Lecture 3

The Shankar strategy
Teach quantum using postulates
Teach the math first then the postulates

Okay, but there is a lot of math!
What math do we need?

We need the math for the postulates
So we have the chicken or egg problem

“tell them what you are going to tell them”

Today
postulates and the related math
Dirac notation
photon polarization

Math Topics

Operators

Observable operators
Observables
Hamiltonian operator
Position operator
Momentum operator
Linear operators
Identity operator
Hermitean operators
AntiHermitean operators
Unitary operators
Projection operators
adjoint operation

Vectors

bra
ket
state vector
inner product
outer product
basis
change basis
characteristic equation
adjoint operation
orthogonal
normal
orthonormal expansion
eigenvalues
eigenvectors
eigensubspace

Measurement
Dirac delta function
Time-dependent Schrodinger equation
Time-independent Schrodinger equation
Dirac notation
Hilbert Space
discrete non-degenerate spectrum
discrete spectrum with degeneracy
continuous non-degenerate spectrum

Preface to the First Edition

Publish and perish—*Giordano Bruno*

Given the number of books that already exist on the subject of quantum mechanics, one would think that the public needs one more as much as it does, say, the latest version of the Table of Integers. But this does not deter me (as it didn't my predecessors) from trying to circulate my own version of how it ought to be taught. The approach to be presented here (to be described in a moment) was first tried on a group of Harvard undergraduates in the summer of '76, once again in the summer of '77, and more recently at Yale on undergraduates ('77–'78) and graduates ('78–'79) taking a year-long course on the subject. In all cases the results were very satisfactory in the sense that the students seemed to have learned the subject well and to have enjoyed the presentation. It is, in fact, their enthusiastic response and encouragement that convinced me of the soundness of my approach and impelled me to write this book.

The basic idea is to develop the subject from its postulates, after addressing some indispensable preliminaries. Now, most people would agree that the best way to teach any subject that has reached the point of development where it can be reduced to a few postulates is to start with the latter, for it is this approach that gives students the fullest understanding of the foundations of the theory and how it is to be used. But they would also argue that whereas this is all right in the case of special relativity or mechanics, a typical student about to learn quantum mechanics seldom has any familiarity with the mathematical language in which the postulates are stated. I agree with these people that this problem is real, but I differ in my belief that it should and can be overcome. This book is an attempt at doing just this.

It begins with a rather lengthy chapter in which the relevant mathematics of vector spaces developed from simple ideas on vectors and matrices the student is assumed to know. The level of rigor is what I think is needed to make a practicing quantum mechanic out of the student. This chapter, which typically takes six to eight lecture hours, is filled with examples from physics to keep students from getting too fidgety while they wait for the "real physics." Since the math introduced has to be taught sooner or later, I prefer sooner to later, for this way the students, when they get to it, can give quantum theory their fullest attention without having to

battle with the mathematical theorems at the same time. Also, by segregating the mathematical theorems from the physical postulates, any possible confusion as to which is which is nipped in the bud.

This chapter is followed by one on classical mechanics, where the Lagrangian and Hamiltonian formalisms are developed in some depth. It is for the instructor to decide how much of this to cover; the more students know of these matters, the better they will understand the connection between classical and quantum mechanics. Chapter 3 is devoted to a brief study of idealized experiments that betray the inadequacy of classical mechanics and give a glimpse of quantum mechanics.

Having trained and motivated the students I now give them the postulates of quantum mechanics of a single particle in one dimension. I use the word "postulate" here to mean "that which cannot be deduced from pure mathematical or logical reasoning, and given which one can formulate and solve quantum mechanical problems and interpret the results." This is not the sense in which the true axiomatist would use the word. For instance, where the true axiomatist would just postulate that the dynamical variables are given by Hilbert space operators, I would add the operator identifications, i.e., specify the operators that represent coordinate and momentum (from which others can be built). Likewise, I would not stop with the statement that there is a Hamiltonian operator that governs the time evolution through the equation $i\hbar\partial|\psi\rangle/\partial t = H|\psi\rangle$; I would say the H is obtained from the classical Hamiltonian by substituting for x and p the corresponding operators. While the more general axioms have the virtue of surviving as we progress to systems of more degrees of freedom, with or without classical counterparts, students given just these will not know how to calculate anything such as the spectrum of the oscillator. Now one can, of course, try to "derive" these operator assignments, but to do so one would have to appeal to ideas of a postulatory nature themselves. (The same goes for "deriving" the Schrödinger equation.) As we go along, these postulates are generalized to more degrees of freedom and it is for pedagogical reasons that these generalizations are postponed. Perhaps when students are finished with this book, they can free themselves from the specific operator assignments and think of quantum mechanics as a general mathematical formalism obeying certain postulates (in the strict sense of the term).

The postulates in Chapter 4 are followed by a lengthy discussion of the same, with many examples from fictitious Hilbert spaces of three dimensions. Nonetheless, students will find it hard. It is only as they go along and see these postulates used over and over again in the rest of the book, in the setting up of problems and the interpretation of the results, that they will catch on to how the game is played. It is hoped they will be able to do it on their own when they graduate. I think that any attempt to soften this initial blow will be counterproductive in the long run.

Chapter 5 deals with standard problems in one dimension. It is worth mentioning that the scattering off a step potential is treated using a wave packet approach. If the subject seems too hard at this stage, the instructor may decide to return to it after Chapter 7 (oscillator), when students have gained more experience. But I think that sooner or later students must get acquainted with this treatment of scattering.

The classical limit is the subject of the next chapter. The harmonic oscillator is discussed in detail in the next. It is the first realistic problem and the instructor may be eager to get to it as soon as possible. If the instructor wants, he or she can discuss the classical limit after discussing the oscillator.

Contents

4

The Postulates—a General Discussion

Having acquired the necessary mathematical training and physical motivation, you are now ready to get acquainted with the postulates of quantum mechanics. In this chapter the postulates will be stated and discussed in broad terms to bring out the essential features of quantum theory. The subsequent chapters will simply be applications of these postulates to the solution of a variety of physically interesting problems. Despite your preparation you may still find the postulates somewhat abstract and mystifying on this first encounter. These feelings will, however, disappear after you have worked with the subject for some time.

4.1. The Postulates‡

The following are the postulates of nonrelativistic quantum mechanics. We consider first a system with one degree of freedom, namely, a single particle in one space dimension. The straightforward generalization to more particles and higher dimensions will be discussed towards the end of the chapter. In what follows, the quantum postulates are accompanied by their classical counterparts (in the Hamiltonian formalism) to provide some perspective.

Classical Mechanics	Quantum Mechanics
I. The state of a particle at any given time is specified by the two variables $x(t)$ and $p(t)$, i.e., as a point in a two-dimensional phase space.	I. The state of the particle is represented by a vector $\|\psi(t)\rangle$ in a Hilbert space. **1**
II. Every dynamical variable ω is a function of x and p: $\omega = \omega(x, p)$.	II. The independent variables x and p of classical mechanics are represented **2**

‡ Recall the discussion in the Preface regarding the sense in which the word is used here.

by Hermitian operators X and P with the following matrix elements in the eigenbasis of X‡

$$\langle x|X|x'\rangle = x\delta(x-x')$$
$$\langle x|P|x'\rangle = -i\hbar\delta'(x-x') \qquad \textbf{2}$$

The operators corresponding to dependent variables $\omega(x, p)$ are given Hermitian operators

$$\Omega(X, P) = \omega(x\to X, p\to P)\S$$

III. If the particle is in a state given by x and p, the measurement‖ of the variable ω will yield a value $\omega(x, p)$. The state will remain unaffected.

III. If the particle is in a state $|\psi\rangle$, measurement‖ of the variable (corresponding to) Ω will yield one of the eigenvalues ω with probability $P(\omega)\propto|\langle\omega|\psi\rangle|^2$. The state of the system will change from $|\psi\rangle$ to $|\omega\rangle$ as a result of the measurement. **3**

IV. The state variables change with time according to Hamilton's equations:

$$\dot{x} = \frac{\partial\mathscr{H}}{\partial p}$$

$$\dot{p} = -\frac{\partial\mathscr{H}}{\partial x}$$

IV. The state vector $|\psi(t)\rangle$ obeys the *Schrödinger equation*

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle \qquad \textbf{4}$$

where $H(X, P) = \mathscr{H}(x\to X, p\to P)$ is the quantum Hamiltonian operator and $\mathscr{H}$ is the Hamiltonian for the corresponding classical problem.

4.2. Discussion of Postulates I–III

The postulates (of classical and quantum mechanics) fall naturally into two sets: the first three, which tell us how the system is depicted at a given time, and the last, which specifies how this picture changes with time. We will confine our attention to the first three postulates in this section, leaving the fourth for the next.

The first postulate states that a particle is described by a ket $|\psi\rangle$ in a Hilbert space which, you will recall, contains *proper vectors* normalizable to unity as well as

‡ Note that the X operator is the same one discussed at length in Section 1.10. Likewise $P=\hbar K$, where K was also discussed therein. You may wish to go over that section now to refresh your memory.

§ By this we mean that Ω is the same function of X and P as ω is of x and P.

‖ That is, in an ideal experiment consistent with the theory. It is assumed you are familiar with the ideal classical measurement which can determine the state of the system without disturbing it in any way. A discussion of ideal quantum measurements follows.

CHAPTER 3

THE POSTULATES OF QUANTUM MECHANICS. OPERATORS, EIGENFUNCTIONS, AND EIGENVALUES

In this chapter we consider four basic postulates of quantum mechanics, which when taken with the Born postulate described in Section 2.8, serve to formalize the rules of quantum mechanics. Mathematical concepts material to these postulates are developed along with the physics. The postulates are applied over and over again throughout the text. We choose the simplest problems first to exhibit their significance and method of application—that is, problems in one dimension.

3.1 OBSERVABLES AND OPERATORS

Postulate I

This postulate states the following: To any self-consistently and well-defined observable in physics (call it A), such as linear momentum, energy, mass, angular momentum, or number of particles, there corresponds an operator (call it $\hat{A}$) such that measurement of A yields values (call these measured values a) which are eigenvalues of $\hat{A}$. That is, the values, a, are those values for which the equation

(3.1) $$\hat{A}\varphi = a\varphi$$

an eigenvalue equation

3.2 MEASUREMENT IN QUANTUM MECHANICS

Postulate II

The second postulate[1] of quantum mechanics is: measurement of the observable A that yields the value a leaves the system in the state φ_a, where φ_a is the eigenfunction of $\hat{A}$ that corresponds to the eigenvalue a.

As an example, suppose that a free particle is moving in one dimension. We do not know which state the particle is in. At a given instant we measure the particle's momentum and find the value $p = \hbar k$ (with k a specific value, say $1.3 \times 10^{10}\ \text{cm}^{-1}$). This measurement[2] leaves the particle in the state φ_k, so immediate subsequent measurement of p is certain to yield $\hbar k$.

Suppose that one measures the position of a free particle and the position $x = x'$ is measured. The first two postulates tell us the following. (1) There is an operator corresponding to the measurement of position, call it $\hat{x}$. (2) Measurement of x that yields the value x' leaves the particle in the eigenfunction of $\hat{x}$ corresponding to the eigenvalue x'.

The operator equation appears as

$$\hat{x}\delta(x - x') = x'\delta(x - x') \tag{3.26}$$

Dirac Delta Function

The eigenfunction of $\hat{x}$ has been written[3] $\delta(x - x')$ and is called the *Dirac delta function*. It is defined in terms of the following two properties. The first are the integral properties

$$\int_{-\infty}^{\infty} f(x')\delta(x - x')\,dx' = f(x)$$

$$\int_{-\infty}^{\infty} \delta(x - x')\,dx' = 1 \tag{3.27}$$

[1] This postulate has been the source of some discussion among physicists. For further reference, see B. S. DeWitt, *Phys. Today* **23**, 30 (September 1970).

