Chapter 9. States of Position and Momentum

1. Introduction

Our use of the probability amplitude has thus far been limited to descriptions of photon polarization experiments and atomic beam experiments of the Stern-Gerlach type. The quantum states that we have studied provide no description of position or momentum measurements. Yet a complete specification of the state of a particle should include information on its motion. In this Chapter we begin the discussion of experiments in which position and momentum play an essential role.*

Quantum concepts are not always necessary for the interpretation of position and momentum measurements. In many situations the motion of particles, even atomic particles, is adequately described by the laws of classical mechanics and electromagnetism. For example, the operation of a cathode ray tube or of a particle accelerator is satisfactorily explained by calculating the classical trajectories of the particles in the electric and magnetic fields that are present. A successful design for either of these devices can be drawn without taking any account of the laws of quantum physics. Even the operation of the devices used below to define a momentum quantum state will be explained in terms of classical physics. But many experiments involving atomic particle do require quantum concepts. No classical analysis can account for the interference and diffraction of particles (Section 3), a phenomenon that readily lends itself to description in terms of quantum amplitudes. Moreover, the internal states of atoms are described by quantum amplitudes (Chapters 12 to 14). From the present analysis of position and momentum amplitudes (Chapters 9 to 11) will come not only an introduction to the amplitudes that govern atomic states but also an appreciation for the limits in which a classical description of motion suffices.

2. Momentum states

Using the notation for quantum states introduced in earlier chapters, we let the state vector $|\vec{p}>$ denote the state of a particle whose momentum is $\vec{p}$. A device that plays the role of a projector for this state is sketched in Figure 1. The direction of $\vec{p}$ is established by slits S_1 and S_2: a particle will not pass through S_2 unless it is incident in the direction determined by the line joining S_1 and S_2. Any charged particle passing through S_2 is deflected by the transverse magnetic field into a circular path, but only those particles whose momentum has

---

* Actually the momentum is not an irrelevant factor in a Stern-Gerlach experiment. The momentum fixes the amount of time a particle spends in the magnetic field and thus determines the magnitude of its deflection. Indeed, we saw in Chapter 7 that it is the spread in momentum that causes great practical difficulty in separating into distinct beams atoms with three or more spin states. But the momentum does not enter the description of the spin states themselves.

magnitude near

$$p = \frac{qBD}{2c} \tag{1}$$

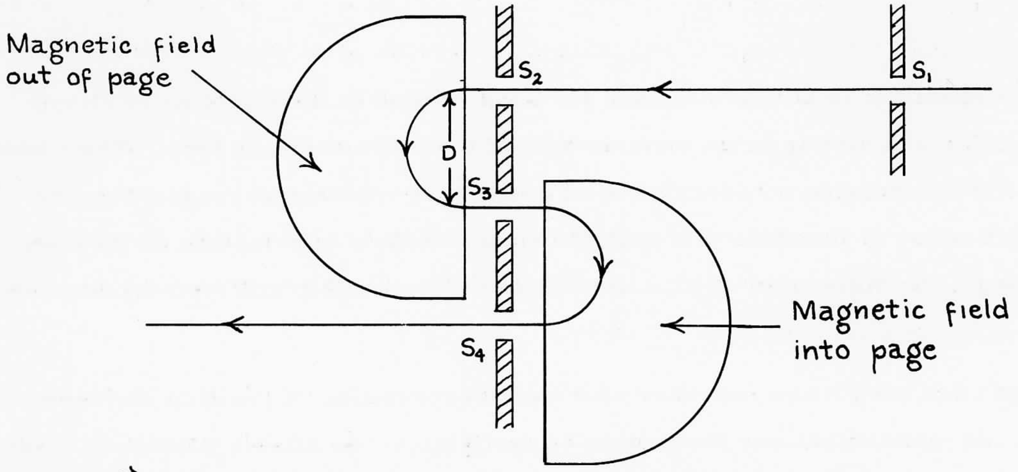


Figure 1. $\vec{p}$-projector for particles of negative charge (Schematic).

pass through S_3. All others run into one of the screens. Thus by adjusting the field strength we can (for a given charge q) select particles of any desired momentum. The purpose of the second magnetic field is merely to restore the original direction of motion.

Since the slits have a finite width, the apparatus of Figure 1 transmits particles with a narrow range of momenta rather than strictly a single value. One might think that by choosing sufficiently narrow slits it should be possible to make the width of the momentum band arbitrarily small. This is not the case because when a particle passes through an extremely narrow slit its direction may be altered appreciably. This is the phenomenon of diffraction, which we discuss later in the present chapter. For an apparatus of given size there is a definite limit to the precision with which a single momentum can be selected by this method. To produce a beam of particles in the ideal state $|\vec{p}\rangle$ would require an apparatus of infinite size. However, even with attainable dimensions it is possible to make the spread of transmitted momenta quite small (see the exercises).

If the particles selected by the projector of Figure 1 enter a second identical device, they are (ideally) all transmitted. This is one of the conditions that must be satisfied by a quantum state. In contrast, if the beam transmitted by the $\vec{p}$-projector passes into a second projector set to select particles of a different momentum $\vec{p}\,' \neq \vec{p}$, then none of the particles get through the second projector.* This is the property of orthogonality, which we write in the usual way:

$$\langle \vec{p}\,' | \vec{p} \rangle = 0 \qquad \vec{p}\,' \neq \vec{p} \tag{2}$$

---

* If $\vec{p}$ is very close to $\vec{p}\,'$, close enough to lie within the band of momenta transmitted by the $\vec{p}\,'$-projector, then of course some of the particles will get through the second projector. This represents an imperfection of the device rather than a failure of Eq. 2, which is correct as an ideal limit.

The value of $\langle \vec{p} | \vec{p} \rangle$ defines the normalization of the state vector, and is discussed below (Section 5). To demonstrate the completeness of the set of states $|\vec{p}\rangle$, we must exhibit an analyzer, an apparatus with a number of exit channels such that for an arbitrary incident beam the sum of the intensities in the output channels equals the intensity of the incident beam. The $\vec{p}$-projector with some screens removed provides a fairly satisfactory analyzer, as shown in Figure 2. Different paths through the analyzer correspond to different values of the momentum, as described by Eq. 1. The apparatus of Fig. 2 falls short of the ideal analyzer in the respect that it analyzes into momentum components only a beam traveling initially in a specific direction. Since the complete set of possible values of $\vec{p}$ includes all directions as well as all magnitudes, we do not have, strictly speaking, a complete analyzer.

One important feature of momentum states distinguishes them from all other quantum states so far discussed: momentum states form a continuous set. The "exit channels" of the analyzer of Figure 2 are infinite in number and are not discrete but rather form a continuum. A free particle can have any momentum whatever. Quantum mechanics demands the "quantization" of some observables (their limitation to certain discrete values). But the momentum of a free particles is not quantized. A complete set of momentum states contains a continuous infinity of members, rather than a finite number as do all sets of states previously discussed. This is a conceptual complication, although it sometimes leads to mathematical simplification. For example, all sums over complete sets of states are replaced by integrals, and integrals are often more easily evaluated than are sums of large numbers of terms.

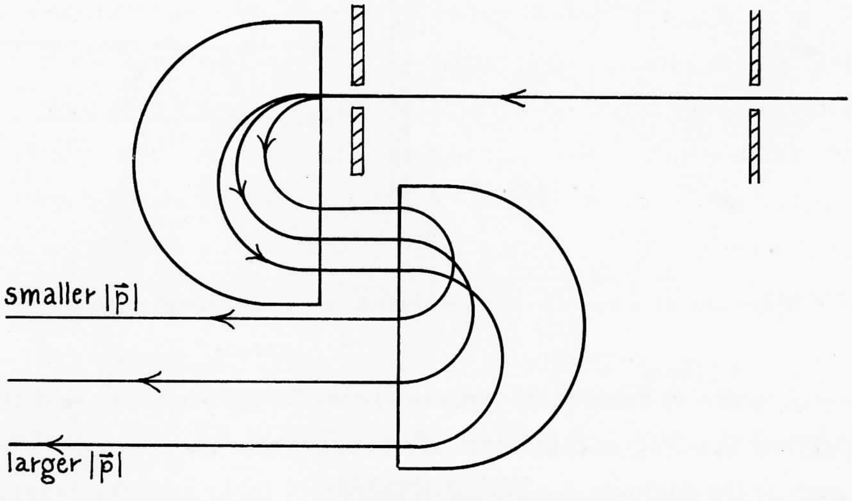


Figure 2. An analyzer for momentum states of charged particles.

The projector and analyzer we have exhibited are effective only on charged particles. Momentum states $|\vec{p}\rangle$ for neutral particles are defined with other devices. For example, one determines the momentum of neutrons typically by using a pulsed source and measuring the time of flight to a detector. A velocity and a known particle mass together determine the momentum.

3. Interference experiments with particles

We now consider experiments that demonstrate the non-classical behavior of material particles. The analysis of the experiments leads us to introduce a new basic amplitude.

As before, the experiments that lead to the introduction of amplitudes involve the phenomenon of interference. The fact that beams of particles such as electrons and neutrons exhibit interference effects has already been mentioned in Chapter 4. Figure 8 of that chapter shows an electron interference pattern obtained using two slits. We now analyze the electron interference experiment in some detail. For convenience of exposition we let the electrons pass through small circular holes rather than rectangular slits, although slits are easier to use and are employed in actual experimental demonstration.

The basic experiment is diagrammed in Figure 3. A $\vec{p}$-projector is located a long distance from the screen, which contains two small circular holes at positions $\vec{r}_1$ and $\vec{r}_2$. The holes are close enough together (and far enough from the source) that they may both be considered to be in the direct path of the incident beam. The detector D is also far from the screen, as determined by a similar criterion: straight lines connecting each of the two holes with point D have effectively the same direction. In other words, all particles that arrive at D are moving in nearly the same direction. In order to minimize any possible effect of interaction between different electrons in the beam, the experiment is performed with a very weak beam. In principle, one could use a beam so weak that on the average only one electron is present in

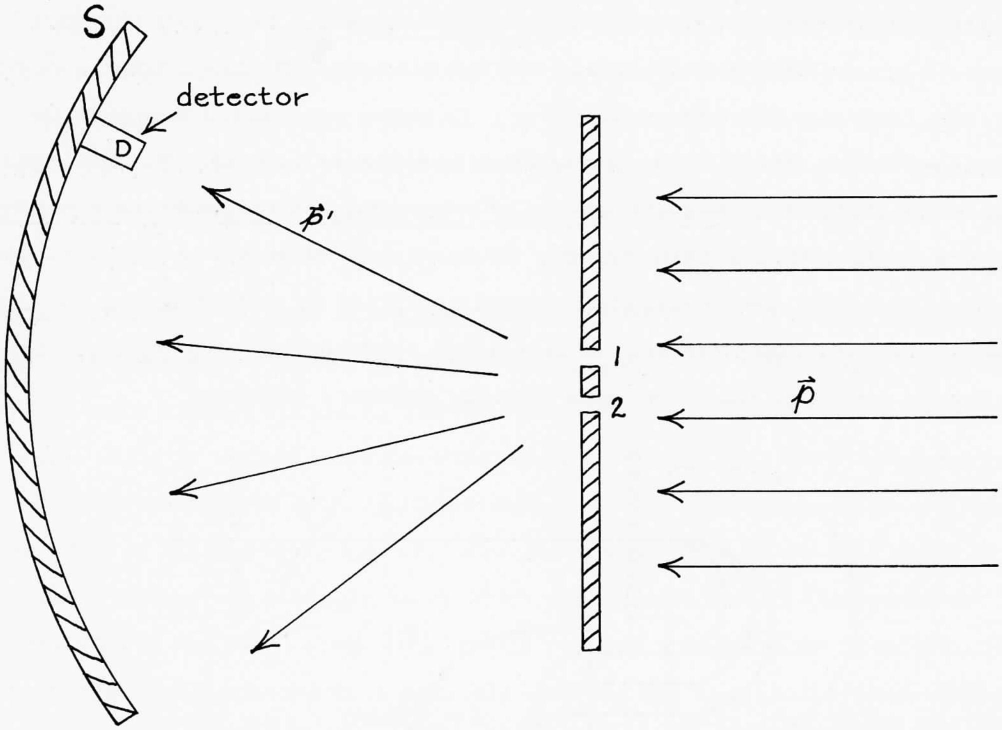


Figure 3. Schematic diagram of a two-hole interference experiment with particles.

the apparatus at a given time (this is the same kind of precaution we took with atomic beams). The observed effects, in this limit, are properties of single electrons.

The experiment consists of measuring the relative counting rate of the detector as this detector is moved to different positions on the sphere S centered at the position of the holes. This is done first (case a) with only one hole open and second (case b) with both holes open. In each case we can verify with a momentum analyzer that $|\vec{p}\,'|$, the magnitude of the momentum of the detected particles, is the same as $|\vec{p}|$, the magnitude of the momentum of the incident particles. Thus the quantity of interest is the dependence of the counting rate on the direction of $\vec{p}\,'$.

As long as the holes are fairly large (say 1 mm in diameter--a more exact criterion is discussed later), the electrons behave as one would expect them to according to Newtonian mechanics: they travel straight ahead. With either one hole or both holes open, the counting rate is zero except in the forward direction $(\vec{p}\,' = \vec{p})$. However, when the experiments are repeated with very small holes, one begins to detect electrons in directions other than forward. The patterns of intensity are very similar to those observed with light passes through the same holes. Just as in the case of light, the pattern spreads as the size of the holes is made smaller. The pattern observed with only one hole open (case a) is called a diffraction pattern. A sketch of intensity as a function of angle for a circular hole is shown in Figure 4. Figure 5 is a picture of an electron diffraction pattern for a single slit. (No picture of a single hole pattern for electrons is available.)

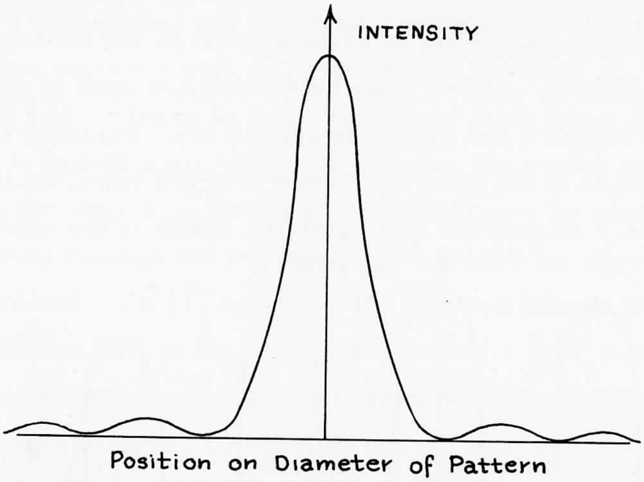


Figure 4. Intensity of diffraction pattern of a single circular hole measured along a diameter across the diffraction pattern.

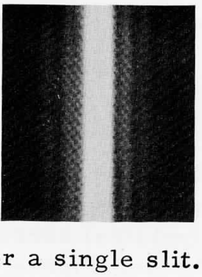


Figure 5. Electron diffraction pattern for a single slit. Photograph taken by Jönssen (Zeitschrift für Physik, 161, 454, (1961)).

The most interesting (and, from a classical viewpoint, unexpected) feature of Figures 4 and 5 is the presence of zeros in the intensity pattern. The fact that some electrons are deflected might be understandable classically as an "edge effect." Passing electrons might induce currents in the material close to the edge of the hole or at least electrically polarize the material, resulting in forces that deflect electrons. A simple spreading of the electron beam might be understandable on this basis. However, the appearance of zeros in the intensity pattern cannot be explained as an edge effect. We know in fact that corresponding diffraction patterns for light are explained in terms of interference between light waves arriving at the screen from different parts of the hole. We shall see later (exercises) that a similar explanation in terms of interference correctly predicts the diffraction pattern of Figure 5. This experiment could be used to introduce the particle amplitudes. However, it is simpler to obtain these amplitudes from an analysis of the two-hole intensity pattern, case b of the experiment under discussion. For this purpose we require only one qualitative feature of the single-hole pattern: for a small enough hole the central maximum encompasses a considerable angular interval. The counting rate is then approximately the same as long as the outgoing momentum $\vec{p}'$ is within a sizeable angular interval enclosing the direction of the incoming momentum $\vec{p}$. For the present we restrict our observations to this region (although there are no conceptual problems in treating the general case). In the limit of an infinitesimally small hole, all directions are included within the central maximum; in this limit the results that we derive are applicable to the entire pattern.

