantum entity is intrinsically indeterminate, with the quantum degree of freedom having no exact and objective existence prior to a measurement being performed. Hence, as one can see from the description of a classical and quantum entity, the words determinate and indeterminate provide a consistent manner of describing the contrasting classical and quantum entities.

The word *indeterminate* is closely related to the concept of trans-empirical. On examining the concept of indeterminacy, one comes to the conclusion that it is another way of describing a trans-empirical entity, since at the root of indeterminacy is not the lack of precise knowledge of an a priori objectively existing entity, but rather indeterminacy refers to the intrinsic nature of an entity that *does not have* an objective existence. While acknowledging that the word "indeterminate" can be defined in many different ways, in this book the concept of "indeterminateness" is defined to mean a trans-empirical entity, namely, a form of existence that is not observable and is trans-empirical.

The words indeterminate and trans-empirical will be used interchangeably, with a specific choice being made depending on the context of the discussion.

Consider the earlier example of a quantum particle discussed in Sect. 2.2 The position degree of freedom of the particle is intrinsically indeterminate in the sense that it does not have any definite position; instead, the single particle exists at every point (centered on some average position), and in this sense the particle's position

[1]The term "virtual" is henceforth avoided since it has become associated with software-based simulations such as "virtual reality" and "virtual machines" that are not related, in any manner, to the virtual states of quantum mechanics.

is indeterminate. Clearly, the quantum particle is in a trans-empirical state since one can never empirically observe the particle being at more than one point.

Consider the indeterminate paths discussed in Sect. 2.10 and shown in Fig. 2.6. In what sense are the paths indeterminate? There are *N*-paths in Fig. 2.6, paths that are determinate and have a definite trajectory. The particle's path being indeterminate means that it simultaneously exists in all the allowed determinate paths; again, the particle taking indeterminate paths of necessity requires that the quantum particle exists in a trans-empirical form since one will never find the particle at more than one point.

In conclusion, the quantum entity's indeterminacy and its trans-empirical state are inseparable since to be indeterminate the quantum entity must exist in a trans-empirical state and vice versa.

3.3 Quantum Mechanics and the Trans-empirical

This book is written within the framework of the Copenhagen interpretation of quantum mechanics, developed primarily by Bohr and Heisenberg [17].

Recall in the Copenhagen interpretation only the experimentally *observed* quantities of the quantum system are taken to be *real*. There are a number of issues on which the Copenhagen interpretation is silent, the main one being the *ontological* status of the quantum system when it is *not* observed.

As argued in Sect. 3.1, clearly the quantum entity, in particular the quantum degree of freedom, continues to *exist* even when it is *not observed*. A primary focus of this book is to try and clarify the ontological condition of the degree of freedom when it is *not observed*. In particular, in what sense does the quantum entity *continue* to exist, both before and after it is observed.

> The Copenhagen interpretation is *augmented* and *enhanced* by expanding the definition of what is to be considered to *exist*. The degree of freedom and the quantum state are not taken to be merely a mathematical representation of nonexistent entities that magically produce the results of quantum mechanics. Rather, this book explores in what quantum mechanical sense can the quantum entity, that is, the degree of freedom and its quantum state, be taken to *exist*. The existence of the quantum entity is taken to replace the concept of "objective reality" of classical mechanics.

The domain of what *exists* is extended from the classical domain of the *empirical* to a new domain of Nature that is *trans-empirical*. The concept that Nature has a trans-empirical domain is introduced in an attempt to explain and understand the range of phenomena revealed by quantum mechanics. This expanded view of what *exists* is postulated with the purpose of clarifying the workings of quantum mechanics.

The trans-empirical domain does not exist in classical physics since all phenomena in classical physics exist objectively and hence are, in principle, empirical and directly observable. Quantum theory has enhanced our conception of Nature by its discovery of the trans-empirical domain.

Trans-empirical	Transition		Measure-ment	Empirical
$\mathcal{F}$	$\mathcal{V}(\mathcal{F})$	$\frac{\partial \psi(t,\mathcal{F})}{\partial t}$	$\hat{O}(\mathcal{F})$	$E_{\mathcal{V}}[\hat{O}(\mathcal{F})]$
Degree of freedom	State space	Dynamics	Operators	Observation

Fig. 3.1 The organization of the superstructure of a quantum entity into the empirical, the trans-empirical, and the transitional domains (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Physics is an empirical science, and all laws of physics are empirical; the relation of a trans-empirical domain of Nature to the empirical realm of physics is discussed in Sect. 3.10.

Quantum Superstructure and the Trans-empirical

The five cardinal pillars of the quantum entity's superstructure, given in Fig. 2.5, can be organized using the criterion of empirical and trans-empirical; in the next two sections, each cardinal pillar of quantum mechanics is analyzed from the point of what is and what is not empirically observable; the results of this analysis are summarized in Fig. 3.1.

- The foundation of the quantum entity, namely, its degree of freedom $\mathcal{F}$, is a *trans-empirical* quantity.
- The empirical manifestation of the quantum entity is observed in experimental readings. The result of experiments is to provide an empirical value for the expectation value of observable operators, namely, $E_{\mathcal{V}}[\mathcal{O}(\mathcal{F})]$.
- The *transitional* domain—connecting the trans-empirical and empirical domains—is *straddled* by quantum state vector. On the one hand, the quantum state $\psi(t,\mathcal{F})$ is *trans-empirical*, is in "contact" with the degree of freedom $\mathcal{F}$, and encodes all the observable properties of $\mathcal{F}$. On the other hand, measurements on the degree of freedom $\mathcal{F}$ cause the quantum state to collapse to its *empirical manifestation*, which after repeated measurements yields the empirical result $E_{\mathcal{V}}[\mathcal{O}(\mathcal{F})]$.
- Operators $\mathcal{O}(\mathcal{F})$ representing physical observables are in the *transition* domain from the side of the empirical domain, causing the quantum state to make a transition from the trans-empirical to the empirical domain. The operators do not have direct overlap or "contact" with the degree of freedom $\mathcal{F}$, but instead, the state vector $\psi(t,\mathcal{F})$ mediates the connection of the operators $\mathcal{O}(\mathcal{F})$ with the degree of freedom $\mathcal{F}$.

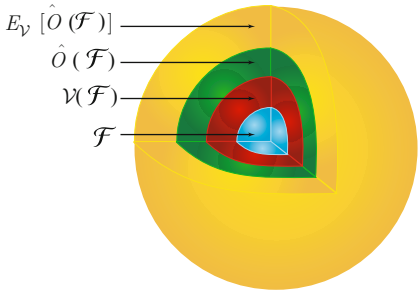


Fig. 3.2 The enveloping theoretical superstructure that describes the nature of a quantum entity (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Figure 3.2 illustrates the nature and organization of the superstructure of a quantum entity. The foundation of the quantum entity is its degree of freedom $\mathcal{F}$, with the entire quantum superstructure built on it. The diagram represents the encompassing nature of the various domains of the quantum entity's superstructure. It is logical to conclude that the characteristic features of a quantum entity are a manifestation of its mode of existence that reflects an ontology far more complex than that of a classical entity.

3.4 Quantum Degree of Freedom $\mathcal{F}$ Is Trans-empirical

The quantum degree of freedom is a quantitative entity that simultaneously exists in all of its possible allowed values. It is intrinsically and inherently indeterminate and trans-empirical and constitutes the degree of freedom space $\mathcal{F}$.

As mentioned earlier, an important conclusion of the Bell analysis, discussed in detail in Chap. 7, is that when a quantum degree of freedom is subjected to an experiment, mathematically represented by operators $\mathcal{O}$, it can be deduced that the degree of freedom *does not* have any precise and determinate value before it is observed.

Clearly, *before* a measurement is carried out, the degree of freedom $\mathcal{F}$ *exists*.[2] The fact that $\mathcal{F}$ *does not* have a precise value before a measurement refers to its *mode of existence*. It has been discussed in Chap. 2 that the mode of existence of a degree of freedom is that it exists in all of its totality, with its different possible values being organized as a geometrical space, namely, $\mathcal{F}$; metaphorically speaking, the degree of freedom simultaneously exists over the entire range of all of its possible values. This mode of existence is not empirically observable, and hence, the degree of freedom is considered to be trans-empirical.

The degree of freedom is an entire space $\mathcal{F}$; expecting the degree of freedom to have a single value is like trying to describe the entire space by giving the coordinates of a single point of that space. The degree of freedom $\mathcal{F}$ existing as an *entire space* is an explanation of why the degree of freedom does not have a precise value before it is observed.

[2] And continues to exist after a measurement as well.

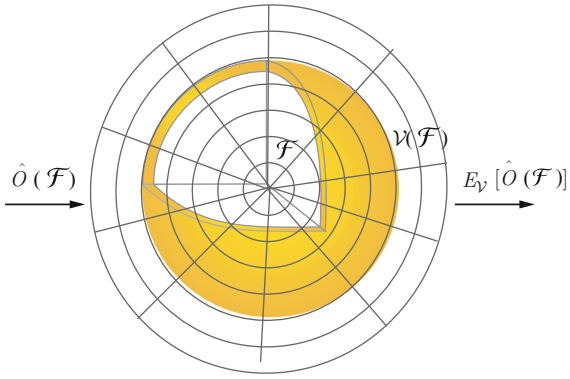


Fig. 3.3 The degree of freedom $\mathcal{F}$ is veiled and masked by state space $\mathcal{V}(\mathcal{F})$ from being directly observed by operators $\mathcal{O}(\mathcal{F})$. A measurement by the operator $\mathcal{O}(\mathcal{F})$ is represented by a "net" that is cast around the quantum entity to extract information about the degree of freedom $\mathcal{F}$. Repeated observations yield the expectation value $E_{\mathcal{V}}[\mathcal{O}(\mathcal{F})]$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

The framework of quantum probability identifies observables with operators and not with the specific values of the degree of freedom; this is unlike classical probability theory for which the observables are specific values of a random variable that are directly observed when the random variable is sampled.

Operators act on the quantum state of the quantum entity, and the expectation value of the operator is determined by the state vector in question. In other words, an act of applying the operator to the degree of freedom leads to the operator indirectly "picking out" the properties of the degree of freedom by collapsing the quantum state—with the degree of freedom $\mathcal{F}$ remaining indeterminate and trans-empirical.

All empirical quantities in quantum mechanics, such as the expectation value of physical operators, are determinate and have a unique value; empirical properties are also determinate for a classical random variable. In classical probability theory, since one can directly observe the random variable, the particular values of the random variable exist objectively. In contrast, the degree of freedom $\mathcal{F}$ always remains concealed, veiled, and masked by the quantum state vector $\psi(\mathcal{F})$, and the specific and determinate "points" of the space $\mathcal{F}$ cannot be directly observed.

Recall that the quantum entity is constituted by its degree of freedom and its state vector. The quantum operators representing the observed quantities act on the quantum state $\psi(\mathcal{F})$ and in general on the state space $\mathcal{V}(\mathcal{F})$. The operators $\mathcal{O}(\mathcal{F})$ can only indirectly detect the existence of the degree of freedom via their action on the quantum state. The quantum state "shields" $\mathcal{F}$ from being directly observed by operators $\mathcal{O}(\mathcal{F})$. The action of an operator on state space and its relation to the underlying degree of freedom are shown in Fig. 3.3.

The degree of freedom thus always exists as a trans-empirical entity, which is an entire space $\mathcal{F}$. In particular, the degree of freedom is *indeterminate* because it (the

degree of freedom) does not exist as a particular value ('point') but as an entire space $\mathcal{F}$, and hence, it ($\mathcal{F}$) cannot be described by a quantity having a single, determinate and fixed value.

3.5 The Quantum State ψ: Transition

The trans-empirical quantum degree of freedom, *in principle*, can never be observed. The *quantum state* (or state in short), in contrast, is at the *interface* of the trans-empirical degree of freedom on the one hand and the empirical experimentally measured properties of the degree of freedom on the other hand. The quantum state *connects* the trans-empirical domain with the empirical and experimental domain.

The *transition*—from the trans-empirical to the empirical domain—is effected by the act of measurement, and the *reverse transition*—from the empirical to the trans-empirical domain—is the result of the preparation of a quantum state.[3]

The empirical and trans-empirical duality is shown in Fig. 3.4. The term "wave-particle duality," used by the founders of quantum mechanics, is in fact the duality of the empirical and trans-empirical domains.

The quantum state is realized in a very specific and particular form in quantum mechanics. Let the quantum state of the degree of freedom be represented by $\psi(t,\mathcal{F})$, which is a complex valued function of $\mathcal{F}$ and is an element of state space $\mathcal{V}(\mathcal{F})$; the quantum state is sometimes referred as a *probability amplitude* to differentiate it from empirically observed probabilities. In the framework of the empirical and trans-empirical dual domains and the Copenhagen interpretation of quantum mechanics, the quantum state has the following formulation:

- The quantum state has *two forms of existence*, namely, a *trans-empirical* form when it is not observed and the form of *empirical manifestation* when it is observed, as shown in Fig. 3.4.
- When the quantum state is *not observed*, the quantum state's existence is *trans-empirical* and given by $\psi(t,\mathcal{F})$; the time evolution of the state function $\psi(t,\mathcal{F})$ takes place entirely in the trans-empirical domain, with all the many different specific values of the degree of freedom $\mathcal{F}$ simultaneously evolving. In particular,

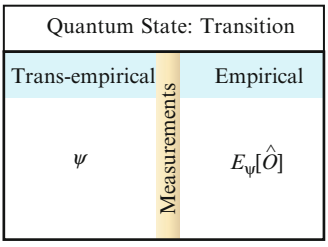


Fig. 3.4 The trans-empirical form of the quantum state vector and its empirical manifestation are connected by the process of measurement (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

[3]The preparation of a quantum state is briefly summarized in Sect. 2.11 and discussed in more detail in Sect. 9.9.

the Schrödinger equation evolves the trans-empirical form of the state by an equation determining $\partial\psi(t,\mathcal{F})/\partial t$ and yields the value of $\psi(t,\mathcal{F})$ for future time t.

- When the quantum state is *observed* by a process of measurement—with the experimental device mathematically represented by an operator $\mathcal{O}$—the state $\psi(t,\mathcal{F})$ collapses, namely, makes a *discontinuous* transition, to its *empirical manifestation*, as shown in Fig. 3.4. Repeated observations of $\psi(t,\mathcal{F})$ yield the empirical average value of the operator given by $E_\psi[\mathcal{O}]$.[4]

The preparation of a state and its subsequent experimental observation is one of the cornerstones of quantum mechanics. All processes used for the preparation and measurement of a quantum state have a finite resolution of points in space and take place over a finite duration of time; hence, the separation of the empirical and trans-empirical domains is not a sharp boundary but rather a narrow gray area, labeled "Measurements" in Fig. 3.4; in the interface, the state of the quantum object is a mixture of the empirical and trans-empirical states. In Sect. 8.11, quantum superposition is studied in some detail; it is shown that there are quantum entities that are in the gray area in Fig. 3.4—being a "mix" of empirical readings and trans-empirical states.

For concreteness, consider a quantum particle, say, an electron, with a position degree of freedom given by $\mathcal{F} = \Re$. Consider observing the state with an *operator* that projects the degree of freedom to the specific *position* $x \in \Re$ of the quantum particle. The *trans-empirical* form of the quantum state is $\psi(t,\Re)$. The process of measurement is the mechanism of going from its trans-empirical form to its empirical manifestation, and the preparation of a quantum state causes a transition from its empirical to its trans-empirical form. The various aspects of the quantum state and its empirical manifestation, for the position degree of freedom $\Re$, are shown in Fig. 3.5.

Applying the position projection operator to the state $\psi(t,\Re)$ collapses it to its *empirical manifestation* given by $|\psi(t,x)|^2$ and yields the likelihood that the projection operator at position x detects the state vector, as shown in Fig. 3.5.

The *preparation* of a quantum system selects a particular quantum state. Once the quantum system is prepared to be in a specific state and not observed anymore, it makes a smooth transition to its trans-empirical form and evolves as a trans-empirical entity, with all the different possible values of the degree of freedom evolving simultaneously. Hence, as shown in Fig. 3.5, the preparation of a quantum state causes a transition from the empirical to the trans-empirical state.

The measurement of the position projection operators picks out *one possible* outcome for the state vector from the many different possible outcomes. It is because the quantum degree of freedom is a trans-empirical quantity—having no fixed value—that identical acts of observing the quantum state lead to many different possible outcomes, as discussed later in Sect. 9.3.

[4]It is shown in Sect. 5.8 that $E_\psi[\mathcal{O}] = \langle\psi|\mathcal{O}|\psi\rangle$.

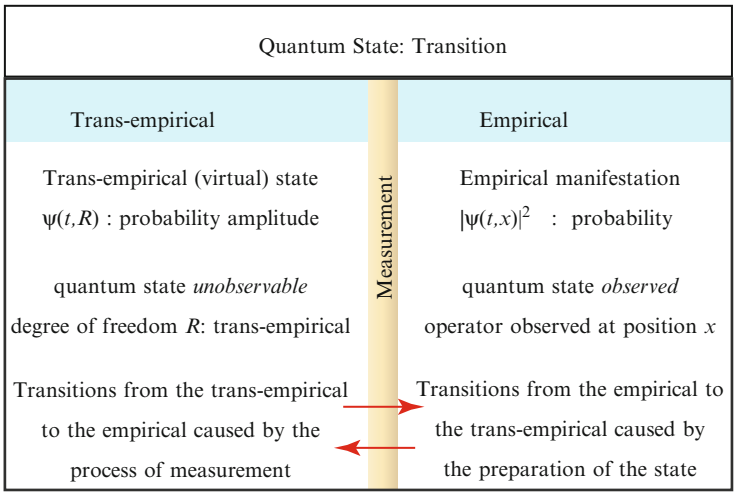


Fig. 3.5 Trans-empirical existence of the quantum state vector and its empirical manifestation for the particle's position degree of freedom (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

In summary, a remarkable property of the quantum state is that it has two forms of existence, a *trans-empirical form* given by $\psi(t,\mathcal{F})$ and its "collapsed" form that yields its *empirical manifestation*, given by $E_\psi[\mathcal{O}]$. Metaphorically speaking, the state vector is "alive" in its trans-empirical and is "dead"—ceases to exist—in its empirical manifestation.

It is due to the quantum state possessing two qualitatively different forms—concretely, the empirical manifestation $|\psi(t,x)|^2$ and the trans-empirical form $\psi(t,x)$—that the quantum state, by being the interface of these two domains, can connect and straddle two qualitatively different domains of Nature.

3.6 Trans-empirical Domain and Laws of Physics

All laws of physics are empirically grounded, even if this be indirectly. Empirical science is defined by the following principle enunciated by Richard Feynman: "Experiment and experiment alone is the sole criterion of scientific truth" [24]. How does a trans-empirical construct such as the quantum degree of freedom $\mathcal{F}$ and the quantum state $\psi(\mathcal{F})$ enter into an empirical science such as physics?

There are many mathematical constructs in physics, such as the electromagnetic Maxwell field A_μ, space-like points in special relativity, and the interior of black holes in general relativity that are not directly observable and are mathematical objects useful for expressing the physical theory. However, all the mathematical constructs of classical physics refer to tangible entities that have determinate values and also have an objective existence (reality) in spacetime.

In contrast, the quantum entity is not determinate but rather is inherently indeterminate—which makes the quantum degree of freedom $\mathcal{F}$ completely unobservable. It is indeterminateness, which is at the foundation of quantum mechanics, that creates an unbridgeable schism between quantum and classical physics. Indeterminate $\mathcal{F}$ is described by trans-empirical quantum state $\psi(\mathcal{F})$ which exists in state space $\mathcal{V}(\mathcal{F})$ that is a superstructure on physical spacetime.

$\psi(\mathcal{F})$ has a finite likelihood of simultaneously "existing" at all of the possible values of the degree of freedom $\mathcal{F}$; this trans-empirical mode of existence of the quantum state cannot be experimentally observed since an experiment always collapses the state to a determinate condition and hence always produces a *single* and *determinate* result.

One can only gain *indirect evidence* of the quantum entity by experimentally interacting with the indeterminate degree of freedom $\mathcal{F}$ using appropriate operators $\mathcal{O}(\mathcal{F})$. The operators repeatedly encounter $\psi(\mathcal{F})$ in the experiments carried out, as shown in Fig. 3.3, with all the information that can be extracted from $\mathcal{F}$ being contained in the quantum state $\psi(\mathcal{F})$.

The result of, in principle, infinitely many observations yields the full range of the possible determinate values of quantum mechanical state vector—together with the *likelihood* of a specific determinate value being observed in a particular experiment; one can theoretically reconstruct the properties of the quantum entity by *imagining* (mathematically modeling) its indeterminate degree of freedom $\mathcal{F}$ together with its trans-empirical quantum state $\psi(\mathcal{F})$.

The quantum entity exists in the trans-empirical domain. There are, however, *measurable* and *empirical consequences* of the trans-empirical domain that yield observable properties of the quantum entity. In fact, it is a matter of supreme irony that physics, the most empirical of all the sciences, is founded on trans-empirical entities that appear, of necessity, in the mathematical formalism of quantum mechanics.

3.7 Quantum Superposition: Trans-empirical Paths

The concept of the *indeterminacy* of *quantum paths* was discussed in Sect. 2.10, and this concept is now studied in some detail using the two-slit experiment. Indeterminate paths can never be directly observed, and hence indeterminate paths are *trans-empirical paths*.

The two-slit experiment goes back to Young (1799), who showed that light going through two slits results in interference, and was crucial in demonstrating that light is a wave. The two-slit experiment is one of the deepest and most important experiments in quantum mechanics and is discussed with the aim of demonstrating the mode of existence of indeterminate paths, as well as the role of measurement in causing a transition from the trans-empirical form of the quantum entity to its empirical manifestation.

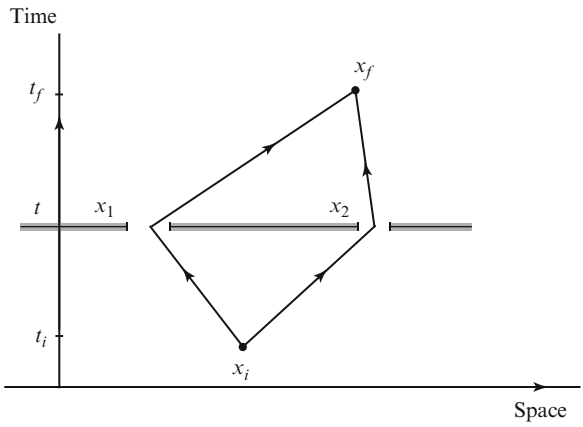


Fig. 3.6 A quantum particle is observed *only* at initial position x_i at time t_i and at final position x_f at time t_f. The paths taken from initial to final position are indeterminate and trans-empirical (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Quantum superposition is one of the bedrocks of quantum mechanics and is discussed in detail in Chap. 8. The two-slit experiment provides one of the simplest illustrations of indeterminate trans-empirical paths and quantum superposition.

As explained by Richard Feynman, the two-slit experiment "has been designed to contain all of the mystery of quantum mechanics, to put you up against the paradoxes and mysteries and peculiarities of nature one hundred percent"[13]. Feynman further explains that all the other paradoxical situations in quantum mechanics can always be explained by this experiment—which reveals "nature in her most elegant and difficult form"[13].

The simplest case of indeterminate trans-empirical paths is for the quantum particle to simultaneously exist in two distinct paths, as shown in Fig. 3.6, and which can be generalized to the case of the N-slit, shown in Fig. 2.6. In this section, the two-slit experiment is analyzed using electrons; it is shown that when a measurement determines the path of the electron, the path is empirical and determinate and the electron behaves like a classical particle; however, when the trajectory is *not observed*, the electron exists in an indeterminate and trans-empirical state and exhibits the phenomenon of quantum interference.

The two-slit experiment is employed for analyzing the following topics:

- The concepts of the empirical and trans-empirical are applied to the time evolution of a quantum entity. In classical mechanics, the classical entity takes a determinate path, going through either the slit at position x_1 or through the slit at position x_2. The result of the two-slit experiment can be explained by postulating that when the quantum entity is *not observed*, the *path* taken by the quantum entity is *indeterminate* and trans-empirical, with the quantum entity simultaneously existing in *both* paths.
- The superposition of the quantum state vector, which is due to the *linearity* of the Schrödinger equation, is also valid for indeterminate paths. For a quantum particle "taking" indeterminate trans-empirical paths, the probability amplitude of going from an initial to a final position is shown to result from the quantum superposition of determinate empirical paths.

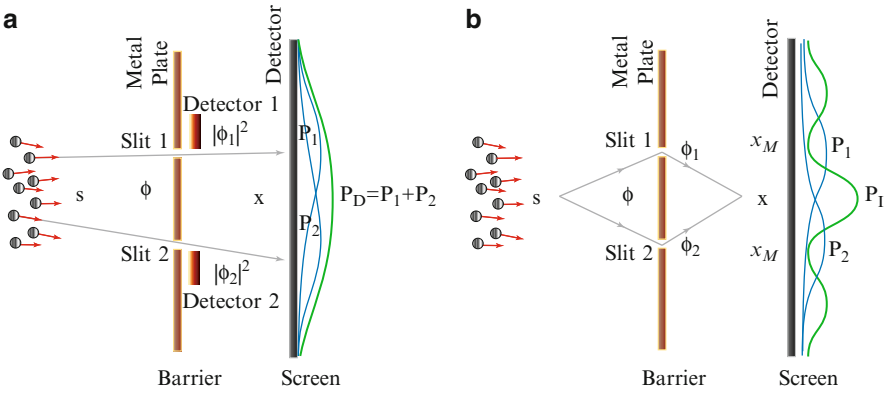


Fig. 3.7 (**a**) Two-slit experiment *with* determination of the path taken to reach the screen. (**b**) Two-slit experiment *without* determination of the path taken to reach the screen (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

The Experiment

A quantum particle going through two slits, as given in Fig. 3.6, is realized experimentally by the arrangement shown in Fig. 3.7. Unlike Fig. 3.6, the time dependence of the paths is not shown in Fig. 3.7, where the emphasis is on the measurements being performed. The information about the paths, in particular whether they are empirical or trans-empirical, is reconstructed from the experimental measurements.

The experiment consists of an electron gun (*source*) that sends identically prepared electrons, through a barrier with *two slits*, to a *screen* where a screen detector keeps track of the point at which the electron hits the screen. Note that the electrons are sent towards the slits *one by one* so that at any given time there is only *one* electron traveling from the electron gun to the screen.

The electron leaves the source, shown in Fig. 3.7a, b, with the initial position of the electron denoted by s; it is then observed at the screen at position denoted by x. There are two possible paths from source to screen, labeled path 1 going through slit 1 and path 2 going through slit 2, and shown in Fig. 3.7b.

The experiment is performed *with* detectors 1 and 2, as shown in Fig. 3.7a and *without* these two detectors, as shown in Fig. 3.7b. In effect, with detectors 1 and 2 switched on, the path taken by the electron is known, whereas in the case without the detectors, the path information is not known.

Experiment with Detectors 1 and 2: No Interference

The experiment is shown in Fig. 3.7a; *both* slits 1 and 2 are open and detectors 1 and 2, at the back of the slits, record which slit the electron passes through. Since

it is known which slit the electron goes through, one can plot the following three distribution curves:

- Distribution curve P_1 for electrons that pass through slit 1
- Distribution curve P_2 for electrons that pass through slit 2
- Distribution curve P_D for electrons that pass through *either* slit 1 *or* 2

The experimental result is the following:

$$P_D = P_1 + P_2 \;\; ; \;\; \int \mathrm{d}x P_D(x) = 1 \tag{3.1}$$

The behavior of the electron *with detection* of which path is taken is exactly the same as one would obtain for bullets being shot through a metal screen with two openings. When the trajectory of the electron from source to screen is *experimentally observed*, the electron received at the screen detector came *either* through slit 1 *or* through slit 2. The two possible ways of getting to the screen are mutually exclusive, and these do not interact in any manner and do not generate any *interference* pattern.

The two-slit experiment has two possible electron paths, namely, path 1 and path 2. The quantum mechanical description of the two slit experiment is given by assigning a *probability amplitude* (a complex number) to each of the two paths. The probability amplitude is derived from the quantum state of the electron's quantum degree of freedom and is discussed (2.8) and in more detail in Sect. 8.3

The following is a notation for the probability amplitude for the different possibilities[5]:

ϕ_1 : the trans-empirical probability amplitude for determinate path 1

ϕ_2 : the trans-empirical probability for amplitude determinate path 2

ϕ : the trans-empirical general probability amplitude

The probability of finding the electron at the screen *with detection* of path taken is labeled by P_D and *without detection* of path taken is labeled by P_I.[6]

The experimental result for P_D is shown in Fig. 3.7a, and the result for P_I is shown in Fig. 3.7b. Note the experimental results obtained for P_D and P_I are *qualitatively different* and need an explanation.

[5]To connect with the notation of Chap. 8, one has

$$\phi_1 = \langle x|1\rangle\langle 1|s\rangle \;\; , \;\; \phi_2 = \langle x|2\rangle\langle 2|s\rangle .$$

[6]All measurements in quantum mechanics are of operators representing a physical quantity. "Measuring the position" of the electron is a shorthand for the more precise statement that measurement is mathematically represented by applying the position projection operator (of the electron) on the electron's state vector. See Sect. 9.2.

The explanation of quantum mechanics for the experiment with detectors 1 and 2 switched on is the following. The electron leaves the source s with *trans-empirical* probability amplitude ϕ. When the probability amplitude encounters the detectors at the barrier, a measurement is performed by either detector 1 or 2, depending at which slit the electron is detected. The trans-empirical probability amplitude ϕ *collapses* to its *empirical manifestation—either* $|\phi_1|^2$ *or* $|\phi_2|^2$, depending on whether it is detected at slit 1 or slit 2, respectively.

The electron arrives at the screen in an empirical state, and the probability of being observed at a point on the screen is given by either $|\phi_1|^2$ or $|\phi_2|^2$. The probability of finding the electron at the screen detector is the result of the two mutually exclusive possibilities and hence is given by their sum, namely,

$$P_D = |\phi_1|^2 + |\phi_2|^2 = P_1 + P_2 \quad : \quad \text{collapse at barrier} \tag{3.2}$$

and one obtains the result given in (3.1).

The result given in (3.1) and (3.2) is equivalent to the *classical result* obtained by shooting classical particles through the two slits. The reason being that observations made by detectors 1 and 2 at the two slits *determines* which slit the electron went through by collapsing the probability amplitude and in doing so causes the electron to take a determinate path, namely, going through *either* slit 1 *or* slit 2.

Having the information of which path is taken is equivalent to the classical description of a particle, since the classical particle always takes a unique path from its initial to final position.

Experiment *Without* Detectors 1 and 2: Trans-empirical

Consider now the same experiment as before but with detectors 1 and 2 at the barrier removed. The experiment is shown in Fig. 3.7b. The electrons are sent in one by one, and *no measurement* is made to determine which slit the electron goes through, and hence the path taken by the electron is no longer known. As shown in Fig. 3.7b, the electron can take two possible paths to reach the point x at the screen.

The electron leaves the source with trans-empirical probability amplitude ϕ. On crossing the slits (barrier), since the electron is *not* observed, the path taken by the electron is not known, and hence the path of the electron is trans-empirical and indeterminate.

One needs to decide as to what is the trans-empirical probability amplitude past the barrier. Since the propagation of the trans-empirical probability amplitude ϕ is determined by the linear Schrödinger equation, from (2.7), the probability amplitude obeys the *superposition principle*. The probability amplitudes for the two possible paths both obey the Schrödinger equation, and hence their linear sum also obeys the Schrödinger equation; the trans-empirical probability amplitude for arriving at the screen is given by quantum superposition, namely, *summing* the probability amplitude for the two determinate paths and yields

$$\phi = \phi_1 + \phi_2 \quad : \quad \text{trans-empirical probability amplitude at the screen}$$

Note that to obtain ϕ, one is not superposing material displacements of a medium as is the case for classical waves, but instead one is superposing trans-empirical probability amplitudes ϕ_1 and ϕ_2 for determinate paths to obtain the trans-empirical probability amplitude $\phi = \phi_1 + \phi_2$ for the electron taking an indeterminate path; ϕ is said to be wavelike since it is spread over space. Quantum superposition is qualitatively different from superposing, for example, water waves in which one is adding the physical displacement of the underlying water.

On reaching the screen, the measurement of the electron at the screen detector collapses the trans-empirical probability amplitude ϕ to its empirical manifestation $|\phi|^2$ and yields the empirical probability P_I. Hence

$$\begin{aligned} P_I = |\phi|^2 = |\phi_1 + \phi_2|^2 \quad &: \quad \text{collapse at the screen} \\ \Rightarrow P_I = |\phi_1|^2 + |\phi_2|^2 + \phi_1\phi_2^* + \phi_1^*\phi_2 \quad &; \quad \int \mathrm{d}x P_I(x) = 1 \end{aligned} \tag{3.3}$$

The probability amplitude for the electron with an indeterminate path has nonlocal information about the likelihood of occurrence everywhere in space. In particular, for the two-slit experiment, there are nodal points (minimas) of interference pattern indicated by X_m in Fig. 3.7b, points for which the probability amplitude is zero, yielding zero likelihood of that the electron will be detected at those points. Since the electrons are sent in one by one, each electron, regardless of where on the screen it is detected, has the information about the entire screen since no electron ever hits the nodal points on the screen.

The interference pattern P_I shown in Fig. 3.7b has been verified by many experiments and shows that on repeatedly sending in the electrons—*sent in one by one*—and detecting the position of the electrons arriving at the screen, results in building up, step-by-step, the interference pattern given by P_I. The probability amplitude at the screen is $\phi = \phi_1 + \phi_2$ and shows that when the path taken by the electron is *not detected*, the electron's path is indeterminate and trans-empirical, showing interference.[7]

The trans-empirical superposition is lost in Fig. 3.7a because one *detects* the passage of the electron by, say, shining light on the electron as it passes through the slit. The shining of light is a measurement process that causes a transition by collapsing the trans-empirical probability amplitude before the barrier, namely, $\phi = \phi_1 + \phi_2$, to the empirical *probability* given by either $|\phi_1|^2$ or $|\phi_2|^2$ after the barrier—depending at which slit the electron is observed. The empirical expression $|\phi_1|^2$ or $|\phi_2|^2$ is said to be particle like since it implies that the electron is following a definite trajectory. In other words, when the electron's path is measured, the nonlocal probability amplitude collapses to an (localized) empirically observed determinate state that is particle like.

[7] It is important to note, as discussed in the next section, that the electron is interfering with *itself*—a completely nonclassical and enigmatic phenomenon.

A cardinal point to note is the crucial role of *measurement* in producing two *qualitatively different* results—namely, of P_D *with* the detection of the electron's path and of P_I *without* detection of its path. Measurement causes a transition of the electron from its trans-empirical state to its empirical manifestation, and the two-slit experiment shows this difference in a stark and clear manner.

In conclusion, the two-slit experiment shows that both quantum states and quantum paths display empirical and trans-empirical behavior. Furthermore, measurement has a central role in determining whether the electron behaves as an empirical entity or as a trans-empirical state. When it is not observed, *the electron evolves on indeterminate paths and displays wavelike behavior* that reflects its trans-empirical form, and when the electron is observed, it is in an empirical condition and behaves like a classical particle.

3.8 Trans-empirical Interpretation of Two-Slit Experiment

The two-slit experiment is interpreted in the framework of trans-empirical and empirical aspects of the quantum entity to clarify the role played by the trans-empirical probability amplitude.

Consider the case when no measurement is made to determine the path taken by the electron for the two-slit experiment. P_I, given in Fig. 3.7b, is the *probability* of finding a particle at a certain point on the screen when a measurement is *not* made. Consider the following term in (3.3), namely,

$$\phi_1^* \phi_2 + \phi_1 \phi_2^* \tag{3.4}$$

that contributes to P_I; how are we to interpret this term?

The expression $\phi_1^* \phi_2$ (the other term in (3.4) is similar to this) is a *cross-term* of the probability amplitude for electron going through slit 1, represented by ϕ_1 with the probability amplitude ϕ_2 representing the electron going through slit 2.

One *interpretation* of the term $\phi_1^* \phi_2$ is that when the trajectory of the electron is *not experimentally observed*, the *single* electron's path is indeterminate; the electron received at the screen detector simultaneously comes through *both* paths, in order words, through both slit 1 *and* through slit 2, and hence, the electron interferes with *itself*! Clearly, one can never directly observe a *single* electron simultaneously going through *both* slits, or equivalently, of having an indeterminate path that requires the single electron to simultaneously exist at two different points.

Although our discussion was in the context of the two-slit experiment, it can be shown that *all* quantum paths and entities have an empirical and a trans-empirical component. When the electron is *not observed*, it is a trans-empirical entity that is indeterminate: The single electron *simultaneously* exists at more than one point; in particular, the time evolution of the electron is indeterminate, evolving along indeterminate paths such that it goes through *both slits* simultaneously.

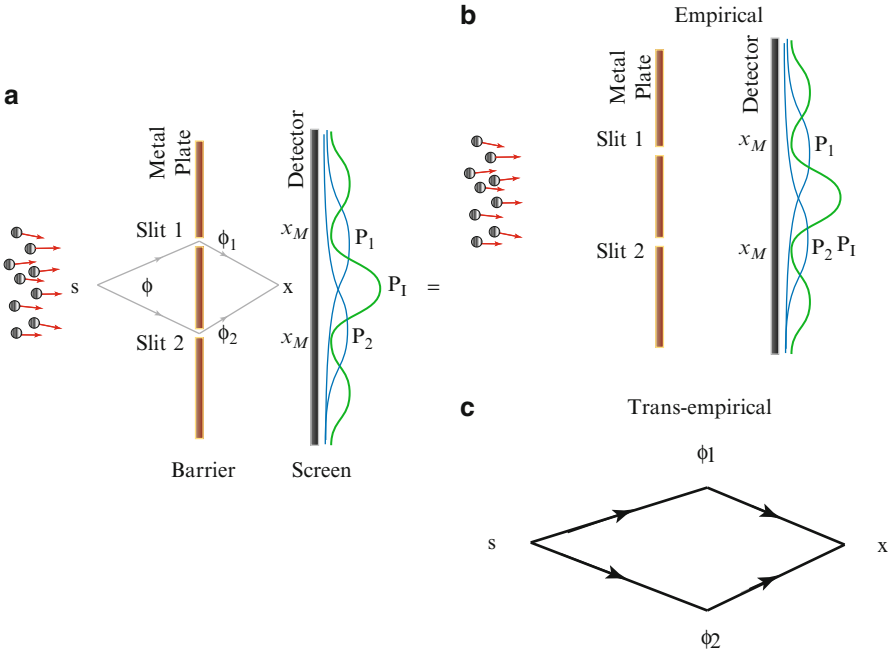


Fig. 3.8 (**a**) Two slits with the path taken is *not known*. (**b**) The empirically observable aspects of the experiment—electrons leaving the source and being detected at the screen. (**c**) The electron's unobservable trans-empirical paths, resulting in quantum superposition (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

The case when the path taken by the electron travels is not known is shown in Fig. 3.8a. The *empirical* part of the experiment is observing the electron leaving the source and arriving at the screen, as shown in Fig. 3.8b. The two paths in Fig. 3.8c are drawn with the understanding that these paths are trans-empirical and can never be experimentally observed; Fig. 3.8c is the *trans-empirical* aspect of the two-slit experiment.

3.9 The Trans-empirical Quantum Principle

From an analysis of the quantum entity, it is seen that quantum paths and state vectors both exhibit two inseparable but distinct forms of behavior, namely, the empirical and trans-empirical, and which constitute two distinct domains of Nature. These two domains themselves are the reflection of various aspects of the quantum entity and are expressed in the trans-empirical quantum principle .

The Copenhagen interpretation is silent on the nature of the quantum entity when it is not being observed. The trans-empirical quantum principle is a proposal that *extends* the Copenhagen interpretation by postulating principles about the mode of

existence of the quantum entity. The extension of the Copenhagen interpretation, given below, is made with the aim of making the inner consistency of quantum mechanics self-evident:

1. The trans-empirical quantum principle states that all entities in Nature have two forms of existence: empirical and trans-empirical.
2. The foundation of the quantum entity is its degree of freedom that is intrinsically trans-empirical and inherently indeterminate, uncertain, and in principle can never be observed. The quantum state of the quantum entity connects the degree of freedom to experimental observations, which are mathematically represented by operators.
3. The quantum state is trans-empirical when it is not observed. The time evolution of the quantum state, determined by the Schrödinger equation, takes place entirely in the trans-empirical domain.
4. The act of observation causes the quantum state to make a transition to a determinate, certain, and empirical condition; the transition is effected by the collapse of the trans-empirical quantum state to its empirical manifestation in the empirical domain. Preparation of a quantum state creates a reverse transition from the empirical preparation of the quantum state to its trans-empirical form of existence.
5. The two domains of the quantum entity, the empirical domain where its empirical manifestation takes place and the trans-empirical domain where its state vector exists, constitute an inseparable pair and form an integral and composite whole. Taken together, the empirical and trans-empirical domains are the appropriate framework for a complete description of the quantum entity.

3.10 Does the Quantum State $\psi(t, \mathcal{F})$ "Exist"?

The symbolic approach for understanding Nature has been championed by no less a quantum theorist than Dirac, who states the following: "The symbolic method .. seems to go more deeply into the nature of things" [10].

Trans-empirical entities cannot be directly observed but, instead, can only be represented by symbols. In particular, the trans-empirical quantum degree of freedom is represented by $\mathcal{F}$, and its state vector is represented by the symbol $\psi(t, \mathcal{F})$; these symbols *go more deeply into the nature of things* by representing the inner trans-empirical nature of things.

The point of view adopted in this book is that trans-empirical constructs, represented by symbols such as $\mathcal{F}$ and $\psi(t, \mathcal{F})$, are not "real" in the sense that they are not empirical entities, but the trans-empirical constructs nevertheless do "exist." $\mathcal{F}$ and $\psi(t, \mathcal{F})$ have the same *ontological status* and have the same quality of being as an empirical quantity such $E_\psi[\mathcal{O}]$, namely, the empirical average of a physical property represented by the operator $\mathcal{O}(\mathcal{F})$ acting on the state vector $\psi(t, \mathcal{F})$ of the quantum degree of freedom $\mathcal{F}$.

It is worth noting that for almost a century, Schrödinger's equation—based on the trans-empirical construct of the quantum state ψ—has successfully faced a vast

range of experimental tests. In spite of all its experimental success, many physicists have an agnostic view that does not attribute any "existence" to the quantum state, and consider $\psi(t, \mathcal{F})$ to be merely a mathematical tool, not having the same quality of "a real existence" as an experimental observation.

In fact, one of the prime distinctions between the empirically observed properties and trans-empirical form of the quantum state lies in the fact that the Schrödinger equation determines the evolution of *only* the trans-empirical state $\psi(t, \mathcal{F})$, with the empirical effects of the quantum state being the result of measurement, and hence *not determined* by the Schrödinger equation.

There is an agnostic view—which is of course mathematically consistent—that one need not consider $\psi(t, \mathcal{F})$ to "exist" in order to carry out all the mathematical calculations of quantum mechanics or to do experiments to verify these calculations. The agnostic view is that if one takes the state vector $\psi(t, \mathcal{F})$ to "exist", then one will apparently be stuck with paradoxes and contradictions that are thought to necessarily follow.

This agnostic view, however, misses a number of crucial points. To start with, there are *no internal contradictions* or inconsistencies in assuming that the degree of freedom and the quantum state have the same level of ontological existence as empirical observations, and this will be demonstrated by many derivations in the ensuing chapters.

Moreover, similar to the observed outcomes of a quantum state that conserve all the conserved quantities, even when the state is *not observed*, the time evolution of the degree of freedom continues to conserve—in a trans-empirical and probabilistic sense—all the empirically conserved quantities, including energy and momentum. Hence, even though the trans-empirical quantum degree of freedom cannot be directly observed, its quantum state described by the trans-empirical state vector $\psi(t, \mathcal{F})$ has an existence, a "being", that is "real enough" to continue to conserve all the experimentally conserved quantities.

Furthermore, there is another even more important possibility that the agnostic view excludes. If there indeed exists a trans-empirical domain in Nature—and there are no empirical grounds for ruling out this possibility—then the only way this domain would be accessible to human cognition would be through the means of mental and theoretical constructs similar to the symbolic constructions of quantum mechanics.

It is only by pursuing the analysis of a possible trans-empirical characteristic of Nature that the existence of such a domain can be studied and investigated. And by gaining a greater and deeper understanding of the symbols representing the trans-empirical domain, appropriate experiments could be designed to discover new and hitherto unknown properties and features of the trans-empirical domain that could have an empirical manifestation.

Noteworthy 3.1: 'Wave- particle duality'

Some authors loosely state that the quantum system "sometimes behaves like a particle and sometimes behaves like a wave." It should be kept in mind that the

concepts of "particle" and "wave" in particle-wave duality do not *both* refer to a classical particle and a classical wave.

For a quantum particle, say, an electron, the "particle" refers to an empirically observed classical behavior, such as the detection of the electron at a definite position in space, whereas the "wave" refers to the trans-empirical quantum state vector of the electron.

In contrast, for a photon, the "particle" refers to a quantum excitation of the quantized electromagnetic field and, when not observed, the photon is in a trans-empirical state—being spread over all space.[8] Only when the photon is absorbed or emitted by an electron does it behave like a particle carrying its own energy and momentum. Furthermore, the "wave" that is associated with the photon is the classical and empirically observable electromagnetic wave, which is a reflection of the *field equation* of the photon field. The empirical electromagnetic wave, unlike the case for the electron, has no relation to the trans-empirical "wavelike behavior" of the photon field's quantum state vector.

The wave-particle duality refers to entities that are ontologically distinct, with one being an empirical entity and the other being a trans-empirical construct.

The confusing and often misleading terminology of "wave-particle duality" is completely eschewed in subsequent chapters; instead, the terms trans-empirical and empirical are employed to demonstrate the consistency, internal structure, and inner workings of quantum mechanics.

3.11 Summary

It is proposed that our conception of a quantum entity can be based on a view of Nature that extends the realm of empirical reality to include the existence of another domain, namely, the trans-empirical domain, and the connecting interface of these two domains. Figure 3.2 graphically shows the various structures that are required for describing a quantum entity.

The quantum entity exists in the form of an indeterminate trans-empirical degree of freedom $\mathcal{F}$ that is described by a quantum state $\psi(t,\mathcal{F})$, which in turn is an element of the state space $\mathcal{V}(\mathcal{F})$. The quantum state $\psi(t,\mathcal{F})$ contains all the possible information that can be experimentally extracted from the quantum degree of freedom $\mathcal{F}$. Knowledge of the existence of the quantum entity, in particular of its trans-empirical degree of freedom and quantum state, can only be inferred from the *mathematical symbols* of quantum mechanics that, in essence, require conscious reasoning; in particular, one can never *directly observe* a trans-empirical entity using any experimental device.

The process of measurement causes a transition from the trans-empirical form of the quantum state to its empirical manifestation. The preparation of a quantum state

[8]The photon is nothing like a classical particle since special relativity requires it to be massless.

creates a reverse transition from the empirical preparation of the quantum state to its trans-empirical form. Due to quantum indeterminacy, quantum states prepared in an identical manner, when subjected to the *same* experiment, will, in general, lead to *many different* observed results.

The quantum state straddles the empirical and trans-empirical domains: When it is *not observed*, it is trans-empirical and described by $\psi(t, \mathcal{F})$; every time the value of an operator $\mathcal{O}(\mathcal{F})$, representing physical quantities, is experimentally determined for the quantum state, the state $\psi(t, \mathcal{F})$ collapses to a particular, determinate, and specific state. The process of measurements is represented by applying operators $\mathcal{O}(\mathcal{F})$ on the quantum state, and repeated experiments yield its empirical expectation value—given by $E_{\psi}[\mathcal{O}(\mathcal{F})]$.

One can question whether the existence of the trans-empirical quantum state can be inferred from the failure of classical physics, or is the postulate that the trans-empirical domain exists due to the mathematical symbols of quantum mechanics? It seems that the empirically observed phenomenon and the mathematical symbols of quantum mechanics both reinforce each other in leading one to conclude that a trans-empirical quantum domain, in fact, does exist.

The interplay of the empirical and trans-empirical domains is very finely balanced and can naively give rise to many apparent inconsistencies—as was evidenced by the debates at the inception of quantum mechanics. The fact that quantum mechanics is not inconsistent is a nontrivial result and in many cases needs to be theoretically and experimentally demonstrated. It can be said that due to its subtle trans-empirical foundations, *quantum mechanics runs on the edge of consistency*—in many cases avoiding inconsistencies in strange, novel, and unexpected ways.

Degree of Freedom $\mathcal{F}$; State Space $\mathcal{V}$

4

In classical mechanics, the description of an entity, say, a billiard ball, requires that, at time t, one specifies its position and velocity—namely, $x(t)$ and $v(t)$. In general, to describe a classical system, one needs to assign unique values to all its dynamical variables.

As discussed in Chap. 2, a physical entity in quantum mechanics consists of its degree of freedom $\mathcal{F}$ together with the entity's state vector $\psi(t,\mathcal{F})$. The quantum generalization of dynamical variables of classical mechanics is the *degree of freedom*, which is a space $\mathcal{F}$ constituted by *all possible values* of the degree of freedom.

A quantum entity is inherently indeterminate and hence requires a mathematical framework appropriate for describing this indeterminacy. A quantum description is given by specifying the state vector $\psi(t,\mathcal{F})$, which encodes all possible observable properties of the degree of freedom.

The state vector $\psi(t,\mathcal{F})$ is an element of the state space $\mathcal{V}(\mathcal{F})$, which is a linear vector state space. As discussed in Chap. 3, the degree of freedom $\mathcal{F}$ is trans-empirical, whereas the state vector $\psi(t,\mathcal{F})$ straddles the trans-empirical and empirical domains.

Realistic quantum systems come in great variety and diversity, and the key concepts in their description are now introduced. This chapter studies the mathematical structure of the degree of freedom $\mathcal{F}$ and its associated linear vector space $\mathcal{V}(\mathcal{F})$. The following are this chapter's main topics:

- The relation of state space to the empirically observed quantum phenomena is discussed.
- The simplest possible degree of freedom is binary, having two possible values is discussed in great detail, and its state space is derived.
- The $(2N+1)$-state degree of freedom is discussed, and the "position" degree of freedom x having continuous values is obtained as the limiting case of $N \to \infty$.
- The coordinate and momentum basis states for the continuous degree of freedom are discussed.

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8_4, © Springer Science+Business Media New York 2013

- The general notion of a linear vector state space is lastly discussed in complete generality, and the notion of the Hilbert space is defined.[1]

4.1 Dirac's Formulation of the Quantum State

Dirac's notation is followed in the book and discussed in detail in Sect. 4.9.

- The term *quantum state* refers to all the descriptions of the quantum entity and includes the terms state, state vector, and state function.
- The term state or *state vector* refers to the quantum state considered as a vector in state space $\mathcal{V}$, usually denoted by $\psi(t,\mathcal{F})$.
- In Dirac's bracket notation, a state vector is denoted by $|\psi(t,\mathcal{F})\rangle$ or $|\psi\rangle$ in short and is called a *ket* vector.
- The dual to the ket vector is denoted by $\langle\psi(t,\mathcal{F})|$ or $\langle\psi|$ in brief and is called a *bra* vector.
- The scalar product of two state vectors χ,ψ is a complex number $\in \mathbb{C}$ and is denoted by the full bracket, namely, $\langle\chi|\psi\rangle = \langle\psi|\chi\rangle^*$, where * denotes complex conjugation.
- The term *state function* refers to the *components* of the state vector and is denoted by $\langle x|\psi(t,\mathcal{F})\rangle \equiv \langle x|\psi_t\rangle \equiv \psi(t,x)$, where $x \in \mathcal{F}$, namely, x is an element of the degree of freedom $\mathcal{F}$.

For degrees of freedom taking discrete values, Dirac's bra and ket vectors are nothing except the row and column vectors of a finite-dimensional linear vector space, with the bracket of two state vectors being the usual scalar product of two vectors.

When the degree of freedom becomes continuous, Dirac's notation carries over into functional analysis and allows for studying questions of the convergence of infinite sequences of state vectors that goes beyond linear algebra.

4.2 State Space and Experiment

There is a view that "Quantum phenomena do not occur in a Hilbert [state] space. They occur in a laboratory" [1]. The statement revolves around the words "phenomena" and "occur." The dictionary meaning of *phenomenon* is *an occurrence, circumstance, or fact that is perceptible by the senses*. The statement is about "quantum phenomena," and given that phenomena must be perceptible, it clearly refers to the detection and measurement of the properties of a quantum entity.

This view is a statement about the empirical aspect of quantum mechanics and expresses the fact that $\psi(t,\mathcal{F})$, the trans-empirical state vector of a quantum degree of freedom, has an empirical manifestation in the laboratory, in an experimental device.

[1] Hilbert space is a special case of state space and is discussed in Sect. 4.10.

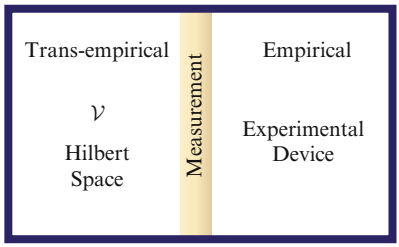


Fig. 4.1 The experimental device is in the empirical domain; the Hilbert space of states $\mathcal{V}$ is trans-empirical (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

The statement about quantum phenomena occurring in a laboratory is technically correct but is only half of the story since it *does not* address the question as to what is the form of existence of the quantum entity when it is *not observed*—namely, when the entity's state vector is not actualized into a tangible phenomenon.

When the quantum entity is *not* being observed, its state vector $\psi(t,\mathcal{F})$ exists in state space $\mathcal{V}(\mathcal{F})$. All elements of the state space are trans-empirical, and hence, state space $\mathcal{V}(\mathcal{F})$ is trans-empirical, with its empirical manifestations being observed in the laboratory. Moreover, as shown in Fig. 4.1 and discussed in Sect. 3.5, the quantum state $\psi(t,\mathcal{F})$ straddles *both* the trans-empirical and empirical domains that are *inseparable* and connected by the process of measurement.

In particular, the degree of freedom for position is the real line $\mathcal{F} = \Re$, and the state vector $\psi(t,\Re)$ of the quantum entity exists in state space. When the degree of freedom $\Re$ is subjected to a measurement in the laboratory by an experimental device that is the projection operator for *specific* position x, discussed later in Sect. 9.2, the probability of registering a signal for that projection operator is given by $|\psi(t,x)|^2$.

Hence, the form of existence of the state vector is quite different in the two domains, being $\psi(t,\Re)$ in the trans-empirical domain and $|\psi(t,x)|^2$ in the empirical domain.

The time evolution of $\psi(t,\mathcal{F})$ is entirely trans-empirical and takes place in state space—being governed by the Schrödinger equation that simultaneously evolves all possible values of the degree of freedom—and shown as a wavy line in Fig. 4.2a. On the other hand, as shown in Fig. 4.2b, when an experiment is carried out—shown by an "**x**"—the state vector $\psi(t,\Re)$ "collapses" to only *one* possible value of the quantum state vector, namely, $|\psi(t,x)|^2$ and indicated in the figure by $|\psi|^2$, and which is empirically observed in the laboratory.

Once the experiment is over, then—after a characteristic time determined by the nature of the experiment—the state vector goes back to its trans-empirical evolution. The interplay of the observed and unobservable aspects of a quantum state vector is shown in Fig. 4.2b.

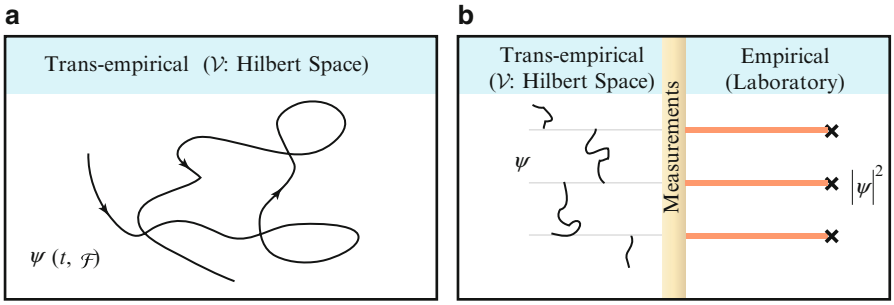


Fig. 4.2 (**a**) The trans-empirical state vector $\psi(t, \mathcal{F})$ evolving in state space. (**b**) The evolving state vector ψ that is measured many times in the laboratory and "collapses" to $|\psi|^2$ after each measurement (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

4.3 Quantum Degree of Freedom $\mathcal{F}$

By its very definition the degree of freedom is trans-empirical and hence cannot be directly observed by any experiment; the conception and construction of the degree of freedom are entirely theoretical. In fact, one of the main objectives of theoretical physics is to identify the degree(s) of freedom of a particular phenomenon.

The degree of freedom, $\mathcal{F}$, comes in a vast variety of forms depending on the nature of the quantum entity. In general, a degree of freedom can be either a discrete set or a continuous space. The simplest degree of freedom has two discrete values; on the other extreme, a single quantum entity such as a quantum field is constituted by infinitely many continuous degrees of freedom.

An example of a *discrete degree of freedom* is the spin degree of freedom. The space $\mathcal{F}$ for spin ℓ degree of freedom consists of $2\ell+1$ number of points, namely, $\mathcal{F} = \{-\ell, -\ell+1, \ldots, -1, 0, 1, \ldots, \ell-1, \ell\}$, and is shown in Fig. 4.3.

An example of a *continuous degree of freedom* is a single quantum particle in a finite three-dimensional box, denoted by B; the particle can be anywhere inside the box and hence the degree of freedom space $\mathcal{F}^1 = B$ since it consists of all points inside B and is shown in Fig. 4.4. Consider a quantum entity that consists of N particles in the box B; each degree of freedom consists of the space B, and hence, the space for N degrees of freedom is given by $\mathcal{F}^N = B \times B \times B \times \cdots \times B = B^N$. If the box is infinitely large, the space constituted by the degree of freedom is $\mathcal{F}^N = \Re^{3N}$, namely, a $3N$-dimensional Euclidean space.

Clearly, the degree of freedom space $\mathcal{F}$ is not a physical space but rather another distinct space that is indirectly linked to physical space by the process of measurements performed on the quantum entity. The state space $\mathcal{V}(\mathcal{F})$ of the quantum entity is built on $\mathcal{F}$. The quantum state vector exists in $\mathcal{V}(\mathcal{F})$, and this is where its time evolution takes place. The relationship of physical space to state space $\mathcal{V}(\mathcal{F})$ is that all processes of measurements and observations take place in physical space by "collapsing" the state vector, as shown in Fig. 4.2b, whereas the unobserved quantum state evolves in state space, as shown in Fig. 4.2a.

Fig. 4.3 Discrete degree of freedom $\mathcal{F}$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

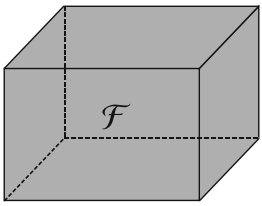


Fig. 4.4 The continuum degree of freedom $\mathcal{F}$ of a single quantum particle in a three-dimensional box B (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

4.4 Binary Degree of Freedom and State Space

The simplest quantum degree of freedom $\mathcal{F}_2$ is binary, being a double-valued system that has only two possible values, namely, $\mathcal{F}_2 = \{-1, +1\}$, and is studied in some detail to motivate the results for more complicated degrees of freedom.

The electron has an *intrinsic* angular momentum called *spin*; the (z-component of the) spin of the electron can either point up or down and hence forms a double-valued system. Another example of a two-state system is a particle that can have only two possible positions.

In quantum information theory, the binary degree of freedom is called a qubit since it is the quantum generalization of the classical "bit" of information science.

A quantum entity having a binary degree of freedom is mathematically described by specifying the state vector $|\psi\rangle$ for its two distinct possible states and which belongs to the state space $\mathcal{V}_2(\mathcal{F}_2)$. The two distinct states should be "orthogonal" to each other, in the sense that being in one state is completely different from being in the other state—and are sometimes called "up" or "down" states.

$$|\psi_{\text{up}}\rangle = |u\rangle; \;\; |\psi_{\text{down}}\rangle = |d\rangle$$

The state vector of a double-valued system can be represented by a collection of complex-valued two-dimensional vectors.

One should note that the state space $\mathcal{V}_2(\mathcal{F}_2)$ has no relation with a physical space, but rather should be viewed as a mathematical construction for describing the state vector of a two-state quantum system. The symbols $|u\rangle$ (for up) and $|d\rangle$ (for down) are used to free one from thinking of these vectors as referring to physical space, which are usually labeled by x- and y-axes.

One possible representation of the quantum basis states is to assign $|u\rangle = e_x$ and $|d\rangle = e_y$, where e_x, e_y are the basis vectors of the two-dimensional vector space, as shown in Fig. 4.5. The distinct basis state vectors and their dual basis vectors have the following representation:

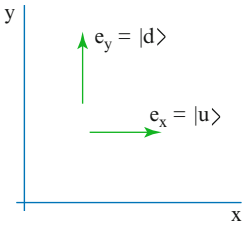


Fig. 4.5 Basis vectors for the two-dimensional state space (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

$$|u\rangle = e_x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \; ; \; |d\rangle = e_y = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{4.1}$$
$$\langle u| = e_x^{\mathrm{T}} = \begin{bmatrix} 1 & 0 \end{bmatrix} \; ; \; \langle d| = e_y^{\mathrm{T}} = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

The *inner product* of two vectors is defined as follows:

$$\langle d|u\rangle = \begin{bmatrix} 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0 \tag{4.2}$$

The completeness equation, which expresses the fact that a collection of basis vectors forms a compete basis of a linear vector space and is given in (6.2), is realized for the two-dimensional state space by the *outer product* of two vector as follows:

$$\begin{aligned} |u\rangle\langle u| + |d\rangle\langle d| &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbb{I}_2 \end{aligned} \tag{4.3}$$

State Space $\mathcal{V}_2$ of $\mathcal{F}_2$

A spin pointing purely up $|u\rangle$ or down $|d\rangle$ is the closest that, in quantum mechanics, one can get to a classical object since, on being observed, it is certain to be in either up or down respectively. The general state vector of a quantum mechanical binary (two-state) system is more subtle: We can *superpose* two states, as discussed in Sect. 3.7, and obtain a trans-empirical state that simultaneously points up and down, but with only a certain *likelihood*.[2] Superposing a quantum spin pointing up with one pointing down yields a state vector given by

$$|\psi\rangle = \alpha|u\rangle + \beta|d\rangle \; ; \; \langle\psi| = \alpha^*\langle u| + \beta^*\langle d|$$

[2]The concept of quantum superposition is discussed, in-depth, in Chap. 8.

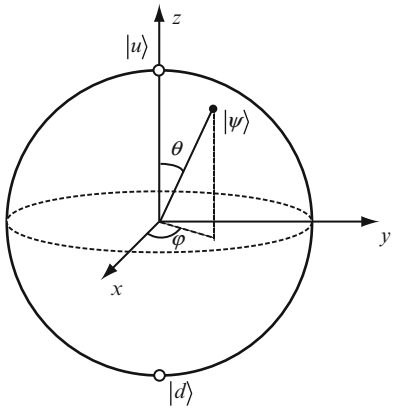


Fig. 4.6 Bloch sphere for the two-dimensional state space (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

$$\langle \psi | \psi \rangle = |\psi|^2 = |\alpha|^2 + |\beta|^2 \tag{4.4}$$

where α and β are *complex numbers*.

For a probabilistic interpretation of $|\psi\rangle$, the total probability has to be unity, and one obtains unit the normalization of the state vector, namely, that

$$\langle \psi | \psi \rangle = |\psi|^2 = |\alpha|^2 + |\beta|^2 = 1 \tag{4.5}$$

The coefficients now have the following physical interpretation:

$$|\alpha|^2 = \text{Probability that the spin is pointing up}$$
$$|\beta|^2 = \text{Probability that the spin is pointing down}$$

The coefficients α, β parametrize the state space $\mathcal{V}_2$ of the two-state system. Note that state vector $|\psi\rangle$ given in (4.4) cannot be in general represented by a two-dimensional unit vector in $\Re^2$, since only vectors with real coefficients can be drawn in Fig. 4.5. In contrast, α, β are complex numbers, and from (4.5), it can be seen that their possible values span out a three-dimensional sphere $\mathcal{S}^3$.

There is, however, a redundancy in this description since the coefficients α, β can both be rescaled by the same constant phase and yield

$$|\tilde{\psi}\rangle = e^{i\phi} |\psi\rangle \quad \Rightarrow \quad \langle \tilde{\psi} | \tilde{\psi} \rangle = \langle \psi | \psi \rangle \tag{4.6}$$

Hence, both state vectors $|\psi\rangle$ and $|\tilde{\psi}\rangle$ provide the same description; in other words, states linked by a global pure phase, namely, $|\psi\rangle \to e^{i\phi} |\psi\rangle$, are *equivalent*.[3]

[3] ϕ being a pure phase means that it is real.

The phase ϕ forms a space that is isomorphic to a circle $\mathcal{S}^1$. Hence, one needs to "divide out" $\mathcal{S}^3$ by $\mathcal{S}^1$ to form equivalence classes of state vectors and yields $\mathcal{S}^3/\mathcal{S}^1 = \mathcal{S}^2$, which is a two-dimensional sphere.[4]

In summary, all possible distinct state vectors of a binary quantum system are parametrized by the angles of the Bloch sphere $\mathcal{S}^2$; each point on the Bloch sphere corresponds to a (unique) state vector of the two-state system state space $\mathcal{V}_2$,

Bloch Sphere

The state space $\mathcal{V}_2$ of a two-state system is isomorphic to the Bloch sphere $\mathcal{S}^2$, shown in Fig. 4.6. The Bloch sphere allows for a specific representation of the most general state vector for the two-state system.

Consider the spherical polar coordinates θ, ϕ shown in Fig. 4.6. An arbitrary normalized two-state ket vector $|\psi\rangle$, shown in Fig. 4.6, is given by the following mapping of the Bloch sphere into two-dimensional state space:

$$|\psi(\theta,\phi)\rangle = \cos\left(\frac{\theta}{2}\right)\begin{bmatrix}1\\0\end{bmatrix} + \mathrm{e}^{\mathrm{i}\phi}\sin\left(\frac{\theta}{2}\right)\begin{bmatrix}0\\1\end{bmatrix}$$

$$0 \le \theta \le \pi \;\; ; \;\; 0 \le \phi \le 2\pi \tag{4.7}$$

The basis states have been chosen to reflect the fact that $\theta = 0$ points in the "up" direction and $\theta = \pi$ points in the "down" direction and shown in Fig. 4.6; more precisely, (4.7) yields the following:

$$|\psi(0,\phi)\rangle = \begin{bmatrix}1\\0\end{bmatrix} = |u\rangle$$

$$|\psi(\pi,\phi)\rangle = \mathrm{e}^{\mathrm{i}\phi}\begin{bmatrix}0\\1\end{bmatrix} \equiv \begin{bmatrix}0\\1\end{bmatrix} = |d\rangle \tag{4.8}$$

Note the ambiguity in the definition of the state vector $|\psi(\theta,\phi)\rangle$ for $\theta = \pi$ is automatically removed for quantum states since states differing by a pure phase are equivalent, as shown in (4.6).

The presence of the phase $\mathrm{e}^{\mathrm{i}\phi}$ in (4.7) shows that vector $|\psi(\theta,\phi)\rangle$ does not correspond to a two-dimensional unit vector in Euclidean space, which has only real-valued components. In fact, every state vector $|\psi(\theta,\phi)\rangle$ corresponds a (unique) *three-dimensional* Euclidean unit vector, as shown in Fig. 4.6, and is uniquely

[4]To prove this result, one needs to construct $\mathcal{S}^3$ by a Hopf fibration, using the mathematics of fiber bundles, by fibrating the base manifold $\mathcal{S}^2$ with fibers given by $\mathcal{S}^1$.

parametrized by the coordinates $\theta \in [0,\pi]$ and $\phi \in [0,2\pi]$. In summary, every point on the surface of the Bloch sphere $\mathcal{S}^2$ corresponds to a unique state vector in $\mathcal{V}_2$.[5]

4.5 Degree of Freedom $\mathcal{F}_{(2N+1)}$: State Space $\mathcal{V}_{(2N+1)}$

The $\mathcal{F}_{(2N+1)}$ degree of freedom has $(2N+1)$ possible values; the notation is chosen to make a systematic transition to the case of $N=\infty$. The basis states of state space $\mathcal{V}_{(2N+1)}$ are chosen to be basis vectors of $\mathfrak{R}^{(2N+1)}$. A convenient basis is state vector $|n\rangle$ (column vector) and its dual state vector basis $\langle n|$ (row vector), given by

$$|n\rangle = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \;;\; \langle n| = \underbrace{\begin{bmatrix} 0 \cdots 1\ 0 \cdots \end{bmatrix}}_{n\text{th position}} \tag{4.9}$$

$$\begin{aligned} \langle n|m\rangle &= \delta_{n-m} \\ &\equiv \begin{cases} 1 & n=m \\ 0 & n\neq m \end{cases} : m,n = 0,\pm 1,\ldots,\pm N \end{aligned} \tag{4.10}$$

Note δ_{n-m} is the Kronecker delta function and has only integer arguments. An arbitrary quantum state is given by

$$|\psi\rangle = \sum_{n=-N}^{+N} c_n |n\rangle$$

where c_n's are complex numbers such that

$$\langle\psi|\psi\rangle = \sum_{n=-N}^{+N} |c_n|^2 = 1$$

Hence, the $\mathcal{V}_{(2N+1)}$-state space, similar to the case of the two-state system, can be shown to be isomorphic to $S^{(2N+1)}/S^1 \simeq \mathcal{V}_{(2N+1)}$.

Similar to (4.3), for state space $\mathcal{V}_{(2N+1)}$, the outer product of $|n\rangle$ with the dual $\langle n|$, namely, $|n\rangle\langle n|$, is given by the following:

[5] The ambiguity for the special value of $\theta=\pi$ is removed by (4.8).

$$|n\rangle\langle n| = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \cdots 1\ 0 \cdots 0 \end{bmatrix} = \begin{bmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

Hence, the completeness equation is given by

$$\sum_{n=-N}^{+N} |n\rangle\langle n| = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix} = \mathbb{I}_{2N+1} \tag{4.11}$$

Noteworthy 4.1: Dirac delta function

The Dirac delta function is required in the study of continuum degrees of freedom, and some of its essential properties are reviewed. Dirac delta functions are not ordinary but rather are generalized functions also called distributions. In essence, the Dirac delta function is the continuum generalization of the discrete Kronecker delta function given in (4.10).

Consider a continuous line labeled by coordinate x such that $-\infty \leq x \leq +\infty$, and let $f(x)$ be an infinitely differentiable function. The Dirac delta function, denoted by $\delta(x-a)$, is defined by the following:

$$\begin{aligned} &\delta(x-y) = \delta(y-x) \equiv \begin{cases} \infty, & x = y \\ 0, & x \neq y \end{cases} \\ &\delta(c(x-y)) = \frac{1}{|c|}\delta(x-y) \\ &\int_{-\infty}^{+\infty} \mathrm{d}x f(x)\delta(x-y) = f(y) \\ &\int_{-\infty}^{+\infty} \mathrm{d}x f(x)\frac{\mathrm{d}^n}{\mathrm{d}x^n}\delta(x-a) = (-1)^n \frac{\mathrm{d}^n}{\mathrm{d}x^n} f(x)|_{x=a} \end{aligned} \tag{4.12}$$

To make the connection with the discrete Kronecker delta function given in (4.10), consider the discretization of the continuous line into a discrete lattice with spacing a. As shown in Fig. 4.7, the continuous degree of freedom x, $-\infty \leq x \leq +\infty$, takes only discrete values at points $x = na$ with $n = 0, \pm 1, \pm 2, \ldots$. The discretization of

(4.12), for $x = na$ and $y = ma$, yields the following:

$$a \sum_{n=-\infty}^{\infty} f(na)\delta(x-y) = f(ma)$$

$$\Rightarrow \delta(x-y) = \lim_{a\to 0} \frac{1}{a}\delta_{m-n} \tag{4.13}$$

Note from above that the Dirac δ-function has dimension of $1/a$.
A representation of the δ-function based on the Gaussian distribution is

$$\delta(x-a) = \lim_{\sigma\to 0} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-a)^2\right\}$$

4.6 Continuous Degree of Freedom

A quantum particle can be found by the position projection operators at any point of space; to simplify the discussion, suppose the particle is moving in a one-dimensional space and can be found on the real line $\Re$, namely, at any point $x \in [-\infty, +\infty]$. Hence, the *degree of freedom* is $\mathcal{F} = \Re$, and specific values of the degree of freedom x are a real *continuous variable*. Since there are infinitely many points on the real line, the quantum particle's degree of freedom has *infinitely* many possible outcomes.