[2] Measurement is taken in the idealized sense. More formal discussions on the theory of measurement may be found in K. Gottfried, *Quantum Mechanics*, W. A. Benjamin, New York, 1966; J. Jauch, *Foundations of Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1968, and E. C. Kemble, *The Fundamental Principles of Quantum Mechanics with Elementary Applications*, Dover, New York, 1958.

[3] More accurately one says that $\delta(x - x')$ is an eigenfunction of $\hat{x}$ in the coordinate representation. This topic is returned to in Section 7.4 and in Appendix A.

3.3 THE STATE FUNCTION AND EXPECTATION VALUES

Postulate III

The third postulate of quantum mechanics establishes the existence of the state function and its relevance to the properties of a system: The state of a system at any instant of time may be represented by a state or wave function ψ which is continuous and differentiable. All information regarding the state of the system is contained in the wavefunction. Specifically, if a system is in the state $\psi(\mathbf{r}, t)$, the average of any physical observable C relevant to that system at time t is

$$\langle C \rangle = \int \psi^* \hat{C} \psi \, d\mathbf{r} \tag{3.32}$$

(The differential of volume is written $d\mathbf{r}$.) The average, $\langle C \rangle$, is called the *expectation value* of C.

The physical meaning of the average of an observable C involves the following type of (conceptual) measurements. The observable C is measured in a specific experiment, X. One prepares a very large number (N) of identical replicas of X. The initial states $\psi(\mathbf{r}, 0)$ in each such replica are all identical. At the time t, one measures C in all these replica experiments and obtains the set of values $C_1, C_2, \ldots, C_N$. The average of C is then given by the rule

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i \qquad (N \gg 1) \tag{3.33}$$

The postulate stated above claims that this experimentally calculated average (3.33) is the same as that given by the integral in (3.32). Another way of defining $\langle C \rangle$ is in terms of the probability $P(C_i)$. This function gives the probability that measurement of C finds the value C_i. For $\langle C \rangle$, we then have

$$\langle C \rangle = \sum_{\text{all } C} C_i P(C_i) \tag{3.34}$$

This is a consistent formula if all the values C may assume comprise a discrete set (e.g., the number of marbles in a box). In the event that the values that C may assume comprise a continuous set (e.g., the values of momentum of a free particle), $\langle C \rangle$ becomes

$$\langle C \rangle = \int C P(C) \, dC \tag{3.35}$$

The integration is over all values of C. Here $P(C)$ is the probability of finding C in the interval $C, C + dC$.

The quantity $\langle C \rangle$ is also called the *expectation value* of C because it is representative of the value one expects to obtain in any given measurement of C. This will

3.4 TIME DEVELOPMENT OF THE STATE FUNCTION

Postulate IV

The fourth postulate of quantum mechanics specifies the time development of the state function $\psi(\mathbf{r}, t)$: the state function for a system (e.g., a single particle) develops in time according to the equation

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \hat{H}\psi(\mathbf{r}, t) \tag{3.45}$$

This equation is called the *time-dependent Schrödinger* equation.[1] The operator $\hat{H}$ is the Hamiltonian operator. For a single particle of mass m, in a potential field $V(\mathbf{r})$, it is given by (3.12). If $\hat{H}$ is assumed to be independent of time, we may write

$$\hat{H} = \hat{H}(\mathbf{r}) \tag{3.46}$$

Under these circumstances, one is able to construct a solution to the time-dependent Schrödinger equation through the technique of separation of variables. We assume a solution of the form

$$\psi(\mathbf{r}, t) = \varphi(\mathbf{r})T(t) \tag{3.47}$$

Substitution into (3.45) gives

$$i\hbar \frac{T_t}{T} = \frac{\hat{H}\varphi}{\varphi} \tag{3.48}$$

The subscript t denotes differentiation with respect to t. Equation (3.48) is such that the left-hand side is a function of t only, while the right-hand side is a function of $\mathbf{r}$ only. Such an equation can be satisfied only if both sides are equal to the same constant, call it E (we do not yet know that E is the energy).

$$\hat{H}\varphi(\mathbf{r}) = E\varphi(\mathbf{r}) \tag{3.49}$$

$$\left(\frac{\partial}{\partial t} + \frac{iE}{\hbar}\right) T(t) = 0 \tag{3.50}$$

The first of these equations is the time-independent Schrödinger equation (3.13). This identification serves to label E, in (3.49), the energy of the system. That is, E, as it appears in this equation, is an eigenvalue of $\hat{H}$. But the eigenvalues of $\hat{H}$ are the allowed energies a system may assume, and we again conclude that E is the energy of the system.

[1] A formulation of the Schrödinger equation that has its origin in the classical principle of least action has been offered by R. P. Feynman, *Rev. Mod. Phys.* **60**, 367 (1948). An elementary description of this derivation may be found in S. Borowitz, *Quantum Mechanics*, W. A. Benjamin, New York, 1967.

Chapter 3

Cohen-Tannoudji

OUTLINE OF CHAPTER III

A. INTRODUCTION

B. STATEMENT OF THE POSTULATES

1. Description of the state of a system
2. Description of physical quantities
3. The measurement of physical quantities
   - *a.* Possible results
   - *b.* Principle of spectral decomposition
   - *c.* Reduction of the wave packet
4. Time evolution of systems
5. Quantization rules
   - *a.* Statement
   - *b.* Important examples

C. THE PHYSICAL INTERPRETATION OF THE POSTULATES CONCERNING OBSERVABLES AND THEIR MEASUREMENT

1. The quantization rules are consistent with the probabilistic interpretation of the wave function
2. Quantization of certain physical quantities
3. The measurement process
4. Mean value of an observable in a given state
5. The root-mean-square deviation
6. Compatibility of observables
   - *a.* Compatibility and commutation rules
   - *b.* Preparation of a state

D. THE PHYSICAL IMPLICATIONS OF THE SCHRÖDINGER EQUATION

1. General properties of the Schrödinger equation
   - *a.* Determinism in the evolution of physical systems
   - *b.* The superposition principle
   - *c.* Conservation of probability
   - *d.* Evolution of the mean value of an observable; relation to classical mechanics
2. The case of conservative systems
   - *a.* Solution of the Schrödinger equation
   - *b.* Stationary states
   - *c.* Constants of the motion
   - *d.* Bohr frequencies of a system. Selection rules
   - *e.* The time-energy uncertainty relation

E. THE SUPERPOSITION PRINCIPLE AND PHYSICAL PREDICTIONS

1. Probability amplitudes and interference effects
   - *a.* The physical meaning of a linear superposition of states
   - *b.* Summation over the intermediate states
   - *c.* Conclusion: the importance of the concept of probability amplitudes
2. Case in which several states can be associated with the same measurement result
   - *a.* Degenerate eigenvalues
   - *b.* Insufficiently selective measurement devices
   - *c.* Recapitulation: must one sum the amplitudes or the probabilities?
   - *d.* Application to the treatment of continuous spectra

the corresponding function $\psi(\mathbf{r}) = \langle \mathbf{r} | \psi \rangle$. Therefore, the quantum state of a particle at a fixed time is characterized by a ket of the space $\mathscr{E}_\mathbf{r}$. In this form, the concept of a state can be generalized to any physical system.

First Postulate: At a fixed time t_0, the state of a physical system is defined by specifying a ket $|\psi(t_0)\rangle$ belonging to the state space $\mathscr{E}$. **1**

It is important to note that, since $\mathscr{E}$ is a vector space, this first postulate implies a superposition principle : a linear combination of state vectors is a state vector. We shall discuss this fundamental point and its relations to the other postulates in §E.

2. Description of physical quantities

We have already used, in §D-1 of chapter I, a differential operator H related to the total energy of a particle in a scalar potential. This is simply a special case of the second postulate.

Second Postulate: Every measurable physical quantity $\mathscr{A}$ is described by an operator A acting in $\mathscr{E}$; this operator is an observable. **2**

COMMENTS:

(*i*) The fact that A is an observable (*cf.* chap. II, §D-2) will be seen below (§3) to be essential.

(*ii*) Unlike classical mechanics (*cf.* §A), quantum mechanics describes in a fundamentally different manner the state of a system and the associated physical quantities : a state is represented by a vector, a physical quantity by an operator.

3. The measurement of physical quantities

a. POSSIBLE RESULTS

The connection between the operator H and the total energy of the particle appeared in §D-1 of chapter I in the following form: the only energies possible are the eigenvalues of the operator H. Here as well, this relation can be extended to all physical quantities.

Third Postulate: The only possible result of the measurement of a physical quantity $\mathscr{A}$ is one of the eigenvalues of the corresponding observable A. **3**

COMMENTS:

(*i*) A measurement of $\mathscr{A}$ always gives a real value, since A is by definition Hermitian.

(*ii*) If the spectrum of A is discrete, the results that can be obtained by measuring $\mathscr{A}$ are quantized (§ C-2).

b. PRINCIPLE OF SPECTRAL DECOMPOSITION

We are going to generalize and discuss in more detail the conclusions of § A-3 of chapter I, where we analyzed a simple experiment performed on polarized photons.

Consider a system whose state is characterized, at a given time, by the ket $|\psi\rangle$, assumed to be normalized to 1:

$$\langle\psi|\psi\rangle = 1 \tag{B-1}$$

We want to predict the result of the measurement, at this time, of a physical quantity $\mathscr{A}$ associated with the observable A. This prediction, as we already know, is of a probabilistic sort. We are now going to give the rules which allow us to calculate the probability of obtaining any given eigenvalue of A.

α. *Case of a discrete spectrum*

First, let us assume that the spectrum of A is entirely discrete. If all the eigenvalues a_n of A are non-degenerate, there is associated with each of them a unique (to within a constant factor) eigenvector $|u_n\rangle$:

$$A|u_n\rangle = a_n|u_n\rangle \tag{B-2}$$

Since A is an observable, the set of the $|u_n\rangle$, which we shall take to be normalized, constitutes a basis in $\mathscr{E}$, and the state vector $|\psi\rangle$ can be written:

$$|\psi\rangle = \sum_n c_n|u_n\rangle \tag{B-3}$$

We postulate that the probability $\mathscr{P}(a_n)$ of finding a_n when $\mathscr{A}$ is measured is:

$$\mathscr{P}(a_n) = |c_n|^2 = |\langle u_n|\psi\rangle|^2 \tag{B-4}$$

> *Fourth Postulate (case of a discrete non-degenerate spectrum)*: When the physical quantity $\mathscr{A}$ is measured on a system in the *normalized* state $|\psi\rangle$, the probability $\mathscr{P}(a_n)$ of obtaining the *non-degenerate* eigenvalue a_n of the corresponding observable A is:
>
> $$\mathscr{P}(a_n) = |\langle u_n|\psi\rangle|^2$$
>
> where $|u_n\rangle$ is the normalized eigenvector of A associated with the eigenvalue a_n.

If, now, some of the eigenvalues a_n are degenerate, several orthonormalized eigenvectors $|u_n^i\rangle$ correspond to them:

$$A|u_n^i\rangle = a_n|u_n^i\rangle; \quad i = 1, 2, \ldots g_n \tag{B-5}$$

$|\psi\rangle$ can still be expanded in the orthonormal basis $\{|u_n^i\rangle\}$:

$$|\psi\rangle = \sum_n \sum_{i=1}^{g_n} c_n^i|u_n^i\rangle \tag{B-6}$$

In this case, the probability $\mathcal{P}(a_n)$ becomes:

$$\mathcal{P}(a_n) = \sum_{i=1}^{g_n} |c_n^i|^2 = \sum_{i=1}^{g_n} |\langle u_n^i | \psi \rangle|^2 \tag{B-7}$$

(B-4) is then seen to be a special case of (B-7), which can therefore be considered to be the general formula.