The result of the two-hole experiment is again similar to the analogous experiment with light. In particular, the intensity pattern when both holes are open is <u>not</u> the sum of the individual intensity patterns of Figure 5 for each hole separately. Figure 6 shows the variation of intensity as a function of angle in the plane of Figure 3. (The behavior in other directions is similar--see the exercises.) Again the most striking result is the appearance of zeros of in-

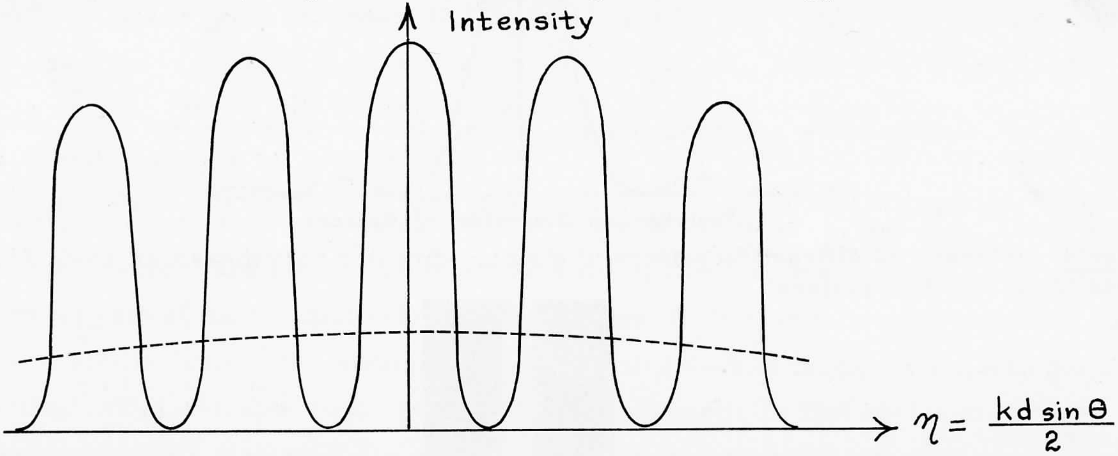


Figure 6. Intensity pattern due to two holes as measured along sphere S in the plane of Figure 3. The variables k, d, and θ that determine the parameter η are defined in Figure 7 and Eq. 10. The envelope of the pattern (dashed line) corresponds to a portion of the central maximum of the diffraction pattern of each hole (see Figure 4). This diffraction effect modulates the interference pattern whose intensity is proportional to $\cos^2\eta$ (Eq. 9).

tensity at points where the intensity is not zero when either hole by itself is open. Such zeros cannot result from the addition of two positive intensities. This is the same dilemma we encountered earlier in connection with photon and atomic beam interference experiments. Our successful analysis of those experiments encourages us to use probability amplitudes to describe the present experiment as well. Accordingly we assume that the probability of detecting an electron, with the detector in a given position and only one hole open, is proportional to the absolute square of some complex amplitude; and when both holes are open and the detector is in the same position, the probability is proportional to the absolute square of the sum of the two single-hole amplitudes.

To find an acceptable expression for the amplitudes thus defined we argue as follows. Consider first the single-hole experiment. The electron motion consists of two sequential steps in the sense defined in Chapter 4 (Figure 10 of that chapter). First the electron with momentum $\vec{p}$ arrives at the hole (position $\vec{r}_1$); then the electron propagates from $\vec{r}_1$ to the detector which measures the momentum $\vec{p}'$. Therefore the amplitude that describes the entire experiment must be the product of the amplitudes for the two sequential steps:

$$\left(\begin{array}{l}\text{amplitude for counting}\\ \text{electron with only hole}\\ \text{number 1 open}\end{array}\right) = \left(\begin{array}{l}\underline{\text{Step 2:}}\text{ amplitude}\\ \text{for electron at } \vec{r}_1\\ \text{to propagate and}\\ \text{be detected with}\\ \text{momentum } \vec{p}'\end{array}\right)\left(\begin{array}{l}\underline{\text{Step 1:}}\text{ amplitude}\\ \text{for electron of}\\ \text{momentum } \vec{p} \text{ to}\\ \text{arrive at } \vec{r}_1\end{array}\right) \tag{3}$$

The amplitude for Step 1 determines the result of a simpler experiment in which the initial beam of momentum $\vec{p}$ impinges directly on a detector at position $\vec{r}_1$. We denote this amplitude by the bracket $\langle \vec{r}_1 | \vec{p} \rangle$. Later in this chapter we shall argue that, as the notation suggests, this quantity is indeed a projection amplitude, the scalar product of two state vectors. Similarly the amplitude for Step 2 on the right side of (3) may be associated with an experiment in which a particle passing through the position $\vec{r}_1$ is tested for momentum $\vec{p}'$. We denote this amplitude by the bracket $\langle \vec{p}' | \vec{r}_1 \rangle$. Then the total amplitude for counting electrons with only hole number 1 open (left side of Eq. 3) is the product $\langle \vec{p}' | \vec{r}_1 \rangle \langle \vec{r}_1 | \vec{p} \rangle$. A similar product of amplitudes describes the experiment with only hole number 2 open. When both holes are open, the resultant amplitude is, by hypothesis, the sum of the amplitudes for each hole separately. The corresponding probability is the square of the magnitude of this resultant amplitude:

$$\left(\begin{array}{l}\text{probability for counting}\\ \text{electron with both holes}\\ \text{open}\end{array}\right) \propto \left| \langle \vec{p}' | \vec{r}_1 \rangle \langle \vec{r}_1 | \vec{p} \rangle + \langle \vec{p}' | \vec{r}_2 \rangle \langle \vec{r}_2 | \vec{p} \rangle \right|^2 \tag{4}$$

4. Particle amplitudes

Our task is to find an expression for the amplitudes appearing in Eq. 4 that will account for the observed interference pattern. In doing this, we are guided by the analogy with Young's experiment in optics and the successful explanation of the latter based on the wave theory of light. The explanation of Young's experiment depends on the difference in optical path along the two possible paths from source to detector. When this difference is an integral multiple of the wavelength of the light, the waves that reach the detector from the two holes interfere constructively, and a maximum of the interference pattern results. If the optical paths from the source to the two holes are equal, the condition for an intensity maximum is (See Figure 7):

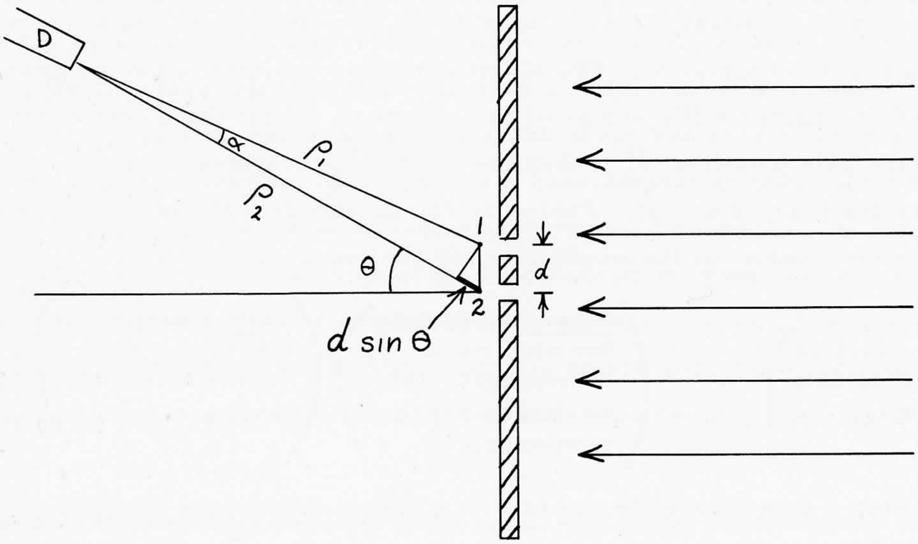


Figure 7. Construction for the analysis of a two-hole interference experiment.

$$\rho_1 - \rho_2 = n\lambda \qquad n = \text{integer} \tag{5}$$

Under the conditions of the problem (detector very far away), the angle α in Figure 7 is very small and Eq. 5 leads to the well-known result for the directions of the intensity maxima:

$$n\lambda = d \sin\theta \tag{6}$$

where d is the separation between the holes.

The successful explanation of Young's interference experiment was instrumental in establishing the acceptance of the wave theory of light in the nineteenth century. In seeking an analogous explanation of electron interference, we lack at this point one essential ingredient: a "wavelength" for the electrons. This missing ingredient is supplied by the famous deBroglie hypothesis. Louis deBroglie in 1924 suggested that a wavelength should be associated with any material particle. His proposal was all the more remarkable because it was made before any of the interference experiments with particles had been performed (and, naturally, before any amplitudes had been discussed). DeBroglie's relation can be obtained by observing that the

Einstein relation for photons (Eq. 1 of Chapter 2) can be written in the form

$$E = pc = h\nu = hc/\lambda \tag{7}$$

and therefore the wavelength associated with a photon is related to its momentum by the equation

$$\lambda = h/p \tag{8}$$

DeBroglie boldly postulated that relation (8) defines a wavelength that can be associated in some way with any particle. This turned out to be an inspired guess. (Actually, deBroglie was led to Eq. 8 by an argument more complicated than the one presented above.) In the present problem Eq. 6, with λ determined by the deBroglie relation (8), gives precisely the directions of the observed maxima in the electron interference pattern.

The deBroglie relation $\lambda = h/p$ has been "verified" in the sense that it leads to predictions confirmed by experiment.* One may wonder what it means to "associate" a wavelength with a particle. This association cannot be with the oscillation of a field in some classical limit, as is the case with electromagnetic radiation; visible particles do not behave in a way describable at all in terms of wavelengths. Yet the interpretation of wavelength is straightforward when carried out in terms of probability amplitudes. For free particles the "deBroglie wavelength" is associated with the <u>phases</u> of these amplitudes.

In order to obtain the desired expressions for the amplitudes $\langle\vec{r}|\vec{p}\rangle$ and $\langle\vec{p}|\vec{r}\rangle$, we require not only the directions of the maxima in the interference pattern (Eqs. 6 and 8) but also the actual dependence on angle of the intensity pattern sketched in Figure 6. The results of experiments with particles conform to those of corresponding experiments with light. The intensity as a function of angle θ in the plane of Figure 3 is

$$\text{intensity} \propto \cos^2\left(\frac{k\,d\sin\theta}{2}\right) \tag{9}$$

where

$$k = 2\pi/\lambda = 2\pi p/h = p/\hbar \tag{10}$$

is called the <u>wave number.</u> (Recall that $k = 2\pi/\lambda$ is also the definition of the wave number in classical optics.) The corresponding <u>vector</u> $\vec{k}$ is called the <u>wave vector.</u> The wave vector points in the direction of the momentum $\vec{p}$ and its magnitude is given by (10). The use of $\vec{k}$ instead of $\vec{p} = \hbar\vec{k}$ to represent particle momentum simplifies many expressions like (9) by incorporating Planck's constant.

Now equate the two expressions for intensity, the theoretical expression (4) and the experimental observation (9), using $\vec{k}$ instead of $\vec{p}$ to label momentum states

$$\left|\langle\vec{k}'|\vec{r}_1\rangle\langle\vec{r}_1|\vec{k}\rangle + \langle\vec{k}'|\vec{r}_2\rangle\langle\vec{r}_2|\vec{k}\rangle\right|^2 \propto \cos^2\left(\frac{k\,d\sin\theta}{2}\right) \tag{11}$$

---

* Historically the first confirmation came in the Davisson-Germer experiment, in which electrons are diffracted by a crystal lattice ratherthan by slits or holes. A film of the Davisson-Germer experiment entitled "Matter Waves" is available from the Education Development Center, Inc., Watertown, Massachusetts.

The absolute square of the product of amplitudes $\langle\vec{k}'|\vec{r}_1\rangle\langle\vec{r}_1|\vec{k}\rangle$ represents the counting rate when only hole 1 is open, while the absolute square of $\langle\vec{k}'|\vec{r}_2\rangle\langle\vec{r}_2|\vec{k}\rangle$ is the counting rate when only hole 2 is open. These two counting rates are equal (the holes are so close together that the detector is at essentially the same angle θ with respect to each hole). Consequently the two products of amplitudes on the left side of Eq. 11 have the same magnitude, and we can write

$$\langle\vec{k}'|\vec{r}\rangle\langle\vec{r}|\vec{k}\rangle = A^2 e^{i\alpha_{k'k}(\vec{r})} \tag{12}$$

Here A^2 represents the common magnitude. Its value depends on the normalization of the states, and is discussed in Section 6. The phase $\alpha_{k'k}(\vec{r})$ is a function of both position $\vec{r}$ and the incoming and outgoing momenta $\vec{k}$ and $\vec{k}'$. When $\vec{r}$ is set equal to $\vec{r}_1$ Eq. 12 gives the phase of $\langle\vec{k}'|\vec{r}_1\rangle\langle\vec{r}_1|\vec{k}\rangle$ and when $\vec{r}$ is set equal to $\vec{r}_2$ it gives the phase of $\langle\vec{k}'|\vec{r}_2\rangle\langle\vec{r}_2|\vec{k}\rangle$. Our task is to find the dependence of this phase on $\vec{k}$, $\vec{k}'$, and $\vec{r}$. A similar procedure was followed in our derivation of the amplitudes for previously analyzed experiments.

Substitute the forms (12) into (11) to obtain, after some manipulation

$$\cos^2\left(\frac{\alpha_{k'k}(\vec{r}_1) - \alpha_{k'k}(\vec{r}_2)}{2}\right) = \cos^2\left(\frac{kd \sin\theta}{2}\right) \tag{13}$$

The argument of the cosine on the right hand side can be expressed as a difference using the following geometrical identity (See Figures 3 and 7)

$$k\, d \sin\theta = (\vec{k}' - \vec{k}) \cdot (\vec{r}_2 - \vec{r}_1) \tag{14}$$

where we make use of the fact that $\vec{k}$ and $\vec{k}'$ have the same magnitude. Therefore

$$\alpha_{k'k}(\vec{r}_2) - \alpha_{k'k}(\vec{r}_1) = \pm(\vec{k}' - \vec{k}) \cdot (\vec{r}_2 - \vec{r}_1) \tag{15}$$

This yields an expression for $\alpha_{k'k}(\vec{r})$.

$$\alpha_{k'k}(\vec{r}) = \pm(\vec{k}' - \vec{k}) \cdot \vec{r} + \delta \tag{16}$$

where δ is an arbitrary constant. To conform with standard convention we choose the negative sign in Eq. 16 and set $\delta = 0$. That is, we write (12) as

$$\langle\vec{k}'|\vec{r}\rangle\langle\vec{r}|\vec{k}\rangle = A^2 e^{-i(\vec{k}' - \vec{k}) \cdot \vec{r}} \tag{17}$$

Equation 17 leads directly to the form of the individual projection amplitudes, since $\langle\vec{r}|\vec{k}\rangle$ must be independent of $\vec{k}'$ and $\langle\vec{k}'|\vec{r}\rangle$ must be independent of $\vec{k}$. The solution that meets this condition is

$$\begin{aligned} \langle\vec{r}|\vec{k}\rangle &= A\, e^{i\vec{k}\cdot\vec{r}} \\ \langle\vec{k}'|\vec{r}\rangle &= A\, e^{-i\vec{k}'\cdot\vec{r}} \end{aligned} \tag{18}$$

If we consider $\vec{k}$ to be an arbitrary wave vector, then the second of these equations can be written

$$\langle\vec{k}|\vec{r}\rangle = A\, e^{-i\vec{k}\cdot\vec{r}} \tag{19}$$

Equations 18 and 19 satisfy the general relation for projection amplitudes for arbitrary states i and j

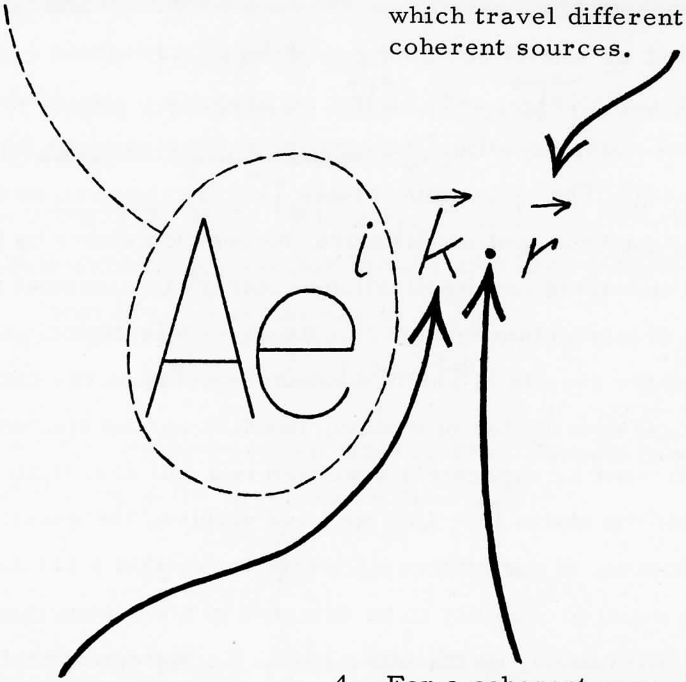
$$\langle i|j\rangle = \langle j|i\rangle^* \tag{20}$$

Figure 8 summarizes the derivation and interpretation of the amplitude $A\,e^{i\vec{k}\cdot\vec{r}}$.