As shown in Fig. 4.7, let the continuous degree of freedom x, $-\infty \le x \le +\infty$, take only discrete values at points $x = na$ with lattice spacing a and with $n = 0, \pm 1, \pm 2, \ldots$; in other words, the lattice is embedded in the continuous line $\Re$ and the lattice point n is identified with the point na in $\Re$. To obtain the continuous position degree of freedom $\mathcal{F}$, let $a \to 0$ and the allowed values of the particle's position x can take any real value, that is, $x \in \Re$, and hence $\mathcal{F} \to \Re$.

Similar to the finite case given in (4.9), the basis states are labeled by $|n\rangle$ and the dual basis states by $\langle n|$. The state space for the continuous degree of freedom x is obtained by considering the state space $\mathcal{F}_{(2N+1)}$ in the limit of $N \to \infty$. Note one

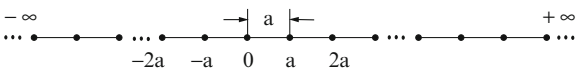


Fig. 4.7 Discretization of a continuous degree of freedom space $\mathcal{F} = \Re$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

needs *first* to take $N \to \infty$ *and then* take $a \to 0$ for obtaining the continuous degree of freedom x.[6]

The discrete basis vectors of the quantum particle's state space $\mathcal{V}$ are represented by infinite column vectors with the only nonzero entry being unity in the n-th position. Hence, taking $N \to \infty$ in (4.9) yields

$$|n\rangle \ : \ n = 0, \pm 1, \pm 2, \ldots, \pm\infty$$

where, more explicitly

$$|n\rangle = \begin{bmatrix} .. \\ 0 \\ 1 \\ 0 \\ .. \end{bmatrix} \ : \ n\text{th position}$$

The basis vectors for the dual state space $\mathcal{V}_{\rm D}$ are given by

$$\begin{aligned} \langle n| &= \begin{bmatrix} \cdots 0 \ 1 \ 0 \cdots \end{bmatrix} \\ \Rightarrow \langle n|m\rangle &= \delta_{n-m} \end{aligned} \tag{4.14}$$

The completeness of the basis states, taking $N \to \infty$ in (4.11), yields the following:

$$\sum_{n=-\infty}^{+\infty} |n\rangle\langle n| = \text{diagonal}(\ldots, 1, 1, \ldots) = \mathbb{I} \ : \ \text{Completeness Equation}$$

where $\mathbb{I}$ above is the infinite-dimensional unit matrix.

The limit of $a \to 0$ needs to be taken to obtain a continuous x; in terms of the underlying lattice, the continuous point x is related to the discrete lattice point n by the following:

$$-\infty \le x \le +\infty \ \ : \ \ x = \lim_{a \to 0} [na]; \ \ n = 0, \pm 1, \pm, \ldots, \pm\infty$$

The state vector for the particle is given by the ket vector $|x\rangle$, with its dual vector given by the bra vector $\langle x|$. The basis state $|n\rangle$ is dimensionless; the ket vector $|x\rangle$ has a dimension of $1/\sqrt{a}$ since, from (4.13), the Dirac δ-function has dimension of $1/a$. Hence due to dimensional consistency,

$$|x\rangle = \lim_{a \to 0} \frac{1}{\sqrt{a}} |n\rangle \ \ ; \ \ \langle x| = \lim_{a \to 0} \frac{1}{\sqrt{a}} \langle n| \tag{4.15}$$

[6]For compact degrees of freedom such as the circle S^1, the infinitesimal a is related to N; see Sect. 5.5.

The position projection operator is equal to the outer product of the position ket vector with the bra vector and is given by

$$|x\rangle\langle x| = \lim_{a\to 0} \frac{1}{a}|n\rangle\langle n| \tag{4.16}$$

The scalar product, for $x = na$ and $x' = ma$, in the limit of $a \to 0$, is given, from (4.14), (4.15), and (4.13), by the Dirac delta function:

$$\langle x|x'\rangle = \lim_{a\to 0} \frac{1}{a}\delta_{m-n} \Rightarrow \langle x|x'\rangle = \delta(x-x') \tag{4.17}$$

The completeness equation has the following continuum limit:

$$\mathbb{I} = \sum_{n=-\infty}^{+\infty} |n\rangle\langle n| = \lim_{a\to 0} a \sum_{n=-\infty}^{+\infty} |x\rangle\langle x| \tag{4.18}$$

$$\Rightarrow \int_{-\infty}^{\infty} \mathrm{d}x|x\rangle\langle x| = \mathbb{I} \ : \ \text{Completeness Equation} \tag{4.19}$$

Equation 4.18 shows that the projection operators given in (4.16) are complete and span the entire of state space $\mathcal{V}$.

$\mathbb{I}$ is the *identity operator* on state space $\mathcal{V}$; namely, for any state vector $|\psi\rangle \in \mathcal{V}$, it follows from the completeness equation that

$$\mathbb{I}|\psi\rangle = |\psi\rangle$$

The completeness equation given by (4.19) is a key equation that is central to the analysis of state space and yields the following:

$$\langle x|\mathbb{I}|x'\rangle = \int_{-\infty}^{\infty} \mathrm{d}z\langle x|z\rangle\langle z|x'\rangle = \int_{-\infty}^{\infty} \mathrm{d}z\delta(x-z)\delta(z-x') = \delta(x-x')$$

that follows from the definition of the Dirac delta function $\delta(x-x')$. The above equation shows that $\delta(x-x')$ is the matrix element of the identity operator $\mathbb{I}$ for the continuous the degree of freedom $\mathcal{F} = \Re$ in the x basis.

A state space $\mathcal{V}(\mathcal{F})$ of a continuous degree of freedom $\mathcal{F}$ is also called a *function space*, and it is for this reason that the subject of functional analysis studies the mathematical properties of quantum mechanics.

Consider the special case of $\mathcal{F} = \Re$; a state vector $|f\rangle$ that is an element of $\mathcal{V}(\Re)$ yields a state function $f(x)$ given by $f(x) = \langle x|f\rangle$; hence all functions of x, namely, $f(x)$, can be thought of as elements of a state space $\mathcal{V}(\Re)$. Of course, being an element of a state space endows the function $f(x)$ with the additional property of linearity that needs to be consistent with all the other properties of $f(x)$. It should be noted that not all functions are elements of a quantum mechanical state space.

4.7 Basis States for State Space

The bra and ket vectors $\langle x|$ and $|x\rangle$ are the basis vectors of the $\mathcal{V}_{\rm D}$ and $\mathcal{V}$, respectively. For $2N+1$-dimensional Euclidean space, it can readily be shown that the $2N+1$ vectors $|n\rangle, n = \pm 1, \pm 2, \ldots, \pm N$ given in (4.9) are linearly independent and form a complete basis for the space: any vector in the $2N+1$-dimensional Euclidean space can be expanded as a linear sum of these basis states. The crucial property of these basis states is that the complete basis vectors satisfy the *completeness equation* given in (4.11), namely, that

$$\sum_{n=-N}^{+N} |n\rangle\langle n| = \mathbb{I} \;\; ; \;\; \langle n|m\rangle = \delta_{n-m}$$

Any $2N+1$-dimensional vector $\mathbf{v}$, denoted by $|v\rangle$, can be expanded as follows:

$$|v\rangle = \mathbb{I}|v\rangle = \sum_{n=-N}^{N} v_n|n\rangle \;\; ; \;\; v_n = \langle n|v\rangle \tag{4.20}$$

For the infinite-dimensional state space, a complete basis set of vectors must satisfy the completeness equation, which for the coordinate basis $|x\rangle$ is given by (4.19) and (4.17), namely,

$$\int_{-\infty}^{\infty} {\rm d}x|x\rangle\langle x| = \mathbb{I} \;\; ; \;\; \langle x|x'\rangle = \delta(x-x')$$

In general, for state vectors $|\psi_n\rangle$ with components give by $\psi_n(x) = \langle x|\psi_n\rangle$, the condition for a complete basis is given by

$$\sum_{n=-\infty}^{+\infty} |\psi_n\rangle\langle\psi_n| = \mathbb{I} \;\; \Rightarrow \;\; \sum_{n=-\infty}^{+\infty} \psi_n(x)\; \psi_n^*(x') = \delta(x-x')$$

The completeness equation is also referred as the *resolution of the identity* since only a complete set of basis states can yield the identity operator on state space.

An element of the state space $\mathcal{V}$ is a ket vector $|\psi\rangle$ and can be thought of as an infinite-dimensional vector with *components* given by $\psi(x) = \langle x|\psi\rangle$. In analogy with (4.20), the vector $|\psi\rangle$ has the following representation in the $|x\rangle$ basis:

$$|\psi\rangle = \int_{-\infty}^{\infty} {\rm d}x|x\rangle\langle x|\psi\rangle = \int_{-\infty}^{\infty} {\rm d}x\psi(x)|x\rangle \;\; ; \;\; \psi(x) = \langle x|\psi\rangle \tag{4.21}$$

The vector $|\psi\rangle$ can be mapped to a unique *dual* vector denoted by $\langle\psi| \in \mathcal{V}_{\rm D}$; in components $\psi^*(x) = \langle\psi|x\rangle$ and

$$\langle\psi| = \int_{-\infty}^{\infty} \mathrm{d}x \langle\psi|x\rangle\langle x| = \int_{-\infty}^{\infty} \mathrm{d}x \psi^*(x)\langle x| \;\; ; \;\; \psi^*(x) = \langle\psi|x\rangle$$

Note the state vector and its dual are related by complex conjugation, namely,

$$\langle\chi|\psi\rangle = \langle\psi|\chi\rangle^* \;\; \Rightarrow \;\; \langle x|\psi\rangle = \langle\psi|x\rangle^* \tag{4.22}$$

The scalar product of two state vectors is given by[7]

$$\langle\chi|\psi\rangle \equiv \int \mathrm{d}x \chi^*(x)\psi(x)$$

The vector $|\psi\rangle$ and its dual $\langle\psi|$ have the important property that they define the "length" $\langle\psi|\psi\rangle$ of the vector. The completeness equation (4.19) yields the following:

$$\langle\psi|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x \psi(x)^*\psi(x) \geq 0$$

4.8 Unitary Transformation: Momentum Basis

The degree of freedom space $\mathcal{F}$ can be assigned coordinates in order to describe it in detail. There are a variety of *equivalent* coordinate systems for a given state space, the various coordinate systems being related by *unitary* transformations, denoted by $\mathbb{U}$; these are transformations that preserve the scalar product of two state vectors. Let $|\psi\rangle$ be a state vector and $\langle\chi|$ be a dual state vector; recall from (4.22) that the dual vector is obtained by complex conjugating the state vector. Consider the transformations

$$|\psi\rangle \to \mathbb{U}|\psi\rangle \;\; ; \;\; \langle\chi| \to \langle\chi|\mathbb{U}^\dagger \tag{4.23}$$

where $\mathbb{U}^\dagger$ is transposition and complex conjugation. For an $N \times N$ matrix, one has $\mathbb{M}^\dagger_{ij} = \mathbb{M}^*_{ji}$.

The transformation given by (4.23) is unitary if it leaves the scalar product unchanged, namely,

$$\begin{aligned} &\langle\chi|\psi\rangle \to \langle\chi|\mathbb{U}^\dagger\mathbb{U}|\psi\rangle = \langle\chi|\psi\rangle \\ &\Rightarrow \;\; \mathbb{U}^\dagger\mathbb{U} = \mathbb{I} \;: \text{Unitary transformation} \end{aligned} \tag{4.24}$$

[7] A more direct derivation of the completeness equation is the following:

$$\langle\chi|\psi\rangle = \langle\chi|\{\int_{-\infty}^{\infty} \mathrm{d}x|x\rangle\langle x|\}|\psi\rangle \;\; \Rightarrow \mathbb{I} = \int_{-\infty}^{\infty} \mathrm{d}x|x\rangle\langle x|$$

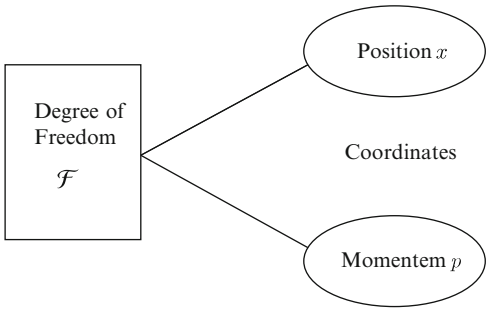


Fig. 4.8 Coordinates of the degree of freedom space $\mathcal{F}$ of a quantum particle (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

For the case of a quantum particle, as shown in Fig. 4.8, one can choose the position coordinate x or equivalently the momentum coordinate p to coordinatize $\mathcal{F}$. The different coordinates used for describing the degree of freedom space $\mathcal{F}$ yield different basis states of state space $\mathcal{V}$.

For finite dimension vector spaces, such as a N-dimensional Euclidean space, a complete set of basis vectors can be rotated to yield a new and equivalent set of complete basis vectors. The rotations are transformations that leave the scalar product unchanged, as in (4.24). The change of coordinates for a Euclidean space is transformations by an orthogonal matrix; for finite-dimensional and real spaces, unitary transformations in fact reduce to the orthogonal matrices.

The change of basis of a finite-dimensional space generalizes to infinite-dimensional state space, with the transformations being unitary transformations. In particular, under a unitary transformation $\mathbb{U}$, a complete coordinate basis of $\mathcal{F}$ yields a new and equivalent complete basis.

For concreteness, consider a quantum particle in one dimension; the degree of freedom space is $\mathcal{F} = \Re$ and the state space is $\mathcal{V}(\Re)$. The *representation* of state space depends on the choice of a coordinate systems for the degree of freedom $\mathcal{F}$. Note, however, that the state space $\mathcal{V}$, as such, is a coordinate independent object and in particular does not depend on the basis chosen to represent it.

Two widely used coordinate systems for $\mathcal{F} = \Re$ are the position x and momentum p coordinates for $\Re$ and are shown in Fig. 4.8. This yields the following two representations of the state vector:

$$|\psi\rangle \to \psi(x) = \langle x|\psi\rangle; \quad x \in \Re$$
$$|\psi\rangle \to \psi(p) = \langle p|\psi\rangle; \quad p \in \Re$$

The different basis chosen to represent the degree of freedom space $\mathcal{F}$ is not simply a mathematical property of $\mathcal{F}$. Rather, the basis chosen has direct experimental significance since the design of the experimental device that studies the degree of freedom $\mathcal{F}$ are based on the representation chosen.

Unitary Transformation $\mathbb{U}$

A unitary transformation connects the basis states of state space $\mathcal{V}(\Re)$ for the two different coordinates of $\mathcal{F} = \Re$. Consider the coordinate and momentum ket basis states $|x\rangle; |p\rangle$ with $x, p \in [-\infty, +\infty]$; a unitary transformation $\mathbb{U}$ maps the coordinate to the momentum basis and yields the following[8]

$$|p\rangle = \int_{-\infty}^{\infty} \mathrm{d}x \mathbb{U}[p,x]|x\rangle$$

where

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi\hbar} \mathbb{U}[p,x]\mathbb{U}^{\dagger}[p,x'] = \delta(x-x') \quad : \quad \text{Unitarity}$$

The Fourier representation of the Dirac delta function yields

$$\delta(x-x') = \langle x|x'\rangle = \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi\hbar} \mathrm{e}^{\mathrm{i}p(x-x')/\hbar} = \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi\hbar} \langle x|p\rangle\langle p|x'\rangle \tag{4.25}$$

The matrix elements of the unitary transformation is seen to be the scalar product of the dual momentum basis state $\langle p|$ with the basis state $|x\rangle$ and is given by

$$\begin{aligned} \langle x|\mathbb{U}^{\dagger}|p\rangle &= \mathbb{U}^{\dagger}[p,x] = \langle x|p\rangle = \mathrm{e}^{\mathrm{i}px/\hbar} \\ \langle p|\mathbb{U}|x\rangle &= \mathbb{U}[p,x] = \langle p|x\rangle = \mathrm{e}^{-\mathrm{i}px/\hbar}. \end{aligned} \tag{4.26}$$

From (4.25), the completeness equation for momentum space basis $|p\rangle$ is given by

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi\hbar} |p\rangle\langle p| = \mathbb{I} \;\; ; \;\; \langle p|p'\rangle = 2\pi\hbar\delta(p-p') \tag{4.27}$$

In summary, a complete basis states for the state space is provided by *either* the coordinate basis $|x\rangle$ *or* the momentum basis $|p\rangle$; the unitary transformation connecting the two bases is given by

$$|x\rangle = \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi\hbar} |p\rangle\langle p|x\rangle = \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi\hbar} \mathrm{e}^{-\mathrm{i}px/\hbar}|p\rangle \tag{4.28}$$

$$|p\rangle = \int_{-\infty}^{\infty} \mathrm{d}x|x\rangle\langle x|p\rangle = \int_{-\infty}^{\infty} \mathrm{d}x \mathrm{e}^{\mathrm{i}px/\hbar}|x\rangle \tag{4.29}$$

[8]The factor of $2\pi\hbar$ in $\mathrm{d}p$ has been introduced for dimensional consistency, with $\mathbb{U}$ being dimensionless and the dimension of $\hbar$ = dimension of $x\times$ dimension of p; furthermore, all the factors of $2\pi\hbar$ are carried by the momentum integration and momentum delta functions. That $\hbar$ is Planck's constant can only be concluded when momentum is defined in Sect. 5.6.

Note the unitary transformation connecting the coordinate and momentum basis states, as given in (4.28) and (4.29), shows that the change of basis is a trans-empirical superposition of basis vectors of state space.

From a mathematical point of view, the coordinate and momentum basis are equivalent, related by a unitary transformation. However, from the point of view of quantum mechanics, the choice of the two basis vectors has very different physical implications and which is related to the theory of quantum measurement.

The state vector $|\psi\rangle$ represents the quantum state of the particle. Choosing a coordinate basis $|x\rangle$ to study the quantum system requires projection operators that detect the quantum state at position x, and with predicted probability $|\langle x|\psi\rangle|^2 = |\psi(x)|^2$, and is discussed in Sect. 9.2. On the other hand, choosing the momentum basis $|p\rangle$ requires measuring the quantum state's momentum, using a different device consisting of projection operators that detect the momentum p, and with predicted probability $|\langle p|\psi\rangle|^2 = |\widetilde{\psi}(p)|^2$.

Note that the function $|\widetilde{\psi}(p)|^2$ is very different from the function $|\psi(x)|^2$ and expresses the fact that the state vector yields very different results, depending on how it is measured.

It will be shown in Sect. 5.7 that the transformations given in (4.28) and (4.29) relating x and p result in the Heisenberg Uncertainty Principle; among many implications of the Heisenberg Uncertainty Principle is that *both* x and p cannot be *simultaneously* measured. Hence, choosing the coordinate or momentum basis for $\mathcal{F} = \Re$ determines how to measure the state vector $|\psi\rangle$—with experiments designed for different basis states giving different, but consistent, results.

4.9 State Space $\mathcal{V}$

One of the most remarkable properties of the quantum description of Nature is that the state vector, denoted by $|\psi\rangle$, is an element of a state space $\mathcal{V}$ that is a *linear vector space*. The precise structure of the linear vector space $\mathcal{V}$ depends on the nature of the quantum degree of freedom. From the simplest quantum system consisting of two possible states to a system having N particles in four-dimensional spacetime to quantum fields having infinite number of degrees of freedom, there is a linear vector space $\mathcal{V}$ and a state vector defined for these degrees of freedom.

Euclidean space $\Re^N$ is a finite-dimensional linear vector space; the linear vector spaces $\mathcal{V}$ that occur in quantum mechanics and quantum field theory are usually state spaces that are an infinite-dimensional generalization of $\Re^N$. Infinite-dimensional linear vector spaces arise in many applications in science and engineering, including the study of partial differential equations and dynamical systems, and many of their properties, such as the addition of vectors, are the generalizations of the properties of finite-dimensional vector spaces.

The following are some of the main properties of a *linear vector space* $\mathcal{V}$:

1. Since they are elements of a linear vector space, a state vector can be added to other state vectors. In particular, *ket vectors* $|\psi\rangle$ and $|\chi\rangle$ are *complex-valued vectors* of $\mathcal{V}$ and can be added as follows:

$$|\eta\rangle = a|\psi\rangle + b|\chi\rangle \tag{4.30}$$

   where a, b are complex numbers $\in \mathbb{C}$ and yield another element $|\eta\rangle$ of $\mathcal{V}$. Vector addition is commutative and associative.
2. For every ket vector $|\psi\rangle \in \mathcal{V}$, there is a dual (bra) vector $\langle\psi|$ that is an element of the *dual linear vector space* $\mathcal{V}_{\mathrm{D}}$. The dual vector space is also linear and yields the following:

$$\langle\eta| = a^*\langle\psi| + b^*\langle\chi|$$

   The collection of all (dual) bra vectors forms the dual space $\mathcal{V}_{\mathrm{D}}$.
3. More formally, $\mathcal{V}_{\mathrm{D}}$ is the collection of *all* linear mappings that take elements of $\mathcal{V}$ to $\mathcal{C}$ by the *scalar product*. In mathematical notation,

$$\mathcal{V}_{\mathrm{D}} : \mathcal{V} \to \mathbb{C}$$

   The vector space and its dual, shown in Fig. 4.9, are not necessarily isomorphic.[9]
4. For any two ket $|\psi\rangle$ and bra $\langle\eta|$ vectors belonging to $\mathcal{V}$ and $\mathcal{V}_{\mathrm{D}}$, respectively, the *scalar product*, namely, $\langle\eta|\psi\rangle$, yields a complex number and has the following property:

$$\langle\eta|\psi\rangle = \langle\psi|\eta\rangle^*$$

   where recall $*$ stands for complex conjugation. The scalar product is linear and yields

$$\langle\eta|\zeta\rangle = a^*\langle\psi|\zeta\rangle + b^*\langle\chi|\zeta\rangle$$

   In particular, $\langle\psi|\psi\rangle \equiv |\psi|^2$ is a real number—a fact of far-reaching consequence in quantum mechanics.
5. One of the fundamental properties of quantum states is that *two states are distinct* if they are *linearly independent*. In particular, two states $|\psi\rangle$ and $|\chi\rangle$ are completely distinct if and only if they are *orthogonal*, namely,

$$\langle\chi|\psi\rangle = 0 \quad : \quad \text{orthogonal} \tag{4.31}$$

   Two states being distinct is central to the theory of quantum measurement and is discussed in Chap. 9.

[9] Two spaces are isomorphic if there is an invertible mapping that maps each element of one space to a (unique) element of the other space.

4.10 Hilbert Space

Starting in the 1900s, Hilbert space was studied by David Hilbert, Erhard Schmidt, and Frigyes Riesz as belonging to the class of infinite-dimensional function space. The main feature that arises in a Hilbert space is the issue of *convergence* of an infinite sequence of elements of Hilbert space, something that is absent in a finite-dimensional vector space.

To allow for the probabilistic interpretation of the state vector $|\psi\rangle$, all state vectors that represent physical systems must have unit norm, that is,

$$\langle\psi|\psi\rangle \equiv |\psi|^2 = 1 \quad : \quad \text{unit norm}$$

A linear vector space $\mathcal{V}$ that is a *normed vector space* is called a *Hilbert space*, shown schematically in Fig. 4.10. For a Hilbert space, the dual state space is isomorphic to the Hilbert space, namely, $\mathcal{V} \simeq \mathcal{V}_{\mathrm{D}}$.

The state space of quantum entities is a Hilbert space. However, there are classical random systems, for example, that occur in finance and for quantum dissipative processes, where the state space is not a Hilbert space and in particular leads to a dual state space $\mathcal{V}_{\mathrm{D}}$ is not isomorphic to the state space $\mathcal{V}$ [3].

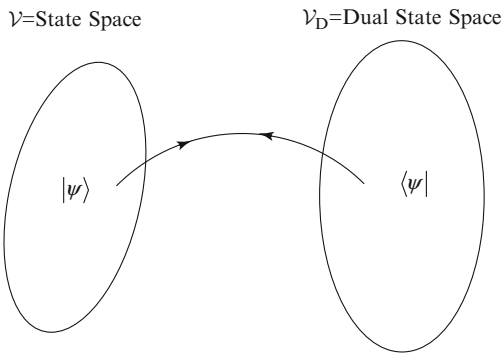


Fig. 4.9 A state space $\mathcal{V}$ and its dual space $\mathcal{V}_{\mathrm{D}}$; note that the dual space is not necessarily isomorphic to the state space (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

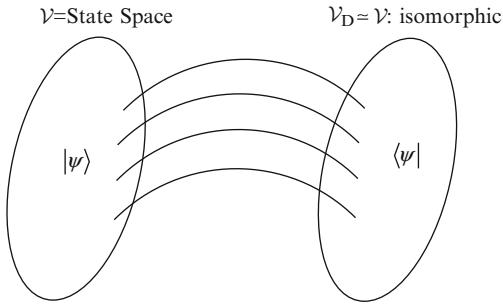


Fig. 4.10 Hilbert space is a unit norm state space with $\mathcal{V} \simeq \mathcal{V}_{\mathrm{D}}$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Recall from Sect. 4.5 that for a quantum entity having spin ℓ, its degree of freedom $\mathcal{F} = \{\ell, \ell-1, \ldots, 1, 0, -1, \ldots, -\ell\}$ is the discrete space and its Hilbert space $\mathcal{V}(\mathcal{F})$ is isomorphic to the space $S^{2\ell+1}/S^1$.

For the continuous degree of freedom $\mathcal{F} = \Re$, an element of $|\psi\rangle$ of Hilbert space has unit norm and hence yields

$$\langle\psi|\psi\rangle \equiv |\psi|^2 = \int_{-\infty}^{+\infty} \mathrm{d}x |\psi(x)|^2 = 1 \quad : \quad \text{unit norm}$$

Noteworthy 4.2: Quantum mathematics

Quantum mathematics originates in quantum physics. One of the essential features of quantum mathematics is the *synthesis* of *calculus* with *linear algebra*. To illustrate this synthesis, note that in linear algebra the representation of a vector in N-dimensional Euclidean space $\Re^N$ is given in terms of N linearly independent basis vectors $\mathbf{e}_i$. An arbitrary vector is expressed by its components, which in Dirac's notation is given as follows:

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{i-j} = \langle i|j\rangle \;\; ; \;\; \mathbf{f} = \sum_{i=1}^{N} f_i \mathbf{e}_i = \sum_{i=1}^{N} f_i |i\rangle$$

In quantum mathematics, an infinite-dimensional generalization of linear algebra is made by generalizing Euclidean space $\Re^N$ to state space $\mathcal{V}$. There are now a continuous infinity of independent basis vectors $|x\rangle$, where x is governed by rules of calculus. The "vector" $|f\rangle$ belonging to $\mathcal{V}$, from (4.21), has the following representation in the $|x\rangle$ basis:

$$|f\rangle = \int_{-\infty}^{\infty} \mathrm{d}x f(x)|x\rangle; \; \langle x|x'\rangle = \delta(x - x')$$

The Dirac notation provides a transparent representation of the infinite-dimensional generalization of linear algebra that naturally combines it with calculus. Function $f(x)$ of a continuous variable x, the mainstay of calculus, in quantum mathematics is endowed with a *linear* structure that is inherited from state (function) space $\mathcal{V}$. The Dirac delta function plays a crucial role in the mathematical realization of the synthesis of linear algebra with calculus.

4.11 Summary

This chapter was focused on the mathematical description of the degree of freedom as well as of its state space.

The collection of all the possible values of a quantum entity constitutes its degree of freedom. The simplest possible quantum degree of freedom, consisting of two possible values, was analyzed is some detail, and its state space was shown to be

given by the Bloch sphere. The degree of freedom with $2N+1$ possible values was studied to show how the continuous degree of freedom emerges in the limit of $N \to \infty$.

For discrete degrees of freedom, the state space is isomorphic to a subspace of an N-dimensional Euclidean space, whereas for a continuous degree of freedom, the state space is an infinite-dimensional generalization of Euclidean space. The basis states for the continuous degree of freedom were derived, and the unitary transformation mapping the coordinate to the momentum basis was obtained.

The general properties of the infinite-dimensional state space were discussed and shown to be a logical development of the concepts of a finite-dimensional vector space. In quantum mechanics, the state space is a (normed) Hilbert space dictated by the probabilistic interpretation of the state vector.

The mathematical thinking about quantum mechanics developed by Dirac, partly expressed in his notation for linear vector spaces, provides a transparent and clear framework for introducing and developing the ideas of the degree of freedom and of state space. Dirac's mathematical formulation of quantum mechanics should not be thought of as being merely notation but, rather, as a major conceptual revolution. Dirac's notation is as far-reaching and groundbreaking as the change in the way of thinking about the real numbers brought about by the shift from the Roman to the decimal notation.

Operators

5

An observation in classical mechanics experimentally determines the unique properties of the system, usually specified by its position, energy, momentum, and so on. A quantum system is indeterminate, and hence one has to define what one means by an observation. Furthermore, all the physical quantities of classical mechanics, such as position, momentum, energy, and angular momentum, need to be generalized to reflect quantum indeterminacy.

Instead of having a unique value, physical quantities for a quantum system take a whole *range of possible values* that depend on the state function of the quantum entity. The state vectors are elements of a state space, which is a Hilbert space due to the probabilistic interpretation in quantum mechanics.[1] All experimental observations carried out on the state vector are represented and realized by operators that are also called *observables* when referring to physical quantities such as energy and momentum.[2]

The following topics are covered in this chapter:

- The essential role of operators in extracting information from and about the degree of freedom.
- Hermitian operators are defined, and their properties are exemplified by the Hamiltonian operator.
- Position and momentum are studied in some detail as these are amongst the most important operators in quantum mechanics.
- Some important properties of observables and their quantum numbers are discussed.
- The Heisenberg commutation equation is derived.

[1] In the application of *quantum mathematics* to finance, the Hamiltonians evolving the stochastic financial instruments are not Hermitian. Furthermore, the state space of financial instruments on which the Hamiltonian acts is not a positive normed Hilbert space but instead is much larger with many financial instruments having a divergent, infinite norm [3].

[2] Operators representing physical quantities have been termed as observables by Dirac, and we use this terminology [10].

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8_5, © Springer Science+Business Media New York 2013

- The framework of state space and Hermitian operators is utilized to briefly discuss the Schrödinger and Heisenberg formulations of quantum mechanics.

5.1 Operators: Trans-empirical to Empirical

All the information that can be obtained from a degree of freedom $\mathcal{F}$ is encoded in its state vector $|\psi\rangle$. Operators $\mathcal{O}$ are mathematical mappings of the state space into itself. The "diagonal" matrix elements of the operators, namely, $\langle\psi|\mathcal{O}|\psi\rangle$, will be shown in Sect. 5.8 to be equal to the average observed value of the operator $\mathcal{O}$ for a given quantum state $|\psi\rangle$.

In other words, operators $\mathcal{O}$ act on the state vector $|\psi\rangle$ and project it to the experimentally measured quantity $\langle\psi|\mathcal{O}|\psi\rangle$ that yields the empirical content of the degree of freedom $\mathcal{F}$. As shown in Fig. 5.1, the operator $\mathcal{O}$ provides a mapping of the trans-empirical state space to the empirically observed results of the experiment that measures the average value of $\mathcal{O}$.

The experimental observation of a quantum system always results in assigning a real value to the physical quantity being measured. Hence, all operators representing physical quantities must be assigned real numbers by the state vector that is being observed, which in turn requires that the observables (operators) are Hermitian (defined precisely in Sect. 5.2). This is the fundamental reason why, in quantum mechanics, all physical quantities such as energy and momentum are represented by Hermitian operators.

The *experimental devices* that measure the observable properties of a degree of freedom $\mathcal{F}$ are mathematically realized in quantum mechanics by Hermitian operators $\mathcal{O}_i$, $i = 1, 2, \ldots, I$. There is no analog of operators $\mathcal{O}_i$ in classical mechanics.

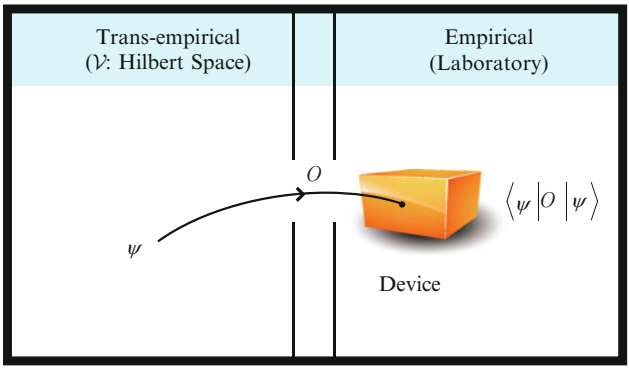


Fig. 5.1 The operator $\mathcal{O}$ acts on state vector $|\psi\rangle$ and causes the state vector to "collapse," resulting in the observed value of the operator—given after repeated observations by $\langle\psi|\mathcal{O}|\psi\rangle$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

In summary, all observables (physical quantities) are represented by Hermitian operators that act on the state vector to yield, after repeated measurements, the average value of the operators for the quantum state.

5.2 Hermitian Operators

Every degree of freedom $\mathcal{F}$ defines a state space $\mathcal{V}$ and operators $\mathcal{O}$ that act on that state space. All operators $\mathcal{O}$ are mathematically defined to be *linear mappings* of the state space $\mathcal{V}$ into itself, shown in Fig. 5.2, and yield

$$\mathcal{O}: \ |\psi\rangle \to \mathcal{O}|\psi\rangle \ \Rightarrow \ \mathcal{O}: \ \mathcal{V} \to \mathcal{V}$$
$$\mathcal{O}\Big(a|\psi_1\rangle + b|\psi_2\rangle\Big) = a\mathcal{O}|\psi_1\rangle + b\mathcal{O}|\psi_2\rangle \ : \ \text{linear}$$

where a, b are constants.

An operator is an element of the space formed by the outer product of $\mathcal{V}$ with its dual $\mathcal{V}_{\mathrm{D}}$, that is,

$$O \in \mathcal{V} \otimes \mathcal{V}_{\mathrm{D}}$$

All physical quantities of a degree of freedom such as position, momentum, energy, angular momentum, spin, and charge are all represented by Hermitian operators. Hermitian operators $\mathcal{O}$ map the space $\mathcal{V}$ onto itself.

For a two-state system, the state space is given by the Bloch sphere, discussed in Sect. 4.4, and operators are 2×2 complex-valued Hermitian matrices. A Hermitian matrix is only defined for complex *square* matrices $N \times N$ with M_{ij} with $i, j = 1, 2, \ldots, N$, and M_{ij} satisfies

$$M^{\dagger} = (M^T)^* = M \ \Rightarrow \ M^*_{ji} = M_{ij} \ ; \ i, j = 1, 2, \ldots, N \ : \ \text{Hermitian} \tag{5.1}$$

Note the crucial point that unless i, j have the same range, the equality in (5.1) cannot hold for all i, j.

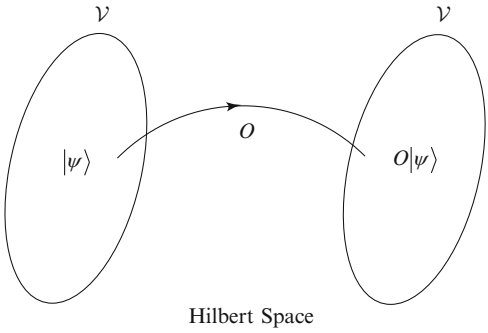


Fig. 5.2 An operator $\mathcal{O}$ acting on element $|\psi\rangle$ of the state space $\mathcal{V}$ and mapping it to $\mathcal{O}|\psi\rangle$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Hermitian operators on linear vector state space are infinite-dimensional generalizations of $N \times N$ matrices, with $N \to \infty$, and have new properties that are absent in finite matrices.

The range over which the finite index i takes values for a finite-dimensional matrix has a generalization for state vectors in Hilbert space $\mathcal{V}$. The domain of the operator $\mathcal{O}$, denoted by $D(\mathcal{O}) \subset \mathcal{V}$, is defined by all elements $|\psi\rangle$ in $\mathcal{V}$ such that $\mathcal{O}|\psi\rangle \in \mathcal{V}$. Similarly, the vector $|\chi\rangle$ is in $D(\mathcal{O}^\dagger)$, the domain of $\mathcal{O}^\dagger$, if $\mathcal{O}^\dagger|\chi\rangle \in \mathcal{V}$.

The analog of Hermitian conjugation being defined only for a square matrix is that for operators on Hilbert space, the adjoint (Hermitian conjugate) operator can be defined for *only* those operators $\mathcal{F}$ for which the domain of the operator and its adjoint are isomorphic or, in other words, $D(\mathcal{O}) = D(\mathcal{O}^\dagger)$.

Hermitian conjugation of operators on state space is defined by

$$\langle\psi|\mathcal{O}^\dagger|\chi\rangle \equiv \langle\chi|\mathcal{O}|\psi\rangle^* \quad : \quad \text{Hermitian conjugation}$$

Once the domains of the operator and its conjugate are isomorphic, the form of the operator has to be invariant under conjugation, that is, $\mathcal{O} = \mathcal{O}^\dagger$, for the operator to be self-adjoint.[3] More precisely, an operator is Hermitian if its Hermitian conjugate operator is equal to the operator itself, that is, if

$$\mathcal{O}^\dagger = \mathcal{O} \;\; \Rightarrow \;\; \langle\psi|\mathcal{O}|\chi\rangle \equiv \langle\chi|\mathcal{O}|\psi\rangle^* \quad : \quad \text{Hermitian operator} \tag{5.2}$$

Note that all the *diagonal elements* of a Hermitian operator $\mathcal{O}$ are real since for any arbitrary state vector $|\psi\rangle$, the diagonal element is as shown below:

$$\langle\psi|\mathcal{O}|\psi\rangle = \langle\psi|\mathcal{O}^\dagger|\psi\rangle = \langle\psi|\mathcal{O}|\psi\rangle^* \quad : \quad \text{Real} \tag{5.3}$$

Furthermore, similar to matrices, Hermitian conjugation is a linear operation and yields, for a sum and products of operators, the following:

$$(c_1\mathcal{O}_1 + c_2\mathcal{O}_2 + \cdots)^\dagger = c_1^*\mathcal{O}_1^\dagger + c_2^*\mathcal{O}_2^\dagger + \cdots \;\; ; \;\; (\mathcal{O}_1\mathcal{O}_2\ldots)^\dagger = \ldots\mathcal{O}_2^\dagger\mathcal{O}_1^\dagger$$

The trace operation for an operator $\mathcal{O}$, similar to matrices, is defined as a sum of all its "diagonal elements." To make this statement more precise, one needs a resolution of the identity operator on state space $\mathcal{V}$. Consider for concreteness the continuous degree of freedom with the completeness equation given by (4.19) as follows:

$$\mathbf{I} = \int_{-\infty}^{\infty} \mathrm{d}x|x\rangle\langle x|$$

[3] One of the reasons for studying the Hermitian conjugate operator is because one can ascertain the space that an operator acts on, namely whether it acts on $\mathcal{V}$ or on its dual $\mathcal{V}_\mathrm{D}$. For non-Hermitian operators, and these are the ones that occur in describing classical random systems such as those that occur in finance [3], the difference is important.

Trace is a *linear* operation on $\mathcal{O}$ and is defined by

$$\mathrm{tr}(\mathcal{O}) = \mathrm{tr}(\mathcal{O}\mathbf{I}) = \int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{tr}\Big(\mathcal{O}|x\rangle\langle x|\Big) = \int_{-\infty}^{\infty} \mathrm{d}x \langle x|\mathcal{O}|x\rangle \tag{5.4}$$

The properties of the trace operation are summarized below:

$$\mathrm{tr}\left[\sum_i c_i \mathcal{O}_i\right] = \sum_i c_i \mathrm{tr}[\mathcal{O}_i]$$

$$\mathrm{tr}\left[\sum_i c_i \mathcal{O}_i\right]^\dagger = \sum_i c_i^* \mathrm{tr}^*[\mathcal{O}_i]$$

$$\mathrm{tr}(\mathcal{O}_1\mathcal{O}_2\mathcal{O}_3) = \mathrm{tr}(\mathcal{O}_3\mathcal{O}_1\mathcal{O}_2) \quad : \quad \text{cyclic}$$

Cyclicity of the trace makes it invariant under unitary transformations U, namely,

$$\mathrm{tr}[U\mathcal{O}U^\dagger] = \mathrm{tr}[\mathcal{O}U^\dagger U] = \mathrm{tr}[\mathcal{O}]$$

A unitary operator, the generalization of the exponential function $\exp \mathrm{i}\phi$, is given in terms of a Hermitian operator $\mathcal{O}$ by the following:

$$U = \mathrm{e}^{\mathrm{i}\phi\mathcal{O}} \;\Rightarrow\; UU^\dagger = \mathbb{I}$$

$$V = \frac{1 - \mathrm{i}a\mathcal{O}}{1 + \mathrm{i}a\mathcal{O}} \;\Rightarrow\; VV^\dagger = \mathbb{I}$$

5.3 Eigenstates: Projection Operators

Consider a Hermitian operator $\mathcal{O}$. The eigenstates and (real) eigenvalues are a very special set of state vectors that are only *rescaled* by its eigenvalues when the operator $\mathcal{O}$ acts on them and are given by the following:

$$\mathcal{O}|\psi_n\rangle = \lambda_n|\psi_n\rangle \;\; ; \;\; \lambda_n^* = \lambda_n : \text{ Real number} \tag{5.5}$$

The eigenstates $|\psi_n\rangle$ of $\mathcal{O}$ are the closest one can come to a classical state in the following sense: *every* measurement of the properties of an eigenstate $|\psi_n\rangle$, in particular of its eigenvalue λ_n, will always yield the *same* value. In this sense the outcome of the experiment is deterministic.

Note, however, that the eigenstate $|\psi_n\rangle$ is nevertheless not a classical entity; the degree of freedom $\mathcal{F}$ remains transempirical and indeterminate even for an eigenstate; what is determinate is *a* property of the degree of freedom, encoded in the eigenvalue of the eigenfunction.

For example, for a hydrogen atom in an energy eigenstate, a measurement of its energy will always yield the energy eigenvalue, but the electron's coordinate

degree of freedom is nevertheless indeterminate—having observed values for the coordinate projection operator only with certain likelihood, which is determined by the Schrödinger equation.

All Hermitian operators have the following important properties:

- The eigenfunctions are orthonormal and complete, namely,

$$\langle \psi_n | \psi_n \rangle = \delta_{n-m} \;\; ; \;\; \sum_n |\psi_n\rangle\langle\psi_n| = \mathbb{I} \tag{5.6}$$

- Every Hermitian operator defines a complete set of projection operators Π_n that are defined below; (5.6) yields

$$\Pi_n = |\psi_n\rangle\langle\psi_n| \;\; ; \;\; \Pi_n \Pi_m = \delta_{n-m}\Pi_n \;\; ; \;\; \sum_n \Pi_n = \mathbb{I} \tag{5.7}$$

- The Hilbert space and its dual are isomorphic and hence $\mathcal{O} \in \mathcal{V} \otimes \mathcal{V}_{\rm D} \equiv \mathcal{V} \otimes \mathcal{V}$; the *spectral decomposition* of a Hermitian operator is given by the following:

$$\mathcal{O} = \sum_n \lambda_n |\psi_n\rangle\langle\psi_n| = \sum_n \lambda_n \Pi_n \tag{5.8}$$
$$\mathcal{O}^\dagger = \sum_n \lambda_n^* \left(|\psi_n\rangle\langle\psi_n|\right)^\dagger = \mathcal{O}$$

since all eigenvalues λ_n are real. In other words, a Hermitian operator is completely equivalent to the set of all of its projection operators (eigenfunctions) and eigenvalues.

The collection of eigenvalues of an operator, called its eigenspectrum, depends on the nature of the operator.

There are quantum degrees of freedom for which observables like momentum and position can have both discrete and continuous eigenvalues; the discussion for integer quantum numbers can be generalized these systems .

For example, consider the eigenspectrum of the hydrogen atom. The electron interacts with the proton via the Coulomb potential. The energy eigenfunctions have a discrete energy spectrum, from $-13.6\,\text{eV}$ to 0, given by $E_n = -13.6/n^2$ eV and correspond to the eigenstates of the hydrogen atom, which is a bound state of the electron and proton; the integer n is called the principal quantum number of the hydrogen atom.

There is also a continuous energy spectrum from 0 out to infinite energy of the electron interacting with the proton via the Coulomb potential, and it corresponds to energy eigenstates of the electron *scattering* off the proton .

The eigenspectrum of an electron interacting with a proton via the Coulomb potential is shown in Fig. 5.3.

- The trace operation, from (5.8), can be defined in terms of the complete eigenfunctions as follows:

$$\text{tr}(\mathcal{O}) = \sum_n \lambda_n \, \text{tr}\Big(|\psi_n\rangle\langle\psi_n|\Big) = \sum_n \lambda_n \, \langle\psi_n|\psi_n\rangle = \sum_n \lambda_n$$

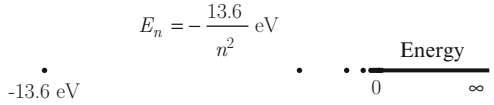


Fig. 5.3 The eigenspectrum of the electron–proton system (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

- Since all eigenvalues of a Hermitian operator are real, let them be labeled in increasing value, that is, $\lambda_{\min}, \ldots, \lambda_n, \ldots, \lambda_{\max}$. For a normalized state vector $|\psi\rangle$ (with $\langle\psi|\psi\rangle = 1$), the expectation of the operator is bounded by its minimum and maximum eigenvalue, namely,

$$\lambda_{\min} \leq \langle\psi|\mathcal{O}|\psi\rangle \leq \lambda_{\max} \;\; ; \;\; \langle\psi|\psi\rangle = 1 \tag{5.9}$$

  Since the projection operator obeys $\Pi_n^2 = \Pi_n$, it has two eigenvalues, namely, 0 and 1; hence from (5.9) above,

$$0 \leq \langle\psi|\Pi_n|\psi\rangle \leq 1 \;\; ; \;\; \langle\psi|\psi\rangle = 1 \tag{5.10}$$

  Equations (5.7) and (5.10) have far-reaching consequences and form the basis of the theory of quantum probability discussed in Chap. 7.
- Unitary transformations discussed in Sect. 4.8 reflect the fact that an observable $\mathcal{O}$ and its eigenfunctions are only defined up to a unitary transformation, with the eigenvalues λ_n being invariant. In particular, under a unitary transformation, the eigenfunction equation given in (5.5) yields the following:

$$\mathcal{O} \to \mathcal{O}_U = U\mathcal{O}U^\dagger = \sum_n \lambda_n U|\psi_n\rangle\langle\psi_n|U^\dagger = \sum_n \lambda_n|\chi_n\rangle\langle\chi_n| \tag{5.11}$$

$$\mathcal{O}_U|\chi_n\rangle = \lambda_n|\chi_n\rangle \;\; \Rightarrow \;\; \mathcal{O}_U U|\psi_n\rangle = \lambda_n U|\psi_n\rangle \;\; \Rightarrow \;\; U|\psi_n\rangle = |\chi_n\rangle$$

- Functions of Hermitian operators are fundamental to a quantum system; for an arbitrary operator-valued function $f(\mathcal{O})$, the spectral resolution given in (5.8) yields

$$f(\mathcal{O}) = \sum_n f(\lambda_n)|\psi_n\rangle\langle\psi_n| \tag{5.12}$$

  where $f(\lambda_n)$ is an ordinary numerical-valued function of the eigenvalues λ_n. Note that from its definition, for Hermitian operator $\mathcal{O}$, the eigenvalues λ_n are all real and yields, for $f(\lambda_n) = f^*(\lambda_n)$, the following:

$$f^\dagger(\mathcal{O}) = f(\mathcal{O}^\dagger) = f(\mathcal{O}) \;\; : \;\; \text{Hermitian operator}$$

Parallel and Orthogonal State Vectors

If two states, namely, $|\chi\rangle$ and $|\eta\rangle$, are parallel, then $|\chi\rangle = a|\eta\rangle$; given that the states are normalized, one has that $a = \exp(\mathrm{i}\phi)$ is a pure phase. Since all measurements take the absolute value of the state vector, one can see that the phase ϕ is removed from all physical measurements. Hence, in quantum mechanics, two state vectors that are *parallel* are *identical*, that is,

$$|\eta\rangle = \mathrm{e}^{\mathrm{i}\phi}|\chi\rangle \;\; \Rightarrow \;\; |\eta\rangle \equiv |\chi\rangle \tag{5.13}$$

It was mentioned in Sect. 4.9 that if two state vectors $|\chi\rangle$ and $|\eta\rangle$ are *orthogonal*, namely,

$$\langle\chi|\eta\rangle = 0$$

then they are *completely distinct* quantum states.

To prove that orthogonal state vectors are completely distinguishable, consider first the special case when the state vectors are two different eigenstates $|\psi_n\rangle$ and $|\psi_m\rangle$. Assuming no degeneracy for the eigenstates, in every measurement, the first state will have energy E_n and the second state, measured by the same apparatus, will have energy $E_m \neq E_n$; hence the two state vectors are completely distinguishable.

To prove the general case of orthonormal vectors $|\chi\rangle$ and $|\eta\rangle$, one can do a unitary transformation of the basis states and choose $|\chi\rangle$ and $|\eta\rangle$ as two of the basis vectors. Then the argument for eigenvectors carries over to the general case.

5.4 Operators and Quantum Numbers

A classical system has conserved quantities such as energy, momentum, and angular momentum; in fact, one usually characterizes a classical system by its conserved quantities, called constants of motion. There is a quantum mechanical generalization of the classically conserved quantities.

Since operators do not commute, define the commutator of two operators by the following:

$$[\mathcal{O}_i, \mathcal{O}_j] = \mathcal{O}_i\mathcal{O}_j - \mathcal{O}_j\mathcal{O}_j \;\; ; \;\; i, j = 1, 2, \ldots, J$$

Operators fall into following two categories depending on whether they commute or do not commute:

- *Non-commuting operators*

$$[\mathcal{O}_i, \mathcal{O}_j] \neq 0 \;\; ; \;\; i, j = J+1, 2, \ldots, I$$

- *Commuting operators*

$$[\mathcal{Q}_i, \mathcal{Q}_j] = 0 \;\; ; \;\; i, j = 1, 2, \ldots, J$$

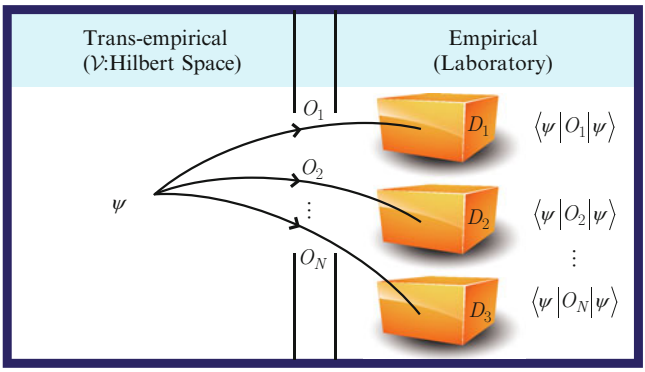


Fig. 5.4 Different devices D_i are required to separately measure the properties of non-commuting operators $\mathcal{O}_i$ for the state vector $|\psi\rangle$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

An experimental device has to be *custom designed* to measure the properties of an operator $\mathcal{O}_i$, with each operator requiring a specific device that is capable of measuring the eigenstates and eigenvalues of operator $\mathcal{O}_i$. For operators $\mathcal{O}_i$ that *do not commute*, a separate device D_i has to be built for each operator, and the operators have to be studied separately, as indicated in Fig. 5.4.

The Hamiltonian H is the most important operator for a quantum system since it evolves the state vector in time. Consider a collection of *commuting operators* $\mathcal{O}_i$; $i = 1, 2, \ldots, N$ that also commute with H, namely,

$$\begin{aligned}[\mathcal{O}_i, H] &= 0; \ \ i = 1, 2, \ldots, N \\ [\mathcal{O}_i, \mathcal{O}_j] &= 0; \ \ i, j = 1, 2, \ldots, N\end{aligned}$$

An important result of Hermitian operators is that all commuting operators can be simultaneously diagonalized with eigenfunctions $|\psi_{n,n_1,n_2,\ldots,n_N}\rangle$ that obey

$$\begin{aligned}H|\psi_{n;n_1,n_2,\ldots,n_N}\rangle &= E_n|\psi_{n;n_1,n_2,\ldots,n_N}\rangle; \ \ n = 1, 2, \ldots \\ \mathcal{O}_i|\psi_{n;n_1,n_2,\ldots,n_N}\rangle &= \lambda^i_{n_i}|\psi_{n;n_1,n_2,\ldots,n_N}\rangle; \ \ i = 1, 2, \ldots, N \ \ ; \ \ n_i = 0, \pm 1, \pm 2, \ldots\end{aligned}$$

Suppose $|\psi_{n_1,n_2,\ldots,n_N}\rangle$ is the initial state vector and is an eigenfunction of all the operators $\mathcal{O}_i$; $i = 1, 2, \ldots, N$—but *not* necessarily an eigenfunction of H; then, from (5.38) and due to the commutativity of $\mathcal{O}_i$ with H,

$$\begin{aligned}|\psi_{t;n_1,n_2,\ldots,n_N}\rangle &= \mathrm{e}^{-\mathrm{i}tH/\hbar}|\psi_{n_1,n_2,\ldots,n_N}\rangle \\ \Rightarrow \mathcal{O}_i|\psi_{t;n_1,n_2,\ldots,n_N}\rangle &= \mathrm{e}^{-\mathrm{i}tH/\hbar}\mathcal{O}_i|\psi_{n,n_1,n_2,\ldots,n_N}\rangle \\ &= \lambda^i_{n_i}\mathrm{e}^{-\mathrm{i}tH/\hbar}|\psi_{n_1,n_2,\ldots,n_N}\rangle \\ &= \lambda^i_{n_i}|\psi_{t;n_1,n_2,\ldots,n_N}\rangle\end{aligned}$$

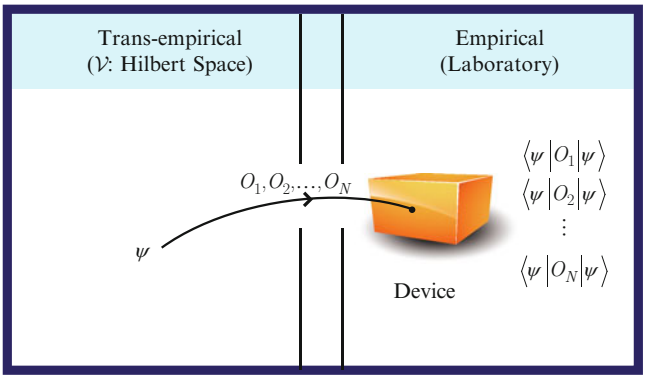


Fig. 5.5 A single device can simultaneously measure the expectation value of commuting operators $\mathcal{O}_i$; $i = 1, 2, \ldots, N$ for the state vector $|\psi\rangle$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

The result above shows that the set of observables which commute with the Hamiltonian and with each other provide eigenvalues and eigenfunctions that are *conserved* in time.

In principle, all commuting operators can be observed simultaneously and again in principle by a single device. Figure 5.5 shows a single device measuring the properties of all the commuting operators.

The integers n_i; $i = 1, 2, \ldots, N$ are called *quantum numbers* and replace the determinate and unique values of the classical observables like momentum, angular momentum, and energy that classically have *continuous* values. The *time-independent* eigenvalues $\lambda^i_{n_i}$ are *constants of motion* and are the generalization of classically conserved quantities. Fixing the values of the various n_i's fully specifies a particular eigenstate of the observables $\mathcal{O}_i$; $i = 1, 2, \ldots, N$. An arbitrary state vector for such a system can be expressed by

$$|\psi(t)\rangle = \sum_{n_1, n_2, \ldots, n_N} c_{n_1, n_2, \ldots, n_N}(t) |\psi_{n_1, n_2, \ldots, n_N}\rangle$$

5.5 Periodic Degree of Freedom

The completeness equation is one of the most important properties of the eigenfunctions of a Hermitian operator. To illustrate this property it is shown how, for a special case, the eigenfunctions yield a resolution of the identity as given in (5.6). The result is derived for a particle moving on a circle, that is, with the degree of freedom being S^1. Another noteworthy aspect is that the coordinate operator for the real line $\Re$ can be seen to arise from the compact degree of freedom that has a finite normalization for the coordinate eigenfunction.

Consider the Hamiltonian operator H, which is the generalization of the classical concept of energy. For concreteness, consider the Hamiltonian in (2.6), with the potential for a free particle given by $V(x) = 0$; hence

$$H = -\frac{1}{2m}\frac{\partial^2}{\partial x^2}; \quad x = x + 2\pi L \tag{5.14}$$

Note x is defined on a *periodic domain* $[0, 2\pi L]$.

Consider the eigenfunctions of H given by

$$H\psi_{\mathrm{E}}(x) = E\psi_{\mathrm{E}}(x); \quad \int_0^{2\pi L} \mathrm{d}x |\psi_{\mathrm{E}}(x)|^2 = 1$$

The ground state (having lowest energy) is nondegenerate and is given by

$$\psi_0(x) = \frac{1}{\sqrt{2\pi L}}$$

All the other energy eigenfunctions are twofold degenerate and given by

$$H\psi_{\pm n}(x) = E\psi_{\pm n}(x); \; E = \frac{n^2}{2mL^2}$$
$$\psi_{\pm n}(x) = \frac{1}{\sqrt{2\pi L}} \mathrm{e}^{\pm \mathrm{i}\frac{nx}{L}}, \quad n = 1, 2, \ldots, +\infty$$

Hence, from the general results given in (5.6), one concludes that the eigenfunctions of H are complete and yield the following completeness equation:

$$\mathbb{I} = |\psi_0\rangle\langle\psi_0| + \sum_{n=1}^{+\infty} (|\psi_{+n}\rangle\langle\psi_{+n}| + |\psi_{-n}\rangle\langle\psi_{-n}|) \tag{5.15}$$

The completeness equation requires *all* the eigenfunctions, including all the degenerate eigenfunctions, of the Hermitian operator. The matrix element of the identity operator, from (4.19), is given by

$$\langle x'|\mathbb{I}|x\rangle = \delta(x - x')$$

Hence, to prove (5.15), we need to find the matrix element of the right-hand side and show that it is equal to $\delta(x - x')$.

Consider the following expression:

$$\langle x'|\{|\psi_0\rangle\langle\psi_0| + \sum_{n=1}^{+\infty} (|\psi_{+n}\rangle\langle\psi_{+n}| + |\psi_{-n}\rangle\langle\psi_{-n}|)\}|x\rangle$$

$$
\begin{aligned}
&= \frac{1}{2\pi L}\left(1+\sum_{n=1}^{+\infty}\left(\mathrm{e}^{\mathrm{i}\frac{n}{L}(x'-x)}+\mathrm{e}^{-\mathrm{i}\frac{n}{L}(x'-x)}\right)\right)\\
&= \frac{1}{2\pi L}\sum_{n=-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}\frac{n}{L}(x'-x)} = \sum_{n=-\infty}^{+\infty}\delta\left(x-x'+2\pi nL\right)\\
&= \delta\left(x-x'\right)\ \text{ since }\ x,x'\in[0,2\pi L]
\end{aligned}
$$

which proves (5.15). To obtain the final expression requires Poisson's summation formula

$$
\frac{1}{2\pi L}\sum_{n=-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}\frac{n}{L}(x'-x)} = \sum_{n=-\infty}^{+\infty}\delta\left(x-x'+2\pi nL\right) \tag{5.16}
$$

Taking the limit of $L\to\infty$ yields the result for $x\in[-\infty,+\infty]$; the completeness equation given in (5.15) for a the circle converges to the unit operator for a continuous degree of freedom given in (4.19) since

$$
\begin{aligned}
\lim_{L\to\infty}\frac{1}{2\pi L}\sum_{n=-\infty}^{+\infty}\mathrm{e}^{\mathrm{i}\frac{n}{L}(x'-x)} &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\mathrm{d}p\mathrm{e}^{\mathrm{i}p(x'-x)}\\
&= \delta(x'-x)\ ;\ x,x'\in[-\infty,+\infty]
\end{aligned}
$$

5.6 Position and Momentum Operators $\hat{x}$ and $\hat{p}$

A quantum particle has a continuous (real) degree of freedom $x\in\mathcal{F}=\Re$. The state space consists of all functions of the single variable x, namely, $\mathcal{V}=\{\psi(x)|x\in\Re\}$, where $\langle x|\psi\rangle=\psi(x)$ [5].

One of the most important observables is the Hermitian coordinate operator $\hat{x}$ that represents the coordinate degree of freedom on the state space of the quantum particle. An experiment to observe the coordinate operator is discussed in Sect. 9.2.

The observable $\hat{x}$ is defined as a multiplication of the state vector $\psi(x)\in\mathcal{V}$ by x, that is,

$$
\hat{x}\psi(x)\equiv x\psi(x)
$$

The operator $\hat{x}$ has a *continuous spectrum* of eigenvalues and eigenstates.

Similar to a $N\times N$ matrix M that is fully specified by its matrix elements $M_{ij}, i,j=1,\ldots,N$, an operator is also specified by its matrix elements. In the bracket notation,

$$
\begin{aligned}
\hat{x}\psi(x) &= x\psi(x)\\
\Rightarrow\langle x|\hat{x}|\psi\rangle &= x\langle x|\psi\rangle = x\psi(x)
\end{aligned}
$$

In other words, the matrix element $\langle x|\hat{x}|\psi\rangle$ of the operator $\hat{x}$ is given by $x\psi(x)$. Choose the function $|\psi\rangle=|x'\rangle$ that yields

$$\begin{aligned}\langle x|\hat{x}|x'\rangle &= x\langle x'|x'\rangle = x\delta(x-x')\\ \Rightarrow \hat{x}|x'\rangle &= x'|x'\rangle\end{aligned} \tag{5.17}$$

From above it follows that the observable $\hat{x}$ has eigenfunctions $|x\rangle$ with eigenvalues $x \in \Re$; hence the spectral resolution of observable $\hat{x}$ and its *completeness equation* are given by

$$\hat{x} = \int_{-\infty}^{\infty} \mathrm{d}x|x\rangle x\langle x| \; ; \; \int_{-\infty}^{\infty} \mathrm{d}x|x\rangle\langle x| = \mathbb{I} \tag{5.18}$$

The last equation above is the completeness equation given earlier in (4.19).

The coordinate position projection operators are given by

$$\Pi_x = |x\rangle\langle x|; \;\; \Pi_x\Pi_y = \delta(x-y)\Pi_x; \;\; \int_{-\infty}^{\infty} \mathrm{d}x\Pi_x = \mathbb{I} \tag{5.19}$$

where last equation follows from (5.18). Since the eigenfunctions $|x\rangle$ are continuous, the normalization of the eigenfunctions as well as of the coordinate position projection operators is divergent; a limit has to be taken, such as the one discussed in Sect. 5.5 for the case of the compact degree of freedom S^1, to obtain finite results for the continuous coordinate operator.

For N particles in three dimensions, one has the following straightforward generalization of the coordinate operator $\hat{x}$:

$$\hat{x} = [\hat{x}_1 \otimes \hat{y}_1 \otimes \hat{z}_1] \otimes [\hat{x}_2 \otimes \hat{y}_2 \otimes \hat{z}_2] \otimes \cdots \otimes [\hat{x}_N \otimes \hat{y}_N \otimes \hat{z}_N] \tag{5.20}$$

The completeness equation is given by

$$\begin{aligned}\mathcal{I}_N = \int_{-\infty}^{\infty} &\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}z_1 \ldots \mathrm{d}x_N\mathrm{d}y_N\mathrm{d}z_N\\ &\times |x_1,y_1,z_1\rangle\langle x_1,y_1,z_1| \otimes \cdots \otimes |x_N,y_N,z_N\rangle\langle x_N,y_N,z_N|\end{aligned}$$

where $|x_1,y_1,z_1\rangle = |x_1\rangle|y_1\rangle|z_1\rangle$, $\langle x_1,y_1,z_1| = \langle x_1|\langle y_1|\langle z_1|$ and so on.

Momentum Operator $\hat{p}$

Momentum is a central concept in classical physics, and important classical quantities—such as energy and angular momentum—depend on momentum. Since the state vector $\psi(x)$ depends on only x, what is the quantum generalization of classical momentum $p = m\mathrm{d}x/\mathrm{d}t$ (quantum particle has mass m)?

In the path integral formulation of quantum mechanics, discussed in Chap. 11, the momentum of a particle has the classical form, namely, $p = m\mathrm{d}x/\mathrm{d}t$, but in the path integral approach, $x(t)$, for each t, is an integration variable. The definition of momentum implies that $p(t)$ is an indeterminate quantity; to see this, suppose we observe $x(t)$ at instant t; the definition of $p(t)$ is given by $p = m\mathrm{d}x/\mathrm{d}t \simeq m(x(t+$

$\epsilon) - x(t))/\epsilon$, but $x(t+\epsilon)$ at the next instant $t+\epsilon$ is not observed and hence is an integration variable and not fixed—leading to a $p(t)$ that is indeterminate.

Another equivalent method consists of defining momentum without any reference to the next instant in time and leads to momentum being realized as an *operator* $\hat{p}$ acting on state space $\mathcal{V}$.

There are many ways of motivating the definition of the momentum operator, but in the final analysis, one has to *postulate* the definition of the momentum operator as there is no way of "deriving" this result from classical physics. Of course, the definition has to be mathematically consistent, and the final test of whether the postulate is correct is experiment; the definition adopted for momentum has been rigorously verified by many experiments.

The momentum operator $\hat{p}$ is *postulated* to be the following[4]:

$$\hat{p} = -\mathrm{i}\hbar\frac{\partial}{\partial x} \tag{5.21}$$

Note Planck's constant $\hbar$ enters due to dimensional consistency, but its actual empirical value is fixed by Nature and has to be obtained by doing an appropriate experiment.

Consider a particle moving in one dimension. The differential operator $\partial/\partial x$ maps $\psi(x) \in \mathcal{V}$ to its derivative $\partial\psi(x)/\partial x \in \mathcal{V}$. The momentum operator is $\hat{p} = -\mathrm{i}\hbar\partial/\partial x$; in Dirac's notation,

$$\langle x|\hat{p}|\psi\rangle = -\mathrm{i}\hbar\left\langle x\middle|\frac{\partial}{\partial x}|\psi\right\rangle = -\mathrm{i}\hbar\frac{\partial\psi(x)}{\partial x} \tag{5.22}$$

An important feature of differential operators, such as $\partial/\partial x$, is that they always act on the *dual space*, as is the case for (5.22).

From (5.2), a Hermitian operator satisfies the following:

$$\hat{p}^\dagger = \hat{p} \quad \Rightarrow \quad \langle\psi|\hat{p}^\dagger|\chi\rangle \equiv \langle\chi|\hat{p}|\psi\rangle^* = \langle\psi|\hat{p}|\chi\rangle \tag{5.23}$$

To prove (5.23) that $\hat{p}$ is a Hermitian operator, doing an integration by parts yields the following:

$$\begin{aligned}\langle\chi|\hat{p}|\psi\rangle^* &= \left[-\int_{-\infty}^{+\infty}\mathrm{d}x\chi^*(x)\mathrm{i}\hbar\frac{\partial\psi(x)}{\partial x}\right]^* \\ &= -\int_{-\infty}^{+\infty}\mathrm{d}x\psi^*(x)\mathrm{i}\hbar\frac{\partial\chi(x)}{\partial x} = \langle\psi|\hat{p}|\chi\rangle \quad : \text{Hermitian}\end{aligned}$$

[4]From (5.20), since the coordinate operator for the $3N$ degrees of freedom is a tensor product of the single degree of freedom, it is sufficient to define the momentum operator for one dimension and build up the momentum for the $3N$ degrees of freedom by an appropriate tensor product.

The eigenfunctions of $\hat{p}$, in the notation of (5.22), are given by

$$\langle x|\hat{p}|p\rangle = p\mathrm{e}^{\mathrm{i}px/\hbar}$$

The operator $\hat{p}$, from the completeness equation for momentum given in (4.27), has the following representation:

$$\hat{p} = \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi\hbar}|p\rangle p\langle p| \tag{5.24}$$

The momentum operator, acting on the state vector, shifts its position in space. More precisely, using (5.22), for a constant a, consider the following *shift operator*:

$$T(a) = \mathrm{e}^{\mathrm{i}\frac{a}{\hbar}\hat{p}};\ \ T(a)T(b) = T(a+b) \tag{5.25}$$

$$\langle x|T(a)|\psi\rangle = \mathrm{e}^{a\frac{\partial}{\partial x}}\psi(x) = \psi(x+a) = \langle x+a|\psi\rangle \tag{5.26}$$

$$\Rightarrow T(a)|x\rangle = |x-a\rangle;\ \ \langle x|T(a) = \langle x+a| \tag{5.27}$$

One *definition*, from first principles, of the momentum operator $\hat{p}$ is of being a translation operator as given in (5.27).

5.7 Heisenberg Commutation Equation

Observables that do not commute occur widely in quantum mechanics and, in fact, are the reason that operator algebras that occur in quantum mechanics are nontrivial. One of the most important case of non-commuting observables is that of the position and momentum operators, for which

$$[\hat{x}, \hat{p}] = \mathrm{i}\hbar\mathbb{I} \neq 0 \tag{5.28}$$

Equation (5.28) defines the famous Heisenberg commutation equation, also called the Heisenberg algebra.

To explore the concept of *non-commuting operators*, (5.28) is derived from first principles. Consider the following representation of the coordinate and momentum observables in one dimension given by (5.18) and (5.24):

$$\hat{x} = \int_{-\infty}^{\infty} \mathrm{d}x|x\rangle x\langle x|;\ \ \hat{p} = \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi\hbar}|p\rangle p\langle p|$$

The commutator of the coordinate and momentum of a quantum particle, from above equations, is given by[5]

[5]There is an elementary derivation of $[\hat{x}, \hat{p}]$ using the chain rule of differentiation; this derivation examines the operator structure of the momentum and coordinate operators.

$$[\hat{x},\hat{p}] \equiv \hat{x}\hat{p} - \hat{p}\hat{x} = \int_{-\infty}^{\infty} \mathrm{d}x \frac{\mathrm{d}p}{2\pi\hbar}\, xp \left[|x\rangle\langle x|p\rangle\langle p| - |p\rangle\langle p|x\rangle\langle x|\right] \tag{5.29}$$

Expressing the commutator entirely in the coordinate basis by transforming the momentum basis, using (4.29) as well as (4.26), yields

$$\begin{aligned}
[\hat{x},\hat{p}] &= \int_{-\infty}^{\infty} \mathrm{d}x\mathrm{d}x' \frac{\mathrm{d}p}{2\pi\hbar} xp \left[\mathrm{e}^{\mathrm{i}px/\hbar}\mathrm{e}^{-\mathrm{i}px'/\hbar}|x\rangle\langle x'| - \mathrm{e}^{\mathrm{i}px'/\hbar}\mathrm{e}^{-\mathrm{i}px/\hbar}|x'\rangle\langle x|\right] \\
&= \frac{\hbar}{i}\int_{-\infty}^{\infty} \mathrm{d}x\mathrm{d}x' x \left[|x\rangle\langle x'|\frac{\partial}{\partial x}\delta(x-x') - |x'\rangle\langle x|\frac{\partial}{\partial x'}\delta(x'-x)\right] \qquad (5.30) \\
&= \frac{\hbar}{i}\int_{-\infty}^{\infty} \mathrm{d}x\mathrm{d}x' \left[(x-x')\frac{\partial}{\partial x}\delta(x-x')\right]|x\rangle\langle x'| \\
&= \mathrm{i}\hbar\int_{-\infty}^{\infty} \mathrm{d}x|x\rangle\langle x| = \mathrm{i}\hbar\mathbb{I} \qquad (5.31)
\end{aligned}$$

where (4.25) yields (5.30) and the last equation follows from the identity[6]

$$(x-x')\frac{\partial}{\partial x}\delta(x-x') = -\delta(x-x')$$

Hence

$$[\hat{x},\hat{p}] = \mathrm{i}\hbar\mathbb{I} \quad : \quad \text{Heisenberg commutation equation} \tag{5.32}$$

For N particles moving in three space dimensions, the degrees of freedom are x_{ai}, p_{ai} with $a = 1,2,\ldots,N$ and $i = 1,2,\ldots,3$. The Heisenberg commutation equation is given by

$$\begin{aligned}
&[\hat{x}_{ai},\hat{p}_{bj}] = \mathrm{i}\hbar\delta_{a-b}\delta_{i-j}\mathbb{I} \\
&[\hat{x}_{ai},\hat{x}_{bj}] = 0; \quad [\hat{p}_{ai},\hat{p}_{bj}] = 0
\end{aligned}$$

Noteworthy 5.1: Position and momentum are incompatible

Since $[\hat{x},\hat{p}] = \mathrm{i}\hbar$ II position and momentum do not commute and there is no state vector that is the simultaneous eigenfunction of both $\hat{x}$ and $\hat{p}$. The quantum particle can have an eigenfunction of *either* the position *or* the momentum operator, but not of both. The non-commutativity of $\hat{x}$ and $\hat{p}$ is an operator expression of the fact that if the position of the quantum particle is known at instant t, its momentum is not known.