Fourth Postulate (case of a discrete spectrum): When the physical quantity $\mathcal{A}$ is measured on a system in the *normalized* state $|\psi\rangle$, the probability $\mathcal{P}(a_n)$ of obtaining the eigenvalue a_n of the corresponding observable A is:

$$\mathcal{P}(a_n) = \sum_{i=1}^{g_n} |\langle u_n^i | \psi \rangle|^2$$

where g_n is the degree of degeneracy of a_n and $\{ |u_n^i\rangle \}$ $(i = 1, 2, ..., g_n)$ is an orthonormal set of vectors which forms a basis in the eigensubspace $\mathcal{E}_n$ associated with the eigenvalue a_n of A.

4'

For this postulate to make sense, it is obviously necessary that, if the eigenvalue a_n is degenerate, the probability $\mathcal{P}(a_n)$ be independent of the choice of the $\{ |u_n^i\rangle \}$ basis in $\mathcal{E}_n$. To verify this, consider the vector:

$$|\psi_n\rangle = \sum_{i=1}^{g_n} c_n^i |u_n^i\rangle \tag{B-8}$$

where the coefficients c_n^i are the same as those appearing in the expansion (B-6) of $|\psi\rangle$:

$$c_n^i = \langle u_n^i | \psi \rangle \tag{B-9}$$

$|\psi_n\rangle$ is the part of $|\psi\rangle$ which belongs to $\mathcal{E}_n$, that is, the projection of $|\psi\rangle$ onto $\mathcal{E}_n$. This is, moreover, what we find when we substitute (B-9) into (B-8):

$$\begin{aligned} |\psi_n\rangle &= \sum_{i=1}^{g_n} |u_n^i\rangle \langle u_n^i | \psi \rangle \\ &= P_n |\psi\rangle \end{aligned} \tag{B-10}$$

where:

$$P_n = \sum_{i=1}^{g_n} |u_n^i\rangle \langle u_n^i| \tag{B-11}$$

is the projector onto $\mathcal{E}_n$ (§ B-3-b of chapter II). Let us now calculate the square of the norm of $|\psi_n\rangle$. From (B-8):

$$\langle \psi_n | \psi_n \rangle = \sum_{i=1}^{g_n} |c_n^i|^2 \tag{B-12}$$

Therefore, $\mathcal{P}(a_n)$ *is the square of the norm of* $|\psi_n\rangle = P_n|\psi\rangle$, the projection of $|\psi\rangle$ onto $\mathcal{E}_n$. From this expression, it is clear that a change in the basis in $\mathcal{E}_n$ does not affect $\mathcal{P}(a_n)$. This probability is written:

$$\mathcal{P}(a_n) = \langle \psi | P_n^\dagger P_n | \psi \rangle \tag{B-13}$$

or, using the fact that P_n is Hermitian ($P_n^\dagger = P_n$) and that it is a projector ($P_n^2 = P_n$):

$$\mathscr{P}(a_n) = \langle \psi | P_n | \psi \rangle \tag{B-14}$$

β. *Case of a continuous spectrum*

Now let us assume that the spectrum of A is continuous and, for the sake of simplicity, non-degenerate. The system, orthonormal in the extended sense, of eigenvectors $|v_\alpha\rangle$ of A:

$$A|v_\alpha\rangle = \alpha|v_\alpha\rangle \tag{B-15}$$

forms a continuous basis in $\mathscr{E}$, in terms of which $|\psi\rangle$ can be expanded:

$$|\psi\rangle = \int d\alpha\, c(\alpha)|v_\alpha\rangle \tag{B-16}$$

Since the possible results of a measurement of $\mathscr{A}$ form a continuous set, we must define a probability density, just as we did for the interpretation of the wave function of a particle (§ B-2 of chapter I). The probability $d\mathscr{P}(\alpha)$ of obtaining a value included between α and $\alpha + d\alpha$ is given by:

$$d\mathscr{P}(\alpha) = \rho(\alpha)\, d\alpha$$

with:

$$\rho(\alpha) = |c(\alpha)|^2 = |\langle v_\alpha|\psi\rangle|^2 \tag{B-17}$$

Fourth Postulate (case of a continuous non-degenerate spectrum): When the physical quantity $\mathscr{A}$ is measured on a system in the *normalized* state $|\psi\rangle$, the probability $d\mathscr{P}(\alpha)$ of obtaining a result included between α and $\alpha + d\alpha$ is equal to:

$$d\mathscr{P}(\alpha) = |\langle v_\alpha|\psi\rangle|^2\, d\alpha$$

where $|v_\alpha\rangle$ is the eigenvector corresponding to the eigenvalue α of the observable A associated with $\mathscr{A}$.

4"

COMMENTS:

(*i*) It can be verified explicitly, in each of the cases considered above, that the total probability is equal to 1. For example, starting with formula (B-7), we find:

$$\sum_n \mathscr{P}(a_n) = \sum_n \sum_{i=1}^{g_n} |c_n^i|^2 = \langle \psi | \psi \rangle = 1 \tag{B-18}$$

since $|\psi\rangle$ is normalized. This last condition is therefore indispensable if the statements we have made are to be coherent. Nevertheless, it is not essential: if it is not fulfilled, it suffices to replace (B-7) and (B-17), respectively, by:

$$\mathscr{P}(a_n) = \frac{1}{\langle \psi | \psi \rangle} \sum_{i=1}^{g_n} |c_n^i|^2 \tag{B-19}$$

over time, that is, enables us to calculate from $|\psi(0)\rangle$ the state $|\psi(t_0)\rangle$ "immediately before" the measurement. If the measurement has yielded the non-degenerate eigenvalue a_n, the state $|\psi'(t_1)\rangle$ at a time $t_1 > t_0$ must be calculated from $|\psi'(t_0)\rangle = |u_n\rangle$, the state "immediately after" the measurement, using the sixth postulate to determine the evolution of the state vector between the times t_0 and t_1 (fig. 1).

(*ii*) If we perform a second measurement of $\mathscr{A}$ immediately after the first one (that is, before the system has had time to evolve), we shall always find the same result a_n, since the state of the system immediately before the second measurement is $|u_n\rangle$, and no longer $|\psi\rangle$.

When the eigenvalue a_n given by the measurement is degenerate, postulate (B-28) can be generalized as follows. If the expansion of the state $|\psi\rangle$ immediately before the measurement is written, with the same notation as in section b:

$$|\psi\rangle = \sum_n \sum_{i=1}^{g_n} c_n^i |u_n^i\rangle \tag{B-29}$$

the modification of the state vector due to the measurement is written:

$$|\psi\rangle \xrightarrow{(a_n)} \frac{1}{\sqrt{\sum_{i=1}^{g_n} |c_n^i|^2}} \sum_{i=1}^{g_n} c_n^i |u_n^i\rangle \tag{B-30}$$

$\sum_{i=1}^{g_n} c_n^i |u_n^i\rangle$ is the vector $|\psi_n\rangle$ defined above [formula (B-8)], that is, the projection of $|\psi\rangle$ onto the eigensubspace associated with a_n. In (B-30), we normalized this vector since it is always more convenient to use state vectors of norm 1 [comment (*i*) of § b above]. With the notation of (B-10) and (B-11), we can therefore write (B-30) in the form:

$$|\psi\rangle \xrightarrow{(a_n)} \frac{P_n|\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}} \tag{B-31}$$

Fifth Postulate: If the measurement of the physical quantity $\mathscr{A}$ on the system in the state $|\psi\rangle$ gives the result a_n, the state of the system immediately after the measurement is the normalized projection, $\frac{P_n|\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}}$, of $|\psi\rangle$ onto the eigensubspace associated with a_n.

5

The state of the system immediately after the measurement is therefore always an eigenvector of A with the eigenvalue a_n. We stress the fact, however, that *it is not an arbitrary ket of the subspace* $\mathscr{E}_n$, but the part of $|\psi\rangle$ which belongs to $\mathscr{E}_n$ (suitably normalized, for convenience). In the light of § 3-b-γ above, equation (B-28) can be seen to be a special case of (B-30). When $g_n = 1$, the summation over i disappears from (B-30), which becomes:

$$\frac{1}{|c_n|} c_n |u_n\rangle = e^{i\,\mathrm{Arg}\,c_n} |u_n\rangle \tag{B-32}$$

This ket indeed describes the same physical state as $|u_n\rangle$.

4. Time evolution of systems

We have already presented, in § B-2 of chapter I, the Schrödinger equation for one particle. Here we shall write it in the general case.

Sixth Postulate: The time evolution of the state vector $|\psi(t)\rangle$ is governed by the Schrödinger equation:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H(t) |\psi(t)\rangle$$

where $H(t)$ is the observable associated with the total energy of the system.

6

H is called the *Hamiltonian operator* of the system, as it is obtained from the classical Hamiltonian (appendix III and §5 below).

5. Quantization rules

We are finally going to discuss how to construct, for a physical quantity $\mathscr{A}$ already defined in classical mechanics, the operator A which describes it in quantum mechanics.

a. STATEMENT

Let us first consider a system composed of a single particle, without spin, subject to a scalar potential. In this case:

With the position $\mathbf{r}(x, y, z)$ of the particle is associated the observable $\mathbf{R}(X, Y, Z)$. With the momentum $\mathbf{p}(p_x, p_y, p_z)$ of the particle is associated the observable $\mathbf{P}(P_x, P_y, P_z)$.

Recall that the components of $\mathbf{R}$ and $\mathbf{P}$ satisfy the canonical commutation relations [chap. II, equations (E-30)]:

$$[R_i, R_j] = [P_i, P_j] = 0$$
$$[R_i, P_j] = i\hbar\delta_{ij} \qquad \text{(B-33)}$$

Any physical quantity $\mathscr{A}$ related to this particle is expressed in terms of the fundamental dynamical variables $\mathbf{r}$ and $\mathbf{p}$: $\mathscr{A}(\mathbf{r}, \mathbf{p}, t)$. To obtain the corresponding observable A, one could simply replace, in the expression for $\mathscr{A}(\mathbf{r}, \mathbf{p}, t)$, the variables $\mathbf{r}$ and $\mathbf{p}$ by the observables $\mathbf{R}$ and $\mathbf{P}$*:

$$A(t) = \mathscr{A}(\mathbf{R}, \mathbf{P}, t) \qquad \text{(B-34)}$$

* See, in complement B_{II}, the definition of a function of an operator.

However, this mode of action would be, in general, ambiguous. Assume, for example, that in $\mathscr{A}(\mathbf{r}, \mathbf{p}, t)$ there appears a term of the form :

$$\mathbf{r} \cdot \mathbf{p} = xp_x + yp_y + zp_z \tag{B-35}$$

In classical mechanics, the scalar product $\mathbf{r} \cdot \mathbf{p}$ is commutative, and one can just as well write :

$$\mathbf{p} \cdot \mathbf{r} = p_x x + p_y y + p_z z \tag{B-36}$$

But when $\mathbf{r}$ and $\mathbf{p}$ are replaced by the corresponding observables $\mathbf{R}$ and $\mathbf{P}$, the operators obtained from (B-35) and (B-36) are not identical [see relations (B-33)] :

$$\mathbf{R} \cdot \mathbf{P} \neq \mathbf{P} \cdot \mathbf{R} \tag{B-37}$$

Moreover, neither $\mathbf{R} \cdot \mathbf{P}$ nor $\mathbf{P} \cdot \mathbf{R}$ is Hermitian :

$$(\mathbf{R} \cdot \mathbf{P})^\dagger = (XP_x + YP_y + ZP_z)^\dagger = \mathbf{P} \cdot \mathbf{R} \tag{B-38}$$

To the preceding postulates, therefore, must be added a symmetrization rule. For example, the observable associated with $\mathbf{r} \cdot \mathbf{p}$ will be :

$$\frac{1}{2}(\mathbf{R} \cdot \mathbf{P} + \mathbf{P} \cdot \mathbf{R}) \tag{B-39}$$

which is indeed Hermitian. For an observable which is more complicated than $\mathbf{R} \cdot \mathbf{P}$, an analogous symmetrization is to be performed.