Figure 8. The amplitude $A\,e^{i\vec{k}\cdot\vec{r}}$: where does it come from?

1. Interpretation of interference experiments with two or more beams of particles requires a complex probability amplitude.

2. Interference patterns result from differences in phase between probability amplitudes of two or more combining beams which travel different distances $\underline{r}$ from coherent sources.



3. Dimensions of the interference pattern (distance between intensity maxima) depend upon the energy--and hence upon the momentum--of the incident particles. The wave vector $\vec{k}$ is the momentum of a particle in the beam converted to units of inverse length by the conversion factor $1/\hbar$:

$$\vec{k} = \vec{p}/\hbar$$

4. For a coherent mono-energetic beam of directed momentum $\vec{p}$ (and wave vector $\vec{k} = \vec{p}/\hbar$) the phase at any point in space is the dot product between the momentum and the radius vector $\vec{r}$ from the origin. (Phase arbitrariliy chosen to be zero at origin.)

5. Position states

The amplitudes $\langle\vec{r}|\vec{k}\rangle$ and $\langle\vec{k}|\vec{r}\rangle$ were defined in the course of interpreting an interference experiment with particles. These amplitudes can also be considered as projection amplitudes between momentum states defined in Section 2 and "position states" denoted by the state vector $|\vec{r}\rangle$ and its dual $\langle\vec{r}|$. The state $|\vec{r}\rangle$ is one in which the particle in question is located at the point $\vec{r}$. It follows that the set of states $|\vec{r}\rangle$ is orthogonal: if a particle is known with certainty to be at position $\vec{r}$ it cannot also be at the different position $\vec{r}\,'$.

$$\langle \vec{r}\,' | \vec{r} \rangle = 0 \qquad \vec{r}\,' \neq \vec{r} \tag{21}$$

the set of states $|\vec{r}\rangle$ is also complete: a particle must be somewhere. One cannot readily demonstrate these properties by means of projectors and analyzers in the way we have done with other states. A position analyzer, for example, would have to have output "channels" that take up all of space, each output channel corresponding to a definite position. As our study of quantum states continues, we shall make less and less use of projectors and analyzers because they are increasingly difficult to realize, even in principle, for the states we wish to discuss. But the concepts "representation," "completeness," and "orthogonality" to which these devices have led us are of continuing usefulness throughout quantum physics.

Our interpretation of $\langle \vec{r} | \vec{k} \rangle$ and $\langle \vec{k} | \vec{r} \rangle$ as projection amplitudes entails that position states and momentum states constitute alternative representations for the same state (here the state of a free particle). The momentum states $|\vec{k}\rangle$ by themselves constitute a complete set for the description of particle motion; likewise the position states by themselves constitute a complete set. This statement has implications contrary to classical intuition: it means that if we know the results of momentum measurements on a collection of identically-prepared particles, we can predict the results of position measurements on the same particles, and vice versa. In the classical description of motion, position and momentum are entirely independent quantities; both must be separately specified in a full description of particle motion.

In particular, for the states $|\vec{k}\rangle$ that we have studied, the quantity $|\langle \vec{r} | \vec{k} \rangle|^2$ is independent of $\vec{r}$; this means, if our interpretation is valid, that a particle whose momentum is known to be $\hbar\vec{k}$ has equal probability to be detected in a position measurement anywhere in space. In classical mechanics, on the other hand, the statement that a particle has momentum $\vec{p}$ implies nothing whatever about the results of position measurements.

The discussion of the preceding paragraph is our first example of an "uncertainty relation" in quantum mechanics: if the state of a particle provides a unique value of momentum, it is necessarily accompanied by complete "uncertainty" as to position. Such a state is of course an idealization which can never be realized in practice. In Chapter 10 we shall see how one may combine a limited localization in space with a finite range of momenta to obtain states that provide a more realistic description of particle motion.

6. Properties of position and momentum states

The development of the preceding sections indicates that the set of momentum states $|\vec{k}\rangle$ and the set of position states $|\vec{r}\rangle$ constitute two representations for the description of an arbitrary spatial state, which we may denote by the ket $|\psi\rangle$. Any such state is represented by its projection amplitudes $\langle \vec{r} | \psi \rangle$ into position states, or alternatively by its projection amplitudes $\langle \vec{k} | \psi \rangle$ into momentum states. These two sets of amplitudes are related by a generalization of the transformation matrices treated earlier (Chapters 5 and 6). In this section we discuss this and other general properties of spatial states. Many specific examples will be treated in the chapters that follow.

Because any position or momentum measurement necessarily comprises a band of values rather than a single value, we do not define a probability for detecting a particle precisely at the point $\vec{r}$, or with precisely the momentum $\vec{k}$. Instead one defines a probability per unit volume that in a position measurement a particle in the state $|\psi\rangle$ will be detected in the neighborhood of the point $\vec{r}$. The quantity

$$\left|\langle\vec{r}\,|\psi\rangle\right|^2 d^3r \tag{22}$$

measures the probability of detection within the small volume element d^3r. Similarly the quantity

$$\left|\langle\vec{k}\,|\psi\rangle\right|^2 d^3k \tag{23}$$

measures the probability of finding a momentum value within the "volume element d^3k in momentum space" in the neighborhood of the momentum $\vec{k}$.

Consider now the expansion of the state $|\psi\rangle$ in terms of the position basis states or in terms of the momentum basis states. Recall that for systems with a finite number of basis states, expansion takes the form (Eq. 16 of Chapter 6)

$$|\psi\rangle = \sum_j |j\rangle\langle j|\psi\rangle \tag{24}$$

where the index j runs over the complete set of basis states. The obvious generalization of this expression consists of writing in the r-representation

$$|\psi\rangle = \int |\vec{r}\rangle\, d^3r\, \langle\vec{r}\,|\psi\rangle \tag{25}$$

or, in the k-representation,

$$|\psi\rangle = \int |\vec{k}\rangle\, d^3k\, \langle\vec{k}\,|\psi\rangle \tag{26}$$

The region of integration in (25) and in (26) runs respectively over all of position space or all of momentum space.

Expansions (25) and (26) are correct exactly as written only if a specific normalization is chosen for the basis states $|\vec{r}\rangle$ and $|\vec{k}\rangle$. To demonstrate this, choose $|\psi\rangle$ in Eq. 25 to be the position state $|\vec{r}\,'\rangle$ and take its inner product with the state $|\vec{k}\rangle$.

$$\langle\vec{k}|\vec{r}\,'\rangle = Ae^{-i\vec{k}\cdot\vec{r}\,'} = \int \langle\vec{k}|\vec{r}\rangle\, d^3r\, \langle\vec{r}|\vec{r}\,'\rangle \tag{27}$$

On account of the orthogonality of the position states (Eq. 21), the factor $\langle\vec{r}|\vec{r}\,'\rangle$ in the integrand vanishes for all $\vec{r} \neq \vec{r}\,'$, i.e., everywhere but at a single point. We have not yet discussed the value of $\langle\vec{r}|\vec{r}\rangle$, but it is clear that $\langle\vec{r}|\vec{r}\rangle$ cannot be equal to one as in the normalization with discrete bases. If it were equal to unity, then the integral on the right side of Eq. 27 would vanish and the equality could not be satisfied. In fact, the integral vanishes if

---

* The projection amplitude $\langle\vec{r}\,|\psi\rangle$, a function of $\vec{r}$, can also be written as $\psi(\vec{r})$. This is called the "wave function." Similarly the projection amplitude $\langle\vec{k}\,|\psi\rangle$ can be written $\psi(\vec{k})$ and called the "wave function in momentum space."

$\langle\vec{r}|\vec{r}\rangle$ has any finite value whatever. In order to satisfy equality 27, the amplitude $\langle\vec{r}|\vec{r}\rangle$ must be infinite. The "function" $\langle\vec{r}|\vec{r}\,'\rangle$ is then a peculiar function indeed: it is zero everywhere except at the single point $\vec{r}\,' = \vec{r}$, where it is infinite in such a way as to satisfy Eq. 27. This function was first proposed by Dirac and is called the "Dirac delta function" (δ-function). Specifically, $\delta(\vec{r} - \vec{r}\,')$ is a function with the following properties:

$$\delta(\vec{r} - \vec{r}\,') = 0 \qquad \vec{r} \neq \vec{r}\,' \tag{28}$$

$$\int d^3r \ \delta(\vec{r} - \vec{r}\,') = 1 \tag{29}$$

$$\int d^3r \ f(\vec{r})\,\delta(\vec{r} - \vec{r}\,') = f(\vec{r}\,') \tag{30}$$

for any well-behaved function $f(\vec{r})$. Equation 27 is just of the form (30), with $f(\vec{r})$ equal to the amplitude $\langle\vec{k}|\vec{r}\rangle$. The delta function finds frequent application in the quantum mechanical description of continuum states. Strictly speaking, it is not a function but is called by mathematicians a "generalized function" or a "distribution." We call this method of treating the amplitude $\langle\vec{r}\,'|\vec{r}\rangle$ "delta function normalization" and write

$$\langle\vec{r}\,'|\vec{r}\rangle = \delta(\vec{r}\,' - \vec{r}) \tag{31}$$

A similar argument can be carried through starting from Eq. 26 and leads to the conclusion that if this equation is to be valid, the momentum states must likewise be normalized to a delta function:

$$\langle\vec{k}\,'|\vec{k}\rangle = \delta(\vec{k}\,' - \vec{k}) \tag{32}$$

The delta function normalization for the basis states does not imply, however, that all spatial states must be normalized in this fashion. On the contrary, the normalization of $|\psi\rangle$ is still completely arbitrary. For on writing the dual of (25) and taking the inner product of $\langle\psi|$ with $|\psi\rangle$, we obtain

$$\langle\psi|\psi\rangle = \int d^3r \int d^3r' \ \langle\psi|\vec{r}\,'\rangle\langle\vec{r}\,'|\vec{r}\rangle\langle\vec{r}|\psi\rangle = \int d^3r \ |\langle\vec{r}|\psi\rangle|^2 \tag{33}$$

To obtain the last form we have put $\langle\vec{r}\,'|\vec{r}\rangle = \delta(\vec{r}\,' - \vec{r})$ and used the property (30) of the delta function. If the integral (33) converges, the state $|\psi\rangle$ is said to be normalizable; it is then convenient to choose a normalization such that the integral in (33) has the value unity. With this normalization, the quantity $d^3r\,|\langle\vec{r}|\psi\rangle|^2$ represents exactly the probability of detecting the particle within the volume element d^3r around $\vec{r}$ in a position measurement. On the other hand, if the integral (33) diverges, such an interpretatiion is not possible. One is then free to choose an arbitrary normalization, keeping in mind the fact that only relative probabilities are correctly given.

In point of fact, only normalizable states can represent the state of a real particle. We have already remarked that the momentum state $|\vec{k}\rangle$, for which the normalization integral (33) clearly diverges, is an idealization. No real particle could be in a state such that the particle is equally likely to be anywhere in all space. If it were, the probability of detecting the particle in any finite region of space would necessarily be zero. Likewise, no real particle could be in a state for which all values of momentum are equally likely, as is the case for the state $|\vec{r}\rangle$. These states are useful for two reasons. First, they represent ideal limits useful in analyzing real states. Second, they can be superposed to provide a description of states that are normalizable.

Our choice of a delta function normalization for the position and momentum basis states implies a definite value for the constant A in the projection amplitudes. $\langle\vec{r}|\vec{k}\rangle$ and $\langle\vec{k}|\vec{r}\rangle$ (eqs. 18, 19). To see this, put $|\psi\rangle = |\vec{r}\rangle$ in expansion (26) and take the inner product with the similar expression for the dual vector $\langle r'|$. The result is

$$\langle\vec{r}\,'|\vec{r}\rangle = \int\langle\vec{r}\,'|\vec{k}'\rangle\, d^3k'\, \langle\vec{k}'|\int|\vec{k}\rangle\, d^3k\, \langle\vec{k}|\vec{r}\rangle$$

or

$$\delta(\vec{r}\,' - \vec{r}) = A^2\int d^3k' d^3k e^{i(\vec{k}\cdot\vec{r} - \vec{k}'\cdot\vec{r}\,')}\,\delta(\vec{k}' - \vec{k}) = A^2\int d^3k e^{i\vec{k}\cdot(\vec{r} - \vec{r}\,')} \tag{34}$$

But it may be shown (see the appendix) that

$$\int d^3k\, e^{i\vec{k}\cdot(\vec{r} - \vec{r}\,')} = (2\pi)^3\,\delta(\vec{r}\,' - \vec{r}) \tag{35}$$

Hence, in order to satisfy eq. (34), A must have the value

$$A = (2\pi)^{-3/2} \tag{36}$$

That is, the basic projection amplitudes are, with this normalization

$$\langle\vec{r}|\vec{k}\rangle = \frac{1}{(2\pi)^{3/2}}\, e^{i\vec{k}\cdot\vec{r}} = \langle\vec{k}|\vec{r}\rangle^*. \tag{37}$$

The transformation law between the amplitudes $\langle\vec{r}|\psi\rangle$ and $\langle\vec{k}|\psi\rangle$ is likewise determined. Taking the inner product of $\langle\vec{k}|$ with expansion (25) for $|\psi\rangle$, we get

$$\begin{aligned}\langle\vec{k}|\psi\rangle &= \int d^3r\, \langle\vec{k}|\vec{r}\rangle\langle\vec{r}|\psi\rangle \\ &= (2\pi)^{-3/2}\int e^{-i\vec{k}\cdot\vec{r}}\, \langle\vec{r}|\psi\rangle\, d^3r\end{aligned} \tag{38}$$

Similarly, one also finds for the inverse transformation,

$$\langle\vec{r}|\psi\rangle = \int d^3k\, \langle\vec{r}|\vec{k}\rangle\langle\vec{k}|\psi\rangle = (2\pi)^{-3/2}\int e^{i\vec{k}\cdot\vec{r}}\, \langle\vec{k}|\psi\rangle\, d^3k \tag{39}$$

The last two relations assert that the position and momentum amplitudes of any state are Fourier transforms of one another.

APPENDIX. SOME PROPERTIES OF DELTA FUNCTIONS

The three-dimensional delta function defined by the properties (28) to (30) is a product of three one-dimensional functions with similar properties:

$$\delta(\vec{r} - \vec{r}\,') = \delta(x - x')\,\delta(y - y')\,\delta(z - z') \tag{A1}$$

where $\delta(x - x')$, for example, has the properties

$$\delta(x - x') = 0 \qquad x \neq x' \tag{A2}$$

$$\int_{-\infty}^{+\infty} \delta(x - x')\,dx = 1 \tag{A3}$$

$$\int_{-\infty}^{+\infty} \delta(x - x')f(x)dx = f(x') \tag{A4}$$

Property (A4) may be taken by itself as the defining property of the delta function, since any expression that contains a delta function becomes a proper mathematical function only after an integration is performed over the argument of the delta function.