[6]The identity results from the equation

$$\frac{\partial}{\partial x}[(x-x')\delta(x-x')] = 0$$

.

This is because the quantum particle *does not* have a unique trajectory, and hence, there is no unique derivative of the position and consequently no unique momentum. Consider observing the particle's position; at the next instant, the quantum particle is in a *trans-empirical state* and "is" at all possible points simultaneously, as shown in Fig. 11.6—and hence has no determinate value for its momentum.

5.8 Expectation Value of Operators

Important operators associated with a particle's degree of freedom are its position, momentum, energy, angular momentum, and so on and are represented by Hermitian operators. Physical quantities are indeterminate; the best that we can do in quantum mechanics is to measure the average value of the operator that represents a physical quantity, termed as the *expectation value* of the operator.

A *fundamental postulate* of quantum mechanics that follows from (2.3) and discussed in detail in Sect. 9.6 is the following: On repeatedly measuring the value of the observable $\mathcal{O}$ for some state $|\chi\rangle$, the expectation value (average value) of the observable is given by

$$E[\mathcal{O}] \equiv \langle\chi|\mathcal{O}|\chi\rangle \tag{5.33}$$

In other words, the expectation value of the observable is the diagonal value of the operator $\mathcal{O}$ for the given state $|\chi\rangle$. The expected value of a physical quantity, from (5.3), is always a real quantity, and this is the reason for representing all observables by Hermitian operators.

Consider some physical quantity, such as a particle's spin, and let it be represented by an operator $\mathcal{O}$ with eigenvalues λ_i and eigenstates ψ_i. A typical physical state is a superposition of the eigenstates with amplitude c_i and is written as

$$|\psi\rangle = \sum_i c_i|\psi_i\rangle; \;\; \mathcal{O}|\psi_i\rangle = \lambda_i|\psi_i\rangle \tag{5.34}$$

The result of measuring the physical quantity $\mathcal{O}$ for the state $\psi(x)$ *always* results in the state vector $\psi(x)$ "collapsing" (being projected), with probability $|c_i|^2$, to one of eigenstates of the operator O, say $\psi_i(x)$—whose eigenvalue λ_i is then physically observed. A detailed discussion is given in Sect. 9.6.

After repeated measurements on the system—each prepared in an identical manner and hence represented by $\psi(x)$—the average value of $\mathcal{O}$ is given by

$$E_\chi[\mathcal{O}] = \sum_i |c_i|^2\lambda_i = \sum_i |c_i|^2\langle\psi_i|\mathcal{O}|\psi_i\rangle = \langle\psi|\mathcal{O}|\psi\rangle \tag{5.35}$$

5.9 The Schrödinger Equation

The Schrödinger equation, introduced in Sect. 2.9, is reexpressed in the language of state space and operators that has been developed in Chap. 4 and in this chapter.

The Schrödinger equation determines the time evolution of the state function $|\psi(t)\rangle$, with the symbol t being the time parameter. To write down the Schrödinger equation, one first needs to specify the degrees of freedom of the system in question that in turn specifies the nature of the state space $\mathcal{V}$; one also needs to specify the Hamiltonian H of the system that describes the range and form of energy the system can have.

The celebrated Schrödinger equation is given by

$$-\frac{\hbar}{i}\frac{\partial|\psi(t)\rangle}{\partial t} = H|\psi(t)\rangle \tag{5.36}$$

For a quantum particle with mass m moving in one dimension in a potential $V(x)$, the Schrödinger equation, stated earlier (2.5), is given as follows:

$$-\frac{\hbar}{i}\left\langle x\middle|\frac{\partial}{\partial t}\middle|\psi(t)\right\rangle = \left\langle x\middle|H\middle|\psi(t)\right\rangle \;\Rightarrow\; -\frac{\hbar}{i}\frac{\partial\psi(t,x)}{\partial t} = H\left(x,\frac{\partial}{\partial x}\right)\psi(t,x) \tag{5.37}$$

where the Hamiltonian operator acts on the dual basis, as in (5.22). In the position basis, the state vector is

$$\langle x|\psi(t)\rangle = \psi(t,x)$$

The Hamiltonian for the important case of a quantum particle moving in one dimension, from (2.6), is given by

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

A variety of techniques have been developed for solving the Schrödinger equation for a wide class of potentials as well as for multiparticle quantum systems [15].

Let $|\psi\rangle$ be the initial value of the state vector at $t = 0$ with $\langle\psi|\psi\rangle = 1$. Equation (5.36) can be integrated to yield the following formal solution:

$$|\psi(t)\rangle = \mathrm{e}^{-\mathrm{i}tH/\hbar}|\psi\rangle = U(t)|\psi\rangle \tag{5.38}$$

Similar to the momentum operator translating the state vector in space, as in (5.26), the Hamiltonian H is an operator that translates the initial state vector in time, as in (5.38). The evolution operator $U(t)$ is defined by

$$U(t) = \mathrm{e}^{-\mathrm{i}tH/\hbar} \;\;;\;\; U^{\dagger}(t) = \mathrm{e}^{\mathrm{i}tH/\hbar}$$

and is unitary since H is Hermitian; more precisely,

$$U(t)U^{\dagger}(t) = \mathbb{I}$$

The unitarity of $U(t)$, and by implication the Hermiticity of H, is crucial for the conservation of probability. The total probability of the quantum system is conserved over time since unitarity of $U(t)$ ensures that the normalization of the state function is time independent; more precisely,

$$\langle \psi(t)|\psi(t)\rangle = \langle \psi|U^{\dagger}(t)U(t)|\psi\rangle = \langle \psi|\psi\rangle = 1$$

The operator $U(t)$ is the exponential of a Hermitian operator that in many cases, as given in (2.6), is a differential operator. The Feynman path integral is a mathematical tool for analyzing $U(t)$ and is discussed in Chap. 11.

The Schrödinger equation has the following remarkable features:

- The Schrödinger equation is a first-order differential equation in time, in contrast to Newton's equation of motion that is a second-order differential equation in time. At $t = 0$, the Schrödinger equation requires the initial state function for *all values* of the degree of freedom be specified, namely, $|\psi(\Re)\rangle$, whereas in Newton's law, only the position and velocity *at the starting point* of the particle are required.
- At each instant, Schrödinger's equation specifies the state function for all values of the *indeterminate* degree of freedom. In contrast, Newton's law of motion specifies only the *determinate* position and velocity of a particle.
- Equation (5.36) is a *linear equation*; two solutions $|\psi(t)\rangle$ and $|\chi_t\rangle$ of the Schrödinger equation can be added to yield yet another solution given by $a|\psi(t)\rangle + b|\chi(t)\rangle$. The linearity of the Schrödinger equation is the reason that all the state vectors $|\psi(t)\rangle$ are elements of a linear vector space $\mathcal{V}$.
- The state vector $|\psi(t)\rangle$ is a complex-valued vector. In fact, the Schrödinger equation is the first equation in natural science for which complex numbers are essential and not just a convenient mathematical tool for representing real quantities.

5.10 Heisenberg Operator Formulation

Every physical property of a degree of freedom is mathematically realized by a Hermitian operator $\mathcal{O}$. Generalizing (5.35) to time-dependent state vectors and from (5.38), the expectation value of an operator at time t, namely, $\mathcal{O}(t)$, is given by

$$\begin{aligned} E_{\psi}[\mathcal{O}(t)] &= \langle \psi(t)|\mathcal{O}|\psi(t)\rangle = \langle \psi|\mathrm{e}^{\mathrm{i}tH/\hbar}\mathcal{O}\mathrm{e}^{-\mathrm{i}tH/\hbar}|\psi\rangle \\ &= \mathrm{tr}\big(\mathcal{O}(t)\rho\big) \quad : \quad \rho = |\psi\rangle\langle\psi| \end{aligned} \tag{5.39}$$

The density matrix ρ is time independent and encodes the initial quantum state of the degree of freedom.

From (5.39), the time-dependent expectation value has two possible interpretations: the state vector is evolving in time, namely, the state vector is $|\psi(t)\rangle$ and the operator $\mathcal{O}$ is constant, or equivalently, the state vector is fixed, namely, $|\psi\rangle$, and instead, the operator is evolving in time and is given by $\mathcal{O}(t)$. The *time-dependent* Heisenberg operators $\mathcal{O}(t)$ are given by

$$\begin{aligned} O(t) &= \mathrm{e}^{\mathrm{i}tH/\hbar}\mathcal{O}\mathrm{e}^{-\mathrm{i}tH/\hbar} \\ \mathrm{i}\hbar\frac{\partial O(t)}{\partial t} &= [\mathcal{O}(t), H] \end{aligned} \tag{5.40}$$

In the Heisenberg formulation of quantum mechanics, quantum indeterminacy is completely described by the algebra of Hermitian operators.

All physical observables of a quantum degree of freedom are elements of the Heisenberg operator algebra and so are the density matrices that encode the initial quantum state of the degree of freedom. Quantum indeterminacy is reflected in the spectral decomposition of the operators in terms of its eigenvalues and projection operators (eigenvectors), as given in (5.8). For example, the single value of energy for a classical entity is replaced by a whole range of eigenenergies of the Hamiltonian operator for a quantum degree of freedom, with the eigenfunctions encoding the inherent indeterminacy of the degree of freedom.

The time dependence of the state vector given by the Schrödinger equation is replaced by the time dependence of the operators given in (5.40). All expectation values are obtained by performing a trace over this operator algebra, namely, by $tr\big(\rho\mathcal{O}(t)\big)$ as given in (5.39).

From the point of quantum probability, as discussed in Chap. 7, Heisenberg's operator formulation goes far beyond just providing a mathematical framework for the mechanics of the quantum, but instead, also lays the foundation of the quantum theory of probability.

5.11 Summary

Observable properties of a degree of freedom are represented in quantum mechanics by Hermitian operators, also termed as observables. The operators mathematically map the transempirical form of the state vector of the quantum entity to its empirical manifestation. The mapping of the operator is empirically realized by an experimental apparatus that is custom-built to model the mathematical operator. The empirical value of an observable quantity is obtained by the operator acting on the underlying quantum state vector.

The structure of an operator is realized by its spectral decomposition, in terms of all of its eigenfunctions and eigenvalues, and which also yields a representation of the completeness relation of the underlying state space. The position and momentum operator were discussed at length as these are the leading exemplars of Hermitian operators as well as among the most important observables.

The set of all mutually commuting observables—and which, in turn, all commute with the Hamiltonian—provide an exhaustive description of all the conserved quantities of a quantum entity and are the quantum generalization of the constants of motion of classical mechanics.

The state space and operators for a given degree of freedom provide the appropriate mathematical framework for the Schrödinger and Heisenberg formulations of quantum mechanics.

Density Matrix: Entangled States

6

The Hilbert space for a quantum system contains states that behave in a manner that is similar to classical objects. There are, however, also states in Hilbert space that are enigmatic and nonclassical in the sense of being forbidden by classical physics. Two leading exemplars of these nonclassical states are the following:

- Superposed states are obtained by adding state vectors, an operation allowed since elements of Hilbert space are vectors that can be added. Addition of state vectors gives rise to the principle of quantum superposition. Superposed states have been discussed in Sect. 3.7 and in more detail in Chap. 8.
- There is another set of trans-empirical and paradoxical results for quantum systems having at least *two or more degrees of freedom*, leading to the existence of what are called *entangled* states. It is this aspect of Hilbert space that is the subject of this chapter.

Nonclassical and trans-empirical states, such as superposed and entangled state vectors, exhibit behavior that are full of surprises and which are radically different from what one observes and expects in classical physics.

Entangled states arise due to the tensor product of state vectors and for which, in particular, the classical concept of separating the quantum system into its component parts does not hold. The density matrix (operator) ρ is a useful mathematical tool for studying the properties of entangled degrees of freedom.

In Chap. 7, it is shown that entangled states are an important resource for studying quantum indeterminacy. The density matrix (operator) ρ plays an essential role in understanding the paradoxes that arise in studying quantum indeterminacy, exemplified by the famous EPR (Einstein-Podolsky-Rosen) paradox.

A formulation of a quantum state that is mathematically more appropriate for describing the quantum entity, before and after a measurement is made, is provided by the density matrix and is discussed in Sect. 9.6.

The main focus of this chapter is to develop the mathematical machinery required for studying the tensor product of states and operators and the various results that follow from it.

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8_6, © Springer Science+Business Media New York 2013

6.1 Tensor Product

Distinct vector spaces $\mathcal{V}$ and $\mathcal{W}$, having different underlying degrees of freedom, yield a tensor product space denoted by $\mathcal{V}\otimes\mathcal{W}$; for ket vectors $|\psi\rangle \in \mathcal{V}$ and $|\chi\rangle \in \mathcal{W}$, elements of $\mathcal{V}\otimes\mathcal{W}$ are given by the tensor product ket vector

$$|\psi\rangle \otimes |\chi\rangle = |\psi\rangle|\chi\rangle$$

The tensor product space $\mathcal{V}\otimes\mathcal{W}$ inherits the linear structure of the constituent vector spaces $\mathcal{V}$ and $\mathcal{W}$.

The finite-dimensional state space is discussed so that the formulas can be explicitly written; the infinite-dimensional case has a similar mathematical structure. Consider an N-dimensional vector $|v\rangle \in \mathcal{V}_N$ and an M-dimensional vector $|w\rangle \in \mathcal{W}_M$; the tensor product state space is $\mathcal{V}_{MN} = \mathcal{V}_M \otimes \mathcal{W}_N$ and is an MN-dimensional vector space.

Consider ket vectors given by

$$|v\rangle = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} \;\; ; \;\; |w\rangle = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{pmatrix}$$

The tensor product of an N-dimensional vector with an M-dimensional vector yields an NM-dimensional vector that is an element of vector space $\mathcal{V}_{\mathrm{MN}}$. The tensor product vector $|v\rangle \otimes |w\rangle$ is defined by multiplying, from the left, each element of $|v\rangle$ into all the elements of $|w\rangle$ and yields the following[1]:

$$|v\rangle|w\rangle \equiv |v\rangle \otimes |w\rangle = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix}_N \otimes \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{pmatrix}_M = \begin{pmatrix} v_1 w_1 \\ v_1 w_2 \\ \vdots \\ v_1 w_M \\ \vdots \\ v_N w_1 \\ v_N w_2 \\ \vdots \\ v_N w_M \end{pmatrix}_{\mathrm{NM}} \tag{6.1}$$

[1]Multiplying from the right by elements of $|w\rangle$ is another, but distinct, way of defining the outer product. One needs to consistently use only the rule one has chosen.

Tensor Product of Operators

Consider two matrices A and B that act on space $\mathcal{V}$ and $\mathcal{W}$, respectively; the tensor product of the matrices $A \otimes B$ acts on the space $\mathcal{V} \otimes \mathcal{W}$. The matrix elements of the tensor product matrix is given by

$$\langle w'|\langle v'|(A \otimes B)|v\rangle|w\rangle = \langle v'|A|v\rangle\langle w'|B|w\rangle$$

Let the matrix elements of A and B be given by

$$A = \begin{bmatrix} a_{11} & a_{12} & .. & a_{1M} \\ a_{21} & a_{22} & .. & a_{2M} \\ .... & & & \\ a_{M1} & a_{M2} & .. & a_{MM} \end{bmatrix} \; ; \; B = \begin{bmatrix} b_{11} & b_{12} & .. & b_{1N} \\ b_{21} & b_{22} & .. & b_{2N} \\ .... & & & \\ b_{N1} & b_{N2} & .. & b_{NN} \end{bmatrix}$$

Similar to the case of tensor product state vector, the matrix elements of the tensor product matrix $A \otimes B$ are given as follows:

$$A \otimes B = \begin{bmatrix} a_{11} & a_{12} & .. & a_{1M} \\ a_{21} & a_{22} & .. & a_{2M} \\ .... & & & \\ a_{M1} & a_{M2} & .. & a_{MM} \end{bmatrix} \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & .. & a_{1M}B \\ a_{21}B & a_{22}B & .. & a_{2M}B \\ .... & & & \\ a_{M1}B & a_{M2}B & .. & a_{MM}B \end{bmatrix}$$

Writing out all the matrix elements explicitly yields

$$A \otimes B = \begin{bmatrix} a_{11}b_{11} & .. & .. & a_{1M}b_{1N} \\ .. & .. & .. & .. \\ .. & .. & .. & .. \\ a_{M1}b_{N1} & .. & .. & a_{MM}b_{NN} \end{bmatrix} \; : \; MN \times MN \text{ matrix}$$

6.2 The Outer Product

The outer product of two linear vector spaces is an essential construct in quantum mechanics and is necessary for studying the Hilbert space of a single degree of freedom as well as the state space for two or more degrees of freedom.

For a *single state space*, the outer product space of the vector space with its dual is given by $\mathcal{V} \otimes \mathcal{V}_{\mathrm{D}}$; the elements of the outer product space are denoted by $|\psi\rangle\langle\eta|$, with $|\psi\rangle \in \mathcal{V}; \langle\eta| \in \mathcal{V}_{\mathrm{D}}$.

Consider a single discrete degree of freedom taking values $1, 2, \ldots, N$ with a finite set of basis state vectors $|i\rangle, \; i = 1, 2, \ldots, N$. The basis set is called complete if any vector in $\mathcal{V}$ can be written as a linear combination of the basis states. In particular, for a Hilbert space, a complete set of basis states satisfy the following identity:

$$\sum_{i=1}^{N} |i\rangle\langle i| = \mathbb{I} \ : \ \text{Completeness Equation} \tag{6.2}$$

where $\mathbb{I}$ is the *identity operator* on $\mathcal{V} \otimes \mathcal{V}_{\rm D}$, defined in general by

$$\mathbb{I}|\psi\rangle = |\psi\rangle \ ; \ \langle\chi|\mathbb{I} = \langle\chi|$$

for all $|\psi\rangle, \langle\chi|$. The completeness equation for a continuous degree of freedom is given in (4.19).

Two *distinct vector spaces* $\mathcal{V}$ and $\mathcal{W}$, having different underlying degrees of freedom, yield an outer product space denoted by $\mathcal{V} \otimes \mathcal{W}$; for ket vectors $|\psi\rangle \in \mathcal{V}$ and $|\chi\rangle \in \mathcal{W}$, elements of $\mathcal{V} \otimes \mathcal{W}$ are given by the outer product of the ket times the bra vector and is written as

$$|\chi\rangle \otimes \langle\psi| \equiv |\chi\rangle\langle\psi|$$

Consider the following N-dimensional ket vector $|\psi\rangle$ and the M-dimensional bra (dual) vector $\langle\chi|$:

$$|\psi\rangle = \begin{bmatrix} a_1 \\ a_2 \\ .. \\ a_N \end{bmatrix} \ ; \ \langle\chi| = \begin{bmatrix} b_1^* & b_2^* & .. & b_M^* \end{bmatrix}$$

The outer product of an N-dimensional vector with an M-dimensional vector yields an NM-dimensional matrix. The outer product is defined by multiplying all the components of the ket (column) vector $|\psi\rangle$ from the left into each element of the bra (row) vector $\langle\chi|$ on the right to yield the following[2]:

$$\begin{aligned} |\psi\rangle\langle\chi| &= \begin{bmatrix} a_1 \\ a_2 \\ .. \\ a_N \end{bmatrix} \otimes \begin{bmatrix} b_1^* & b_2^* & .. & b_M^* \end{bmatrix} \\ &= \begin{bmatrix} a_1 b_1^* & a_1 b_2^* & .. & a_1 b_M^* \\ a_2 b_1^* & a_2 b_2^* & .. & a_2 b_M^* \\ .... & & & \\ a_N b_1^* & a_N b_2^* & .. & a_N b_M^* \end{bmatrix} \ : \ N \times M \text{ matrix} \end{aligned} \tag{6.3}$$

The NM-dimensional matrices yield linear transformations of underlying vector space $\mathcal{V}_{\rm NM}$.

[2]One can equivalently define the outer product by multiplying the entire row vector from the right into each element of the column vector on the left and obtain the same result.

6.3 Partial Trace for Outer Products

Consider the outer product of two state vectors given by

$$\mathcal{O} = |\psi\rangle\langle\chi| \;\; ; \;\; \langle x'|\mathcal{O}|x\rangle = \psi(x')\chi^*(x)$$

In the position basis, for a continuous degree of freedom x, trace is defined in (5.4) as follows:

$$\mathrm{tr}[\mathcal{O}] = \int \mathrm{d}x \langle x|\mathcal{O}|x\rangle \tag{6.4}$$

In particular, the trace operation for the outer product of two states is, from (6.4), the following:

$$\mathcal{O} = |\psi\rangle\langle\chi| \;\; \Rightarrow \;\; \mathrm{tr}[\mathcal{O}] = \int \mathrm{d}x \psi(x)\chi^*(x)$$

If the state vectors are elements of an N-dimensional Euclidean space, then $\mathcal{O}$ is simply an $N \times N$ matrix, and the trace of $\mathcal{O}$ is the sum of its diagonal elements, since for matrix M_{ij}, trace is defined by $\sum_i M_{ii}$; hence, it follows that

$$\mathrm{tr}[\mathcal{O}] = \sum_{i=1}^{N} \langle i|\mathcal{O}|i\rangle = \sum_{i=1}^{N} \psi(i)\chi^*(i)$$

Consider a system with two degrees of freedom with state vectors $|\psi_1\rangle|\psi_2\rangle$; the outer product is given by

$$\mathcal{O} = |\psi_1\rangle\langle\psi_1| \otimes |\psi_2\rangle\langle\psi_2|$$

One can now perform a *partial trace* on $\mathcal{O}$, say over system 2, and yields

$$\mathrm{tr}_2(\mathcal{O}) = |\psi_1\rangle\langle\psi_1|\big[\langle\psi_2|\psi_2\rangle\big] = c|\psi_1\rangle\langle\psi_1| \;\; \text{with} \;\; \langle\psi_2|\psi_2\rangle = c$$

One can further generalize the concept of a partial trace; consider the following linear sum of the outer product of states:

$$\mathcal{O} = \sum_{i=1}^{N} p_i |\psi_1^i\rangle|\langle\psi_1^i| \otimes |\psi_2^i\rangle\langle\psi_2^i|$$

where p_i are numbers. The partial trace over system 2, using the linearity of trace as given in (5.4), is defined as follows:

$$\begin{aligned} \mathrm{tr}_2(\mathcal{O}) &= \sum_{i=1}^{N} p_i |\psi_1^i\rangle\langle\psi_1^i|\langle\psi_2^i|\psi_2^i\rangle \\ &= \sum_{i=1}^{N} p_i c_i |\psi_1^i\rangle\langle\psi_1^i| \quad \text{with} \;\; \langle\psi_2^i|\psi_2^i\rangle = c_i \end{aligned} \tag{6.5}$$

6.4 Density Matrix ρ

The density matrix plays a central role in the Heisenberg operator formulation, discussed in Sect. 5.10, with the Schrödinger state vector being replaced by the density matrix. The density matrix is a special Hermitian operator that has many applications and provides a quantum mechanical generalization of the concept of conditional probabilities for a quantum mechanical degrees of freedom.[3]

Pure Density Matrix

The pure density matrix is a Hermitian operator that is equivalent to the state vector and provides an operator description of the quantum entity.

For a state vector $|\chi\rangle$, the pure density matrix is defined by

$$\rho_{\rm P} = |\chi\rangle\langle\chi| \tag{6.6}$$

A pure density matrix $\rho_{\rm P}$ is a projection operator and has the following properties:

$$\begin{aligned} \rho_{\rm P} = |\chi\rangle\langle\chi| \;\; ; \;\; \rho_{\rm P}^2 = \rho_{\rm P} \\ {\rm tr}(\rho_{\rm P}^2) = {\rm tr}(\rho_{\rm P}) = 1 \;\; : \;\; \text{Pure state} \end{aligned} \tag{6.7}$$

Expressing the state vector in terms of a complete basis state given by $|\chi_i\rangle$ yields the following:

$$\begin{aligned} |\chi\rangle = \sum_i c_i|\chi_i\rangle \;\; ; \;\; \rho_{\rm P} \equiv |\chi\rangle\langle\chi| \\ \Rightarrow \rho_{\rm P} = \sum_{ij} c_i c_j^* |\chi_i\rangle\langle\chi_j| = \sum_i |c_i|^2 |\chi_i\rangle\langle\chi_j| + \sum_{ij;\, i\neq j} c_i c_j^* |\chi_i\rangle\langle\chi_j| \end{aligned} \tag{6.8}$$

The off-diagonal terms $i \neq j$ given in (6.8) are completely quantum mechanical in origin and are due to correlations between two different eigenstates $|\chi_i\rangle$ and $|\chi_j\rangle$.

The expectation value of any operator $\mathcal{O}$ in a state $|\psi\rangle$ can be obtained from the pure state density matrix and is discussed in Sect. 6.11. The density matrix for a pure state, namely, $\rho_{\rm P}$, is *equivalent* to the state vector $|\psi\rangle$ and encodes the result of all observations that can be made on the quantum system.

[3]The density matrix should be termed the density operator since, in general, it is not a finite or infinite matrix; however, the term density matrix is so widely used that its proper definition is implicitly understood.

Mixed Density Matrix

A *mixed density matrix* is defined for a collection of orthonormal projection operators $|\psi_i\rangle\langle\psi_i|$ and is given by

$$\rho_{\mathrm{M}} = \sum_{i=1}^{N} p_i |\psi_i\rangle\langle\psi_i| \tag{6.9}$$

$$0 < p_i < 1 \;\; ; \;\; \sum_{i=1}^{N} p_i = 1 \;\; \Rightarrow \;\; \mathrm{tr}(\rho_{\mathrm{M}}) = \sum_{i=1}^{N} p_i = 1$$

Note the cardinal point that for the mixed density matrix, there are no off-diagonal terms such as the terms $|\chi_i\rangle\langle\chi_j|,\ i \neq j$ given in (6.8).

The mixed density matrix has the following defining property:

$$\rho_{\mathrm{M}}^2 = \sum_i p_i^2 |\psi_i\rangle\langle\psi_i| \quad \Rightarrow \mathrm{tr}(\rho_{\mathrm{M}}^2) = \sum_i p_i^2 < 1 \tag{6.10}$$

Only for a pure state, where only one of the p_i is 1, is $\mathrm{tr}(\rho^2) = 1$. Hence, a definition of a mixed state is

$$\mathrm{tr}(\rho_{\mathrm{M}}^2) < 1 \;\; : \;\; \text{Mixed state} \tag{6.11}$$

In (6.5), it was shown that if one starts with a pure density matrix and a partial trace is performed over one of the degrees of freedom, then one obtains a mixed density matrix. Performing a partial trace erases information about the degree of freedom, and hence, the density matrix of a mixed state contains less information than a pure state.

The density matrix for a mixed state is required for mathematically representing the result of quantum measurements, discussed in Sect. 9.6, and is a precise measure of how much information is lost in performing an observation on a quantum system.

Another important application of the mixed density matrix is in the description of quantum mechanical states that, in addition to quantum indeterminacy, also have classical randomness—as is the case for the thermodynamics of a quantum system—and is discussed in Sect. 6.11.

Density Matrix for a Two-State System

The general expression for a ket vector $|\psi\rangle$ of a two-state, parametrized by the Bloch sphere, is given by (4.7) as follows:

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)\begin{bmatrix}1\\0\end{bmatrix} + \mathrm{e}^{\mathrm{i}\phi}\sin\left(\frac{\theta}{2}\right)\begin{bmatrix}0\\1\end{bmatrix}$$

The density matrix for the pure state is

$$\rho_{\mathrm{P}} = |\psi\rangle\langle\psi| = \frac{1}{2}\left[\mathbb{I} + i\sum_{i=1}^{3}\hat{n}_i\sigma_i\right] \;\; ; \;\; \mathrm{tr}\rho_{\mathrm{P}}^2 = \hat{n}^2 = 1 \tag{6.12}$$

where the σ_i are the Pauli spin matrices given by

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} ; \; \sigma_2 = \begin{bmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{bmatrix} ; \; \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} ; \; \mathrm{tr}(\sigma_i\sigma_j) = 2\delta_{i-j} \tag{6.13}$$

The unit vector $\hat{n}$ is an arbitrary three-dimensional vector that lies *on* the Bloch sphere, shown in Fig. 4.6, and is given by

$$\hat{n} = (\sin\theta\cos\phi, \cos\theta\cos\phi, \sin\phi) \; ; \; 0 \leq \theta \leq \pi \; ; \; 0 \leq \phi \leq 2\pi$$

A vector lying *inside* the Bloch sphere is given by

$$a\hat{n} \;\; ; \;\; a \in [0,1]$$

It can be shown that the most general two-state mixed density matrix is given by

$$\rho_{\mathrm{M}} = \frac{1}{2}\left[\mathbb{I} + ia\sum_{i=1}^{3}\hat{n}_i\sigma_i\right] \;\; ; \;\; \mathrm{tr}\rho_{\mathrm{M}}^2 = a^2\hat{n}^2 = a^2 < 1 \tag{6.14}$$

For a mixed state, the density matrix is ρ_{M} with $0 \leq a < 1$, and hence, all the density matrices for *mixed states* lie *inside* the Bloch sphere. For a pure state $\mathrm{tr}\rho_{\mathrm{P}}^2 = 1$ and which yields $a = 1$. Hence, all the density matrices for *pure states* are on the *surface* of the Bloch sphere.

The two-state density matrix has a major application in the study of quantum information, in particular on studying the effect of measurements on qubits.

6.5 The Schmidt Decomposition

A general normalized state vector for two degrees of freedom, including both entangled and non-entangled states, is given by

$$|\Psi\rangle = \sum_{i,j=1}^{N} c_{ij}|e_i\rangle|e_j\rangle \;\; ; \;\; \sum_{i,j=1}^{N}|c_{ij}|^2 = 1 \tag{6.15}$$

$$\langle e_i|e_j\rangle = \delta_{i-j}$$

where $|e_i\rangle$ is a complete basis for both the degrees of freedom.[4]

The general expression for the state vector given in (6.15) can be simplified by choosing a new basis. In particular, one can choose the Schmidt basis, which depends on the state vector $|\Psi\rangle$, to obtain a much simpler representation that requires only a single summation and is given by

$$|\Psi\rangle = \sum_{i=1}^{N} c_i |\psi_i^I\rangle |\psi_i^{II}\rangle \;\; ; \;\; \sum_i |c_i|^2 = 1 \tag{6.16}$$
$$\langle \psi_i^I | \psi_j^I \rangle = \delta_{i-j} = \langle \psi_i^{II} | \psi_j^{II} \rangle$$

The coefficients c_i and state vectors $|\psi_i^I\rangle, |\psi_i^{II}\rangle$ are all functions of $|\Psi\rangle$.

The expression given in (6.15) is shown to be reducible to (6.16). The pure density matrix for the state vector given in (6.15) is the following:

$$|\Psi\rangle\langle\Psi| = \sum_{i,j=1}^{N} \sum_{i',j'=1}^{N} c_{ij} c_{i'j'}^* |e_i\rangle\langle e_{i'}| \otimes |e_j\rangle\langle e_{j'}|$$

Consider the following operator and its eigenfunctions:

$$\rho = \sum_{i=1}^{N} \sum_{i'=1}^{N} \left[\sum_{j=1}^{N} c_{ij} c_{i'j}^* \right] |e_i\rangle\langle e_{i'}| \;\; ; \;\; \mathrm{tr}(\rho) = 1 \tag{6.17}$$
$$\rho |\psi_i^I\rangle = \alpha_i |\psi_i^I\rangle \;\; ; \;\; \sum_{i=1}^{N} |\psi_i^I\rangle\langle\psi_i^I| = \mathbb{I}$$

From (6.17)

$$\sum_{k=1}^{N} \alpha_k = \mathrm{tr}(\rho) = 1 \tag{6.18}$$

Using the completeness equation of $|\psi_i^I\rangle$ yields the following:

$$|\Psi\rangle = \sum_{i,j=1}^{N} c_{ij} \left[\sum_{k=1}^{N} |\psi_k^I\rangle\langle\psi_k^I| \right] |e_i\rangle|e_j\rangle$$
$$= \sum_{k=1}^{N} |\psi_k^I\rangle |\Lambda_k\rangle \tag{6.19}$$

where

[4] One can take the complete basis states of the degrees of freedom that is of larger dimension N and use it for the other degree of freedom, with some left over unused basis vectors.

$$|\Lambda_k\rangle = \sum_{i,j=1}^{N} c_{ij}\langle\psi_k^I|e_i\rangle|e_j\rangle \;\; ; \;\; k = 1,2,\ldots,N$$

The state vectors $|\Lambda_k\rangle$ are orthogonal; to prove this note that

$$\begin{aligned}
\langle\Lambda_k|\Lambda_l\rangle &= \sum_{i',j'=1}^{N}\sum_{i,j=1}^{N} c_{i'j'}^* c_{ij}\langle e_{i'}|\psi_k^I\rangle\langle e_{j'}|e_j\rangle\langle\psi_l^I|e_i\rangle \\
&= \sum_{i',i'=1}^{N}\sum_{j=1}^{N} c_{i'j}^* c_{ij}\langle e_{i'}|\psi_k^I\rangle\langle\psi_l^I|e_i\rangle \\
&= \langle\psi_l^I|\left[\sum_{i',i'=1}^{N}\sum_{j=1}^{N} c_{i'j}^* c_{ij}|e_i\rangle\langle e_{i'}|\right]|\psi_k^I\rangle \\
&= \langle\psi_l^I|\rho|\psi_k^I\rangle = \alpha_k\langle\psi_l^I|\psi_k^I\rangle = \alpha_k\delta_{k-\ell}
\end{aligned} \tag{6.20}$$

Hence, it follows from (6.20) that

$$\langle\Lambda_k|\Lambda_\ell\rangle = \alpha_k\delta_{k-\ell}$$

Defining the orthonormal basis states by

$$|\psi_k^{II}\rangle = \frac{1}{\sqrt{\alpha_k}}\Lambda_k$$

yields, from (6.18) and (6.19), the representation

$$|\Psi\rangle = \sum_{i=1}^{N}\sqrt{\alpha_i}|\psi_i^I\rangle|\psi_i^{II}\rangle \;\; ; \;\; \sum_{k=1}^{N}\alpha_k = 1$$

$$\langle\psi_i^I|\psi_j^I\rangle = \delta_{i-j} = \langle\psi_i^{II}|\psi_j^{II}\rangle$$

and is the result stated in (6.16) with $c_i = \sqrt{\alpha_i}$.

6.6 Reduced Density Matrix

The concept of reduced density matrix can be defined for a system having two or more degrees of freedom. Consider an experiment in which the projection operators for only one of the degrees of freedom are measured, with the projection operators for the other degrees of freedom being completely ignored. Clearly, there is a loss of information regarding the state of the other degrees of freedom. The reduced density matrix provides a precise measure on how much information is lost in such a "partial" experiment.

Consider a quantum entity with only two different degrees of freedom, with its state vector, using the Schmidt decomposition given in (6.16), as follows:

$$|\Psi\rangle = \sum_{i=1}^{N} c_i |\psi_i^I\rangle |\psi_i^{II}\rangle \;\; ; \;\; \sum_i |c_i|^2 = 1$$
$$\langle \psi_i^I | \psi_j^I \rangle = \delta_{i-j} = \langle \psi_i^{II} | \psi_j^{II} \rangle$$

The pure density matrix for the system is given by

$$\rho = |\Psi\rangle\langle\Psi| = \sum_{ij=1}^{N} c_i c_j^* |\psi_i^I\rangle\langle\psi_i^I| \otimes |\psi_i^{II}\rangle\langle\psi_i^{II}| \tag{6.21}$$

If measurements are made on only the state vectors ψ_i^I, then the loss of information encoded in state vectors ψ_i^{II} is mathematically realized by performing a partial trace over the II degrees of freedom, as discussed in Sect. 6.3. Performing the partial trace in (6.21) yields the reduced density matrix ρ_{R}, namely,

$$\begin{aligned} \rho_{\mathrm{R}} &= \mathrm{tr}_{II}(\rho) = \mathrm{tr}_{II}\big(|\Psi\rangle\langle\Psi|\big) \\ &= \sum_{ij=1}^{N} c_i^* c_j |\psi_i^I\rangle\langle\psi_j^I| \Big[\langle \psi_j^{II} | \psi_i^{II} \rangle\Big] \\ \Rightarrow \rho_{\mathrm{R}} &= \sum_{i=1}^{N} |c_i|^2 |\psi_i^I\rangle\langle\psi_i^I| \end{aligned} \tag{6.22}$$

Hence, (6.22) shows that the loss of information for a pure density matrix, given in (6.21), yields a reduced density matrix that is a mixed density matrix, defined in (6.9).

The analysis carried out for discrete degrees of freedom to obtain the reduced density matrix given in (6.22) can also be done for continuous degrees of freedom. Consider, for concreteness, a quantum system with two degrees of freedom, for example, two particles with degrees of freedom x_1, x_2 (coordinates in one dimension), respectively, and with state vector $\psi(x_1, x_2)$. Consider a non-factorizable state vector and its density matrix given by

$$\begin{aligned} \langle x_1, x_2 | \psi \rangle &= \psi(x_1, x_2) \neq \psi_1(x_1)\psi_2(x_2) \\ \rho &= |\psi\rangle\langle\psi| \;\; ; \;\; \langle x_1, x_2 | \rho | x_1', x_2' \rangle = \psi(x_1, x_2)\psi^*(x_1', x_2') \end{aligned} \tag{6.23}$$

One can sum over one of the degrees of the freedom—in general, by performing a partial trace of ρ over a degree of freedom as was done in (6.5)—say over the coordinate x_2 and obtain the *reduced density matrix* ρ_{R} that provides a complete description for all measurement carried out on only the degree of freedom x_1; in symbols

$$\rho_{\mathrm{R}} = \mathrm{tr}_2(\rho) = \mathrm{tr}_2(|\psi\rangle\langle\psi|)$$
$$\langle x_1|\rho_{\mathrm{R}}|x_1'\rangle = \int \mathrm{d}x_2 \psi(x_1,x_2)\psi^*(x_1',x_2) \tag{6.24}$$

Comparing (6.24) and (7.12), it can be seen that the reduced density matrix provides a quantum mechanical generalization of the concept of the marginal probability distribution.

6.7 Separable Quantum Systems

A *separable* quantum system is defined to be a system in which the degrees of freedom can be *exactly factorized* in the sense that the state vector for the degrees of freedom is a tensor product, as given below:

$$|\psi\rangle = |\psi_I\rangle|\psi_{II}\rangle \;\; ; \;\; \langle x_1,x_2|\psi\rangle = \psi_I(x_1)\psi_{II}(x_2) \tag{6.25}$$

and yields the pure density matrix given by

$$\rho_{\mathrm{P}} = |\psi\rangle\langle\psi| = |\psi_I\rangle\langle\psi_I| \otimes |\psi_{II}\rangle\langle\psi_{II}|$$

The reduced density matrix, as discussed in (6.22), is obtained by performing a partial trace over the x_2 degree of freedom and for the separable quantum system is given by[5]

$$\rho_{\mathrm{P,R}} = \mathrm{tr}_2\big(|\psi\rangle\langle\psi|\big) = |\psi_I\rangle\langle\psi_I|\big(\langle\psi_{II}|\psi_{II}\rangle\big)$$
$$= |\psi_I\rangle\langle\psi_I| \;\; ; \;\; \mathrm{tr}(\rho_{\mathrm{P,R}}) = 1 = \mathrm{tr}(\rho_{\mathrm{P,R}}^2)$$

In other words, the reduced density matrix of a separable system is *also* a pure density matrix.

Consider two different systems with their own degrees of freedom with density matrices ρ_i^A and ρ_i^B such that

$$\mathrm{tr}(\rho_i^A) = 1 = \mathrm{tr}(\rho_i^B)$$

One can think of the density matrices as projection operators for the two different systems.

A general representation of a composite system consisting of two *separable subsystems* is given by the following bipartite (mixed) density matrix:

[5] A similar result holds for taking a partial trace over the x_1 degree of freedom.

$$\rho_{AB} = \sum_{i=1}^{N} p_i \rho_i^A \otimes \rho_i^B \;\; \Rightarrow \;\; \mathrm{tr}(\rho_{AB}) = \sum_{i=1}^{N} p_i = 1 \;\; ; \;\; p_i \in [0,1] \tag{6.26}$$

It is the condition of $\sum_{i=1}^{N} p_i = 1$ that implies that the system is separable, with a complete description of system A and B being contained solely in ρ_i^A and ρ_i^B, respectively.

The bipartite density matrix represents a *separable quantum system* for which the degrees of freedom for A and B can be considered in isolation from each other. In other words, one can unambiguously separately measure the degrees of freedom for A and B and still obtain the correct result for the expectation value of all observables pertaining to only one of the systems.

The reduced density matrix for the separable system is given by

$$\rho_{A,R} = \mathrm{tr}_B(\rho_{AB}) = \sum_{i=1}^{N} p_i \rho_i^A \;\; ; \;\; \rho_{B,R} = \mathrm{tr}_A(\rho_{AB}) = \sum_{i=1}^{N} p_i \rho_i^B \tag{6.27}$$

6.8 Entangled Quantum States

In classical mechanics, the point particles obeying Newton's laws are *always* distinct entities. In contrast, the distinct "identity" of a particular quantum mechanical degree of freedom is only meaningful for special cases.

More precisely, if the state vector for the two degrees of freedoms can be completely *factorized*, namely, if the joint state vector is a tensor product of the individual state vectors of each degree of freedom, then one of the degrees of freedom can be observed independently from the other. However, if the joint state vectors cannot be factorized, which are called *entangled states*, the two degrees of freedom become *inseparable*, and one cannot consider either of the degrees of freedom independently of the other. The state given in (6.23) is an example of an entangled state.

One needs a quantum system with two or more degrees of freedom to obtain an entangled state.

An entangled state vector does not have any dynamics, and the property of entanglement is purely kinematic, namely, it pertains entirely to the structure of the state vector and not to how it evolves in time (dynamics). The quantum entity represented by an entangled state does not exist in classical physics and shows the rich structure of quantum mechanics.

Recall from Sect. 4.7 that the basis states of state space are only defined up to a unitary transformation. Hence, a state vector that is apparently not separable could, in fact, be separable if the basis states are transformed to a new basis. To provide a precise *basis-independent* formulation of entangled states, one needs to express the quantum system in the language of the density matrix. Just such a general criterion is provided by the reduced density matrix and is derived below.

The *pure density matrix* for the state vector $|\Psi_{\rm E}\rangle$ given in (6.16) is the following:

$$\rho_{\rm E} = |\Psi_{\rm E}\rangle\langle\Psi_{\rm E}| = \sum_{ij=1}^{N} c_i c_j^* |\psi_i^I\rangle\langle\psi_j^I| \otimes |\psi_i^{II}\rangle\langle\psi_j^{II}| \;\; ; \;\; {\rm tr}(\rho_{\rm E}^2) = 1$$

As was the case for (6.22), performing a partial trace over the degree of freedom II yields, from (6.5) and (6.16), the *reduced density matrix* for the entangled state as follows:

$$\rho_{\rm E,R} = {\rm tr}_{II}\Big(|\Psi_{\rm E}\rangle\langle\Psi_{\rm E}|\Big) = \sum_{i=1}^{N} |c_i|^2 |\psi_i^I\rangle\langle\psi_i^I| \tag{6.28}$$

$${\rm tr}(\rho_{\rm E,R}^2) = \sum_{i=1}^{N} |c_i|^4 < 1 \tag{6.29}$$

${\rm tr}(\rho_{\rm E,R}^2) < 1$ is a basis-independent result, since a unitary change of basis, as discussed in Sect. 4.7, leaves ${\rm tr}(\rho_{\rm E,R}^2)$ invariant.

${\rm tr}(\rho_{\rm E,R}^2) < 1$ leads to the conclusion that the state $|\Psi_{\rm E}\rangle$ itself cannot be written, in *any* set of basis states, as a product state $|\psi_I\rangle|\psi_{II}\rangle$. This is because a partial trace of the product state would lead to a reduced matrix $\rho_{\rm E,R}$ that would be a pure density matrix—and thus contradict the result that ${\rm tr}(\rho_{\rm E,R}^2) < 1$, obtained in (6.29).

In conclusion, for $c_i \neq 0$

$$|\Psi_{\rm E}\rangle = \sum_{i=1}^{N} c_i |\psi_i^I\rangle|\psi_i^{II}\rangle \neq |\chi_I\rangle|\chi_{II}\rangle : \;\; \text{Entangled}$$

$|\Psi_{\rm E}\rangle$ is an entangled state; in general, the two or more degrees of freedom for an entangled state need to be treated as one indecomposable and inseparable system, with the identities of the individual degrees of freedom, taken in isolation, being meaningless. In contrast, for a separable system, each degree of freedom can be considered to be a distinct entity and separate from the other degree of freedom.

Entanglement for Composite Systems

The criterion of entanglement for a pure density matrix $\rho_{\rm P} = |\psi\rangle\langle\psi|$ is given by examining its reduced density matrix $\rho_{\rm R}$; if ${\rm tr}(\rho_{\rm R}^2) < 1$, then the state $|\psi\rangle$ is entangled. This criterion does not hold for density matrix of systems that are the composite of two or more different systems. In particular for bipartite states,

$$\rho_{AB} = \sum_{i=1}^{N} p_i \rho_i^A \otimes \rho_i^B \;\; \Rightarrow \;\; {\rm tr}\big((\rho_{AB})^2\big) = \sum_{i=1}^{N} p_i^2 < 1$$

Although one has ${\rm tr}((\rho_{AB})^2) < 1$, this *does not* necessarily imply that either system A or B is entangled. Separable systems have been proven to satisfy, using definitions

given in (6.27), the following two inequalities:

$$\mathbb{I}_A \otimes \rho_{B,R} - \rho_{AB} \geq 0 \;\; ; \;\; \rho_{A,R} \otimes \mathbb{I}_B - \rho_{AB} \geq 0$$

The operator inequality means that all the eigenvalues of the operator are nonnegative. If any one of these two conditions are violated, then ρ_{AB} represents a composite system that is entangled. This is called the reduction criterion [27].

6.9 A Pair of Entangled Spins

Consider a pair of spins (two-state systems) with basis states $|u_1\rangle, |d_1\rangle$ and $|u_2\rangle, |d_2\rangle$ defined in (4.1).

A general expression for a *separable* product state for the pair of spins is the following:

$$|\Psi_S\rangle = \Big[a|u_1\rangle + b|d_1\rangle\Big]\Big[\alpha|u_2\rangle + \beta|d_2\rangle\Big] \;\; ; \;\; |a|^2 + |b|^2 = 1 = |\alpha|^2 + |\beta|^2$$

In contrast, an example of an *entangled state* for the two spins, using the rules of tensor product of vectors given in (6.1), is given by

$$\begin{aligned} |\Psi_E\rangle &= a|u_1\rangle|d_2\rangle + b|d_1\rangle\ |u_2\rangle \;\; ; \;\; |a|^2 + |b|^2 = 1 \\ &= a\begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} + b\begin{pmatrix}0\\1\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\a\\b\\0\end{pmatrix} \end{aligned} \tag{6.30}$$

The entangled state vector $|\Psi_E\rangle$ has been studied extensively and plays a central role in the EPR paradox as well as in empirical tests of Bell's theorem and is discussed in Sect. 7.6.

The proof that (6.30) is an entangled state requires the evaluation of the reduced density matrix. The density matrix is given by a outer product and, using the rules given in (6.3), yields

$$\begin{aligned} \rho_E &= |\Psi_E\rangle\langle\Psi_E| \\ &= |a|^2|u_1\rangle\langle u_1| \otimes |d_2\rangle\langle d_2| + |b|^2|d_1\rangle\langle d_1| \otimes |u_2\rangle\langle u_2| + \text{off-diagonal} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & |a|^2 & ab^* & 0 \\ 0 & a^*b & |b|^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned} \tag{6.31}$$

The *reduced density matrix* is defined by taking the partial trace over the degree of freedom of the second spin; under the partial trace, the off-diagonal terms in (6.31)

are all zero. The result is the following:

$$\begin{aligned}\rho_{\mathrm{ER}} &= \mathrm{tr}_2(\rho_{\mathrm{E}}) \\ &= |a|^2|u_1\rangle\langle u_1| + |b|^2|d_1\rangle\langle d_1| \\ &= \begin{bmatrix} |a|^2 & 0 \\ 0 & |b|^2 \end{bmatrix}\end{aligned} \tag{6.32}$$

Taking the trace of the reduced matrix over the degree of freedom of the first spin yields, from (6.32), the following:

$$\begin{aligned}\mathrm{tr}\rho_{\mathrm{ER}} &= |a|^2 + |b|^2 = 1 \quad : \quad \text{Normalization} \\ \mathrm{tr}(\rho_{\mathrm{ER}}^2) &= |a|^4 + |b|^4 = 1 - 2|ab|^2 < 1\end{aligned} \tag{6.33}$$

Note if either a or b is zero, (6.33) shows that there is no entanglement, as indeed is the case since the state vector given in (6.30) becomes a product state and is separable.

Hence, we conclude from (6.33) that since the reduced density matrix $\mathrm{tr}(\rho_{\mathrm{ER}}^2) < 1$, the state vector given in (6.30) is entangled, namely,

$$|\Psi_{\mathrm{E}}\rangle = a|u_1\rangle|d_2\rangle + b|d_1\rangle\ |u_2\rangle : \quad \text{Entangled}$$

Noteworthy 6.1: Quantum superposed states and entangled states

Since the state vector is an element of a linear vector space, it can be added to other state vectors as well as "multiplied" with other state vectors to form tensor product state vectors. Both the quantum superposed states and entangled states have no classical analog and are states of the quantum system that are trans-empirical.

- The addition of quantum state vectors $a|\psi\rangle + b|\chi\rangle$ yields a trans-empirical state and is a result of the *quantum superpostion principle*.
- Tensor product of two states of the kind $|\psi_1\rangle|\chi_2\rangle + |\psi_2\rangle|\chi_1\rangle$ is a trans-empirical state and is an example of *entangled states*.

Some of the important nonclassical results are discussed in Chap. 7 on Quantum Indeterminacy, in Chap. 8 on Quantum Superposition, and in Chap. 9 on Quantum Measurement; in particular, it will be shown that the process of measurement crucially hinges on the formation of entangled states.

6.10 Quantum Entropys

Entropy is a measure of the ignorance regarding a system. The concept of entropy in statistical physics has a natural analog for quantum systems and, following von Neumann, is defined as follows:

$$\mathcal{S} = -\text{tr}(\rho \ln \rho) = -\sum_{i=1}^{N} p_i \ln p_i \tag{6.34}$$

$$\rho = U \text{diag}(p_1, p_2, \ldots, p_N) U^\dagger \;\; ; \;\; UU^\dagger = \mathbb{I}$$

Consider a pure state with $p_1 = 1$ and $p_i = 0$; $i \neq 1$; then

$$\rho = |\psi\rangle\langle\psi| \;\; \Rightarrow \;\; \mathcal{S} = -\text{tr}(\rho \ln \rho) = 0$$

A pure state yields zero entropy since, as expected, there is no ignorance in knowing the state of the system. In contrast to a pure state, if one has no information about a system, then one expects that entropy should be a maximum.

The entropy of a mixed state, from (6.9) and (6.34), is the following:

$$\mathcal{S} = -\text{tr}(\rho_\text{M} \ln \rho_\text{M}) = -\sum_{i=1}^{N} p_i \ln p_i$$

$$\rho_\text{M} = V \text{diag}(p_1, p_2, \ldots, p_N) V^\dagger \;\; ; \;\; VV^\dagger = \mathbb{I}$$

To find the density matrix that yields a maximum value of entropy $\mathcal{S}$, we maximize $\mathcal{S}$ with respect to all the p_i's, with the constraint that $\sum_{i=1}^{N} p_i = 1$; using Lagrange multiplier λ yields the maximization problem:

$$L = \mathcal{S} + \lambda \left[\sum_{i=1}^{N} p_i - 1 \right]$$

$$0 = \frac{\partial L}{\partial p_I} = -k_\text{B}(\ln p_I + 1) + \lambda \;\; \Rightarrow \;\; p_I = \text{constant}$$

$$0 = \frac{\partial L}{\partial \lambda} = \sum_{i=1}^{N} p_i - 1 \;\; \Rightarrow \;\; p_I = \frac{1}{N}$$

The result above shows that maximum entropy state is one for which all the states are equally likely. The fact that all states are equally likely is precisely what one expects for a system about which one is totally ignorant.

The density matrix is proportional to the identity operator $\mathbb{I}$ since $\sum_{i=1}^{N} |\psi_i\rangle \langle\psi_i| = \mathbb{I}$; hence, for an N-state maximally uncertain system

$$\rho_\text{max} = \frac{1}{N}\mathbb{I} \;\; \Rightarrow \;\; \text{tr}(\rho_\text{max}) = 1$$

$$\mathcal{S}_\text{max} = -\text{tr}(\rho \ln \rho) = \frac{1}{N} \ln(N) \text{tr}(\mathbb{I})$$

$$\Rightarrow \mathcal{S} = \ln(N) \;\; : \;\; \text{maximum entropy} \tag{6.35}$$

Maximally Entangled States

For a state vector with two degrees of freedom consider, from (6.16), the following entangled state in the Schmidt representation:

$$|\Psi_{\rm E}\rangle = \sum_{i=1}^{N} c_i |\psi_i^I\rangle |\psi_i^{II}\rangle \;\; ; \;\; \sum_{i=1}^{N} |c_i|^2 = 1$$
$$\langle \psi_i^I | \psi_j^I \rangle = \delta_{i-j} = \langle \psi_i^{II} | \psi_j^{II} \rangle$$

that yields, from (6.36), the reduced density matrix for the entangled state as follows:

$$\rho_{\rm ER} = {\rm tr}_2 \Big(|\Psi_{\rm E}\rangle\langle\Psi_{\rm E}| \Big) = \sum_{i=1}^{N} |c_i|^2 |\psi_i^I\rangle\langle\psi_i^I|$$

The maximally entangled state has the maximum entropy and hence yields

$$\rho_{\rm ER}\Big|_{\rm Maximal} = \left[\sum_{i=1}^{N} |c_i|^2 |\psi_i^I\rangle\langle\psi_i^I| \right]_{\rm Maximal} = \frac{1}{N}\mathbb{I} \tag{6.36}$$

since the completeness of the eigenfunctions of a Hermitian operator gives a resolution of the identity operator.[6] Hence one obtains

$$|c_i|^2 = \frac{1}{N} \;\; \Rightarrow \;\; c_i = \frac{1}{\sqrt{N}} {\rm e}^{{\rm i}\phi_i}$$

and yields the *maximally entangled state* given by[7]

$$|\Psi_{\rm E}\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} {\rm e}^{{\rm i}\phi_i} |\psi_i^I\rangle |\psi_i^{II}\rangle \tag{6.37}$$

An example of a pair of maximally entangled spins is given by the following density matrix of a nonseparable system:

$$\rho_{\rm NS} = \frac{1}{4}\mathbb{I}\otimes\mathbb{I} \;\; ; \;\; {\rm tr}(\rho_{\rm NS}) = 1 \;\; ; \;\; \mathbb{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The reduced density matrix shows that the pair of spins is entangled since

$$\rho_{\rm NS,R} = {\rm tr}_2(\rho_{\rm NS}) = \frac{1}{2}\mathbb{I} \;\; ; \;\; {\rm tr}(\rho_{\rm NS,R}^2) = \frac{1}{2} \;\; : \;\; \text{Maximally entangled}$$

[6] Namely, $\sum_{i=1}^{N} |\psi_i^I\rangle\langle\psi_i^I| = \mathbb{I}$.

[7] The maximally entangled state is the same whether the partial trace is performed over quantum system I or system II.

An Entangled State of Two Spins

An entangled state for two spin degrees of freedom, from (6.30), is given by

$$|\Psi_{\rm E}\rangle = a|u_1\rangle|d_2\rangle + b|d_1\rangle\ |u_2\rangle\ \ ;\ \ |a|^2 + |b|^2 = 1 \tag{6.38}$$

with the reduced density matrix, from (6.32), given by

$$\begin{aligned}\rho_{\rm ER} &= {\rm tr}_2(\rho_{\rm E}) \\ &= |a|^2|u_1\rangle\langle u_1| + |b|^2|d_1\rangle\langle d_1| \\ &= \begin{bmatrix} |a|^2 & 0 \\ 0 & |b|^2 \end{bmatrix} \\ &\Rightarrow\ p_1 = |a|^2\ ;\ p_2 = |b|^2\end{aligned} \tag{6.39}$$

Hence, from (6.34), the entropy of this state is given by

$$\mathcal{S} = -{\rm tr}(\rho_{\rm ER}\ln\rho_{\rm ER}) = -p_1\ln p_1 - p_2\ln p_2 \tag{6.40}$$

$$= -|a|^2\ln(|a|^2) - |b|^2\ln(|b|^2) \tag{6.41}$$

For the following special case, and from (6.35),

$$\begin{aligned}|a| &= \frac{1}{\sqrt{2}} = |b| \\ \Rightarrow \mathcal{S} &= \ln(2) : \text{Maximum entropy}\end{aligned} \tag{6.42}$$

Hence, a *maximally entangled* state of two spins is given by

$$|\Psi_{\rm E}\rangle = \frac{1}{\sqrt{2}}\left[{\rm e}^{{\rm i}\phi}|u_1\rangle|d_2\rangle + |d_1\rangle\ |u_2\rangle\right] \tag{6.43}$$

6.11 Pure and Mixed Density Matrix

The density matrix, introduced in Sect. 6.4, is a Hermitian operator closely related to the state vector; recall from (6.45) that the pure density matrix for a state vector $|\chi\rangle$ is defined by

$$\rho_{\rm P} = |\chi\rangle\langle\chi| \tag{6.44}$$

From (5.33), the measurement of the expectation value of observable $\mathcal{O}$ can be expressed in terms of the density matrix of a pure state $\rho_{\rm P}$ as follows:

$$E_\chi[\mathcal{O}] \equiv \langle\chi|\mathcal{O}|\chi\rangle = {\rm tr}(\mathcal{O}|\chi\rangle\langle\chi|) = {\rm tr}(\mathcal{O}\rho_{\rm P}) \tag{6.45}$$

Equation (5.35) for the expectation value of an operator $\mathcal{O}$ with eigenvectors $\mathcal{O}|\psi_i\rangle = \lambda_i|\psi_i\rangle$ can be rewritten in terms of the mixed density matrix $\rho_{\rm M}$ as follows:

$$E_\psi[\mathcal{O}] = \langle\psi|O|\psi\rangle = \mathrm{tr}(\sum_i |c_i|^2 \mathcal{O}|\psi_i\rangle\langle\psi_i|) = \mathrm{tr}\,(\mathcal{O}\rho_{\rm M})$$

$$\Rightarrow \rho_{\rm M} = \sum_i p_i|\psi_i\rangle\langle\psi_i| \;\;;\;\; p_i = |c_i|^2$$

The mixed density matrix $\rho_{\rm M}$ can be used for evaluating the expectation value of any function of the operator $\mathcal{O}$. However, if one uses $\rho_{\rm M}$ for evaluating the expectation value of *another* operator $\mathcal{Q}$ that does not commute with $\mathcal{O}$, namely, $[\mathcal{O},\mathcal{Q}] \neq 0$, then there are *unavoidable* errors. The magnitude of these errors is set by the Heisenberg Uncertainty Principle and is discussed in Sect. 9.10.

Consider a quantum mechanical system to be in thermal equilibrium with a heat bath at temperature T. The system now has a quantum mechanical indeterminacy as well as classical uncertainty due to thermal randomness. The behavior of the quantum system is described by the canonical ensemble's probability distribution of energy eigenstates—given by the Boltzmann distribution.

Let H be the quantum mechanical Hamiltonian with the following spectral decomposition in terms of the energy eigenfunctions $|\psi_i\rangle$ and eigenvalues E_i

$$H = \sum_i E_i|\psi_i\rangle\langle\psi_i|$$

A quantum system with *thermal uncertainty* is described by the density matrix $\rho_{\rm T}$ given by

$$\rho_{\rm T} = \frac{1}{Z}\mathrm{e}^{-H/k_{\rm B}T} = \frac{1}{Z}\sum_i \mathrm{e}^{-E_i/k_{\rm B}T}|\psi_i\rangle\langle\psi_i| \;\;;\;\; Z = \mathrm{tr}\mathrm{e}^{-H/k_{\rm B}T} \tag{6.46}$$

$$\Rightarrow \rho_{\rm T} = \sum_i p_i|\psi_i\rangle\langle\psi_i| \;\;;\;\; \mathrm{tr}(\rho_{\rm T}) = \sum_i p_i = 1 \;\;;\;\; p_i = \frac{1}{Z}\mathrm{e}^{-E_i/k_{\rm B}T}$$

where $k_{\rm B}$ is the Boltzmann constant. The thermal density matrix $\rho_{\rm T}$ for the canonical ensemble is a mixed state since

$$\mathrm{tr}(\rho_{\rm T}^2) = \sum_i p_i^2 \;< 1$$

The reason that $\rho_{\rm T}$ is a mixed state is because thermal randomness leads to a *classical uncertainty* in the state of the system; this in turn entails that all the quantum state vectors $|\psi_i\rangle$ must be decoherent since there are no quantum correlations between the different quantum states—unlike the case for a pure density matrix $\rho_{\rm P}$ that has off-diagonal terms as given in (6.8).

The expectation value of an operator $\mathcal{O}$, for which $[\mathcal{O},H] \neq 0$, and that is in equilibrium with a heat bath is given by

$$E_T[\mathcal{O}] = \mathrm{tr}(\mathcal{O}\rho_\mathrm{T}) = \sum_i p_i \langle\psi_i|\mathcal{O}|\psi_i\rangle$$
$$= \sum_i p_i \alpha_i \;\; ; \;\; \alpha_i = \langle\psi_i|\mathcal{O}|\psi_i\rangle$$

The thermal density matrix ρ_T encodes both thermal and quantum uncertainty, reflected in probability p_i that the quantum system is in eigenstate E_i and the expectation value α_i of the operator in this eigenstate.

6.12 Summary

The tensor and outer product of states and operators were discussed to prepare the mathematical tools for discussing various properties of the tensor product states and the density matrix.

The density matrix provides an alternative formulation of a quantum entity and is equivalent to the description of the quantum entity by a state vector. Moreover, the density matrix has many advantages over the state vector in describing the properties of tensor product states of a system having many degrees of freedom.

Entangled states are purely *kinematical* in that these are quantum entities with state vectors that exist in Hilbert space that have remarkable nonclassical properties. Nowhere in the discussion on entangled states was the role of the Hamiltonian required.

Of course, if one would like to create an entangled state from a state vector that is initially a product (non-entangled) state, then one needs to subject it to interactions—and for which a Hamiltonian is required; an example of such case is discussed in Chap. 10 on the Stern-Gerlach experiment.

For a pure density matrix, the reduced density matrix being mixed provides a criterion for deciding whether the (two or more) degrees of freedom of a state vector are entangled. The condition that the reduced density matrix be mixed works only for pure states of the composite system given by state vector $|\psi\rangle$. If the whole system is in a pure state, then indeed one could conclude from the condition $\mathrm{tr}(\rho_\mathrm{R}^2) < 1$ for the reduced density matrix that the composite system is indeed entangled.

On the other hand, the general state of a bipartite composite separable system is given by $\rho_{AB} = \sum_i p_i \rho_i^A \otimes \rho_i^B$. When the composite system is mixed, the reduced density matrices are also mixed and one cannot conclude anything about the entanglement from the purity, or otherwise, of the subsystem's reduced density matrix $\mathrm{tr}(\rho_{A,R}^2)$ or $\mathrm{tr}(\rho_{B,R}^2)$. Entanglement for a bipartite system is more complicated and cannot be read from the properties of the reduced density matrix; a "reduction criterion" for deciding whether the composite system is entangled was discussed for this case.

Quantum entropy provides a measure for the degree of information that a density matrix holds. A maximally entangled state was shown to have maximum entropy.

Quantum Indeterminacy

7

Quantum mechanics superseded classical physics due to its empirical success and leads to a conception of Nature based on quantum indeterminacy and uncertainty. The paradigm of quantum mechanics is that Hermitian operators $\mathcal{B}$ extract information from the quantum state vector $\psi(\mathcal{F})$ that describes the indeterminate and trans-empirical degree of freedom $\mathcal{F}$; the particular and specific values of the degrees of freedom are, in principle, experimentally inaccessible and can never be directly observed. Every specific experiment measures a particular possible value of the operator $\mathcal{B}$; repeated measurements yield its average value $E_\psi[\mathcal{B}]$. Figure 7.1 illustrates the paradigm of quantum mechanics.

Many of the paradoxes of quantum mechanics result from the fact that the quantum degree of freedom is *indeterminate*. One may wonder how is quantum indeterminacy different from classical randomness, since there are many classical systems that are random but don't have the paradoxes of quantum mechanics. For example, it is well known since the nineteenth century that statistical mechanics, based as it is on classical mechanics and *classical probability*, discussed in Sect. 7.3, works very well for large thermodynamical systems.

In essence, one needs to address the question as to whether classical physics, together with classical probability, can explain the results of all experiments and thus negate the need for quantum mechanics.

To compare classical randomness with quantum indeterminacy, one needs to carefully define what is the difference of these two concepts [35]. It would seem that such a subtle question as to which of these two is realized in Nature might be difficult to define theoretically and to resolve experimentally. Ironically enough, the difference between classical randomness and quantum indeterminacy was brought into a sharp focus by a landmark paper of Einstein, Podolsky, and Rosenfeld (EPR) [11] in which it was claimed that quantum mechanics is incomplete; an example was presented that seemed to contradict the indeterminateness that is at the foundation of quantum mechanics.

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8_7, © Springer Science+Business Media New York 2013

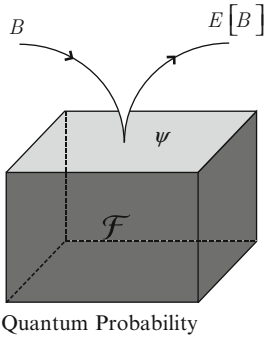


Fig. 7.1 The paradigm of quantum mechanics (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

7.1 The EPR Paradox

The EPR paradox is based on the analysis of *two* back-to-back photons—in a net spin zero state—created by the annihilation of an electron-positron pair. For simplicity, following Bohm, consider two identical spin 1/2 particles, say, two electrons e^-, e^-, in a net spin zero state, emanating from the decay of a spin zero particle, as shown in Fig. 7.2.

The EPR state vector factorizes into a space-dependent and a spin-dependent component; one needs to only analyze the spin-dependent component, denoted by $|\psi\rangle_{\rm EPR}$. Let the z-component of the spin pointing up be represented by $|u\rangle$ and pointing down be represented by $|d\rangle$; the net spin zero state, indicated in Fig. 7.2, is given by[1]

$$|\psi\rangle_{\rm EPR} = \frac{1}{\sqrt{2}}\Big[|u_1\rangle \otimes |d_2\rangle - |d_1\rangle \otimes |u_2\rangle\Big] \tag{7.1}$$

and which follows from the conservation of angular momentum.

Consider the two particles being well separated, for example, one of them being on Earth and the other being on the Moon.

The z-component of the spin of the two particles given in $|\psi\rangle_{\rm EPR}$ can be measured independently by two separate devices, namely, device 1 and device 2. From the state vector given in (7.1), it follows that if detector 1 measures spin up, namely, $|u_1\rangle$, then one can *predict* that the other spin in detector 2 is down, namely, $|d_2\rangle$—*without performing a measurement*. This feature of the state vector $|\psi\rangle_{\rm EPR}$ led EPR to conclude that the z-component of spin 2 has an *objectively reality* independent of any measurement.

> If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists **an element of reality** corresponding to that quantity [11].

[1] $|\psi\rangle_{\rm EPR}$ is an example of a *maximally entangled* state vector, as discussed in (6.43).

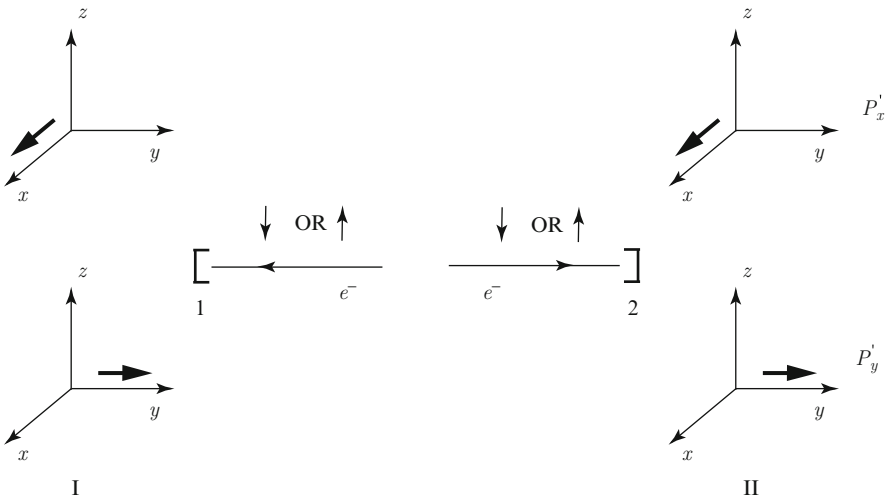


Fig. 7.2 Experimental arrangement for the EPR paradox. There are two detectors, labeled 1 and 2 on end of each arm. Each arm of the device has only one electron, and the figure is drawn to indicate that the electron can point either up or down in either arm (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

The EPR paper has two paradoxes:

- The possibility that *all components* of a spin can be known precisely, thus contradicting the Heisenberg Uncertainty Principle.
- The nonlocal collapse of the state vector.

All Spin Components Can Be Measured?

The EPR paradox seemed to provide an example that contradicted the quantum view that the quantum degree of freedom is inherently indeterminate; the state of spin 2 seemed to be in a determinate state since its value can be predicted without the need to make a measurement.

The pair of spins is in a net zero angular momentum state, and hence, the quantum state is spherically symmetric. Consider the two spins having a large space-like separation when they are observed. Suppose the z-component of the spin 1 is observed; then the z-component of spin 2 is fixed due to angular momentum conservation.

According to the EPR reasoning, the first observer could have chosen to measure the x- or y-component of spin 1 instead of the z-component, which would have meant that the x- or y-component of spin 2 was fixed. Since space-like separated events should not influence each other, the freedom of experimenter 1 to choose

to measure either z or x or y spin components implies that *all* the components of spin 2 should have been fixed beforehand, leading to a violation of the Heisenberg Uncertainty Principle.

From this analysis, EPR concluded (incorrectly) that quantum theory is incomplete as it does not allow for any state vector to have fixed values for all the components of spin; they then concluded that quantum theory needs to be completed by a more comprehensive theory that, in particular, would allow fixing all the components of spin.

Bohr's Response

In response to the EPR paradox, Bohr, in his typically opaque and elliptic manner, pointed out that the EPR view of the state of spin 2 being determinate is an illusion and reflects what one chooses to measure. If, instead of measuring the z-component of spin 2, detector 2 measures the x- or y-component of spin 2, as shown in Fig. 7.2, then quantum indeterminateness will be seen to exist for these components of the spin degree of freedom. Furthermore, Bohr pointed out that the second observer cannot set up any experiment that can actually measure *all* the components of spin 2 [15].

Classical Versus Quantum Correlations: Special Relativity

The EPR highlighted the correlations of the state vector across space-like distances that apparently seemed to lead to a contradiction between the nonlocal and instantaneous collapse of the state vector and special relativity.

If one measures the z-component of spin 1 as being up, then one can conclude that the z-component of spin 2 is down even without performing a measurement. This correlation in of itself is not surprising since even in classical physics, if two projectiles fly away from an explosion that conserves momentum, then the momentum of one projectile is exactly opposite that of the other. What special relativity demands is that, once the projectiles have a space-like separation, any change made on the momentum of one projectile cannot affect the momentum of the other projectile.

In the case of quantum mechanics, the paradox lies in the fact that *until a measurement* is made, the z-components of both the spins are indeterminate, as shown in Fig. 7.3. It is only *after* a measurement is made on the z-component of spin 1 that the corresponding value of spin 2 is fixed by the nonlocal and instantaneous collapse of the state vector. The nonlocal EPR correlations are also called quantum correlations.