> The observable A which describes a classically defined physical quantity $\mathscr{A}$ is obtained by replacing, in the suitably symmetrized expression for $\mathscr{A}$, $\mathbf{r}$ and $\mathbf{p}$ by the observables $\mathbf{R}$ and $\mathbf{P}$ respectively.

R

We shall see, however, that there exist quantum physical quantities which have no classical equivalent and which are therefore defined directly by the corresponding observables (this is the case, for example, for particle spin).

COMMENT :

The preceding rules, and commutation rules (B-33) in particular, are valid only in cartesian coordinates. It would be possible to generalize them to other coordinate systems ; however, they would no longer have the same simple form as they do above.

b. IMPORTANT EXAMPLES

α. *The Hamiltonian of a particle in a scalar potential*

Consider a (spinless) particle of charge q and mass m, placed in an electric field derived from a scalar potential $U(\mathbf{r})$. The potential energy of the particle is therefore $V(\mathbf{r}) = qU(\mathbf{r})$, and the corresponding classical Hamiltonian is written [appendix III, formula (29)] :

$$\mathscr{H}(\mathbf{r}, \mathbf{p}) = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) \tag{B-40}$$

CHAPTER 3

FORMALISM

3.1 HILBERT SPACE

In the last two chapters we have stumbled on a number of interesting properties of simple quantum systems. Some of these are "accidental" features of specific potentials (the even spacing of energy levels for the harmonic oscillator, for example), but others seem to be more general, and it would be nice to prove them once and for all (the uncertainty principle, for instance, and the orthogonality of stationary states). The purpose of this chapter is to recast the theory in a more powerful form, with that in mind. There is not much here that is genuinely *new*; the idea, rather, is to make coherent sense of what we have already discovered in particular cases.

Quantum theory is based on two constructs: *wave functions* and *operators*. The state of a system is represented by its wave function, observables are represented by operators. Mathematically, wave functions satisfy the defining conditions for abstract **vectors**, and operators act on them as **linear transformations**. So the natural language of quantum mechanics is **linear algebra**.[1]

But it is not, I suspect, a form of linear algebra with which you are immediately familiar. In an N-dimensional space it is simplest to represent a vector, $|\alpha\rangle$, by the N-tuple of its components, $\{a_n\}$, with respect to a specified orthonormal basis:

$$|\alpha\rangle \rightarrow \mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}. \qquad [3.1]$$

[1] If you have never studied linear algebra, you should read the Appendix before continuing.

The **inner product**, $\langle\alpha|\beta\rangle$, of two vectors (generalizing the dot product in three dimensions) is the complex number,

$$\langle\alpha|\beta\rangle = a_1^* b_1 + a_2^* b_2 + \cdots + a_N^* b_N. \qquad [3.2]$$

Linear transformations, T, are represented by **matrices** (with respect to the specified basis), which act on vectors (to produce new vectors) by the ordinary rules of matrix multiplication:

$$|\beta\rangle = T|\alpha\rangle \rightarrow \mathbf{b} = \mathbf{T}\mathbf{a} = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1N} \\ t_{21} & t_{22} & \cdots & t_{2N} \\ \vdots & \vdots & & \vdots \\ t_{N1} & t_{N2} & \cdots & t_{NN} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}. \qquad [3.3]$$

But the "vectors" we encounter in quantum mechanics are (for the most part) *functions*, and they live in *infinite*-dimensional spaces. For them the N-tuple/matrix notation is awkward, at best, and manipulations that are well-behaved in the finite-dimensional case can be problematic. (The underlying reason is that whereas the *finite* sum in Equation 3.2 always exists, an *infinite* sum—or an integral—may not converge, in which case the inner product does not exist, and any argument involving inner products is immediately suspect.) So even though most of the terminology and notation should be familiar, it pays to approach this subject with caution.

The collection of *all* functions of x constitutes a vector space, but for our purposes it is much too large. To represent a possible physical state, the wave function Ψ must be *normalized*:

$$\int |\Psi|^2 \, dx = 1.$$

The set of all **square-integrable functions**, on a specified interval,[2]

$$f(x) \quad \text{such that} \quad \int_a^b |f(x)|^2 \, dx < \infty, \qquad [3.4]$$

constitutes a (much smaller) vector space (see Problem 3.1(a)). Mathematicians call it $L^2(a, b)$; physicists call it **Hilbert space**.[3] In quantum mechanics, then,

Wave functions live in Hilbert space. 1 [3.5]

[2] For us, the limits (a and b) will almost always be $\pm\infty$, but we might as well keep things more general for the moment.

[3] Technically, a Hilbert space is a **complete inner product space**, and the collection of square-integrable functions is only *one example* of a Hilbert space—indeed, every finite-dimensional vector space is trivially a Hilbert space. But since L^2 is the arena of quantum mechanics, it's what physicists generally *mean* when they say "Hilbert space." By the way, the word **complete** here means that any Cauchy sequence of functions in Hilbert space converges to a function that is also in the space: it has no "holes" in it, just as the set of all real numbers has no holes (by contrast, the space of all *polynomials*, for example, like the set of all *rational* numbers, certainly *does* have holes in it). The completeness of a *space* has nothing to do with the completeness (same word, unfortunately) of a *set of functions*, which is the property that any other function can be expressed as a linear combination of them.

We define the **inner product of two functions**, $f(x)$ and $g(x)$, as follows:

$$\langle f|g\rangle \equiv \int_a^b f(x)^* g(x)\,dx. \qquad [3.6]$$

If f and g are both square-integrable (that is, if they are both in Hilbert space), their inner product is guaranteed to exist (the integral in Equation 3.6 converges to a finite number).[4] This follows from the integral **Schwarz inequality**:[5]

$$\left|\int_a^b f(x)^* g(x)\,dx\right| \leq \sqrt{\int_a^b |f(x)|^2\,dx \int_a^b |g(x)|^2\,dx}. \qquad [3.7]$$

You can check for yourself that Equation 3.6 satisfies all the conditions for an inner product (Problem 3.1(b)). Notice in particular that

$$\langle g|f\rangle = \langle f|g\rangle^*. \qquad [3.8]$$

Moreover, the inner product of $f(x)$ with *itself*,

$$\langle f|f\rangle = \int_a^b |f(x)|^2\,dx, \qquad [3.9]$$

is *real* and non-negative; it's *zero* only[6] when $f(x) = 0$.

A function is said to be **normalized** if its inner product with itself is 1; two functions are **orthogonal** if their inner product is 0; and a *set* of functions, $\{f_n\}$, is **orthonormal** if they are normalized and mutually orthogonal:

$$\langle f_m|f_n\rangle = \delta_{mn}. \qquad [3.10]$$

Finally, a set of functions is **complete** if any *other* function (in Hilbert space) can be expressed as a linear combination of them:

$$f(x) = \sum_{n=1}^{\infty} c_n f_n(x). \qquad [3.11]$$

---

[4] In Chapter 2 we were obliged on occasion to work with functions that were *not* normalizable. Such functions lie *outside* Hilbert space, and we are going to have to handle them with special care, as you will see shortly. For the moment, I shall assume that all the functions we encounter *are* in Hilbert space.

[5] For a proof, see F. Riesz and B. Sz.-Nagy, *Functional Analysis* (Unger, New York, 1955), Section 21. In a *finite* dimensional vector space the Schwarz inequality, $|\langle\alpha|\beta\rangle|^2 \leq \langle\alpha|\alpha\rangle\langle\beta|\beta\rangle$, is easy to prove (see Problem A.5). But that proof *assumes* the existence of the inner products, which is precisely what we are trying to *establish* here.

[6] What about a function that is zero everywhere except at a few isolated points? The integral (Equation 3.9) would still vanish, even though the function itself does not. If this bothers you, you should have been a math major. In physics such pathological functions do not occur, but in any case, in Hilbert space two functions that have the same square integral are considered equivalent. Technically, vectors in Hilbert space represent **equivalence classes** of functions.

Now, the outcome of a measurement has got to be *real*, and so, *a fortiori*, is the *average* of many measurements:

$$\langle Q \rangle = \langle Q \rangle^*. \qquad [3.14]$$

But the complex conjugate of an inner product reverses the order (Equation 3.8), so

$$\langle \Psi | \hat{Q}\Psi \rangle = \langle \hat{Q}\Psi | \Psi \rangle. \qquad [3.15]$$

and this must hold true for any wave function Ψ. Thus operators representing *observables* have the very special property that

$$\langle f | \hat{Q} f \rangle = \langle \hat{Q} f | f \rangle \quad \text{for all } f(x). \qquad [3.16]$$

We call such operators **hermitian.**

Actually, most books require an ostensibly stronger condition:

$$\langle f | \hat{Q} g \rangle = \langle \hat{Q} f | g \rangle \quad \text{for all } f(x) \text{ and all } g(x). \qquad [3.17]$$

But it turns out, in spite of appearances, that this is perfectly equivalent to my definition (Equation 3.16), as you will prove in Problem 3.3. So use whichever you like. The essential point is that a hermitian operator can be applied either to the first member of an inner product or to the second, with the same result, and hermitian operators naturally arise in quantum mechanics because their expectation values are real:

Observables are represented by hermitian operators. **2** [3.18]

Well, let's *check* this. Is the momentum operator, for example, hermitian?

$$\langle f | \hat{p} g \rangle = \int_{-\infty}^{\infty} f^* \frac{\hbar}{i} \frac{dg}{dx}\, dx = \frac{\hbar}{i} f^* g \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \left(\frac{\hbar}{i} \frac{df}{dx} \right)^* g\, dx = \langle \hat{p} f | g \rangle. \qquad [3.19]$$

I used integration by parts, of course, and threw away the boundary term for the usual reason: If $f(x)$ and $g(x)$ are square integrable, they must go to zero at $\pm\infty$.[8]

---

space into a function *outside* it (see Problem 3.2(b)), and in this case the domain of the operator may have to be restricted.

[8] Actually, this is not quite true. As I mention in Chapter 1, there exist pathological functions that are square-integrable but do *not* go to zero at infinity. However, such functions do not arise in physics, and if you are worried about it we will simply restrict the domain of our operators to exclude them. On *finite* intervals, though, you really *do* have to be more careful with the boundary terms, and an operator that is hermitian on $(-\infty, \infty)$ may *not* be hermitian on $(0, \infty)$ or $(-\pi, \pi)$. If you're wondering about the infinite square well, it's safest to think of those wave functions as residing on the infinite line—they just happen to be *zero* outside $(0, a)$.

Notice how the complex conjugation of i compensates for the minus sign picked up from integration by parts—the operator d/dx (without the i) is *not* hermitian, and it does not represent a possible observable.

---

*Problem 3.3 Show that if $\langle h|\hat{Q}h\rangle = \langle \hat{Q}h|h\rangle$ for all functions h (in Hilbert space), then $\langle f|\hat{Q}g\rangle = \langle \hat{Q}f|g\rangle$ for all f and g (i.e., the two definitions of "hermitian"—Equations 3.16 and 3.17—are equivalent). *Hint:* First let $h = f + g$, and then let $h = f + ig$.

Problem 3.4

(a) Show that the *sum* of two hermitian operators is hermitian.

(b) Suppose $\hat{Q}$ is hermitian, and α is a complex number. Under what condition (on α) is $\alpha\hat{Q}$ hermitian?

(c) When is the *product* of two hermitian operators hermitian?

(d) Show that the position operator ($\hat{x} = x$) and the hamiltonian operator ($\hat{H} = -(\hbar^2/2m)d^2/dx^2 + V(x)$) are hermitian.

Problem 3.5 The **hermitian conjugate** (or **adjoint**) of an operator $\hat{Q}$ is the operator $\hat{Q}^\dagger$ such that

$$\langle f|\hat{Q}g\rangle = \langle \hat{Q}^\dagger f|g\rangle \quad \text{(for all } f \text{ and } g\text{)}. \tag{3.20}$$

(A hermitian operator, then, is equal to its hermitian conjugate: $\hat{Q} = \hat{Q}^\dagger$.)