The following properties can be demonstrated directly from the defining relations:

$$\delta(-x) = \delta(x) \tag{A5}$$

$$\delta(ax) = \frac{1}{a}\,\delta(x) \tag{A6}$$

$$x\,\delta(x) = 0 \tag{A7}$$

Equation 35 of the text is obtained using the following identity, which is one form of the Fourier integral theorem:*

$$f(x') = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dk \int_{-\infty}^{+\infty} dx\; f(x)e^{ik(x'-x)} \tag{A8}$$

This "implies" that

$$\int_{-\infty}^{+\infty} e^{ik(x'-x)}dk = 2\pi\,\delta(x-x') \tag{A9}$$

On multiplying three expressions of the form (A9) using x, y, and z, one obtains Eq. 35 of the text. Let us look at the integral (A9) to verify that it behaves as a delta function. It is obvious that theintegral is infinite for $x = x'$. For $x \neq x'$ the limit

$$\lim_{a\to\infty}\int_{-a}^{a} e^{ik(x'-x)}dk \tag{A10}$$

does not exist. (The integral oscillates.) Therefore property (A2) is not satisfied. However,

---

* See, for example, page 247 of The Mathematics of Physics and Chemistry by H. Margenau and G. Murphy, D. van Nostrand Company, New York, 1951.

the defining property of the delta function (A4) is satisfied. If we multiply (A10) by some function f(x) and integrate over x, and if it is permissible to interchange the order of integration, we obtain

$$\lim_{a\to\infty}\int_{-a}^{a} dk \int_{-\infty}^{\infty} dx\, f(x) e^{ik(x'-x)} \qquad \text{(A11)}$$

which is the integral (A8) and does converge.

Chapter 10. The Motion of Free Particles

1. Introduction

In the preceding chapter we introduced momentum states $|\vec{k}\rangle$ and position states $|\vec{r}\rangle$, and argued that these provide alternative bases for the description of particle states. As has already been remarked, no actual particle could ever be observed precisely in the state $|\vec{k}\rangle$, inasmuch as this state is non-normalizable and implies equal probability for detecting the particle anywhere in space. Likewise, the precise state $|\vec{r}\rangle$ is not one in which we could expect to find an actual particle. In the present chapter we show how appropriate superpositions of the basis states may be constructed which are physically realizable, and investigate some of their properties. We begin by constructing a state whose position amplitude is localized near the point $\vec{r}_o$, and whose momentum amplitude is localized near the value $\vec{k}_o$. The general relation between position and momentum amplitudes (Eq. 9-38 and 9-39) leads to a restriction on the product of the possible position and momentum localizations. This restriction is the so-called "uncertainty relation." By introducing a time-dependent factor in the amplitudes, we then show how such localized "packet states" change with time and can be used to describe particle motion.

2. Localized States

As a first step toward the construction of a localized state, consider a superposition of two distinct momentum states:

$$|\psi\rangle = |\vec{k}_1\rangle\alpha + |\vec{k}_2\rangle\beta \qquad (1)$$

where α and β are complex constants. Since the individual states $|\vec{k}_1\rangle$ and $|\vec{k}_2\rangle$ are themselves non-normalizable, their superposition is likewise non-normalizable. However it exhibits, for the first time, a probability distribution that varies with position. From Eq. 37 of Chapter 9, the position amplitudes for the state (1) are given by

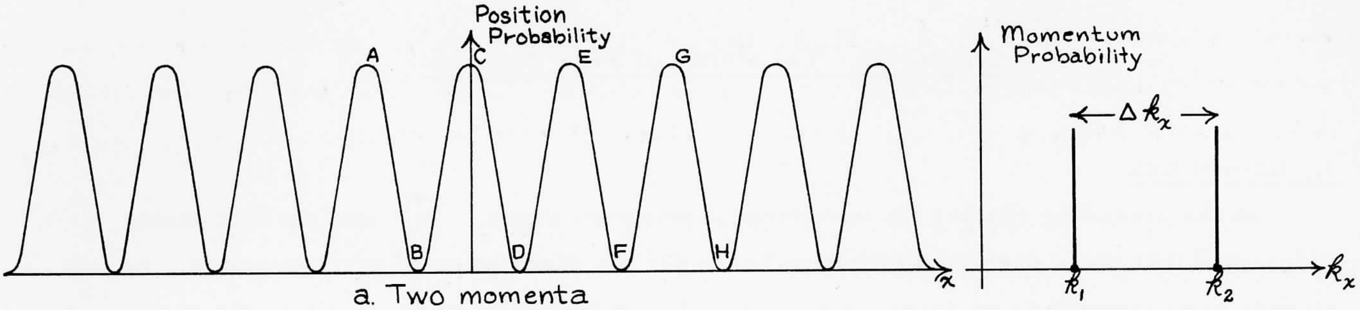
$$\langle\vec{r}|\psi\rangle = (2\pi)^{-3/2}(\alpha e^{i\vec{k}_1\cdot\vec{r}} + \beta e^{i\vec{k}_2\cdot\vec{r}}) \qquad (2)$$

The relative probability for detecting a particle in this state within the volume element d^3r about the point $\vec{r}$ is therefore

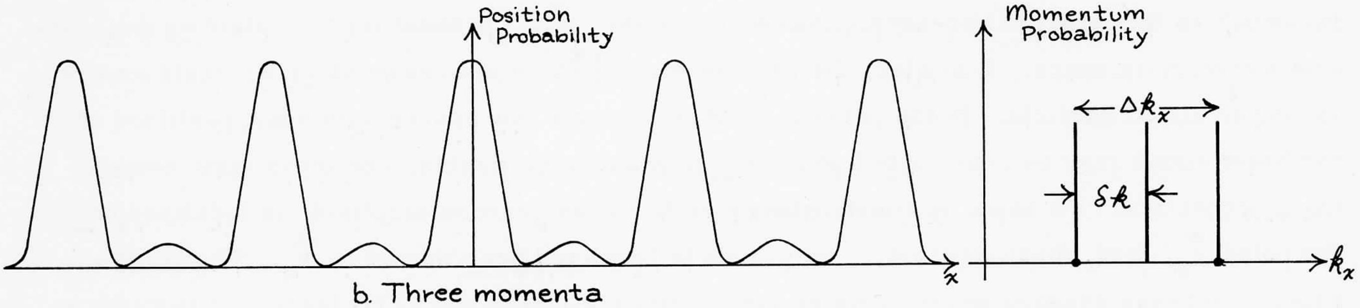
$$|\langle\vec{r}|\psi\rangle|^2 d^3r = (2\pi)^{-3}\left(|\alpha|^2 + |\beta|^2 + 2\mathrm{Re}[\alpha^*\beta e^{(\vec{k}_2-\vec{k}_1)\cdot\vec{r}}]\right)d^3r \qquad (3)$$

As a specific example, let $\alpha = \beta = 2\pi^{3/2}$. Then Eq. (3) reads

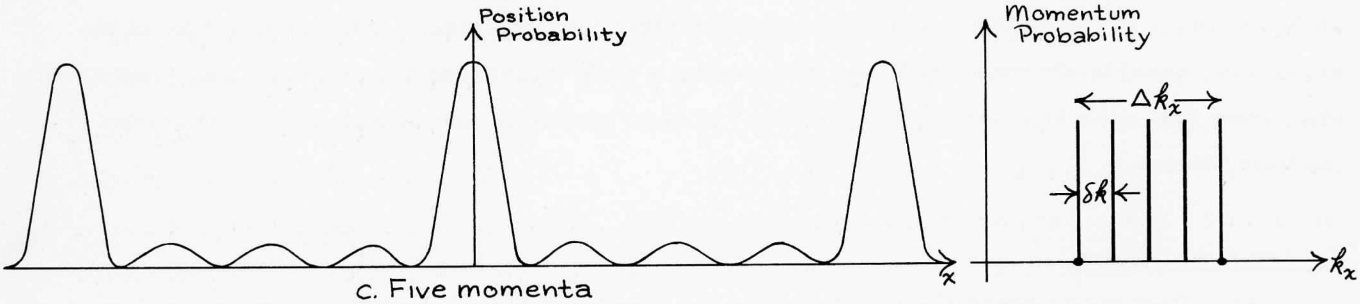
$$|\langle\vec{r}|\psi\rangle|^2 = 1 + \cos(\vec{k}_2 - \vec{k}_1)\cdot\vec{r} = 2\cos^2\left(\frac{\Delta\vec{k}\cdot\vec{r}}{2}\right) \qquad (4)$$



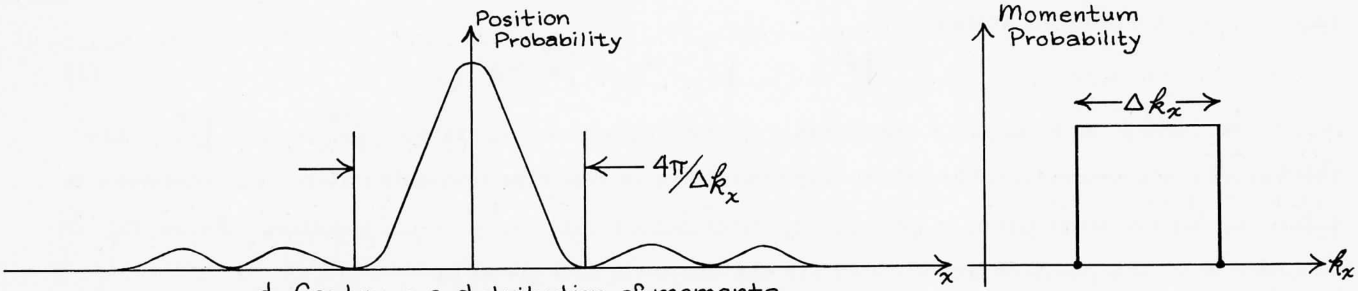
a. Two momenta



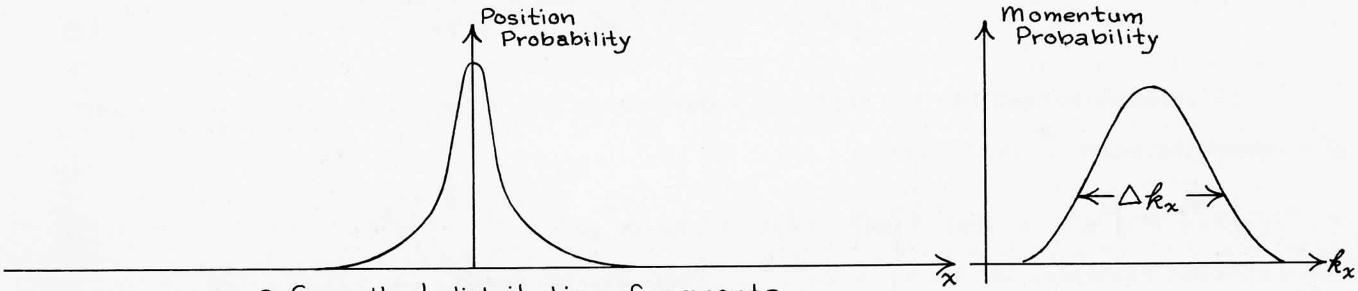
b. Three momenta



c. Five momenta



d. Continuous distribution of momenta



e. Smoothed distribution of momenta

Figure 1. For caption, see bottom of facing page.

where we have put $\Delta\vec{k} = \vec{k}_2 - \vec{k}_1$. The function (4) is plotted in Figure 1a; for concreteness we suppose that $\Delta\vec{k}$ points in the x direction. A particle in state $|\psi>$ would be most likely to be detected near points A, C, E, G, etc., and least likely to be detected near points B, D, F, etc.

One would be hard put to prepare a state that even closely resembles the one described by Figure 1a. Even apart from the problem of non-normalizability, the presence of two distinct momentum components is very hard to achieve. In particular, we emphasize that this state can not be approximated by mixing two beams, one containing particles with momenta near $\vec{k}_1$ and the other containing particles with momenta near $\vec{k}_2$. (Just as the spin state $|+x>$ cannot be prepared by mixing two beams in states $|+z>$ and $|-z>$ from different sources.)

The superpositions of three distinct momentum states, or of any finite number, exhibit properties not qualitatively different from those of the state just considered. Figure 1b shows the pattern of probability distribution corresponding to a state that contains three momentum components; Figure 1c is a similar plot for a state with five components. In each of these cases the maximum and minimum momentum components have been chosen to be respectively the same as in Figure 1a, differing by the fixed amount $\Delta\vec{k}$. As the figures show, increasing the number of intermediate momentum components has the effect of pushing the principal peaks in the spatial probability distribution farther and farther apart. However, a superposition of any finite number of momentum components has a spatial probability distribution that is necessarily periodic. No such superposition can represent a state in which the probability is concentrated near one point; in a genuinely localized state $|<r|\psi>|^2$ must be vanishingly small outside some finite region of space. A localized state can be obtained by going to the limit of a superposition of a continuous band of momentum states, as we now proceed to verify.

Caption for Figure 1. Superposition of momentum basis states to construct localized packet states. The right hand set of curves shows the momentum basis states used in each superposition in the special case that all momenta lie along the x-direction. The vertical scale is proportional to the coefficient of the corresponding momentum state, e.g., α and β in Eq. 1. The left hand set of curves show the relative probability of finding the particle at different positions along the x-axis. The vertical scales are not the same for curves a, b, c, etc. The width of all intensity peaks for these special cases is given approximately by the expression $\Delta x \sim 1/\Delta k$.

a. Superposition of two momentum basis states. The probability curve is given by Eq. 4. The distance between adjacent maxima is $2\pi/\Delta k$.

b. Superposition of three momentum basis states. Probability peaks are the same width, but major peaks are now separated by twice the distance of case a: $2\pi/\delta k = 4\pi/\Delta k$.

c. Superposition of five momentum basis states. Major probability peaks are now separated by four times the distance of case a: $2\pi/\delta k = 8\pi/\Delta k$. As long as the number of momentum components is finite, the probability pattern will be periodic.

d. By superposing a continuous range of momentum basis states, one can reduce the number of major probability peaks to one. This is the case treated in Eqs. 5 to 9 of the text. Subsidiary maxima in probability result from the abrupt cutoffs in the momentum distribution.

e. Smoothed out momentum distribution leads to a single maximum in the probability distribution.

Consider first an equal superposition of a band of momentum states, all in the same direction (say x), with magnitudes between k_1 and k_2. Using the notation $|k_x, k_y, k_z>$ to express explicitly the three components of $\vec{k}$, we can write the state in question as

$$|\psi> = \int_{k_1}^{k_2} dk_x \, |k_x, 0, 0> \tag{5}$$

This corresponds to momentum amplitudes

$$<\vec{k}|\psi> = \delta(k_y)\delta(k_z)f(k_x) \tag{6}$$

where

$$\begin{aligned} f(k_x) &= 1 \qquad k_1 < k_x < k_2 \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{7}$$

The position amplitudes are then given by the integral

$$<\vec{r}|\psi> = (2\pi)^{-3/2} \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y \int_{-\infty}^{\infty} dk_z \, \delta(k_y)\delta(k_z)f(k_x)e^{i(k_x x + k_y y + k_z z)} \tag{8}$$

From the property (9-A3) of the δ-function the k_y and k_z integrations in (8) give unity, and we are left with an elementary integral over k_x:

$$<\vec{r}\,|\,\psi> = (2\pi)^{-3/2} \int_{k_1}^{k_2} e^{ik_x x} dx = 2(2\pi)^{-3/2} e^{ik_{ox}x} \left[\frac{\sin(\Delta k_x x/2)}{x}\right] \tag{9}$$

where we have put $k_{ox} = \frac{1}{2}(k_1 + k_2)$ and $\Delta k_x = k_2 - k_1$.

The absolute square of the position amplitude (9) is ploted in Figure 1d. A measurement of the x-component of position is most likely to give a result within the central maximum in this figure. The appearance of subsidiary maxima in the position probability curve is a consequence of the sharp cutoff in the momentum probability distribution. This sharp momentum cutoff was introduced to make the integral in (9) easy to evaluate. A more physically realistic momentum distribution is obtained by "rounding the corners" of the momentum probability distribution, as in Figure 1e. The resulting position probability distribution has a single peak and no subsidiary maxima. A form often used to describe a localized probability distribution is the Gaussian, which has the interesting property that its Fourier transform is likewise a Gaussian. Thus if the k_x dependence of $<\vec{k}\,|\,\psi>$ is proportional to $e^{-ak_x^2}$, then the x dependence of $<\vec{r}/\psi>$ is proportional to e^{-bx^2}, where $\underline{a}$ and $\underline{b}$ are related constants.