The question arises, how is the "information" that a measurement has been performed on spin 1 communicated instantaneously to spin 2? The conventional answer is that the one has no control on the outcome of the measurement performed on spin

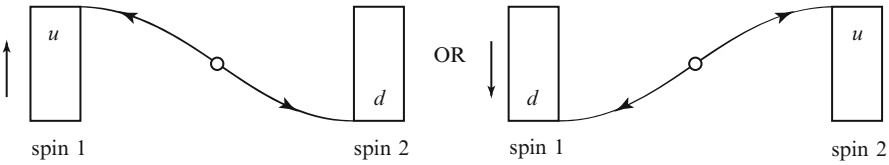


Fig. 7.3 EPR correlation (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

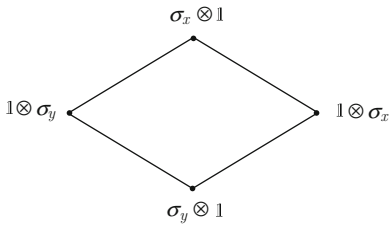


Fig. 7.4 Four operator for Bell's theorem. Each operator commutes with its nearest neighbor (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

1 and consequently one does not known in advance what will be the corresponding value for spin 2. Hence, no information is transferred in the instantaneous collapse of the state vector.

It is only after the result of the two spin measurements are physically sent to a common point, and which can be done only at a speed less or equal to the speed of light, that can one see the exact correlation of the result of measurements made on the two spins.

7.2 The Bell-CHSH Operator

In a seminal paper published in 1964, John Bell proposed—in response to the EPR paradox—a precise experiment to experimentally differentiate quantum indeterminacy from classical randomness; based on his insight, an experiment was designed to decide whether Nature is best described by classical or quantum indeterminacy. Experiments showed that quantum indeterminacy is the appropriate description of Nature.

The result of Bell, however, goes far beyond the EPR paradox and leads to a criterion for demarcating quantum indeterminacy from classical probability. Due to the generality of the Bell's result, it is now referred to as the Bell theorem.

Since the electrons are *well separated* so that no light signal can connect them, the operators for the two electrons commute. The state vector for the two electrons is taken to be an entangled state, and hence, their spins are quantum correlated. Suppose the x- and y-components of the spin of each of the electron are measured, as shown in Fig. 7.2; the operators for measuring the spins, shown in Fig. 7.4, are given by the following:

$$\text{Electron 1}: \quad \sigma_x \otimes \mathbb{I}; \quad \sigma_y \otimes \mathbb{I}$$
$$\text{Electron 2}: \quad \mathbb{I} \otimes \sigma_x; \quad \mathbb{I} \otimes \sigma_y$$

The Pauli σ-matrices are given in (6.13), with the notation[2]

$$\sigma_x = \sigma_1, \qquad \sigma_y = \sigma_2, \qquad \sigma_z = \sigma_3 \tag{7.2}$$

The spin operators $\sigma_i \otimes \mathbb{I}$ for the first electron *commute* with the spin operators $\mathbb{I} \otimes \sigma_i$ of the second electron; the four operators are shown in Fig. 7.4.

Define the operators (2×2 matrices) a, b, a', b' by the following:

$$a = \sigma_x; \quad b = \sigma_y; \quad a' = \sigma_x; \quad b' = \sigma_y \tag{7.3}$$
$$\Rightarrow a^2 = 1 = b^2; \quad (a')^2 = 1 = (b')^2 \tag{7.4}$$

Define the (Hermitian) Bell-CHSH operator [9] :

$$\mathcal{B} = a \otimes (a' + b') + b \otimes (b' - a') \tag{7.5}$$

From (7.3),

$$\mathcal{B} = \sigma_x \otimes (\sigma_x + \sigma_y) + \sigma_y \otimes (\sigma_y - \sigma_x) \tag{7.6}$$

An experimental device measures the expectation value of the (Hermitian) operator $\mathcal{B}$. Let the quantum state of the electrons be described by the pure density matrix ρ given by

$$\rho = |\psi\rangle\langle\psi|; \qquad \text{tr}(\rho) = 1$$

R_{q} is equal to the absolute value of the *quantum expectation value* of $\mathcal{B}$ and is given by the following:

$$R_{\text{q}} = \left|E_{\text{q}}[\mathcal{B}]\right| = \left|\text{tr}\{\rho\mathcal{B}\}\right| \tag{7.7}$$

Note that since

$$\text{tr}\{\rho\left(\mathcal{B} - R_{\text{q}}\mathbb{I}\right)^2\} \geq 0$$

we obtain the important identity

$$R_{\text{q}}^2 = \text{tr}^2(\rho\mathcal{B}) \leq \text{tr}(\rho\mathcal{B}^2) \tag{7.8}$$

[2]Note $[\sigma_i, \sigma_j] = 2i\sum_{k=1}^{3} \epsilon_{ijk}\sigma_k$ and ϵ_{ijk} is the completely antisymmetric tensor.

The definition of $\mathcal{B}$ from (7.10) yields, using $[\sigma_x, \sigma_y] = 2i\sigma_z$, the following diagonal representation:

$$\begin{aligned} \mathcal{B}^2 &= 4\mathbb{I} + [a,b] \otimes [a',b'] \qquad (7.9) \\ &= 4\mathbb{I} - 4\sigma_z \otimes \sigma_z = 4 - 4 \cdot \mathrm{diag}(1,-1,-1,1) \\ &\Rightarrow \mathcal{B}^2 = \mathrm{diag}(0,8,8,0) \qquad (7.10) \end{aligned}$$

Note the crucial commutator term $[a,b] \otimes [a',b']$ in (7.9) that results from the operator structure of quantum mechanics. It is this term that leads to a violation of the Bell inequality (to be derived later in Sect. 7.4), an inequality that holds for all classical probabilistic systems.

For any quantum state, the largest expectation value of $\mathcal{B}^2$—in fact of any Hermitian operator—is bounded by it's largest eigenvalue as in (5.9); hence, (7.10) implies, together with (7.8), that

$$\begin{aligned} R_{\mathrm{q}}^2 &= \mathrm{tr}^2(\rho\mathcal{B}) \leq \mathrm{tr}(\rho\mathcal{B}^2) \leq 8 \\ &\Rightarrow R_{\mathrm{q}} \leq 2\sqrt{2} = 2.82842712\ldots: \quad \text{Quantum inequality} \qquad (7.11) \end{aligned}$$

7.3 Classical Probability: Objective Reality

A precise definition needs to be given of classical probability theory so that we can address the question of whether the readings of an experimental device measuring a given physical system can be explained using the framework of classical probability.

In statistical mechanics, a large collection of gas molecules are described by assuming the position and velocity of each molecule is a classical random variable. Each molecule *objectively exists* in some state, and the *uncertainty in the knowledge* of the state of the molecule is attributed to our *ignorance* of the microscopic state of a very large collection of molecules.

Classical probability is based on the concept of a random variable, which takes a range of values, and that exists objectively regardless of whether it is measured (sampled) or not. A unique probability, called the joint probability distribution, is assigned to a collection of random variables and predicts how frequently will a collection of specific values appear when the random variables are sampled.

Following Kolomogorov, classical probability theory is defined by the following postulates:

- A collection of all possible allowed random sample values labeled by ω, which forms a *sample space* Ω
- A *joint probability distribution* function $P(\omega)$ that determines the probability for the simultaneous occurrence for these random events and provides an exhaustive and complete description of the random system

The events can be enumerated by *random variables*, say, $\mathbf{X} = (X, Y, Z, \ldots)$, that map the random events ω of the sample space Ω to real numbers (Fig. 7.5), namely,

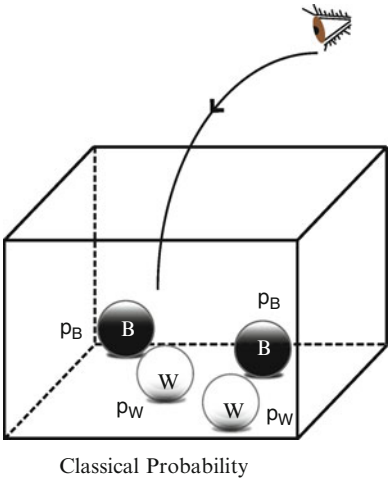


Fig. 7.5 Classical random events: the number of black and white balls inside a closed box *exist objectively*, independent of being observed or not. Furthermore, there exists a probability $p_{\mathrm{B}}, p_{\mathrm{W}}$ that is *intrinsic* to each possible outcome and is assigned to *each* black and white ball (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

$$\begin{aligned}
&\mathbf{X} : \Omega \ \rightarrow \ \Re^N \\
&X, Y, Z : \omega \rightarrow \Re \otimes \Re \otimes \Re; \quad \omega \in \Omega \\
&P(X, Y, Z) : \ \text{joint probability distribution}
\end{aligned}$$

Every element of the sample space Ω is assigned a likelihood of occurrence that is given by the joint probability distribution function $P(\omega)$; for the mapping of ω by random variables X, Y, Z to the real numbers, the joint probability distribution function is $P(X, Y, Z)$.

The assignment of a likelihood of occurrence $P(\omega)$ to *each element* of the sample space, namely, to each $\omega \in \Omega$, is the *defining property* of classical probability theory; this assignment implicitly assumes that *each element* ω of Ω *exists objectively*—regardless of being observed or not—and an experiment finds it in its *preexisting state* with probability specified by the probability distribution. It is precisely on this point that quantum probability will be seen to differ from classical probability.

Joint, Marginal, and Conditional Probabilities

The joint probability distribution function obeys all the laws of classical probability. Consider random variables X, Y, Z. Their joint probability distribution is given by

$$1 \geq P(X, Y, Z) \geq 0; \quad \int_{-\infty}^{\infty} \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z P(x, y, z) = 1$$

In other words, $P(X=x,Y=y,Z=z)$ yields the probability for the *simultaneous occurrence* of the sample values x,y, *and* z of the random variables X,Y,Z. Consider a function H that depends on the random variables X,Y,Z; its (average) classical expectation value is given by

$$E_{\rm c}[H]=\int {\rm d}x\,{\rm d}y\,{\rm d}z H(x,y,z)P(x,y,z)$$

If the random variables are independent, the joint probability distribution function factorizes and yields

$$P(x,y,z)=P_1(x)P_2(y)P_3(z)$$

For many random variables, one can form various *marginal* and *conditional probability distributions*. The probability that random variables are observed having random values X,Y, regardless of the value of Z, is given by the marginal distribution for two random variables, namely,

$$P(X,Y)=\int_{-\infty}^{\infty}{\rm d}z P(X,Y,z);\qquad \int_{-\infty}^{\infty}{\rm d}x\,{\rm d}y P(x,y)=1 \tag{7.12}$$

The *conditional probability* for events A,B is defined as follows. Let $P(A,B)$ be the joint probability distribution that events A and B both occur. The conditional probability $P(A|B)$ that A occurs, given that B has definitely occurred, is given by conditional probability

$$P(A|B)=\frac{P(A,B)}{P(B)}\ \ \Rightarrow\ \ P(A|B)P(B)=P(B|A)P(A)$$

For the case of a classical random particle such as a gas molecule in a room, the probability of finding the classical particle at point x,y, given that it has been definitely observed at z, is given by the *conditional probability*

$$P(X,Y|Z)=\frac{P(X,Y,Z)}{P(Z)}=\frac{P(X,Y,Z)}{\int_{-\infty}^{\infty}{\rm d}x\,{\rm d}y P(x,y,Z)};\quad \int_{-\infty}^{\infty}{\rm d}x\,{\rm d}y P(x,y|Z)=1$$

7.4 The Bell Inequality

Consider the Bell-CHSH operator that is being measured, namely,

$$\mathcal{B}=a\otimes(a'+b')+b\otimes(b'-a')$$

From (7.4), $a^2=b^2=(a')^2=(b')^2=1$; on being experimentally observed, the quantities a,b,a',b' take two possible values, namely, ±1. Since the values of

a, b, a', b' are indeterminate and random, one can consider explaining the result of the experiment in terms of *classical random variables* f, g, f', g' that take random values of ± 1, namely,[3]

$$a \to f; \quad b \to g; \quad a' \to f'; \quad b' \to g'$$
$$f^2 = g^2 = (f')^2 = (g')^2 = 1$$

The random variables are discrete, taking only two values and hence yield a sample space $\Omega = 2^4$: a four dimensional lattice of 16 discrete points.

The probability of the random variables f, g, f', g' simultaneously taking different values is given by the joint probability distribution $P(f, g, f', g')$.

In terms of the classical random variables, the Bell-CHSH operator that is being observed is modeled by a classical random function H. The binary valued operators a, b, a', b' are replaced by binary valued classical random variables f, g, f', g'; the tensor product $\otimes$ is replaced by ordinary multiplication since classical random variables do not have any linear structure that a quantum operator is endowed with by the underlying Hilbert space.

Hence, the representation of the Bell-CHSH operator based on classical random variables is as follows:

$$\mathcal{B} \to H = f(f' + g') + g(g' - f') \tag{7.13}$$

The expectation value of H is given by

$$E_c[H] = \sum_{f, f', g, g' = \pm 1} H(f, g, f', g') P(f, g, f', g')$$

Since all the random variables f, g, f', g' *commute*, similar to (7.9), one has the following:

$$H^2 = 4 \tag{7.14}$$

Note that $H^2 = 4$ differs from Q^2 given in (7.10) because H is a scalar, not having any operator structure. Similar to (7.8), using (7.14) yields the Bell inequality

$$\begin{aligned} R_c^2 &= |E_c[H]|^2 \leq E(H^2) = 4 \\ &\Rightarrow R_c = |E_c[H]| \leq 2 : \quad \text{Bell inequality} \end{aligned} \tag{7.15}$$

[3]The random variables f, g, f', g' take only two discrete values; the value of ± 1 is not fundamental but convenient; the proof goes through for the more general case of the random variables f, g, f', g' being uniformly distributed on [0,1].

In summary, there are two inequalities: one for the expectation value of $\mathcal{B}$, based as it is on quantum mechanics, with $R_q \leq 2\sqrt{2}$ given in (7.11), and another inequality based on classical probability given by $R_c \leq 2$ as in (7.15), namely, the Bell inequality.

An experiment, performed by Aspect in 1982, showed that the value of $E[\mathcal{B}]$ is just less than $2\sqrt{2}$, violating the limit set by the Bell inequality and demonstrating that quantum mechanics gives the correct description of Nature [2].

The validity of the inequality based on quantum mechanics implies that although the joint probability distribution $P(f, g, f', g')$ *does exist,* it, nevertheless, *cannot* correctly produce the experimental violation of the Bell inequality.

Furthermore, the result of Bell in fact shows that *no classical probability theory*, based as it is on a joint probability distribution, can produce the violation of Bell's inequality.

Noteworthy 7.1: Hidden variables

The EPR paper was written in the context of a debate whether quantum mechanics was a complete theory or had some elements missing that led to the apparent indeterminacy that is its hallmark. What EPR proposed is that there are variables not being accounted for in the current formulation of quantum mechanics, and for this reason called *hidden variables*.

It is postulated that there are hidden variables, and are classical dynamical variables, and are required for describing a quantum system. Since the hidden variables are not observed, it is postulated that the observed magnitude of a system's physical property is the average value over all possible values of the hidden variables, with the different specific values of the hidden variables being random and governed by a classical probability distribution. Every specific value of the hidden variable exists *objectively*, independent of any observation. It can be shown that hidden variables can explain the quantum behavior of spin 1/2 (degree of freedom with two values), but not for degrees of freedom with three or more distinct values [15].

It was Einstein's view that, on including hidden classical variables, quantum mechanics could be shown to be equivalent to a classical random system, with quantum indeterminacy being similar to the randomness that appears in statistical mechanics. Bell's analysis shows that this view of Einstein is, in fact, not valid.

7.5 The Bell Inequality Non-violation

The quantum inequality, given in (7.11), states that the expectation value of the Bell-CHSH quantum operator obeys the inequality

$$R_q = |\mathrm{tr}(\rho\mathcal{B})| \leq 2\sqrt{2}: \quad \text{Quantum inequality}$$

whereas for classical random variables, from (7.15), the expectation value has a smaller lower bound, namely,

$$R_c = |E_c[H]| \leq 2: \quad \text{Bell inequality}$$

It is shown in this section that non-entangled states and separable systems *do not violate* the classical bound; hence, as far as the Bell criterion is concerned, they are indistinguishable from a system described by classical random variables.

Non-entangled States

Recall, from (7.10), the Bell-CHSH operator is defined

$$\mathcal{B} = a \otimes (a' + b') + b \otimes (b' - a')$$

If the quantum system is not entangled, as discussed in Sect. 6.7, its state vector factorizes and is given by

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$$

and density matrix also factorizes and can be represented as follows:

$$\rho = |\psi\rangle\langle\psi| = \rho_1 \otimes \rho_2; \qquad \rho_1 = |\psi_1\rangle\langle\psi_1|; \qquad \rho_2 = |\psi_2\rangle\langle\psi_2|$$

Hence, the expectation value of the Bell-CHSH operator is given by

$$\begin{aligned} \text{tr}(\rho\mathcal{B}) &= \text{tr}\Big((\rho_1 \otimes \rho_2)\mathcal{B}\Big) \\ &= \text{tr}(\rho_1 a) \cdot \text{tr}(\rho_2(a' + b')) + \text{tr}(\rho_1 b) \cdot \text{tr}(\rho_2(b' - a')) \end{aligned} \tag{7.16}$$

Recall from (7.4) that

$$\begin{aligned} & a^2 = 1 = b^2; \qquad (a')^2 = 1 = (b')^2 \\ \Rightarrow & -1 \leq \text{tr}(\rho_1 a),\ \text{tr}(\rho_1 b),\ \text{tr}(\rho_2 a'),\ \text{tr}(\rho_2 b') \leq +1 \end{aligned} \tag{7.17}$$

Hence, from (7.16) and (7.17),

$$\begin{aligned} R_q &\equiv \Big|\text{tr}\Big((\rho_1 \otimes \rho_2)\mathcal{B}\Big)\Big| \\ &\leq |\text{tr}(\rho_1 a) \cdot \text{tr}(\rho_2(a' + b'))| + |\text{tr}(\rho_1 b) \cdot \text{tr}(\rho_2(b' - a'))| \\ &\leq |\text{tr}(\rho_2(a' + b'))| + |\text{tr}(\rho_2(b' - a'))| \\ &\leq |s + s'| + |s - s'| \end{aligned} \tag{7.18}$$

where

$$s = \text{tr}(\rho_2 b'); \qquad s' = \text{tr}(\rho_2 a'); \qquad s, s' \in [-1, +1]$$

The right-hand side $|s+s'|+|s-s'|$ is bounded by 2; for example, if $s=s'=1$, the first term is 2 and second term is zero, and when $s=-s'=1$ the first term is zero and the second term is 2.

Hence, as given in (7.15),

$$R_{\mathrm{q}} \leq 2: \quad \text{non-entangled states obey the Bell inequality} \tag{7.19}$$

One can conclude that a non-entangled quantum system can be described by classical probability and in fact, using the Bell inequality as a criterion, cannot be distinguished from a classical random system.

Separable Systems

The proof that a separable system does not violate Bell's inequality is similar to the one given above for non-entangled states.

Consider a general separable state shared by two parties A and B, with states ρ_i^A and ρ_i^B are held by parties A and B, respectively. From (6.26), this state is given by the following bipartite (mixed) density matrix:

$$\rho_{AB} = \sum_{i=1}^{N} p_i \rho_i^A \otimes \rho_i^B; \quad \sum_{i=1}^{N} p_i = 1;\ p_i \in [0,1]$$

Recall, from (7.10), the Bell-CHSH operator is defined

$$\mathcal{B} = a \otimes (a'+b') + b \otimes (b'-a')$$

The average value of the Bell-CHSH operator $\mathcal{B}$ for a separable bipartite system is given by

$$\mathrm{tr}(\mathcal{B}\rho_{AB}) = \sum_{i=1}^{N} p_i \mathrm{tr}\big(\mathcal{B}(\rho_i^A \otimes \rho_i^B)\big)$$

The trace in the equation factorizes due to the tensor product of operators in $\mathcal{B}$ and, using $\left|\mathrm{tr}\big(\sigma_x \rho_i^B\big)\right| \leq 1$ and $\left|\mathrm{tr}\big(\sigma_y \rho_i^B\big)\right| \leq 1$, yields the following for a separable bipartite system:

$$\begin{aligned}
R_{\mathrm{q}}^S &= \Big|\mathrm{tr}(\mathcal{B}\rho_{AB})\Big| \\
&\leq \sum_{i=1}^{N} p_i \Big|\mathrm{tr}\big(\mathcal{B}(\rho_i^A \otimes \rho_i^B)\big)\Big| \\
&\leq \sum_{i=1}^{N} p_i \left\{ \left|\mathrm{tr}\big(\sigma_x \rho_i^B\big)\right| \left|\mathrm{tr}\big(\sigma_x \rho_i^B\big) + \mathrm{tr}\big(\sigma_y \rho_i^B\big)\right| + \left|\mathrm{tr}\big(\sigma_y \rho_i^B\big)\right| \left|\mathrm{tr}\big(\sigma_x \rho_i^B\big) - \mathrm{tr}\big(\sigma_y \rho_i^B\big)\right| \right\} \\
&\leq \sum_{i=1}^{N} p_i \left\{ \left|\mathrm{tr}\big(\sigma_x \rho_i^B\big) + \mathrm{tr}\big(\sigma_y \rho_i^B\big)\right| + \left|\mathrm{tr}\big(\sigma_x \rho_i^B\big) - \mathrm{tr}\big(\sigma_y \rho_i^B\big)\right| \right\}
\end{aligned}$$

Since $\sigma_x^2 = 1 = \sigma_y^2$, we have

$$s_i = \mathrm{tr}(\rho_i^B \sigma_y); \qquad s_i' = \mathrm{tr}(\rho_i^B \sigma_x); \qquad s_i, s_i' \in [-1, +1]$$

Hence,

$$R_{\mathrm{q}}^S \leq \sum_{i=1}^{N} p_i \left\{ \left| s_i' + s_i \right| + \left| s_i' - s_i \right| \right\}$$

Equation above is the generalization of the result obtained in (7.18) for a single quantum system.

Similar to the reason for obtaining result given in (7.19), the equation above yields

$$R_{\mathrm{q}}^S \leq 2: \quad \text{separable states obey the Bell inequality}$$

The proof that a separable bipartite system and a non-entangled do not violate Bell's inequality does not mean that these are equivalent to classical states. The only statement we can make for these systems is that the Bell inequality cannot differentiate them from a classical system. As we shall discuss later, in Sect. 7.7, the criterion of "contextuality" is more general and can be used to show that the behavior of no quantum state, except for the spin 1/2 case, can be explained by classical probability theory.

7.6 Bell Inequality Violation: Entangled States

The question naturally arises as to what are the quantum states that violate the classical bound encoded in the Bell inequality; namely, what are the quantum states for which

$$2 \leq R_{\mathrm{q}} = |\mathrm{tr}(\rho \mathcal{B})| \leq 2\sqrt{2}: \quad \text{Quantum regime}$$

Entangled states yield a *violation* of the Bell inequality and hence cannot be explained by classical probability theory. Instead of giving a general proof of the Bell inequality for entangled states, the expectation value of the Bell-CHSH operator is explicitly calculated for a specific form of an entangled state, namely, a pair of spins discussed in Sect. 6.9.

In particular, the value of the quantity R_{q} is derived in terms of the parameters of the entangled state. The extreme limit of $R_{\mathrm{q}} = 2\sqrt{2}$ will be shown to be the value of R_{q} for a maximally entangled state.

From (6.30), a class of entangled states for a system of two spin 1/2 degrees of freedom, in tensor product notation, is given by

$$|\psi\rangle = \alpha|u_1\rangle \otimes |d_2\rangle + \beta|d_1\rangle \otimes |u_2\rangle; \qquad |\alpha|^2 + |\beta|^2 = 1; \qquad \rho = |\psi\rangle\langle\psi| \quad (7.20)$$

The Bell-CHSH operator, from (7.6), has the following explicit representation:

$$\mathcal{B} = \sigma_x \otimes (\sigma_y + \sigma_x) + \sigma_y \otimes (\sigma_y - \sigma_x) \quad (7.21)$$

and the expectation value of the Bell-CHSH operator is given by

$$\mathrm{tr}(\rho\mathcal{B}) = \langle\psi|\Big[\sigma_x \otimes (\sigma_y + \sigma_x) + \sigma_y \otimes (\sigma_y - \sigma_x)\Big]|\psi\rangle \quad (7.22)$$

The representation for the σ matrices is given in (7.2), and let the basis states be given, from (4.1), as follows:

$$|u\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \qquad |d\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \qquad \langle u| = \begin{bmatrix} 1 & 0 \end{bmatrix}; \qquad \langle d| = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

The representation chosen yields the following nonzero terms:

$$\begin{aligned} \mathrm{tr}(\rho\mathcal{B}) &= \alpha\beta^*\langle d_1| \otimes \langle u_2|\Big[\sigma_x \otimes (\sigma_y + \sigma_x) + \sigma_y \otimes (\sigma_y - \sigma_x)\Big]|u_1\rangle \otimes |d_2\rangle \\ &\quad + \text{complex conjugate} \\ &= \alpha\beta^*\Big[\langle d_1|\sigma_x|u_1\rangle\Big\{\langle u_2|\sigma_y|d_2\rangle + \langle u_2|\sigma_x|d_2\rangle\Big\} \\ &\quad + \langle d|_1\sigma_y|u_1\rangle\Big\{\langle u_2|\sigma_y|d_2\rangle - \langle u_2|\sigma_x|d_2\rangle\Big\}\Big] + \text{c.c.} \\ &= \alpha\beta^*[1\cdot(i+1) - i\cdot(i-1)] + \text{c.c.} \\ &\Rightarrow \mathrm{tr}(\rho\mathcal{B}) = 2(\alpha\beta^* + \alpha^*\beta) + 2i(\alpha\beta^* - \alpha^*\beta) \end{aligned}$$

Choose the following parametrization[4]:

$$\alpha = \mathrm{e}^{\mathrm{i}\phi}\sin(\theta); \qquad \beta = \cos(\theta); \qquad \theta \in [0, \pi]; \qquad \phi \in [0, 2\pi]$$

that yields the state vector

$$|\psi\rangle = \mathrm{e}^{\mathrm{i}\phi}\sin(\theta)|u_1\rangle \otimes |d_2\rangle + \cos(\theta)|d_1\rangle \otimes |u_2\rangle \quad (7.23)$$

Hence, we obtain the following violation of the Bell inequality:

[4]Since state vectors are only defined up to a phase, this is the most general parametrization.

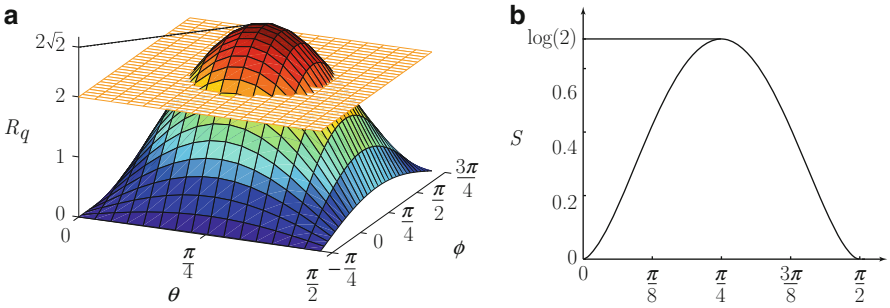


Fig. 7.6 (**a**) The values of R_q for different entangled states. (**b**) The quantum entropy $\mathcal{S}$ of the reduced density matrix for the entangled states (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

$$\begin{aligned}\mathrm{tr}(\rho\mathcal{B}) &= 4\sin(\theta)\cos(\theta)[\sin(\phi)+\cos(\phi)] = 2\sqrt{2}\sin(2\theta)\sin(\phi+\frac{\pi}{4})\\ \Rightarrow R_{\mathrm{q}} &= |\mathrm{tr}(\rho\mathcal{B})| = 2\sqrt{2}\left|\sin(2\theta)\sin(\phi+\frac{\pi}{4})\right| \in [0,2\sqrt{2}]\end{aligned}$$

Figure 7.6a shows the value of R_{q} for different entangled states labeled by θ, ϕ; there is an entire range of θ and ϕ for which R_{q} is greater than the classical limit of 2. The maximum value of R_{q} is given by the following parameters:

$$\theta = \phi = \frac{\pi}{4} \quad \Rightarrow R_{\mathrm{q}} = 2\sqrt{2}$$

The maximum value of $R_{\mathrm{q}} = 2\sqrt{2}$ corresponds, from (7.23), to the state vector

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left[\mathrm{e}^{\mathrm{i}\pi/4}|u\rangle \otimes |d\rangle + |d\rangle \otimes |u\rangle\right] \tag{7.24}$$

which is a *maximally entangled state*, as discussed in (6.43). It is noteworthy that the maximum value of $R_{\mathrm{q}} = 2\sqrt{2}$—furthest from the classical limit $R_{\mathrm{q}} = 2$—is given for a state that is also the most nonclassical, namely, a maximally entangled state.

The degree of entanglement, from (6.34), is given by quantum entropy of the *reduced* density matrix

$$\mathcal{S} = -\mathrm{tr}(\rho_{\mathrm{R}} \ln \rho_{\mathrm{R}})$$

For the state given in (7.20), the entropy $\mathcal{S}$, as shown in Fig. 7.6b, is given by (6.40) and yields the following:

$$\begin{aligned}\mathcal{S}(\theta) &= -2|\alpha|^2 \ln(|\alpha|) - 2|\beta|^2 \ln(|\beta|)\\ &= -2\sin^2\theta \ln(|\sin\theta|) - 2\cos^2\theta \ln(|\cos\theta|)\end{aligned} \tag{7.25}$$

As given in (6.42), the maximally entangled state has the maximum quantum entropy given by

$$\mathcal{S}_{\text{Max}} = S\left(\frac{\pi}{4}\right) = \ln(2)$$

It can be seen from Fig. 7.6a that only for values for θ, ϕ for which $R_{\text{q}} > 2$—near the maximally entangled state—is the Bell inequality violated. This does not mean that entangled states with $R_{\text{q}} \leq 2$ do not violate the Bell inequality.

In fact, *all* entangled states violate the Bell inequality. If the Bell-CHSH operator does not show a violation for an entangled state, then one needs to find another operator that is *appropriate* for the entangled state in question—and an analysis similar to the one given for the states given in (7.20) will show the expected violation. Furthermore, the maximal violation of the Bell inequality does not necessarily need maximally entangled states, with $R_{\text{q}} = 2\sqrt{2}$ being the value for some non-maximally entangled states as well.

7.7 The Bell–Kochen–Specker Inequality

A system is said to be contextual if the outcome of a measurement of one of the observables depends upon what other measurements are performed alongside with it. The outcomes for classical probability theory are predefined and non-contextual; all observations in deterministic classical physics are similarly non-contextual; classical properties are inherent to the entity and do not depend on the observations of other properties.

In contrast, the properties of a quantum entity are not intrinsic; the observed value of an operator for a given state vector depends on what other operators are being measured as well, namely, on the context of the measurement. Contextuality is a reflection of quantum indeterminateness since a quantum entity does not exist objectively in a determinate state but instead has many possible outcomes depending on how it is observed.

The Kochen–Specker (KS) theorem states that the empirical predictions of quantum mechanics cannot be produced by any non-contextual theory [21].

A maximal set of commuting observables define a context. Consider the case of three Hermitian operators A, B, and C such that $[A, B] = 0 = [A, C]$ but $[B, C] \neq 0$. The KS theorem states that the value of $E[A]$ depends on whether it is measured together with experiments performed to measure B or C or neither. In other words, the value of $E[A]$ depends on the *context* of its measurement, defined as performing—simultaneous with the experiment on A—*other* experiments that measure other operators that commute with A but not necessarily with each other.

The generalized Bell inequality, due to Kochen and Specker, namely, the Bell–Kochen–Specker (BKS) inequality, analyzes under what condition a *single* degree of freedom exhibits the inequivalence between explanations based on classical and quantum probabilities. There is no need to consider entangled states, which requires

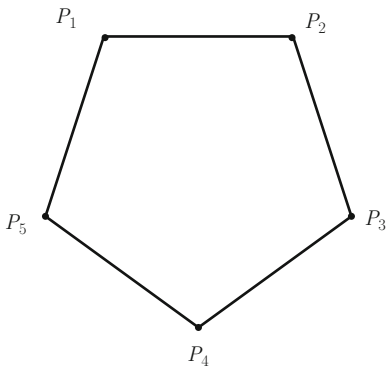


Fig. 7.7 Five-operator generalization of Bell's theorem. Each operator commutes with its nearest neighbor (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

at least two or more degrees of freedom, and neither is it necessary to have spins with space-like separation (so that the operators operating on the different spins commute). All that is needed for the BKS theorem is the existence of a certain collection of Hermitian operators.

In this section, the BKS inequality is applied to degrees of freedom with three and four distinct values, namely, a single spin 1 two spin 1/2 systems, respectively.[5]

Spin 1

The BKS inequality can be derived for a spin 1 system that is located at *a single point*; so the issue of space-like separation of the degrees of freedom is excluded. Furthermore, since there is only one degree of freedom, the quantum state cannot be an entangled state [20].

Consider the case of P_1, P_2, P_3, P_4, P_5, namely, five commuting and non-commuting operators that are arranged in Fig. 7.7. Let the operators be numbered periodically with $P_6 \equiv P_1$; then the commutation equations are given by the following:

$$[P_n, P_{n+1}] = 0; \qquad [P_n, P_{n+2}] \neq 0; \quad n = 1, 2, \ldots 5$$

and is a generalization of the diagram given for the four operator Bell inequality in Fig. 7.4.

It can be shown that the BKS inequality for this case is given by classical probability theory and yields

$$\sum_i E_c[P_i] \leq 2 : \quad \text{BKS inequality} \tag{7.26}$$

[5] A single spin 1/2 (with degree of freedom having only two distinct values) can be fully explained by classical hidden variables and hence has no contextuality [15].

The "contextual" inequality, obtained by evaluating the expectation value of P_i in a quantum state, is given by

$$\sum_i E_{\mathrm{q}}[P_i] \leq \sqrt{5} \tag{7.27}$$

and violates the BKS inequality given in (7.26).

State-Independent Violation of BKS Inequality

The pentagram inequality is state dependent since only a restricted set of states obey (7.27) and violate the inequality given in (7.26). On the other hand, there are state-independent inequalities that hold for all spin 1 quantum states; in fact, it can be shown that for a spin 1 degree of freedom, one requires *at least* 13 operators to show the violation of the BKS inequality for all spin 1 states [37].

Consider 13 operators Q_n, which are 3×3 matrices such that $Q_n^2 = 1$. Similar to Fig. 7.7, the 13 operators are placed on a planar orthogonality graph, but with a more complicated layout, with some operators having up to four neighbors. If one represents these operators by classical discrete random variables q_n taking only two values of ± 1 with $q_n^2 = 1$, then an analysis similar to the Bell-CHSH analysis can be carried out for this case.

The classical expectation value for the 13 random variables q_n, taken over all classical probability distribution functions, yields the following:

$$\sum_{n=1}^{13} E_{\mathrm{c}}[q_n] - \frac{1}{4} \sum_{m,n=1}^{13} \Gamma_{mn} E_{\mathrm{c}}[q_m q_n] \leq 8 : \quad \text{BKS inequality} \tag{7.28}$$

Γ_{mn} is the adjacency matrix, with $\Gamma_{mn} = 1$ if the random variables q_m and q_n are neighbors on the orthogonality graph, and zero otherwise.

A quantum expectation value taken for *all possible* spin 1 state vectors yields the following result:

$$\sum_{n=1}^{13} E_{\mathrm{q}}[Q_n] - \frac{1}{4} \sum_{m,n=1}^{13} \Gamma_{mn} E_{\mathrm{q}}[Q_m Q_n] \leq 8\frac{1}{3} \tag{7.29}$$

Comparing (7.28) and (7.29) shows that the quantum case violates the BKS inequality and cannot be explained by any non-contextual theory.

Recall spin 1/2 has no contextual property and, as mentioned earlier, can be explained by hidden variables. Spin 1 is of fundamental importance since it is irreducible to spin 1/2 and is a minimal three-valued degree of freedom; the fact that spin 1 exhibits contextuality leads to the expectation that all quantum degrees of freedom, except for the spin 1/2, are contextual and hence cannot be explained by classical probability theory.

Two Spin 1/2

It can be shown that one needs *at least* four operators for a two spin 1/2 system, such as the Bell-CHSH operator with commutation relations given in Fig. 7.7, to show that a classical probability distribution function cannot yield the correct quantum result for a two spin 1/2 system.

It is shown below that any collection of three operators is inadequate for showing the violation of the Bell inequality.

Consider the case of three Hermitian operators A, B, and C such that $[A,B] = 0 = [A,C]$ but $[B,C] \neq 0$ and constructed from the two spin 1/2 degrees of freedom. Since A can be simultaneously measured with other operators that commute with it, the joint probability distribution functions $p_1(A,B)$ and $p_2(A,C)$ can be measured, and which are theoretically also obtainable from quantum mechanics.

Although not within the framework of quantum mechanics, a classical joint probability distribution function does in fact exist for A,B,C considered as classical random variables and is given as follows:

$$p(A,B,C) = \frac{p_1(A,B)p_2(A,C)}{p(A)} \tag{7.30}$$

where

$$\sum_B p_1(A,B) = p(A) = \sum_C p_2(A,C)$$

This construction reproduces the experimentally measurable marginal probability distribution function. One recovers, for instance, the experimentally observed $p_1(A,B)$ by summing over the outcomes for C in $p(A,B,C)$ and which results in a cancellation of $p(A)$ on the right-hand side of (7.30), leading to the required probability $p_1(A,B)$.

A similar analysis can be done for the case of three Hermitian operators A, B, and C such that $[A,B] = 0$ but $[B,C] \neq 0$ and $[A,C] \neq 0$; it can be shown for this case that a classical $\tilde{p}(A,B,C)$ exists for this case as well. For the case of all three operators A,B, and C either all commuting with each other or none of them commuting, a classical probability distribution function exists for both cases as discussed in general in Sect. 7.8 .

Hence, since the expectation of three operators can be fully described by a classical probability distribution function, no operator constructed out of their combination can show a violation of the Bell inequality.

Recall in Sect. 7.6 a violation of the Bell inequality was shown for a specific entangled state using the Bell-CHSH operator. In general, it is not known what is the minimum collection of operators that can show a Bell violation for *all* entangled quantum states.

For a degree of freedom with four values, which includes the two spin 1/2 case, it is possible to show, using contextuality, that every state vector is essentially quantum and inequivalent to any classical system; the proof needs a set of 18 operators [22, 32].

Discussion

The tests to decide whether a system can be described by quantum or classical probability are more and more stringent.

The most basic and fundamental is contextuality; this criterion works for all quantum entities, and the operators constructed to show contextuality act on all the degrees of freedom. The second test is entanglement that requires two or more degrees of freedom. The third test is a violation of the Bell inequality and is the most special case since it requires that the quantum entity be in the form of a tensor product structure of two subsystems A and B—and with the Bell-CHSH type of operators having a similar tensor product structure.

For many applications in quantum information such as cryptography and teleportation, the Bell inequality criterion and entanglement are more useful because they have more structure.

The violation of the BKS inequality is essentially due to the indeterminacy of the degree of freedom that is at the foundation of a quantum entity. State vectors representing a quantum entity and physically observable quantities (of the quantum entity) being determined by the expectation values of Hermitian operators acting on the degree of freedom's Hilbert space of states is the schema of quantum mechanics. Classical probability is unable to explain or describe the physics of the quantum.

In conclusion, in the BKS scheme there always exists an appropriate collection of operators that can decide at the most basic level whether a system can be described by quantum or classical probability theory.[6]

7.8 Commuting and Non-commuting Operators

To prove the inequivalence of classical and quantum probability, one needs to have a collection of *both* commuting and non-commuting operators, as seen in Figs. 7.4 and 7.7. It is shown in this section that experiments that measure a collection of *only* commuting or *only* non-commuting operators are indistinguishable from a classical random system. Let the quantum system consists of N degrees of freedom x_i ; $i = 1, 2, \ldots, N$ with state vector given by $\psi(x_1, x_2, \ldots, x_N) = \langle x_1, x_2, \ldots, x_N | \psi \rangle$.

For the case of all commuting operators $\mathcal{O}_i$; $[\mathcal{O}_i, \mathcal{O}_j] = 0$; $i, j = 1, 2, \ldots, N$, all the operators can be measured simultaneously. For concreteness, let the operators be the coordinate operators represented by

$$\mathcal{O}_i = |x_i\rangle\langle x_i| \equiv \underbrace{\mathbb{I} \otimes \mathbb{I} \otimes \ldots |x\rangle\langle x| \ldots \otimes \mathbb{I}}_{i\text{th position}}$$

$$[\mathcal{O}_i, \mathcal{O}_j] = 0; \quad i, j = 1, 2, \ldots, N$$

[6] I thank Ravishankar Ramanathan for many useful discussions on the Bell inequality, entanglement, and contextuality.

and define the operator

$$\mathcal{O} = \mathcal{O}_1\mathcal{O}_2\ldots\mathcal{O}_N = |x_1\rangle\langle x_1| \otimes |x_2\rangle\langle x_2| \ldots \otimes |x_N\rangle\langle x_N|$$

The joint probability distribution is given by

$$P(x_1, x_2, \ldots, x_N) = \mathrm{tr}(\rho\mathcal{O}) = |\psi(x_1, x_2, \ldots, x_N)|^2; \quad \rho = |\psi\rangle\langle\psi|$$

Hence, if a quantum system is observed with only commuting operators, then it can be described by classical probability since there exists a joint probability distribution $P(x_1, x_2, \ldots, x_N)$ that yields all the expectation values and is provided by quantum mechanics itself. And conversely, the entire field of classical probability is seen to be equivalent to that subset of quantum probability that explains the behavior of commuting operators.

For the case of all non-commuting operators, consider, for example, observing non-commuting operators x and p; each has to be measured separately leading a probability distribution given by $P_1(x)$ and $P_2(p)$, respectively. A joint probability distribution, given by $P_1(x)P_2(p)$, provides the marginal distributions that yield the result for the measurement of either x or p. Similarly, consider a collection of non-commuting operators $\mathcal{J}_i$; $[\mathcal{J}_i, \mathcal{J}_j] \neq 0$; $i, j = 1, 2, \ldots, N$ such that no two operators can be observed simultaneously; a joint probability distribution that yields all the observed marginal distributions is given by

$$P(x_1, x_2, \ldots, x_N) = \prod_i P_i(x_i) = \prod_i |\psi_i(x_i)|^2$$

The joint probability distribution $\prod_i P_i(x_i)$ is a theoretical construct that is outside the framework of quantum mechanics and is based on classical probability.

From the above discussion, it can be seen that it is only an appropriate collection of *commuting and non-commuting* operators that can show a possibility of violating the Bell inequality. The result can be shown to follow from the BKS theorem [21]. In particular, for systems that violate the Bell inequality, it follows that there *does not exist* a joint probability distribution $P(x_1, x_2, \ldots, x_N)$ that can correctly yield all the expectation values and marginal probability distributions required by quantum indeterminacy.

7.9 Quantum Probability

The framework of quantum probability identifies Hermitian operators $\mathcal{O}(\mathcal{F})$ with observables—instead of with random variables as is the case for classical probability theory. Observables act on the state space $\mathcal{V}(\mathcal{F})$ of the quantum degree of freedom $\mathcal{F}$, and the expectation value of the observable $\langle\psi|\mathcal{O}(\mathcal{F})|\psi\rangle$ is determined by

the state vector $|\psi\rangle$ of the degree of freedom. In other words, it is the operators that "pick out" the properties of the degree of freedom that is being observed by collapsing the state vector—with the degree of freedom remaining indeterminate and trans-empirical.

A remarkable conclusion of the Bell analysis is that a quantum degree of freedom does not have any precise and determinate value before it is observed. Clearly, the degree of freedom exists—before and after it is subjected to an experiment—and not being able to observe its specific values refers to its *mode of existence*.

It has been discussed in Sect. 3.4 that the mode of existence of a degree of freedom is to exist in all of its totality as an entire space; metaphorically speaking, the degree of freedom simultaneously exists over the entire range of all its possible values. For example, x, the position degree of freedom of a quantum particle, simultaneously exists at all points and forms the degree of freedom space $\Re = \{x : -\infty \leq x \leq +\infty\}$: the real line. Hence, the position degree of freedom never "takes" a particular value—unlike the classical random variable X that actually *takes* particular values x with probability $P(x)$.

The mode of existence of the degree of freedom is *not empirically observable*, and hence, the degree of freedom is trans-empirical.

The fundamental reason that classical probability fails to explain quantum indeterminacy is because the particular and specific values of a quantum degree of freedom *do not exist objectively*, that is, do not exist independent of a measurement. The degree of freedom does not have a particular value and is intrinsically an entire space $\mathcal{F}$ that is trans-empirical; this is an explanation of why the particular values of the degree of freedom do not exist objectively. Since the degree of freedom *does not exist objectively* for a quantum system, a sample space Ω *does not exist* for a quantum degree of freedom; hence, a probability distribution function cannot be assigned to the particular values of the degree of freedom space $\mathcal{F}$.

Quantum probability—represented in Fig. 7.1—necessarily follows from Heisenberg's idea that all physical quantities are mathematically represented by Hermitian operators $\mathcal{O}$ and experiments measure the expectation value of these operators.

According to the quantum theory of measurement, one can only assign a likelihood to the outcome of a *process of measurement*, which mathematically consists of applying a *projection operator* on the state vector of the degree of freedom. All the observable properties of the degree of freedom are represented by operators that act on the Hilbert space of states. On repeatedly applying an operator to the quantum state, one obtains the average value of the operator for that state.

The algebra of all the Hermitian projection operators defined on state space, namely, $\{\Pi_n;\ n = 1, 2, 3, \ldots\}$, replaces the sample space Ω of classical probability theory. Quantum probability *assigns probabilities to a complete set of Hermitian projection operators* Π_n—and consequently to any function of these operators as well. The assignment of probabilities to Hermitian operators fully takes into account the non-commutative structure and contextual nature of quantum mechanics. The projection operators also play a central role in the theory of measurement as discussed in Chap.9.

Every Hermitian (or more generally self-adjoint) operator is the sum of projection operators, which in turn obeys the completeness equation, discussed in (5.7), and given as follows[7]:

$$\mathcal{O} = \sum_{in=1}^{N} \lambda_n \Pi_n; \qquad \Pi_m \Pi_n = \delta_{m-n} \Pi_n; \quad \sum_n \Pi_n = \mathbb{I} \tag{7.31}$$

Since $\Pi_n^2 = \Pi_n$, (5.10) shows that the expectation value of Π_n, namely, $p_n \in [0,1]$. The process of measurement yields the expectation value of the projection operator Π_n for the state ψ; hence,

$$E_\psi[\Pi_n] = p_n; \qquad \Pi_n^2 = \Pi_n \;\; \Rightarrow \;\; 0 \leq p_n \leq 1$$

The expectation value of the completeness equation of Π_n given in (5.7)

$$\sum_n \Pi_n = \mathbb{I}$$

yields the following result:

$$\Rightarrow \;\; \sum_{n=1}^{N} p_n = \sum_n E_\psi[\Pi_n] = E_\psi[\sum_n \Pi_n] = E_\psi[\mathbb{I}] = 1 \tag{7.32}$$

The remarkable property $\Pi_n^2 = \Pi_n$ of the projection operators yields $p_n \in [0,1]$, and hence, p_n can be interpreted as the *probability* of Π_n *detecting* the state vector $|\psi\rangle$, as shown in Fig. 7.8, by *collapsing* the state vector to $\langle\psi|\Pi_n|\psi\rangle$. Furthermore, Π_n's obeying the completeness equation yields the required result that the total probability is 1, as given in (7.32).

The average value of the projection operator Π_n for a given quantum state $|\psi\rangle$ is equal to the probability p_n.[8] The projection operators yield the following quantum probabilities:

$$p_n = E_\psi[\Pi_n] \in [0,1]: \quad \text{probability of } \Pi_n \text{ observing the quantum state} \tag{7.33}$$

$$\sum_{n=1}^{N} p_n = 1; \qquad \{p_n \in [0,1];\, n = 1,2,\ldots,N\}: \quad \text{Quantum probabilities}$$

[7]The most general formulation of quantum measurement theory is in terms of self-adjoint operators that are not necessarily projection operators and which are called positive-operator valued measure (POVM) [1].

[8]Note that probability $p_n = |c_n|^2$, where the state vector $|\psi\rangle = \sum_n c_n|\psi_n\rangle$ and $\Pi_n = |\psi_n\rangle\langle\psi_n|$; see discussion in Sect. 9.4.

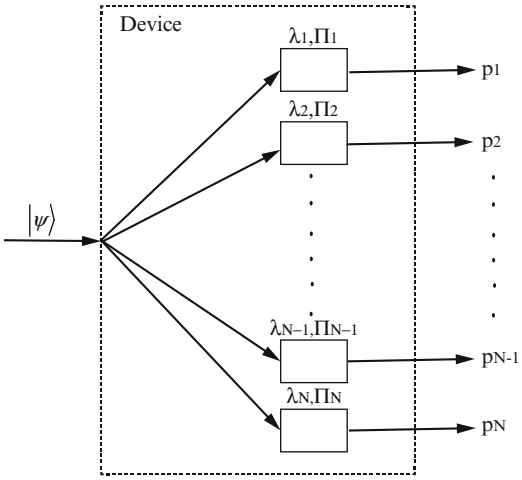


Fig. 7.8 Measurement and projection operators (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Figure 7.8 illustrates the role played by the projection operators. Consider making a measurement of the quantum state $|\psi\rangle$ using the Hermitian operator $\mathcal{O}$; a device is designed to embody the properties of operator $\mathcal{O}$, which entails constructing the projection operators Π_n that are characterized by their eigenvalues λ_n.

Every time $|\psi\rangle$ is measured by the device, as shown in Fig. 7.8, the state is detected by collapsing it to only *one* of the projection operators Π_n; on repeatedly measuring $|\psi\rangle$, it is found that there is a probability p_n that projection operator Π_n will detect the state $|\psi\rangle$. All the observable properties of the state $|\psi\rangle$ can be expressed in terms of p_n and λ_n.

In summary, quantum mechanics defines a theory of probability by *assigning probabilities* p_n to the expectation value of *projection operators* Π_n and *not* to the specific values of the degree of freedom. When the state vector is experimentally observed by the projection operator Π_n, the *probability* that the projection operator Π_n will *detect* the state vector is given by p_n.

Every Hermitian operator has a spectral representation in terms of the projection operators Π_n and its eigenfunction $|\psi_n\rangle$ and eigenvalues λ_n, as given in (5.8); more precisely, one has, for operator $\mathcal{O}$ and its function $f(\mathcal{O})$, the following:

$$\mathcal{O} = \sum_n \lambda_n \Pi_n; \quad \Pi_n = |\psi_n\rangle\langle\psi_n|; \quad \langle\psi_n|\psi_n\rangle = 1$$

$$f(\mathcal{O}) = \sum_n f(\lambda_n)\Pi_n$$

All expectation values can then be evaluated as follows:

$$E_\psi[\mathcal{O}] = \sum_n \lambda_n E_\psi[\Pi_n] = \sum_n \lambda_n p_n$$

$$E_\psi[f(\mathcal{O})] = \sum_n f(\lambda_n) p_n$$

Quantum probability is a *synthesis* of *operator algebra* and *probability theory*, similar to the synthesis of linear algebra and calculus in defining Hilbert state space. The random variables of classical probability theory, which can be integers or real numbers, are replaced by Hermitian projection operators. Every projection operator has eigenvalues and eigenstates—and hence has an underlying linear vector space structure that is absent for the random variables of classical probability theory.

Position Degree of Freedom

Consider the important example of the position degree of freedom x; quantum probability assigns a probability to the position projection operators detecting the state vector $|\psi\rangle$ and not to the occurrence of a particular value of a quantum degree of freedom.

There is a common misconception that $|\psi(x_0)|^2$ is the probability of finding a quantum particle's degree of freedom, namely, its position x, as having a specific value of x_0; in fact, $|\psi(x_0)|^2$ is the probability of the observation occurring at the projection operator $|x_0\rangle\langle x_0|$ that belongs to the measuring device, such as the photographic plate discussed in Sect. 9.2.

What is measured and observed is the *expectation value* of the position projection operator $|x_0\rangle\langle x_0|$. As discussed in detail in Sect. 9.2, the probability of the quantum state collapsing at projection operator $|x_0\rangle\langle x_0|$ is given by

$$p(x) = E_\psi\Big(|x_0\rangle\langle x_0|\Big) = \mathrm{tr}\Big(\rho|x_0\rangle\langle x_0|\Big) = |\psi(x_0)|^2; \quad \rho = |\psi\rangle\langle\psi|$$

The completeness equation is given by $\int \mathrm{d}x\,|x\rangle\langle x| = \mathbb{I}$ and yields

$$\mathrm{tr}(\rho) = \mathrm{tr}(\rho\mathbb{I}) = \int \mathrm{d}x\,\mathrm{tr}\Big(\rho|x\rangle\langle x|\Big) = \int \mathrm{d}x\,\psi^*(x)\psi(x)$$

$$\Rightarrow \int \mathrm{d}x|\psi(x)|^2 = \int \mathrm{d}x\,p(x) = 1: \quad \text{Total probability is 1.}$$

Noteworthy 7.2: Paradox of probability in quantum mechanics

Every quantum state has (nonlocal) information about the possible outcomes of all the *projection operators* $|x\rangle\langle x|$; for example, for a particle in a box, the observed value of the projection operator at *all* of the forbidden nodal points is *always* zero. Nonlocal information seems to contradict the relativistic principle that information can only propagate at the speed of light or slower, but quantum mechanics avoids this contradiction; however, it is not clear what is the mechanism by which the particle has nonlocal information.

Every outcome is statistically independent of all the other outcomes; every time a measurement is performed, the projection operators can, in principle, detect the

state vector *anywhere* that the degree of freedom exists. Quantum mechanics *forbids* the observer from having *any à priori knowledge* about the result of a particular *outcome*, since each outcome is statistically independent. The state vector encodes all information about the degree of freedom, but it is not clear how the state vector can build up the observed pattern.

For a particle confined inside a box, the probability distribution is seen to emerge by the repeated observations by the position projection operator of the particle's state vector. Similarly, in the two-slit experiment without detecting which path is taken, electrons are sent in one by one and projection operators at the screen detect each incoming electron, one by one. How does the interference pattern emerge? Namely, what is keeping count?

To appreciate the paradox of quantum mechanics, one needs to contrast it with classical probability theory. There is no need to keep count of the outcomes of the sampling of a classical random variable since each possible outcome has its own *intrinsic* probability of occurrence. In contrast, the particular values of a quantum degree of freedom do not have any intrinsic probability of occurrence.

Since each measurement is identical and every particular outcome is the result of a measurement, one needs to keep count of the outcomes to form the quantum probability distribution. The question arises as to *what is the mechanism in Nature that is keeping count* of all the experimental outcomes? It is, after all, only the result of many repeated measurements that yields the expected probability distribution. Quantum mechanics has no answer to this question.

7.10 A Metaphor

Heisenberg made the following observation: *It is a trite saying that 'Analogies cannot be pushed too far,' yet they may be used to describe things for which our language has no words* [17]. Since the quantum entity, and the quantum degree of freedom in particular, is something that human language does not ordinarily encounter, a metaphorical description of the quantum entity, and of quantum probability in general, is given.

Consider a quantum degree of freedom that is inside a safe box; the safe is unlike anything classical in the sense that the door of the safe, in principle, *can never be opened*—as shown in Fig. 7.9. Keys to the safe are metaphors for operators $\mathcal{B}$; these keys are special in the sense that they *fit* the locks of the safe.

On repeatedly applying the keys to the locks to open the safe, it is found that the safe does not open but the keys *themselves* undergo a transformation! The transformed state of the keys yields all the information that can be extracted from and about the degree of freedom inside the box. The locks are a metaphor for the state vector $|\psi\rangle$ of the quantum degree of freedom, and the transformed state of the key is a metaphor for the expectation value of the operators, namely, $E_\psi[\mathcal{B}]$.

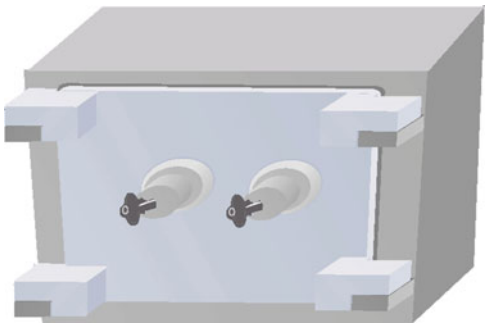


Fig. 7.9 A safety box that, in principle, *can never be opened* is a metaphor for the mode of existence of a quantum degree of freedom $\mathcal{F}$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

One can push this analogy further and ask how, in the first place, did the degree of freedom get inside a safe that cannot be opened? This comes back to the question as to what is a quantum entity? One of the conclusions drawn in Sect. 2.12 is that the quantum entity *exists* as the inseparable combination of the degree of freedom and its quantum state, which describes the quantum entity to an observer. Following Heisenberg's advice, one cannot push the analogy too far since any physical safe can be opened, if necessary by being broken up—whereas the degree of freedom can never be directly observed. However, the analogy remains useful.

Since it is permanently enveloped by the "veil" of the quantum state, the degree of freedom is not directly empirically accessible; in other words, being empirically inaccessible is the *mode of existence* of the degree of freedom. Quantum mechanics tries to explain how Nature exists; as with all explanations in physics, there is no attempt to address why Nature is the way it is. Moreover, unlike classical physics that ontologically provides a description of what Nature is as such, in itself, independent of any observer, quantum mechanics instead takes observation to be a fundamental property of Nature and formulates a description of Nature based on what an *observer* can *know* about Nature.

7.11 Summary

In summary, quantum mechanics is a theory of probability that is a synthesis of the linear structure of Hermitian operators with the concept of uncertainty and indeterminateness.

It was shown that quantum mechanics defines a theory of quantum probability that is a generalization of classical probability. One of the conclusions reached in this chapter is that although the degree of freedom space $\mathcal{F}$ superficially looks similar to the sample space Ω of a classical random variable, this identification is *not* correct. The fundamental reason that quantum probability is more general than classical probability is because, unlike sample space Ω, the space $\mathcal{F}$ has an associated linear vector (Hilbert) space $\mathcal{V}(\mathcal{F})$ as well as Hermitian operators $\mathcal{O}(\mathcal{F})$ representing physical observables.

A classical random variable X takes a range of values. The different values x of the classical random variable, namely, the sample values of the random variable X, exist intrinsically and objectively, and hence, an intrinsic probability $p(x)$ can be assigned to each value of the classical random variable; in particular, being observed or not makes no difference to this assignment of classical probabilities to the elements of the sample space Ω.

The BKS theorem shows, conclusively and clearly, that *no classical system*—including one that contains classical random variables—can explain the behavior of a quantum entity.[9] In particular, the *quantum degree of freedom*—unlike a classical random variable—has *no precise intrinsic value before it is observed*. The violation of the BKS inequality demonstrates that, at the most fundamental level, a quantum degree of freedom—unlike the case of a classical random variable—is intrinsically indeterminate and can never be observed directly. It shows that quantum indeterminacy is different from the classical randomness of a classical random entity.

The violation of the BKS inequality results from representing the quantum system by a state vector and operators acting on the state vector—and by defining probabilities based on projection operators; in particular, the *dynamics* of a quantum state—given by its time evolution driven by the Hamiltonian operator—does not enter the Bell or BKS analysis; what is at issue is the nature of the quantum entity itself. Violation of the Bell inequality is one of the more special criteria for deciding if a physical system can be described by quantum or classical probability.

The quantum degree of freedom is a trans-empirical quantity that simultaneously "exists" in all of its possible states; the trans-empirical existence of the degree of freedom, as the term denotes, cannot be directly observed. The experimentally observable properties of the degree of freedom are described by its Schrödinger *state vector* (or equivalently by its density matrix ρ).

As shown in Fig. 7.1, a quantum entity is a pair: the degree of freedom and its state vector ψ. By interposing the *state vector* between the indeterminate degree of freedom and its (degree of freedom's) observable properties—encoded by the operators $\mathcal{B}$—quantum mechanics interprets the notion of quantum *indeterminacy* as a form of *quantitative uncertainty*.

When a property of a quantum degree of freedom is experimentally observed, quantum probability assigns a *likelihood* that the state vector's collapse will be detected by a particular *projection operator*. By assigning probabilities to projection operators, an entirely new and distinct field of quantum probability was created by the founders of quantum mechanics.

With the wisdom of hindsight, it can now be said that the genius of Werner Heisenberg was not just to have discovered the operator formulation of quantum mechanics but to have discovered quantum probability as well and that too in 1925—several years before the theory of classical probability was rigorously defined in 1933 by Kolmogorov [35].

[9] Ignoring the special case of spin 1/2.

Quantum Superposition

8

There are two forms of quantum indeterminacy, namely, that of the indeterminate degree of freedom reflected in the state vector and that of indeterminate paths expressed in the superposition of probability amplitudes, introduced in Sects. 2.7 and 2.10. For both cases, the linearity of the Schrödinger equation, combined with the intrinsic indeterminacy of the quantum entity, leads to the physics of superposition—with each case having distinctive and specific features.

In the subsequent sections, the following topics are studied:

- The superposition of state vectors
- The superposition of probability amplitudes for determinate and indeterminate paths
- The Mach–Zehnder interferometer and the superposition of photons
- The dependence of quantum superposition on measurements
- The quantum eraser and quantum superposition

Classical Superposition

To highlight the nonclassical properties of quantum superposition, consider first the case of classical superposition.

In classical physics, superposition arises in the study of waves. A typical wave, such as a shallow water wave, obeys a linear wave equation and consists of periodic oscillations of an underlying classical medium, namely, water.

A wave is an extended object, spread over space, and is described by the amplitude $h_1(t,r)$, defined as the value (of the displacement) of the periodically oscillating medium at point r. The question naturally arises: what happens when two waves, described by amplitudes $h_1(t,r)$ and $h_2(t,r)$, overlap—as shown in Fig. 8.1a? For linear waves, the answer is very simple: The resultant wave is obtained by *adding* the amplitude of the two waves at every point of space, called *superposition*, with the following superposed wave:

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8_8, © Springer Science+Business Media New York 2013

$$h(t,r) = h_1(t,r) + h_2(t,r)$$

Note the classical superposed wave at any given point, as shown in Fig. 8.1a, has only *one value*. The value of the resultant wave is *different* from the two component waves—with the values at different points taking a value that lies in between the sum and difference of the values of the two component waves. In summary, for classical waves, once the component waves have been superposed, there is only *one value* for the displacement of the resultant wave.

8.1 Superposing State Vectors

Let state vectors ψ_1 and ψ_2 be energy eigenfunctions with energy E_1 and E_2, as shown in Fig. 8.1b. The state vector embodies the indeterminacy of the underlying degree of freedom $\mathcal{F}$. Furthermore, due to the linearity of the Scrhödinger equation, the superposed state vector is given by *adding* the two state vectors

$$|\psi\rangle = a|\psi_1\rangle + b|\psi_2\rangle; \qquad |a|^2 + |b|^2 = 1 \tag{8.1}$$

and is *also* a solution of the Schrödinger equation. The superposed state vector $|\psi\rangle$ encodes both the indeterminacy of $\mathcal{F}$ and the dynamics arising from the Schrödinger equation.

What is the interpretation of the superposed state vector $|\psi\rangle$? What are its physical properties?

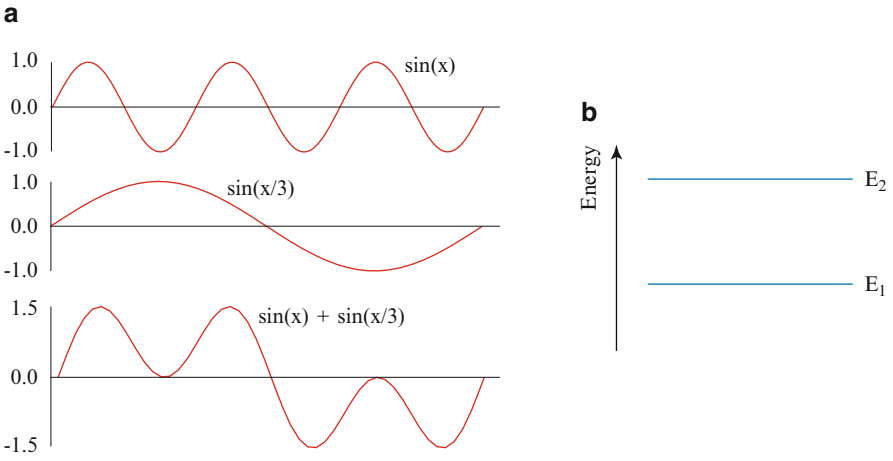


Fig. 8.1 (**a**) Superposition of classical waves. (**b**) Superposition of two quantum state vectors, with energy eigenvalues of E_1 and E_2 (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

The superposed state vector is a trans-empirical state, with the degree of freedom existing simultaneously in two distinct quantum states. Its discrete and trans-empirical nature is revealed in the measurement of energy of the superposed state ψ: *Every* measurement of the state's energy results in *only* the discrete energy E_1 *or* E_2 being observed. Hence, unlike the classical case, the superposed state vector $|\psi\rangle$ "remembers" the component state vectors out of which it is composed.

If the energy measurement is *repeated* many times, it will be found that the *average* energy $\bar{E}$ of ψ, as derived in (9.5), is given by

$$\bar{E} = |a|^2 E_1 + |b|^2 E_2 \tag{8.2}$$

The intermediate value of energy $\bar{E}$ is realized by observing energies E_1 or E_2 with *probability* $|a|^2$ or $|b|^2$, respectively—and *not* by the value $\bar{E}$ being directly observed as the energy of the state ψ. Hence, in contrast to the classical wave that has displacement very different from the component waves, *every time* the energy of ψ is measured, it is either E_1 or E_2—but *never* any other value.

In summary, classical superposition produces an empirical wave from the component empirical waves. In contrast, quantum superposition creates a trans-empirical state $|\psi\rangle$ with the superposed state being simultaneously in two possible states, specified by the component state vectors $|\psi_1\rangle$ and $|\psi_2\rangle$, which continue to preserve their identity. When a measurement is performed on $|\psi\rangle$, the energy of each component has a *likelihood* of being detected.

As discussed in Sect. 5.8, to measure the average value of an operator $\mathcal{O}$ for the state $|\psi\rangle$, one needs to first represent $|\psi\rangle$ as a trans-empirical superposition of the eigenfunctions $|\chi_n\rangle$ of $\mathcal{O}$, namely, as in (5.34)

$$|\psi\rangle = \sum_n c_n |\chi_n\rangle$$

The decomposition of $|\psi\rangle$ into a superposed trans-empirical state is a fundamental precursor to the process of quantum measurement.

Superposition for Spin $1/2$

Consider the superposition of a two-state spin $1/2$ state vectors. The eigenstates of σ_z, the operator for measuring the z-component of the spin, are the following:

$$|+\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \qquad |-\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \;\Rightarrow\; \sigma_z|\pm\rangle = \pm\frac{\hbar}{2}|\pm\rangle$$

where

$$\sigma_z = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

An example of a superposed state given in (8.1), for the case of the spin $1/2$ state vectors, is the following:

$$|\Psi\rangle = \alpha|+\rangle + \beta|-\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}; \quad \langle\Psi| = \begin{bmatrix} \alpha^* & \beta^* \end{bmatrix} \tag{8.3}$$

All measurements for the value of spin observable σ_z on state vector $|\Psi\rangle$ will produce either $+\hbar/2$ or $-\hbar/2$ and no other value. The average value of the z-component of spin for the superposed state—which is a concrete realization of (8.2)—is given by

$$E_\Psi[\sigma_z] = \langle\Psi|\sigma_z|\Psi\rangle = \frac{\hbar}{2}\Big[|\alpha|^2 - |\beta|^2\Big]; \quad |\alpha|^2 + |\beta|^2 = 1 \tag{8.4}$$

8.2 Probability and Probability Amplitudes

The quantum degree of freedom $\mathcal{F}$ is indeterminate and trans-empirical; a description of the quantum degree of freedom at a particular instant is given by the state vector, namely, $|\psi\rangle$. The likelihood of the projection operator $|x\rangle\langle x|$ observing the degree of freedom $\mathfrak{R} = \mathcal{F}$ is given by $\mathrm{tr}(|x\rangle\langle x|\rho) = |\langle x|\psi\rangle|^2$, where $\rho = |\psi\rangle\langle\psi|$.

Another class of quantum indeterminacy, as discussed in Sect. 2.10, consists of a quantum particle taking *indeterminate trans-empirical paths* in evolving from an arbitrary initial state vector to an arbitrary final state vector, as shown in Fig. 8.2. The *probability amplitude* is used for describing the indeterminate and trans-empirical paths of a quantum particle.

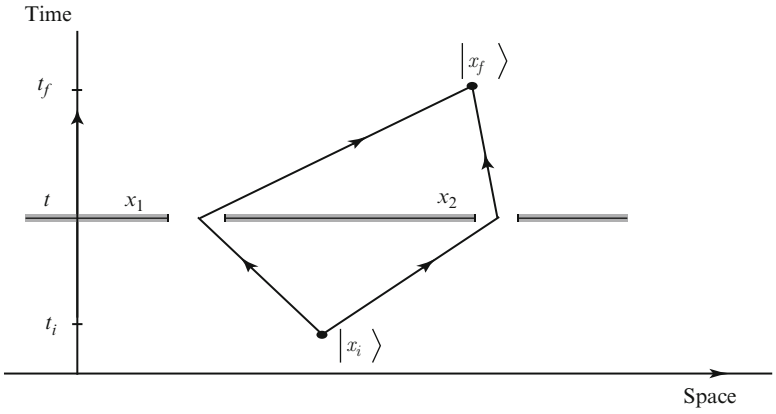


Fig. 8.2 A quantum particle's space–time trajectory is empirical (determinate) or trans-empirical (indeterminate), depending on whether a measurement is performed or not performed to determine the path taken, respectively (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

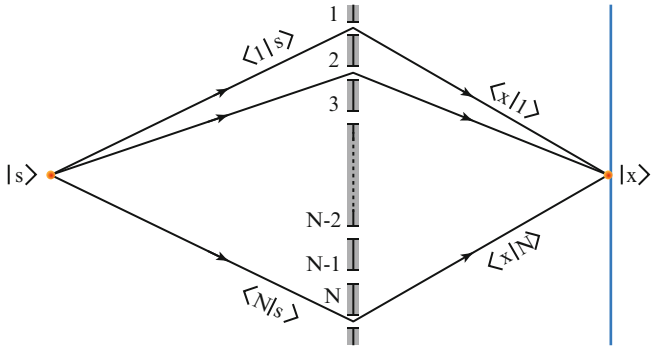


Fig. 8.3 Probability amplitudes for transition from initial state vector $|s\rangle$ to final state vector $|x\rangle$ via N different possible intermediate paths (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Consider the case of an electron making a transition from the initial position eigenstate $|x_\mathrm{i}\rangle$ at time t_i to the final position eigenstate $|x_\mathrm{f}\rangle$ at time t_f, via *two slits*, as shown in Fig. 8.2.[1] In all subsequent discussions in this chapter the space–time trajectories, be they determinate or indeterminate, are not shown explicitly, but instead are assumed.

In particular, in Fig. 8.3, the layout of the N-slits only in space is shown, with the time evolution of the electron from its initial to final position being assumed. The focus is on measurements of the initial and final states of the electron and, in particular, if any measurements are made to determine the path taken by the electron. From these measurements (or lack of them thereof) the space–time trajectories—be they determinate or indeterminate—can be deduced.

Consider a quantum system making a transition from an arbitrary initial state function $|\psi\rangle$ to an arbitrary final state function $|\eta\rangle$; the probability amplitude for this transition is given by the scalar product $\langle\eta|\psi\rangle$, and the *likelihood* for this transition is given by $|\langle\eta|\psi\rangle|^2$. In other words, the probability and probability amplitude for the transition are given by the following:

$$\text{Initial state function}: \ |\psi\rangle; \ \text{Final state function}: \ |\eta\rangle$$

$$\text{Probability Amplitude}: \ \langle\eta|\psi\rangle; \ \text{Probability of transition}: \ |\langle\eta|\psi\rangle|^2$$

A quantum particle making a transition from an initial to a final state can go through N intermediate *paths*, as shown in Fig. 8.3. The law for combining the effect of these paths, and which yields the probability amplitude, has two different cases:

- No measurement is made between the initial and final positions of the electron, and the information on which path has been taken by the quantum particle is

[1] Position measurements are made by projection operators $|x\rangle\langle x|$ and discussed in Sect. 9.2.

not known. Hence, the intermediate paths are *indistinguishable*; the electron's path from its initial and final position is *indeterminate* and trans-empirical; the quantum particle exists in all possible paths.