(a) Find the hermitian conjugates of x, i, and d/dx.

(b) Construct the hermitian conjugate of the harmonic oscillator raising operator, a_+ (Equation 2.47).

(c) Show that $(\hat{Q}\hat{R})^\dagger = \hat{R}^\dagger\hat{Q}^\dagger$.

3.2.2 Determinate States

Ordinarily, when you measure an observable Q on an ensemble of identically prepared systems, all in the same state Ψ, you do *not* get the same result each time—this is the *indeterminacy* of quantum mechanics.[9] *Question:* Would it be possible to prepare a state such that *every* measurement of Q is certain to return the *same* value (call it q)? This would be, if you like, a **determinate state**, for the observable Q. (Actually, we already know one example: Stationary states are determinate states of the Hamiltonian; a measurement of the total energy, on a

---

[9] I'm talking about *competent* measurements, of course—it's always possible to make a *mistake*, and simply get the wrong answer, but that's not the fault of quantum mechanics.

particle in the stationary state Ψ_n, is certain to yield the corresponding "allowed" energy E_n.)

Well, the standard deviation of Q, in a determinate state, would be *zero*, which is to say,

$$\sigma^2 = \langle(\hat{Q} - \langle Q\rangle)^2\rangle = \langle\Psi|(\hat{Q} - q)^2\Psi\rangle = \langle(\hat{Q} - q)\Psi|(\hat{Q} - q)\Psi\rangle = 0. \quad [3.21]$$

(Of course, if every measurement gives q, their average is also q: $\langle Q\rangle = q$. I also used the fact that $\hat{Q}$, and hence also $\hat{Q} - q$, is a *hermitian* operator, to move one factor over to the first term in the inner product.) But the only function whose inner product with itself vanishes is 0, so

$$\hat{Q}\Psi = q\Psi. \quad [3.22]$$

This is the **eigenvalue equation** for the operator $\hat{Q}$; Ψ is an **eigenfunction** of $\hat{Q}$, and q is the corresponding **eigenvalue**. Thus

Determinate states are eigenfunctions of $\hat{Q}$. **3** [3.23]

Measurement of Q on such a state is certain to yield the eigenvalue, q.

Note that the eigen*value* is a *number* (not an operator or a function). You can multiply any eigenfunction by a constant, and it is still an eigenfunction, with the same eigenvalue. Zero does not count as an eigenfunction (we exclude it by definition—otherwise *every* number would be an eigenvalue, since $\hat{Q}\,0 = q\,0 = 0$ for any operator $\hat{Q}$ and all q). But there's nothing wrong with zero as an eigen*value*. The collection of all the eigenvalues of an operator is called its **spectrum**. Sometimes two (or more) linearly independent eigenfunctions share the same eigenvalue; in that case the spectrum is said to be **degenerate**.

For example, determinate states of the total energy are eigenfunctions of the Hamiltonian:

$$\hat{H}\psi = E\psi, \quad [3.24]$$

which is precisely the time-independent Schrödinger equation. In this context we use the letter E for the eigenvalue, and the lower case ψ for the eigenfunction (tack on the factor $\exp(-iEt/\hbar)$ to make it Ψ, if you like; it's still an eigenfunction of H).

Example 3.1 Consider the operator

$$\hat{Q} \equiv i\frac{d}{d\phi}, \quad [3.25]$$

where ϕ is the usual polar coordinate in two dimensions. (This operator might arise in a physical context if we were studying the bead-on-a-ring; see Problem 2.46.) Is $\hat{Q}$ hermitian? Find its eigenfunctions and eigenvalues.

That's why the stationary states of the infinite square well, for example, or the harmonic oscillator, are orthogonal—they are eigenfunctions of the Hamiltonian with distinct eigenvalues. But this property is not peculiar to them, or even to the Hamiltonian—the same holds for determinate states of *any* observable.

Unfortunately, Theorem 2 tells us nothing about degenerate states ($q' = q$). However, if two (or more) eigenfunctions share the same eigenvalue, any linear combination of them is itself an eigenfunction, with the same eigenvalue (Problem 3.7(a)), and we can use the **Gram-Schmidt orthogonalization procedure** (Problem A.4) to *construct* orthogonal eigenfunctions within each degenerate subspace. It is almost never necessary to do this explicitly (thank God!), but it can always be done in principle. So *even in the presence of degeneracy* the eigenfunctions can be *chosen* to be orthogonal, and in setting up the formalism of quantum mechanics we shall assume that this has already been done. That licenses the use of Fourier's trick, which depends on the orthonormality of the basis functions.

In a *finite*-dimensional vector space the eigenvectors of a hermitian matrix have a third fundamental property: They span the space (every vector can be expressed as a linear combination of them). Unfortunately, the proof does not generalize to infinite-dimensional spaces. But the property itself is essential to the internal consistency of quantum mechanics, so (following Dirac[11]) we will take it as an *axiom* (or, more precisely, as a restriction on the class of hermitian operators that can represent observables):

> **Axiom:** The eigenfunctions of an observable operator are *complete*: Any function (in Hilbert space) can be expressed as a linear combination of them.[12]

Problem 3.7

(a) Suppose that $f(x)$ and $g(x)$ are two eigenfunctions of an operator $\hat{Q}$, with the same eigenvalue q. Show that any linear combination of f and g is itself an eigenfunction of $\hat{Q}$, with eigenvalue q.

(b) Check that $f(x) = \exp(x)$ and $g(x) = \exp(-x)$ are eigenfunctions of the operator d^2/dx^2, with the same eigenvalue. Construct two linear combinations of f and g that are *orthogonal* eigenfunctions on the interval $(-1, 1)$.

---

[11] P. A. M. Dirac, *The Principles of Quantum Mechanics*, Oxford University Press, New York (1958).

[12] In some specific cases completeness is provable (we know that the stationary states of the infinite square well, for example, are complete, because of Dirichlet's theorem). It is a little awkward to call something an "axiom" that is *provable* in some cases, but I don't know a better way to handle it.

Chapter 1

Operator Methods In Quantum Mechanics

1.1 Introduction

The purpose of the first two lectures is twofold. First, to review the mathematical formalism of elementary non-relativistic quantum mechanics, especially the terminology. The second purpose is to present the basic tools of operator methods, commutation relations, shift operators, etc. and apply them to familiar problems such as the harmonic oscillator. Before we get down to the operator formalism, let's remind ourselves of the fundamental postulates of quantum mechanics as covered in earlier courses. They are:

1

- **Postulate 1:** The state of a quantum-mechanical system is completely specified by a function $\Psi(\mathbf{r}, t)$ (which in general can be complex) that depends on the coordinates of the particles (collectively denoted by $\mathbf{r}$) and on the time. This function, called the *wave function* or the state function, has the important property that $\Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t)\, d\mathbf{r}$ is the probability that the system will be found in the volume element $d\mathbf{r}$, located at $\mathbf{r}$, at the time t.

2

- **Postulate 2:** To every observable A in classical mechanics, there corresponds a *linear Hermitian operator* $\hat{A}$ in quantum mechanics.

3

- **Postulate 3:** In any measurement of the observable A, the only values that can be obtained are the *eigenvalues* $\{a\}$ of the associated operator $\hat{A}$, which satisfy the eigenvalue equation

$$\hat{A}\Psi_a = a\Psi_a$$

where Ψ_a is the *eigenfunction* of $\hat{A}$ corresponding to the eigenvalue a.

- **Postulate 4:** If a system is in a state described by a normalised wavefunction Ψ, and the eigenfunctions $\{\Psi_a\}$ of $\hat{A}$ are also normalised, then the probability of obtaining the value a in a measurement of the observable A is given by

$$P(a) = \left| \int_{-\infty}^{\infty} \Psi_a^* \Psi \, d\mathbf{r} \right|^2$$

(Recall that a function $\Phi(\mathbf{r})$ such that

$$\int_{-\infty}^{\infty} \Phi^* \Phi \, d\mathbf{r} = 1$$

is said to be *normalised*.)

- **Postulate 5:** As a result of a measurement of the observable A in which the value a is obtained, the wave function of the system becomes the corresponding eigenfunction Ψ_a. (This is sometimes called the *collapse of the wave function*.)

- **Postulate 6:** Between measurements, the wave function evolves in time according to the *time-dependent Schrödinger equation*

$$\frac{\partial \Psi}{\partial t} = -\frac{i}{\hbar} \hat{H} \Psi$$

where $\hat{H}$ is the Hamiltonian operator of the system.

The justification for the above postulates ultimately rests with experiment. Just as in geometry one sets up axioms and then logically deduces the consequences, one does the same with the postulates of QM. To date, there has been no contradiction between experimental results and the outcomes predicted by applying the above postulates to a wide variety of systems.

We now explore the mathematical structure underpinning quantum mechanics.

1.1.1 Mathematical foundations

In the standard formulation of quantum theory, the state of a physical system is described by a vector in a Hilbert space H over the complex numbers. The observables and dynamical variables of the system are represented by linear operators which transform each state vector into another (possibly the same) state vector. Throughout this course (unless stated otherwise) we will adopt Dirac's notation: thus a state vector is denoted by a *ket* $|\Psi\rangle$. This ket provides a complete description of the physical state. In the next section we will explore the mathematical properties of the Hilbert space and learn why it plays such a central role in the mathematical formulation of quantum mechanics.

LECTURE 7 7-9-08

CLASSICAL POLARIZATION APPLETS

DIRAC NOTATION

IN QM, PHOTONS HAVE 2 DEGREES OF FREEDOM

TWO STANDARD WAYS TO REPRESENT THESE 2 DOF

LINEAR POLARIZATION

CIRCULAR POLARIZATION

CAN WRITE BOTH AS A TWO COMPONENT VECTOR

$$\begin{pmatrix} a \\ b \end{pmatrix}$$

COMPONENTS DEPEND ON THE BASIS BUT THE VECTOR DOES NOT

USING LINEAR POL BASIS

$$|x\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

can make any vector in the 2d space

$$|y\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$|R\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$$

can also make any vector in the 2d space

$$|L\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

USING THE CIRCULAR POL BASIS

$$|R\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$|L\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