The state (5) is localized only in x. Since the y and z components of momentum were specified to be zero, the spatial probability distribution is uniform in y and z. However, the development above is readily generalized to construct a state localized in all three dimensions. Such a state is

$$|\psi> = A \int_{k_{1x}}^{k_{2x}} dk_x \int_{k_{1y}}^{k_{2y}} dk_y \int_{k_{1z}}^{k_{2z}} dk_z \, |k_x, k_y, k_z> \tag{10}$$

That is, $<\vec{k}\,|\,\psi>$ has the constant value A for $\vec{k}$ within a rectangular box of dimensions

$\Delta k_x = k_{2x} - k_{1x}$, $\Delta k_y = k_{2y} - k_{1y}$, $\Delta k_z = k_{2z} - k_{1z}$, and is zero elsewhere. The corresponding position amplitudes are

$$\begin{aligned}\langle\vec{r}\,|\,\psi\rangle &= (2\pi)^{-3}\int d^3k\, e^{i\vec{k}\cdot\vec{r}}\,\langle\vec{k}\,|\,\psi\rangle \\ &= A(2\pi)^{-3}\int_{k_{1x}}^{k_{2x}} dk_x \int_{k_{1y}}^{k_{2y}} dk_y \int_{k_{1z}}^{k_{2z}} dk_z\, e^{i(k_x x + k_y y + k_z z)} \\ &= A(\pi)^{-3} e^{i\vec{k}_o\cdot\vec{r}}\;\frac{\sin(\Delta k_x x/2)}{x}\;\frac{\sin(\Delta k_y y/2)}{y}\;\frac{\sin(\Delta k_z z/2)}{z} \end{aligned} \qquad (11)$$

where k_{oy} and k_{oz} are defined analogously to k_{ox}. State (10) is localized around the origin in space. The spatial probability distribution varies in the x, y, and z directions according to Figure 1d. Furthermore, the state is normalizable; since the integral $\int_{-\infty}^{\infty}\frac{\sin^2(ax)}{x^2}\,dx$ converges, we can choose the constant A in Eq. 10 so as to make $\int|\langle\vec{r}\,|\,\psi\rangle|^2 d^3r = 1$. The more reasonable smoothed momentum distribution of Figure 1e can likewise be generalized to three dimensions and leads to a result similar to (11) but without any secondary maxima. Any such normalizable localized state is called a packet state. The corresponding position probability amplitude is usually called a wave packet.

One can think of preparing a packet state by allowing a steady beam of particles to impinge on a shutter that opens and closes. The probability for finding a particle at a position in space beyond the shutter is approximated by the left hand curve of Figure 1e, while the probability of detecting one or another momentum in the packet state is approximated by the right hand curve of the same figure. The relation between position and momentum localization in a packet state is developed in the following section.

3. Position-momentum uncertainty relations.

The two curves of Figure 1d specify the relative probabilities of various possible results in measurements of x and of k_x carried out on particles in the state (10). As in all quantum states previously considered, experimental determination of these probabilities requires that measurements be carried out on many particles, identically prepared. Any measurement of k_x carried out on a particle in the state in question is sure give a result between k_1 and k_2. Thus the quantity $k_2 - k_1 = \Delta k$ may be called the momentum width or momentum uncertainty of the state. In a position measurement there is a non-zero probability of obtaining arbitrarily large results, but it is clear that most measurements will yield a position within the central maximum of the figure. Thus we may take the extent of the central maximum as an approximate measure of the position width or position uncertainty Δx. Eq. (11) gives, for Δx thus defined,

$$\Delta x = 4\pi/\Delta k_x$$

or

$$\Delta x\,\Delta k_x = 4\pi \qquad (12)$$

Evidently, there is a certain degree of arbitrariness in the definition of Δx.* For example, we could have chosen it to be the full width at half maximum of the probability distribution, or the interval in x that contains half the total probability. Each of these choices would have led to a slightly different numerical value for the width. By any reasonable definition, however, Δx is inversely proportional to Δk_x and their product is of order unity.

By a similar analysis, one can assign position and momentum widths (uncertainties) to any packet state. The transformation law between position and momentum amplitudes ensures that a packet narrow in a given position coordinate must be broad in the corresponding momentum coordinate, and vice versa. In every case the "uncertainty product" (12) is at least of order unity:

$$\Delta x \, \Delta k_x \gtrsim 1 \tag{13a}$$

or, in terms of p,

$$\Delta x \, \Delta p_x \gtrsim \hbar \tag{14a}$$

The minimum uncertainty product is achieved by a smooth distribution such as the one of Figure 1e; this fulfills the approximate equality in Eqs. (13), (14). Eq. (14) is one of the Heisenberg uncertainty relations. It is a direct consequence of the Fourier transform relations (9-38, 9-39) between position and momentum amplitudes. Similar relations obviously hold between Δy and Δp_y, and between Δz and Δp_z, analogously defined.

$$\Delta y \, \Delta k_y \gtrsim 1 \tag{13b}$$

$$\Delta z \, \Delta k_z \gtrsim 1 \tag{13c}$$

and

$$\Delta y \, \Delta p_y \gtrsim \hbar \tag{14b}$$

$$\Delta z \, \Delta p_z \gtrsim \hbar \tag{14c}$$

Other uncertainty relations will arise later.

We remark that the position state $|\vec{r}\rangle$ represents the limiting case of a packet state of zero position width; the momentum width, as we have seen, is then infinite. The state $|\vec{k}\rangle$ represents the opposite limit of zero momentum width and infinite position width. No matter how one approaches these limiting cases, the uncertainty relation (14) must be satisfied for every intermediate case.

Considerable confusion exists concerning the significance of the uncertainty relations. The relations do not restrict the precision with which either position or momentum of an individual particle may be measured. In fact, they say nothing about the result of a single measurement. Rather, they relate the spread in the results of many position measurements (each carried out on a different particle in the same state) to the spread in results of many momentum measurements carried out on particles in the same state. In brief, the uncertainty relations assert that no state exists for which both a position and a momentum measurement are

---

* In general, also in the definition of Δk_x. Only because of the particular choice of a rectangular shape is Δk_x in the present example uniquely defined.

sure to give results within arbitrarily small intervals.

The uncertainty relations are helpful in deciding whether a classical description of particle motion is adequate in a given situation. The classical notion of a trajectory, in which both position and momentum are precisely specified at every instant, is in principle incompatible with the uncertainty relations. However, even in a classical description the trajectory need in general be specified only within certain tolerances, whose magnitudes depend on the particular circumstances. If the product of the tolerable classical "widths" in position and momentum already far exceeds the value of Planck's constant $\hbar$, then it is possible to construct a packet state which for all practical purposes duplicates the classical trajectory.* Under such circumstances the classical treatment provides a satisfactory description.

A couple of examples illustrate the argument. Suppose we wish to construct a p-projector (Chapter 9, Figure 1) which will select protons of momentum 10^{-16} gm cm/sec ($\frac{v}{c} \approx .002$) with a precision of 0.1%. With a magnetic field of 1000 gauss the distance D between entrance and exit slits is (Eq. 9-1) 12.5 cm. The orbit diameter for a proton whose momentum is 0.1% too high is 0.1% higher than that for one with the prescribed momentum. Therefore, in order to reject such a proton, which follows the dotted path in Figure 2, the slit openings must be no wider than 0.005 cm. If a measurement of y position is carried out on protons just after they have passed through the entrance slit, the results will all lie within the interval $\Delta y = .005$ cm. The uncertainty relation tells us that if we try to construct a packet state to describe the situation, such a state must contain a spread of y-momenta

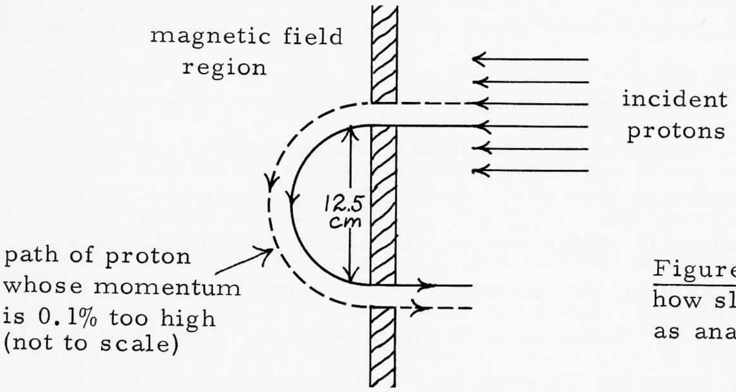


Figure 2. First half of a $\vec{p}$-projector, showing how slit widths determine momentum spread as analyzed classically. Not to scale.

$$\Delta p_y \gtrsim \hbar/\Delta y \approx 2 \times 10^{-23} \text{gm cm/sec} \qquad (15)$$

This means that even a proton which was incident in an absolutely perpendicular direction must be expected to acquire, as a result of passing through the slit, a y-momentum of the order given by Eq. (15). However, since this is only about 10^{-7} times the x-momentum, the resultant broadening of the beam is insignificant, and the uncertainty relations place no important restrictions on the operation of the projector. On the other hand, if we were to try to increase

---

* Packet states that describe particle motion are discussed in the next section.

the precision of the device by several orders of magnitude, the same analysis would lead to an unacceptably high value of Δp_y. Physically, this means that the entrance slit would have become so narrow that diffraction effects broaden the beam by an amount greater than the slit opening. The projector could not then operate with the precision demanded.

A similar analysis can be used to demonstrate that a classical orbit description cannot possibly apply to the motion of the electron in the ground state of the hydrogen atom. A packet state that resembles the classical orbit would have to have a position width substantially smaller than the dimensions of the orbit, i.e., $< 10^{-8}$ cm. Likewise, the momentum width would have to be smaller than the magnitude of the momentum, which is $\sim 10^{-19}$ gm cm/sec. ($v \sim 10^{8}$ cm/sec). This would require

$$\Delta x \Delta p_x \ll 10^{-27} \text{ gm m}^2/\text{sec} \tag{16}$$

which is excluded by the uncertainty relation. Consequently, no packet state can be constructed which in any way resembles the classical orbit under discussion. The ground state of the hydrogen atom must be analyzed by quantum mechanics (Chapter 12).

4. Time-dependent phases

In the first part of this chapter we have examined the spatial probability distributions of various superpositions of momentum states, and seen how particular superpositions represent "packet states" localized in space. One important part of the description of a state is still lacking, namely the time dependence. In general, states evolve with time: the results of a measurement carried out at one time are in general not the same as they are if one waits to carry out the same measurement at a later time. Since each of the probability distributions we have studied is time-independent, it can describe the state in question only at one fixed instant of time.

In finding the correct generalization to include time dependence we are guided by the results of Chapter 8. There we found that we could account for the "precession" of spin states by associating with each of the states $|\pm z\rangle$ in a magnetic field B_z the time dependence $e^{-iEt/\hbar}$, where E is the energy of the "stationary state" in question.

In the present case of free particles, the energy involved is kinetic; a momentum state $|\vec{k}\rangle$ can be considered a stationary state for a free particle, with energy

$$E_k = p^2/2m = \hbar^2 k^2/2m . \tag{17}$$

We are therefore led by this inductive reasoning to associate with the state $|\vec{k}\rangle$ the time-dependent phase factor $e^{-i(\hbar k^2/2m)t}$. As in the case of spin states, this phase factor does not imply any observable time-dependent results for a stationary state*; but for packets states it

---

* Nor for any superposition of states that have the same energy, i.e., states for which the magnitude of momentum is the same. This explains why we were able to describe the interference patterns of Chapter 9 without including any time dependence.

leads to a probability distribution that moves with time, as we shall see below.

To indicate explicitly when we are discussing time-dependent states, we use the notation $|\psi(t)\rangle$ already introduced in Chapter 8. In particular, the time-dependent state for a single momentum $\vec{k}$ will be denoted by $|\psi_k(t)\rangle$. That is, we write

$$|\psi_k(t)\rangle = |\vec{k}\rangle e^{-iE_k t/\hbar} \tag{18}$$

The appearance of the time dependence $e^{-iEt/\hbar}$, postulated here on the basis of the analogy with spin states, can be inferred also from a relativistic argument presented in Box 1.

Box 1. Derivation of the phase factor $e^{i(\vec{p}\cdot\vec{r} - Et)/\hbar}$ using relativity

Suppose one demands that the phase of the spatial amplitude for a state of momentum $\vec{p}$ be the same at a particular place and time (relativistic terminology: a particular event) as observed in all frames of reference. In the laboratory frame the coordinates of the event are $\vec{r}$ and t. An observer in a rocket that zooms past designates the same event by the coordinates $\vec{r}\,'$ and t', related to $\vec{r}$ and t by the Lorentz transformation equation. (See, for example, Physics, a New Introductory Course, Part III, page 3-13.) Momentum and energy in the two frames are also related by a Lorentz transformation equation. (PANIC, Part III, page 7-13.) Using these relations, one can verify that $\vec{p}\cdot\vec{r} = p_x x + p_y y + p_z z$ does not have the same value as $\vec{p}\,'\cdot\vec{r}\,' = p_{x'}x' + p_{y'}y' + p_{z'}z'$. In relativistic terminology the product $\vec{p}\cdot\vec{r}$ is not invariant under a Lorentz transformation. Hence the phase factor $e^{i\vec{p}\cdot\vec{r}/\hbar}$ does not have the same value for the laboratory observer as for the rocket observer, contrary to our requirement. A quantity that includes the term $\vec{p}\cdot\vec{r}$ and has the same value for all observers is the invariant expression $\vec{p}\cdot\vec{r} - Et$, which equals $\vec{p}\,'\cdot\vec{r}\,' - E't'$, as can be demonstrated directly from the Lorentz transformation equations. Thus the inclusion of the phase factor $e^{-iEt/\hbar}$ leads to a relativistic invariant phase.

In brief, the derivation starts from the assumption that any amplitude for a particular state at a given point of space-time must be the same in any frame of reference. This condition is not satisfied by the phase factor $e^{i\vec{k}\cdot\vec{r}}$, since $\vec{k}\cdot\vec{r}$ is not a relativistic invariant. However, the quantity $\vec{k}\cdot\vec{r} - (E/\hbar)t$, where E is the total energy of the particle, is an invariant. Consequently the inclusion of the factor $e^{-i(E/\hbar)t}$ leads to a relativistically invariant phase.

The relativistic argument is, at best, suggestive: The theory we have developed is a nonrelativistic one, and it is not clear what role, if any, a relativistic argument can play. Eventually, the inclusion of the time-dependent phase factor must be justified by the experimental results it predicts.

5. The motion of a free particle

We now examine the consequences that result from including the time-dependent phase factors in states that are superpositions of basis states with different energies. We begin once again with a simple superposition of two different momentum states. Let

$$|\psi(t)\rangle = 2\pi^{3/2}\left[|\psi_{k_1}(t)\rangle + |\psi_{k_2}(t)\rangle\right] \qquad (19)$$

where the multiplicative factor $2\pi^{3/2}$ is chosen for numerical simplicity. From Eq. 18 we obtain

$$\langle \vec{r}|\psi(t)\rangle = \frac{1}{\sqrt{2}}\left[e^{i(\vec{k}_1\cdot\vec{r} - E_1 t/\hbar)} + e^{i(\vec{k}_2\cdot\vec{r} - E_2 t/\hbar)}\right] \qquad (20)$$

As before, put

$$\begin{aligned}\vec{k}_2 &= \vec{k}_o + \Delta\vec{k}/2\\ \vec{k}_1 &= \vec{k}_o - \Delta\vec{k}/2\end{aligned} \qquad (21)$$

and use (17). Then (20) takes the form

$$\langle\vec{r}|\psi(t)\rangle = \frac{1}{\sqrt{2}}\, e^{i\left[\vec{k}_o\cdot\vec{r} - \frac{\hbar}{2m}(k_o^2 + \Delta k^2)t\right]}\left[2\cos\frac{1}{2}(\Delta\vec{k}\cdot\vec{r} - \Delta E t/\hbar)\right] \qquad (22)$$

where

$$\Delta E = E_2 - E_1 = \frac{\hbar^2}{m}\vec{k}_o\cdot\Delta\vec{k} \qquad (23)$$

The spatial probability distribution is proportional to

$$|\langle\vec{r}|\psi\rangle|^2 = 2\cos^2\frac{1}{2}(\Delta\vec{k}\cdot\vec{r} - \Delta E t/\hbar) \qquad (24)$$

This is an "interference pattern" in both space and time. The pattern at any fixed time is as shown in Figure 1a, but the entire pattern moves in the direction of $\Delta\vec{k}$ at a velocity given by

$$v = \frac{\Delta E}{\hbar\Delta k} = \frac{\hbar k_o}{m} \qquad (25)$$

This velocity is just p_o/m, the classical velocity of a particle whose momentum is the mean of the two momentum components in the state.