- The intermediate paths are *distinguishable* due to measurements carried out to determine which path is taken. All the paths are now *determinate*, exist objectively, and are empirical; the quantum particle takes a unique path from initial to final state.

8.3 Empirical and Trans-Empirical Paths

The case of the two slits has been discussed in detail in Sect. 3.7 to illustrate the paradoxes of quantum superposition. The discussion is now extended to the general N-slit case.

Consider the case of an initial state vector $|s\rangle$ making a transition to a final state vector $|x\rangle$ via N intermediate slits given by $|i\rangle$; $i = 1, 2, \ldots, N$ and shown in Fig. 8.3. The experimental realization of this transition is for electrons to start at point $|s\rangle$ and then, later, be detected at final position $|x\rangle$ with the appropriate projection operators.

In going from $|s\rangle$ to $|x\rangle$, the particle can go through any of the N-slits. The probability amplitude for going from state observed at $|s\rangle$ to the one observed at $|x\rangle$ is given by $\langle x|s\rangle$. What is the probability that the particle starting at $|s\rangle$ arrives at $|x\rangle$?

There are two very different expressions for the transition probability P. Unlike classical physics, the answer depends on whether the intermediate states are *distinguishable* or *indistinguishable*, namely, whether a *measurement* is performed to ascertain which one of the N slits the particle went through.

- When the intermediate paths are *distinguished* by performing a measurement, the path taken is a determinate and *empirical path*; the *probability* for going through the different empirical paths *is added* and yields

$$P_{\mathrm{D}} = |\langle x|s\rangle|^2 = \sum_{i=1}^{N} \left|\langle x|i\rangle\langle i|s\rangle\right|^2$$

$$\Rightarrow P_{\mathrm{D}} = \sum_{i=1}^{N} P_{xi}P_{is};\ \ P_{xi} = |\langle x|i\rangle|^2;\ \ P_{is} = |\langle i|s\rangle|^2 \tag{8.5}$$

  Classical probability theory yields the transition probability P_{D} and corresponds to the case when all the intermediate states are distinguished and empirical and which in turn implies that the paths exist objectively.
- For the case when the intermediate paths are *indistinguishable*, the paths taken are indeterminate and *trans-empirical*. The probability amplitudes for the transition via one of the trans-empirical paths *are added* to yield the total probability amplitude; this yields for $\psi_{xi} = \langle x|i\rangle$ the following:

$$\langle x|s\rangle = \sum_{i=1}^{N} \langle x|i\rangle\langle i|s\rangle = \sum_{i=1}^{N} \psi_{xi}\psi_{is} \tag{8.6}$$

$$P_{\mathrm{I}} = \left|\langle x|s\rangle\right|^2 = \left|\sum_{i=1}^{N} \langle x|i\rangle\langle i|s\rangle\right|^2 = \left|\sum_{i=1}^{N} \psi_{xi}\psi_{is}\right|^2 \tag{8.7}$$

$$\Rightarrow P_{\mathrm{I}} = \sum_{i=1}^{N} P_{xi}P_{is} + \sum_{ij;i\neq j}^{N} \langle x|i\rangle\langle i|s\rangle\langle x|j\rangle^*\langle j|s\rangle^*$$

There is no analog of (8.7) in classical probability theory since the indeterminate paths do not objectively exist. The paths are indeterminate and trans-empirical, with the electron simultaneously existing in all the possible paths.

The quantum result P_{I} is nonclassical and involves the *interference* of the different paths due to the inclusion of off-diagonal elements in P_{I}. The superposition of indistinguishable and *indeterminate* paths is an expression of quantum probability discussed in Chap. 7; there is no classical probability distribution function that can explain the results of a transition mediated by indistinguishable and indeterminate paths.

In summary, the fundamental difference between the two cases, having distinguishable and indistinguishable intermediate paths, is how the probability amplitude is composed.

- For the *distinguishable paths*, all the paths taken by the electron exist objectively, and the *probabilities* for all the paths are composed as per the rules of classical probabilities and yield $P_{\mathrm{D}} = \sum_{i=1}^{N} P_{xi}P_{is}$.
- For the *indistinguishable paths*, the paths are indeterminate; the *probability amplitudes* for all the indeterminate paths are composed, namely, $\sum_{i=1}^{N} \psi_{xi}\psi_{is}$ the probability is given by $P_{\mathrm{I}} = |\sum_{i=1}^{N} \psi_{xi}\psi_{is}|^2$—thus yielding interference terms in P_{I}.

To simplify the notation, define ϕ_i, the probability amplitude to go through slit i—shown in Fig. 8.4—as given by

$$\phi_i = \langle x|i\rangle\langle i|s\rangle \equiv \psi_{xi}\psi_{is}$$

For the distinguishable paths, the probability of going from the state observed at $|s\rangle$ to the one observed at $|x\rangle$, from (8.5), is given by

$$P_{\mathrm{D}} = \sum_{i=1}^{N} |\phi_i|^2 \tag{8.8}$$

For the case of indistinguishable and indeterminate paths, the probability amplitudes for *all* the various paths through N-slits are added, that is, are superposed. From (8.6) and (8.7), the indistinguishable case yields

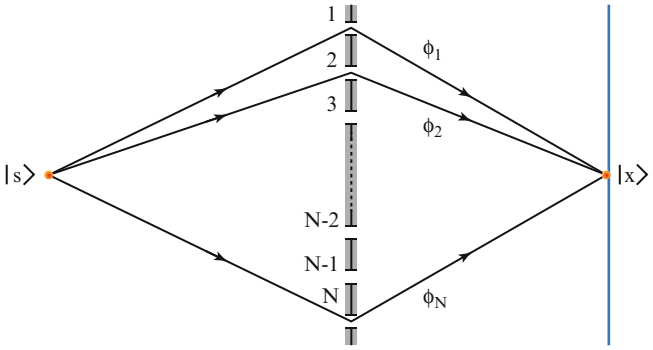


Fig. 8.4 The probability amplitude $\phi_1, \phi_2, \ldots, \phi_n$ for the different paths (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

$$\langle x|s\rangle = \sum_{i=1}^{N} \phi_i \tag{8.9}$$

$$P_{\mathrm{I}} = |\langle x|s\rangle|^2 = |\sum_{i=1}^{N} \phi_i|^2 = \sum_{i=1}^{N} |\phi_i|^2 + \sum_{ij;i\neq j}^{N} \phi_i \phi_j^* \tag{8.10}$$

For the case of indistinguishable paths, the indeterminate paths consist of all the possible determinate paths taken together, as one single unit (collection). The particle's path being indeterminate means that it simultaneously exists in all the possible determinate paths. The superposition of probability amplitudes for the indistinguishable case is obtained by adding all these determinate paths to obtain $\langle x|s\rangle$ as in (8.9). Although summing over indeterminate paths looks similar to the superposition of classical waves, it is fundamentally different: one is adding the probability amplitudes for different *possible determinate paths*, whereas for classical waves, one is adding the physical amplitude of oscillation of the material medium. Furthermore, as discussed above, the probability amplitude is no ordinary wave, but rather embodies all the *information* that can be extracted from the indeterminate quantum paths.

For the N-slit case, the interference term $\sum_{ij;i\neq j}^{N} \phi_i \phi_j^*$ given in (8.10) shows that — for indistinguishable paths a *single electron* simultaneously exists in *all* paths that go through the N slits. One can go even further and remove all the slits; the result is that the single electron *simultaneously exists in all possible paths* going from $|s\rangle$ to $|x\rangle$, with the probability for the transition $|\langle x|s\rangle|^2$ given by the interference of all the paths with each other. The approach of summing over all possible indeterminate paths yields the Feynman path integral and is discussed in detail in Chap. 11.

8.4 Successive Slits

One can have a situation where the quantum particle successively encounters a number of slits.

Figure 8.5 shows the case of the quantum particle encountering two successive double slits, one after the other. There are four possible probability amplitudes, corresponding to four possible paths to go from $|s\rangle$ to $|x\rangle$ given by the following:

$$\phi_1 = \langle x|a\rangle\langle a|1\rangle\langle 1|s\rangle; \quad \phi_2 = \langle x|a\rangle\langle a|2\rangle\langle 2|s\rangle \tag{8.11}$$

$$\phi_3 = \langle x|b\rangle\langle b|1\rangle\langle 1|s\rangle; \quad \phi_4 = \langle x|b\rangle\langle b|2\rangle\langle 2|s\rangle \tag{8.12}$$

- *Determinate paths*
  If the path taken by the particle is observed, then the probability of particle going from initial to final state is given, as in (8.8), by the addition of the probabilities for the individual paths and yields

$$P_{\mathrm{D}} = |\phi_1|^2 + |\phi_2|^2 + |\phi_3|^2 + |\phi_4|^2$$

- *Indeterminate paths*
  If all the four paths are indistinguishable, the paths taken by the quantum particle are trans-empirical, and hence the probability amplitude is given by the *superposition* of the probability amplitudes for the different paths and yields, as in (8.9), the following:

$$P_{\mathrm{I}} = \left|\phi_1 + \phi_2 + \phi_3 + \phi_4\right|^2$$

The generalization of the probability amplitude for many successive slits is straightforward; the quantum superposition of trans-empirical (indistinguishable) paths is the basis of the Feynman path integral that is discussed in Chap. 11.

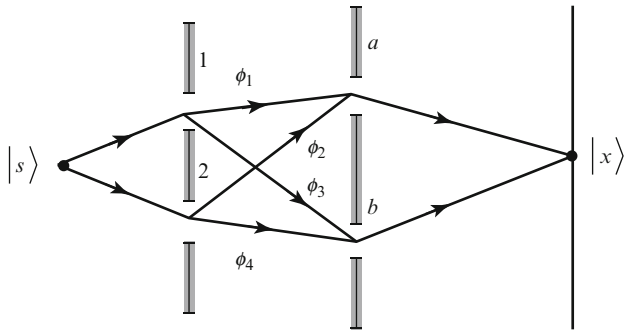


Fig. 8.5 Superposition of probability amplitude for indistinguishable possibilities (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

In summary, for the case when the *path* taken is *not known*, the total probability amplitude is obtained by superposing the probability amplitudes $\sum_i \phi_i$ for the indistinguishable paths and results in a trans-empirical state; the probability of transition is given by $P_{\mathrm{I}} = |\sum_i \phi_i|^2$.

For the case when the *path* taken is *known*, the quantum system does not obey the quantum superposition principle; instead, the transition probability is given by the result of classical probability, namely, $P_{\mathrm{D}} = \sum_i |\phi_i|^2$, which is the sum of $|\phi_i|^2$, namely, the probability of the electron taking the empirical and objectively existing path through slit i.

8.5 The Mach–Zehnder Interferometer

The discussion of the two-slit experiment in Sect. 3.7 using electrons is repeated using photons instead of electrons and for which one needs to employ the Mach–Zehnder interferometer. The two-slit experiment is revisited using photons to examine the role of measurement in causing a transition between the empirical and trans-empirical paths of a quantum entity.

Let the photon traveling in the x- and y-direction be denoted by state vector $|x\rangle$ and $|y\rangle$, respectively—as shown in Fig. 8.6—with

$$\langle x|x\rangle = 1 = \langle y|y\rangle; \qquad \langle y|x\rangle = 0$$

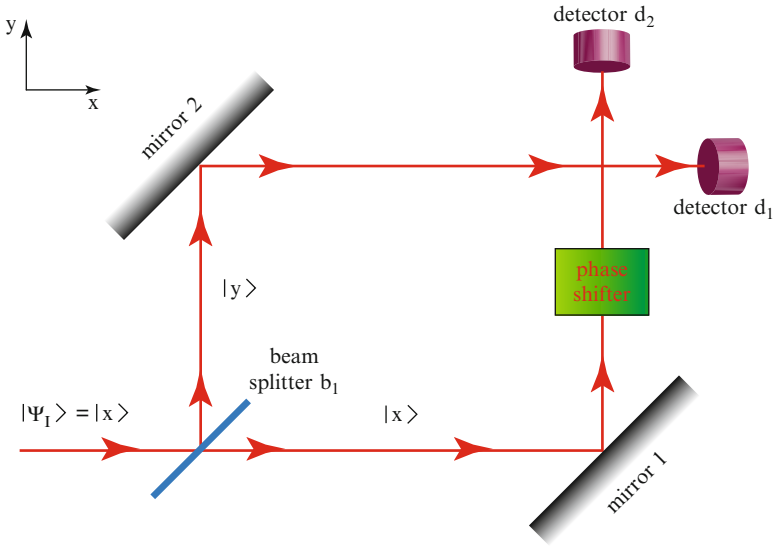


Fig. 8.6 A Mach–Zehnder interferometer arranged to show no interference (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

In this section, all the experiments consist of only a *single photon*—that is either taking (superposed) trans-empirical paths or a definite empirical path.

The detectors d_1 and d_2 are *projection operators* in the x- and y-direction and given by

$$d_1 = |x\rangle\langle x|; \qquad d_2 = |y\rangle\langle y| \tag{8.13}$$

In the Mach–Zehnder interferometer, as shown in Fig. 8.6, a *single photon*, in an initial quantum state denoted by $|\Psi_I\rangle = |x\rangle$, is directed toward a *beam splitter*,[2] namely, b_1, where the photon has an equal probability of passing through the mirror—or of being reflected upwards and hence transformed to state $|y\rangle$. The beam splitter conserves probability and performs the following unitary transformation $\mathcal{B}$ on the photon:

$$\begin{bmatrix} |x'\rangle \\ |y'\rangle \end{bmatrix} = \mathcal{B} \begin{bmatrix} |x\rangle \\ |y\rangle \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} |x\rangle \\ |y\rangle \end{bmatrix} \tag{8.14}$$

After going through the beam splitter, the single photon is in a trans-empirical superposed state and *simultaneously* travels along paths in the x- and y-direction, as shown in Fig. 8.6, and is subsequently reflected by the mirrors. Mirror 1 reflects the photon going along the x-axis to the y-axis, and mirror 2 reflects the photon going along the y-axis to the x-axis, as shown in Fig. 8.6. The unitary operator $\mathcal{M}$ representing the mirrors is given by

$$\begin{bmatrix} |x'\rangle \\ |y'\rangle \end{bmatrix} = \mathcal{M} \begin{bmatrix} |x\rangle \\ |y\rangle \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} |x\rangle \\ |y\rangle \end{bmatrix}$$

After being reflected by the mirror 1 and going through a phase shifter, the photon ends up at detector d_2; similarly, the photon reflected off mirror 2 ends up at detector d_1.

The phase shift is the result of slightly changing the photon's path length so that the photon reflecting off mirrors 1 travels a distance slightly different from the photon reflecting off mirror 2. A photon traveling along the two paths has a (slightly) different relative phase when it arrives at the detectors and which is denoted by ϕ. In symbols

$$\begin{aligned} |y\rangle \rightarrow \text{Phase shifter} \rightarrow e^{i\phi}|y\rangle : &\quad \text{after mirror 1} \\ |x\rangle \rightarrow \text{Phase shifter} \rightarrow |x\rangle : &\quad \text{after mirror 2} \end{aligned}$$

The unitary operator for the phase shift is given by

[2] A beam splitter is usually a half-silvered mirror that has an equal probability of the photon being reflected or of being transmitted without reflection.

$$\begin{bmatrix} |x'\rangle \\ |y'\rangle \end{bmatrix} = \mathcal{P} \begin{bmatrix} |x\rangle \\ |y\rangle \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix} \begin{bmatrix} |x\rangle \\ |y\rangle \end{bmatrix}$$

8.6 Determinate Empirical Paths: No Interference

The experiment shown in Fig. 8.6 detects the path taken by the photon: detector $d_1(d_2)$ will click only if the photon was reflected off mirror 1(2). The setup in Fig. 8.6 is analogous to the case of an electron traveling through the two slits *with detection*; continuously changing the phase shift (leading to a possible change in the intensity of photons received if there is interference) is analogous to continuously moving the detector up and down the screen, as in Fig. 3.7.

In terms of the state vectors, for this process, the initial state undergoes the following transformations:

$$\begin{aligned} |\psi_I\rangle &= |x\rangle \\ &\to \text{Beam splitter } b_1 \to \frac{1}{\sqrt{2}}[|x\rangle + |y\rangle] \\ &\to \text{Mirrors} \quad \to \frac{1}{\sqrt{2}}[|y\rangle + |x\rangle] \\ &\to \text{Phase shifter} \to \frac{1}{\sqrt{2}}[e^{i\phi}|y\rangle + |x\rangle] = |\psi_F\rangle \end{aligned} \tag{8.15}$$

The probability of the projection operators (detectors) d_1 and d_2 detecting the photon, denoted by P_1 and P_2, respectively, is given from (8.13) by the following:

$$\begin{aligned} P_1 &= \text{tr}(d_1\rho) = \text{tr}(|x\rangle\langle x|\rho) = |\langle x|\psi_F\rangle|^2 = \frac{1}{2} \\ P_2 &= \text{tr}(d_2\rho) = \text{tr}(|y\rangle\langle y|\rho) = |\langle y|\psi_F\rangle|^2 = \frac{1}{2} \\ P_1 &+ P_2 = 1; \quad \rho = |\psi_F\rangle\langle\psi_F| \end{aligned}$$

Since there is only a single photon in the system, only one of the detectors receives a photon and is the reason that $P_1 + P_2 = 1$. The result of the experiment is shown in Fig. 8.7; for a given phase difference ϕ, both detectors d_1 and d_2, on the average, receive the same number of photons.

As the relative phase ϕ is varied, the detector readings are constant and there is *no interference*, as shown in Fig. 8.7; the reason being that the information of the path taken by the photon is known. Hence, for all values of the phase difference, the path taken by the photon in the interferometer is empirical. The lack of interference is shown by the fact that the counts in detectors d_1 and d_2 are independent of the phase shift ϕ.

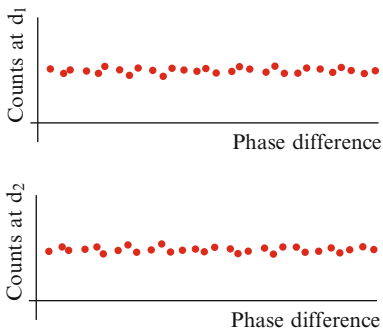


Fig. 8.7 No interference; which path taken by the photon is known (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

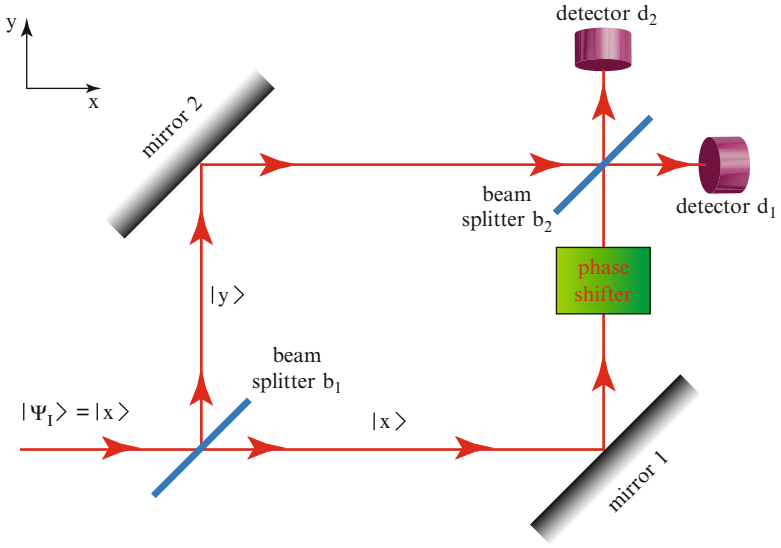


Fig. 8.8 A Mach–Zehnder interferometer arranged to show interference (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

8.7 Indeterminate Trans-Empirical Paths: Interference

Suppose that, as shown in Fig. 8.8, a *second* beam splitter b_2 is put in just before the photon reaches the detectors. By putting the second beam splitter b_2, the information about which path the photon took is lost since the photon received in detector d_1 and d_2 could have come from *either* the *x*- *or* *y*-direction. The photon is now in a superposed trans-empirical state. More precisely, as shown in Fig. 8.8 and using (8.14), the final state is the following:

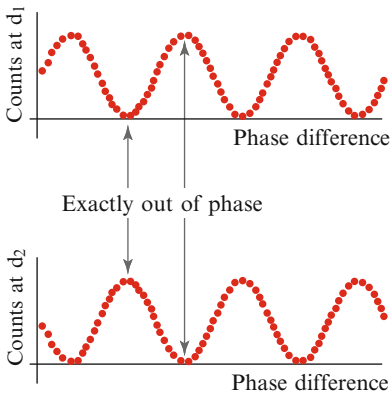


Fig. 8.9 Interference; which path that is taken by the photon is not known (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

$$
\begin{aligned}
|\psi_{\mathrm{I}}\rangle \;&=\; |x\rangle \\
&\to \text{Beam splitter } b_1 \to \frac{1}{\sqrt{2}}[|x\rangle + |y\rangle] \\
&\to \text{Mirrors} \quad \to \quad \frac{1}{\sqrt{2}}[|y\rangle + |x\rangle] \\
&\to \text{Phase shifter} \to \frac{1}{\sqrt{2}}[\mathrm{e}^{\mathrm{i}\phi}|y\rangle + |x\rangle] \\
&\to \text{Beam splitter } b_2 \to \frac{1}{2}\left(\mathrm{e}^{\mathrm{i}\phi}[|y\rangle - |x\rangle] + [|x\rangle + |y\rangle]\right) \\
&\to \frac{1}{2}\left[(1-\mathrm{e}^{\mathrm{i}\phi})|x\rangle + (1+\mathrm{e}^{\mathrm{i}\phi})|y\rangle\right] = |\psi_{\mathrm{F}}\rangle
\end{aligned}
\tag{8.16}
$$

The photon arriving at the detectors could have come from either path and hence what is received at the detector is a *superposed* state of the photon, with equal amplitude to have taken either path. The detectors hence now show interference, with the probability for the photon—having relative phase difference of ϕ—being detected at detectors d_1 and d_2 given by

$$P_1 = \mathrm{tr}\left(d_1\rho\right) = |\langle x|\psi_{\mathrm{F}}\rangle|^2 = \sin^2\left(\frac{\phi}{2}\right) \tag{8.17}$$

$$P_2 = \mathrm{tr}\left(d_2\rho\right) = |\langle y|\psi_{\mathrm{F}}\rangle|^2 = \cos^2\left(\frac{\phi}{2}\right) \tag{8.18}$$

$$P_1 + P_2 = 1 \tag{8.19}$$

The result of the experiment is shown in Fig. 8.9; for a given phase difference ϕ, detectors d_1 and d_2, on the average, receive different number of photons that is the result of the photon's probability amplitudes for the trans-empirical paths interfering. Since there is only one photon, the total probability of detection has to be unity and is shown in Fig. 8.9.

For the case of two beam splitters b_1 and b_2, the classical explanation is that the photon takes an empirical and objectively existing path—taking *either* path reflected off mirror 1 *or* path that reflects off mirror 2; the classical analysis predicts that intensity of photons received at the detectors should have no interference, which is shown by experiments to be incorrect.

In contrast, the quantum mechanical explanation is that the *single photon* exists in indeterminate trans-empirical paths—for which the single photon *simultaneously* exists in *both paths*. Since the two paths have a phase difference of ϕ, the photon existing in the *two paths* interferes with *itself*—either constructively or destructively—thus giving rise to the interference pattern. As expressed by Dirac "each photon .. interferes only with itself. Interference between two different photons never occurs" [10].

The interference that results from the superposed trans-empirical state has been experimentally confirmed time and again for a great variety of cases.

Needless to say, one can never *experimentally* observe the photon existing in both paths simultaneously; instead, one can only infer the existence of the trans-empirical paths due to the photon making a transition from its trans-empirical state to an empirical photon—that in turn is actually detected in an experimental apparatus.

8.8 Quantum Eraser

We take a closer look at the process of measurement to decide if we can erase the information on the path taken by the photon even *after* an experimental arrangement has apparently measured the path taken. The earlier experiments are reviewed, and then modifications are made to introduce the idea of erasure of information.

Recall in the experimental setup in Fig. 8.8 there are two beam splitters b_1 and b_2 that have been placed between the source of the photon and its detection. The paths taken by the photon arriving at detectors d_1 and d_2 are indeterminate and trans-empirical. Hence the photon exhibits interference and yields the interference pattern shown in Fig. 8.9.

One can ask the question: can the interference pattern for the case of a *single photon*, as discussed in Sect. 8.7, be erased after the photon has left the source? The quantum eraser is such a device. It is important to note that the final state of the quantum eraser consists of *two photons*, namely, the initial photon and the second photon that is due to the down conversion of the initial photon.

The quantum eraser can *erase* the interference of a single photon with itself but at expense of creating *another* photon in the device that contains information on the path taken by the original photon.

Moreover, the quantum eraser also *creates* new forms of interference, but involving not one but, instead, *two photons*.

8.9 Erasing Interference

The quantum eraser is shown in Fig. 8.10.

- The photon from the source, *after* going through beam splitter b_1 but *before* reflecting off either mirror 1 or mirror 2, goes through a device, called a down converter, where a nonlinear process *splits* the *single* photon—conserving energy and momentum—into *two* photons. One of the resulting photons continues along the original path, and the other photon heads in an orthogonal direction, as shown in Fig. 8.10, with detectors d_3 and d_4 detecting the down-converted photon.
- Consider the case of there being two photons in the Mach–Zehnder interferometer. Even though we still have beam splitters b_1 and b_2 in place at the source of light and at detectors d_1 and d_2, respectively, there will *no interference*. This is because the down conversion render's the photon—detected by either d_1 or d_2—into an empirical photon that takes a well-defined and determinate path.
- Consider, for example, the case when a photon is detected by say detector d_4; we then know that the photon detected in *either* d_1 *or* d_2 has taken the path of being reflected off mirror 1, since only a photon traveling toward mirror 1 could have undergone the down conversion and sent a photon to detector d_4. Hence, in spite of having beam splitter b_2 in place, detectors d_1 and d_2 will *not* show any interference! And this is precisely what experiments confirm to be the case.

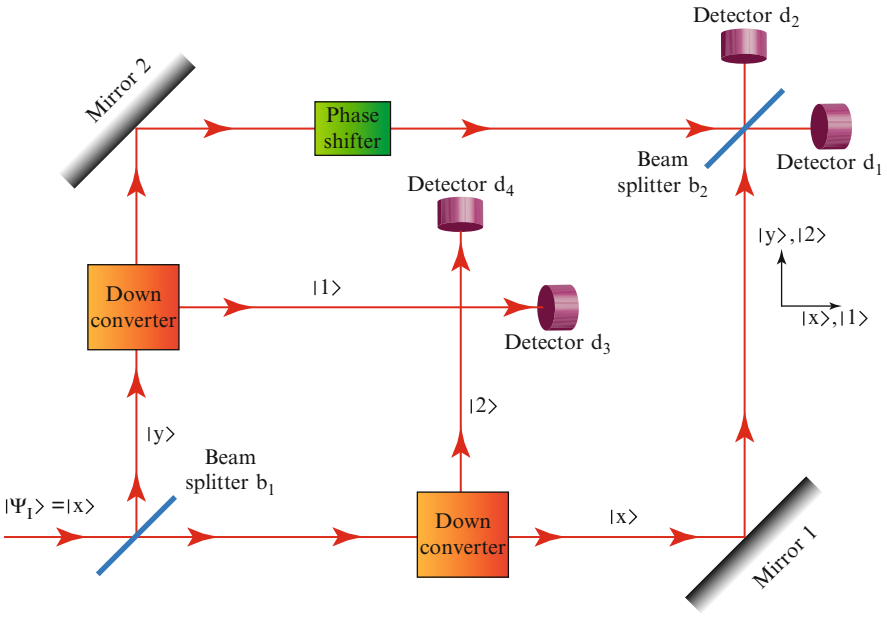


Fig. 8.10 Quantum eraser: the down-converted photon yields information on the path taken by the photon (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Let $|x\rangle, |y\rangle$ denote the photons traveling in the x, y-direction towards detectors d_1, d_2, respectively. A down-converted photon $|1\rangle$ traveling towards detector d_3 takes place only when there is a photon traveling towards mirror 2, with no down-converted photon traveling towards detector d_4. Similarly, a down-converted photon $|2\rangle$ exists only if a photon is traveling towards mirror 1.

A precise definition of the state vectors of the down-converted photons is the following:

- $|1\rangle$: *One* down-converted photon traveling in the x-direction towards detectors d_3; *no* down-converted photon traveling in the y-direction.
- $|2\rangle$: *One* down-converted photon traveling in the y-direction towards detectors d_4; *no* down-converted photon traveling in the x-direction.

With these definitions, detectors d_3 and d_4 are *projection operators* in the one- and two-direction and given by

$$d_3 = |1\rangle\langle 1|; \qquad d_4 = |2\rangle\langle 2| \tag{8.20}$$

The state vector yields the following realization of the quantum eraser:

$$\begin{aligned}
|\psi_I\rangle &= |x\rangle \\
&\to \text{Beam splitter } b_1 \to \frac{1}{\sqrt{2}}[|x\rangle + |y\rangle] \\
&\to \text{Down converter} \to \frac{1}{\sqrt{2}}[|x\rangle|2\rangle + |y\rangle|1\rangle] \\
&\to \text{Mirrors and phase shift} \to \frac{1}{\sqrt{2}}[|y\rangle|2\rangle + \mathrm{e}^{\mathrm{i}\phi}|x\rangle|1\rangle] \\
&\to \text{Beam splitter } b_2 \to \frac{1}{2}\left([|y\rangle - |x\rangle]|2\rangle + \mathrm{e}^{\mathrm{i}\phi}[|x\rangle + |y\rangle]|1\rangle]\right) \\
&= |\psi_N\rangle
\end{aligned}$$

One can perform a *coincident measurement* by simultaneously recording the photons received in detectors d_1, d_2 and d_3, d_4. Consider the pure state density matrix

$$\rho = |\psi_N\rangle\langle\psi_N| \tag{8.21}$$

The density matrix given in (8.21) has encoded in it the down conversion of the photon.

The information of the path taken by the photon is known by the experiment, regardless of whether the information is recorded or not. To know whether the original interference pattern of the single photon considered in Sect. 8.5—obtained without down conversion—is still valid or not, one needs to only record the result obtained in detectors d_1, d_2, ignoring the results recorded by detectors d_3, d_4. This, in effect, requires that a partial trace be taken of ρ over the states detected by detectors

d_3, d_4; performing a partial trace over the detector states $|1\rangle, |2\rangle$ using (6.5) yields the following:

$$\rho_{N,\mathrm{R}} = \mathrm{tr}_{|1\rangle,|2\rangle}\left(|\psi_N\rangle\langle\psi_N|\right);\ \ \langle i|j\rangle = \delta_{i-j}\ ;\ i,j = 1,2$$

$$\Rightarrow \rho_{N,\mathrm{R}} = \frac{1}{2}\left(|x\rangle\langle x| + |y\rangle\langle y|\right);\ \ \mathrm{tr}\,\rho_{N,\mathrm{R}}^2 = \frac{1}{2} < 1$$

$$P_1 = \mathrm{tr}\left(\rho_{N,\mathrm{R}} d_1\right) = \mathrm{tr}\left(\rho_{N,\mathrm{R}}|x\rangle\langle x|\right) = \frac{1}{2}$$

$$P_2 = \mathrm{tr}\left(\rho_{N,\mathrm{R}} d_2\right) = \mathrm{tr}\left(\rho_{N,\mathrm{R}}|y\rangle\langle y|\right) = \frac{1}{2}$$

The reduced density matrix $\rho_{N,\mathrm{R}}$ is mixed, since information contained in detectors d_3, d_4 has been traced over, and shows that there is *no interference*—since the probability for finding a photon for both detectors is equal to 1/2, regardless of the relative phase angle ϕ.

8.10 Restoring Interference

Consider placing a *third beam* splitter b_3 between detectors d_3 and d_4, as shown in Fig. 8.11.

As was the case for the erasure of interference, the photon from the source, *after* going through beam splitter b_1 but *before* reflecting off either mirror 1 or mirror 2, goes through the down converter, where a nonlinear process *splits* the *single* photon into *two* photons. One of the resulting photons continues along the original path, and the other photon heads in an orthogonal direction, as shown in Fig. 8.11, with detector d_3 or d_4 detecting the down-converted photon.

After down conversion, the Mach–Zehnder interferometer has *two photons* in the apparatus, unlike the case discussed in Sect. 8.5 and similar to that discussed in Sect. 8.9. Since we now have yet another beam splitter b_3, the down-converted photon must go through b_3, as shown in Fig. 8.11, before the photon is detected by *either* detector d_3 *or* d_4.

After the beam splitter b_3 is put in place, it will be seen that interference is restored. Detecting a photon in either detector d_3 or d_4 no longer yields the information on which path was taken by the photons that are detected by d_1 or d_2; however, the interference now involves a coincident measurement of *two* photons.

The down-converted photon that is received at d_3 or at d_4 could have come from *either* of the two paths taken by the photon. Hence, detecting the down-converted photon no longer yields the which-path information that led to the absence of interference for a photon detected by detectors d_1 and d_2.

Hence, *interference* is *restored* since the path taken's knowledge has been *erased* by beam splitter b_3. However, as will be seen below, the measurements that yield interference are quite different than the case for the interference of a single photon.

To see how interference is restored, consider the following analysis:

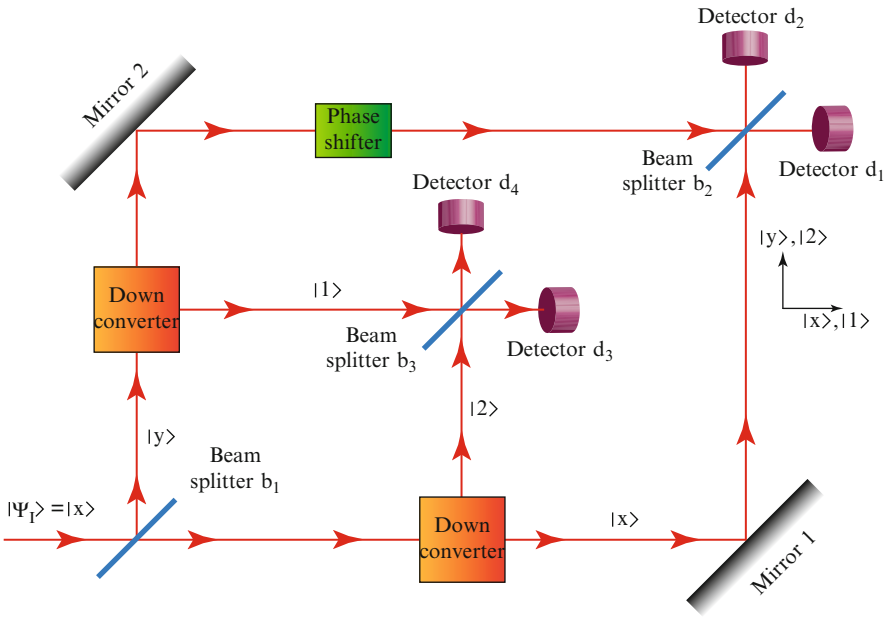


Fig. 8.11 Quantum eraser with beam splitter b_3: a coincident measurement of the down-converted photon yields and the original photon yields interference (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

$$
\begin{aligned}
&|\psi_1\rangle = |x\rangle \\
&\rightarrow \text{Beam splitter } b_1 \rightarrow \frac{1}{\sqrt{2}}[|x\rangle + |y\rangle] \\
&\rightarrow \text{Down converter} \rightarrow \frac{1}{\sqrt{2}}[|x\rangle|2\rangle + |y\rangle|1\rangle] \\
&\rightarrow \text{Beam splitter } b_3 \rightarrow \frac{1}{2}\left(|x\rangle[-|1\rangle + |2\rangle] + |y\rangle[|1\rangle + |2\rangle]\right) \\
&\rightarrow \text{Mirrors and phase shift} \rightarrow \frac{1}{2}\left(|y\rangle[-|1\rangle + |2\rangle] + \mathrm{e}^{\mathrm{i}\phi}|x\rangle[|1\rangle + |2\rangle]\right) \\
&\rightarrow \text{Beam splitter } b_2 \rightarrow \frac{1}{2\sqrt{2}}\left([|y\rangle - |x\rangle][-|1\rangle + |2\rangle] + \mathrm{e}^{\mathrm{i}\phi}[|x\rangle + |y\rangle][|1\rangle + |2\rangle]\right) \\
&= \frac{1}{2\sqrt{2}}\left[(1+\mathrm{e}^{\mathrm{i}\phi})|x\rangle|1\rangle + (1+\mathrm{e}^{\mathrm{i}\phi})|y\rangle|2\rangle - (1-\mathrm{e}^{\mathrm{i}\phi})|x\rangle|2\rangle - (1-\mathrm{e}^{\mathrm{i}\phi})|y\rangle|1\rangle\right] \\
&= |\psi_F\rangle
\end{aligned}
\tag{8.22}
$$

Coincident measurements consist of measuring two detectors simultaneously; for the quantum eraser, there are four possible coincident measurements. Detectors

d_1, d_2 given in (8.13) are measured simultaneously with detectors d_3, d_4 given in (8.20) and correspond to readings by the following combination of detectors, given below together with their likelihood of occurrence:

$$\begin{aligned} P_{x1}: &\quad d_1 \otimes d_3 = |x\rangle\langle x| \otimes |1\rangle\langle 1|; \qquad P_{y1}: \quad d_2 \otimes d_3 = |y\rangle\langle y| \otimes |1\rangle\langle 1| \\ P_{x2}: &\quad d_1 \otimes d_4 = |x\rangle\langle x| \otimes |2\rangle\langle 2|; \qquad P_{y2}: \quad d_2 \otimes d_4 = |y\rangle\langle y| \otimes |2\rangle\langle 2| \end{aligned}$$

All the computations have the following generic derivation; let $\rho = |\psi_F\rangle\langle\psi_F|$; then, from (8.22)

$$\begin{aligned} P_{x1} &= \mathrm{tr}\left[\left(d_1 \otimes d_3\right)\rho\right] \\ &= \frac{1}{8}\left|\left(1 + \mathrm{e}^{\mathrm{i}\phi}\right)\right|^2 \mathrm{tr}\left[\left(d_1 \otimes d_3\right)\left(|x\rangle\langle x| \otimes |1\rangle\langle 1|\right)\right] \\ &= \frac{1}{2}\cos^2\left(\frac{\phi}{2}\right)\left|\langle x|x\rangle\langle 1|1\rangle\right|^2 = \frac{1}{2}\cos^2\left(\frac{\phi}{2}\right) \end{aligned}$$

Hence, repeating the derivation for other cases leads to the following result:

$$\begin{aligned} P_{x1} &= \frac{1}{2}\cos^2\left(\frac{\phi}{2}\right) = P_{y2}; \qquad P_{x2} = \frac{1}{2}\sin^2\left(\frac{\phi}{2}\right) = P_{y1} \\ &\Rightarrow P_{x1} + P_{x2} + P_{y2} + P_{y1} = 1 \end{aligned} \tag{8.23}$$

Equation (8.23) states the following: It is certain that one photon will be found by either detector d_1 or d_2 *and* another photon will certainly be detected by either detector d_3 or d_4.

If one does a coincidence measurement by counting only those cases when photons are simultaneously detected by *both* detectors d_1 and d_3 (one photon in each detector)—or equivalently, by *both* detectors d_2 and d_4—then, as one varies the relative phase ϕ, one obtains an interference pattern from P_{x1}, P_{x2}—or P_{y2}, P_{y1}—which is similar to one given in (8.17) and (8.18).

One may object that in the case of the quantum eraser, one has shifted one's attention from single photon interference to the case of the interference involving two photons. This shift is unavoidable since the path of the initial photon is being probed with another (down-converted) photon, and, hence, both photons have to be observed to obtain the path information.

The experimental arrangement of the quantum eraser is not simply to detect whether the initial photon exists in an empirical or in a trans-empirical state; in fact, after down conversion, there are now two photons in the apparatus, and they are jointly in an empirical or a trans-empirical state depending on the measurements being carried out. The fact that interference exists for the coincidence counts shows that it is the *two-photon* system that is now in the trans-empirical state and which gives rise to the interference.

To see that the interference does not exist for the single photon detected by either detector d_1 or d_2, consider the density matrix formed from the state vector given in (8.22), namely,

$$\rho = |\psi_F\rangle\langle\psi_F|$$

Performing a partial trace over the states $|1\rangle, |2\rangle$—namely ignoring the readings of the detectors d_3 and d_4—using (6.5) yields, after some algebra, the following the reduced matrix

$$\rho_R = \text{tr}_{|1\rangle,|2\rangle}\left(|\psi_F\rangle\langle\psi_F|\right) = \frac{1}{2}\left(|x\rangle\langle x| + |y\rangle\langle y|\right)$$

and which, from the earlier discussion, does not have any interference.

In other words, if one were to only measure a single photon using detectors d_1 and d_2, then one would completely miss the interference effect.

8.11 Partial Quantum Eraser

For the case of both the two-slit experiment, the Mach–Zehnder interferometer and the quantum eraser, one either had *full interference* or *no interference*. We now examine the case for which the path information in only partially erased.

The Mach–Zehnder interferometer is used for setting up an experiment that smoothly interpolates between having full knowledge of the path taken—and thus destroying quantum interference, to the case of having no knowledge of the path taken—and thus fully restoring interference.

For the case of partial erasure, the Mach–Zehnder interferometer has two beam splitters b_1 and b_2 and two detectors d_1 and d_2, with a relative phase shifter ϕ. The general idea is to couple a *measuring device*—which consists of another quantum system, say a spin variable $|s\rangle$—to the photon's path.

Measuring the state of the spin $|s\rangle$ determines whether the photon has taken a particular path. One can gradually turn on the spin-photon coupling (interaction) and examine the influence of the coupling on the interference pattern. The experimental arrangement is shown in Fig. 8.12.

The spin $|s\rangle$ has an interaction U with the photon. If the photon travels on the path that reflects off mirror 2, then it interacts with $|s\rangle$—and causes $|s\rangle$ to undergo a transition to state $|s'\rangle$—given by the following:

$$|s\rangle \to |s'\rangle = U|s\rangle$$

$$\langle s'|s\rangle = \langle s|U^\dagger|s\rangle \equiv |\alpha|\exp(i\chi); \quad |\alpha| \in [0,1]$$

The parameter α measures the strength of the coupling of the spin to the photon. There is a similar coupling that one can introduce in the quantum eraser. If beam

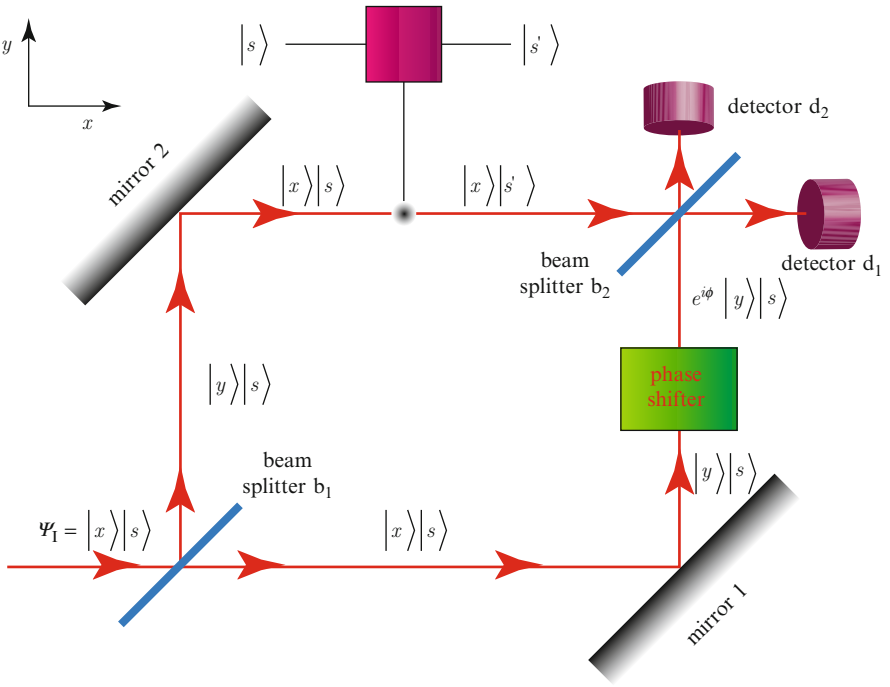


Fig. 8.12 Partial quantum erasure of the path information (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

splitter b_3, shown in Fig. 8.11, transmits the photon with only some finite likelihood, then the parameter $|\alpha|$ is a measure of the degree of transparency of beam splitter b_3.

The value of $|\alpha| = 0$ corresponds to the beam splitter b_2 having no effect on the photon and $|\alpha| = 1$ corresponding to the full action of the beam splitter, as given in (8.14). It can be shown that quantum eraser shown in Fig. 8.11 is equivalent to the case of partial quantum superposition discussed in this Sect. 8.11.

The initial state of the system $|\psi_I\rangle$ is equal to the tensor product $|x\rangle|s\rangle$. The initial state evolves in the following manner to the final state $|\psi_F\rangle$:

$$
\begin{aligned}
|\psi_I\rangle &= |x\rangle|s\rangle \\
&\rightarrow (\text{beam splitter } b_1) \rightarrow \frac{1}{\sqrt{2}}\Big[|x\rangle + |y\rangle\Big]|s\rangle \\
&\rightarrow \text{Mirrors} \qquad \rightarrow \frac{1}{\sqrt{2}}\Big[|y\rangle + |x\rangle\Big]|s\rangle \\
&\rightarrow \text{Phase shifter} \rightarrow \frac{1}{\sqrt{2}}\Big[e^{i\phi}|y\rangle + |x\rangle\Big]|s\rangle
\end{aligned}
$$

$$\rightarrow \text{Interaction} \rightarrow \frac{1}{\sqrt{2}}\left(\mathrm{e}^{\mathrm{i}\phi}|y\rangle|s\rangle + |x\rangle|s'\rangle\right) \tag{8.24}$$

$$\begin{aligned}
&\rightarrow \text{Beam splitter } b_2 \rightarrow \frac{1}{2}\left(\mathrm{e}^{\mathrm{i}\phi}\{|y\rangle - |x\rangle\}|s\rangle + \{|x\rangle + |y\rangle\}|s'\rangle\right)\\
&\equiv |a\rangle|s\rangle + |b\rangle|s'\rangle\\
&= |\psi_{\mathrm{F}}\rangle
\end{aligned} \tag{8.25}$$

where

$$|a\rangle \equiv \frac{1}{2}\mathrm{e}^{\mathrm{i}\phi}(|y\rangle - |x\rangle); \qquad |b\rangle \equiv \frac{1}{2}(|x\rangle + |y\rangle) \tag{8.26}$$

Note the formation of an *entangled state* given in (8.24), which is the result of the photon interacting with the spin degree of freedom.

The process of the interaction of the spin degree of freedom $|s\rangle$ with the device represented by U is irreversible; for example, if the interaction leads to a state $|s'\rangle = U|s\rangle$ such that $\langle s|s'\rangle = 0$, then one can be certain that the photon was reflected off mirror 2; in this case, the information extracted by the spin degree of freedom is reflected in the absence of interference for the photon—even if only the photon's state is recorded.

The density matrix for the final state is given by

$$\begin{aligned}
\rho &= |\psi_{\mathrm{F}}\rangle\langle\psi_{\mathrm{F}}|\\
&= |a\rangle\langle a| \otimes |s\rangle\langle s| + |a\rangle\langle b| \otimes |s\rangle\langle s'| + |b\rangle\langle a| \otimes |s'\rangle\langle s| + |b\rangle\langle b| \otimes |s'\rangle\langle s'|
\end{aligned}$$

that encodes the information for an experiment in which *both* the photon and spin degree of freedom's state are measured.

The reduced density matrix ρ_{R}, obtained by tracing over the spin degree of freedom, has less information, namely, what is required for an experiment in which the state of *only* the photon degree of freedom is measured, with no measurement being performed to ascertain the state of the spin degree of freedom. Since we are only interested in studying whether there is interference for the photon, consider the experiment in which only the final state of photon is recorded at detectors d_1 and d_2. The result of the experiment is encoded in the reduced density matrix ρ_{R} for which the spin degrees of freedom have been traced over and given by

$$\begin{aligned}
\rho_{\mathrm{R}} &= \mathrm{tr}_{s,s'}(\rho) = \mathrm{tr}_{s,s'}\left(|\psi_{\mathrm{F}}\rangle\langle\psi_{\mathrm{F}}|\right)\\
&= |a\rangle\langle a| + |\alpha|\left(\mathrm{e}^{\mathrm{i}\chi}|a\rangle\langle b| + \mathrm{e}^{-i\chi}|b\rangle\langle a|\right) + |b\rangle\langle b|
\end{aligned}$$

since

$$\langle s|s\rangle| = 1 = \langle s'|s'\rangle; \qquad \langle s'|s\rangle = |\alpha|\mathrm{e}^{\mathrm{i}\chi}$$

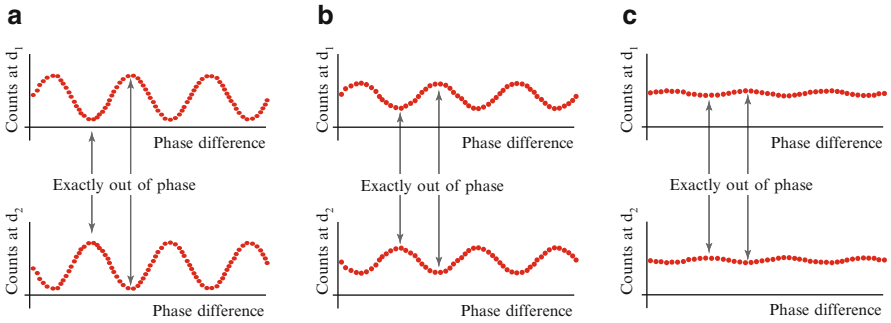


Fig. 8.13 (**a**) The case of $|\alpha| = 1$ that yields full interference; experiment has no information about the path that is taken by the photon. (**b**) Case of $0 < |\alpha| < 1$; experiment has only partial information about which path is taken by the photon—leading to partial interference. (**c**) No interference for the case of $|\alpha| = 0$, since the path is taken by the photon, is experimentally known with 100 % certainty (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Consider measurements carried out by detectors d_1 and d_2, with P_1, P_2 being the likelihood that a photon is detected by d_1 and d_2, respectively. Then, setting $\chi = 0$ without any loss of generality, we have[3]

$$\begin{aligned} P_1 &= \mathrm{tr}(d_1 \rho_R) = \mathrm{tr}(|x\rangle\langle x|\rho_R) = \langle x|\rho_R|x\rangle \\ &= |\langle x|a\rangle|^2 + |\alpha|\Big(\langle x|a\rangle\langle b|x\rangle + \langle a|x\rangle\langle x|b\rangle\Big) + |\langle x|b\rangle|^2 \\ &= \frac{1}{2}\big[1 - |\alpha|\cos(\phi)\big] \end{aligned} \tag{8.27}$$

where (8.26) has been used to obtain (8.27). Similarly

$$P_2 = \mathrm{tr}(d_2 \rho_R) = \mathrm{tr}(|y\rangle\langle y|\rho_R) = \langle y|\rho_R|y\rangle = \frac{1}{2}\big[1 + |\alpha|\cos(\phi)\big] \tag{8.28}$$

$$P_1 + P_2 = 1$$

The result obtained in (8.27) and (8.28) shows that the constant $|\alpha|$ smoothly interpolates between interference and no interference. The result is illustrated in Fig. 8.13—with the case of full interference given by $|\alpha| = 1$ gradually going into the case of no interference with $|\alpha| = 0$.

- $|\alpha| = 0$. No interference

[3]The equations below are obtained by displacing $\phi \to \phi - \chi$; the result is independent of χ since it simply redefines the origin of ϕ, which in the first place is arbitrary.

Since $\langle s|s'\rangle = \langle s|U|s\rangle = 0$, the final and initial states are different with 100 % certainty.[4] Hence, the photon's path is fully known. As expected

$$|\alpha| = 0; \qquad \langle s|s'\rangle = 0$$
$$P_1 \to \frac{1}{2}; \qquad P_2 \to \frac{1}{2}$$
$$P_1 + P_2 = 1$$

- $|\alpha| = 1$. Full interference
  The case of $|\alpha| = 1$ yields that $|s'\rangle = |s\rangle$ implying that there is no interaction of the photon with the spin—leading to the complete loss of knowledge about which path was taken. As expected

$$|\alpha| = 1; \qquad |s'\rangle = |s\rangle$$
$$P_1 \to \sin^2\frac{\phi}{2}; \qquad P_2 \to \cos^2\frac{\phi}{2}$$
$$P_1 + P_2 = 1$$

In summary, as one varies the strength of the interaction by varying the value of $|\alpha|$, one smoothly interpolates between the photon taking trans-empirical paths (interference) to the limit of the photon taking an empirical path (no interference).

- For the case when $0 < |\alpha| < 1$ the photon exists in a state that is a *mixture* of indeterminate and determinate paths corresponding to a partial erasure of information. We conclude that, for $0 < |\alpha| < 1$, the quantum system exists in a state that is partly empirical and partly trans-empirical, as expressed in (8.27) and (8.28).

For the general case $|\alpha|$ given in (8.27) and (8.28), the quantum entity is partly empirical and partly trans-empirical. This result illustrates that the detailed structure of the transitional domain between the empirical and trans-empirical domains, shown as the *gray area* and labeled as "Measurement" in Fig. 3.4, depends on the accuracy and other details of the experiment being used to study the quantum entity.

8.12 Summary

Quantum mechanics is based on the trans-empirical quantum principle that a quantum entity has an indeterminate and trans-empirical form that makes a transition to its empirical form when the quantum entity is observed. Superposition in quantum

[4] In quantum mechanics, two states are distinguishable with 100 % confidence if the states are orthogonal, that is, if $\langle s|s'\rangle = \langle s|U|s\rangle = 0$. See discussion in Sect. 5.3.

mechanics yields the results of classical probability when empirical paths are composed and nonclassical results when indeterminate paths are superposed.

> Quantum superposition adds the probability amplitudes for indeterminate and trans-empirical paths. All the paradoxes of quantum superposition reflect the trans-empirical nature of the indeterminate paths.

The superposition of the state vector and probability amplitudes were analyzed to show that the role of measurement is central to the workings of quantum superposition. For the case of the superposed state vector, measurements reveal its constituent state vectors and their respective eigenvalues.

A detailed analysis was carried out of the physics of determinate and indeterminate paths, with quantum superposition of indeterminate paths being the realization of the linearity of quantum mechanics for probability amplitudes. It was shown that the probability amplitude has two forms depending on whether the quantum entity takes an empirical path or trans-empirical paths. Performing or not performing a measurement determines whether the path taken is empirical or trans-empirical, respectively.

The Mach–Zehnder interferometer was used to analyze cases where the path taken by a photon traveling through the interferometer was *either* not observed *or* observed—leading to interference or the lack of it. Interference for a quantum particle is a nonclassical result, and the lack of interference is the expected classical behavior. Partial and imprecise measurements were studied to show how interactions can be used to obtain information on which path was taken by a photon. And finally, the quantum eraser was discussed to show how information can be preserved or erased even after information on the path taken has apparently been generated in one segment of the apparatus.

Quantum superposition is nonclassical due to the trans-empirical nature of the state vector and the probability amplitude. Unlike symbols of classical physics such as position and momentum that can be directly related to measurable quantities, the physical content of the symbols for the trans-empirical quantities of quantum mechanics can never be directly observed. For example, it is the mathematical expression for the probability amplitude of a quantum entity "taking" indeterminate paths that yields predictions for the observable consequences of quantum superposition.

One of the conclusions from our study of quantum superposition is that it is the *analysis* of the symbols of quantum mechanics which leads us to infer that the trans-empirical form of the quantum entity does indeed exist.

Quantum Theory of Measurement

9

> A careful analysis of the process of observation in atomic physics has shown that the subatomic particles have no meaning as isolated entities, but can only be understood as interconnections between the preparation of an experiment and the subsequent measurement. Erwin Schrödinger.

A quantum entity, as discussed in Sect. 2.12, consists of an inseparable pair, namely, its degree of freedom and state vector. One of the most enigmatic features of a quantum entity is that its degree of freedom can *never* be directly observed. On attempting to observe its degree of freedom, what one ends up observing is the state vector of the degree of freedom, as illustrated in Fig. 3.3. In fact, as discussed in Sect. 7.9, one does not even observe the state vector; every experiment ultimately observes only the *effect* of the state vector on the projection operators, which are physical detectors. Quantum probability assigns probabilities to the likelihood of a projection operator detecting the state vector.

The role of measurement in experimentally observing the properties of a quantum entity is its fundamental difference from a classical thing—since it is the exceptional role of measurement in quantum mechanics that gives meaning to the concept and existence of the state vector.

This chapter studies exactly what do we mean by detecting a state vector and precisely what constitutes a projection operator.

In classical physics, Nature is considered to exist objectively and to be in an exact and determinate state, and whether one observes it or not does not matter. In contrast, the quantum entity's observed empirical form of existence is fundamentally different and distinct from its unobservable, indeterminate, and trans-empirical form. Measurement is the connecting link between the empirical and trans-empirical forms of a quantum state vector, as shown in Fig. 3.4.

In classical physics, all the changes caused by experimentally observing an object can, in principle, be accounted for exactly; in particular, all experimental error can, in principle, be reduced to zero. In quantum mechanics, experimental precision is intrinsically limited due to the nonzero value of $\hbar$ and reflects the indeterminate characteristic of the degree of freedom that yields uncontrollable and unpredictable

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8_9, © Springer Science+Business Media New York 2013

results for the measurement; the Heisenberg Uncertainty principle encodes this aspect of quantum measurements.

The role of measurement is central to the interpretation of quantum mechanics. An experimental observation carried out on the state vector $\psi(\mathcal{F})$ of the degree of freedom $\mathcal{F}$ *causes* a *discontinuous transition* from the trans-empirical form of the state vector to its empirical manifestation.

- The measurement process is mathematically represented by applying the operator $\mathcal{O}(\mathcal{F})$ on the trans-empirical state $|\psi(\mathcal{F})\rangle$ to yield another trans-empirical state $\mathcal{O}(\mathcal{F})|\psi(\mathcal{F})\rangle$.
- The act of measurement causes a collapse—a discontinuous and irreversible change—of the trans-empirical state $\mathcal{O}(\mathcal{F})|\psi(\mathcal{F})\rangle$ to an empirical reading of a measuring device.
- Repeated measurements yield the average value of the operator for the given quantum state, namely, $E_\psi[\mathcal{O}(\mathcal{F})] = \langle\psi|\mathcal{O}(\mathcal{F})|\psi\rangle$.

Note the discontinuous change caused by the physical act of measurement cannot be produced by the Schrödinger equation, which evolves the state vector $|\psi\rangle$ by a continuous unitary evolution. Hence, a consistent interpretation of quantum mechanics requires *both* the Schrödinger equation's unitary evolution *as well as* a non-unitary discontinuous transition.

All experiments and experimental detectors used for studying quantum phenomena have the following characteristics:

- The repeated and independent *preparations* of the same quantum state, which is subsequently subjected to repeated measurements.
- The detector greatly *amplifies* the quantum quantity being studied so that its value can be inferred from macroscopic quantities that are observable in the macroscopic world. Amplification is required when studying the quantum realm.
- The process of measurement *entangles* the quantum system's degrees of freedom with the degrees of freedom of the detector, creating a joint entangled state vector of the quantum system and detector.
- The act of measurement culminates in the quantum state *collapsing* to one of the states of the detector—in other words, causing a *discontinuous* change in the state vector—and in doing so brings about an *irreversible* change in the quantum system.
- Due to the entanglement of the quantum state and the detector, the final (macroscopic) state of the *detector*—the result of the observation—allows us to conclude with certainty on the value of the microscopic *quantum* quantity that is being observed.

In summary, all quantum measurements require four ingredients, namely, *preparation, amplification, entanglement*, and *collapse (irreversibility)*. This is the view of experiment that is practically followed by the majority of physicists. However, there are physicists who consider this interpretation of quantum mechanics as being incomplete and inadequate. Various theories have been proposed to resolve the problem of measurement and are discussed in Sect. 9.11.

To clarify the concept of measurement in quantum mechanics, the following diverse aspects of the process of measurement are discussed:

- A photographic plate's exposure as a form of quantum measurement.
- Measuring the expectation value of an operator for a state vector.
- The experimental device.
- The process of measurement and the preparation of a quantum state.
- The measurement process, in particular the collapse of the state vector, in terms of the density matrix.
- The Heisenberg Uncertainty Principle is motivated by analyzing the measurement of two non-commuting observables.

Noteworthy 9.1: Macroscopic quantum systems

There are macroscopic quantum entities such as a superconductor or a superfluid for which the amplification needed for making a quantum measurement is not necessary. A superconductor has zero resistance for a D.C. current, which is a macroscopic classical quantity; similarly, viscosity is a macroscopic property of classical fluids, which is zero for a superfluid—and hence there is no need for any amplification. However, these macroscopic quantum systems tend to be very special.

Even for macroscopic quantum systems, amplification is required for observing the microscopic and quantum basis for their macroscopic features. At low enough temperatures, in some metals pairs of electrons form a bound state—which are microscopic in nature—and that condense to form a superconductor; similarly, magnetic flux quantization of magnetic field trapped in a superfluid needs amplification to be observed.

9.1 Measurement: Trans-Empirical to Empirical

Consider the coordinate degree of freedom $x \in \Re = \mathcal{F}$. Measurement plays an indispensable role in quantum mechanics since it connects the trans-empirical form of the quantum state, as described by the state vector $|\psi(t)\rangle$, with the expectation value of the position projection operator $|x\rangle\langle x|$ given by $|\psi(t,x)|^2$.

The nonclassical nature of $|\psi(t)\rangle$ can be seen from the following fact: The moment the quantum particle is observed (measured) by the position projection operator $|x\rangle\langle x|$, the state vector of the quantum particle instantaneously "collapses," that is, becomes zero *everywhere in space*—except where it is observed—since once the particle is detected by operator $|x\rangle\langle x|$, there is zero likelihood of a projection operator at any other point finding it.

No classical wave collapses on being observed driving home the point that the state vector $|\psi(t)\rangle$ is not like any ordinary classical wave. A classical wave, like the waves on an ocean, exists objectively and remains in its state whether one observes it or not.

Some of the new concepts that have been introduced by quantum mechanics are the following:

- The quantum state embodies the exhaustive and complete description of a quantum entity and, in particular, describes the trans-empirical state vector as well as the empirical manifestation of the quantum entity.
- Once a quantum state is prepared and no longer observed, it makes a transition from the empirical to the trans-empirical state. In contrast, measurement causes a discontinuous transition of the quantum state from the trans-empirical to the empirical domain.

9.2 Position Projection Operator

The process of measurement is illustrated in Fig. 9.1. An incoming particle—represented by a state vector $|\psi\rangle$, with state function given by $\psi(x) = \langle x|\psi\rangle$—is shown in Fig. 9.1a to be spread out to represent the fact that the particle has a finite probability of being detected by the projector operators located at many different points of space.

The photographic plate represents a typical detector carrying out a position measurement; *every* grain of the photographic plate at position x represents a projection operator $|x\rangle\langle x|$. On hitting the photographic plate, the particle deposits all of its energy at a single point x (or more precisely, in a volume having a dimension much, much smaller than the spread of its state function), as shown in Fig. 9.1b.

It is concluded that the process of measurement has resulted in the particle's state vector being *observed* by the projection operator $|x\rangle\langle x|$ at point x.

Figure 9.2 shows the empirical form $|\psi(x_0, y_0)|^2$ and the trans-empirical form $\psi(x,y)$ of the state function for a particle being observed by the photographic plate. The empirical manifestation $|\psi(x_0, y_0)|^2$ of the quantum state function is nonzero *only* at a given position x_0, y_0; the trans-empirical and indeterminate form $\psi(x,y)$ of the same state function is nonzero for *all* x, y.

The quantum concept of repeated measurements is discussed in Sect. 9.3. In the case of a particle being detected at the photographic plate, of course once it has been detected, the particle is "lost." For this case, one repeats the experiment using identical particles and with identical preparations.

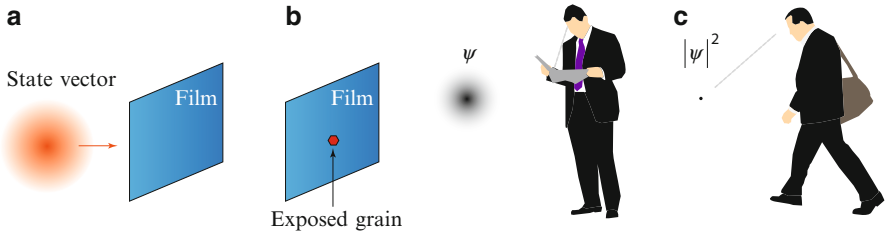


Fig. 9.1 (**a**) A quantum particle approaches a photographic plate. (**b**) The state vector collapses at the position where the grain is exposed. (**c**) An observer not observing and observing the particle (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

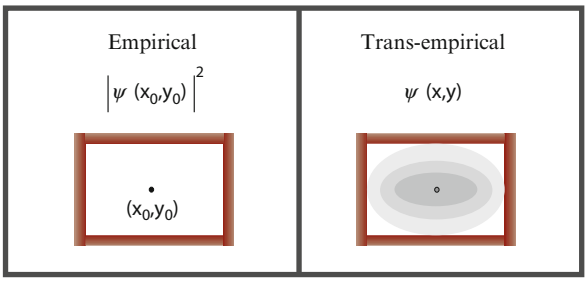


Fig. 9.2 Representation of the photographic plate: the empirical manifestation $|\psi(x_0, y_0)|^2$ of the quantum state function and its trans-empirical and indeterminate form $\psi(x, y)$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

a

Trans-empirical	Transition		Measure-ment	Empirical
$x \in \mathcal{F}=\Re$	$\psi(t,x) \in \mathcal{V}(\Re)$	$\frac{\partial \psi(t,x)}{\partial t}$	$\hat{O}(x, \frac{\partial}{\partial x})$	$E_{\psi}[\hat{O}] = \langle \psi \vert O \vert \psi \rangle$

b

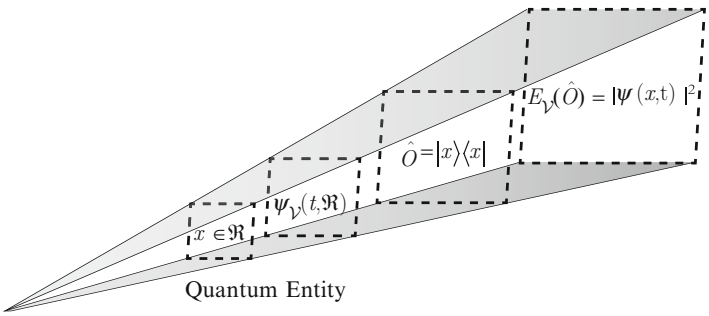


Fig. 9.3 (**a**) The theoretical superstructure of quantum mechanics for a continuous degree of freedom $\mathcal{F}$. (**b**) The superstructure of a quantum particle that is being measured by the position projection operator $|x\rangle\langle x|$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

If the same experiment is now *repeated* many, many times, it will be found that the average value of the position projection operator $|x\rangle\langle x|$ at different points x is given by $\mathrm{tr}(\rho|x\rangle\langle x|) = |\psi(x)|^2$, where $\rho = |\psi\rangle\langle\psi|$. The result $|\psi(x)|^2$ can also be interpreted as the *probability* that the operator $|x\rangle\langle x|$, located at position x on the photographic plate, is likely to detect the state vector.

There is another class of quantum measurements in which the *same* quantum entity can be repeatedly observed. Take the case of shining light on an atom and

observing the light it emits; once the emission is over, the atom returns to its original state, and one can again shine light on the atom—with the subsequent emitted light being observed again, and so on.

Quantum Mechanical Superstructure for Position Operator

The superstructure of the quantum entity has a concrete and transparent form when degree of freedom and the projection operators performing the measurements are specified. The various domains of the quantum entity, as given in Figs. 2.4 and 2.5, are expressed in a more concrete manner, shown in Fig. 9.3a and b, using the concepts and notation developed so far.

For the case of the position projection operator $|x\rangle\langle x|$, as shown in Fig. 9.3b, the degree of freedom is the continuous Euclidean space $\mathcal{F} = \Re$ with $x \in \Re$. For the continuous degree of freedom, the Hilbert space $\mathcal{V}(\mathcal{F})$ consists normalized state vectors that are functions of the coordinate x, namely, $\psi(t,x)$ with $\int \mathrm{d}x|\psi(t,x)|^2 = 1$.

The observables are Hermitian differential operators $\mathcal{O} = O(x, \partial/\partial x)$ that act on the state function $\psi(t,x)$. The expectation value of the operator for the quantum state is what is empirically measured and is a diagonal matrix element given by

$$E_\psi[\mathcal{O}] = \langle\psi|\mathcal{O}|\psi\rangle = \int \mathrm{d}x\, \psi^*(t,x) \left[O\left(x, \frac{\partial}{\partial x}\right) \psi(t,x) \right]$$

For the position projection operator $|x\rangle\langle x|$, the probability of the state vector collapsing at $|x\rangle\langle x|$ is given by $E_\psi[|x\rangle\langle x|] = |\psi(t,x)|^2$.

9.3 Repeated Observations in Quantum Mechanics

In quantum mechanics, the term "repeated observations" does not mean that one successively observes the same quantum system, say a particle, many times. Instead, what it means is that one first experimentally prepares a quantum entity to be in a *particular* quantum state that is of interest; once the preparation is completed, the quantum entity is then experimentally observed. After the observation is over, the quantum entity is again experimentally prepared to be in exactly the *same particular* quantum state and then again experimentally observed—with the outcome not necessarily being same as the previous case.

Note that quantum mechanics generalizes the concept of causality. The preparation of the same particular quantum state leads to a multiplicity of outcomes when observed. In other words, the *same cause*, namely, the preparation of the state, leads on being observed to a *multiplicity of effects*.

This process of preparation and observation of the quantum entity needs to be carried out many many times to be able to observe all the possible values of the

operators in quantum state, together with the likelihood for the various values. Hence, in quantum mechanics, it is necessary to repeatedly observe a quantum entity to determine the *probability* of the different possible outcomes.

A crucial point to note is that each preparation of a particular quantum state must be *independent* of the other preparations. Furthermore, each experimental observation is *independent* of all other observations. In other words, each preparation and observation must be *statistically independent* of all the rest. The statistical independence ensures that the outcomes are not correlated and the result yields the probability distribution.

The process of repeated measurement in effect creates an ensemble of states, each having the possibility of yielding a different result on being observed. Each process of measurement collapses the state vector to a definite value of the observing projection operator. One can think of collecting all the identically prepared quantum states and carrying out all measurements simultaneously. The result of such a procedure, for infinitely many identically prepared states, yields all possible outcomes for the quantum state, of course with different likelihoods. In effect, the repeated observations yield a classical random ensemble, and the results of classical probability can be applied to the result of repeated measurements. Before the measurements are carried, what we have prepared is a quantum ensemble (collection) of state vectors; this ensemble has no relation to the classical random ensemble that consists of the different values of the random variable; in contrast, the quantum ensemble is a collection of *state vectors* with the degree of freedom remaining indeterminate and trans-empirical.

Repeated measurements make sense only in quantum mechanics, where each measurement can potentially give a new result. In classical physics, there is a rigid *one-to-one* correspondence between cause and effect, with one cause giving rise to one and only one effect. In classical physics—which includes Einstein's general theory of relativity—one would expect to obtain exactly the same result, within experimental errors, on observing states that were prepared in an identical manner. Hence, in classical physics there is no a priori reason for making repeated measurements of the same system.

9.4 Expectation Value of Projection Operators

The expectation value of operators has been discussed in Sects. 5.8 and 7.9. The main results are re-derived in the context of measurement theory.

Consider carrying out observations on a state vector $|\chi\rangle$ using the Hermitian operator $\mathcal{O}$. Since $\mathcal{O}$ is a Hermitian operator (observable), it has a complete set of eigenstates $|\psi_n\rangle$ with eigenvalues λ_n and yields the spectral resolution, from (5.8), given by

$$\mathcal{O} = \sum_{n=1}^{N} \lambda_n |\psi_n\rangle\langle\psi_n|; \qquad \mathcal{O}|\psi_n\rangle = \lambda_n |\psi_n\rangle \tag{9.1}$$

$$\Rightarrow \mathcal{O} = \sum_{n=1}^{N} \lambda_n \Pi_n; \;\; \Pi_n = |\psi_n\rangle\langle\psi_n|; \qquad \Pi_n^2 = \Pi_n \tag{9.2}$$

All eigenvalues λ_n are real because operator $\mathcal{O}$ is Hermitian.