INNER PRODUCT make BRA KET

SCALAR PRODUCT

KET $|c\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$

KET $|d\rangle = \begin{pmatrix} \gamma \\ \delta \end{pmatrix}$

$$\langle c | d \rangle = (\alpha \quad \beta)^* \begin{pmatrix} \gamma \\ \delta \end{pmatrix}$$

$$= \alpha^* \gamma + \beta^* \delta$$

$$\langle d | c \rangle = \langle c | d \rangle^*$$

in general inner product is a complex number

LENGTH OF A VECTOR

$$\vec{a} \cdot \vec{b} = |a|\,|b| \cos\theta$$

$$\vec{a} \cdot \vec{a} = |a|^2$$

$$\text{length} = \sqrt{\vec{a} \cdot \vec{a}}$$

QM $\qquad |c\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$

$$\langle c | c \rangle = \begin{pmatrix} \alpha^* & \beta^* \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

$$= \alpha^* \alpha + \beta^* \beta$$

(real, real)

$$= \text{REAL NUMBER}$$

CM unit vectors $|\vec{a}| = 1$

QM normalized vectors

$$\langle x | x \rangle = \begin{pmatrix} 1 & 0 \end{pmatrix}^* \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$= 1$$

$$\langle y | y \rangle = \begin{pmatrix} 0 & 1 \end{pmatrix}^* \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1$$

$$\langle R | R \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \end{pmatrix}^* \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}$$

$$= \frac{1}{2} (1 + 1) = 1$$

$$\langle L | L \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \end{pmatrix}^* \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

$$= \frac{1}{2} (1 + 1) = 1$$

inner product is a map from two vectors into numbers

general: complex

special: real

OUTER PRODUCT

$$|x\rangle\langle x| = \begin{pmatrix} 1 \\ 0 \end{pmatrix}^* \begin{pmatrix} 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$|y\rangle\langle y| = \begin{pmatrix} 0 \\ 1 \end{pmatrix}^* \begin{pmatrix} 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

outer product is a map from 2 VECTORS INTO MATRIX (OPERATOR)

$$|\psi\rangle = \begin{pmatrix} a \\ b \end{pmatrix}$$

outer product first

$$(|x\rangle\langle x|)\,|\psi\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}^* \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1\,0) \begin{pmatrix} a \\ b \end{pmatrix} = a \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

inner product first

can evaluate this Dirac expression two ways: (1) do the outer product first or (2) do the inner product first

so $|x\rangle\langle x|$ picks out the x component of $|\psi\rangle$

ORTHO NORMAL BASES

$$\hat{i}\cdot\hat{j}=0 \qquad \hat{j}\cdot\hat{i}=0$$

$$\langle x|y\rangle = 0 \qquad \langle y|x\rangle = 0$$

$$\langle R|L\rangle = 0 \qquad \langle L|R\rangle = 0$$

~~$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ i\end{pmatrix}^*$~~

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & i\end{pmatrix}^*\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ -i\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & -1\end{pmatrix} = 0$$

in circular basis

$$|R\rangle = \begin{pmatrix}1\\ 0\end{pmatrix}$$

$$|L\rangle = \begin{pmatrix}0\\ 1\end{pmatrix}$$

in circ pol basis IS OBVIOUS

$$\langle R|L\rangle = \langle L|R\rangle = 0$$

GENERAL VECTOR

$$|\psi\rangle = N \begin{pmatrix} a e^{i\alpha} \\ b e^{i\beta} \end{pmatrix}$$

$$\langle\psi|\psi\rangle = \left[N \begin{pmatrix} a e^{i\alpha} & b e^{i\beta} \end{pmatrix}\right]^* \left[N \begin{pmatrix} a e^{i\alpha} \\ b e^{i\beta} \end{pmatrix}\right]$$

$$= N^* N \left(a^* a\, e^{-i\alpha} e^{i\alpha} + b^* b\, e^{-i\beta} e^{+i\beta} \right)$$

$$= |N|^2 \left(|a|^2 + |b|^2 \right)$$

$$1 = |N|^2 \left(|a|^2 + |b|^2 \right)$$

$$N = \sqrt{\frac{1}{|a|^2 + |b|^2}}$$

HOW MUCH x POL ?

$$\langle x | \psi \rangle = (1 \; 0)^* \; N \begin{pmatrix} a\, e^{i\alpha} \\ b\, e^{i\beta} \end{pmatrix}$$

$$= N \, a \, e^{i\alpha}$$

HOW MUCH y POL ?

$$\langle y | \psi \rangle = (0 \; 1)^* \; N \begin{pmatrix} a\, e^{i\alpha} \\ b\, e^{i\beta} \end{pmatrix}$$

$$= N \, b \, e^{i\beta}$$

HOW MUCH R POL ?

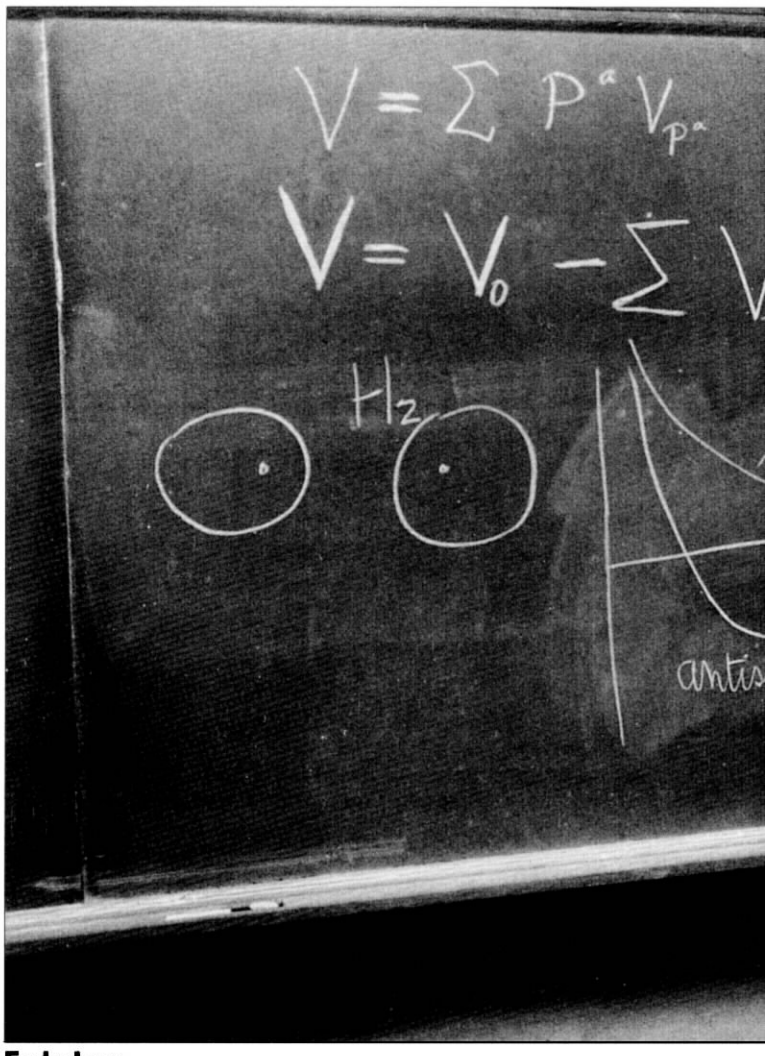
$$\langle R | \psi \rangle = \left[\frac{1}{\sqrt{2}} (1 \; i) \right]^* \; N \begin{pmatrix} a\, e^{i\alpha} \\ b\, e^{i\beta} \end{pmatrix}$$

$$= \frac{N}{\sqrt{2}} \left(a\, e^{i\alpha} - i\, b\, e^{i\beta} \right)$$

Paul Dirac published the first of his papers on "The Quantum Theory of the Electron" seventy years ago this month. The Dirac equation, derived in those papers, is one of the most important equations in physics

Paul Dirac: the purest soul in physics

Michael Berry



EACH day, I walk past the road where Paul Adrien Maurice Dirac lived as a child. It is pleasant to have even this tenuous association with one of the greatest intellects of the 20th century. Paul Dirac was born at 15 Monk Road in Bishopston, Bristol, on 8 August 1902, and educated at the nearby Bishop Road Primary School. The family later moved to Cotham Road, near the University of Bristol, and in 1914 the young Dirac joined Cotham Grammar School, formerly the Merchant Venturers.

Dirac was a student at Bristol University between 1918 and 1923, first in electrical engineering and then in applied mathematics. Much later, he said: "I owe a lot to my engineering training because it [taught] me to tolerate approximations. Previously to that I thought...one should just concentrate on exact equations all the time. Then I got the idea that in the actual world all our equations are only approximate. We must just tend to greater and greater accuracy. In spite of the equations being approximate, they can be beautiful."

Because Dirac was a quiet man – famously quiet, indeed – he is not well known outside physics, although this is slowly changing. In 1995 a plaque to Dirac was unveiled at Westminster Abbey in London and last year Institute of Physics Publishing, which is based in Bristol, named its new building Dirac House.

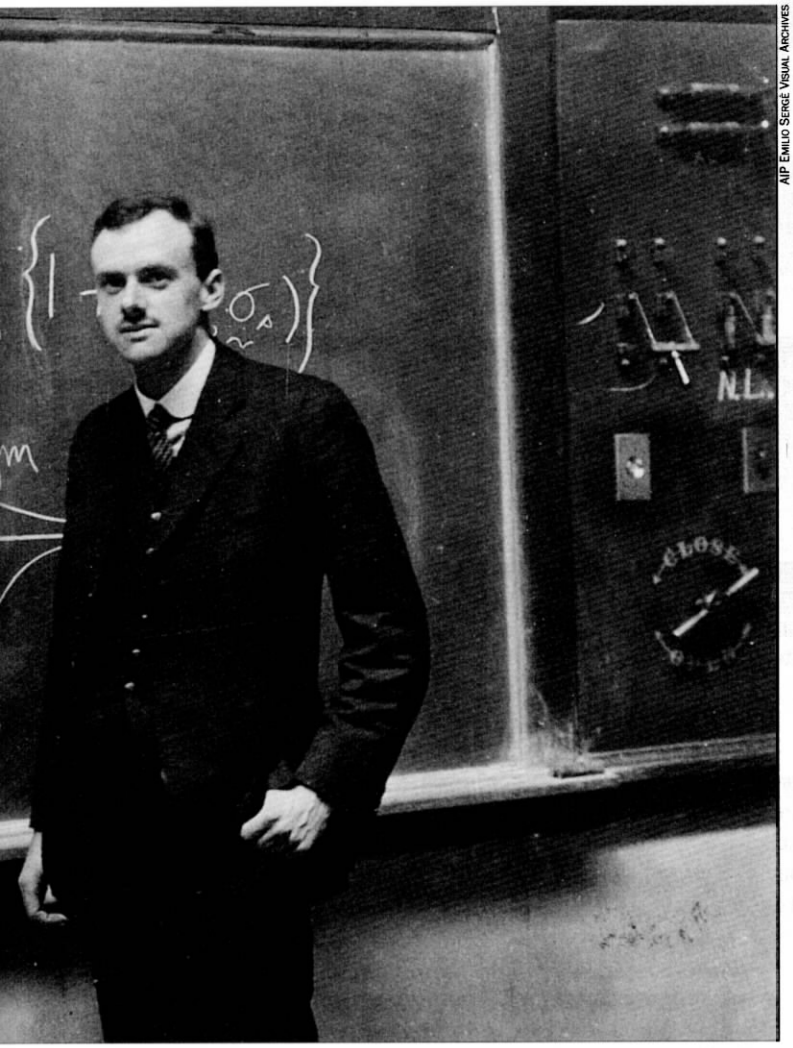
It is hard to give the flavour of Dirac's achievements in a non-technical article, because his work was so mathematical. He once said: "A great deal of my work is just playing with equations and seeing what they give."

Early days

When Dirac went to Cambridge in 1923, the physics of matter on the smallest scales – in those days this was the physics of the atom – was in ferment. It had been known for more than a decade that the old mechanics of Newton – "classical" mechanics, as it came to be called – does not apply in the microscopic world. In particular, evidence from the light coming out of atoms seemed to indicate that some quantities that in classical mechanics can take any values are actually restricted to a set of particular values: they are "quantized". One of these quantities is the energy of the electrons in an atom. This was strange and shocking. Imagine being told that when your car accelerates from 0 to 70 miles per hour it does so in a series of jumps from one speed to another (say in steps of one thousandth of a mph), with the intermediate speeds simply not existing. It did not make sense, and yet observations seemed to demand such an interpretation.

In the first attempts at a theoretical understanding, physicists tried to find the general rules for imposing these restrictions on classical mechanics – that is rules for quantization. It seemed that in order to quantize, it was necessary first to identify those quantities that do not change when their environment is slowly altered. If a pendulum is slowly shortened, for example, it swings farther and also faster, in such a way that its energy divided by its frequency stays constant. These rules worked for simple atoms and molecules but failed for complicated ones.

Dirac entered physics at the end of this baroque period. One of his first papers was an attempt at a general theory of



AIP Emilio Segrè Visual Archives

these unchanging quantities. This is a delicate problem in classical mechanics, not solved even now. It is amazing today to read that paper. In its mathematics it is quite unlike any of Dirac's later works (for example, he brings in fine differences between rational and irrational numbers), and "pre-invents" techniques developed by other people only decades later. (I say pre-invents because the paper was forgotten until recently.)