Consider now the time development of a general superposition of momentum states. If each momentum component has the time dependence given by (18), the general time-dependent state vector $|\psi(t)\rangle$ can be expanded in the form

$$|\psi(t)\rangle = \int d^3k\, |\vec{k}\rangle C_k\, e^{-iE_k t/\hbar} \qquad (26)$$

where we have written C_k for the time-independent projection amplitude $\langle\vec{k}|\psi(0)\rangle$. The spatial distribution at time t is

$$|\langle\vec{r}|\psi(t)\rangle|^2 = \frac{1}{(2\pi)^3}\left|\int d^3k C_k e^{i(\vec{k}\cdot\vec{r} - \hbar k^2 t/2m)}\right|^2 \qquad (27)$$

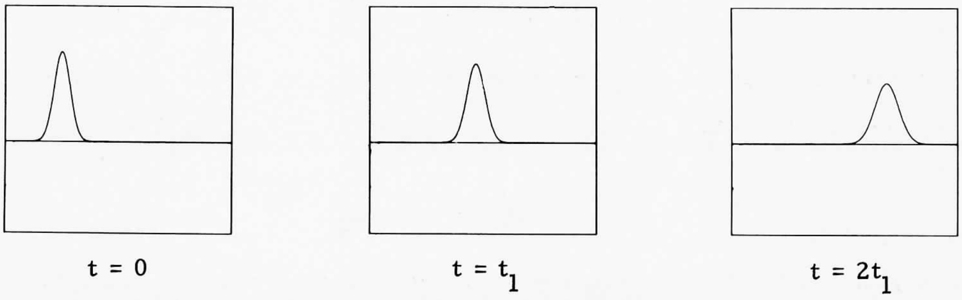
Because of the quadratic term in the exponent, (27) is not an elementary integral even in the simple case treated in Eq. 11. However, the integral can always be evaluated, numerically if necessary. The result of this analysis is that the maximum of the probability distribution (27) moves with the speed p_o/m, which is exactly the same as the classical speed of a particle of momentum p_o equal to an average momentum of the packet. Figure 3a shows the result of a computer calculation of the propagation of such a probability distribution.

It is a general property of packet states that the probability distribution (27) spreads as it propagates; that is, the position width increases with time. This spreading is shown by the propagating packet in Figure 3a. (The momentum width, of course, remains unchanged -- Figure 3b.) The probability amplitude used in the calculations that led to Figure 3 is a Gaussian, mentioned earlier in this chapter. The integral (27) can be evaluated in closed form for a Gaussian. For this state the spatial probability distribution remains Gaussian as it spreads. Taking a one-dimensional packet for simplicity, the spatial amplitude at $t = 0$ is

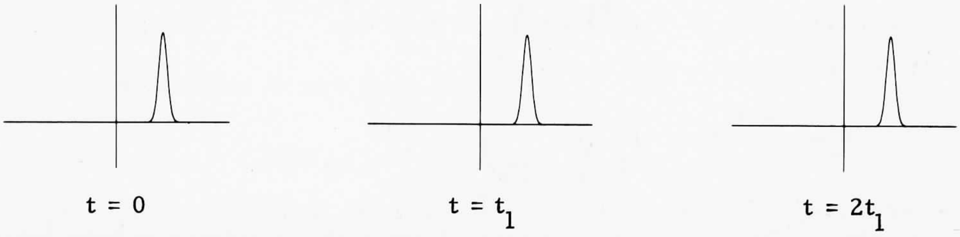
$$\langle x, y, z | \psi(0) \rangle = (\mathrm{const}) e^{ik_o x - x^2/(2a_o^2)} \tag{28}$$

which is a Gaussian packet of mean momentum k_o and spatial width $\approx a_o$. As time passes, the phase of the amplitude changes and the constant a_o in (28) is replaced by the function

$$a(t) = a_o \left(1 + \frac{\hbar^2 t^2}{m^2 a_o^4} \right)^{1/2} \tag{29}$$



a. Spatial probability distribution. Horizontal axis measures position, vertical axis measures position probability.



b. Momentum probability distribution. Horizontal axis measures momentum, vertical axis measures momentum probability.

Figure 3. Time development of position and momentum probability distributions of a packet state as calculated and displayed by a computer. The position probability maximum propagates and spreads with time. The momentum probability distribution for a free particle does not change with time. Courtesy of Dr. Judah Schwartz.

This result illustrates two important general properties: (1) an initially narrow packet spreads more rapidly than an initially broad one (the coefficient of t^2 is inversely proportional to the 4th power of a_o), and (2) no packet state remains localized within a specified position width for an indefinite period. This means that the approximate correspondence between a moving packet and a classical trajectory must eventually break down. However, in many cases of interest the period during which the packet remains localized is so long that the eventual breakdown is not of practical importance (see the exercises).

The spreading of a packet state with time is analogous mathematically to the spreading of an electromagnetic pulse in a dispersive medium, i.e., a medium such as glass in which the velocity of light depends upon the frequency. That such a pulse changes its shape is easily demonstrated by Fourier analyzing the pulse into sinusoidal components. If each component propagates at a different speed, then when we reconstitute the pulse from the components at a later time, the shape will have changed. This analogy breaks down when we consider light pulses traveling in vacuum; in this case all frequencies travel with the same speed, so the pulse does not change shape. In contrast, probability pulses corresponding to free particle packet states <u>always</u> change shape as they propagate.

Chapter 11. The Schrodinger Equation

1. Introduction

Up to this point we have described the quantum states of various systems, taking the amplitudes for specific states either from experiments (sometimes real experiments, sometimes idealized ones) or from classical analogies. We now ask the important question, "What law or laws determine the probability amplitudes for a particular system?" For a wide variety of systems the probability amplitudes are solutions of a partial differential equation called the Schrodinger equation. This equation is the dynamical law for a system; it is analogous to Newton's second law ($\vec{F} = m\vec{a}$) in classical mechanics. In the present chapter we exhibit a differential equation satisfied by the free-particle states studied in the preceding two chapters, and present a plausibility argument that leads to the generalization required to describe particles under the influence of a potential. In following chapters we shall investigate some solutions of the equation.

2. The Schrodinger equation for free particles

In the preceding chapter we introduced the time-dependent free-particle basis states $|\psi_k(t)\rangle$, whose position amplitudes are

$$\langle\vec{r}\,|\psi_k(t)\rangle = \frac{1}{(2\pi)^{3/2}} e^{i(\vec{k}\cdot\vec{r} - \omega_k t)} \tag{1}$$

where

$$\omega_k = E_k/\hbar = \hbar k^2/2m \tag{2}$$

A general free-particle state at $t = 0$ can be expanded in terms of momentum basis states

$$|\psi(0)\rangle = \int d^3k\,|\vec{k}\rangle C_k \tag{3}$$

where

$$C_k = \langle\vec{k}\,|\psi(0)\rangle$$

This state evolves with time according to the expression

$$|\psi(t)\rangle = \int d^3k\,|\vec{k}\rangle C_k e^{-i\omega_k t} \tag{4}$$

The spatial amplitudes at time t are thus

$$\langle\vec{r}\,|\psi(t)\rangle = \frac{1}{(2\pi)^{3/2}} \int d^3k C_k e^{i(\vec{k}\cdot\vec{r} - \omega_k t)} \tag{5}$$

It is our present goal to exhibit a differential equation satisfied by the amplitudes (5), one which will be generalizable to include the case of particles subject to external forces. To this end we first investigate the effect of differentiating the amplitudes (1) and look for a single differential

---

equation that is satified by all of these amplitudes (all values of $\vec{k}$). Such an equation will also be satisfied by the superposition (5) that describes a general free-particle state.

The effect of differentiating the basis amplitudes (1) either with respect to time or with respect to any single space coordinate is merely to multiply the amplitude by a factor that is independent of $\vec{r}$. In particular, differentiating with respect to time one obtains

$$\frac{\partial}{\partial t} \langle \vec{r} | \psi_k(t) \rangle = -i\omega_k \langle \vec{r} | \psi_k(t) \rangle$$

or

$$i\hbar \frac{\partial}{\partial t} \langle \vec{r} | \psi_k(t) \rangle = E_k \langle \vec{r} | \psi_k(t) \rangle \qquad (6)$$

Equation (6) is satified by each of the basis amplitudes (1), each with its own value of energy E_k. This same equation is therefore not satisfied by the superposition (5), because each different value of $\vec{k}$ in the superposition carries a different value of E_k. Therefore we must look further for a differential equation satisfied by (5).

The derivative of (1) with respect to x gives

$$\frac{\partial}{\partial x} \langle \vec{r} | \psi_k(t) \rangle = ik_x \langle \vec{r} | \psi_k(t) \rangle \qquad (7)$$

with similar expressions for

$$\frac{\partial}{\partial y} \langle \vec{r} | \psi_k(t) \rangle \text{ and } \frac{\partial}{\partial z} \langle \vec{r} | \psi_k(t) \rangle$$

All three space derivatives can be combined using the vector operator

$$\vec{\nabla} \equiv \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z} :$$

$$\vec{\nabla} \langle \vec{r} | \psi_k(t) \rangle = i(\hat{x}k_x + \hat{y}k_y + \hat{z}k_z) \langle \vec{r} | \psi_k(t) \rangle = i\vec{k} \langle \vec{r} | \psi_k(t) \rangle$$

or, equivalently,

$$-i\hbar \vec{\nabla} \langle \vec{r} | \psi_k(t) \rangle = \vec{p} \langle \vec{r} | \psi_k(t) \rangle \qquad (8)$$

There is no way to combine Eq. 8 with Eq. 6 to eliminate the dependence on energy E_k or momentum $\vec{p} = \hbar\vec{k}$. But notice that for a free particle $E = p^2/(2m)$. Therefore we continue with the second space derivatives, expecting to find an equation that depends upon p^2 that can be combined with Eq. 6. The second derivative of (1) with respect to x is

$$\frac{\partial^2}{\partial x^2} \langle \vec{r} | \psi_k(t) \rangle = -k_x^2 \langle \vec{r} | \psi_k(t) \rangle \qquad (9)$$

with similar expressions for second derivatives with respect to y and z. Summing the three equations of the form (9), one gets an expression for the Laplacian of the function $\langle \vec{r} | \psi_k(t) \rangle$:

$$\vec{\nabla}\cdot\vec{\nabla} \langle \vec{r} | \psi_k(t) \rangle = \nabla^2 \langle \vec{r} | \psi_k(t) \rangle = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \langle \vec{r} | \psi_k(t) \rangle$$

$$= -k^2 \langle \vec{r} | \psi_k(t) \rangle \qquad (10)$$

which can be written as

$$-\frac{\hbar^2}{2m}\nabla^2\langle\vec{r}\,|\psi_k(t)\rangle = \frac{p^2}{2m}\langle\vec{r}\,|\psi_k(t)\rangle \qquad (11)$$

Equation (8) asserts that the differential operator $-i\hbar\vec{\nabla}$, acting on the amplitude $\langle\vec{r}\,|\psi_k(t)\rangle$, multiplies this amplitude by the momentum for the state.* Likewise, Eq. 11 asserts that the differential operator $-(\hbar^2/2m)\nabla^2$ multiplies the amplitude $\langle\vec{r}\,|\psi_k(t)\rangle$ by the kinetic energy for the state in question, and Eq. 6 asserts that the differential operator $ih\frac{\partial}{\partial t}$ multiplies it by the total energy of the state. Of course, in the present example of a free particle, kinetic energy and total energy are identical. However, the statements as given here are the ones that can be generalized beyond the case of free particles: in a quite general sense the differential operator $-ih\vec{\nabla}$ can be related to the momentum, while $-(\hbar^2/2m)\nabla^2$ is related to the kinetic energy and $i\hbar\frac{\partial}{\partial t}$ is related to the total energy.

Now we are able to combine Eqs. 6 and 11 to eliminate the E_k.

$$-\frac{\hbar^2}{2m}\nabla^2\langle\vec{r}\,|\psi_k(t)\rangle = i\hbar\frac{\partial}{\partial t}\langle\vec{r}\,|\psi_k(t)\rangle \qquad (12)$$

Observe that neither $\vec{k}$ nor ω_k appears in (12). This single differential equation is satisfied by every amplitude $\langle\vec{r}\,|\psi_k(t)\rangle$ for every value of $\vec{k}$. Therefore the equation is also satisfied by any sum of such amplitudes for different $\vec{k}$, and by the linear combination (5) for the general state $|\psi(t)\rangle$. That is, the space amplitudes $\langle\vec{r}\,|\psi(t)\rangle$ for <u>any</u> free-particle state $|\psi(t)\rangle$ satisfy the differential equation

$$-\frac{\hbar^2}{2m}\nabla^2\langle\vec{r}\,|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\langle\vec{r}\,|\psi(t)\rangle \qquad (13)$$

Equation 13 is the Schrodinger equation for free particles, written in terms of position amplitudes.** It merely re-expresses already-known relations. In the next section we discuss the generalization to include external forces, which leads to new results.

One can write also a Schrodinger equation in terms of momentum amplitudes. For free particles this equation is very simple. Expression (4) implies

$$\langle\vec{k}\,|\psi(t)\rangle = C_k e^{-i\omega_k t} \qquad (14)$$

Hence, on differentiating with respect to time one obtains

$$i\hbar\frac{\partial}{\partial t}\langle\vec{k}\,|\psi(t)\rangle = E_k\langle\vec{k}\,|\psi(t)\rangle \qquad (15)$$

---

* We distinguish between a <u>differential operator</u>, such as $\vec{\nabla}$, that acts on amplitudes and the operators discussed in previous chapters that act on state vectors.

** Equation 13 is sometimes called the <u>Schrodinger wave equation</u>. Notice, however, that since only a <u>first</u> time derivative appears, it is not a true wave equation in the sense in which this term is generally employed. The equations that describe electromagnetic waves, elastic waves, etc., all involve the <u>second</u> time derivative.

This is the Schrodinger equation in terms of momentum amplitudes. Despite the resemblance of Eq. 15 to Eq. 6, these equations make very different statements. Equation 6 is in the position representation, and applies to the position amplitudes of the particular basis state $|\psi_k(t)\rangle$. In contrast, Eq. 15 is in the momentum representation and applies to a general state $|\psi(t)\rangle$.

3. The Schrodinger equation for a particle in a potential

Thus far we have considered only free-particle states. But most problems of interest involve interaction of particles, either with other particles or with externally-applied fields. In a hydrogen atom the electron and proton exert forces on one another. In a scattering experiment a free particle strikes a target and is deflected due to interaction with atoms or nuclei in the target. The free-particle Schrodinger equation obtained in the preceding section is useful primarily as a starting point from which one can proceed to the equation satisfied by the amplitudes that describe interacting particles.

We have already pointed out that the differential operation $\left(-\frac{\hbar^2}{2m}\right)\nabla^2$ represents the kinetic energy in the sense that acting on a single amplitude $\langle\vec{r}|\psi_k(t)\rangle$ it multiplies that amplitude by the kinetic energy of a particle in the state $|\psi_k(t)\rangle$. Acting on a linear superposition of such amplitudes, it multiplies each component by the appropriate kinetic energy. We could write Eq. 13 symbolically as

$$i\hbar\frac{\partial}{\partial t}\langle\vec{r}|\psi(t)\rangle = (\text{K.E.})_{op}\langle\vec{r}|\psi(t)\rangle \tag{16}$$

where

$$(\text{K.E.})_{op} \equiv -\frac{\hbar^2}{2m}\nabla^2$$

Suppose now that the particle in question is subject to an external interaction which can be described by a potential energy $V(\vec{r})$. The generalization of Eq. 16, proposed by Schrodinger, is to add a "potential energy operator" and write instead of (16)

$$i\hbar\frac{\partial}{\partial t}\langle\vec{r}|\psi(t)\rangle = \left[(\text{K.E.})_{op} + (\text{P.E.})_{op}\right]\langle\vec{r}|\psi(t)\rangle \tag{18}$$

The "potential energy operator" is not a differential operator but merely a multiplicative one. It multiplies any amplitude $\langle\vec{r}|\psi(t)\rangle$ by the number $V(\vec{r})$. Equation 18 therefore reads explicitly

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\right]\langle\vec{r}|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}\langle\vec{r}|\psi(t)\rangle \tag{19}$$

Equation (19), the nonrelativistic Schrodinger equation for a particle of mass m, is the basic dynamical law of quantum mechanics.* It has contributed to the complete or partial under-

---

* Equation 19 is the Schrodinger equation in the r-representation. The equation can also be written in the momentum representation, generalizing Eq. 15. However, including the potential is not so straight-forward a step in this case. A potential $V(\vec{r})$, which is merely a multiplicative factor in the position representation, has a more complicated form in momentum space: In the momentum representation the Schrodinger equation for the above potential is an integro-differential equation. We shall not develop the momentum space formulation any further here.

standing of a vast body of phenomena on the atomic and sub-atomic levels. Like Newton's law, the Schrodinger equation cannot be "derived"; we have given here a plausibility argument similar to the one which led Schrodinger to the equation. One postulates the equation, and tests the validity of the postulate by comparing the predictions of the theory with experiment. This program is begun in the next chapter.