Since eigenstates $|\psi_n\rangle$ are complete, the state vector $|\chi\rangle$ has an eigenfunction decomposition

$$|\chi\rangle = \sum_n c_n |\psi_n\rangle \tag{9.3}$$

The state $|\chi\rangle$ is repeatedly prepared, in an *identical manner*, and the preparation entails fixing the values of all the coefficients c_n in the decomposition given in (9.3). Every time the experiment is repeated, measurements are made *independently* of all the other measurements.

Every time the value of the operator $\mathcal{O}$ is observed for the state $|\chi\rangle$, the operator $\mathcal{O}$ is *always* found to have *one* of its eigenvalues λ_n (with probability $|c_n|^2$)—and never any other value. This is one of the salient features of quantum superposition, as discussed in Sect. 8.1, of which the expression of χ in terms of the eigenstates as in (9.3) is a particular case.

As discussed in Sect. 5.8, the quantum mechanical interpretation of $|c_n|^2$ is that it is the *probability* of finding the state $|\chi\rangle$ in the state specified by eigenstate $|\psi_n\rangle$, which has the eigenvalue λ_n.

In the formulation of quantum probability, as discussed in Sect. 7.9, the *fundamental postulate* of quantum mechanics made by Max Born can be restated as assigning *probabilities* to the *projection operators* Π_n. It is *postulated* that observing the operator $\mathcal{O}$ entails, as shown in Fig. 7.8, applying the projection operator Π_n on the quantum state vector and which yields, after repeated applications, the average value of the projection operator Π_n for state $|\chi\rangle$, given by the following:

$$|c_n|^2 = E_\chi[\Pi_n] = |\langle\psi_n|\chi\rangle|^2 = \mathrm{tr}(\rho\Pi_n); \quad \rho = |\chi\rangle\langle\chi| \tag{9.4}$$

The quantum probabilities p_n given in (7.33) are determined by the state vector $|\chi\rangle$ and given by

$$p_n = |c_n|^2 = E_\chi[\Pi_n]; \qquad \sum_n p_n = 1$$

The observed values of λ_n can be *different* each time the identically prepared states are observed. Carrying out observations many times results in finding the average value of $\mathcal{O}$, namely, $E[\mathcal{O}]$, for a state vector $|\chi\rangle$, which from (9.2) is given by

$$E_\chi[\mathcal{O}] = \sum_n \lambda_n E_\chi[\Pi_n] \tag{9.5}$$

where $E_\chi[\Pi_n]$ is the *expectation value* of Π_n for quantum state $|\chi\rangle$. Hence, from (9.4) and (9.5),

$$E_\chi[\mathcal{O}] = \sum_n \lambda_n |c_n|^2 = \sum_n \lambda_n |\langle \psi_n|\chi\rangle|^2 = \sum_n \lambda_n \langle \chi|\psi_n\rangle\langle\psi_n|\chi\rangle$$
$$= \langle \chi|\mathcal{O}|\chi\rangle \tag{9.6}$$

where (9.1) yields the final result. From above, it is clear that the diagonal value of the operator $\mathcal{O}$ for a given state, namely, $\langle\chi|\mathcal{O}|\chi\rangle$, is the average value of the eigenvalues λ_n of the operator $\mathcal{O}$ for the state vector $|\chi\rangle$.

Hence, in general the value of any physical quantity $\mathcal{O}$ for a given quantum state $|\chi\rangle$ is obtained by finding the diagonal matrix element of the operator, namely, $\langle\chi|\mathcal{O}|\chi\rangle$. From (9.6), the expectation value of the observable $\mathcal{O}$ is given by

$$\langle\chi|\mathcal{O}|\chi\rangle = \textstyle\sum_{i=1}^{N} |c_i|^2 \lambda_i = \mathrm{tr}\,(\mathcal{O}|\chi\rangle\langle\chi|) \tag{9.7}$$
$$\Rightarrow \langle\chi|\mathcal{O}|\chi\rangle = \mathrm{tr}\,(\rho\mathcal{O}); \quad \rho = |\chi\rangle\langle\chi| \tag{9.8}$$

The projection operators Π_n are orthonormal since

$$\Pi_n = |\psi_n\rangle\langle\psi_n|; \quad \Pi_n\Pi_m = \delta_{n-m}\Pi_n \tag{9.9}$$
$$\textstyle\sum_n \Pi_n = \mathbb{I} \tag{9.10}$$

where the last equation follows from the completeness equation given in (5.6).

On being observed by the operator $\mathcal{O}$, the state vector $|\chi\rangle$ "collapses" to $|\psi_n\rangle$ or, equivalently, is projected to state $|\psi_n\rangle$; from (9.4), the process of measurement yields the following:

$$\mathcal{O} \to |\chi\rangle \to |\langle\psi_n|\chi\rangle|^2 = \langle\chi|\Pi_n|\chi\rangle = E_\chi[\Pi_n]$$

Decoherence

From an operational point of view, the process of measurement applies the projection operator Π_n on the state $|\chi\rangle$ and projects it to the state vector $|\psi_n\rangle$; in symbols

$$|\chi\rangle \to \text{Measurement} \to \frac{\Pi_n|\chi\rangle}{\sqrt{\langle\chi|\Pi_n|\chi\rangle}} = \left[\frac{\langle\psi_n|\chi\rangle}{\sqrt{\langle\chi|\Pi_n|\chi\rangle}}\right]|\psi_n\rangle$$
$$= \mathrm{e}^{\mathrm{i}\phi_n}|\psi_n\rangle \tag{9.11}$$

where ϕ_n is a pure phase.

It was shown in (5.13) that $\mathrm{e}^{\mathrm{i}\phi_n}|\psi_n\rangle$ and $|\psi_n\rangle$ are parallel and thus are identical since they differ by only a constant phase; the measuring process given in (9.11) projects the state vector $|\chi\rangle$ to $|\psi_n\rangle$.

Hence, from (9.11), the process of measurement discontinuously projects the state $|\chi\rangle$ to state $|\psi_n\rangle$ as follows:

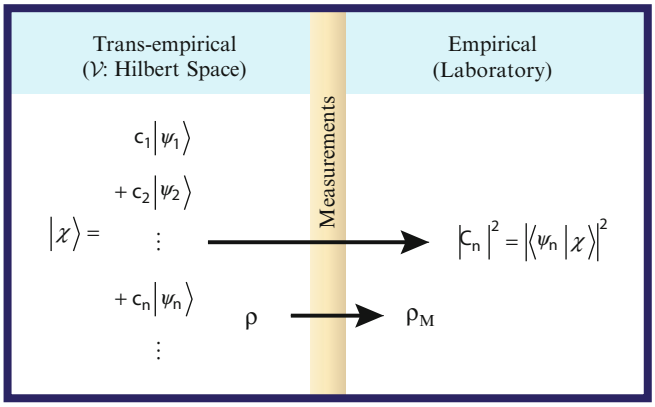


Fig. 9.4 Collapse of the state vector $|\psi\rangle = \sum_n c_n|\psi_n\rangle$ to $|c_n|^2 = |\langle\psi_n|\psi\rangle|^2$ and the collapse of pure density matrix ρ to the mixed density matrix ρ_M, which is equivalent to a classical random ensemble (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

$$|\chi\rangle \to e^{i\phi_n}|\psi_n\rangle$$

Every time a measurement is made, a *different* and random ϕ_n results from the process of measurement. To explain decoherence, von Neumann [36] *postulated* that repeated measurements result in the experiment taking the *average* over all the random ϕ_n's and yield the following result[1]:

$$\rho = \sum_{ij=1}^{N} c_i c_j^* |\psi_i\rangle\langle\psi_j| : \quad \text{coherent} \tag{9.12}$$

$$\to \text{Measurements} \to \prod_{n=1}^{N} \int_{-\pi}^{+\pi} \frac{d\phi_n}{2\pi} \sum_{ij=1}^{N} e^{i(\phi_i - \phi_j)} c_i c_j^* |\psi_i\rangle\langle\psi_j|$$

$$= \sum_{i=1}^{N} |c_i|^2 |\psi_i\rangle\langle\psi_i| = \rho_M : \quad \text{decoherence} \tag{9.13}$$

State vector $|\chi\rangle$ is a trans-empirical state and, as shown in Fig. 9.4, is a quantum superposition of the basis eigenstates $\{|\psi_n\rangle; n = 1, 2, \ldots, N\}$. The process of measurement collapses ρ to ρ_m and the entire state vector $|\chi\rangle$ to the state $|\psi_n\rangle$ with probability $|c_n|^2$, as shown in Fig. 9.4.

The collapse of the state vector and of the density matrix is illustrated in Fig. 9.4.

Decoherence is defined as the loss of information due to the cancellation of the off-diagonal terms in the sum $\sum_{ij=1}^{N} c_i c_j^* |\psi_i\rangle\langle\psi_j| = |\chi\rangle\langle\chi|$. The process of

[1] It should be noted that there is no experimental proof that von Neumann's postulate is correct.

measurement reduces the pure density matrix ρ to the *mixed density matrix* ρ_{M} [33]. The relation of the mixed density matrix to the process of measurement is discussed in Sect. 9.7.

9.5 The Experimental Device

To study the operator $\mathcal{O}$ from an experimental point of view, one has to construct a device that measures the eigenfunctions and eigenvalues of the operator $\mathcal{O}$. The device, in particular, can determine the likelihood of $|\chi\rangle$ being observed in a particular eigenstate of $\mathcal{O}$.

Every observable (Hermitian operator) $\mathcal{O}$ needs a *specific* experimental device to measure its properties. The detector (device) has a number of states $|D_n\rangle$ that can be coupled to the quantum system. The detector states $|D_n\rangle$ are *macroscopically distinguishable* configurations of the device and form a complete orthonormal basis of the detector's Hilbert space $\mathcal{V}_{\mathrm{D}}$.

$$\sum_{n=1}^{M} |D_n\rangle\langle D_n| = \mathbb{I}_{\mathrm{D}} \; ; \;\; \langle D_n|D_m\rangle = \delta_{n-m}$$

One can think of the device states $|D_n\rangle$ as the needle of a counter pointing at position x_n that is well separated from all other counterpositions x_i; $i \neq n$; in particular, one can represent the device states by sharply peaked Gaussian functions given by the following:

$$\langle x|D_n\rangle = \left(\frac{1}{\pi\sigma^2}\right)^{1/4} \exp\left\{-\frac{1}{\sigma^2}(x-x_n)^2\right\} \tag{9.14}$$

It can be shown that, for small enough σ^2 and large enough M, the detector states $|D_n\rangle$ yield the completeness equation

$$\sum_{n=1}^{M} \langle x|D_n\rangle\langle D_n|x'\rangle \simeq \delta(x-x')$$

The experimental apparatus is schematically represented in Fig. 9.5. Note the similarity of the experimental device with the schema for the device given in terms of projection operators in Fig. 7.8. In fact, each counter reading x_n corresponds to a projection operator Π_n of the operator $\mathcal{O}$—defined in (9.2)—since the experimental apparatus has been designed based on the correspondence $\Pi_n \leftrightarrow |D_n\rangle\langle D_n|$.

Let the state vector of the quantum system belong to a Hilbert space $\mathcal{V}_{\mathrm{Q}}$. A subset of the device state vectors are in one-to-one correspondence with the eigenstates and eigenvalues of the operator $\mathcal{O}$ given in (9.1), namely,

$$|\psi_n\rangle \leftrightarrow |D_n\rangle; \;\; n = 1, 2, \ldots, N < M \tag{9.15}$$

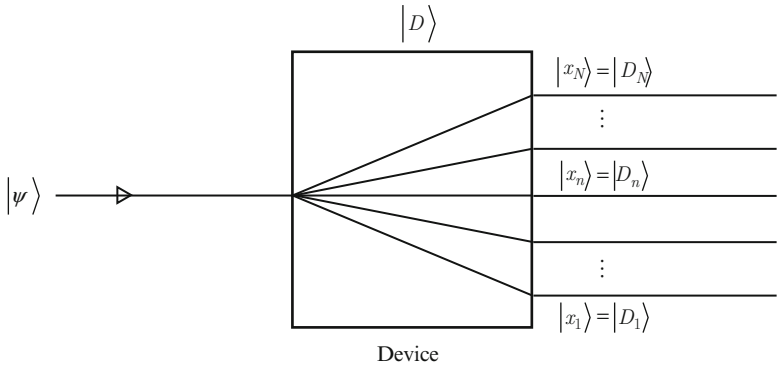


Fig. 9.5 The experimental device: the incoming quantum state $|\psi\rangle$ goes through the experimental apparatus, interacts with the device states $|D\rangle$, and emerges being resolved into the detector states realized as the counter pointing to position $|x_n\rangle$, $n = 1, 2, \ldots, N$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

$$\lambda_n \leftrightarrow x_n \tag{9.16}$$

The correspondence of the eigenvalues λ_n with the state of the device, symbolized by the position of the counter being at x_n, is taken to mean that the eigenvalues of an operator $\mathcal{O}$ can be recovered from the readings of the counter.[2] The correspondence of eigenstates and eigenvalues with the device states and counterpositions is based on the premise that the experimental device is designed to measure the properties of the said operator $\mathcal{O}$.

For example, for a system having angular momentum ℓ, the eigenstates are given by $|\ell, m\rangle$, with the expectation value of the z-component of angular momentum L_z being given by $m = 0, \pm 1, \pm 2, \ldots, \pm\ell$. For a given ℓ, the device is designed to measure the $2\ell + 1$ eigenstates and eigenvalues. Each pointer position x_n of the experimental device, shown in Fig. 9.5, can be taken to be the reading of the counter for a specific values of m, with $2\ell + 1 = N$. The special case for spin $s = 1/2$, with states $|1/2, m\rangle$ having $m = \pm 1/2$, will be discussed in Sect. 10.6 for the Stern–Gerlach experiment.

Another example is a device measuring the eigenenergies of the hydrogen atom, with the experimental detector designed to measure the absorption and emission radiation by the hydrogen atom.

To describe the state vector of the quantum entity that belongs to Hilbert space $\mathcal{V}_{\mathrm{Q}}$, the device Hilbert space $\mathcal{V}_{\mathrm{D}}$ needs to contain the Hilbert space $\mathcal{V}_{\mathrm{Q}}$, namely,

$$\mathcal{V}_{\mathrm{D}} \supset \mathcal{V}_{\mathrm{Q}}$$

[2]The more general statement is that the eigenvalues λ_n can be reconstructed from the collection of all the counter readings x_i; $i = 1, 2, \ldots, N$. The assumption that λ_n depends only on x_n is made for simplicity.

The states of the quantum entity and those of the device are described by elements of the tensor product space Hilbert space

$$\mathcal{V}_{\mathrm{Q}} \otimes \mathcal{V}_{\mathrm{D}}$$

Observables $\mathcal{O}$ that represent the physical properties of the quantum degree of freedom act only on the state vector $|\chi\rangle \in \mathcal{V}_{\mathrm{Q}}$ and do not have any action on the experimental device; hence, on the enlarged Hilbert space $\mathcal{V}_{\mathrm{Q}} \otimes \mathcal{V}_{\mathrm{D}}$, the observables are extended to act trivially on the device Hilbert space and represented by

$$\mathcal{O} \to \mathcal{O}_E = \mathcal{O} \otimes \mathbb{I}_{\mathrm{D}} \tag{9.17}$$

where $\mathbb{I}$ is the identity operator on device Hilbert space $\mathcal{V}_{\mathrm{D}}$.

The quantum degree of freedom and the experimental device are *coupled* so that the states of the device can provide a measurement of the properties of the quantum entity. Let H_{Q} be the Hamiltonian of the quantum system that acts only on $\mathcal{V}_{\mathrm{Q}}$; one has to add an interaction Hamiltonian H_{QD} that evolves the device states as well as *couples* the quantum degrees of freedom to the device; H_{QD} acts on the larger Hilbert space $\mathcal{V}_{\mathrm{Q}} \otimes \mathcal{V}_{\mathrm{D}}$.

Hence, the Hamiltonian H required for performing a measurement is given by

$$H = H_{\mathrm{Q}} \otimes \mathbb{I}_{\mathrm{D}} + H_{\mathrm{QD}} : \quad \mathcal{V}_{\mathrm{Q}} \otimes \mathcal{V}_{\mathrm{D}} \to \mathcal{V}_{\mathrm{Q}} \otimes \mathcal{V}_{\mathrm{D}} \tag{9.18}$$

9.6 The Process of Measurement

The subject of what constitutes a "measurement" in quantum mechanics is a vast subject, and the underlying principles are still being debated. Instead of discussing this subject in full generality, only two specific questions given below are addressed. The discussion of this section is revisited in Sect. 10.6 in the concrete context of the measurement of the electron's spin in the Stern–Gerlach experiment.

The fundamental role of the experimental device is to (a) amplify the quantum properties so that they can be made to correspond to macroscopically observable quantities and (b) to "collapse" the quantum state and in doing so causing an irreversible change. The process of measurement is illustrated in Fig. 9.6; the quantum state ψ is generated by a source in the laboratory and then sent to the device, where the quantum state and device state D are entangled, with the quantum state being amplified; a measurement is then performed on the quantum state thus bringing about an irreversible change, with the result being recorded by the detector.

Figure 9.6 illustrates the roles played by the empirical and trans-empirical domains in an experiment. The quantum state ψ is prepared empirically so as to create a non-entangled product state of the quantum entity and the device, namely, $|\psi\rangle|D\rangle$. The time evolution of both, the quantum state and the detector states, then takes place for a duration of t_* entirely in the trans-empirical domain and leads to

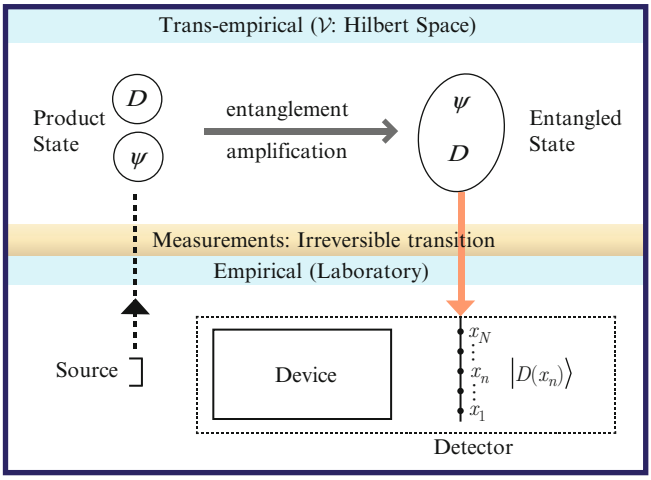


Fig. 9.6 The process of quantum measurement: empirical preparation of the quantum state and the transition from its trans-empirical state to the empirical manifestation. The collapse of the state vector $|\psi\rangle \otimes |D\rangle$ to a specific state $|\psi_n\rangle \otimes |D_n(t_*)\rangle$ is shown by the downward-pointing arrow (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

the amplification of the quantum state as well as its *entanglement* with the detector states. At later time t_*, the quantum state is empirically detected by collapsing the quantum state at the detector state $|D_n(t_*)\rangle$.

Consider the complete orthonormal eigenstates $|\psi_n\rangle$ of $\mathcal{O}$, as given in (9.1), and that provide its spectral representation

$$\mathcal{O} = \sum_{n=1}^{N} \lambda_n |\psi_n\rangle\langle\psi_n|; \qquad \mathcal{O}|\psi_n\rangle = \lambda_n |\psi_n\rangle \tag{9.19}$$

We address, in this and the following Sect. 9.8, the following two issues:

- Given a given state vector

$$|\chi\rangle = \sum_{n=1}^{N} c_n |\psi_n\rangle \tag{9.20}$$

how can one experimentally determine the coefficients $|c_n|^2$?

- How can one experimentally determine the expectation value of $\mathcal{O}$ for the state $|\chi\rangle$, namely, $\langle\chi|\mathcal{O}|\chi\rangle$? In particular, this entails determining all the eigenvalues λ_n of the operator $\mathcal{O}$.

The measurement process has the following three components:

- *Initial product state*: The initial state of the quantum system $|\chi\rangle$ and device $|D\rangle$ is a non-entangled product state as the two are not yet coupled; this yields the following initial product state and pure density matrix:

$$|\Phi_{\text{in}}\rangle = |\chi\rangle \otimes |D\rangle; \qquad \langle D|D\rangle = 1$$
$$\rho_{\text{in}} = |\Phi_{\text{in}}\rangle\langle\Phi_{\text{in}}| \tag{9.21}$$

- *Amplification and entanglement*: The initial state vector's evolution is driven by the Hamiltonian H given in (9.18); the coupled quantum entity and device evolve for a time t_*, which is a characteristic time scale for amplifying the quantum quantities by coupling *distinct* microscopic quantum states to *distinct* macroscopic device states. The Hamiltonian H is designed to entangle the quantum and device states.

  At the end of time t_*, for an ideal device, one has a *maximally entangled state*, as derived in (6.37), given by the following:

$$\begin{aligned} |\Phi_{\text{out}}\rangle &= \mathrm{e}^{-\mathrm{i}t_* H}|\Phi_{\text{in}}\rangle \\ &= \sum_{n=1}^{N} c_n |\psi_n\rangle \otimes |D_n(t_*)\rangle \\ \rho_{\text{out}} &= |\Phi_{\text{out}}\rangle\langle\Phi_{\text{out}}| \end{aligned} \tag{9.22}$$

  Note that (9.22) is an expression of the so-called Schrödinger cat problem: the state $|\Phi_{\text{out}}\rangle$ consists of a quantum *superposition* of the *macroscopic states* of the device—the correctness of which is still being debated.[3]
- *Collapse of state vector*: The final stage of the measurement process, according to Bohr and Heisenberg, is to collapse the state vector; this collapse entails making an *irreversible change* in the coupled quantum system-device arrangement—by, for example, recording the result of the measurement.

9.7 Mixed Density Matrix $\rho_{\mathbf{M}}$

The collapse of the state vector yields a discontinuous transition from the pure ρ_{out} to the mixed density matrix ρ_{M} given by

$$\begin{aligned} &\rho_{\text{out}} \to \text{Measurement} \to \tilde{\rho}_{\text{M}} \\ &\tilde{\rho}_{\text{M}} = \sum_{n=1}^{N} |c_n|^2 |\psi_n\rangle\langle\psi_n| \otimes |D_n(t_*)\rangle\langle D_n(t_*)| \\ &\text{tr}(\tilde{\rho}_{\text{M}}^2) < 1 \end{aligned} \tag{9.23}$$

[3] Schrödinger illustrated the paradox of superposing macroscopic states by the famous cat example. A device releases a poison, if triggered by the (uncertain) alpha decay of an unstable atom, that kills a cat. Before the cat is observed, the state of the cat is the following superposed state, namely, $|\text{cat}\rangle = |\text{cat; dead}\rangle| + |\text{cat; alive}\rangle$. The paradox lies in explaining how can the cat be dead and alive at the same time.

Recall from (9.12) and (9.13) that the result of every measurement, without any reference to the experimental device, results in decoherence and yields the mixed density matrix $\rho_{\rm M}$ for the state vector given by

$$\rho = \sum_{ij=1}^{N} c_i c_j^* |\psi_i\rangle\langle\psi_j| \rightarrow \text{Measurement} \rightarrow \rho_{\rm M} = \sum_{n=1}^{N} |c_n|^2 |\psi_i\rangle\langle\psi_i|$$

Every measurement records the outcome $|\psi_i\rangle\langle\psi_i|$, which corresponds to a specific eigenstate and has a likelihood of $|c_i|^2|$ of occurrence. The fundamental indeterminacy of the quantum degree of freedom is reflected in the fact that it is *impossible* to predict which projector $|\psi_i\rangle\langle\psi_i|$ will detect the collapse; all one can say is that there is a certain probability of the collapse occurring at only one particular detector.

The off-diagonal terms $|\psi_i\rangle\langle\psi_j|$, $i \neq j$ are purely quantum mechanical and their abscence reflects that the interference of trans-empirical states is completely absent. The mixed density matrix is an observable quantity and exists in the empirical domain in contrast to the pure density matrix that exists in the trans-empirical domain, and is illustrated in Fig. 9.4.

Obtaining the mixed density matrix $\tilde{\rho}_{\rm M}$ given in (9.23) completes the process of measurement and yields a result that is similar to (9.13), except now the device eigenstates are maximally entangled with the state vector. The quantum superposition cross-terms $c_i c_j^*$ that appear in (9.22) for $\rho_{\rm out}$ have all been removed from $\tilde{\rho}_{\rm M}$ by the process of measurement—and leading to the collapse of the state vector (decoherence).

The mixed density matrix $\tilde{\rho}_{\rm M}$ is a critical link in the interpretation of quantum mechanics—its crucial feature being the *absence* of cross-terms of detector states, namely, $|\psi_m\rangle|D_m(t_*)\rangle\langle D_n(t_*)|\langle\psi_n|, m \neq n$; these terms would imply the collapse of the state vector being simultaneously recorded by *two* detectors—and would invalidate the interpretation of $|c_n|^2$ as the probability p_n of observing the collapsed state at detector state $|D_n(t_*)\rangle$.

What is remarkable is that all measurements performed on a quantum entity are consistent with the postulate of the collapse of the state vector—with only a *single* detector recording the collapse of the state vector.[4]

Noteworthy 9.2: Classical random ensemble

Classical probability theory is discussed in Sect. 7.3. Consider a classical random variable X that takes only two values, represented by two balls that are either B and W; on observing (sampling) the random variable, one obtains its different possible values with probability given by p_B and p_W, with $p_B + p_W = 1$. One way of representing the sampling of the random variable is to theoretically construct a

[4]More advanced books on the foundations of quantum mechanics take issue with this interpretation arguing that a measurement in fact leads to $U\rho_{\rm M}U^\dagger$ that brings back the mixing of all the detector states [4]. The transformation by U can, in principle, be undone by a rotation of the basis states being used to make the measurements.

classical ensemble that has infinitely many B and W balls, and with the number of the balls in the ensemble being in proportion to p_B and p_W. When a ball is picked from the ensemble, every ball has an equal likelihood of being chosen; hence, when a ball is chosen from the ensemble, the likelihood of choosing either B or W will converge to p_B and p_W as the number of times one chooses goes to infinity.

The fundamental difference between classical randomness and quantum indeterminacy lies in the nature of the classical random ensemble; a similar construction is invalid for the quantum degree of freedom since no experiment can sample the specific values of the degree of freedom, which before the quantum entity is observed has no preexisting value. This aspect of quantum mechanics has been discussed in Chap. 7.

Since there is no quantum superposition of the different states of the detector every time the state vector is detected, there will be one and only one reading of the detectors, and by repeating the experiment, one can obtain the probability $p_n = |c_n|^2$ that the collapse of the state vector is detected by the detector state $|D_n(t_*)\rangle$. Hence, both the mixed density matrices ρ_{M} and $\tilde{\rho}_{\mathrm{M}}$ are equivalent to empirical quantities. This conclusion is discussed further in Sect. 9.8.

It is essential that there be no quantum cross-terms in the mixed density matrix. Since the basis states being used to make the measurements are only defined up to a unitary transformation, as discussed in (5.11), to interpret the mixed density matrix as being equivalent to a classical random ensemble, one only needs to prove that there exist complete basis states in which the mixed density matrix is diagonal, as given in (9.13) or (9.23).

Both ρ_{M} and $\tilde{\rho}_{\mathrm{M}}$ are equivalent to classical random ensembles and for which all the random outcomes, unlike the case of quantum indeterminacy, exist objectively before they are observed. What this means is that there is a classical ensemble that is equivalent to ρ_{M} and $\tilde{\rho}_{\mathrm{M}}$; the classical ensemble consists of a collection of possible outcomes, say $|\psi_i\rangle\langle\psi_i|$; the key point is that one can *objectively* assign a probability $p_n = |c_n|^2$ that the outcome $|\psi_i\rangle\langle\psi_i|$ will be observed when the ensemble is sampled.

9.8 Reduced Density Matrix $\rho_{\mathbf{R}}$

Consider the quantum entity-device being described by the mixed density matrix $\tilde{\rho}_{\mathrm{M}}$ as given in (9.23), namely,

$$\tilde{\rho}_{\mathrm{M}} = \sum_{n=1}^{N} |c_n|^2 |\psi_n\rangle\langle\psi_n| \otimes |D_n(t_*)\rangle\langle D_n(t_*)|$$

If one takes a reading of the device after one run of the experiment has been completed, one will find the device to be in a state $|D_n(t_*)\rangle$, with the pointer of the dectector at position x_n; if one does a thought experiment in which one subsequently (or simultaneously with the measurement of the state by the device) measures the

state of the quantum entity, one is *certain* to find the quantum entity to be in state $|\psi_n\rangle$ due to the entanglement of the quantum states and the detector states.[5]

Hence knowing the state of the device automatically gives us full knowledge of the quantum state vector. This is the reason one does not need to directly observe the quantum state $|\psi_n\rangle$; instead, it is sufficient to observe only the detector states $|D_n(t_*)\rangle$ and infer the properties of the quantum state [15].

Determining the Coefficients $|\mathbf{c_n}|^2$

Due to entanglement of the state vector and detector states, only the state of the device $|D_n(t_*)\rangle$ is observed, regardless of the state of the quantum entity. Completely ignoring the quantum degree of freedom results in a loss of information and is expressed by summing the mixed density matrix ρ_{M} over the degrees of freedom of Hilbert space $\mathcal{V}_{\mathrm{Q}}$ and yields the *reduced density matrix* ρ_{R} given by

$$\rho_{\mathrm{R}} = tr_{\mathcal{V}_{\mathrm{Q}}}\left[\tilde{\rho}_{\mathrm{M}}\right] = tr_{\mathcal{V}_{\mathrm{Q}}}\Big[\sum_{n=1}^{N} |c_n|^2 |\psi_n\rangle\langle\psi_n| \otimes |D_n(t_*)\rangle\langle D_n(t_*)|\Big]$$

$$= \sum_{n=1}^{N} |c_n|^2 |D_n(t_*)\rangle\langle D_n(t_*)| \tag{9.24}$$

$$\Rightarrow \mathrm{tr}(\rho_{\mathrm{R}} |D_n(t_*)\rangle\langle D_n(t_*)|) = |c_n|^2 = p_n \tag{9.25}$$

Since the mixed density matrix $\tilde{\rho}_{\mathrm{M}}$ represents a *classical random system*, on being observed, the system will be found in the state $|\psi_n\rangle \otimes |D_n(t_*)\rangle$ with likelihood p_n; in symbols

$$tr\big(\rho_{\mathrm{R}} |D_n(t_*)\rangle\langle D_n(t_*)|\big) = p_n \tag{9.26}$$

From the fundamental postulate of quantum mechanics due to Max Born, we can conclude that $|c_n|^2 = p_n$, and this completes the determination of the coefficients c_n.

Determining the Expectation Value of $\mathcal{O}$ and Eigenvalues $\lambda_{\mathbf{n}}$

The operator $\mathcal{O}$ and its expectation value are given, from (9.6), by

$$\mathcal{O} = \sum_{n=1}^{N} \lambda_n |\psi_n\rangle\langle\psi_n|; \;\; \rho = |\chi\rangle\langle\chi; \;\; |\chi\rangle = \sum_n c_n |\psi_n\rangle$$

[5] One can create a more complicated experiment where a subsequent measurement is performed on the final state with *another* device and come to the same conclusion as the thought experiment.

$$E_\chi[\mathcal{O}] \equiv \langle\chi|\mathcal{O}|\chi\rangle = \sum_n \lambda_n |c_n|^2 = \mathrm{tr}(\rho\mathcal{O})$$

On being coupled to the device, the state vector $|\chi\rangle$ is extended to $|\chi\rangle \otimes |D\rangle$; the expectation value is given on the tensor product state vector of the extended operator $\mathcal{O}_E$ (given in (9.17)) and yields the following:

$$E_\chi[\mathcal{O}] = E_{\chi\otimes D}[\mathcal{O}_E] = \mathrm{tr}\Big\{(\mathcal{O}\otimes\mathbb{I}_\mathrm{D})\rho_\mathrm{in}\Big\}$$

On evolving the density matrix ρ_in for time t_* yields

$$E_{\chi\otimes D}[\mathcal{O}_E] = \mathrm{tr}\Big\{(\mathcal{O}\otimes\mathbb{I}_\mathrm{D})\rho_\mathrm{out}\Big\}$$

and completing the measurement process by collapsing the state vector gives

$$\rho_\mathrm{out} \to \tilde{\rho}_\mathrm{M} \quad \Rightarrow \quad E_{\chi\otimes D}[\mathcal{O}_E] = \mathrm{tr}\Big\{(\mathcal{O}\otimes\mathbb{I}_\mathrm{D})\tilde{\rho}_\mathrm{M}\Big\}$$

where $\tilde{\rho}_\mathrm{M}$ is given by (9.23).

The expectation value of the operator $\mathcal{O}_E$ is evaluated in two steps; first a partial trace is performed over the quantum entity's Hilbert space $\mathcal{V}_\mathrm{Q}$ and one is left with performing the trace over the device Hilbert space $\mathcal{V}_\mathrm{D}$. Using the notation of $\mathcal{O}_{E|\chi}$ to denote the partial trace of operator $\mathcal{O}_E$ over $\mathcal{V}_\mathrm{Q}$ yields[6]

$$\begin{aligned}
\mathcal{O}_{E|\chi} \equiv E_\chi[\mathcal{O}_E] &= \mathrm{tr}_{\mathcal{V}_\mathrm{Q}}\Big\{(\mathcal{O}\otimes\mathbb{I}_\mathrm{D})\rho_\mathrm{M}\Big\} \\
&= \mathrm{tr}_{\mathcal{V}_\mathrm{Q}}\Big\{(\mathcal{O}\otimes\mathbb{I}_\mathrm{D})\sum_{n=1}^{N}|c_n|^2|\psi_n\rangle\langle\psi_n|\otimes|D_n(t_*)\rangle\langle D_n(t_*)|\Big\} \\
&= \sum_{n=1}^{N}|c_n|^2\langle\psi_n|\mathcal{O}|\psi_n\rangle\cdot\mathbb{I}_\mathrm{D}|D_n(t_*)\rangle\langle D_n(t_*)| \\
\Rightarrow \mathcal{O}_{E|\chi} &= \sum_{n=1}^{N}\lambda_n|c_n|^2|D_n(t_*)\rangle\langle D_n(t_*)|
\end{aligned}$$

The final result is given by tracing over the detector Hilbert space $\mathcal{V}_\mathrm{D}$ and yields

$$\begin{aligned}
E_\chi[\mathcal{O}] = E_{\chi\otimes D}[\mathcal{O}_E] &= E_\mathrm{D}[\mathcal{O}_{E|\chi}] \\
&= \mathrm{tr}_{\mathcal{V}_\mathrm{D}}\{\mathcal{O}_{E|\chi}\} = \sum_{n=1}^{N}\lambda_n|c_n|^2
\end{aligned}$$

[6]Note the notation used implies that $\mathcal{O}_{E|\chi} = E_\chi[\mathcal{O}_E]$ is an operator on $\mathcal{V}_\mathrm{D}$ and not equal to $E_\chi[\mathcal{O}]$, which is a real number.

One can observe the coefficients $\lambda_n|c_n|^2$ of the operator $\mathcal{O}_{E|\chi}$ by measuring its expectation value for the state vector $|D_n(t_*)\rangle$ given by

$$\langle D_n(t_*)|\mathcal{O}_{E|\chi}|D_n(t_*)\rangle = \lambda_n(x_n)|c_n|^2 \tag{9.27}$$

Each device state $|D_n\rangle$ is in one-to-one correspondence with the eigenstates $|\psi_n\rangle$ as in (9.15); the value of the counter reading x_n is put into a one-to-one correspondence with the eigenvalue λ_n, that is, as discussed in (9.16)

$$\lambda_n = \lambda_n(x_n)$$

Summing over the readings of all the device states yields the sought for $E[\mathcal{O}]$, namely,

$$E_\chi[\mathcal{O}] = \sum_n \langle D_n(t_*)|\mathcal{O}_{E|\chi}|D_n(t_*)\rangle = \sum_n \lambda_n|c_n|^2$$

The eigenvalues λ_n can be obtained from the results obtained. The coefficients $|c_n|^2$ are equal to the empirically observed probability p_n that have been obtained in (9.26), and using them with (9.27) yields

$$\frac{\langle D_n(t_*)|\mathcal{O}_{\mathrm{R}}|D_n(t_*)\rangle}{p_n} = \lambda_n \tag{9.28}$$

where λ_n is the eigenvalue of the operator $\mathcal{O}$, thus completing the determination of all the eigenvalues of $\mathcal{O}$.

Mixed Density Matrix and Decoherence

Note that measurement leads to an irreversible change since all the off-diagonal terms in (9.12)—which are responsible for quantum interference and other nonclassical effects—are zero in (9.13), and reflect the loss of information.

In other words, the process of measurement leads to decoherence by collapsing a *pure state* into a *mixed state* and, as given earlier in (9.23), yields the following:

$$\begin{aligned} \rho &\to \tilde{\rho}_{\mathrm{M}} \\ \tilde{\rho}_{\mathrm{M}} &= \sum_{n=1}^{N} |c_n|^2 |\psi_n\rangle\langle\psi_n| \otimes |D_n(t_*)\rangle\langle D_n(t_*)| \end{aligned}$$

The result of measurement, as in (9.23), is to collapse ρ to $\tilde{\rho}_{\mathrm{M}}$.

The mixed density matrix $\tilde{\rho}_{\mathrm{M}}$ represents *classical probability*: the probability of observing a state $|\psi_n\rangle \otimes |D_n(t_*)\rangle$ is given by $p_n = |c_n|^2$, with the p_n's obeying the rules of classical probability.

- The discontinuous transformation of ρ to $\tilde{\rho}_{\mathrm{M}}$ is *irreversible* and is described by an increase in quantum entropy, defined in (6.34). More precisely, for the pure state,

$$\mathcal{S} = -\mathrm{tr}(\rho \ln \rho) = 0$$

whereas the mixed state, for $p_i = |c_i|^2 \neq 1$, yields

$$\mathcal{S}_{\mathrm{M}} = -\mathrm{tr}(\tilde{\rho}_{\mathrm{M}} \ln \tilde{\rho}_{\mathrm{M}}) = -k_{\mathrm{B}} \sum_{i=1}^{N} p_i \ln p_i > 0 \tag{9.29}$$

- Any unitary transformation on the state vector, such as the time evolution due to the Schrödinger equation, results in $|\chi\rangle \to U|\chi\rangle$, where $UU^{\dagger} = \mathbb{I}$. This in turn induces the following unitary transformation on the pure density matrix $\rho \to U\rho U^{\dagger}$ and results in the entropy of ρ being invariant.

  However, from (9.29), it is seen that the measurement process *increases* the entropy. Hence the change of ρ to $\tilde{\rho}_{\mathrm{M}}$ cannot be brought about by a unitary transformation, and as mentioned earlier, the process of measurement cannot result from the time evolution driven by the Schrödinger equation.
- From the analysis of measurement, one reaches the conclusion that was mentioned in the introduction to this chapter, namely, that quantum mechanics is founded on *two assumptions*. *Firstly*, the time evolution of the state vector is unitary and determined by the Schrödinger equation, and *secondly*, the process of measurement brings about a non-unitary change in the state vector, and this change is not determined by the Schrödinger equation.

In summary, the measurement process causes an irreversible change in the state vector $|\psi\rangle$—by projecting its density matrix from pure state ρ to mixed state ρ_{M}, which describes a classical random system. The very process of measurement results in decoherence being induced on the state vector. Hence

$$\rho \to \text{Measurement} \to \rho_{\mathrm{M}}: \quad \text{Decoherence} \tag{9.30}$$

Noteworthy 9.3: "Collapse" of state vector and nonlocality

The process of the collapse of the state vector causes an irreversible change in the system and is called decoherence. The nonlocal collapse of $\psi(t,x)$ has puzzled physicists since the beginning of quantum mechanics.

For the collapse to take place instantaneously apparently requires that "information" about the state vector being detected by a particular projection operator is communicated at infinite speed to the rest of space—and would seem to violate the special theory of relativity according to which the velocity of light is the fastest speed for any and every form of communication.

The nonlocal nature of the state vector $\psi(t,x)$ is consistent with all the experiments that have been devised to test this aspect of quantum mechanics. In particular, the famous EPR paradox is a test of the nonlocal characteristic of the state vector, and experiments have shown that it is, in fact, nonlocal. Moreover, a detailed analysis shows that quantum measurement theory is consistent with the special theory of relativity [6].

A central unresolved mystery of quantum mechanics is the instantaneous collapse of the state vector. It is, in fact, the collapse of the state vector that gives meaning to the intrinsically probabilistic nature of quantum mechanics and to the concomitant existence of the trans-empirical domain that embodies all the nonclassical behaviors of the quantum entity.

My own view is that a possible explanation of the instantaneous collapse of the state vector is to take the "picture" of the state vector existing in the trans-empirical domain as being literally true. The state vector $|\psi\rangle$ has a symbol-like reality since it is a mathematical construct that contains the likelihood of all possible outcomes. A precise measurement of the state of the system is equivalent to the complete erasure of the symbol-like structure.[7]

Hence, the collapse of the state vector due to an observation can be thought of as the analog of the "symbol" ceasing to exist; this erasure of the symbol takes place for the entirety of the symbol and is expressed as the instantaneous collapse of the state vector. Since the state vector as a symbol exists only in the trans-empirical domain, the laws of relativity do not apply as relativity limits the speed of transfer of information only for classical signals that exist solely in the empirical domain.

The physics of the trans-empirical domain is determined by laws that are fundamentally nonclassical, symbolic, and mathematical in essence.

9.9 Preparation of a Quantum State

The preparation of a quantum state has features that are similar to the measurement of a state vector but also has some differences. A measurement of a quantum state ascertains the quantitative values of properties of a state vector that exist *before* the measurement is made; a measurement may destroy the measured state, as is the case of a particle being detected by a photographic plate and being lost in the plate. In contrast, the preparation of a state is made with the intention of studying the degree of freedom's state vector *after* its preparation. Furthermore, it is essential that the state be left intact after the preparation so that it may be studied further [4, 19].

Consider one of the simplest, but nevertheless, one of the most important quantum states, namely, a "free" electron inside a box. The preparation of the state is illustrated in Fig. 9.7.

[7] A partial erasure of the symbol is also possible, as discussed in Sect. 8.8 on the Quantum Erasure.

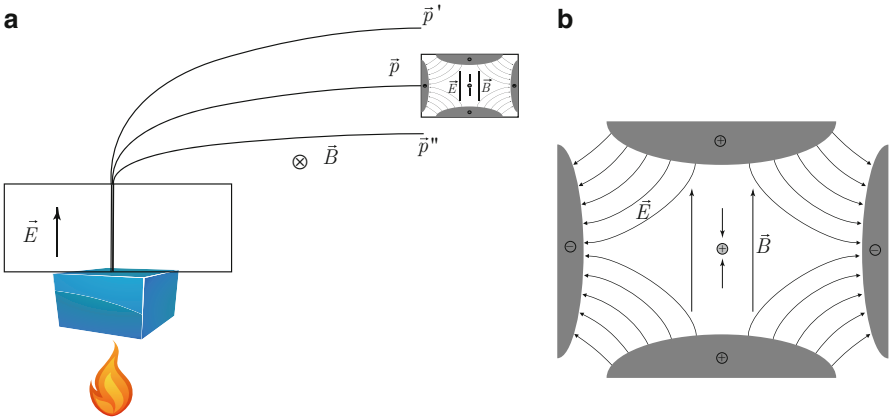


Fig. 9.7 (**a**) Electrons emitted by a source are post-selected for obtaining an electron with a definite momentum $\vec{p}'$. (**b**) The electron inside a cavity (*box*) is kept confined to the cavity using a three-dimensional Penning trap (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Free electrons are obtained from a source, which can be a metal that is heated by a flame. An electric field is applied to the electrons to collimate the beam of electrons as well as to increase their energy. For electrons moving at velocities of a few km/sec, the quantum states are sharply localized wave packets with the motion of the wave packet being well approximated by classical electromagnetism. Hence, one can use a semiclassical analysis for analyzing the evolution of the wave packet.

Electrons with different momentum $\vec{p}'$, $\vec{p}''$ $\vec{p}'''$ and so on are separated by imposing a magnetic field on the electrons. An electron with momentum $\vec{p}'$ is chosen to enter into a cavity, also called a box. In symbols, the initial state of the electron is given by $\rho = |\chi\rangle\langle\chi|$; the procedure for separating the state into different momentum states $|\psi_{\vec{p}'}\rangle$ leads to the following:

$$\rho \to |\psi_{\vec{p}'}\rangle\langle\psi_{\vec{p}'}|$$

The *selection* of a momentum eigenstate for the electron is shown in Fig. 9.7a: the magnetic field causes the electron wave packets to move in a circular motion determined by its momentum, with an electron with a momentum $\vec{p}'$, up to some small error, entering the aperture to the cavity. The electron is in a momentum eigenstate when it enters the cavity. Once the electron crosses the aperture, the shutter is closed, and the electron is now inside the cavity with state vector

$$|\psi_{\vec{p}'}\rangle : \quad \text{state vector entering cavity}$$

The process of selection of momentum eigenstate is a measurement of the momentum of a quantum state: Any momentum measurement made after the selection will yield, with probability 1, momentum $\vec{p}'$.

If, inside the cavity, the electron is not acted upon with some form of interaction, its momentum will lead it to collide with the wall of the cavity and the electron will be lost.

To confine the electron to the cavity, the electron is subjected to new (energy and momentum conserving) interactions. There are various devices that can confine an electron to a cavity, the most widely used ones being the Paul trap and the Penning trap [26].

The Paul trap is an electric quadrupole ion trap that can confine a charged particle inside a cavity in both one and three dimensions.[8] The ion trap uses a *time-dependent* electric field to confine the ion to a cavity; in particular, it uses a constant gradient electric field and an electric field oscillating at the radio frequency to trap the ions.

The Penning trap, in contrast, confines charged particles inside a cavity using *both* a homogeneous static magnetic field and a spatially inhomogeneous static electric field. Figure 9.7b shows a drawing of the Penning trap confining an electron to the cavity.

The state preparation yields the state vector of the electron inside the cavity box. Once the electron enters the cavity, it is subjected to new electromagnetic interactions that changes the nature of its state vector. Being confined to the cavity leads to the boundary condition that the state vector be zero on the boundary of the cavity.

Once the electron enters the cavity, one no longer has any knowledge of its quantum state. Since the electron inside the cavity is no longer being observed, the electron makes a smooth transition from being an empirical entity—for which it is known to be in definite state—to a trans-empirical and indeterminate state. The transition of the electron from the empirical condition to its trans-empirical state is smooth in contrast to its transition from its trans-empirical state to its observed empirical manifestation.

Inside the cavity, the electron's behavior is described by its trans-empirical state vector given by

$$|\psi\rangle_{\text{Box}} = |\psi_{\vec{p}}\rangle_{\text{Zero on boundary of cavity}}: \quad \text{state vector inside cavity}$$

For a box with dimensions $L \times L \times L$ and momentum $\vec{p} = \frac{2\pi\hbar}{L}(n_x, n_y, n_z)$ with n_x, n_y, n_z being integers, the state vector for a free particle in a box is given by

$$\begin{aligned}\langle x,y,z|\psi\rangle_{\text{Box}} &= \langle x,y,z|\psi_{\vec{p}}\rangle_{\text{Zero on boundary of cavity}} \\ &= \left(\frac{2}{L}\right)^{3/2} \sin\left(\frac{2\pi x n_x}{L}\right) \sin\left(\frac{2\pi y n_y}{L}\right) \sin\left(\frac{2\pi z n_z}{L}\right)\end{aligned}$$

[8] Static electric fields alone cannot act as a trap since the electron will drift along the direction of the electric field and finally hit the electric charge that is the source of the electric field.

9.10 The Heisenberg Uncertainty Principle

The Heisenberg Uncertainty Principle is one of pillars of quantum mechanics and is a fundamental result required for avoiding many apparent inconsistencies. Chapter 7 discusses the EPR paradox that seems to imply that the Heisenberg Uncertainty Principle is violated; in spite of what one would naively think, the uncertainty principle turns out to be valid in a remarkable and nontrivial manner.

The expectation value of commuting operators $\mathcal{O}_n, n = 1, 2, \ldots, N$ can be simultaneously evaluated since there exist state vectors that are simultaneously the eigenstate of all the commuting operators. Hence, one can design an experiment, as illustrated in Fig. 5.5, in which all the N eigenvalues are measured simultaneously, and as given in (9.8), one can simultaneously evaluate $\mathrm{tr}(\rho\mathcal{O}_n)$ for all $n = 1, 2, \ldots, N$.

What happens when one tries to simultaneously evaluate, for a given state vector $|\psi\rangle$, the expectation values of two *non-commuting* operators, say A and B?

Let us consider an experiment that can *exactly* measure the expectation value of A and then examine what this measurement process yields for the expectation value of B. The experimental device is designed to measure the eigenstates of A, with no reference being made to the eigenstates of the operator B. The state vector $|\psi\rangle$ is decomposed in terms of the eigenstates of A, namely, $|n\rangle$ with eigenvalue a_n; the action of the operators is given by the following:

$$A|n\rangle = a_n|n\rangle \; ; \; B|n\rangle = \sum_m b_{nm}|m\rangle \; ; \; |\psi\rangle = \sum_n c_n|n\rangle \; ; \; \rho = \sum_{nm} c_n c_m^* |n\rangle\langle m|$$
$$\Rightarrow E[A] = \mathrm{tr}(\rho A) = \sum_n |c_n|^2 a_n \; ; \; E[B] = \mathrm{tr}(\rho B) = \sum_{nm} b_{nm} c_n c_m^* \tag{9.31}$$

Equation (9.31) provides exact expressions for both A and B, so it would seem one should be able to measure the expectation value of both non-commuting operators. This, however, is not possible due to *decoherence* (collapse of the state vector) that is an essential component of the process of measurement. As shown in (9.30), the process of measurement yields

$$\rho \rightarrow \text{Measurement} \rightarrow \rho_{\mathrm{M}}$$
$$\Rightarrow E[A] = \mathrm{tr}(\rho A) \rightarrow \text{Measurement} \rightarrow \mathrm{tr}(\rho_{\mathrm{M}} A) = \sum_n |c_n|^2 a_n = E[A] \tag{9.32}$$
$$E[B] = \mathrm{tr}(\rho B) \rightarrow \text{Measurement} \rightarrow \mathrm{tr}(\rho_{\mathrm{M}} B) = \sum_n b_{nn}|c_n|^2 \neq E[B] \tag{9.33}$$

Equation (9.32) shows that measurement yields the *correct* expectation value for A. But as seen from (9.33), the same measurement yields an *incorrect* value for the expectation value for the non-commuting operator B—since all the off-diagonal terms in (9.31) required for correctly evaluating $E[B]$ have been canceled out due to the collapse of the state vector—namely, due to decoherence. Clearly, one cannot

find the expectation of *both* operators A and B since there is no state vector that is the simultaneous eigenstate of both the operators.

In the case considered, the expectation value of A was evaluated exactly, with the errors for operator B being uncontrolled. One may ask, can one choose a complete basis—with the basis states not necessarily being eigenstates of either A or B—so that the *joint error* in the evaluation of *both* $E[A]$ and $E[B]$ is *minimized*? The answer to this question is given by the Heisenberg Uncertainty Principle.

An Operator Inequality

Let A, B be Hermitian operators; to analyze the error in their expectation values, define the following:

$$\begin{aligned}
&\bar{A} = A - E[A]; \qquad \bar{B} = B - E[B]; \qquad \rho = |\psi\rangle\langle\psi| \\
&(\Delta A)^2 = E\left(\bar{A}\right)^2 = \mathrm{tr}(\rho A^2) - (\mathrm{tr}(\rho A))^2 \\
&(\Delta B)^2 = E\left(\bar{B}\right)^2 = \mathrm{tr}(\rho B^2) - (\mathrm{tr}(\rho B))^2 \\
&\text{and} \quad E\big([A,B]\big) = \mathrm{tr}(B\rho A) - \mathrm{tr}(A\rho B)
\end{aligned} \tag{9.34}$$

The derivation given below yields the following result:

$$\Delta A \Delta B \geq \frac{1}{2}|E[A,B]| : \quad \text{Generalized Uncertainty Principle} \tag{9.35}$$

Define the following state vectors:

$$\begin{aligned}
&\bar{A}|\psi\rangle = |\psi_A\rangle; \quad \bar{B}|\psi\rangle = |\psi_B\rangle \\
&\Rightarrow (\Delta A \Delta B)^2 = \langle\psi_A|\psi_A\rangle\langle\psi_B|\psi_B\rangle
\end{aligned}$$

For any two vectors $\mathbf{a}$ and $\mathbf{b}$ in an N-dimensional Euclidean space, one has the triangle inequality given by $|\mathbf{a}||\mathbf{b}| \geq \mathbf{a}\cdot\mathbf{b}$, where $|\mathbf{a}| = \sqrt{a_1^2 + \cdots + a_N^2}$. The generalization of the result for Euclidean space to Hilbert space operators is the Schwarz inequality given by

$$(\Delta A \Delta B)^2 \geq \left|E\left[\bar{A}\bar{B}\right]\right|^2 \tag{9.36}$$

The identity

$$\bar{A}\bar{B} = \frac{1}{2}[A,B] + \frac{1}{2}\{A,B\}$$

yields from the Schwarz inequality (9.36), for Hermitian operator $C = -i[A,B]$, the following:

$$(\Delta A \Delta B)^2 \geq \frac{1}{4}\left|E[(\{A,B\})^2] + i\big(E[C]\big)^2\right|$$

Choosing $|\psi\rangle$ such that $|\psi_A\rangle = |\psi_B\rangle$ makes the term $E[(\{A,B\})^2] = 0$ and yields the following generalized Heisenberg inequality:

$$(\Delta A \Delta B) \geq \frac{1}{2}\left|E[C]\right| = \frac{1}{2}\left|E([A,B])\right| \tag{9.37}$$

The generalized Heisenberg Uncertainty Principle states that there will always be errors $\Delta A, \Delta B$ when the expectation values of two non-commuting operators A, B are experimentally measured for any state vector; the minimum joint error is given in (9.37) and is equal to $\frac{1}{2}|E([A,B])|$. It is appropriate that only the commutator of A and B, namely, $[A,B]$, determines the minimum error since the commutator is a precise measure of the extent to which the two operators do not commute.

Position Momentum Uncertainty Relation

For the position and momentum operators $\hat{x}$ and $\hat{p}$, the commutation equation is given by

$$[\hat{x}, \hat{p}] = i\hbar$$

The generalized Heisenberg Uncertainty Principle, given in (9.35), yields the celebrated result of Heisenberg, namely, that

$$\Delta\hat{x}\Delta\hat{p} \geq \frac{1}{2}\hbar: \quad \text{Heisenberg Uncertainty Principle} \tag{9.38}$$

Consider the ground state vector of the simple harmonic oscillator, which is not an eigenstate of either $\hat{x}$ or $\hat{p}$, given by

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left\{-\left(\frac{m\omega}{\hbar}\right)\frac{x^2}{2}\right\}$$

It can be shown that $\psi_0(x)$ yields the minimum value for the joint errors of $\hat{x}$ and $\hat{p}$, namely,

$$\Delta\hat{x}\Delta\hat{p}\Big|_{\psi_0} = \frac{1}{2}\hbar$$

Time–Energy Uncertainty Relation

The role and definition of time in quantum mechanics is quite different from position; as discussed in some detail in Chap. 2, the position of a quantum entity becomes the position degree of freedom, which is a dynamical variable, while time remains an external parameter.[9] In contrast, time is not a degree of freedom and has the following threefold interpretations in quantum mechanics [28]:

- *Intrinsic time* is the parameter t that occurs in the Schrödinger equation, and every degree of freedom is a dynamical variable with a time evolution that marks the passage of intrinsic time.
- Experiments take place in space and time. The detectors and switches for electric and magnetic fields in the laboratory, the recording times of detectors, and so on consist of measurements of *external time* carried out by clocks in the laboratory.
- The observation of quantum events, such as the time of arrival of decay products from a radioactive source or the time of tunneling of an electron through a barrier, are examples of *observable time* that are random and have a statistical distribution.

Intrinsic time cannot be placed on the same basis as the position degree of freedom; the reason being that time is conjugate to energy, and an unbounded Hermitian operator representing time, due to the Heisenberg Uncertainty Principle, would require an unbounded energy operator—resulting in a Hamiltonian having a ground state unbounded from below, with minus infinite energy. This would lead to all quantum systems being unstable [15].

All references to time in this book are to intrinsic time, namely, the time that appears in the Schrödinger equation.[10] The time–energy uncertainty is for intrinsic time versus the observed energy of a quantum degree of freedom.

Consider an operator A; the Heisenberg equation of motion as given in (5.40) (and which is equivalent to the Schrödinger equation) is given by

$$i\hbar \frac{\mathrm{d}A}{\mathrm{d}t} = AH - HA = [A, H]$$

The parameter t is the intrinsic time of the quantum system. The general result given in (9.34) and (9.35) yields

$$\Delta A \Delta H \geq \frac{1}{2}\left|E\left([A, H]\right)\right| = \frac{\hbar}{2}\left|E\left(\frac{\mathrm{d}A}{\mathrm{d}t}\right)\right| \tag{9.39}$$

[9] This asymmetry of time and position is resolved in relativistic quantum field theory by "demoting" position from being a degree of freedom, as is the case for nonrelativistic quantum mechanics, to being a parameter like time; both space and time coordinates label the quantum field's degrees of freedom. Thus, having both $t, \mathbf{x}$ as parameters allows one to have exact relativistic invariance under Lorentz transformations on $t, \mathbf{x}$.

[10] All three concepts of time, namely, external, intrinsic, and observed time, can be employed to study a quantum process [28].

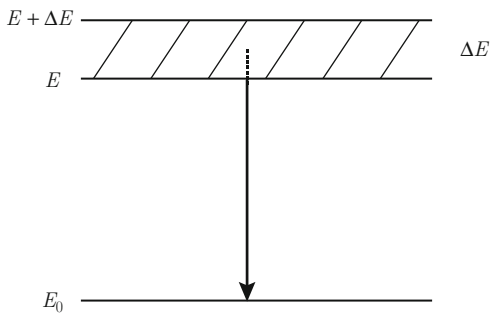


Fig. 9.8 Time–energy uncertainty (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Define the characteristic time for the operator A by the following equation:

$$\Delta t = \frac{\Delta A}{|\frac{\mathrm{d}}{\mathrm{d}t} E[A]|} \tag{9.40}$$

The uncertainty in energy is given by $\Delta H = \Delta E$. Equation (9.39) then yields the time–energy uncertainty as being given by

$$\Delta t \Delta E \geq \frac{1}{2}\hbar \tag{9.41}$$

To illustrate the time–energy uncertainty relation, consider state vectors of atoms that have two well-separated domains for the energy eigenstates: a ground state with energy E_0 and a continuum of eigenstates with eigenenergies in the interval $[E, E + \Delta E]$, as shown in Fig. 9.8.

Let the operator A represent, at time t, the number of atoms $N(t)$ that are described by a state vector which is a superposition of eigenstates with eigenenergies in the range $[E, E + \Delta E]$; the time–energy uncertainty relation states the rate of decay of the number of states is given by

$$N(t) = N_0 \mathrm{e}^{-t/\Delta t}; \quad \Delta t = \frac{\hbar}{2\Delta E}$$

In other words, the uncertainty in the energy ΔE of the state vector results in the state being *unstable*, having a lifetime of Δt and steadily decaying to the ground state. It also follows that if there is no uncertainty in the energy, that is, if $\Delta E = 0$, then the lifetime of the state is infinite, which indeed it must be since a state vector with $\Delta E = 0$ is an eigenstate of the Hamiltonian that never decays.

Heisenberg Uncertainty Principle and the Quantum Entity

The Heisenberg Uncertainty Principle is a tangible and measurable consequence of the indeterminacy of the underlying degree of freedom $\mathcal{F}$.

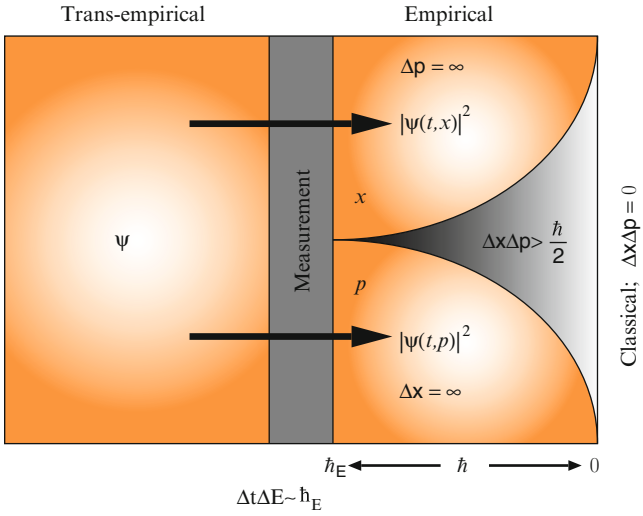


Fig. 9.9 The coordinate and momentum representations of the degree of freedom and the respective state vectors. The empirical value of Planck's constant is denoted by $\hbar_{\mathrm{E}}$. The classical limit emerges as $\hbar \to 0$ and is shown in the diagram (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

The transition from the trans-empirical state vector to its empirical manifestation requires a finite time Δt and is illustrated in Fig. 9.9. Consider a thought experiment carried out on a quantum entity with a perfect device. Every such experiment has an inherent inaccuracy Δt in the experimental determination of the intrinsic parameter of time t that appears in the Schrödinger equation. The limitation in accurately determining t is related to the unavoidable experimental inaccuracy, namely, ΔE, with which the energy E of the quantum state can be determined. These two limitations are related by the Heisenberg Uncertainty Principle—given in (9.41).

The quantum mechanical description of the degree of freedom by a state vector and Hermitian operators necessarily leads to the uncertainty principle. The degree of freedom $\mathcal{F}$ can have mathematical representations that are related by unitary transformations, and there is a different experimental device for ascertaining the properties of $\mathcal{F}$ for each representation. For non-commuting operators representing incompatible properties of the degree of freedom, the uncertainty principle sets a limit of how accurately the operators can be simultaneously measured.

In particular, due to (9.38), the uncertainty principle states the following. One can do a thought experiment that can *exactly measure* the values of *either* the coordinate projection operators $|x\rangle\langle x|$ (with $\Delta\hat{p} = \infty$) based on state vector $\psi(t,x)$ *or* the momentum projection operators $|p\rangle\langle p|$ (with $\Delta\hat{x} = \infty$) based on state vector $\psi(t,p)$, as shown in Fig. 9.9.

Consider a hypothetical scenario of the Planck's constant $\hbar$ varying smoothly and one tracks the quantum description of the degree of freedom, as shown in

Fig. 9.9. As $\hbar \to 0$, the purely quantum description—in terms of state vectors that are eigenstates of either the coordinate operators $|x\rangle\langle x|$ or momentum operators $|p\rangle\langle p|$—goes to zero.

The mixed description of the quantum system, with partial information about *both* the coordinate and momentum representations, is shown in the intermediate domain, with the accuracy of the description given by $\Delta\hat{x}\Delta\hat{p} \geq \hbar/2$. As one takes $\hbar \to 0$, the values of both $\Delta\hat{x}$ and $\Delta\hat{p}$ go to zero, and one obtains the classical description for which the values of *both* x, p are known exactly that is with $\Delta\hat{x}\Delta\hat{p} = 0$ and is shown as the classical limit in Fig. 9.9.

In summary, the mathematical representation of the quantum entity by a degree of freedom, state vectors, and operators—and the experimental determination of the physical properties of the quantum entity—necessarily leads to the Heisenberg Uncertainty Principle.

9.11 Theories of Quantum Measurement

There is at present no consensus amongst quantum theorists as to what constitutes a measurement. We have used the concept of the "collapse of the state vector" to explain the result of quantum measurements, but all the results can be obtained without invoking the collapse concept. There are five main schools of thought on what constitutes a quantum measurement.

- The Copenhagen interpretation is that, on being experimentally observed, the quantum state undergoes an inexplicable collapse to one of the allowed eigenstates; this collapse is not described by the Schrödinger equation. More technically, the state vector collapsing to one of its allowed eigenstates is the detection of the quantum state by one of the projection operators that constitute the measuring device [10].
- All measurements of a quantum system invariably need a device that greatly amplifies the signal from the state vector. Consider using a Geiger counter for measuring the emission of charged particles from a radioactive material. The Geiger counter has an internal high voltage such that the presence of a single charged particle causes a cascade of electrons from the detector's material, releasing up to 10^7–10^8 electrons. The recording of the cascade of electrons is the signal indicating the presence of a charged particle.

  There is a view that a large detector—consisting of $\approx 10^8$ degrees of freedom and obeying the laws of quantum mechanics—creates decoherence by its uncontrollable interactions with the quantum entity and leads to an observed classical state [33].
- In the "many-world" interpretation of quantum mechanics, every measurement is thought to bifurcate the state vector into new branches, depending on the outcome of the measurement. In other words, if one measures a superposed state, such as $\psi = \psi_u + \psi_d$, then sometimes one observes ψ_u and at other times one observes ψ_d. According to this interpretation, an observation leads to two

possible Universes, the first Universe when the state vector ψ_u is observed and the other Universe when the state vector ψ_d is observed.

Each Universe is taken to be equally real, and the subsequent temporal evolution of the state vector takes place—determined by the Schrödinger equation—in the Universe that has been chosen by the experiment, until the next experiment is performed [8].

- Another approach is to introduce a random (fluctuating) classical force that acts on macroscopic objects and causes superposed quantum states to continually fluctuate. This random force causes the transempirical superposed states to rapidly evolve into empirical states that are governed by classical probability.
- There is view that a measurement does not take place in any apparatus, which is in any case governed by quantum principles, but instead, a measurement occurs only when a consciousness-like entity becomes aware of the measurement. In this view, the human *brain* is thought to be described by quantum mechanics, but the human *mind*—taken to be an exemplar of consciousness—is thought to be outside the workings of physical laws. A "recording" of the result of a measurement by the human mind causes an irreversible change by collapsing the state vector and brings the process of measurement to a completion [12].
- My own view, and the one taken in this book, is close to the Copenhagen interpretation that the Schrödinger equation by itself *cannot* explain the process of measurement. Measurement is a projection of the trans-empirical quantum state onto a unique empirical reading of a device; this projection, namely, the collapse of the quantum state, is an assumption that needs to be made in addition to the Schrödinger equation. A dynamical process is needed to make this projection and which would explain the state vector collapse; this mechanism needs to operate on the interface of trans-empirical and empirical domains.

  A *new* equation is required that combines aspects of the Schrödinger equation, which is operational only when the quantum entity exists as a trans-empirical state vector, with another equation that represents the process of measurement. This (new) equation needs to be applied to the quantum state every time an experiment is carried out (on the quantum state) and operates at the interface of the trans-empirical and the empirical domains—causing the quantum state to make a transition from its trans-empirical state to its empirical manifestation.

9.12 Summary

To extract information from the state vector and about the observables, repeated measurements have to be performed. For a given state vector, the degree of freedom is of central significance; although the degree of freedom can have many mathematical representations, for a given problem there is usually a natural choice.

To measure the properties of an observable operator $\mathcal{O}$, a given state vector has to be expressed as a superposition of the eigenstates of $\mathcal{O}$; furthermore, an experiment

has to be designed to measure the eigenvalues and eigenstates of $\mathcal{O}$ to ascertain the value of $\mathcal{O}$ for the state vector. The experiment needs to entangle the detector's state vector with that of the quantum state being observed.

What is remarkable in the physics of measurement is the fundamental role played by the quantum states in Hilbert space. In particular, the existence of entangled states is crucial in coupling the macroscopic experimental device to the microscopic quantum entity.

Measurement lies at the heart of quantum mechanics, but paradoxically, even after over 100 years of Planck's quantum postulate, the process of measurement still remains an enigma. Although the *procedures* for carrying out measurements on a quantum state are garden variety, well understood, and carried out every day, the *theoretical* understanding of measurement is far from clear.

The device can be thought of as a quantum mechanical system in its own right; the combined state vector of the quantum system (which is being observed) and the device evolves according to the Schrödinger equation and *never* undergoes any collapse. So where is the collapse? The fundamental conundrum in quantum measurement is the *mechanism* by which the state vector collapses, or what is the same thing, how does the trans-empirical quantum state "choose" a particular empirical outcome when observed by the projection operators.

All reasoning so far regarding the collapse of the state vector, which in essence is an irreversible change in the combined quantum entity-detector, rests on the fact that the detector is a large system and hence undergoes decoherence and which in turn induces decoherence on the quantum entity. It has been proposed that any collection of atoms larger than 10^8, the number of electrons that cascade in a Geiger counter, can function as a quantum mechanical detector [6].

The fundamental *indeterminacy* of the quantum degree of freedom is reflected in the fact that when a quantum state $|\chi\rangle$ is measured by a process of measurement using projector operators $|D_n\rangle\langle D_n|$, it is *impossible* to predict which projector will detect the state vector's collapse; all one can say is that there is a certain probability of the collapse occurring at a particular projector operator.

There are some "pragmatic" physicists who, focusing on the operational side of measurements, consider theoretical questions regarding "what is a measurement" as being unimportant. But the fact remains that the process of measurement is currently outside the Schrödinger equation and points to a need for either a deeper understanding of quantum mechanics or for a new theory that goes beyond quantum mechanics.

The Stern–Gerlach Experiment

10

The discussion in Chap. 9 shows that the theory of measurement in quantum mechanics is a complex subject. Although widely studied, there is, however, no agreement as to what is the crux of a quantum measurement. Given the central importance of measurements in quantum mechanics, this chapter studies the *Stern–Gerlach experiment*, which is one of the few experiments that can be examined in great detail and can help to further our understanding of the subtleties of quantum measurements.

Spin is a quantum degree of freedom with only two possible outcomes for its z-component and is one of the simplest quantum system; the Stern–Gerlach experiment measures σ_z, the z-component of the spin of an electron.

The mathematics of this experiment is comparatively simple, leaving us to focus on the physics of measurement. The Stern–Gerlach experiment is modeled by a simplified Hamiltonian proposed by Gottfried and Yan [15]; the advantage of this Hamiltonian is that all the results can be obtained exactly. The measurement process is studied by using the Schrödinger equation to evolve the state vector of the electron through all the stages of the experiment.

This chapter is more technical than others because a qualitative discussion cannot address the controversies that surround the problem of measurement; instead, there is a need to quantitatively study the problem so that one makes generalizations that have a precise mathematical basis.

10.1 The Experiment

An electron consists of its position and spin degrees of freedom, with the term "electron" referring to its degrees of freedom. The Stern–Gerlach experiment measures the z-component of the electron's spin, which are eigenvalues of the eigenstates of σ_z, given in Sect. 8.1. For ease of reference, the eigenvalues and eigenstates of σ_z are given below:

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8_10,

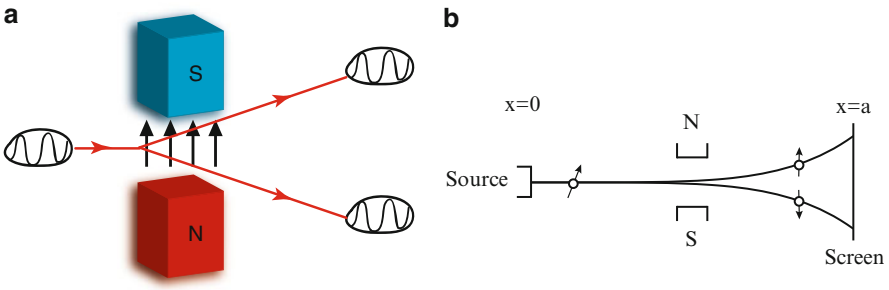


Fig. 10.1 The Stern–Gerlach experiment. The *arrows* pointing up represent an inhomogeneous magnet field. (**a**) The incoming electron is represented by a wave packet. (**b**) The incoming electron's spin is a superposed state; after crossing the magnetic field, the up and down spin eigenstates are well separated (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

$$|\text{up}\rangle = |+\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \qquad |\text{down}\rangle = |-\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \sigma_z|\pm\rangle = \pm\frac{\hbar}{2}|\pm\rangle$$

where

$$\sigma_z = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

The trans-empirical superposed state vector for spin $1/2$, from (8.3), is given by

$$|\Psi\rangle = \alpha|+\rangle + \beta|-\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}; \quad \langle\Psi| = \begin{bmatrix} \alpha^* & \beta^* \end{bmatrix} \tag{10.1}$$

Figure 10.1 shows a schematic representation of the process of measurement: The incoming state vector is an electron wave packet traveling in an inhomogeneous magnetic field, as shown in Fig. 10.1a; the spin is in a superposed state $|\Psi\rangle$ of the up and down spin states. The magnet field separates the up and down spin states, and the electron's state vector finally hits the detector, which is a screen that records the collapse of the electron's state vector, and in effect yielding the position of the electron at screen, as shown in Fig. 10.1b.