At this time the situation in atomic physics resembled that at the end of the 16th century, when the old Earth-centred astronomy had to be made ever more elaborate in the face of more accurate observations. The difficulties of the 16th and 20th centuries were resolved in the same way: by a complete shift of thought. In atomic physics this happened suddenly, in 1925, with the discovery by Heisenberg of quantum mechanics. This seemed to throw out classical mechanics completely, though it was built in as a limiting case to ensure that, on larger scales, the new mechanics agreed with more familiar experience. The quantum rules emerged automatically, but from a mathematical framework that was peculiar. For example, it involved multiplication where the result depends on the order in which the multiplication is done. It is as though 2 multiplied by 3 is different from 3 multiplied by 2. Heisenberg found this ugly and unsatisfactory. Dirac disagreed, and just a few months after Heisenberg he published the first of a series of papers in which quantum mechanics took the definitive form we still use today.

The main idea is that the multiplied objects – objects that represent variables we can measure in experiments – should be thought of as operations. An experiment is an operation, of course, even though its result is a number. With this interpretation, it is not surprising that the order matters: we all know that putting on our socks and then our shoes gives a result different from putting on our shoes and then our socks. Dirac found the one simple rule by which a multiplied by b differed from b multiplied by a, and from which the whole of quantum mechanics follows.

The same unification was soon found to include Schrödinger's way of doing quantum mechanics, where the state of a system is represented by a wave whose strength gives the probabilities of the different possible results of measurements on it. For a while this seemed completely different from the framework that Heisenberg had used, but it quickly emerged that in fact each represents Dirac's operators in a different way. It seemed miraculous.

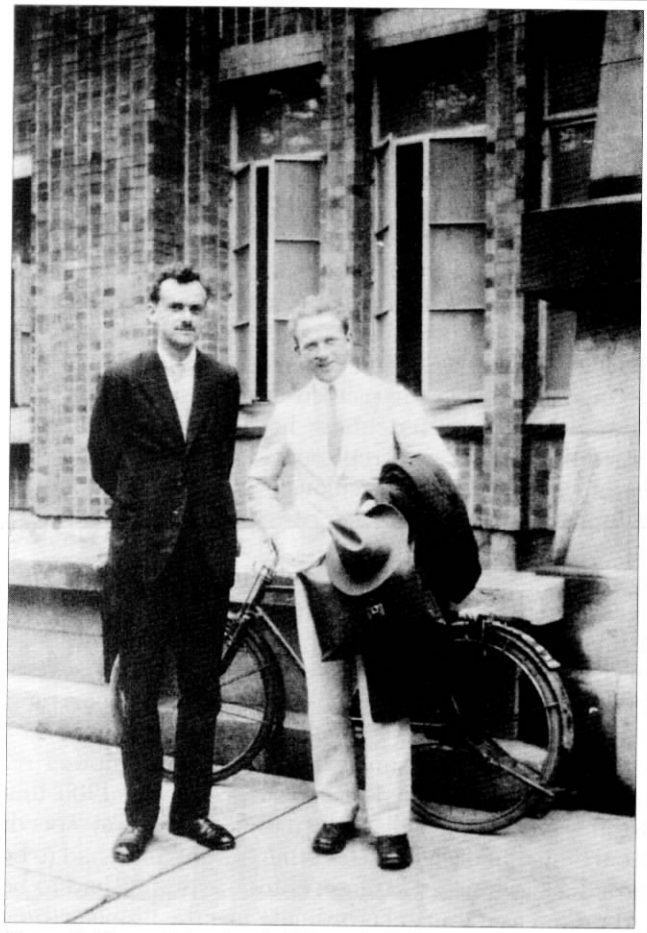
The Dirac equation

Although brilliant – in Einstein's words, "the most logically perfect presentation of quantum mechanics" – this was a reformulation of physics that had, admittedly only just, been discovered. Dirac's main contribution came several years later, when (still in his mid-twenties) he made his most spectacular discovery.

Before quantum mechanics, there had been another revolution in physics, with Einstein's discovery in 1905 that Newton's mechanics fails for matter moving at speeds approaching that of light. To get things right, time had to be regarded as no longer absolute: before-and-after had to be incorporated as a fourth co-ordinate like the familiar three spatial co-ordinates that describe side-to-side, forward-and-backward and up-and-down. Just as what is side-to-side and what is forward-and-backward change when you turn, so time gets mixed in with the other three co-ordinates when you move fast. Now, in the 1920s, came quantum mechanics, showing how Newton's mechanics failed in a different way: on microscopic scales. The question arose: what is the physics of particles that are at the same time small *and* moving fast?

This was a practical question: the electrons in atoms are small, and they move fast enough for the new quantum mechanics to be slightly inaccurate, since it had been constructed to have as its large-scale limit Newton's mechanics rather than Einstein's. From the start people tried to construct a quantum theory concordant with relativity, but failed to overcome technical obstructions: in particular, their attempts gave probabilities that were negative numbers – something that is nonsense, at least in the usual meaning of probability. The question boiled down to this: what are the right sort of quantum waves describing electrons? And what is the wave equation that governs the dynamics of these waves, while satisfying the requirements of relativity and giving sensible physical predictions?

Dirac's construction of his wave equation for the electron – published in two papers in the *Proceedings of the Royal Society (London)* in February and March 1928 – contained one of those outrageous leaps of imagination shared by all great advances in thought. He showed that the simplest wave satisfying the requirements was not a simple number but had four components (see box overleaf). This seemed like a complication, especially to minds still reeling from the unfamiliarity of the "ordinary" quantum mechanics. Four components! Why should anybody take Dirac's theory seriously?



Dirac with Werner Heisenberg in Chicago in 1929.

The Dirac equation

The Dirac equation for an electron moving in an arbitrary electromagnetic field can be written in many ways. In Dirac's original papers it is written as

$$\left[p_0+\frac{e}{c}A_0+\alpha_1\left(p_1+\frac{e}{c}A_1\right)+\alpha_2\left(p_2+\frac{e}{c}A_2\right)+\alpha_3\left(p_3+\frac{e}{c}A_3\right)+\alpha_4 mc\right]\psi=0$$

where $p_0 = ih\partial/c\partial t$ (the energy operator), e is the charge on the electron, A_0 is the scalar potential associated with the electromagnetic field, c is the speed of light, α_i are 4×4 matrices derived from the Pauli matrices, $p_1 = -ih\partial/\partial x$ is a momentum operator ($p_2 = -ih\partial/\partial y$, $p_3 = -ih\partial/\partial z$), A_i are the three components of the electromagnetic vector potential, m is the mass of the electron and ψ is the wavefunction of the electron.

The wavefunction ψ is a 4×1 column vector (also known as a spinor) and each element is a function of space and time, representing the spin state (up or down) of the electron and the associated positron solution. As explained in the main text, the equation was able to explain the results of all of the experiments at the time, to explain the origin of electron spin and to predict the existence of antimatter.

The equation can be written in more compact form. In §67 of *The Principles of Quantum Mechanics* (4th edn, Oxford University Press) it is written as

$$\left\{p_0+\frac{e}{c}A_0-\rho_1\left(\sigma,\mathbf{p}+\frac{e}{c}\mathbf{A}\right)-\rho_3 mc\right\}\psi=0$$

where ρ_1 and ρ_3 are 4×4 matrices (related to α_i and the Pauli matrices), σ is a three-component vector of 4×4 matrices, and $\mathbf{p}$ is a three-component vector of momentum operators. The version of the equation in Westminster Abbey is even more compact and reads $i\gamma\cdot\partial\psi = m\psi$ where γ is a 4×4 matrix and ∂ is a 4-vector.

First, and above all for Dirac, the logic that led to the theory was, although deeply sophisticated, in a sense beautifully simple. Much later, when someone asked him (as many must have done before) "How did you find the Dirac equation?" he is said to have replied: "I found it beautiful." Second, it agreed with precise measurements of the energies of light emitted from atoms, in particularly where these differed from ordinary (non-relativistic) quantum mechanics.

There are two more reasons why the Dirac equation was compelling as the correct description of electrons. To understand them, you should realize that any great physical theory gives back more than is put into it, in the sense that as well as solving the problem that inspired its construction, it explains more and predicts new things. Before the Dirac equation, it was known that the electron spins. The spin is tiny on the scale of everyday but is always the same and plays a central part in the explanation through quantum mechanics of the rules of chemistry and the structure of matter. This spin was a property of the electron, like its mass and its electric charge, whose existence simply had to be assumed before quantum mechanics could be applied. In Dirac's equation, spin did not have to be imported: it emerged – along with the magnetism of the electron – as an inevitable property of an electron that was both a quantum particle and a relativistic one.

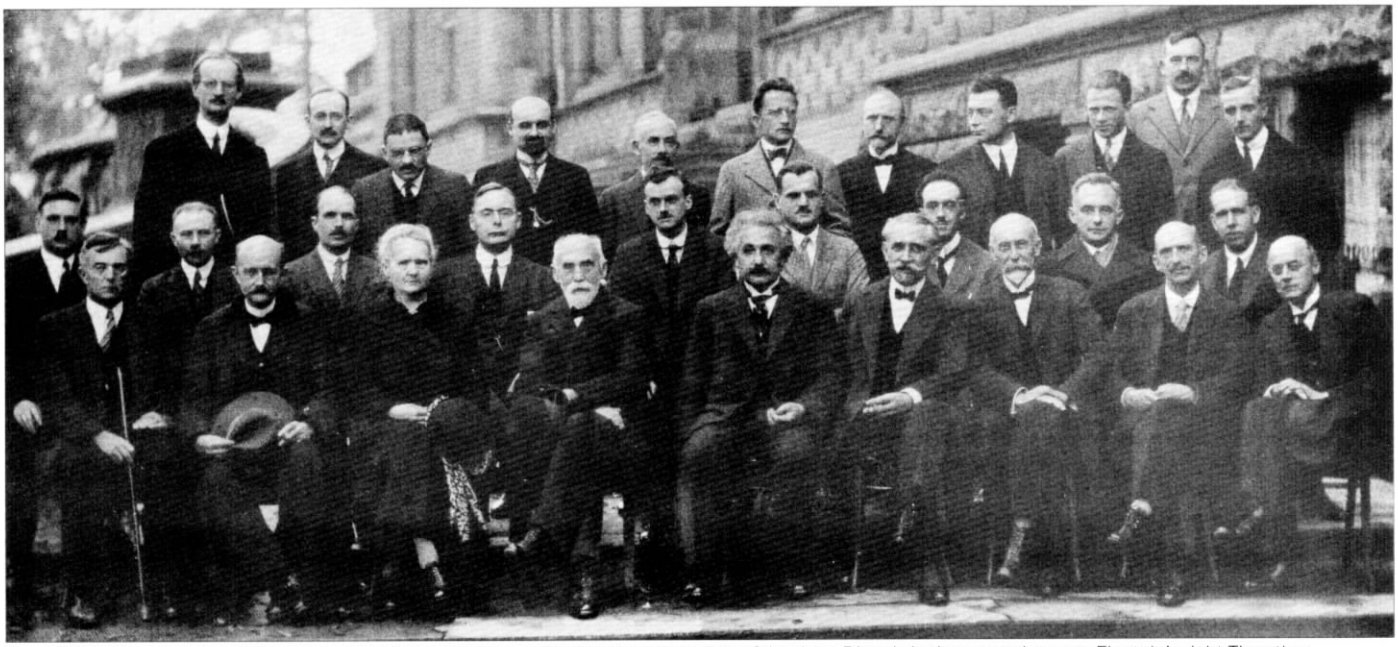
So, electron spin was the third reason for believing Dirac's mathematically inspired equation. The fourth came from a consequence of the equation that was puzzling for a few years at first. Related to its four components was the fact that any solution of the equation where the electron had a positive energy had a counterpart where the energy was negative. It gradually became clear that these counterpart solutions could be interpreted as representing a new particle, similar to the electron but with positive rather than negative charge; Dirac called it an "anti-electron", but it soon came to be known as the positron. If an electron encounters a positron, Dirac predicted, the two charges cancel and the pair annihilates, with the combined mass transforming into radiation in the most dramatic expression of Einstein's celebrated equation $E=mc^2$. Thus was antimatter predicted. When the positron was discovered by Anderson in 1932, Dirac's immortality was assured. Dirac and Schrödinger shared the Nobel Prize for Physics in 1933.