As written in the form (19), the Schrodinger equation contains only spatial amplitudes $\langle\vec{r}|\psi\rangle$. Therefore it does not include a description of particle spin. In section 5 below we shall show that the spin amplitudes studied in Chapters 7 and 8 also satisfy a very similar differential equation.

The most important limitation on the applicability of the Schrodinger equation is that it is not relativistically invariant, being based on the energy-momentum relation $E = p^2/2m$. Therefore one can apply it only to problems in which particle velocities are small compared to c. The electron velocity in the hydrogen atom is of the order 0.01c. The constants of hydrogen have been measured so accurately that the relativistic effects, though small, cannot be entirely neglected. For the "inner orbits" of atoms farther along in the periodic table, these corrections become large.

Because the Schrodinger equation is nonrelativistic, it cannot describe the creation or annihilation of particles, phenomena which are intrinsically relativistic. The equation is readily generalized to include systems of two or more particles, but the number of particles is a "constant of the motion."*

Schrodinger himself proposed, at almost the same time as his non-relativistic equation, a generalization that is relativistically invariant. However, this could not be successfully interpreted as a single-particle theory. A satisfactory relativistic theory was obtained somewhat later by Dirac.

Much of quantum mechanics consists of the study of the solutions of Eq. 19 for different potentials $V(\vec{r})$ and different initial conditions. We postulate that every quantum state of a system is a solution of the equation. This does not imply that every solution of the equation represents a possible physical state. In order to be physically admissible, the solution must satisfy certain conditions. For example, the state vector for a bound state must be normalizable, which requires that the amplitude $\langle\vec{r}|\psi\rangle$ be square-integrable. Other boundary conditions of continuity may be imposed by the geometry of a particular problem. Some of these are discussed in the exercises and in the following chapter.

---

* Some problems involving the emission and absorption of photons can be handled by a semiclassical treatment which, although not rigorous, leads to relatively accurate descriptions of the phenomena.

4. The time-independent Schrodinger equation.

An important class of problems to which the Schrodinger equation is applied is that of finding the stationary states of various systems. Such states are characterized by a definite value for the total energy E and accordingly carry the time dependence $e^{-iEt/\hbar}$. We look for solutions for Eq. 19 that have this simple time dependence. That is, we put

$$|\psi(t)\rangle = |\psi(0)\rangle e^{-iEt/\hbar} \tag{20}$$

For the sake of brevity we shall write $|\psi(0)\rangle$ merely as $|\psi\rangle$, wherever this leads to no confusion. With the substitution (20), Eq. 19 becomes

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r})\right]\langle\vec{r}|\psi\rangle = E\langle\vec{r}|\psi\rangle \tag{21}$$

Equation 21 is called the time-independent Schrodinger equation. This is the form most useful in many applications, particularly when one is interested in the bound states of atoms or other systems. In such problems, one generally finds that for an arbitrary value of E no solution of Eq. 21 exists that satisfies the physically demanded boundary conditions. Those discrete values of E for which physically admissible solutions do exist constitute the "allowed" energy levels (stationary states) of the system. As we have already seen (Chapter 8), a system need not be in a stationary state. However, the stationary states constitute a complete orthogonal set, and generally provide the most convenient basis for the description of an arbitrary state. Let $|\psi_n(t)\rangle$ denote the state vector for the nth stationary state, and E_n the corresponding energy. Then we may expand an arbitrary state at time t = 0 in the form

$$|\psi(0)\rangle = \sum_n |\psi_n\rangle a_n \tag{22}$$

where

$$a_n = \langle\psi_n|\psi(0)\rangle \tag{23}$$

The time evolution of the state $|\psi\rangle$ can then be written directly:

$$|\psi(t)\rangle = \sum_n |\psi_n\rangle a_n e^{-iE_n t/\hbar} \tag{24}$$

One is not obliged to expand in terms of stationary states. For example, the momentum states $|\vec{k}\rangle$, which are stationary only for a free particle, are nonetheless an acceptable set of basis states in terms of which a state vector at a given time may be expanded. However, the time dependence of a momentum state in the presence of interaction is complicated, and such an expansion will not easily tell us how the state evolves with time.

5. Schrodinger equation for spin $\frac{1}{2}$ particle in a magnetic field.

We show in this section that the quantum states of a spin $\frac{1}{2}$ particle in a uniform magnetic field, discussed in Chapter 8, satisfy a differential equation which is similar in form to the Schrodinger equation for spatial states developed earlier in this chapter. The equation for the present case is written naturally in matrix form and resembles most closely the momentum space Schrodinger equation, Eq. 15.

Let a magnetic field $\vec{B}$ point in the z' direction, defined by the polar angles θ, ϕ; and let

$$|m'\rangle \qquad m' = \pm\frac{1}{2}$$

denote the set of basis states $|\pm z\rangle$. These are the stationary states for the problem, with energies

$$E_{m'} = \hbar(m'\omega_p) \tag{25}$$

where $\omega_p = g(eB/2mc)$ is the classical precession frequency (Eq. 1 of Chapter 8). The stationary states have the time dependence $e^{-iE_{m'}t/h}$; for an arbitrary state $|\psi(t)\rangle$, the projection amplitudes in the z' representation have therefore a simple time dependence:

$$\langle m'|\psi(t)\rangle = \langle m'|\psi(0)\rangle e^{-iE_{m'}t/\hbar} \tag{26}$$

We obtain a differential equation satisfied by these amplitudes by taking the time derivative of Eq. 26

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\langle m'|\psi(t)\rangle &= E_{m'}\langle m'|\psi(t)\rangle \\ &= \hbar\omega_p m'\langle m'|\psi(t)\rangle \end{aligned} \tag{27}$$

The two equations contained in Eq. 27, (one for each value of m') can be written as a matrix equation:

$$ih\frac{\partial}{\partial t}\begin{pmatrix}\langle +z'|\psi(t)\rangle \\ \langle -z'|\psi(t)\rangle\end{pmatrix} = \frac{1}{2}\hbar\omega_p\begin{pmatrix}1 & 0 \\ 0 & -1\end{pmatrix}\begin{pmatrix}\langle +z'|\psi(t)\rangle \\ \langle -z'|\psi(t)\rangle\end{pmatrix} \tag{28}$$

The matrix that appears on the right side of Eq. 28 is one of the Pauli matrices introduced in Chapter 7; in fact, it is just the matrix we have called $\sigma_{z'} = \vec{\sigma}\cdot\hat{z}'$. (Eq. 34 of Chapter 7). If we want to write the same equation in the z representation, (this will give a differential equation for the amplitudes $\langle\pm z|\psi(t)\rangle$), the transformed matrix of $\sigma_{z'}$ is given by Eq. 36 of Chapter 7. The equation is then

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\langle +z|\psi(t)\rangle \\ \langle -z|\psi(t)\rangle\end{pmatrix} = \frac{1}{2}\hbar\omega_p\begin{pmatrix}\cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta\end{pmatrix}\begin{pmatrix}\langle +z|\psi(t)\rangle \\ \langle -z|\psi(t)\rangle\end{pmatrix} \tag{29}$$

Note that the pair of equations 29 is coupled (the time derivative of $\langle +z|\psi(t)\rangle$ involves both $\langle +z|\psi(t)\rangle$ and $\langle -z|\psi(t)\rangle$), whereas the corresponding pair of equations in the z'-representation (Eq. 28) is uncoupled. We can write a formal equation for the time derivative of the state vector $|\psi(t)\rangle$, valid in all representations, as follows:

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \frac{1}{2}\hbar\omega_p\frac{\vec{\sigma}\cdot\vec{B}}{B}|\psi(t)\rangle \tag{30}$$

This includes both (28) and (29). Here we have written $\sigma_{z'}$ as $\vec{\sigma}\cdot\vec{B}/B$, since the z' direction was defined as the direction of $\vec{B}$. We call the operator $\frac{1}{2}\hbar\omega_p(\vec{\sigma}\cdot\vec{B})/B$ that appears in

Eq. 30 the Hamiltonian operator for this system and write the equation simply as

$$i\hbar \frac{\partial}{\partial t} |\psi(t)> = H|\psi(t)> \tag{31}$$

Note that $\pm\frac{1}{2}\hbar\omega_p = \pm\frac{1}{2}\hbar g(eB/2mc)$ are just the energies of a classical dipole of angular momentum $\hbar/2$ and magnetic moment $\mu = \frac{1}{2} g\mu_B$ aligned antiparallel or parallel to the magnetic field $\vec{B}$ (here $\mu_B = e\hbar/(2mc)$). Hence the operator H in Eq. 30 may be interpreted as the potential energy operator for the system in the following sense: The matrix element $<+z'|H|+z'>$ measures the potential energy of the particle when it is in the state $|+z'>$. Similarly the matrix element $<-z'|H|-z'>$ measures the potential energy of the particle in state $|-z'>$. In any other state the potential energy, according to our interpretation, does not have a "sharp" value; that is, a measurement of the energy does not always lead to the same result. However, for any state $|\psi>$, the matrix element $<\psi|H|\psi>$ measures the expectation value of the potential energy, that is, the average of many energy measurements.* The matrix element $<\psi|H|\psi>$ can be evaluated from the equation

$$<\psi|H|\psi> = \sum_{m_1', m_2'} <\psi|m_1'><m_1'|H|m_2'><m_2'|\psi> \tag{32}$$

The interpretation of Eq. 31 presented here, with H interpreted as an energy operator, leads to a broad generalization. The equation was derived for the special case of a uniform field. The first generalization is to suppose that Eq. 30 describes the quantum states of a spin 1/2 particle in an arbitrary magnetic field. Going even beyond this, we conjecture that under very general conditions the quantum states of a system obey Eq. 31, with H an operator that represents the total energy of the system in the sense discussed above. The conjecture is valid for a wide variety of physical systems. The precise form of the Hamiltonian depends on the nature of the problem: the number and kinds of particles involved, and their interactions with each other as well as with external fields. Often one can obtain the form of the Hamiltonian from a classical expression for the energy; this is the case in the examples mentioned so far. However, even in cases where no classical analogue exists (for example, problems involving the creation and/or absorption of particles) the Schrodinger equation in the form (31) is found to have validity. In such an analysis the first problem is just to discover the form of the Hamiltonian. One tries various possibilities, guided by intuition and by general requirements such as relativistic invariance. One then calculates the predictions of Eq. 31 with the particular Hamiltonian under study, and compares the predictions with experiment.

---

* In an analogous manner one interprets $\vec{S} = 1/2\,\hbar\vec{\sigma}$ as the operator for (spin) angular momentum and $-\mu\vec{\sigma}$ as the operator for magnetic moment.

Chapter 12. The Hydrogen Atom: States with Spherical Symmetry

1. Introduction

In this chapter and the next we use the Schrodinger equation to study the stationary states of the hydrogen atom. We shall find that the physically acceptable solutions of the time-independent Schrodinger equation are epitomized by three "quantum numbers" n, ℓ, and m, which are restricted to integer values. These quantum numbers can be used to label the stationary states. Their values for a given state determine respectively the energy, the angular momentum and the z-component of angular momentum for the state in question. Heretofore we have defined the quantum states of systems empirically by means of projectors. In the present discussion, for the first time, the very definition of the quantum state emerges from the solution of a dynamical equation. The good agreement with experiment constitutes strong support for the validity of the Schrodinger equation, postulated in the last chapter on a largely speculative basis.

Of all the atoms in the periodic table, hydrogen is the only one for which analytic solutions are possible. Direct numerical integrations of the Schrodinger equation have been carried out for the ground state of helium. The other atoms in the periodic table are analyzed using approximations based on the hydrogen-like amplitudes. It is this general usefulness of the hydrogen atom solutions that makes the analysis of the hydrogen atom central to the understanding we have of the entire periodic table of elements. The exact hydrogen atom solutions can also be applied with only minor modification to other two-particle systems: singly ionized helium (helium nucleus and an electron), positronium (positive electron and negative electron bound together), muonium (positive mu-meson and negative mu-meson bound together), mu-, pi-, and K-mesic atoms (in which the electron in hydrogen is replaced with a negative mu-meson, negative pi-meson, or negative K-meson respectively).

The present chapter is concerned with the class of solutions of the Schrodinger equation that is analytically most simple, namely those solutions for which the spatial probability distribution is spherically symmetric. The most general solution is developed in Chapter 13.

2. Schrodinger equation for a two-particle system.

The hydrogen atom, consisting of one electron and one proton, is the simplest atomic system. Because the electron is so much lighter than the proton,* a quite satisfactory approximate solution of the problem may be obtained by treating the proton as fixed and considering only the motion of the electron. In such an approach one simply solves the time-independent Schrodinger

---

* Ratio of masses $m_p/m_e = 1836$.

equation, Eq. 21 of Chapter 11, with V(r) the potential energy of the electron in the Coulomb field of the fixed proton. If we put the origin at the position of the proton, the potential depends only on the magnitude of the electron coordinate r, and the equation becomes

$$\left[-\frac{\hbar^2}{2m_e}\nabla^2 - \frac{e^2}{r}\right] \langle\vec{r}|\psi\rangle = E\langle\vec{r}|\psi\rangle. \tag{1}$$

If one does not neglect the motion of the proton, one must start with an equation more complicated than (1), an equation that depends on the coordinates of both particles. However, this two-body equation is readily reduced to an "equivalent" one-body equation; moreover, this reduction illustrates the technique of "separation of variables," which is required later in order to obtain a complete solution of the one-body equation. Therefore we shall begin with the full two-body problem. This requires a generalization of the Schrodinger equation. Let $|\psi\rangle$ now denote a state of a system consisting of one electron and one proton.* A position measurement of this system requires the determination of the location of each particle; hence the position basis states may be written in an obvious notation as $|\vec{r}_e, \vec{r}_p\rangle$. Similarly the momentum basis states may be written $|\vec{k}_e, \vec{k}_p\rangle$, and the general state $|\psi\rangle$ is described either by its position amplitudes $\langle\vec{r}_e, \vec{r}_p|\psi\rangle$ or by the momentum amplitudes $\langle\vec{k}_e, \vec{k}_p|\psi\rangle$. These amplitudes are subject to the usual interpretation: $|\langle\vec{r}_e, \vec{r}_p|\psi\rangle|^2\, d^3r_e\, d^3r_p$ measures the probability that the electron will be detected in a position measurement in the volume element d^3r_e near $\vec{r}_e$, and simultaneously the proton will be detected in the volume element d^3r_p near $\vec{r}_p$.