The intuitive concept of the Stern–Gerlach experiment is the following. The state vector subjected to an inhomogeneous magnet field propagates (on the average) along two paths: For "up" spin, case the state vector goes upwards and for the "down" spin case, goes downwards. Measuring the trajectory of the electron's wave packet is equivalent to determining the electron's spin. This is the essence of the Stern–Gerlach experiment, and the remaining sections express this intuitive idea in the mathematical framework of quantum mechanics.

The Stern–Gerlach experiment has the following arrangement:

- The electrons travel along the x-axis and with $z = 0$. The electron's state vector is a wave packet with a well-localized position.
- An inhomogeneous magnetic field points in the z-direction, with the field getting stronger for increasing z. The electron is in the magnetic field for a distance a along the horizontal direction.
- After leaving the magnetic field, the electron's state vector is found by the detectors at the screen to be either above or below $z = 0$.
- The final vertical position z of the electron is measured at horizontal position $x = a$, as shown in Fig. 10.1, where a screen is placed.
- The distance a that the electron travels in the magnetic field is adjusted so that there is a clear separation of the up and down paths of the electron. In particular, the separation of the two wave packets is much greater than the spread of the wave packets.
- From the position measurement of the electron's state vector at the screen, it is concluded whether the electron's spin is pointing up or down.
- The probabilities P_{u} and P_{d} are obtained by repeating the experiment many times and counting how many times the electron's state vector is found at the screen in the up or down positions.

In more general terms, the electron has two distinct degrees of freedom, namely, its position degrees of freedom x, y, z and its spin degree of freedom. The Stern–Gerlach experiment uses the electron's position degree of freedom z as the macroscopic variable that is measured in the laboratory and plays the role of the macroscopic counter reader.

The microscopic values of the z-component of spin, namely, $\pm\frac{\hbar}{2}$, are inferred from the position readings. In essence, due to the effect of the magnetic field, once the two trajectories for the "up" and "down" states have separated out more than the spread of the incident wave packet, the position of the final state yields the value of the observed eigenvalue of the spin operator σ_z.

10.2 Classical and Quantum Predictions

As shown in Fig. 10.1, electrons produced at the source are collimated into an electron beam that travels across an inhomogeneous magnetic field and is finally observed on a screen. The electrons coming out of the source have their spins in an arbitrary superposed state. In the classical picture, the spins are pointing in arbitrary directions. Classical physics predicts that the position of the electron on the screen will lie *everywhere* between the maximum up position, for a spin pointing up, and the minimum down position for a spin pointing down. Classical physics predicts that the observed electrons will *lie on a continuous line*, as shown in Fig. 10.2a.

The quantum mechanical solution for a spin moving in an inhomogeneous magnetic field is dramatically unlike the classical result. The electron coming out of the source has a state vector given by the quantum superposition of the up and down spin states and given by (8.3), namely,

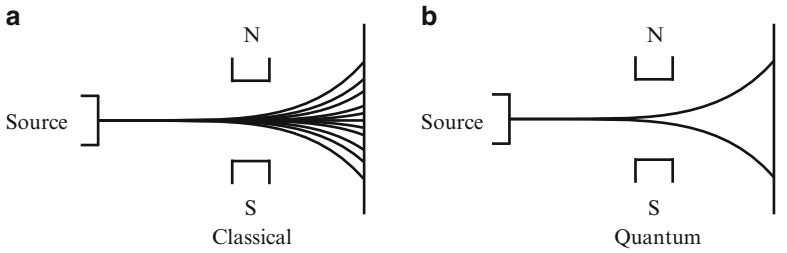


Fig. 10.2 (**a**) Classical prediction for the Stern–Gerlach experiment. (**b**) Quantum prediction for the Stern–Gerlach experiment (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

$$|\Psi\rangle = \alpha|+\rangle + \beta|-\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}; \qquad |\alpha|^2 + |\beta|^2 = 1$$

Since the electrons coming from the source have an arbitrary orientation, the coefficients α, β can take all the allowed values on the Bloch sphere discussed in Sect. 4.4. The electron is in a trans-empirical state, simultaneously being in both the up and down state, and *each* electron has an indeterminate and trans-empirical path, simultaneously propagating along two different possible paths, as shown in Fig. 10.2b and later in Fig. 10.5b.

Quantum superposition predicts that every experiment will only obtain the eigenvalues of the components of the superposed state. The analysis for the quantum superposition of the spin 1/2 degree of freedom is discussed in Sect. 8.1. For the spin 1/2 case, every measurement of σ_z, the z-component of spin, will result in *either* the up value of the spin *or* the down value, namely, $\hbar/2$ or $-\hbar/2$, and with *no other value* for the spin (that is in between the up and down values), as shown in Fig. 10.2b; this follows from the principle of superposition discussed in Sect. 8.1.

In other words, on the screen, all the electrons will be observed at *only two points*, either in the up position for the case of σ_z equal to $\hbar/2$ or in the down position for the case of σ_z equal to $-\hbar/2$, and nowhere else.

The predictions of classical and quantum mechanics are in stark contrast and shown in Fig. 10.2a, b. Experiments confirm the prediction of quantum superposition.

The average value of the z-component of spin, as in (8.4), is given by

$$\langle\Psi|\sigma_z|\Psi\rangle = \frac{\hbar}{2}\left[|\alpha|^2 - |\beta|^2\right]$$

10.3 The Stern–Gerlach Hamiltonian

The electron's spin is measured by greatly amplifying the (different) effect of the experimental device on the two spin states. The process of amplification is realized

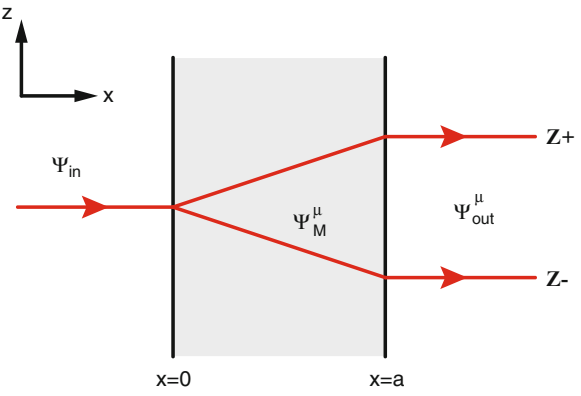


Fig. 10.3 Two possible trajectories of the wave packet under the influence of the magnetic field: cross-sectional view (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

by the electron's spin interacting with the apparatus that is represented by a quantum Hamiltonian.

The process of measuring the electron's spin is modeled by a linearized version of the Stern–Gerlach Hamiltonian—so that all the steps can be carried out exactly. For ease of notation, we set $\hbar = 1$ and we will restore its value if necessary. Following Gottfried and Yan [15], the Stern–Gerlach Hamiltonian is given by

$$\begin{aligned} H &= -\frac{1}{2m}\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right] - i\alpha\sigma_z f(x)\frac{\partial}{\partial z} \\ \sigma_z &= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}; \qquad f(x) = \begin{cases} 1 & x \in [0,a] \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{10.2}$$

where the coupling of the magnetic field to the spin is given by α.[1] The term $\partial/\partial z$ reflects the increasing strength of the magnetic field with increasing horizontal distance z. The function $f(x)$ reflects the fact that the inhomogeneous magnetic field is nonzero only for the horizontal distance $x \in [0,a]$.

The classical solution to this Hamiltonian has two possible trajectories, with the classical spin initially pointing up or down and with the z-coordinate of the particle linearly rising or falling in the interval $x \in [0,a]$, as is shown in Fig. 10.3, for the two spin cases, respectively.

Consider a steady flux of electrons going through the apparatus, one by one. The expected (average) position of the electron follows a time-dependent trajectory, being incident as a free particle on the magnetic field from the left, at $x = 0$, propagating in the magnetic field until the point of $x = a$, and then propagating as a free particle along the x-axis for $x > a$. The time-dependent Schrödinger equation, for $\mathbf{r} = (x,y,z)$, is given in (2.4) ($\hbar = 1$)

[1] $\alpha = \mu_e B$, where μ_e is the magnetic moment of the electron and B has dimension of the magnetic field.

$$i\frac{\partial \psi(t,\mathbf{r})}{\partial t} = H\psi(t,\mathbf{r}) \tag{10.3}$$

and needs to be solved for the Stern–Gerlach system.

The time-dependent state function $\psi(t,x,y,z)$ is expanded in the basis provided by the stationary eigenstates of H. Let

$$\xi_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \qquad \xi_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \qquad \sigma_z \xi_\mu = \frac{\mu}{2}\xi_\mu; \qquad \mu = \pm 1$$

The energy eigenstates of the Hamiltonian H [given in (10.2)] are labeled by $\mathbf{p} = (p_x, p_y, p_z)$ and μ, shown in Fig. 10.3, and are given by

$$H\psi_E^\mu(\mathbf{r};\mathbf{p})\xi_\mu = E(\mu,\mathbf{p})\psi_E^\mu(\mathbf{r};\mathbf{p})\xi_\mu; \quad \mu = \pm 1 \tag{10.4}$$

$$\psi_E^\mu(\mathbf{r};\mathbf{p}) = \begin{cases} \Psi_{\text{in}}(\mathbf{r};\mathbf{p}); & x < 0 \\ \Psi_{\text{M}}^\mu(\mathbf{r};\mathbf{p}); & 0 < x < a \\ \Psi_{\text{out}}^\mu(\mathbf{r};\mathbf{p}); & x > a \end{cases} \tag{10.5}$$

For $x < a$ and $x > a$, there is no potential, and hence the eigenstates are plane waves. It can be verified that the energy eigenstates are given by

$$\begin{aligned} \Psi_{\text{in}}(\mathbf{r}) &= \exp \mathrm{i}\{xp_x + yp_y + zp_z\}; \quad x < 0 \\ \Psi_{\text{M}}^\mu(\mathbf{r}) &= \exp \mathrm{i}\{x\tilde{p}_x + yp_y + zp_z\}; \quad 0 < x < a \\ \Psi_{\text{out}}^\mu(\mathbf{r}) &= \exp \mathrm{i}\{xp_x + yp_y + (z - \bar{z}_\mu)p_z\}; \quad x > a \end{aligned} \tag{10.6}$$

where

$$\begin{aligned} E(\mu,\mathbf{p}) &= \frac{1}{2m}[p_x^2 + p_y^2 + p_z^2] = \frac{1}{2m}[\tilde{p}_x^2 + p_y^2 + p_z^2] + \alpha\mu p_z \\ \tilde{p}_x &= \sqrt{p_x^2 - 2m\alpha\mu p_z} \end{aligned}$$

Note that the incoming eigenfunction $\Psi_{\text{in}}(\mathbf{r})$ does not depend on μ. The effect of the magnetic field is encoded in the eigenstate $\Psi_{\text{M}}^\mu(\mathbf{r})$, with its momentum $\tilde{p}_x$ being modified from the free particle case. The constant phase $\bar{z}_\mu$ in $\Psi_{\text{out}}^\mu(\mathbf{r})$ is fixed by requiring continuity of the state function as a function of t and $\mathbf{r}$.[2]

An arbitrary time-dependent state vector that satisfies the Schrödinger equation given in (10.3) has the following expansion:

[2]The value of $\bar{z}_\mu$ is given in (10.17).

$$\psi(t,\mathbf{r}) = \sum_{\mu=\pm 1} c_\mu \xi_\mu \int \frac{\mathrm{d}^3 p}{(2\pi)^3} g(\mathbf{p}) \psi_E^\mu(\mathbf{r};\mathbf{p}) \mathrm{e}^{-\mathrm{i}tE(\mu,\mathbf{p})/2m} \tag{10.7}$$

The coefficients c_μ and $g(\mathbf{p})$ are fixed by the initial condition at initial time $t = 0$.

10.4 Electron's Time Evolution

To analyze the electron's propagation in time, let t_* be the time required by the electron to travel across the region of the magnetic field, from $x = 0$ to $x = a$. The following notation is used for the state vector and for the different intervals of time and regions of space.

$$\psi(t,\mathbf{r}) = \begin{cases} \psi_{\mathrm{in}}(t,\mathbf{r}); & t < 0;\ x < 0 \\ \psi_{\mathrm{M}}(t,\mathbf{r}); & 0 < t < t_*;\ 0 < x < a \\ \psi_{\mathrm{out}}(t,\mathbf{r}); & t > t_*;\ x > a \end{cases} \tag{10.8}$$

An initial incoming electron wave packet with energy $E = \mathbf{p}^2/2m$ is given by

$$\psi_{\mathrm{in}}(t,\mathbf{r}) = \chi(t,\mathbf{r}) \sum_{\mu=\pm 1} c_\mu \xi_\mu;\ x < 0$$

$$\chi(t,\mathbf{r}) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} g(\mathbf{p}) \mathrm{e}^{\mathrm{i}xp_x + \mathrm{i}yp_y + zp_z} \mathrm{e}^{-\mathrm{i}t\mathbf{p}^2/2m} \tag{10.9}$$

Note the position and spin degrees of freedom for $\psi_{\mathrm{in}}(t,\mathbf{r})$ are *not entangled* since they are in the form of a product: The state vectors for the spin and position degrees of freedom of freedom are completely *factorized* in $\psi_{\mathrm{in}}(t,\mathbf{r})$.

As indicated in Fig. 10.1, the incident state function needs to be a wave packet that is well localized in space so that the detector can observe the motion of the center of the wave packet as it traverses the magnetic field—with the motion of the center of the wave packet yielding information about the spin of the electron.

To create the incident wave packet that is well localized in space, the following function $g(\mathbf{p})$ is chosen:

$$g(\mathbf{p}) = \left(\frac{2\pi}{\beta^2}\right)^{3/2} \mathrm{e}^{-\frac{1}{2\beta^2}(\mathbf{p}-\mathbf{K})^2};\quad \mathbf{K} = (K,0,0) \tag{10.10}$$

The wave packet is chosen to be traveling in the x-direction. Choosing $\beta \ll K$ ensures that the momentum of the wave packet is peaked around $\mathbf{K}$ and does not spread as it traverses the magnetic field. The average momentum K is chosen to be large enough so that there is *no reflected* component of the state vector at $x = 0$. For this choice of $g(\mathbf{p})$, at $t = 0$, the wave packet is well localized around $x = y = z = 0$.

Wave Packet Inside the Magnetic Field: Formation of Entanglement

For $0 < t < t_*$, the wave packet is inside the magnetic field and yields the following time-dependent state function:

$$\begin{aligned}\psi_{\mathrm{M}}(t,\mathbf{r}) &= \sum_{\mu} c_\mu \xi_\mu \chi_\mu(t,\mathbf{r}); \quad x > 0;\ 0 < t < t_* \\ \chi_\mu(t,\mathbf{r}) &= \int \frac{\mathrm{d}^3 p}{(2\pi)^3} g(\mathbf{p}) \Psi_{\mathrm{M}}^{\mu}(\mathbf{r}) \mathrm{e}^{-\mathrm{i}t\mathbf{p}^2/2m};\ \ 0 < x < a \\ &= \int \frac{\mathrm{d}^3 p}{(2\pi)^3} g(\mathbf{p}) \mathrm{e}^{\mathrm{i}x\tilde{p}_x + \mathrm{i}(yp_y + zp_z)} \mathrm{e}^{-\mathrm{i}t\mathbf{p}^2/2m} \end{aligned} \tag{10.11}$$

As the electron propagates along the x-axis, the position and spin degrees of freedom for $\psi_{\mathrm{M}}(t,\mathbf{r})$ become entangled. We study how this entanglement is brought about by the interaction Hamiltonian H given in (10.2).

The momentum of the incident wave packet, due to $g(\mathbf{p})$, is peaked at K along the x-direction; hence, in the momentum integral given in (10.11), p_x can be expanded about the momentum K and yields for $\tilde{p}_x$ the following:

$$\begin{aligned}\tilde{p}_x &= \sqrt{p_x^2 - 2m\alpha\mu p_z} \simeq p_x - m\mu\alpha p_z/p_x \simeq p_x - \mu\alpha p_z/v \\ v &= p_x/m \simeq K/m \end{aligned} \tag{10.12}$$

Equation (10.12) yields, in (10.11), the following:

$$\mathrm{e}^{\mathrm{i}x\tilde{p}_x} \mathrm{e}^{\mathrm{i}zp_z} \simeq \mathrm{e}^{\mathrm{i}xp_x} \mathrm{e}^{\mathrm{i}p_z(z - \mu\alpha x/v)} \tag{10.13}$$

Equation (10.13) is the key to understanding the mechanism of entanglement. Due to the magnetic field, the propagation in the z-direction now has a contribution $\mu\alpha x/v$; furthermore, the propagation is now *entangled* with the spin of the electron, since the trajectory for $\mu = 1$ separates out from the one for $\mu = -1$; in other words, the position and spin degrees of freedom become entangled due to the interaction of the electron's degrees of freedom with the magnetic field.

Performing the Gaussian integrations in (10.11) using (10.10) and (10.13) yields, for $0 < x < a$ and $0 < t < t_*$, the following[3]:

$$\begin{aligned}\chi_\mu(t,\mathbf{r}) &= \int \frac{\mathrm{d}^3 p}{(2\pi)^3} \left(\frac{2\pi}{\beta^2}\right)^{3/2} \mathrm{e}^{-\frac{1}{2\beta^2}(\mathbf{p}-\mathbf{K})^2} \mathrm{e}^{\mathrm{i}xp_x + \mathrm{i}yp_y} \mathrm{e}^{\mathrm{i}p_z(z - \mu\alpha x/v)} \mathrm{e}^{-\mathrm{i}t\mathbf{p}^2/2m} \\ &= \mathcal{N}' \exp\left\{-\frac{m\beta^2}{2(m + \mathrm{i}t\beta^2)} \left[y^2 + \left(z - \frac{\mu\alpha x}{v}\right)^2\right]\right\} \mathrm{e}^{\mathrm{i}Kx} \int \mathrm{d}p \mathrm{e}^{-\frac{1}{2\beta^2}p^2 - \frac{\mathrm{i}t}{2m}(p+K)^2} \mathrm{e}^{\mathrm{i}px} \end{aligned}$$

[3] $\mathcal{N}', \mathcal{N}$ are normalization constants.

$$= \mathcal{N} \exp\left\{ -\frac{m\beta^2}{2(m+\mathrm{i}t\beta^2)} \left[y^2 + \left(z - \frac{\mu\alpha x}{v} \right)^2 + (x - vt)^2 \right] \right\} \mathrm{e}^{\mathrm{i}Kx} \mathrm{e}^{-\frac{\mathrm{i}t}{2m}K^2} \tag{10.14}$$

The average positions are given by (10.14) and yield, to leading order,

$$E_\chi[x] = \int_0^a \mathrm{d}x \int dy\, \mathrm{d}z\, x\, |\chi_\mu(t,\mathbf{r})|^2 \simeq vt$$
$$E_\chi[z] = \int_0^a \mathrm{d}x \int \mathrm{d}y\, \mathrm{d}z\, z\, |\chi_\mu(t,\mathbf{r})|^2 \simeq \frac{\mu\alpha}{v} E_\chi[x] = \mu\alpha t$$

The trajectory of the wave packet is obtained by the expectation value of the position of the wave packet; for $\mathbf{r} = (x,y,z)$, the equations above yield, to leading order, the expected position, given by $\bar{\mathbf{r}}_\mu(t)$, as follows:

$$\bar{\mathbf{r}}_\mu(t) = \langle \chi(t) | \mathbf{r} | \chi(t) \rangle = E_\chi[\mathbf{r}] = (vt, 0, \mu\alpha t) \tag{10.15}$$

The trajectory of the wave packet is the motion of its expected position $\bar{\mathbf{r}}_\mu(t)$. For $0 < x < a$, the trajectories $\bar{\mathbf{r}}_+(t)$ and $\bar{\mathbf{r}}_-(t)$, given in Fig. 10.3, are the upward and downward sloping straight lines, as expected from (10.15).

Wave Packet After Crossing the Magnetic Field: Entangled State

The domain of the magnetic field, defined by $0 < x < a$, is chosen so that—on crossing the distance a—the separation of the electron wave packet for $\mu = \pm 1$ is much greater than its spread, schematically shown in Fig. 10.1.

On reaching $x = a$ in time $t_* = a/v$, the wave packet is out of the magnetic field, and there is no longer any average motion in the z-direction. The trajectory for $t > t_* = a/v$ is determined by the state function $\Psi^\mu_{\mathrm{out}}(\mathbf{r})$ given by (10.6), which together with (10.5) yields

$$\psi_{\mathrm{out}}(t,\mathbf{r}) = \sum_\mu c_\mu \xi_\mu \zeta_\mu(t,\mathbf{r}); \;\; x > a; \;\; t > t_* = a/v$$
$$\zeta_\mu(t,\mathbf{r}) = \int \frac{\mathrm{d}^3 p}{(2\pi)^3} g(\mathbf{p}) \Psi^\mu_{\mathrm{out}}(\mathbf{r}) \mathrm{e}^{-\mathrm{i}t\mathbf{p}^2/2m}$$
$$= \int \frac{\mathrm{d}^3 p}{(2\pi)^3} g(\mathbf{p}) \exp \mathrm{i}\{xp_x + yp_y + (z - \bar{z}_\mu)p_z\} \mathrm{e}^{-\mathrm{i}t\mathbf{p}^2/2m} \tag{10.16}$$

Recall the constant phase $\bar{z}_\mu$ is fixed by the requirement that the state function $\psi(t,\mathbf{r})$ is continuous at $x=a$ and for time $t_*=a/v$; hence, from (10.14),

$$\bar{z}_\mu = \frac{\mu\alpha a}{v} \tag{10.17}$$

Similar to the derivation given in (10.14), in the approximation that $K \gg \beta$, from (10.16), the state function for $x>a$ and $t>t_*$ is given by

$$\zeta_\mu(t,\mathbf{r}) = \mathcal{N} \mathrm{e}^{-\frac{m\beta^2}{2(m+\mathrm{i}t\beta^2)}[y^2+(z-\frac{\mu\alpha a}{v})^2+(x-vt)^2]} \mathrm{e}^{\mathrm{i}Kx} \mathrm{e}^{-\frac{\mathrm{i}t}{2m}K^2} \tag{10.18}$$

$$\Rightarrow \bar{\mathbf{r}}_\mu(t) = \int_a^\infty \mathrm{d}x \int \mathrm{d}y\,\mathrm{d}z\, \mathbf{r}\, |\zeta_\mu(t,\mathbf{r})|^2 \simeq \left(vt, 0, \frac{\mu\alpha a}{v}\right) \tag{10.19}$$

Note that the two state functions ζ_μ are orthogonal, that is, $\langle\zeta_\nu|\zeta_\mu\rangle = \delta_{\mu-\nu}$.

The trajectory of the wave packet is given by the motion of its expected position, given by $\bar{\mathbf{r}}_\mu(t)$; after time t_*, the trajectories $\bar{\mathbf{r}}_+(t)$ and $\bar{\mathbf{r}}_-(t)$ *travel in a straight line* parallel to the x-axis and at a height of $\bar{z}_\pm$—and are shown, in Fig. 10.3 for $x>a$, as two lines parallel to the x-axis.

10.5 Entanglement of Spin and Device

To understand the measurement process, recall from (10.9) that the incident state function is given by

$$\psi_{\text{in}}(t,\mathbf{r}) = \chi(t,\mathbf{r}) \sum_{\mu=\pm1} c_\mu \xi_\mu; \quad x<0;\, t<0$$

The incident state function is a product state—of the space state function $\chi(t,\mathbf{r})$ multiplied by the superposed state of the spins $\sum_\mu c_\mu\xi_\mu$. There is no correlation between the position of the electrons and its spin; measuring the position of the electron yields no information about the spin of the electron.

The Hamiltonian given in (10.2) introduces an interaction of the spin with the external magnetic field and creates a macroscopic amplification between the two microscopic spin eigenstates. The up spin eigenstate has an average trajectory, given by $\bar{\mathbf{r}}_+(t)$, that is clearly different from the average trajectory of the down spin eigenstate $\bar{\mathbf{r}}_+(t)$.

After passing through the magnet, the electron's position degree of freedom becomes *entangled* with its spin degree of freedom, with the deflection of the electron in the z-direction depending on its spin state defined by μ. More precisely, from (10.16),

$$\psi_{\text{out}}(t,\mathbf{r}) = \sum_\mu c_\mu \xi_\mu \zeta_\mu(t,\mathbf{r}); \quad x>a;\, t>t_*=a/v$$

$$\Rightarrow |\psi_{\text{out}}\rangle = \sum_{\mu} c_{\mu} |\xi_{\mu}\rangle |\zeta_{\mu}\rangle; \qquad \langle \zeta_{\nu} | \zeta_{\mu} \rangle = \delta_{\mu - \nu} \tag{10.20}$$

On leaving the magnetic field, the electron is in an entangled state, for which the electron takes the up and down paths that are *exactly correlated* with the state of the spin: Measuring the position of the electron is tantamount to measuring its spin. Or put differently, due to entanglement, immediately after the position measurement of $\bar{z}_{\mu}$, the spin is certain to be in a spin eigenstate ξ_{μ}; if an independent measurement of spin is carried out immediately *after* the measurement of the position yields $\bar{z}_{\mu}$, then the spin is certain to be found in the state ξ_{μ}.

Hence, a measurement of position in effect is also a measurement of spin and can be represented by the projection to state function $|\xi_{\mu}\rangle|\zeta_{\mu}\rangle$; (10.20) then yields

$$|\langle \zeta_{\mu} | \langle \xi_{\mu} | \psi_{\text{out}} \rangle|^2 = |c_{\mu}|^2$$

Equivalently, note that the Stern–Gerlach experiment does not directly observe the spin states of the electron. Hence, a measurement of only the position degree of freedom results in a partial trace over the spin degrees of freedom and yields the following reduced density matrix:

$$\rho_{\text{R}} = tr_{spin}\left(|\psi_{\text{out}}\rangle\langle\psi_{\text{out}}|\right) = \sum_{\mu} |c_{\mu}|^2 |\zeta_{\mu}\rangle\langle\zeta_{\mu}| \tag{10.21}$$

The reduced density matrix ρ_{R} represents a classical random system. After it crosses the magnetic field, the probability of finding the electron in device state $|\zeta_{+}\rangle$, with the pointer at position $\bar{z}_{+}$ is P_{u} and, similarly, the probability of finding the device in state $|\zeta_{-}\rangle$, with pointer at position $\bar{z}_{-}$, is P_{d}; (10.21) then states that

$$P_{\text{u}} = |c_{+}|^2; \qquad P_{\text{d}} = |c_{-}|^2$$

The experiment determines the average value of the z-component of spin, as in (8.4), and is given by

$$\langle \Psi | \sigma_z | \Psi \rangle = \frac{\hbar}{2}\left[|c_{+}|^2 - |c_{-}|^2\right] = \frac{\hbar}{2}\left[P_{\text{u}} - P_{\text{d}}\right]$$

10.6 Summary of Spin Measurement

We recapitulate the process of measurement to highlight its conceptual underpinnings and connect the specific example of the spin measurement to the general features of a quantum measurement discussed in Sects. 9.6–9.8.

The Stern–Gerlach experiment measures the spin magnetic moment, in short the spin, of the quantum spin degree of freedom. The measurement is carried out by entangling the spin degree of freedom with the degrees of freedom of the

experimental device. The z-coordinate of the electron's position degree of freedom is, in fact, a representation of the experimental device, since the z-position of the electron is what the device measures. Hence, the z-coordinate degree of freedom of the electron plays the role of the degree of freedom of the device.

The counterposition of the device is the point on the screen at which the electron's state vector is detected. Hence, the position of the electron on the screen is the degree of freedom of the device. The state of the device has the following three values:

- The counterposition $z = 0$ indicates the neutral position of the counter and contains no information about the spin of the system and is represented by state $|D_0\rangle$.
- The counterposition $z = \bar{z}_+ = \alpha a/v$ shows that the spin is in the up state and is represented by state $|D_+\rangle$.
- The counterposition $z = \bar{z}_- = -\alpha a/v$ shows that the spin is in the down state and is represented by state $|D_-\rangle$.

The three states of the device, represented by the state vectors $|D_0\rangle, |D_+\rangle, |D_-\rangle$, are given by

$$|D_0\rangle = \chi(t, \mathbf{r}); \qquad |D_+\rangle = \zeta_+(t, \mathbf{r}); \qquad |D_+\rangle = \zeta_-(t, \mathbf{r})$$

The measurement process is constituted by the following stages:

- The initial quantum state is prepared at the source to be in a superposed state of the spin degree of freedom. The initial quantum state and device state are in a joint product state.

$$|\psi_{\text{in}}\rangle = |D_0\rangle \left(c_1|\xi_1\rangle + c_2|\xi_2\rangle\right): \quad \text{device and spin not entangled}$$

- After the interaction of the device and the electron's spin, the final state is an entangled state given by

$$|\psi_{\text{in}}\rangle \to |\psi_{\text{out}}\rangle = c_1|\xi_1\rangle|D_+\rangle + c_2|\xi_2\rangle|D_-\rangle: \quad \text{device and spin } \textit{entangled}$$

- The third stage in measuring the spin of the electron is to perform the measurement by *recording* the quantum state of the electron—hence bringing about an irreversible change in the spin-device system by collapsing the state function. This process yields the mixed density matrix ρ_{M} given by

$$\begin{aligned} \rho &= |\psi_{\text{out}}\rangle\langle\psi_{\text{out}}| \to \text{ Measurement} \to \rho_{\text{M}} \\ \rho_{\text{M}} &= |c_1|^2|\xi_1\rangle\langle\xi_1| \otimes |D_+\rangle\langle D_+| + |c_2|^2|\xi_2\rangle\langle\xi_2| \otimes |D_-\rangle\langle D_-| \end{aligned}$$

- Since the value of the spin degree of freedom is not measured, the Stern–Gerlach in effect performs a partial trace over the spin degree of freedom and yields the reduced density matrix ρ_{R} that gives the final result:

$$\rho_{\mathrm{R}} = \mathrm{tr}_{\mathrm{spin}}\left(\rho_{\mathrm{M}}\right) = |c_1|^2 |D_+\rangle\langle D_+| + |c_2|^2 |D_-\rangle\langle D_-|$$

10.7 Irreversibility and Collapse of State Vector

As discussed in Chap. 9, measurement involves four steps, namely, preparation, amplification, entanglement, and irreversibility. The process of measurement is brought to a *closure* by recording the outcome of the measurement, and this process of recording the result brings up about an irreversible change in the detector as well as in the quantum system being observed.

Recall for the Stern–Gerlach experiment, the applied magnetic field brought about an *amplification* of the up and down quantum states of the electron's spin by separating these states into macroscopically separated electron (average) trajectories. Furthermore, the electron's position became *entangled* with its spin degree of freedom, thus allowing for the unambiguous determination of the state of the spin degree of freedom by measuring only the position of the electron.

However, both amplification and entanglement do not bring an irreversible change in the system. As shown schematically in Fig. 10.4, one can apply a reversed magnetic field to the entangled state of the electron and disentangle it and deamplify the separation of the quantum states so that the original state of the electron is restored. According to Wigner [12], the passage of the electron through the magnetic field is not a complete measurement, and it is only when the position of the electron is *recorded*, for example, by a photographic plate as in Fig. 9.1, that an *irreversible* change is made and the process of measurement is completed.

Registering the electron's position is a process that *causes* a collapse of the state vector and brings about an irreversible change in the system—called decoherence. The concept of quantum entropy as discussed in (6.34) provides an appropriate mathematical description of irreversibility in quantum mechanics.

This view of Wigner has been contested by some physicists pointing to the necessity replacing the classical magnetic field with the quantized electromagnetic field [31]; suffice to say, the relevant point in this discussion is that an irreversible change needs to be made for completing the measurement process. If the preparation of the state vector or its subsequent propagation through the magnetic field brings about such an irreversible change, then it proves Wigner's point.

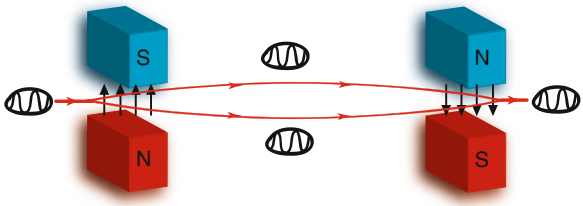


Fig. 10.4 Reversing amplification and entanglement (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

10.8 Interpretation of Spin Measurement

The empirical and trans-empirical interpretation of the Stern–Gerlach experiment is given in Fig. 10.5.

Figure 10.5a–c is a representation of the different aspects of the measurement process of the electron's spin. Figure 10.5a is the usual drawing of the experiment indicating both the possible paths for the electron; Fig. 10.5b and c resolves the experiment into two components: the empirical part shown in Fig. 10.5b and the trans-empirical part shown in Fig. 10.5c.

The stages of the measurement process are the following:

- At the "source," the quantum state to be measured is prepared together with the experimental device that performs the measurement and is shown in Fig. 10.5a. $\psi(\mu, \mathbf{r})$ is the state function of the electron that is emitted by the source, where $\mu = \pm 1$ (up, down) is the spin of the electron; $D(z_0)$ is the state function of the detector with z_0 being the initial value of the detector pointer. The preparation results in a product state $\psi(\mu, \mathbf{r}) \otimes D(z)$.
- Figure 10.5b and c is an empirical and trans-empirical interpretation of the experiment. The experimental device is the "screen," which is in the empirical domain, where both the source and screen (detectors) are placed and shown in Fig. 10.5b. The superposed state of the electron's spin is simultaneously in both the up and down states and hence is trans-empirical. Figure 10.5c shows the trans-empirical state of the electron that does not exist in (physical) space but rather exists in Hilbert space.
- During the transit of the electron from the source to the screen, the state vectors of the electron and the device become entangled, as shown in Fig. 9.6; moreover, the difference between the up and down state of the electron's

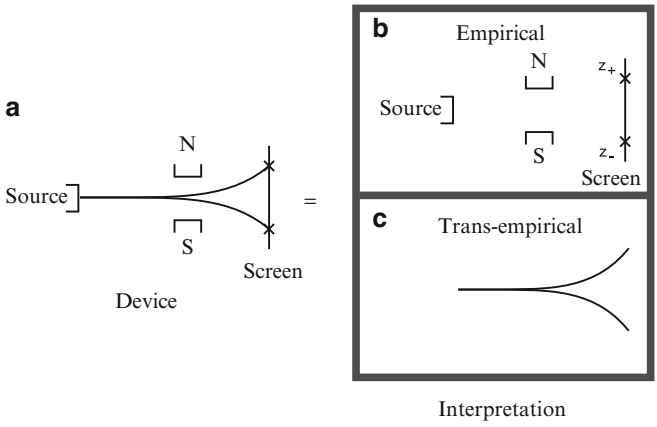


Fig. 10.5 (**a**) Stern–Gerlach experiment. (**b**) Empirical and (**c**) trans-empirical interpretation of the Stern–Gerlach experiment (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

spin is vastly amplified. The trans-empirical entangled state vector is given by $\sum_{s=\pm} c_s \psi(s, \mathbf{r}) \otimes D(z_s)$.

- *Both* the eigenstates arrive at the screen as a trans-empirical superposed state vector.
- When the entangled state vector ψ hits the screen, it makes a discontinuous and irreversible transition from its trans-empirical form ψ to its empirical form $|\psi|^2$. *Only one* of the trans-empirical states is actualized at either z_+ or z_-, with a likelihood given by $|c_+|^2$ or $|c_-|^2$, respectively.
- Suppose the electron's state vector is detected at position z_+ that is shown in Fig. 10.5b in the empirical setup of the experiment. The detector's state vector is put into a definite state $D(z_+)$ when the electron's state vector is detected; due to entanglement, one can conclude that the electron's spin is also put into a definite state $\psi(+, z_+)$.
- Repeating the experiment many times gives the probability for the electron's spin to be in the different possible eigenstates, namely, yields $|c_\pm|^2$.

10.9 Summary

Measurement of the properties of a quantum degree of freedom lies at the heart of quantum mechanics. The measurement of the state vector of a spin 1/2 degree of freedom was studied in great detail, using the Stern–Gerlach experiment, to examine each step in the process of measurement.

A model Hamiltonian was used to obtain the state vector of the electron and evolve it through the experimental apparatus, in particular, from the source of electrons, through the inhomogeneous magnetic field and finally to the screen where the electron's state vector is detected.

The time-dependent state vector yields explicit expressions on how the entanglement of the spin degree of freedom with the position degree of freedom develops due to the interaction of the electron's degrees of freedom with the magnetic field. The state vector also demonstrates how the amplification of the microscopic difference between the up and down spin states is a function of time, with the time spent in the magnetic field determining the degree of macroscopic separation of the two possible paths of the electron.

Recording the collapse of the state vector of the electron on the screen causes the state vector to collapse. This collapse of the state vector has to be *postulated* and brings to a conclusion the Stern–Gerlach experiment.

Measuring only the position of the electron is mathematically realized by a partial trace of the electron's density matrix over the spin degree of freedom and yields the reduced density matrix, which in turn yields the likelihood of finding the spin to be in the up or down state.

The Stern–Gerlach experiment was lastly analyzed to determine the empirical and trans-empirical aspects of the experiment. It was seen that only the preparation of the initial state and the measurement of the final state are empirical events, with the evolution of the electron's state vector, the formation of entanglement, and the

amplification of the physical effect of the spin degree of freedom all being trans-empirical processes.

In summary, the Stern–Gerlach experiment illustrates all four ingredients of a quantum measurement, namely, the preparation of the quantum state, entanglement of the degree of freedom being measured with the device, amplification of the quantum quantity to a macroscopic magnitude, and the irreversible collapse of the state vector.

The Feynman Path Integral

11

The following two independent and equivalent formulations of quantum mechanics have been discussed:

- The Schrödinger equation is a partial differential equation based on the concept of the state vector ψ. It determines the expectation value for the operators of the underlying quantum degree of freedom, discussed in Sect. 5.9.
- Heisenberg's operator formulation in which operators representing observable quantities are evolved using the Heisenberg operator equations, as discussed in Sect. 5.10.

A formulation of quantum mechanics has been given by Dirac and Feynman, which is a third formulation of quantum mechanics that is independent of, and equivalent to, the other two.

In the Dirac–Feynman approach, the inherent indeterminacy of the quantum entity is realized by the time evolution of the degree of freedom having *indeterminate* trans-empirical paths discussed in Sect. 3.7. A complex number is assigned to the probability amplitude for each empirical path. For a quantum degree of freedom evolving from an *observed* initial state to the *observed* final state—and with no other observations made—the Feynman path integral is a mathematical construction that computes the probability amplitudes by summing over *all* the possible empirical paths [5, 14].

11.1 Probability Amplitude and Time Evolution

Recall that the description of a quantum entity, at *a particular instant*, is given by its state vector, namely, $|\psi\rangle$, that yields $\rho = |\psi\rangle\langle\psi|$. The likelihood of observing the degree of freedom's position operator $|x\rangle\langle x|$ for the quantum state is given by $|\langle x|\psi\rangle|^2 = \mathrm{tr}\big(|x\rangle\langle x|\rho\big)$.

To avoid confusion with the state vector, the term *probability amplitude* is used for describing a quantum entity undergoing transitions in time.

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8_11,

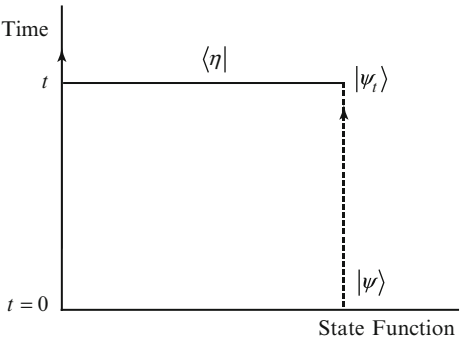


Fig. 11.1 Evolving state vector $|\psi\rangle$ through time t to state vector $|\psi_t\rangle$ to find the probability amplitude $\langle\eta|\psi_t\rangle$ (published with permission of)

Consider a quantum entity making a transition from an arbitrary initial state function $|\psi\rangle$ at time $t_{\rm i}=0$ to an arbitrary final state function $|\eta\rangle$ at final time $t_{\rm f}=t$.

To find the probability amplitude for this transition requires the initial state vectors to be at the same time as the final state vector and hence is written as $\langle\eta|\psi_t\rangle$. To simplify the notation, the probability amplitude and probability for the transition are given by the following:

$$\text{Initial state function}: \quad |\psi\rangle \text{ at time=0}; \quad \text{Final state function}: \quad |\eta\rangle \text{ at time=}t$$
$$\text{Probability Amplitude}: \quad \langle\eta|\psi_t\rangle; \quad \text{Probability of transition}: \quad |\langle\eta|\psi_t\rangle|^2$$

The state vector $|\psi\rangle$ must be evolved for a duration of time t to reach the time at which the final state vector is located, as shown in Fig. 11.1. The initial state vector $|\psi\rangle$ is evolved by applying the evolution operator on it. From (5.38), the time evolution of the state vector is given by

$$|\psi_t\rangle = {\rm e}^{-{\rm i}tH/\hbar}|\psi\rangle = U(t)|\psi\rangle \tag{11.1}$$

Hence, the probability amplitude to go in time t from an arbitrary initial state $|\psi\rangle$ to another arbitrary final state $|\chi\rangle$ is given by

$$\langle\eta|\psi_t\rangle = \langle\chi|{\rm e}^{-{\rm i}tH/\hbar}|\psi\rangle: \quad \text{Probability amplitude } (\psi\to\chi;t) \tag{11.2}$$

and the probability is given by

$$|\langle\eta|\psi_t\rangle|^2 = |\langle\chi|{\rm e}^{-{\rm i}tH/\hbar}|\psi\rangle|^2: \quad \text{Probability}(\psi\to\chi;t)$$

The initial state $|\psi\rangle$ is the "history state" of the transition amplitude, and the final state $\langle\chi|$ is its "destiny state." In quantum theory both the *history state* and the *destiny state* can be *independently* specified.

Consider superposed states given by

$$|\psi\rangle = \sum_i c_i |\psi_i\rangle; \qquad |\chi\rangle = \sum_i b_i |\chi_i\rangle$$

From (11.2), the probability amplitude is given by

$$\langle \eta | \mathrm{e}^{-\mathrm{i}tH/\hbar} | \psi \rangle = \sum_{i,j} b_i^* c_j \langle \chi_i | \mathrm{e}^{-\mathrm{i}tH/\hbar} | \psi_j \rangle$$

Hence, the fundamental expression that needs to be evaluated is the general matrix element

$$\langle \chi_i | \mathrm{e}^{-\mathrm{i}tH/\hbar} | \psi_j \rangle$$

All probability amplitudes can be evaluated from the general matrix element of $\mathrm{e}^{-\mathrm{i}tH/\hbar}$ given above.

Consider the important special case of a quantum particle with the degree of freedom given by the coordinate x. From (11.2), the probability amplitude, using the completeness equation (4.19), is given by

$$\langle \eta | \mathrm{e}^{-\mathrm{i}tH/\hbar} | \psi \rangle = \int \mathrm{d}x_\mathrm{f}\, \mathrm{d}x_\mathrm{i}\, \eta^*(x_\mathrm{f}) \langle x_\mathrm{f} | \mathrm{e}^{-\mathrm{i}tH/\hbar} | x_\mathrm{i} \rangle \psi(x_\mathrm{i})$$

11.2 Evolution Kernel

The time evolution, also called the transition amplitude, of the state vector is determined by the operator $\mathrm{e}^{-\mathrm{i}tH/\hbar}$. The transition amplitude or evolution kernel is the matrix element of $\mathrm{e}^{-\mathrm{i}tH/\hbar}$ and is given by

$$\begin{aligned} K(x_\mathrm{f}, t_\mathrm{f}; x_\mathrm{i}, t_\mathrm{i}) &= \langle x_\mathrm{f} | U(t) | x_\mathrm{i} \rangle = \langle x_\mathrm{f}, t_\mathrm{f} | x_\mathrm{i}, t_\mathrm{i} \rangle; \quad t = t_\mathrm{f} - t_\mathrm{i} \\ \Rightarrow K(x_\mathrm{f}, t_\mathrm{f}; x_\mathrm{i}, t_\mathrm{i}) &= \langle x_\mathrm{f}, t_\mathrm{f} | x_\mathrm{i}, t_\mathrm{i} \rangle = \langle x_\mathrm{f} | \mathrm{e}^{-\mathrm{i}(t_\mathrm{f} - t_\mathrm{i})H/\hbar} | x_\mathrm{i} \rangle \end{aligned} \tag{11.3}$$

Simplifying the notation, the evolution kernel is written as

$$K(x, x'; t) \equiv \langle x | \mathrm{e}^{-\mathrm{i}tH/\hbar} | x' \rangle \tag{11.4}$$

Note the time evolution of a state vector, from (5.38), and using the completeness equation given in (4.19) yields the following:

$$\begin{aligned} \psi(t, x) &= \langle x | \mathrm{e}^{-\mathrm{i}tH/\hbar} | \psi_0 \rangle = \int \mathrm{d}x' \langle x | \mathrm{e}^{-\mathrm{i}tH/\hbar} | x' \rangle \langle x' | \psi_0 \rangle \\ &= \int \mathrm{d}x' K(x, x'; t) \psi_0(x') \end{aligned} \tag{11.5}$$

and is illustrated in Fig. 11.2.

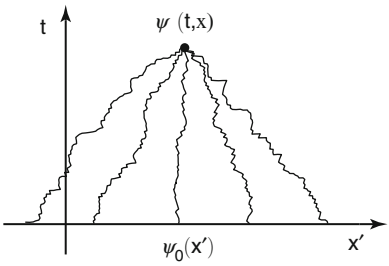


Fig. 11.2 The evolution kernel $K(x,x';t)$ propagates the initial state vector $\psi_0(x)$ through time t to state vector $\psi(t,x)$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

The solution of Schrödinger's equation given in (5.38) and (11.5) is formal since one needs to evaluate the matrix elements of the evolution operator $U(t)$, which in turn requires solving for all the eigenenergies and eigenfunctions of the Hamiltonian H. Let the eigenfunctions of the Hamiltonian be given by

$$H|\psi_n\rangle = E_n|\psi_n\rangle \tag{11.6}$$

The completeness equation, from (5.8), is then given by

$$\mathbb{I} = \sum_n |\psi_n\rangle\langle\psi_n| \tag{11.7}$$

Using the completeness equation given in (11.7) yields the evolution kernel

$$K(x,x';t) = \langle x|\mathrm{e}^{-\mathrm{i}tH/\hbar}|x'\rangle = \sum_n \mathrm{e}^{-\mathrm{i}tE_n/\hbar}\langle x|\psi_n\rangle\langle\psi_n|x'\rangle$$

Even though the expression for the transition amplitude $K(x_\mathrm{f},x_\mathrm{i};t)$ has greatly simplified, the sum over all eigenstates is still quite nontrivial for many realistic cases such as the harmonic oscillator.

All the eigenfunctions and eigenvalues of H are seldom known, and hence equation above is only a formal expression for the transition amplitude; one would like to have other avenues for approximately computing $K(x_\mathrm{f},x_\mathrm{i};t)$.

The *conditional probability* for the particle that starts from the coordinate eigenstate $|x_\mathrm{i}\rangle$ to end up, after evolving for time $t = t_\mathrm{f} - t_\mathrm{i}$, at the coordinate $|x_\mathrm{f}\rangle$, is given by the following:

$$\mathrm{P}(x_\mathrm{f}|x_\mathrm{i};t) = \frac{|\langle x_\mathrm{f}|\mathrm{e}^{-\mathrm{i}tH/\hbar}|x_\mathrm{i}\rangle|^2}{\int \mathrm{d}x_\mathrm{f}|\langle x_\mathrm{f}|\mathrm{e}^{-\mathrm{i}tH/\hbar}|x_\mathrm{i}\rangle|^2} \tag{11.8}$$

The normalization is necessary since one is comparing the likelihood of the particle going from x_i to x_f with the particle ending up at any other position. In particular,

$$\int \mathrm{d}x_\mathrm{f}\mathrm{P}(x_\mathrm{f}|x_\mathrm{i};t) = 1$$

Free Particle

Consider the Hamiltonian of a free quantum particle given by

$$H = \frac{\vec{p}^{\,2}}{2m} = -\frac{\hbar^2}{2m}\frac{\partial}{\partial \tilde{x}}\cdot\frac{\partial}{\partial \tilde{x}}$$

The eigenstates of H are given by the plane wave eigenstates $|\vec{p}\rangle$

$$H|\vec{p}\rangle = \frac{\vec{p}^{\,2}}{2m}|\vec{p}\rangle \tag{11.9}$$

that yield, from (4.27), the completeness equation

$$\int_{-\infty}^{\infty}\frac{\mathrm{d}\vec{p}}{(2\pi\hbar)^3}|\vec{p}\rangle\langle\vec{p}| = \mathbb{I};\;\; \langle\vec{p}|\vec{p}\,'\rangle = (2\pi\hbar)^3\delta(p-p') \tag{11.10}$$

The evolution kernel of the free particle, from (11.3), is given by

$$\begin{aligned} K(x,x';\tau) &= \langle\vec{x}|\mathrm{e}^{-\mathrm{i}t\frac{\vec{p}^{\,2}}{2m\hbar}}|x'\rangle = \int\frac{\mathrm{d}\vec{p}}{(2\pi\hbar)^d}\mathrm{e}^{-\mathrm{i}t\frac{\vec{p}^{\,2}}{2m\hbar}}\mathrm{e}^{\mathrm{i}\vec{p}\cdot(\vec{x}-\vec{x}\,')/\hbar} \\ &= \left(\sqrt{\frac{m}{2\pi\mathrm{i}\hbar t}}\right)^3\exp\left\{\frac{\mathrm{i}}{\hbar}\frac{m(\vec{x}-\vec{x}\,')^2}{2t}\right\} \end{aligned} \tag{11.11}$$

Noteworthy 11.1: Boundary conditions in classical and quantum mechanics

In classical mechanics, a particle is fully described by its position and velocity. Since Newton's equations of motion is based on acceleration that requires the second time derivative, one needs to specify *two* boundary conditions to uniquely specify a classical trajectory. Once the boundary conditions are specified, Newton's equations of motion yield a *determinate* and unique trajectory. In particular, if the position and velocity of a particle is specified at some instant, its future trajectory is fully determined.

In quantum mechanics the situation is quite different. The quantum degree of freedom is described by a state vector that yields the likelihood of being experimentally observed by a particular projection operator. The Schrödinger equation involves only the first time derivative of the state vector; hence, one needs to specify *either* the initial *or* the final state vector. Quantum mechanics, unlike classical mechanics, is a theory of probabilities. The initial state vector $|\psi\rangle$ evolves into a state $|\psi_t\rangle$ that has nonzero probability amplitude $\langle\chi|\psi_t\rangle$ with *many different* state vectors $\langle\chi|$—showing that the time evolution of the quantum particle is *indeterminate* and trans-empirical, with a likelihood of evolving from its initial state to many different possible final states.

11.3 Superposition of Trans-Empirical Paths

The probability amplitude for making a quantum transition from an initial to a final state can go through many intermediate paths and has been discussed in detail in Sect. 8.3; the discussion did not take into account the time dependence of the probability amplitudes, which is addressed in this section; the relevant expressions are re-derived to highlight the time evolution of the quantum entity.

Recall that the probability amplitude has two very different and distinct cases:

- The intermediate paths are *indistinguishable* and trans-empirical, namely, the information on which path has been taken by the quantum entity is *not known*.
- The intermediate paths are *distinguishable* and the path taken is known and empirical.

When the path is *not known*, the intermediate state of the quantum entity is *trans-empirical*; and *empirical* when the path taken is *known*.

The *nonclassical* content of the probability of a quantum event comes out in a remarkable manner for the case of an initial state vector making a transition to a final state via many *indistinguishable* and trans-empirical paths.

Consider the case of an initial state vector $|x_\mathrm{i}, t_\mathrm{i}\rangle$ making a transition to a final state vector $|x_\mathrm{f}, t_\mathrm{f}\rangle$ via N intermediate slits given by $|i\rangle$; $i = 1, 2, \ldots, N$ and shown in Fig. 11.3. In going from $|x_\mathrm{i}, t_\mathrm{i}\rangle$ to $|x_\mathrm{f}, t_\mathrm{f}\rangle$, the particle can go through any of the N-slits. The probability amplitude, in the notation of (11.3), is given by $K(x_\mathrm{f}, t_\mathrm{f}; x_\mathrm{i}, t_\mathrm{i}) = \langle x_\mathrm{f}, t_\mathrm{f} | x_\mathrm{i}, t_\mathrm{i} \rangle$.

The probability P of the transition from an initial state $|x_\mathrm{i}, t_\mathrm{i}\rangle$ to a final state $|x_\mathrm{f}, t_\mathrm{f}\rangle$, as shown in Fig. 11.3, has two different expressions:

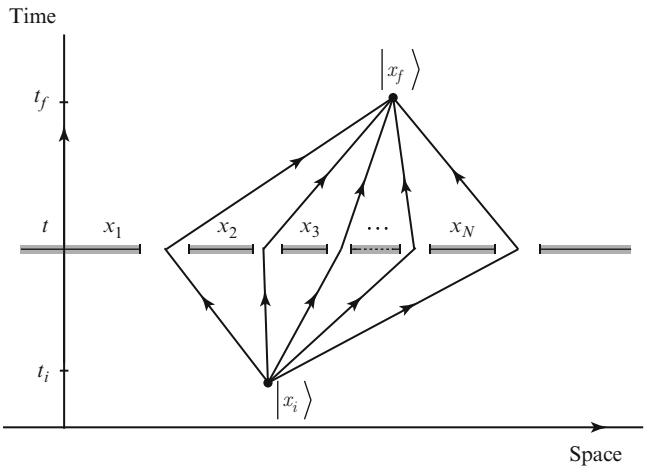


Fig. 11.3 Probability amplitudes for transition from initial state vector $|x_\mathrm{i}\rangle$ at time t_i to final state vector $|x_\mathrm{f}\rangle$ at time t_f—for N different possible intermediate paths (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

- The paths taken for the transition are empirical and determinate due to a measurement being made at time t to ascertain which intermediate position x_i is taken by the particle; in this case, the *probability* for the different paths *are added* and yields the probability of transition P_D given by

$$\begin{aligned} P_\mathrm{D} &= \sum_{n=1}^{N} |\langle x_\mathrm{f}, t_\mathrm{f}|x_n, t\rangle\langle x_n, t|x_\mathrm{i}, t_\mathrm{i}\rangle|^2 \\ &= \sum_{n=1}^{N} P_{x_\mathrm{f},x_n} P_{x_n,x_\mathrm{i}} \\ &\quad \text{where } P_{x_\mathrm{f},x_n} = |\langle x_\mathrm{f}, t_\mathrm{f}|x_n, t\rangle|^2;\ P_{x_n,x_\mathrm{i}} = |\langle x_n, t|x_\mathrm{i}, t_\mathrm{i}\rangle|^2 \end{aligned} \tag{11.12}$$

  The result for P_D follows from the classical composition of conditional probabilities, with the intermediate states being the allowed intermediate states.
- For the case when the intermediate paths are trans-empirical and indistinguishable, the probability amplitudes for the different determinate paths *are added* to yield the transition probability amplitude $\langle x_\mathrm{f}, t_\mathrm{f}|x_\mathrm{i}, t_\mathrm{i}\rangle$ given by

$$\langle x_\mathrm{f}, t_\mathrm{f}|x_\mathrm{i}, t_\mathrm{i}\rangle = \sum_{n=1}^{N} \langle x_\mathrm{f}, t_\mathrm{f}|x_n, t\rangle\langle x_n, t|x_\mathrm{i}, t_\mathrm{i}\rangle \tag{11.13}$$

  The probability amplitude yields the following observable transition probability P_I:

$$\begin{aligned} P_\mathrm{I} &= \left|\langle x_\mathrm{f}, t_\mathrm{f}|x_\mathrm{i}, t_\mathrm{i}\rangle\right|^2 = \left|\sum_{n=1}^{N} \langle x_\mathrm{f}, t_\mathrm{f}|x_n, t\rangle\langle x_n, t|x_\mathrm{i}, t_\mathrm{i}\rangle\right|^2 \\ &= \sum_{n=1}^{N} P_{x_\mathrm{f},x_n} P_{x_n,x_\mathrm{i}} \\ &\quad + \sum_{n\neq m}^{N} \langle x_\mathrm{f}, t_\mathrm{f}|x_n, t\rangle\langle x_n, t|x_\mathrm{i}, t_\mathrm{i}\rangle\langle x_\mathrm{f}, t_\mathrm{f}|x_m, t\rangle^*\langle x_m, t|x_\mathrm{i}, t_\mathrm{i}\rangle^* \end{aligned} \tag{11.14}$$

  There is no analog of (11.14) in classical probability theory.

11.4 The Dirac–Feynman Formula

Consider the case of a determinate and *discrete path*, with infinitesimal steps ϵ—as shown in Fig. 11.4a—that goes from x_i at t_i to final position x_f at time t_f. Let the intermediate points in the path be denoted by the following:

$$x_\mathrm{i} = x_0\ ;\ x_1\ ;\ x_2\ ;\ x_3\ ;\ x_n\ ;\ldots\ ;\ x_{N-1}\ ;\ x_N = x_\mathrm{f};\ t_n = t_\mathrm{i} + \epsilon n;\ t_N = t_\mathrm{f}$$

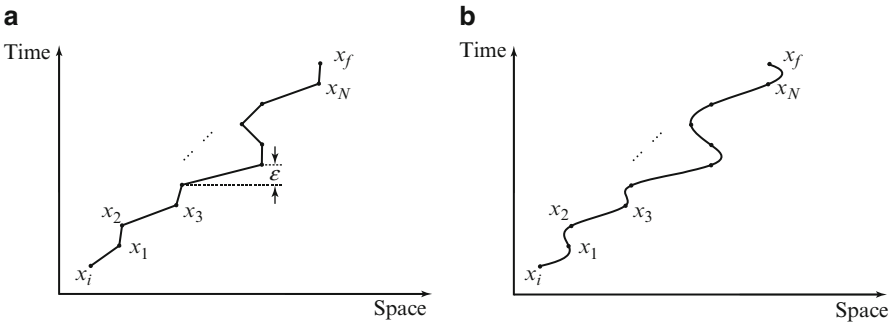


Fig. 11.4 (**a**) A single determinate path, discretized by time steps ϵ, from initial to final position. (**b**) The $\epsilon \to 0$ continuum limit of the discretized path (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

The path is *empirical* since all the intermediate points x_n are known (by a hypothetical experimental observation). Hence the principle of quantum superposition for successive steps, discussed in Sect. 8.4, tells us that the net amplitude ϕ_i is equal to the product of the probability amplitude for each infinitesimal determinate step and yields

$$\phi[\text{path}] = \langle x_N; t_N | x_{N-1}; t_{N-1} \rangle \dots \langle x_{n+1}; t_{n+1} | x_n; t_n \rangle \dots \langle x_1; t_1 | x_0; t_0 \rangle$$

Writing the probability amplitude in product notation yields

$$\phi[\text{path}] = \prod_{n=0}^{N-1} \langle x_{n+1}; t_{n+1} | x_n; t_n \rangle \tag{11.15}$$

For each infinitesimal determinate step $\epsilon = t_{n+1} - t_n$, the probability amplitude is postulated to be given by the Dirac–Feynman formula:

$$\langle x_{n+1}; t_{n+1} | x_n; t_n \rangle = \mathcal{N}(\epsilon) \exp\left\{ \epsilon \frac{i}{\hbar} \mathcal{L}(x_n, x_{n+1}; \epsilon) \right\} \tag{11.16}$$

where $\mathcal{N}(\epsilon)$ is a normalization constant and $\mathcal{L}$ is the *Lagrangian* of the particle given in (2.1). Using the Dirac–Feynman formula, given in (11.16), the probability amplitude for the discretized determinate path given in (11.15) has the following form:

$$\begin{aligned} \phi[\text{path}] &= \prod_{n=0}^{N-1} \langle x_{n+1}; t_{n+1} | x_n; t_n \rangle \\ &= \mathcal{N} \exp\left\{ \epsilon \frac{\mathrm{i}}{\hbar} \sum_{n=0}^{N-1} \mathcal{L}(x_n, x_{n+1}; \epsilon) \right\} \end{aligned} \tag{11.17}$$

$$= \mathcal{N} \exp\left\{\frac{\mathrm{i}}{\hbar} S[\text{path}]\right\} \tag{11.18}$$

where the discrete and determinate path that appears in $S[\text{path}]$ is shown in Fig. 11.4a. $\mathcal{N}$ is a path-independent normalization.

S is the *action functional*, defined earlier in (2.1), and, for the discrete paths, is given from (11.18) and (11.17) as follows:

$$S[\text{path}] = \epsilon \sum_{n=0}^{N-1} \mathcal{L}(x_n, x_{n+1}; \epsilon) \tag{11.19}$$

11.5 The Lagrangian

Let the total time interval $t_N - t_0 = t_\mathrm{f} - t_\mathrm{i}$ be kept fixed and let $N = (t_\mathrm{f} - t_\mathrm{i})/\epsilon$. In the continuum limit, $\epsilon \to 0$ and the paths become continuous. The discretized path shown in Fig. 11.4a converges to the continuous path shown in Fig. 11.4b.

The continuum limit $\epsilon \to 0$ ($N \to \infty$) yields

$$\begin{aligned}
&\frac{x_{n+1} - x_n}{\epsilon} \to \frac{\mathrm{d}x}{\mathrm{d}t}; \;\; t = n\epsilon \\
&\mathcal{L}(x_n, x_{n+1}; \epsilon) \to \mathcal{L}(x, \mathrm{d}x/\mathrm{d}t) \\
&S[\text{path}] \to S[x(t)] = \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}t\, \mathcal{L}(x, \mathrm{d}x/\mathrm{d}t)
\end{aligned}$$

The quantum particle's Lagrangian for continuous for time is $\mathcal{L}(x, \mathrm{d}x/\mathrm{d}t)$ and $S[x(t)]$ is the action functional for the continuous path $x(t)$; the notation used for the action is to indicate that the action depends on the *entire path* $x(t)$ with $t \in (t_\mathrm{i}, t_\mathrm{f})$.

The probability amplitude for the determinate continuous path $x(t)$—going from x_i at time t_i to final position x_f at time t_f—is given by the continuum limit of (11.18), and hence, the continuum action replaces the discretized action, namely, $S[\text{path}]$. Hence

$$\phi[x(t)] = \mathcal{N} \exp\left\{\frac{\mathrm{i}}{\hbar} S[x(t)]\right\} \tag{11.20}$$

The action $S[x(t)]$ has the dimensions of $\hbar$, and dividing it by $\hbar$ is required since only the dimensionless quantity $S/\hbar$ can be exponentiated; it is an empirical result that $\hbar$ is given by Planck's constant.

Equation (11.20) gives the probability amplitude for the quantum particle making a transition from the initial to its final position via *a specific possible path*. In other words, the path $x(t)$ can be *any determinate path* from the initial to the final position and not necessarily the path determined by classical mechanics.

The Hamiltonian given in (2.6) yields the following Lagrangian and action:

$$\mathcal{L}(x, \mathrm{d}x/\mathrm{d}t) = \tfrac{1}{2}m\left(\tfrac{\mathrm{d}x(t)}{\mathrm{d}t}\right)^2 - V(x(t)) \tag{11.21}$$

$$S[x(t)] = \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}t\, \mathcal{L}(x, \mathrm{d}x/\mathrm{d}t) \tag{11.22}$$

Although the Lagrangian and action given in (11.21) and (11.22) look like classical expressions, they are vastly different from the classical case. The reason being that, for quantum mechanics, the symbol $x(t)$ that appears in the Lagrangian and action is an *integration variable* for each t; in contrast in classical mechanics, $x(t)$ in the Lagrangian and action is restricted to the *classical path* $x_\mathrm{c}(t)$, for which the particle's path is a numerical function of time t.

Infinite Divisibility of Quantum Paths

Equation (11.17) is the reason that probability amplitude for a determinate path has a finite limit when the path is infinitely divided. As one makes the step size ϵ smaller, the Dirac–Feynman formula for each infinitesimal determinate step, say, from x_n to x_{n+1}, yields the correspondingly smaller expression in the exponential, namely, $\exp\{\mathrm{i}\epsilon\mathcal{L}(x_n, x_{n+1}; \epsilon)/\hbar\}$; this property of the paths leads, for $N \to \infty$, to a well limit of the infinite product for probability amplitude given in (11.15), namely, $\phi[\text{path}] \to \phi[x(t)]$, given in (11.20). The Dirac–Feynman formula given in (11.16) is the reason that the continuum limit of $\epsilon \to 0$ exists for the probability amplitude.

One can turn the above discussion around and argue that, for quantum mechanics to exist for continuous time, the probability amplitude for an infinitesimal step in time, of necessity, needs to be an exponential of an infinitesimal—of the form given by the Dirac–Feynman formula. This is because any determinate path is infinitely divisible in continuous time and hence requiring a concomitant convergent probability amplitude. The fact that for an infinitesimal step the probability amplitude is an exponential, which is proportional to the Lagrangian, is another deep insight of Dirac [10] and further developed by Feynman [14].

The requirement for a convergent probability amplitude for continuous paths answers a fundamental question as to why the action S that appears in classical physics (and which determines the dynamics of a classical system as in (2.1)) needs to be exponentiated in quantum mechanics as in $\exp\{\mathrm{i}S/\hbar\}$ given in (11.19). The classical to quantum transition is schematically given by

$$S[x_\mathrm{c}(t)] \to \exp\left\{\frac{\mathrm{i}}{\hbar}S[x(t)]\right\}$$

One explanation provided by the probability amplitude is that the requirement of quantum processes taking place in continuous time necessitates the exponentiation of the action. One may even state that the exponentiation of the action in quantum

mechanics is also the reason why quantum mechanics is qualitatively different from, and "exponentially" more complex, than classical mechanics.

11.6 The Feynman Path Integral

The result of the previous section provides an expression for the probability amplitude for the quantum entity to take a specific and *determinate path* in going from its initial to its final position. What is the probability amplitude if the quantum particle is only observed at its initial and final position? Due to the quantum indeterminacy of a quantum entity, we expect that the entity's degree of freedom's path will be indeterminate and hence it will "take" all possible *paths* simultaneously.

How many indeterminate paths are there between the initial and final positions? Clearly, there are many paths, and to develop a sense of these paths, consider putting barriers between the initial and final position to *limit* the number of possible paths, as shown in Fig. 11.5, so that we can enumerate the indeterminate paths. Once the procedure for enumerating the indeterminate paths becomes clear, the barriers will be removed, and all the indeterminate paths will then be included in our analysis.

Figure 11.5 shows a quantum particle going from initial state x_i at time t_i to final position x_f at time t_f through barriers that restrict the number of paths available to the quantum particle. Let an *entire continuous path*—going from initial state x_i to final state x_f through the successive slits as shown in Fig. 11.5—be denoted by path(n), with the probability amplitude denoted by $\phi[\mathrm{path}(n)]$. One can take path(n) to be straight lines from $x_\mathrm{i}, t_\mathrm{i}$ to the successive slit positions and another straight line from the last slit to $x_\mathrm{f}, t_\mathrm{f}$, as shown in Fig. 11.5.

Consider the case where the particle is observed at initial time t_i to be at x_i and then another measurement is only performed at final time t_f—with the particle being detected at x_f. The barriers are placed between the initial and final positions, and let there be N total number of different paths going from x_i to x_f. There are N

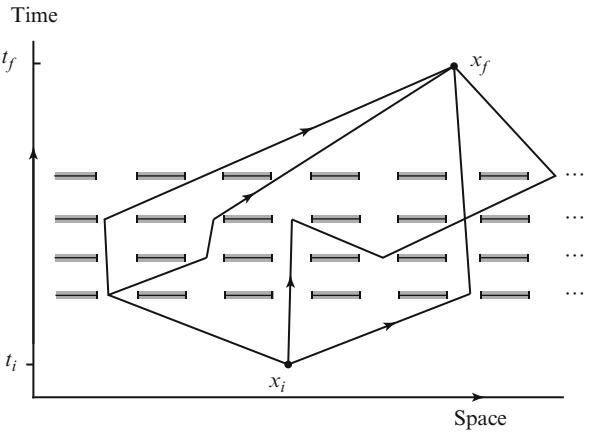


Fig. 11.5 Probability amplitudes for transition from initial state vector $|x_\mathrm{i}, t_\mathrm{i}\rangle$ to final state vector $|x_\mathrm{f}, t_\mathrm{f}\rangle$ for many successive slits with *indistinguishable* trans-empirical paths (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

indeterminate paths from $x_\mathrm{i}, t_\mathrm{i}$ to $x_\mathrm{f}, t_\mathrm{f}$ that are all indistinguishable and hence trans-empirical. From the superposition principle given in (11.13), the total probability amplitude is given by *adding* the probability amplitudes for all the indistinguishable determinate paths and yields

$$\langle x_\mathrm{f}, t_\mathrm{f} | x_\mathrm{i}, t_\mathrm{i} \rangle = \sum_{n=1}^{N} \phi[\mathrm{path}(n)] \tag{11.23}$$

The probability amplitude $\phi[\mathrm{path}(n)]$ for *each determinate path* is given by (11.20) and yields the following

$$\phi[\mathrm{path}(n)] = \mathcal{N} \exp\{\mathrm{i} S_n/\hbar\} \tag{11.24}$$

$$S_n = S[\mathrm{path}(n)]; \;\; n = 1, 2, \ldots, N \tag{11.25}$$

where $S[\mathrm{path}(n)]$ is the action for the continuous path(n) and $\mathcal{N}$ is a path-independent normalization.

Hence, from (11.23) and (11.24), the total space–time probability amplitude that the initial state vector $|x_\mathrm{i}, t_\mathrm{i}\rangle$ makes a transition to the final state vector $|x_\mathrm{f}, t_\mathrm{f}\rangle$—via *trans-empirical paths*—is given by superposing the amplitude for all the trans-empirical paths and yields

$$\begin{aligned} \langle x_\mathrm{f}, t_\mathrm{f} | x_\mathrm{i}, t_\mathrm{i} \rangle &= \mathcal{N} \sum_{n=1}^{N} \mathrm{e}^{\mathrm{i} S[\mathrm{path}(n)]/\hbar} = \mathcal{N} \sum_{n=1}^{N} \mathrm{e}^{\mathrm{i} S_n/\hbar} \\ &= \mathcal{N} \{ \mathrm{e}^{\mathrm{i} S_1/\hbar} + \mathrm{e}^{\mathrm{i} S_2/\hbar} + \cdots \} \end{aligned}$$

From (11.3), the evolution kernel (total transition amplitude) has the following representation:

$$K(x_\mathrm{f}, t_\mathrm{f}; x_\mathrm{i}, t_\mathrm{i}) = \langle x_\mathrm{f}, t_\mathrm{f} | x_\mathrm{i}, t_\mathrm{i} \rangle = \mathcal{N} \sum_{n=1}^{N} \mathrm{e}^{\mathrm{i} S[\mathrm{path}(n)]/\hbar} \tag{11.26}$$

One can successively shrink the barriers between the initial and final positions of the quantum particle, as discussed in Sect. 8.4, and as shown in Fig. 11.5, there will be great proliferation of possible paths. When there are no longer any slits, one has the limit of $N \to \infty$ or what is the same thing, there are infinitely many trans-empirical paths.

The transition amplitude is given by the sum over *all possible paths*, going from the initial position x_i at time t_i to the final state x_f at time t_f, as shown in Fig. 11.6, and yields the following :

$$K(x_\mathrm{f}, t_\mathrm{f}; x_\mathrm{i}, t_\mathrm{i}) = \mathcal{N} \sum_{\text{all paths}} \mathrm{e}^{\mathrm{i} S[\mathrm{path}]\hbar} : \quad \text{Feynman path integral} \tag{11.27}$$

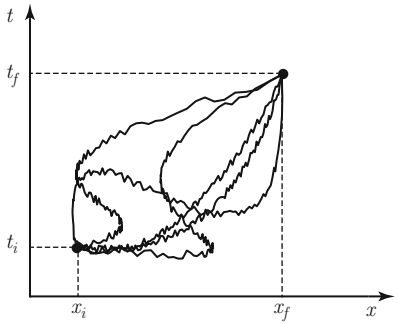


Fig. 11.6 All possible trans-empirical paths from the initial to the final state vector (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

The sum in (11.27) looks more figurative than a precise mathematical expression. After all, how are we supposed to actually perform a sum over infinitely many paths? Equation (11.27) is recast into a precise and mathematical expression in Sect. 11.7.

In summary, all the paths going $x_{\rm i}$ to final state $x_{\rm f}$ are *indistinguishable and hence trans-empirical* since no measurement is performed for the duration, from time $t_{\rm i}$ to time $t_{\rm f}$. The total probability amplitude to make a transition from initial state $x_{\rm i}$ to final state $x_{\rm f}$ is equal to the sum over all the individual probability amplitude $\mathrm{e}^{\mathrm{i}S/\hbar}$ for a specific path from initial state $x_{\rm i}$ to final state $x_{\rm f}$.