Nowadays, positrons are used every day in medicine, in PET (positron emission tomography) scanners that pinpoint interesting places in the brain (e.g. places where drugs are chemically active). These work by detecting the radiation as the positrons emitted from radioactive nuclei annihilate with ordinary electrons nearby.

Other achievements

Having explained spin, it was natural for Dirac to try to explain electric charge, and in particular the mysterious fact that it is quantized: all charges found in nature are multiples of the charge on the electron. In classical electricity, there is no basis for this: charges can have any value.

In 1931 Dirac gave a solution of this problem in an application of quantum mechanics so original that it still astounds us to read it today. He combined electricity with magnetism, in a return to the 18th-century notion of a magnet being a combination of north and south magnetic poles (magnetic



The 1927 Solvay Congress in Brussels was attended by most of the leading physicists of the time. Dirac is in the second row, on Einstein's right.The other delegates are (left to right): front row; I Langmuir, M Planck, Madame Curie, H A Lorentz, A Einstein, P Langevin, Ch E Guye, C T R Wilson, O W Richardson; second row; P Debye, M Knudsen, W L Bragg, H A Kramers, P A M Dirac, A H Compton, L V de Broglie, M Born, N Bohr; back row; A Piccard, E Henriot, P Ehrenfest, E D Herzen, T H de Donder, E Schrödinger, E Verschaffelt, W Pauli, W Heisenberg, R H Fowler, L Brillouin.

charges), in the same way that a charged body contains positive and negative electric charges. That symmetry was lost in the 19th century with the discoveries of Oersted, Ampère and Faraday, culminating in Maxwell's synthesis of all electromagnetic and – in another example of getting out more than you put in – optical phenomena. In its place came a greater simplicity: there are only electric charges, whose movement generates magnetism (and now the motive power for much of our civilisation). The absence of isolated magnetic poles – magnetic monopoles – was built into classical electromagnetism, and also the quantum mechanics that grew out of it.

Dirac wondered if there was any way that magnetic monopoles could be brought into quantum physics without spoiling everything that had grown out of assuming that they did not exist. He found that this could be done, but only if the strength of the monopole (the "magnetic charge") was linked to that of the electric charge, and if both were quantized. This solved the original problem: for consistency with quantum mechanics, the existence of even one monopole anywhere in the universe would suffice to ensure that electric charge must be quantized. The implication is compelling: to account for the quantization of electricity, magnetic poles must exist. After this, Pauli referred to Dirac as "Monopoleon".

Alas, no magnetic monopole has ever been found. Perhaps they do not exist, or perhaps (and there are hints of this in the theory) positive and negative monopoles are so tightly bound together that they have not been separated. Much later, Dirac referred to this theory as "just a disappointment". However, the mathematics he invented to study the monopole – combining geometry with analysis – now forms the basis of the modern theories of fundamental particles.

There were two other seminal contributions to physics in those early years. I have space only to mention them. Dirac applied quantum mechanics to the way light and matter interact. This made him realize that it was necessary to quantize not only particles but the electromagnetic field itself, and led him to the first consistent theory of photons (which had been discovered several decades previously in the beginnings of quantum mechanics). This led to the elaborate and thriving quantum field theories of today.

Dirac also showed how quantum waves for many electrons had to be constructed, incorporating the philosophically intriguing fact that any two of these particles are absolutely identical and so cannot be distinguished in any way. This produced the definitive understanding of earlier rules about how quantum mechanics explains the periodic table of the elements, and provided the basis for the theory of metals and the interior of stars.

Like all scientists at the highest level, Dirac was not afraid to descend from the pinnacle and discuss more down-to-earth matters. Here are two examples. Much of our knowledge comes from light scattered by matter; in particular, that is how we see. In a clever stroke of lateral thinking, Dirac realized that the quantum symmetry between waves of light and waves of matter implied that it is also possible for material particles to be scattered by light, a ghostly possibility that could be observed, as he showed in 1933 in a paper with Peter Kapitza. This was observed for the first time about ten years ago and the manipulation of atoms by laser beams is now a thriving area of applied quantum mechanics – a fact recognized with a Nobel prize last year (*Physics World* November 1997 p51).

The second example is his Second World War work. In the Manhattan Project to develop the first nuclear bombs, it was necessary to separate isotopes of uranium. One class of methods involved the centrifugal effects of fluid streams that were made to bend. Dirac put the theory of these techniques on a firm basis, and indeed his work in this field has been described as seminal.

Dirac stories

It is not my intention to write about what sort of person Dirac was. But I must mention the genre of "Dirac stories". He was so unusual in the logic and precision of his interaction with

we now know there are no magnetic monopoles because of gauge symmetry

610

The Quantum Theory of the Electron.

By P. A. M. Dirac, St. John's College, Cambridge.

(Communicated by R. H. Fowler, F.R.S.—Received January 2, 1928.)

The new quantum mechanics, when applied to the problem of the structure of the atom with point-charge electrons, does not give results in agreement with experiment. The discrepancies consist of "duplexity" phenomena, the observed number of stationary states for an electron in an atom being twice the number given by the theory. To meet the difficulty, Goudsmit and Uhlenbeck have introduced the idea of an electron with a spin angular momentum of half a quantum and a magnetic moment of one Bohr magneton. This model for the electron has been fitted into the new mechanics by Pauli,* and Darwin,† working with an equivalent theory, has shown that it gives results in agreement with experiment for hydrogen-like spectra to the first order of accuracy.

The question remains as to why Nature should have chosen this particular model for the electron instead of being satisfied with the point-charge. One would like to find some incompleteness in the previous methods of applying quantum mechanics to the point-charge electron such that, when removed, the whole of the duplexity phenomena follow without arbitrary assumptions. In the present paper it is shown that this is the case, the incompleteness of the previous theories lying in their disagreement with relativity, or, alternatetively, with the general transformation theory of quantum mechanics. It appears that the simplest Hamiltonian for a point-charge electron satisfying the requirements of both relativity and the general transformation theory leads to an explanation of all duplexity phenomena without further assumption. All the same there is a great deal of truth in the spinning electron model, at least as a first approximation. The most important failure of the model seems to be that the magnitude of the resultant orbital angular momentum of an electron moving in an orbit in a central field of force is not a constant, as the model leads one to expect.

* Pauli, 'Z. f. Physik,' vol. 43, p. 601 (1927).
† Darwin, 'Roy. Soc. Proc.,' A, vol. 116, p. 227 (1927).

351

The Quantum Theory of the Electron. Part II.

By P. A. M. Dirac, St. John's College, Cambridge.

(Communicated by R. H. Fowler, F.R.S.—Received February 2, 1928.)

In a previous paper by the author* it is shown that the general theory of quantum mechanics together with relativity require the wave equation for an electron moving in an arbitrary electromagnetic field of potentials, A_0, A_1, A_2, A_3 to be of the form

$$F\psi \equiv \left[p_0 + \frac{e}{c}A_0 + \alpha_1\left(p_1 + \frac{e}{c}A_1\right) + \alpha_2\left(p_2 + \frac{e}{c}A_2\right) + \alpha_3\left(p_3 + \frac{e}{c}A_3\right) + \alpha_4 mc\right]\psi = 0. \quad (1)$$

The α's are new dynamical variables which it is necessary to introduce in order to satisfy the conditions of the problem. They may be regarded as describing some internal motion of the electron, which for most purposes may be taken to be the spin of the electron postulated in previous theories. We shall call them the spin variables.

The α's must satisfy the conditions

$$\alpha_\mu^2 = 1, \qquad \alpha_\mu\alpha_\nu + \alpha_\nu\alpha_\mu = 0. \qquad (\mu \neq \nu.)$$

They may conveniently be expressed in terms of six variables ρ_1, ρ_2, ρ_3, σ_1, σ_2, σ_3 that satisfy

$$\left.\begin{aligned} &\rho_r^2 = 1, \quad \sigma_r^2 = 1, \quad \rho_r\sigma_s = \sigma_s\rho_r, \quad (r, s = 1, 2, 3) \\ \text{and} \quad &\rho_1\rho_2 = i\rho_3 = -\rho_2\rho_1, \quad \sigma_1\sigma_2 = i\sigma_3 = -\sigma_2\sigma_1 \end{aligned}\right\}, \quad (2)$$

together with the relations obtained from these by cyclic permutation of the suffixes, by means of the equations

$$\alpha_1 = \rho_1\sigma_1, \qquad \alpha_2 = \rho_1\sigma_2, \qquad \alpha_3 = \rho_1\sigma_3, \qquad \alpha_4 = \rho_3.$$

The variables σ_1, σ_2, σ_3 now form the three components of a vector, which corresponds (apart from a constant factor) to the spin angular momentum vector that appears in Pauli's theory of the spinning electron. The ρ's and σ's vary with the time, like other dynamical variables. Their equations of motion, written in the Poisson Bracket notation [ ], are

$$\dot{\rho}_r = c\,[\rho_r, F], \qquad \dot{\sigma}_r = c\,[\sigma_r, F].$$

* 'Roy. Soc. Proc.,' A, vol. 117, p. 610 (1928). This is referred to later by *loc. cit.*

Dirac's papers on the quantum theory of the electron were published in the *Proceedings of the Royal Society (London) A* in 1928 (see further reading).

the world, both in and out of physics, that tales have become attached to him and have acquired a life of their own. I suppose it matters to a historian whether they are true or apocryphal (or as Norman Mailer says, "factoids"), but to us they have a deeper resonance that transcends fact. Resisting temptation, I retell just two less well known ones.

Like many scientists, Dirac was known to sleep during (other people's) lectures, and then wake and suddenly make a penetrating remark. Once, a speaker stopped, scratched his head and declared: "Here is a minus where there should be a plus. I seem to have made an error of sign." Dirac opened one eye and said: "Or an odd number of them." Another time, Dirac was at a meeting in a castle, when another guest remarked that a certain room was haunted: at midnight, a ghost appeared. In his only reported utterance on matters paranormal, Dirac asked: "Is that midnight Greenwich time, or daylight saving time?"

Dirac's writing was famous for its clarity and simplicity. Every physicist knows his *Principles of Quantum Mechanics* – such a perfect and complete summary of his views that in later years his lectures consisted of readings from it. There is the story that he was once present when Niels Bohr was writing a scientific paper – with many hesitations and redraftings, as was his custom. Bohr stopped: "I do not know how to finish this sentence." Dirac replied: "I was taught at school that you should never start a sentence without knowing the end of it."

Many physicists have spoken of Dirac with awe. John Wheeler, referring to the sharp light of his intelligence, said "Dirac casts no penumbra." Niels Bohr said: "Of all physicists, Dirac has the purest soul." He is also reported as saying (I cannot now find this quotation): "Dirac did not have a trivial bone in his body."

The mathematician Mark Kac divided geniuses into two classes. There are the ordinary geniuses, whose achievements one imagines other people might emulate, with enormous hard work and a bit of luck. Then there are the magicians, whose inventions are so astounding, so counter to all the intuitions of their colleagues, that it is hard to see how any human could have imagined them. Dirac was a magician.

Further reading

P A M Dirac 1928 The quantum theory of the electron *Proc. R. Soc. (London)* **117** 610–612

P A M Dirac 1928 The quantum theory of the electron. Part II *Proc. R. Soc. (London)* **118** 351–361

R H Dalitz (ed) 1995 *The Collected Works of P A M Dirac 1924–1948* (Cambridge University Press)

Sir Michael Berry is Royal Society Research Professor at the H H Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK. This is a version of a talk delivered in September 1997 at the official opening of Dirac House, the headquarters of Institute of Physics Publishing