Noteworthy 11.2: Euclidean time τ

For physical (Minkowski) time, the Schrödinger's equation yields

$$\psi_t = \mathrm{e}^{-\mathrm{i}tH/\hbar}\psi_0$$

Minkowski time is analytically continued to Euclidean time τ, defined by

$$t = -i\tau\hbar \tag{11.28}$$

Propagation in Euclidean time is effected by operator $T = \mathrm{e}^{-\tau H}$ and yields

$$\psi_\tau = \mathrm{e}^{-\tau H}\psi_0$$

The reason for studying quantum systems in Euclidean time is to have a well-defined operator $\exp\{-\tau H\}$ that is convergent and not oscillatory. In Minkowski time, one is faced with a similar oscillatory expression such as $\mathrm{e}^{\mathrm{i}S/\hbar}$, which needs to be defined using the theory of distributions when $S \to \infty$.

Analytic continuation to Euclidean time entails *no loss of information* since T and H have the same eigenfunctions, with the Euclidean Hamiltonian equal to the original Minkowski Hamiltonian given by

$$T = \mathrm{e}^{-\tau H}; \qquad H = \frac{\vec{p}^{\,2}}{2m} + V(\vec{x}); \qquad \vec{p} = \frac{\partial}{i\partial\vec{x}}$$

The eigenvalues T_n, E_n are related by

$$E_n = -\frac{1}{\tau} \ln T_n$$

11.7 Path Integral for Evolution Kernel

Recall from (11.4), the evolution kernel (probability amplitude) is defined by

$$K(x_\mathrm{i}, x_\mathrm{f}; t) = \langle x_\mathrm{f} | \mathrm{e}^{-\mathrm{i}tH/\hbar} | x_\mathrm{i} \rangle$$

To render the sum over all paths in continuous space, namely, $\sum_{\text{all paths}}$ and given in (11.27), into a well-defined mathematical quantity, a derivation is given of the path integral starting from the Schrödinger equation. A corollary result will be to show that the definition given of $K(x_\mathrm{i}, x_\mathrm{f}; t)$ given in (11.4) is equivalent to the one derived in (11.27).

The evolution kernel is evaluated in Euclidean time since the expressions are mathematically more rigorous and transparent than in Minkowski time. The Euclidean (imaginary time) evolution kernel is given by

$$K(x, x'; \tau) = \langle x | \mathrm{e}^{-\tau H} | x' \rangle \tag{11.29}$$

Note $[\hat{p}^2, \hat{V}] \neq 0$, and it is this non-commutativity that poses the main problem in quantum mechanics. Ignoring the non-commutativity yields $\mathrm{e}^{-\tau H} \simeq \mathrm{e}^{-\tau \frac{p^2}{2m}} \mathrm{e}^{-\tau V}$, and for this case

$$\begin{aligned} K(x, x'; \tau) &\simeq \langle x | \mathrm{e}^{-\tau \frac{p^2}{2m}} \mathrm{e}^{-\tau V} | x' \rangle \\ &\simeq \mathrm{e}^{-\tau V(x')} \langle x | \mathrm{e}^{-\tau \frac{p^2}{2m}} | x' \rangle \end{aligned} \tag{11.30}$$

and the evolution kernel K requires the kernel for the free particle Hamiltonian $p^2/2m$ given in (11.11).

Note the remarkable fact that for non-commuting operators A and B

$$\mathrm{e}^A \mathrm{e}^B = \mathrm{e}^{A+B+\frac{1}{2}[A,B]+\cdots} \tag{11.31}$$

For $\tau = \epsilon$, infinitesimal time, one has the following result:

$$\begin{aligned} \mathrm{e}^{-\epsilon\hat{H}} &= \mathrm{e}^{-\epsilon(\frac{\hat{p}^2}{2m}+\hat{V})} \\ &= \mathrm{e}^{\frac{-\epsilon\hat{p}^2}{2m}}\mathrm{e}^{-\epsilon\hat{V}} + O(\epsilon^2) \end{aligned} \tag{11.32}$$

Hence, for infinitesimal time ϵ, from (11.30), the transition amplitude $K(x,x';\epsilon)$ can be evaluated exactly to $O(\epsilon^2)$.

The path integral approach is employed fundamentally to build up the finite time transition amplitude by composing the infinitesimal time transition amplitude by repeatedly using the resolution of the identity operator given in (4.19).

The evolution kernel for particle to go from initial position x_i to final position x_f in time τ can be written as follows:

$$\begin{aligned} K(x_\mathrm{f},x_\mathrm{i};\tau) &= \langle x_\mathrm{f}|\mathrm{e}^{-\tau H}|x_\mathrm{i}\rangle \\ &= \langle x|(\mathrm{e}^{-\frac{\tau}{N}H})^N|x'\rangle \end{aligned} \tag{11.33}$$

where for $\epsilon = \frac{\tau}{N}$, we have

$$K = \langle x_\mathrm{f}|\underbrace{\mathrm{e}^{-\epsilon H}\mathrm{e}^{-\epsilon H}\cdots\mathrm{e}^{-\epsilon H}}_{N\text{-times}}|x_\mathrm{i}\rangle \tag{11.34}$$

Inserting $N-1$ times, the completeness equation given in (4.19)

$$\mathbb{I} = \int_{-\infty}^{\infty}\mathrm{d}x|x\rangle\langle x|$$

yields the following:

$$\begin{aligned} K(x_\mathrm{f},x_\mathrm{i};\tau) = \int \mathrm{d}x_1\mathrm{d}x_2\ldots\mathrm{d}x_{N-1}\langle x_\mathrm{f}|\mathrm{e}^{-\epsilon H}|x_{N-1}\rangle\langle x_{N-1}|\mathrm{e}^{-\epsilon H}|x_{N-2}\rangle \\ \cdots\langle x_{n+1}|\mathrm{e}^{-\epsilon H}|x_n\rangle\cdots\langle x_1|\mathrm{e}^{-\epsilon H}|x_\mathrm{i}\rangle \end{aligned} \tag{11.35}$$

Consider the matrix element

$$\langle x|\mathrm{e}^{-\epsilon H}|x'\rangle = \int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi}\langle x|\mathrm{e}^{-\epsilon H}|p\rangle\langle p|x'\rangle \tag{11.36}$$

Since $\langle x|p\rangle = \mathrm{e}^{\mathrm{i}px}$, one has from (11.32)

$$\begin{aligned} \langle x|\mathrm{e}^{-\epsilon H}|x'\rangle &= \mathrm{e}^{-\epsilon V(x')}\int\frac{\mathrm{d}p}{2\pi}\mathrm{e}^{-\frac{\epsilon p^2}{2m}}\mathrm{e}^{\mathrm{i}p(x-x')} \\ &= \sqrt{\frac{m}{2\pi\epsilon}}\mathrm{e}^{-\frac{m}{2\epsilon}(x-x')^2-\epsilon V(x')} \end{aligned} \tag{11.37}$$

Recall from (11.16), the Dirac–Feynman formula, for each infinitesimal determinate step ϵ, is given by

$$\langle x_{n+1};t_{n+1}|x_n;t_n\rangle = \mathcal{N}(\epsilon)\exp\left\{\epsilon\frac{\mathrm{i}}{\hbar}\mathcal{L}(x_n,x_{n+1};\epsilon)\right\}$$

Recall from (11.3), the evolution kernel (transition amplitude) is defined by

$$K(x_\mathrm{f},t_\mathrm{f};x_\mathrm{i},t_\mathrm{i}) = \langle x_\mathrm{f},t_\mathrm{f}|x_\mathrm{i},t_\mathrm{i}\rangle = \langle x_\mathrm{f}|\mathrm{e}^{-\mathrm{i}(t_\mathrm{f}-t_\mathrm{i})H/\hbar}|x_\mathrm{i}\rangle$$

From (11.16) and (11.3), and simplifying the notation yields the Lagrangian, defined for infinitesimal Euclidean time ϵ, given by

$$\mathcal{N}(\epsilon)\mathrm{e}^{\epsilon\mathcal{L}(x,x';\epsilon)} = \langle x|\mathrm{e}^{-\epsilon H}|x'\rangle : \quad \text{Dirac–Feynman formula} \tag{11.38}$$

Hence, from (11.37)

$$\begin{aligned}\mathcal{N}(\epsilon)\mathrm{e}^{\epsilon\mathcal{L}(x,x';\epsilon)} &= \langle x|\mathrm{e}^{-\epsilon H}|x'\rangle \\ &= \sqrt{\frac{m}{2\pi\epsilon}}\mathrm{e}^{-\frac{m}{2\epsilon}(x-x')^2-\epsilon V(x')}\end{aligned} \tag{11.39}$$

For the particle degree of freedom, the Hamiltonian, for potential $V(x)$, from (2.6), is given by

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

and its Lagrangian is given by (11.39), for discrete time $t = n\epsilon$, by

$$\mathcal{L} = -\frac{m}{2}\left(\frac{x-x'}{\epsilon}\right)^2 - V(x) \tag{11.40}$$

The Euclidean Lagrangian is sometimes written more symmetrically as

$$\mathcal{L} = -\frac{m}{2}\left(\frac{x-x'}{\epsilon}\right)^2 - \frac{1}{2}[V(x)+V(x')] \tag{11.41}$$

and to $O(\epsilon)$ is same as the one given in (11.40).

Hence, the transition amplitude is given by

$$\begin{aligned}K(x_\mathrm{f},x_\mathrm{i};\tau) &= \int\prod_{n=1}^{N-1}\mathrm{d}x_n\prod_{n=0}^{N-1}\langle x_{n+1}|\mathrm{e}^{-\epsilon H}|x_n\rangle \\ &\equiv \int \mathrm{D}X\mathrm{e}^{\epsilon\sum_{n=0}^{N-1}\mathcal{L}(x_{n+1},x_n)} \equiv \int \mathrm{D}X e^{S}\end{aligned} \tag{11.42}$$

$$\text{Boundary conditions} \ : \ x_0 = x_\mathrm{i}; \ \ x_N = x_\mathrm{f} \tag{11.43}$$

where the lattice action and path integral integration measure are given by

$$S = -\epsilon \sum_{n=0}^{N-1} \frac{m}{2} \left(\frac{x_{n+1} - x_n}{\epsilon} \right)^2 - \epsilon \sum_{n=0}^{N-1} V(x_n)$$

$$\int \mathrm{D}X = \left(\frac{m}{2\pi\epsilon} \right)^{\frac{N}{2}} \prod_{n=1}^{N-1} \int_{-\infty}^{+\infty} \mathrm{d}x_n$$

In the continuum limit $\epsilon \to 0$, one obtains the following:

$$S = \int_0^\tau \mathcal{L}\,\mathrm{d}t; \quad \mathcal{L} = -\frac{m}{2} \left(\frac{\mathrm{d}x}{\mathrm{d}t} \right)^2 - V(x)$$

$$\int \mathrm{D}X = N \prod_{t=0}^{\tau} \int \mathrm{d}x(t) \tag{11.44}$$

The continuum Euclidean path integral representation for the evolution kernel is given by

$$\begin{aligned} K(x_\mathrm{f}, x_\mathrm{i}; \tau) &= \langle x_\mathrm{f} | \mathrm{e}^{-\tau H} | x_\mathrm{i} \rangle \\ &= \int_{\mathrm{B.C.}} \mathrm{D}X \mathrm{e}^S : \quad \text{Feynman path integral} \qquad (11.45) \\ \text{Boundary condition} \ : \ & x(0) = x_\mathrm{i} \ , \ x(\tau) = x_\mathrm{f} \end{aligned}$$

All paths between the initial and final position, figuratively shown in Fig. 11.6, are summed over in the $\int \mathrm{D}X \mathrm{e}^S$ path integration given in (11.45). The figurative summation over all paths $\sum_{\text{all paths}} \mathrm{e}^{\mathrm{i}S/\hbar}$ given in (11.27) is given a mathematical realization in (11.45), which is a Euclidean integration over all paths.

At each instant, the position degree of freedom takes all its values; at instant t, the degree of freedom is equal to the real line $\mathfrak{R}_t$; the total space of all paths is given by a tensor product over all instants and yields the total space of all paths equal to $\otimes_t \mathfrak{R}_t$. In general, for degree of freedom space given by $\mathcal{F}$, the path space is given by $\otimes_t \mathcal{F}_t$.

Similar to the arbitrariness in choosing complete basis states, as discussed in Sect. 4.7, the choice of the coordinates for the integration variables in the path integral is arbitrary. Just as one can make a change in the basis states of Hilbert space by a unitary transformation, one can make a change of the integration variables; the analog of the requirement of unitarity in changing the basis states is that the change of the path integration variables needs to be invertible, and which in turn yields a positive Jacobian of the transformation. More precisely, let the new integration variables be defined by $y = y(x)$; then,

$$\mathrm{d}y(t) = \int \mathrm{d}t' C(t, t') \mathrm{d}x(t') \quad \Rightarrow \quad \mathrm{D}Y = J[x] \mathrm{D}X$$

where $J = \det[C]$ is the Jacobian of the transformation.

In summary, the *Feynman path integral* is an efficient mathematical instrument for evaluating the finite time matrix elements (of the Euclidean continuation) of the unitary operator $U(t)$, namely, of $\langle x_\mathrm{f}|\mathrm{e}^{-\tau H}|x_\mathrm{i}\rangle$.

Free Particle Path Integral

A path integral derivation is given of the evolution kernel for a free particle degree of freedom moving in one dimension ($d = 1$) and which was obtained earlier in (11.11) using the eigenfunctions of the free particle Hamiltonian.

Let $\epsilon = t/N$; from (11.35), the path integral for finite ϵ is given by a multiple integral. Using the infinitesimal form of (11.11), which also directly follows from the Dirac–Feynman formula, yields (set $\hbar = 1$)

$$\begin{aligned} K(x,x';\tau) &= \prod_{n=0}^{N-1} K(x_{n+1},x_n;\epsilon) \\ &= \left(\sqrt{\frac{m}{2\pi\epsilon}}\right)\prod_{n=1}^{N-1}\sqrt{\frac{m}{2\pi\epsilon}}\int_{-\infty}^{+\infty}\mathrm{d}x_n \mathrm{e}^{-\frac{m}{2\epsilon}\sum_{n=0}^{N-1}(x_{n+1}-x_n)^2} \\ &\equiv \int \mathrm{D}X\, \mathrm{e}^{-\frac{m}{2\epsilon}(x-x_{N-1})^2}\cdots \mathrm{e}^{-\frac{m}{2\epsilon}(x_2-x_1)^2}\mathrm{e}^{-\frac{m}{2\epsilon}(x_1-x')^2} \end{aligned}$$

$$\text{Boundary conditions} \;:\; x = x_N;\;\; x' = x_0 \tag{11.46}$$

and where

$$\int \mathrm{D}X \equiv \left(\sqrt{\frac{m}{2\pi\epsilon}}\right)\prod_{n=1}^{N-1}\sqrt{\frac{m}{2\pi\epsilon}}\int_{-\infty}^{+\infty}\mathrm{d}x_n \tag{11.47}$$

Note the identity

$$\sqrt{\frac{m}{2\pi\epsilon}}\int_{-\infty}^{+\infty}\mathrm{d}x_1 \mathrm{e}^{-\frac{m}{2\epsilon}(x_2-x_1)^2-\frac{m}{2\epsilon}(x_1-x')^2} = \sqrt{\frac{1}{2}}\mathrm{e}^{-\frac{m}{2}\cdot\frac{1}{2\epsilon}(x_2-x')^2}$$

One can evaluate the path integral exactly by performing the DX integrations recursively, starting from the end with x_1. The successive integrations over the variables $x_1 \to x_2 \to x_3 \cdots \to x_{N-1}$ yields

$$K(x,x';\tau) = \sqrt{\frac{m}{2\pi N\epsilon}}\mathrm{e}^{-\frac{m}{2}\cdot\frac{1}{N\epsilon}(x-x')^2} = \sqrt{\frac{m}{2\pi\tau}}\mathrm{e}^{-\frac{m}{2\tau}(x-x')^2} \tag{11.48}$$

and is the result obtained in (11.11).

The case for a free particle in arbitrary d-dimensional space follows from (11.48), since the d-dimensional transition amplitude factorizes into separate one-dimensional components.

To obtain the evolution kernel in Minkowski time, recall from (11.28) that

$$\tau = \frac{\mathrm{i}}{\hbar} t \tag{11.49}$$

Hence, from (11.48), the Minkowski time evolution kernel, denoted by subscript M, and restoring $\hbar$ in the formula yields the following:

$$K_M(x,x';t) = \sqrt{\frac{m}{2\pi\tau}} \mathrm{e}^{-\frac{m}{2\tau\hbar}(x-x')^2} = \sqrt{\frac{m}{2\pi \mathrm{i}\hbar t}} \mathrm{e}^{\mathrm{i}\frac{m}{2t\hbar}(x-x')^2} \tag{11.50}$$

and was obtained in (11.11) using the free particle Hamiltonian.

11.8 Composition Rule for Probability Amplitudes

Consider the case of a particle going through N-slits, as shown in Fig. 11.3, with all the paths being *indistinguishable*. Equation (11.14) yields the following probability amplitude:

$$\langle x_\mathrm{f}, t_\mathrm{f} | x_\mathrm{i}, t_\mathrm{i} \rangle = \sum_{n=1}^{N} \langle x_\mathrm{f}, t_\mathrm{f} | x_n, t \rangle \langle x_n, t | x_\mathrm{i}, t_\mathrm{i} \rangle$$

Suppose the slits have spacing a, so that $x_n = na$, with $n = 0, \pm 1, \pm 2, \ldots, \pm\infty$, that is, the slits extend over the entire x-axis. The probability amplitude, extending (11.14) to the entire x-axis, is given by

$$\langle x_\mathrm{f}, t_\mathrm{f} | x_\mathrm{i}, t_\mathrm{i} \rangle = \sum_{n=-\infty}^{+\infty} \langle x_\mathrm{f}, t_\mathrm{f} | x_n, t \rangle \langle x_n, t | x_\mathrm{i}, t_\mathrm{i} \rangle \tag{11.51}$$

To take the continuum limit of (11.51), the bra and ket vectors $|x_n, t\rangle$, $\langle x_n, t|$ defined on a discrete set $x_n = na$ need to be written in continuum notation; for $a \to 0$, let $z = na$. The connection of the continuous and discrete state vector is given by (4.15):

$$\begin{aligned} \lim_{a\to 0}: \quad & |x_n, t\rangle \to \sqrt{a}|z, t\rangle; \quad x_n = na \\ & \langle x_n, t| \to \sqrt{a}\langle z, t|; \quad -\infty \le z \le \infty \end{aligned} \tag{11.52}$$

Also, let $t_\mathrm{i} = 0$, $t = \tau$, and $t_\mathrm{f} = \tau + \tau'$. The initial and final state vectors are defined for continuous initial and final positions and hence have the limit

$$|x_\mathrm{i}, t_\mathrm{i}\rangle \to |x, 0\rangle; \quad \langle x_\mathrm{f}, t| \to \langle x', \tau + \tau'| \tag{11.53}$$

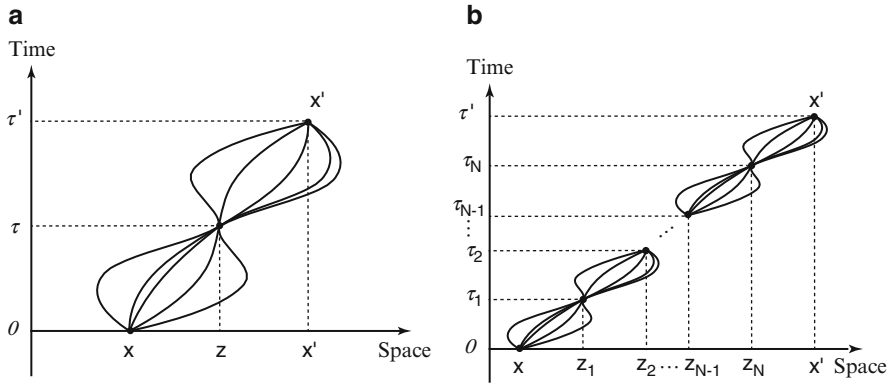


Fig. 11.7 (**a**) Probability amplitudes for transition from initial state vector $|x\rangle$ to final state vector $|x'\rangle$, summing over all indistinguishable paths passing through position z at time τ. (**b**) The probability amplitude with path going through many intermediate positions $z_1, z_2, \ldots, z_n$ at times $\tau_1, \tau_2, \ldots, \tau_N$ (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

As shown in Fig. 11.7a, taking the $a \to 0$ limit, from (11.51), (11.52), and (11.55), yields

$$\langle x', \tau + \tau' | x, 0 \rangle = a \sum_{n=-\infty}^{+\infty} \langle x', \tau + \tau' | z, \tau \rangle \langle z, \tau | x, 0 \rangle \tag{11.54}$$

$$\Rightarrow \langle x', \tau + \tau' | x, 0 \rangle \to \int_{-\infty}^{+\infty} \mathrm{d}z \langle x', \tau + \tau' | z, \tau \rangle \langle z, \tau | x, 0 \rangle \tag{11.55}$$

Writing the transition amplitude in (11.55) in terms of the evolution kernel given in (11.3), for Euclidean time, yields the following:

$$K(x', x; \tau + \tau') = \int_{-\infty}^{+\infty} \mathrm{d}z K(x', z; \tau') K(z, x; \tau) \tag{11.56}$$

Equation (11.56), illustrated in Fig. 11.7a, shows that the definition of the evolution kernel is consistent with the rules for the composition of probability amplitudes by summing over all indistinguishable paths.

In writing (11.56), only the property of the action was used. Consider a finite time slice $[0, \tau + \tau']$, as shown in Fig. 11.7a; due to the term $(\mathrm{d}x/\mathrm{d}t)^2$ in the action given in (11.44), one needs to specify the initial value x at $t = 0$ and a final value z at $t = \tau$, since two boundary conditions are required to specify paths going from x to z. The state space appears in the path integral via the boundary conditions imposed on the paths over which the path integration is defined.

Condition given in (11.56) is a fundamental property of probability amplitudes that allows one to define the state space—since the path integral can be interpreted as the matrix element connecting the initial and final state vectors. The fundamental reason that the action satisfies the composition law is because, writing the action in terms of its initial and final boundary variables as $S[x_{\rm f}, x_{\rm i}]$, the action given in (11.44), in the notation of Fig. 11.7a, has the form

$$S[x', x] = S[x', z] + S[z, x] \tag{11.57}$$

Interestingly enough, the above equation holds only for state space expressed in terms of coordinate state vectors $|x\rangle$; unlike the Schrödinger equation that holds equally in momentum space, the composition law given in (11.56) does not hold when expressed in terms of Fourier transformed variables essentially because (11.57) does not hold for Fourier transformed variables.

In many complicated cases such as quantum mechanical systems that are higher order in time and for quantum field theory on curved space–time, the quantum theory is defined directly in terms of the action, and a Hamiltonian may not exist; in such cases, one can directly base the existence of the state space on the properties of the Lagrangian and action.

For the case where there is a well-defined Hamiltonian, (11.56) follows directly from the definition of the evolution kernel in terms of the Hamiltonian given in (11.29) and the completeness equation; more precisely,

$$\begin{aligned} K(x', x; \tau + \tau') &= \langle x' | \mathrm{e}^{-(\tau'+\tau)\hat{H}} | x \rangle \\ &= \int_{-\infty}^{+\infty} \mathrm{d}z \langle x | \mathrm{e}^{-\tau'\hat{H}} | z \rangle \langle z | \mathrm{e}^{-\tau\hat{H}} | x' \rangle \\ &= \int_{-\infty}^{+\infty} \mathrm{d}z K(x', z; \tau') K(z, x; \tau) \\ \text{since } \mathrm{e}^{-(\tau'+\tau)H} &= \mathrm{e}^{-\tau' H} \mathrm{e}^{-\tau H} \end{aligned}$$

11.9 Trans-Empirical Paths and Path Integral

The Feynman path integral is a summation over all the trans-empirical paths that go from the initial position $x_{\rm i}$ to the final position $x_{\rm f}$ and yields the transition amplitude $K(x_{\rm f}|x_{\rm i}; t) = \langle x_{\rm f} | \mathrm{e}^{-\mathrm{i}tH/\hbar} | x_{\rm i} \rangle$; (11.8) yields the conditional probability given by

$$\mathrm{P}(x_{\rm f}|x_{\rm i}; t) = \frac{|\langle x_{\rm f} | \mathrm{e}^{-\mathrm{i}tH/\hbar} | x_{\rm i} \rangle|^2}{\int \mathrm{d}x_{\rm f} |\langle x_{\rm f} | \mathrm{e}^{-\mathrm{i}tH/\hbar} | x_{\rm i} \rangle|^2}$$

What is actually measured in the laboratory are only the initial and final position projection operators for $x_{\rm i}$ and $x_{\rm f}$, with an elapsed time between the measurements given by $t = t_{\rm f} - t_{\rm i}$. The particle's state is prepared to be at position $x_{\rm i}$ and observed

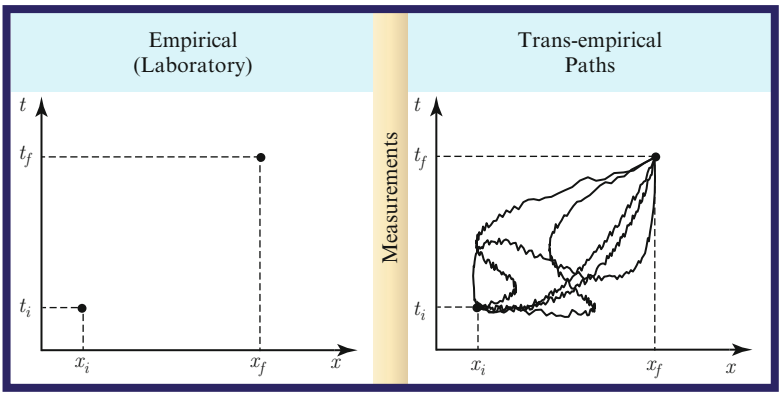


Fig. 11.8 (**a**) The experiment measuring the initial prepared at time t_i and the final position x_f at time t_f. (**b**) The quantum mechanical superstructure of trans-empirical indeterminate paths to explain the observed results (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

at time t_i. The particle then evolves without any measurement till time t_f, and a second measurement is made to determine the particle's position, which is found by projection operators at position x_f.

One then repeats the experiment many times, each time starting with the same prepared state at position x_i and finding that the quantum particle is observed to be at different final positions x_f. On making a frequency table of the frequency for the various final positions and normalizing it, one experimentally obtains the conditional probability $\mathrm{P}(x_\mathrm{f}|x_\mathrm{i};t)$.

Figure 11.6 showing the trans-empirical paths that are summed over in the path integral is resolved into Fig. 11.8a and b. Figure 11.8a shows what is actually measured in an experiment; Fig. 11.8b represents the superstructure of quantum mechanics that is required for *explaining* the results of the experiment. Figure 11.8a exists in the laboratory, whereas Fig. 11.8b shows that the trans-empirical paths exist in path space, which does not exist in physical space.

The trans-empirical paths clearly show the indeterminate nature of the quantum degrees of freedom, since all the degrees of freedom $x(t)$, for $t_\mathrm{i} < t < t_\mathrm{f}$, are *integration variables*, taking all the allowed values in defining the path integral. The path integral formulation of quantum mechanics clearly shows the interplay between observables and unobservables: all integration variables are in principle unobservable since they are summed over for producing any observable effect.

For the path integral, the observables are the initial and final states, namely, the boundary conditions of the path integral, together with time parameter t that labels the degrees of freedom. In quantum field theory, due to the extensive nature of quantum fields, the only observables are parameters of the theory as well as functions of space–time points as these are the labels of the quantum fields' degrees of freedom.

11.10 State Vector and Trans-Empirical Paths

One can connect the indeterminate degrees of freedom, viewed as integration variables, with the Schrödnger formulation based on the quantum state. The state vector $\psi(t,x)$ provides a description of the degree of freedom at time t. In contrast, from (11.4), the evolution kernel is given by

$$K(x,x';t_f-t_i)=\langle x|\mathrm{e}^{-\mathrm{i}(t_\mathrm{f}-t_\mathrm{i})H/\hbar}|x'\rangle$$

The relation to the state vector $\psi(t,x)$ and indeterminate and trans-empirical space–time paths is the following. At some intermediate time $t_0 \in (t_\mathrm{i},t_\mathrm{f})$, consider setting up an experiment that consists of projection operators at every point of space, namely, $|x\rangle\langle x|,\ x\in\Re$, to detect the quantum state. The quantum state of the degree of freedom at time t_0, according to the Schrödnger formulation, is described by the state vector $\psi(t_0,x)$, with density matrix $\rho=|\psi\rangle\langle\psi|$. The Schrödinger state vector $\psi(t_0,x)$ is shown in Fig. 11.9.

The likelihood of the projection operator at t_0,x_0 recording the degree of freedom is given by

$$E[|x_0\rangle\langle x_0|\rho]=|\psi(t_0,x_0)|^2$$

Hence, the probability the projection operator makes a recording is given by

$$P(t_0,x_0)\mathrm{d}x=|\psi(t_0,x_0)|^2\mathrm{d}x=\mathrm{Prob}(x_0<x<x_0+\mathrm{d}x)$$

The coordinate basis obeys the completeness equation, and consequently it is certain that the passage of the particle's degree of freedom must be measured one of the projection operators; hence, the total probability must be 1 and leads to

$$1=\int_{-\infty}^{+\infty}\mathrm{d}xP(t_0,x)$$

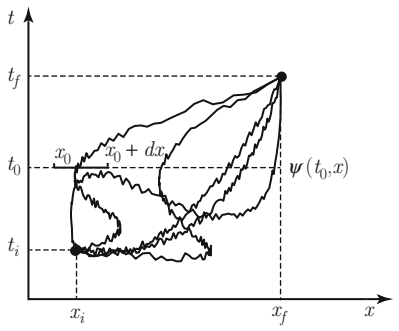


Fig. 11.9 The Schrï£¡dinger wave function yields the likelihood of the quantum particle being near some point (published with permission of © Belal E. Baaquie 2012. All Rights Reserved)

Hence, the path integral yields the Schrödinger state function $\psi(t_0,x)$ together with the probabilistic interpretation required by quantum mechanics, as shown in Fig. 11.9.

The state function $\psi(t_0,x)$ obeys the Schrödinger equation with boundary condition given by the initial condition, namely

$$\begin{aligned}\langle x|\psi(t_0;x')\rangle &\equiv \psi(t_0;x',x) = \langle x|\mathrm{e}^{-\mathrm{i}(t_0-t_i)H/\hbar}|\psi(t_i;x')\rangle \\ \langle x|\psi(t_i;x')\rangle &= \delta(x-x')\end{aligned} \tag{11.58}$$

The state vector and the evolution kernel is given by

$$K(x,x';t_f-t_i) = \langle x|\mathrm{e}^{-\mathrm{i}(t_f-t_0)H/\hbar}|\psi(t_0;x')\rangle$$

The problem with the expression above is that one does not obtain a normalizable state vector $|\psi(t)\rangle$ with the definitions given above.

The state vector $\psi(t_i,x)$ needs to smeared out a bit in order to obtain a normalizable state vector $\psi(t_0,x)$. For example, one can take the initial state vector to be a Gaussian sharply peaked at x as given below:

$$\chi(z;x') = \langle z|\chi(x')\rangle = \left(\frac{1}{2\pi\sigma^2}\right)^{1/4}\exp\left\{-\frac{1}{4\sigma^2}(z-x')^2\right\}$$

$$|\tilde{\psi}(t_0;x')\rangle = \mathrm{e}^{-\mathrm{i}(t_0-t_i)H/\hbar}|\chi(x')\rangle;\ \ \langle\tilde{\psi}(t_0;x')|\tilde{\psi}(t_0;x')\rangle = 1$$

and the evolution kernel is given by

$$K(x,x';t) \simeq \langle x|\mathrm{e}^{-\mathrm{i}(t_\mathrm{f}-t_0)H/\hbar}|\tilde{\psi}(t_0;x')\rangle$$

The limit of $\sigma^2 \to 0$ needs to be taken for recovering the exact expression given for $K(x,x';t_f-t_i)$.

11.11 Path Integral Quantization: Action

The result obtained in (11.45) can be taken to be the starting point of path integral quantization. Instead of starting from a Hamiltonian, as was done in Chap. 11, one can instead model the quantum phenomena in question by *postulating* a Lagrangian.

The starting point is the Lagrangian, and one needs to choose an appropriate potential $V(x)$ and to obtain the (Euclidean) Lagrangian and action given by

$$\mathcal{L} = -\frac{1}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 - V(x)$$

$$S = \int_{t_\mathrm{i}}^{t_\mathrm{f}} \mathrm{d}t\, \mathcal{L}$$

One *postulates* that the system takes all possible trans-empirical paths from the initial to the final state. The evolution kernel is given by the superposition of all the trans-empirical (indistinguishable) paths and is equal to sum of e^S over all possible paths; hence,

$$\langle x_\mathrm{f}|\mathrm{e}^{-\tau H}|x_\mathrm{i}\rangle = \int_{\mathrm{B.C.}} \mathrm{D}X\, \mathrm{e}^S \ : \text{Feynman path integral} \tag{11.59}$$
$$\text{Boundary condition}: \ x(0) = x_\mathrm{i} \ , \ x(\tau) = x_\mathrm{f}$$

One can derive *both*, the state space and the Hamiltonian, from (11.59).

Path integral quantization is more general than starting from the Schrödinger equation for two reasons:

- In the Schrödinger approach, one needs to postulate the properties of state space *in addition* to the Hamiltonian driving the Schrödinger equation.
- The space–time symmetries of the quantum system are explicit in the Lagrangian-based path integral approach, whereas in the Schrödinger approach these are implicit and need to be extracted using the properties of the Hamiltonian and state space. In particular, one has to derive the symmetry operators that commute with the Hamiltonian.

These considerations come to the forefront for complicated systems like non-Abelian gauge fields, where the starting point is the Lagrangian—and path integral quantization turns out to be more efficient than the Schrödinger approach.

11.12 Hamiltonian from Lagrangian

Recall in Sect. 11.6, the Lagrangian was derived from the Hamiltonian using the Dirac–Feynman formula.

The question naturally arises that if the Lagrangian is known, how would one derive its Hamiltonian; the purpose of this section is to carry out this derivation using quantum mechanical techniques. A Lagrangian that is more general than the one discussed in Sect. 11.6, and which arises in the study of option theory in finance [3], is chosen to illustrate some new features. Let the degree of freedom be the real variable ϕ.

Consider the following Lagrangian and action:

$$\mathcal{L}(t) = -\frac{1}{2}\left[m\mathrm{e}^{-2v\phi}\left\{\frac{d\phi}{\mathrm{d}t} + \alpha(\phi,t)\right\}^2 + V(\phi)\right]$$
$$S = \int_0^\tau \mathrm{d}t\, \mathcal{L}(t) = -\frac{1}{2}\int_0^\tau \mathrm{d}t\left[m\mathrm{e}^{-2v\phi}\left\{\frac{d\phi}{\mathrm{d}t} + \alpha(\phi,t)\right\}^2 + V(\phi)\right] \tag{11.60}$$

For greater generality, mass has been taken to depend on the degree of freedom ϕ, namely, that mass $= m\mathrm{e}^{-2\nu\phi}$ and a ϕ dependent drift term $\alpha(\phi,t)$ has also been included in the Lagrangian.

The path integral is given by the following generalization of (11.44):

$$K(\phi_\mathrm{i},\phi_\mathrm{f};\tau) = \int \mathrm{D}\phi \mathrm{e}^{-\nu\phi}\mathrm{e}^{S}$$

$$\int \mathrm{D}\phi \mathrm{e}^{-\nu\phi} \equiv \prod_{t=0}^{\tau}\int_{-\infty}^{+\infty} \mathrm{d}\phi(t)\mathrm{e}^{-\nu\phi(t)}$$

$$\text{Boundary conditions } \phi(\tau)=\phi_\mathrm{f}\ ;\ \phi(t=0)=\phi_\mathrm{i} \tag{11.61}$$

Note that, up to a normalization the path integral integration measure $\int \mathrm{D}\phi$ has a factor of $\mathrm{e}^{-\nu\phi}$ needed to obtain a well-defined Hamiltonian.

Recall from the discussion of the evolution kernel in Sect. 11.6, the path integral is related to the Hamiltonian H by (11.45), namely,

$$K(\phi_\mathrm{i},\phi_\mathrm{f};T) = \int \mathrm{D}\phi\, \mathrm{e}^{-\nu\phi}\mathrm{e}^{S} = \langle\phi_\mathrm{f}|\mathrm{e}^{-\tau H}|\phi_\mathrm{i}\,\rangle \tag{11.62}$$

One needs to extract the Hamiltonian H from the left-hand side of (11.62). Since the Hamiltonian propagates the system through infinitesimal time, it is discretized into a lattice of spacing ϵ, with $t = n\epsilon$ and $N = T/\epsilon$. The path integral reduces to a finite $(N-1)$-fold multiple integral, analogous to what was obtained in (11.35). Discretizing the time derivative by $\mathrm{d}\phi/\mathrm{d}t \to (\phi_{n+1}-\phi_n)/\epsilon$ yields the lattice action and Lagrangian given by

$$\langle\phi_N|\mathrm{e}^{-\epsilon NH}|\phi_0\rangle = \prod_{n=1}^{N-1}\int \mathrm{d}\phi_n \mathrm{e}^{-\nu\phi_n}\mathrm{e}^{S(\epsilon)}$$

$$S(\epsilon) = \epsilon\sum_{n=0}^{N-1} L(n)$$

$$L(n) = -\frac{m\mathrm{e}^{-2\nu\phi_n}}{2\epsilon^2}\left[\phi_{n+1}-\phi_n+\epsilon\alpha_n\right]^2 - \frac{1}{2}[V(\phi_{n+1})+V(\phi_n)] \tag{11.63}$$

As in Sect. 11.6, the completeness equation $\int \mathrm{d}\phi|\phi\,\rangle\langle\phi| = \mathbb{I}$ is used $N-1$ times to write out the expression for $\mathrm{e}^{-\epsilon NH}$, and the Hamiltonian is identified as

$$\begin{aligned}\langle\phi_{n+1}|\mathrm{e}^{-\epsilon H}|\phi_n\,\rangle &= \mathrm{e}^{-\nu\phi_n}\mathrm{e}^{\epsilon L_n}\\ &= \mathrm{e}^{-\nu\phi_n}\exp\left\{-\frac{m\mathrm{e}^{-\nu\phi}}{2\epsilon}\left[\phi_{n+1}-\phi_n+\epsilon\alpha_n\right]^2 - \frac{\epsilon}{2}[V(\phi_{n+1})+V(\phi_n)]\right\}\end{aligned}$$

Since the matrix elements of the Hamiltonian depend on the value of ϕ at two different instants, to simplify notation, let

$$\phi_{n+1} = \phi;\ \ \phi_n = \phi';\ \ \alpha_n = \alpha$$

Ignoring terms that are of $O(\epsilon)$ in (11.63) yields the following:

$$\langle\phi|\mathrm{e}^{-\epsilon H}|\phi'\rangle = \mathrm{e}^{-\nu\phi}\exp\Big\{-\frac{m\mathrm{e}^{-2\nu\phi}}{2\epsilon}\big[\phi-\phi'+\epsilon\alpha\big]^2-\epsilon V(\phi)\Big\} \tag{11.64}$$

Note that unlike (11.38), for which the Hamiltonian is known and the Lagrangian was derived from it, in (11.64) one needs to *derive* the Hamiltonian *from* the known Lagrangian.

The key feature of the Lagrangian that in general allows one to derive its Hamiltonian is that the Lagrangian contains only first-order time derivatives; hence, on discretization the Lagrangian involves only ϕ_n that are nearest neighbors in time, thus allowing it to be represented as the matrix element of $\mathrm{e}^{-\epsilon H}$, as in (11.64). Furthermore, the time derivatives appear in a quadratic form; hence, one can use Gaussian integration to rewrite (11.64) in the following manner[1]:

$$\begin{aligned}\langle\phi|\mathrm{e}^{-\epsilon H}|\phi'\rangle &= \mathrm{e}^{-\nu\phi}\mathrm{e}^{-\epsilon V(\phi)}\int_{-\infty}^{+\infty}\frac{\mathrm{d}p}{2\pi}\exp\big\{-\frac{\epsilon}{2m}p^2+\mathrm{i}p[\phi-\phi'+\epsilon\alpha]\mathrm{e}^{-\nu\phi}\big\}\\ &= \mathrm{e}^{-\epsilon V(\phi)}\int_{-\infty}^{+\infty}\frac{\mathrm{d}p}{2\pi}\exp\Big\{-\frac{\epsilon\mathrm{e}^{2\nu\phi}}{2m}p^2+\mathrm{i}p\big(\phi-\phi'+\epsilon\alpha\big)\Big\}\end{aligned} \tag{11.65}$$

where the pre-factor of $\mathrm{e}^{-\nu\phi}$ has been canceled by rescaling the integration variable $p \to p\mathrm{e}^{\nu\phi}$.

The Hamiltonian $H = H(\phi, \partial/\partial\phi)$ is a differential operator and acts on the dual coordinate ϕ, as is required for all differential operators, as mentioned earlier after (5.22). Hence, for the state function $|\psi\rangle$, which is an element of the state space, the Hamiltonian acts on the *dual basis state* $\langle\phi|$, and yields $\langle\phi|H|\psi\rangle = H(\phi, \partial/\partial\phi)\psi(\phi)$, similar to the result given in (5.37).

The Hamiltonian is hence given by the following representation[2]:

$$\langle\phi|\mathrm{e}^{-\epsilon H}|\phi'\rangle = \mathrm{e}^{-\epsilon H(\phi,\partial/\partial\phi)}\langle\phi|\phi'\rangle = \mathrm{e}^{-\epsilon H(\phi,\partial/\partial\phi)}\int_{-\infty}^{+\infty}\frac{\mathrm{d}p}{2\pi}\mathrm{e}^{\mathrm{i}p(\phi-\phi')} \tag{11.66}$$

since $\langle\phi|\phi'\rangle = \delta(\phi-\phi')$. Ignoring overall constants and using the property of the exponential function under differentiation, one can rewrite (11.65) as

$$\langle\phi|\mathrm{e}^{-\epsilon H}|\phi'\rangle = \exp\Big\{\frac{1}{2m}\epsilon\mathrm{e}^{2\nu\phi}\frac{\partial^2}{\partial\phi^2}+\epsilon\alpha\frac{\partial}{\partial\phi}-\epsilon V(\phi)\Big\}\int_{-\infty}^{+\infty}\frac{\mathrm{d}p}{2\pi}\mathrm{e}^{\mathrm{i}p(\phi-\phi)} \tag{11.67}$$

[1] Ignoring irrelevant constants.

[2] From (4.26), the convention for scalar product is $\langle p|\phi_n\rangle = \exp(-\mathrm{i}p\phi_n)$, and the sign of the exponential in (11.66) reflects this choice. The definition of H requires it to act on the dual state vector $\langle\phi'|$; if one chooses to write the Hamiltonian as acting of the state vector $|\varphi'\rangle$, $H^\dagger$ would then have been obtained. Since H is not Hermitian, this would lead to an incorrect result.

Comparing (11.67) above with (11.66) yields the Hamiltonian

$$H = -\frac{1}{2m} e^{2\nu\phi} \frac{\partial^2}{\partial \phi^2} - \alpha(\phi) \frac{\partial}{\partial \phi} + V(\phi) \tag{11.68}$$

The Hamiltonian is quite general since both $V(\phi)$ and $\alpha(\phi)$ can be functions of the degree of freedom ϕ. Note that the Hamiltonian H is non-Hermitian for a general value of α and is Hermitian only for a pure imaginary α. The path integral has a nontrivial integration measure $\exp\{-\nu\phi\}$ that needs to be specified in addition to the Hamiltonian.

11.13 Summary

The path integral is an independent formulation of quantum mechanics. To show the path integral's connection to the underlying foundations of indeterminacy of the quantum degree of freedom, the path integral was derived from Schrödinger's state vector formulation. The probability amplitude for a given determinate path was evaluated by breaking up the path into a series of infinitesimal paths; the Dirac–Feynman formula yields the probability amplitude for each infinitesimal path; composing all the infinitesimal paths yields the probability amplitude to be proportional to $\exp\{iS/\hbar\}$, where S is the action for the quantum degree of freedom.

The transition of a quantum entity from its initial to final state, without any observations during the interregnum, is made by the degree of freedom *simultaneously* taking all the trans-empirical indeterminate paths, which is a collection of many indistinguishable determinate paths. The principle of quantum superposition yields the transition amplitude as the sum of the probability amplitudes of all the indistinguishable determinate paths and leads to the summing of $\exp\{iS/\hbar\}$ over all the trans-empirical paths and yields the Feynman path integral.

The sum over all paths is given a precise mathematical expression by directly evaluating the transition amplitude using techniques based on the Hilbert space and, in particular, by the repeated employment of the completeness equation; the Lagrangian for the quantum entity was derived from its Hamiltonian. For a quantum particle, the Feynman path integral is seen to be the summation over all the continuous paths from the initial to the final position. From a mathematical point of view, the Feynman path integral is an integration over all the values of the degree of freedom for each instant between the initial and final states and is an infinite dimensional functional integral.

The path integral can directly provide the quantum theory for a given entity and instead of starting from its Hamiltonian operator, as is the case for the Schrödinger and Heisenberg formulations, one starts from its Lagrangian. The path integral is postulated to represent the quantum behavior of the given quantum entity. To show the equivalence of path integral to Schrödinger's state vector formulation, the Hamiltonian and state space of the quantum entity were derived from the Lagrangian.

The Feynman path integral is the point of departure for analyzing more complex systems and in particular for studying quantum field theory. The formulation of non-Abelian gauge fields and fermions and the study of renormalization are most effectively expressed in terms of the path integral. The crowning achievement of high energy physics, namely, the Standard Model of particles and forces, is formulated in terms of the Feynman path integral. The theory of superstrings is exploring new frontiers of physics and mathematics using the mathematical tool of the path integral [5].

Conclusions

12

Quantum mechanics is an empirical science, with experimental observations being the final and sole criterion of what is true and what is false. The founders of quantum mechanics, in particular Niels Bohr and Werner Heisenberg, were at pains to emphasize that theoretical physics should and could explain only the results of experiments. They stayed away from trying to explain what is Nature *as such*, independent of observations, with the implicit message that such an explanation would have no appropriate basis.

The theoretical superstructure of quantum mechanics—to which Heisenberg made unequaled contributions—is a mathematical construct of the human mind. The symbols and icons that are indispensable for explaining quantum phenomena are the free creations of human consciousness. There is a complex web of interpretations that finally relates the mathematical symbols to experimentally observable quantities.

How should we view the symbols of quantum mechanics? What is the ontological essence of these symbols? This question was not clearly addressed by the founders of quantum mechanics. The symbols were taken to be mathematical tools, a sort of a "mathematical language," necessary for describing and explaining quantum entities and processes. The ontological significance of the mathematical structures of quantum mechanics was not addressed, leaving a major conceptual gap in the theoretical edifice of quantum mechanics.

The pioneers of quantum mechanics probably considered the ontological significance of the symbols to be a question that could not be addressed experimentally and hence did not pursue this question. However, after a century of unbroken experimental success of quantum mechanics, the symbols are now seen to be a crucial link in connecting the structures of quantum mechanics to observed phenomena. This holds out the prospect that the *symbols themselves* contain more information than simply what they predict for experiments. In other words, the symbols themselves reflect a truth about Nature and need to be studied from this point of view.

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8_12,

There is an agnostic view that the symbols of quantum mechanics don't have any "existence" but, rather, are mathematical tools to be used solely for computing physically observable effects. Although the agnostic view is consistent, it is incomplete. My own view, and the one consistently taken in this book, is quite the opposite of the agnostic view. The theoretical symbols of quantum mechanics are not considered as being "imaginary" and fictitious, having no ontological reality, having no "being." In fact, the entire thrust of this book is to postulate the existence of a trans-empirical domain that only be represented by mathematical symbols and theoretical constructs.

The symbolic trans-empirical realm is another form of existence of Nature, a realm that goes beyond the empirical domain. Quantum mechanics is the expression and manifestation of the symbolic and trans-empirical realm of Nature.

The quantum degree of freedom, which is at the foundation of the quantum entity, can *never* be directly observed in any experiment, being trans-empirical and indeterminate and having a purely symbolic form of existence. The quantum state that describes the properties of the degree of freedom *straddles* the trans-empirical and empirical domains; the enigmatic "collapse" of the state vector, precipitated by experimentally observing the quantum state, is a transition of the quantum state from its trans-empirical form to its empirical manifestation.

The experimental apparatus is mathematically represented by Hermitian operators, which act on the state vector and cause it to collapse to one of many possible final states. It is *impossible* to predict what is the outcome of any particular experiment, reflecting the intrinsic indeterminacy of the quantum degree of freedom. Repeated experiments yield the underlying statistical regularities of the quantum entity. In the final analysis, the only quantity that one can measure is the likelihood of a specific detector, representing a projection operator, detecting the collapse of the state vector. Probabilities are assigned to the likelihood of the various projection operators detecting the state vector's collapse, and this is all that experiments can observe; the theoretical structure of quantum mechanics computes these observed probabilities.

This, in essence, is the theoretical schema of quantum mechanics that has successfully withstood the tests, for over a century, of numerous and varied experiments.

12.1 Three Formulations of Quantum Mechanics

Quantum mechanics has the following three independent, but equivalent, mathematical formulations for describing quantum indeterminacy:

- The Schrödinger equation for the state vector postulates that a quantum state vector encodes all the information that can be extracted from a quantum degree of freedom. The degree of freedom forever remains trans-empirical since all measurements only encounter the quantum state vector, causing it to collapse to an observed manifestation.

- The Heisenberg operator formalism. The state vector is completely dispensed with and a pure density matrix, which is an operator, represents the quantum entity. All observations consist of detecting of the collapse of the density matrix, which makes a transition from the pure to a mixed density matrix; the detection of the mixed density matrix by projection operators results in the experimental determination of the probability of the various projection operators detecting the quantum entity.

  Quantum probability assigns probabilities to projection operators. The trans-empirical nature of the degree of freedom is reflected in that it is never detected by any of the operators. The BKS inequality shows that the quantum indeterminacy cannot be explained by classical probability theory; in particular, the degree of freedom has no objective existence before an observation showing its trans-empirical nature.
- The Feynman path integral. The quantum degrees of freedom appear as integration variables in the path integral and provide the clearest representation of the trans-empirical degree of freedom. An integration variable has no fixed value but, rather, takes values over its entire range; for the degree of freedom, this means that the entire degree of freedom space $\mathcal{F}$ is integrated over. The freedom to change variables for path integration is equivalent to changing the representation chosen for the degree of freedom and is similar to the freedom in choosing basis states for Hilbert space.

  The path integral was derived as the sum over all the trans-empirical paths, from the initial to the final state, and reflects quantum indeterminacy that is at the foundation of quantum mechanics. The state vector appears as initial and final conditions for the trans-empirical paths that are being summed over.

Each framework has its own advantages, throwing light on different aspects of quantum mechanics that would be otherwise difficult to express. For example, the Schrödinger equation is most suitable for studying the bound sates of a quantum entity; the Heisenberg formulation is most suited for studying the structure of quantum probability, and the Feynman path integral is most appropriate for studying indeterminate and trans-empirical paths.

12.2 Interpretations of Quantum Mechanics

Some of the main interpretations of quantum mechanics have been discussed in various chapters, in particular in Sect. 9.11, and are summarized below to gain a perspective on the different approaches to quantum mechanics.

- *The Copenhagen interpretation*. This view, pioneered by Heisenberg and Bohr, is the standard approach and taught in most textbooks. This interpretation holds that quantum mechanics does not provide a description of an "objective reality"—namely, a reality that exists independent of experiment—but rather is a theory that provides a probabilistic prediction of the results of experiments. The measurement process causes one of the possible outcomes to be actualized.

The outcome of quantum experiments cannot be explained by using a solely "particle" or a "wave" description, leading to the famous "wave-particle" duality; the wave-particle duality is discussed in Sect. 3.10. Bohr further developed this duality into the law of complementarity.

- *The many-worlds interpretation*. In this view, there is no quantum uncertainty, rather, the Universe has potentially infinite many branches; an apparent random outcome of an experiment in effect results in the Universe choosing a particular branch. Every experiment results in the bifurcation of the Universe into branches.
- *Bohm's interpretation*. In this approach, there is no indeterminacy, but rather, the Universe is taken to be determined by the laws of classical mechanics. To explain the results of experiments it is assumed that every particle has an associated "pilot wave" and that results in the "wave-particle" duality of quantum mechanics.
- *'t Hooft's Planckian determinism*. In a more recent development, Gerard 't Hooft developed the idea of a deterministic theory at the Planck scale that results in the apparent quantum randomness at the macroscopic scale. He introduces the idea of "beables" and "changeables" to explain the observed behavior of quantum phenomena.
- *The trans-empirical interpretation*. The approach followed in this book. Nature is taken to have two distinct realms, namely, the empirical realm that is observed in daily life, with all entities appearing to be determinate and particular, and the trans-empirical realm that, in principle, cannot be experimentally observed and is represented by the symbols of quantum mechanics.

  The quantum entity is an inseparable pair, consisting of the trans-empirical degree of freedom and the state vector that straddles the empirical and trans-empirical domains; an experimental observation causes a transition of the quantum state from the trans-empirical to the empirical domain.

The mathematical and symbolical representation of Nature, as exemplified in quantum mechanics, provides a means for understanding of Nature that direct perception using our five senses can never provide. The process of reasoning, reflection, and symbolical thinking comes to the fore in our encounter with physical phenomena that are far removed from everyday life. The study of quantum mechanics leads to the conclusion that Nature at the deepest and most fundamental level is indeed amenable to only representations using symbols and mathematics.

The proposal presented in this book is to interpret the symbols of quantum mechanics as being expressions of a realm of Nature that can never be directly empirically observed; this realm, termed as trans-empirical, has an existence as fundamental as the empirical and observed realm. The trans-empirical realm can be grasped only by the human mind—using theory, symbols, signs, and icons that are mathematical in nature and form the superstructure of quantum mechanics.

The best result of this interpretation would be to provide a perspective on quantum mechanics that is different from the current mainstream view and which, in turn, could lead to new experiments and novel insights on the inner workings of quantum mechanics.

In conclusion, in quantum mechanics the trans-empirical realm becomes ‘visible’ to consciousness in the form of the state function, which straddles both the trans-empirical and empirical domains. The degree of freedom is entirely trans-empirical. The quantum entity is an inseparable pair, namely the degree of freedom and its state function. The trans-empirical realm exists as such in Nature; it can be cognized only by human consciousness, using signs and icons; this realm cannot be directly observed by our five senses or by any experimental device. The mathematical symbols of quantum mechanics provide a specific and particular representation of the trans-empirical realm.

Glossary of Terms

Action. The time integral of the Lagrangian.
Bra and ket vectors. Dirac's notation with the "ket" vector $|\chi\rangle$ representing an element of the state space and the "bra" vector $\langle\psi|$ representing a vector from the dual state space and with $\langle\psi|\chi\rangle$ being a complex number.
Completeness equation. Equation is a statement that the basis states for a state space are linearly independent and span the entire state space.
Contextuality. The observed properties of an entity depend on what other properties are measured. A purely quantum mechanical effect; all classical properties are non-contextual.
Determinate. An entity that is in a definite state; an entity that is empirical.
Density matrix. A description of the quantum entity using operators and which is equivalent to the state vector description.
Dual state space. A space associated with a vector space, consisting of all mappings of the state space into the complex numbers.
Entangled state. A quantum mechanical entity for which its two or more degrees of freedom cannot be viewed in isolation from each other.
Eigenfunctions. Special state vectors that are associated with an operator such that under the action of the operator, they are only changed up to a multiplicative constant, called the **eigenvalues**.
Exist. Describes any entity that "is," namely, has being, and does not necessarily have an objective and empirical existence.
Empirical. Empirical quantities are based on observations. Empirical entities are accessible to direct experimental observations.
Hamiltonian. A Hermitian operator H that is the quantum mechanical generalization of energy. H is the differential operator that evolves the system in time.
Hermitian operators are invariant under conjugation.
Hilbert space. A linear vector space for which all the state vectors have unit norm.
Indeterminate. An indeterminate entity is trans-empirical, namely, has a form of existence that is not directly observable.
Indeterminacy. The property of indeterminate entities. Quantum uncertainty is termed indeterminate to differentiate it from classical randomness.
Indeterminate path. An entity's path being indeterminate means that it simultaneously exists in all of its allowed determinate paths.
Lagrangian. A function of a determinate path.

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8, © Springer Science+Business Media New York 2013

Measurement. The collapse of a state vector by the application of projection operators that correspond to an experimental device.
Ontology. From the Greek term for "being"; that which "is," the present participle of the verb "be"; the term is used for the nature of being, of existence, or of reality.
Operators. The generalization of matrices that act on the state space. Empirically observable quantities are represented by Hermitian operators.
Operator conjugation. The transposition and complex conjugation of operators.
Path. A trajectory in time, usually denoted by $x(t)$, where t is time.
Path integral. An infinite-dimensional integral over all the possible indeterminate paths taken by a quantum entity.
Probability. The theory for explaining random and uncertain behavior.
Quantum degree of freedom. A quantity that exists in many possible states simultaneously, inherently indeterminate and trans-empirical.
Quantum degree of freedom space. The degree of freedom constitutes the space $\mathcal{F}$, which is invariant and unchanging over time.
Quantum entity. A quantum entity is constituted by a pair, namely, the degree(s) of freedom $\mathcal{F}$ and the state vector $\psi(\mathcal{F})$ that encodes all of its properties.
Random variables. Random variables describe classical random phenomena and are described by a joint probability distribution.
Real. Refers only to the result of observations, to what is empirical. Real entities exist objectively.
State vector. The state vector is a function of the degree of freedom space $\mathcal{F}$ and carries all the information that can be extracted from $\mathcal{F}$.
State space. A linear vector space, the generalization of a finite-dimensional vector space, that contains the state vectors of a quantum entity.
Superposition. The adding of state vectors; the adding of paths that are indeterminate.
Trans-empirical. The trans-empirical domain is inaccessible to direct observation and is accessible only to theory or to symbolic representations.
Uncertainty. A term reserved for describing the intrinsic indeterminateness and lack of definiteness of quantum phenomena.

List of Symbols

Only new symbols introduced in a chapter are listed. A consistent system of notation has been used as far as possible.

Chapter 2: The Quantum Entity and Quantum Mechanics

S	Action
$\mathcal{L}$	Lagrangian
$\mathcal{F}$	Degree of freedom space
$\psi(\mathcal{F})$	State vector of the degree of freedom $\mathcal{F}$
$\mathcal{O}(\mathcal{F})$	Operators of the degree of freedom $\mathcal{F}$
$P(t,x)$	Probability of an observation by an operator at x and at time t
$\psi(t,\mathcal{F})$	Time-dependent state vector
$E_\psi[\mathcal{O}(\mathcal{F})]$	Expectation value of $\mathcal{O}(\mathcal{F})$ for $\psi(t,\mathcal{F})$
H	Hamiltonian operator
ϕ_n	Probability amplitude for a determinate path labeled by n
$\phi(x_\mathrm{f},t_\mathrm{f};x_\mathrm{i},t_\mathrm{i})$	Transition amplitude from x_i at time t_i to x_f at time t_f

Chapter 3: Quantum Mechanics: Empirical and Trans-empirical

P_D	Probability of detection of state vector at screen with path taken by electron being known
P_I	Probability of detection of state vector at screen with path taken by electron *not* being known

Chapter 4: Degree of Freedom $\mathcal{F}$; State space $\mathcal{V}$

$\lvert\psi\rangle$	Ket state vector
$\langle\chi\rvert$	Bra state vector
$\mathbb{C}$	Complex numbers
$\langle\chi\lvert\psi\rangle$	Scalar product $\in \mathbb{C}$
$\mathcal{F}^N$	Space of N-degrees of freedom

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8, © Springer Science+Business Media New York 2013

$\Re^{3N}$	$3N$-dimensional Euclidean space
$\|n\rangle$	Column vector
$\langle n\|$	Row vector
δ_{n-m}	Kronecker delta function
$\|n\rangle\langle n\|$	Matrix with single entry at diagonal position n,n
$\delta(x-y)$	Dirac delta function
$\|x\rangle$	Ket vector at position x
$\langle x\|$	Bra state vector at position x
$\|x\rangle\langle x\|$	Projection operator at x
$\int_{-\infty}^{\infty} \mathrm{d}x\|x\rangle\langle x\|$	Sum over all position projection operators
$\mathcal{U}$	Unitary operator
$\mathcal{U}^{\dagger}$	Hermitian conjugate of operator

Chapter 5: Operators

$\mathcal{V}\otimes\mathcal{V}_{\mathrm{D}}$	Tensor product of state space with its dual
$D(\mathcal{O})$	Domain of $\mathcal{V}$ on which $\mathcal{O}$ acts
$D(\mathcal{O}^{\dagger})$	Domain of $\mathcal{V}$ on which $\mathcal{O}^{\dagger}$ acts
$\langle\chi\|\mathcal{O}\|\psi\rangle$	Matrix element of $\mathcal{O}$ for state vectors $\langle\chi\|$ and $\|\psi\rangle$
$\mathrm{tr}(\mathcal{O})$	Trace of operator $\mathcal{O}$
$\|\psi_n\rangle$	Eigenstate of $\mathcal{O}$
λ_n	Eigenvalue of $\mathcal{O}$
$\Pi_n = \|\psi_n\rangle\langle\psi_n\|$	Projection operator
$\|\psi_{n;n_1,n_2,\ldots,n_N}\rangle$	Energy eigenstate n with quantum numbers $n_1,n_2,\ldots,n_N$
$\|\psi_{t;n_1,n_2,\ldots,n_N}\rangle$	Time-dependent state vector with quantum numbers $n_1,n_2,\ldots,n_N$
$\hat{x}$	Position operator
$\hat{p}$	Momentum operator
$T(x)$	Unitary position shift operator
$\rho = \|\psi\rangle\langle\psi\|$	Density matrix for state vector $\|\psi\rangle$

Chapter 6: Density Matrix: Entangled States

$\mathcal{V}\otimes\mathcal{W}$	Tensor product of two state spaces
$\|\psi\rangle\otimes\|\chi\rangle$	Tensor product of two state vectors
$\|\psi\rangle\|\chi\rangle$	Tensor product of two state vectors
$\|\chi\rangle\otimes\langle\psi\|$	Outer product of two state vectors
ρ_{P}	Pure density matrix
ρ_{M}	Mixed density matrix
ρ_{R}	Reduced density matrix
ρ_{T}	Thermal density matrix
$\sigma_1,\sigma_2,\sigma_3$	Pauli 2×2 spin matrices
$\|\Psi_{\mathrm{E}}\rangle$	Entangled state vector
$\mathcal{S}$	Quantum entropy
$\exp\{-H/kT\}$	Boltzmann distribution

Chapter 7: Quantum Indeterminacy

$\mathcal{B}$	The Bell-CHSH operator
R_q	Absolute value of the expectation value of the Bell-CHSH operator
ω	Classical random sample value
Ω	Classical sample space
X, Y, Z	Classical random variables
$P(X, Y, Z)$	Classical joint probability distribution
$P(A\|B)$	Classical conditional probability distribution
$P(X, Y\|Z)$	Classical conditional probability distribution
R_c	Absolute value of the expectation value of the classical H random function
R_q^{S}	Absolute value of the expectation value of the Bell-CHSH operator for separable systems
ρ_{AB}	Bipartite density matrix
Γ_{mn}	Adjacency matrix for spin 1 BKS inequality
Q_n	Operators for spin 1 BKS inequality
q_n	Classical random variables for spin 1 BKS inequality
$\mathcal{O}_i$	Commuting operators
$\mathcal{J}_i$	Non-commuting operators

Chapter 8: Quantum Superposition

$\langle x\|s\rangle$	Probability amplitude for going from initial state $\|s\rangle$ to final state $\langle x\|$
$\langle x\|i\rangle\langle i\|s\rangle$	Probability amplitude for taking determinate path from initial state $\|s\rangle$ to final state $\langle x\|$ via the slit at $\|i\rangle$
$d_1,\ d_2$	Detectors for observing single photons
$\mathcal{B}$	Unitary operator representing the beam splitter
$\mathcal{M}$	Unitary operator representing the mirror
$\mathcal{P}$	Unitary operator representing the phase shifter
$\|\Psi_{\mathrm{I}}\rangle$	Initial state vector
$\|\Psi_{\mathrm{F}}\rangle$	Final state vector
$d_1 \otimes d_3$	Detectors for observing two coincident photons
U	Interaction of spin states $\|s\rangle$ with device

Chapter 9: Quantum Theory of Measurement

$E_\chi[\mathcal{O}]$	Expectation value of $\mathcal{O}$ for state vector $\|\chi\rangle$
$\|D_n\rangle$	Detector states
x_n	Detector readings
$\mathcal{V}_{\mathrm{Q}}$	Hilbert space of the quantum entity
$\mathcal{V}_{\mathrm{D}}$	Hilbert space of the detector
$\mathcal{V}_{\mathrm{Q}} \otimes \mathcal{V}_{\mathrm{D}}$	Hilbert space of the quantum entity and detector

$\mathcal{O}_{\mathrm{E}}$	Operator $\mathcal{O}$ extended to space $\mathcal{V}_Q \otimes \mathcal{V}_{\mathrm{D}}$
H_{Q}	Hamiltonian of the quantum entity
H_{QD}	Hamiltonian coupling the quantum entity and to the detector
$\lvert\Phi_{\mathrm{in}}\rangle$	Initial state of the quantum entity and the detector
$\lvert\Phi_{\mathrm{out}}\rangle$	Final state of the quantum entity and the detector
$\tilde{\rho}_{\mathrm{M}}$	Mixed density matrix of the quantum entity and the detector
$\mathcal{O}_{\mathrm{E}\lvert\chi}$	Partial trace of $\mathcal{O}_{\mathrm{E}}$ over $\mathcal{V}_Q$
$\lvert\phi\rangle_{\mathrm{box}}$	State vector for particle in a box
ΔA	Uncertainty in quantity A

Chapter 10: The Stern-Gerlach Experiment

ξ_+, ξ_-	Spin eigenstates
$\psi^{\mu}_{\mathrm{E}}(\mathbf{r};\mathbf{p})$	Energy eigenstates for the Stern-Gerlach Hamiltonian
$\Psi_{\mathrm{in}}(\mathbf{r})$,	Incoming state vector for the Stern-Gerlach experiment
$\Psi^{\mu}_{\mathrm{M}}(\mathbf{r})$,	State vector propagating in the magnetic field for the Stern-Gerlach experiment
$\Psi^{\mu}_{\mathrm{out}}(\mathbf{r})$	The final state vector for the Stern-Gerlach experiment
$g(\mathbf{p})$	Gaussian wave packet
χ_{μ}	State vector propagating in the magnetic field
ζ_{μ}	State vector after crossing the magnetic field

Chapter 11: The Feynman Path Integral

$\langle x_{\mathrm{f}}, t_{\mathrm{f}} \lvert x_{\mathrm{i}}, t_{\mathrm{i}}\rangle$	Probability amplitude for transition from initial x_{i} at time t_{i} to final position x_{f} at time t_{f}
$K(x, x'; t)$	Evolution kernel; transition amplitude
$\mathrm{P}(x_{\mathrm{f}} \lvert x_{\mathrm{i}}; t)$	Conditional probability for the occurrence of x_{f} given x_{i} occurred at earlier time t
$\phi[\mathrm{path}]$	Probability amplitude for discrete and determinate path
$S[\mathrm{path}]$	Action for discrete path
$S[x(t)]$	Action for continuous path $x(t)$
$x_{\mathrm{c}}(t)$	Classical path
$\int DX$	Path integral measure

References

1. Peres, A.: Quantum Theory: Concepts and Methods. Kluwer, Holland (1998)
2. Aspect, A.: Bell's inequality test: more ideal than ever. Nature **398**(189), 1408–1427 (1999)
3. Baaquie, B.E.: Quantum Finance. Cambridge University Press, Cambridge (2004)
4. Ballentine, L.E.: Quantum Mechanics: A Modern Development. World Scientific, Singapore (1998)
5. Baaquie, B.E.: Path Integrals in Quantum Mechanics, Quantum Field Theory and Superstrings. Cambridge University Press, Cambridge (2013)
6. Bell, J.: Speakable and Unspeakable in Quantum Mechanics. Cambridge University Press, Cambridge (2004)
7. Odom, B., Hanneke, D., D'Urso, B., Gabrielse, G.: New measurement of the electron magnetic moment using a one-electron quantum cyclotron. Phys. Rev. Lett. **97**, 030801 (2006)
8. DeWitt, B.S., Graham, N.: The Many-Worlds Interpretation of Quantum Mechanics. Princeton University Press, Princeton (1973)
9. Clauser, J.F., Horne, M.A., Shimony, A., Holt, R.A.: Proposed experiment to test local hidden-variable theories. Phys. Rev. Lett. **23**, 880–884 (1969)
10. Dirac, P.A.M.: The Principles of Quantum Mechanics. Oxford University Press, Oxford (1999)
11. Einstein, A., Podolsky, B., Rosen, N.: Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777–780 (1935)
12. Wigner, E.P.: The problem of measurement. Am. J. Phys. **31**, 6–15 (1963)
13. Feynman, R.P.: The Character of Physical Law. Penguin Books, Baltimore (2007)
14. Feynman, R.P., Hibbs, A.R.: Quantum Mechanics and Path Integrals. McGraw Hill, New York (1960)
15. Gottfried, K., Yan, T.-M.: Quantum Mechanics. Springer, Germany (2003)
16. Greenstein, G., Zajonc, A.G.: The Quantum Challenge, 2nd edn. Jones and Bartlett, Boston (2006)
17. Heisenberg, W.: The Physical Principals of the Quantum Theory. Dover, New York (1949)
18. Heisenberg, W.: Physics and Philosophy: The Revolution in Modern Science. Prometheus Books, New York (1999)
19. Isham, C.J.: Lectures on Quantum Theory. Imperial College Press, London (1995)
20. Klyachko, A.A., Can, M.A., Binicioğlu, S., Shumovsky, A.S.: Simple test for hidden variables in spin-1 systems. Phys. Rev. Lett. **101**, 020403 (2008)
21. Kochen, S., Specker, E.P.: The problem of hidden variables in quantum mechanics. J. Math. Mech. **17** (1967)
22. Kurzyński, P., Ramanathan, R., Kaszlikowski, D.: Entropic test of quantum contextuality. Phys. Rev. Lett. **109**, 020404 (2012)
23. Lawden, D.F.: The Mathematical Principles of Quantum Mechanics. Dover, New york (2005)
24. Feynman, R.P., Leighton, R.B., Sands, M.: The Feynman Lectures on Physics. Addison-Wesley, Reading (1964)
25. Mackey, G.W.: Mathematical Foundations of Quantum Mechanics. Dover, New York (2004)

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8, © Springer Science+Business Media New York 2013

26. Major, F.G., Gheorghe, V.N., Werth, G.: Charged Particle Traps: Physics and Techniques of Charged Particle Field Confinement. Springer, Germany (2010)
27. Nielsen, M.A., Chang, I.L.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2000)
28. Muga, G.: Time in Quantum Mechanics. Springer, Berlin (2008)
29. Muller, H., Peter, A., Chew, S.: A precision measurement of the gravitational redshift by the interference of matter waves. Nature **463**(3), 926–929 (1983)
30. Newton, R.G.: The Truth of Science: Physical Theories and Reality. Harvard University Press, Cambridge (1997)
31. Healy, R.: The Philosophy of Quantum Mechanics. Cambridge University Press, Cambridge (2008)
32. Ramanathan, R., Soeda, A.A., Kurzyński, P., Kaszlikowsk, D.: Generalized monogamy of contextual inequalities from the no-disturbance principle. Phys. Rev. Lett. **109**, 050404 (2012)
33. Schlosshauer, M.A.: Decoherence: and the Quantum-to-Classical Transition. Springer, Germany (2010)
34. Stapp, H.P.: The Copenhagen interpretation. Am. J. Phys. **40** (1963)
35. Streater, R.F.: Classical and quantum probability. J. Math. Phys. **41**, 3556–3603 (2000)
36. von Neumann, J.: Mathematical Foundations of Quantum Mechanics. Princeton University Press, Princeton (1983)
37. Yu, S., Oh, C.H.: State-independent proof of Kochen-Specker theorem with 13 rays. Phys. Rev. Lett. **108**, 030402 (2012)

Index

B.E. Baaquie, *The Theoretical Foundations of Quantum Mechanics*,
DOI 10.1007/978-1-4614-6224-8, © Springer Science+Business Media New York 2013