Suppose for the moment that both electron and proton are free particles. (This requires that we "turn off" the Coulomb force between them.) Then the probability of detecting the electron at some point must be independent of any measurements carried out on the proton, and vice versa. This requires that the full amplitude $\langle\vec{r}_e, \vec{r}_p|\psi\rangle$ be a product of factors, each referring to one of the particles:

$$\langle\vec{r}_e, \vec{r}_p|\psi\rangle = \langle\vec{r}_e|\psi_e\rangle\langle\vec{r}_p|\psi_p\rangle \qquad \text{free-particle states} \tag{2}$$

Moreover, under the assumed conditions (no interaction), both $\langle\vec{r}_e|\psi_e\rangle$ and $\langle\vec{r}_p|\psi_p\rangle$ must be solutions of the free-particle Schrodinger equation:

$$-\frac{\hbar^2}{2m_e}\nabla_e^2\langle\vec{r}_e|\psi_e\rangle = i\hbar\frac{\partial}{\partial t}\langle\vec{r}_e|\psi_e\rangle \tag{3a}$$

$$-\frac{\hbar^2}{2m_p}\nabla_p^2\langle\vec{r}_p|\psi_p\rangle = i\hbar\frac{\partial}{\partial t}\langle\vec{r}_p|\psi_p\rangle \tag{3b}$$

We have written subscripts on the Laplacians to indicate explicitly that the differentiations are with respect to the coordinates of the corresponding particle. By virtue of Eqs. 2 and 3 the two-body free-particle amplitude $\langle\vec{r}_e, \vec{r}_p|\psi\rangle$ satisfies the equation:

---

* Cf. the discussion of two-photon states in Chapter 6 - Appendix.

$$\left(-\frac{\hbar^2}{2m_e}\nabla_e^2 - \frac{h^2}{2m_p}\nabla_p^2\right)\langle\vec{r}_e,\vec{r}_p|\psi\rangle = i\hbar\frac{\partial}{\partial t}\langle\vec{r}_e,\vec{r}_p|\psi\rangle \tag{4}$$

free-particle states

This may be verified directly by carrying out the indicated differentiations.

The generalization of Eq. 4 to include interactions is the same as the one we have already indicated for the single-particle problem. Letting V_e and V_p denote any potential energies of electron and proton respectively due to external causes, and V_{ep} denote the mutual potential energy of the two particles, one writes:

$$\left[-\frac{\hbar^2}{2m_e}\nabla_e^2 - \frac{\hbar^2}{2m_p}\nabla_p^2 + V_e + V_p + V_{ep}\right]\langle\vec{r}_e,\vec{r}_p|\psi\rangle = i\hbar\frac{\partial}{\partial t}\langle\vec{r}_e,\vec{r}_p|\psi\rangle \tag{5}$$

The further generalization of Eq. 5 to any number of particles is clearcut: one merely adds the appropriate Laplacians and includes all potential energies.

We consider the problem of the hydrogen atom, with no external potentials; hence V_e and V_p are absent, and V_{ep} is the Coulomb energy.* Thus Eq. (5) becomes

$$\left[-\frac{\hbar^2}{2m_e}\nabla_e^2 - \frac{\hbar^2}{2m_p}\nabla_p^2 - \frac{e^2}{|\vec{r}_e - \vec{r}_p|}\right]\langle\vec{r}_e,\vec{r}_p|\psi\rangle = i\hbar\frac{\partial}{\partial t}\langle\vec{r}_e,\vec{r}_p|\psi\rangle \tag{6}$$

The corresponding time-independent equation for the stationary states is (cf. Eq. 21 of Chapter 11):

$$\left[-\frac{\hbar^2}{2m_e}\nabla_e^2 - \frac{\hbar^2}{2m_p}\nabla_p^2 - \frac{e^2}{|\vec{r}_e - \vec{r}_p|}\right]\langle\vec{r}_e,\vec{r}_p|\psi\rangle = E\langle\vec{r}_e,\vec{r}_p|\psi\rangle \tag{7}$$

where E is the total energy of the state. Equation 7 is the one whose solutions we must study. The presence of the interaction term, a function of

$$|\vec{r}_e - \vec{r}_p| = (r_e^2 + r_p^2 - 2r_e r_p \cos\theta_{ep})^{1/2} \tag{8}$$

prevents us from separating (7) into independent equations in $\vec{r}_e$ and $\vec{r}_p$. However, a simple change of variables makes a separation possible, as we show in the next section.

3. Reduction of the two-body equation.

The Schrodinger equation (7) for the hydrogen atom is complicated because the potential term is a complicated function of the coordinates in which the problem has been formulated. A similar difficulty is encountered in a classical treatment of the two-body problem with central forces. The way out of the difficulty in both the quantum and classical analysis is to make a change of variables such that the interparticle vector becomes one of the coordinates. The other coordinate is the center-of-mass of the system. That is, one makes the change of independent

---

* We ignore, at this point, such energies as that of the spin of one particle in the magnetic field due to the motion of the other. These are small effects.

variables from the set $\vec{r}_e$ and $\vec{r}_p$ to the set

$$\vec{r} = \vec{r}_e - \vec{r}_p \tag{9}$$

and

$$\vec{R} = \frac{m_p\vec{r}_p + m_e\vec{r}_e}{m_e + m_p} \tag{10}$$

In terms of these variables the classical energy of the system,

$$E = \frac{p_e^2}{2m_e} + \frac{p_p^2}{2m_p} + \frac{e^2}{|r_e - r_p|} \tag{11}$$

takes the form

$$E = \frac{P^2}{2M} + \frac{p^2}{2\mu} + \frac{e^2}{r} \tag{12}$$

where

$$M = m_e + m_p \tag{13a}$$

is the total mass of the system, and

$$\mu = \frac{m_e m_p}{m_e + m_p} \tag{13b}$$

is the <u>reduced</u> <u>mass</u> (in the present case practically equal to the electron mass.) The quantity

$$\vec{P} = M\dot{\vec{R}} = \vec{p}_e + \vec{p}_p \tag{14}$$

is the total momentum of the system, while

$$\vec{p} = \mu\dot{\vec{r}} = \frac{m_p\vec{p}_e - m_e\vec{p}_p}{M} \tag{15}$$

is the "relative momentum," i.e., the momentum of a fictitious particle of mass μ, moving with the relative velocity $\dot{\vec{r}}$. Likewise, the combination of Laplacians that appears in the Schrodinger equation (7) takes the following form when expressed in terms of the new coordinates $\vec{r}$ and $\vec{R}$:

$$-\frac{\hbar^2}{2m_e}\nabla_e^2 - \frac{\hbar^2}{2m_p}\nabla_p^2 = -\frac{\hbar^2}{2M}\nabla_R^2 - \frac{\hbar^2}{2\mu}\nabla_r^2 \tag{16}$$

where ∇_R^2 and ∇_r^2 are Laplacians for the coordinates $\vec{R}$ and $\vec{r}$ respectively. The identity (16) can either be considered to follow from (11) and (12) by translating a classical momentum into a differential operator, or else the identity can be demonstrated directly (if somewhat tediously) by carrying out the transformation (9) and (10). The amplitude $\langle\vec{r}_e, \vec{r}_p|\psi\rangle$ can be considered a function of $\vec{r}$ and $\vec{R}$. We shall write this amplitude as $\Psi(\vec{r}, \vec{R})$. Thus Eq. 7 now becomes:

$$\left(-\frac{\hbar^2}{2M}\nabla_R^2 - \frac{\hbar^2}{2\mu}\nabla_r^2 - \frac{e^2}{r}\right)\Psi(\vec{r},\vec{R}) = E\,\Psi(\vec{r},\vec{R}) \tag{17}$$

The fact that the potential term in Eq. 17 depends only on one coordinate makes possible a separation of the equation. We look for solutions which are products of a function of $\vec{r}$ alone and another function of $\vec{R}$ alone:

$$\Psi(\vec{r},\vec{R}) = \psi(\vec{r})\phi(\vec{R}) \tag{18}$$

The standard technique for effecting a separation of variables is to substitute the form (18) into (17), keeping in mind the fact that each Laplacian acts only on its coordinates, and then divide both sides of the equation by $\Psi(r,R)$. In this way one obtains the equation

$$\left[-\frac{1}{\psi(\vec{r})}\frac{\hbar^2}{2\mu}\nabla_r^2\psi(\vec{r}) + V(r)\right] + \left[-\frac{1}{\phi(\vec{R})}\frac{\hbar^2}{2M}\nabla_R^2\phi(\vec{R})\right] = E \tag{19}$$

The left side of Eq. 19 is the sum of two terms, the first of which is a function of $\vec{r}$ only, and the second a function of $\vec{R}$ only. Yet their sum is the constant E. This can be true only if each "function" is a constant.* We have, therefore,

$$-\frac{\hbar^2}{2M}\nabla_R^2\phi(\vec{R}) = E_o\phi(\vec{R}) \tag{20}$$

$$\left(-\frac{\hbar^2}{2\mu}\nabla_r^2 + V(r)\right)\psi(\vec{r}) = (E - E_o)\psi(\vec{r}) \ , \tag{21}$$

where E_o is a constant whose significance is immediately clear: Eq. 20 is a Schrodinger equation for a single free particle, whose mass is the total mass of the electron-proton system, and whose motion describes the center of mass of the system. The solutions of this equation have already been studied: they consist of "plane-wave" states (i.e., single-momentum states), described by the amplitude

$$\phi_{\vec{K}}(\vec{R}) = e^{i\vec{K}\cdot\vec{R}} \tag{22a}$$

where $\vec{K}$ is any vector whose magnitude is

$$K = \left(\frac{2mE_o}{\hbar^2}\right)^{1/2} \tag{22b}.$$

(A superposition of such states also satisfies Eq. 21.) Thus the "separation constant" E_o is the energy associated with the motion of the center of mass, which behaves like a free particle; this is to be expected since there are no external forces acting on the system. The remaining energy $E - E_o$ can be called E_{rel}, the relative energy.

The center-of-mass equation, (20), is of no further interest and we henceforth drop the factor $\phi(\vec{R})$. Equation 21 is the one that contains the potential and on whose solution we focus our attention. A separation of this type is evidently possible for a two-body system whenever the potential energy depends only on the relative coordinate between the particles. Equation 21 has the form of a one-body Schrodinger equation, for a particle of mass μ, described by the coordinate $\vec{r}$. Since μ is practically equal to m_e, and $\vec{r}$ is practically the same as $\vec{r}_e$, the equation differs little from Eq. 1. But, as already remarked, the procedure we have

---

* If $f(\vec{r}) + g(\vec{R}) = E$, then $f(\vec{r}_1) + g(\vec{R}) = f(\vec{r}_2) + g(\vec{R})$, and $f(\vec{r}_1) = f(\vec{r}_2)$ for any $\vec{r}_1$, $\vec{r}_2$. Similarly $g(\vec{R}_1) = g(\vec{R}_2)$ for any $\vec{R}_1$, $\vec{R}_2$.

followed in obtaining (21) from the full two-body equation is instructive and will find repeated use later.

4. Spherically symmetric solutions for the hydrogen atom.

We rewrite Eq. 21, dropping all subscripts for simplicity:

$$\left(-\frac{\hbar^2}{2\mu}\nabla^2 - \frac{e^2}{r}\right)\psi(\vec{r}) = E\psi(\vec{r}) \tag{22}$$

We are interested in solutions of this equation that describe bound states of the system. In such a solution the energy E is expected to be negative; furthermore, we must require that the solution be normalizable, i.e., that the integral

$$\int |\psi(\vec{r})|^2 \, d^3r \tag{23}$$

be finite. This implies, in particular, that

$$\psi(\vec{r}) \longrightarrow 0, \qquad r \longrightarrow \infty \tag{24}$$

Other boundary conditions will appear later on. We shall find that solutions which satisfy the boundary conditions do not exist for every value of E; solutions exist only for a particular set of "energy eigenvalues."*

Equation (22) is still a three-dimensional partial differential equation. Since the potential depends only on the scalar r, it is most convenient to formulate the solution in spherical coordinates. The potential is then independent of θ and ϕ, and the equation can be separated into three one-dimensional equations. This procedure is carried out in the following chapter. Here we shall look for a particularly simple type of solution, namely one in which $\psi(\vec{r})$ is itself independent of θ and ϕ. Not only do such solutions exist; but, it turns out, the energies of this restricted class of spherically symmetric solutions already gives the complete set of allowed energies. That is, the energy of each of the non-spherically symmetric states, to be derived in Chapter 13, coincides with the energy of one of the spherically symmetric solutions. (This "degeneracy" is peculiar to the Coulomb potential.) Accordingly, we look for solutions of Eq. 22 in which ψ depends only on the distance r.

The Laplacian in spherical coordinates has a complicated appearance. However, for the solutions in which we are presently interested we require only those terms that include no derivatives with respect to θ or ϕ. In Box 1 it is shown that

$$\nabla^2\psi = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) + \left(\begin{array}{l}\text{terms that involve deriva-}\\ \text{tives with respect to } \theta \text{ and } \phi\end{array}\right) \tag{25}$$

---

* It should be remarked that physically interesting solutions of Eq. 22 exist also for positive energies. These solutions are "unbound" states and describe electron-proton scattering. In a scattering state (which is properly described only by the time-dependent Schrodinger equation), the boundary condition (24) does not apply, since before and after the scattering the particles are found far apart. As a consequence, scattering solutions exist for any (positive) value of the energy.

Box 1. The Laplacian for Spherically Symmetric Functions

Assume $\psi(r)$ is a function of the spherical coordinate r only, and not a function of the angle θ or the angle ϕ. We want to write the Laplacian of this function. Any derivatives with respect to θ or ϕ will yield zero when applied to $\psi(r)$. Recall that in Cartesian coordinates the Laplacian is

$$\nabla^2\psi = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}$$

and that

$$r = (x^2 + y^2 + z^2)^{1/2}$$

Ignoring derivatives in θ and ϕ we have

$$\frac{\partial\psi}{\partial x} = \frac{\partial r}{\partial x}\frac{\partial\psi}{\partial r} = \frac{x}{(x^2 + y^2 + z^2)^{1/2}}\frac{\partial\psi}{\partial r} = \frac{x}{r}\frac{\partial\psi}{\partial r}$$

and

$$\frac{\partial^2\psi}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{x}{r}\frac{\partial}{\partial r}\right) = \left(\frac{\partial x}{\partial x}\right)\left(\frac{1}{r}\frac{\partial\psi}{\partial r}\right) + x\frac{\partial}{\partial x}\left(\frac{1}{r}\frac{\partial\psi}{\partial r}\right) =$$

$$= \frac{1}{r}\frac{\partial}{\partial r} + \frac{x^2}{r}\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial}{\partial r}\right)$$

By an identical argument

$$\frac{\partial^2\psi}{\partial y^2} = \frac{1}{r}\frac{\partial}{\partial r} + \frac{y^2}{r}\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\psi}{\partial r}\right)$$

and

$$\frac{\partial^2\psi}{\partial z^2} = \frac{1}{r}\frac{\partial}{\partial r} + \frac{z^2}{r}\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\psi}{\partial r}\right)$$

Add these expressions together to obtain the Laplacian in spherical coordinates for spherically symmetric functions

$$\nabla^2\psi \longrightarrow \frac{3}{r}\frac{\partial}{\partial r} + \frac{(x^2 + y^2 + z^2)}{r}\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\psi}{\partial r}\right) \tag{26}$$

$$= \frac{3}{r}\frac{\partial}{\partial r} + r\frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial\psi}{\partial r}\right)$$

which can be written as

$$\nabla^2\psi = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) \qquad \text{Spherically symmetric } \psi \tag{27}$$

Hence, on multiplying 25 by r one obtains the following equation for the function $u(r) = r\psi(r)$;

$$-\frac{\hbar^2}{2\mu}\frac{d^2u}{dr^2} - \frac{e^2}{r}u = E\,u \tag{28}$$

Finally, we can avoid carrying along all the constants by expressing (28) in dimensionless units. This change of variables also puts into evidence the natural scales of distance and energy for the problem. Let

$$r = \frac{\hbar^2}{\mu e^2}\rho \tag{29}$$

and

$$E = \frac{\mu e^4}{2\hbar^2}\,\epsilon \tag{30}$$

Then equation (28) becomes simply:

$$\frac{d^2u}{d\rho^2} + (\epsilon + \frac{2}{\rho})\,u = 0 \tag{31}$$

The substitutions (29) and (30) mean that we are measuring distance in units of

$$a_o = \hbar^2/\mu e^2 \tag{32}$$

and energy in units of

$$E_R = \mu e^4/\,2\hbar^2 \tag{33}$$

a_o is called the Bohr radius, and E_R is called the Rydberg energy.

Numerically, $a_o = 5.28 \times 10^{-9}$ cm

and $E_R = 13.6$ electron volts.

These units arise naturally as combinations of the constants in the original Schrodinger equation (e, $\hbar$, and μ) that have the dimensions of length and of energy. Even before we have solved any equation, the appearance of these units is an encouraging sign because these are indeed the units which are appropriate for the hydrogen atom.

a. Ground State. We seek those solutions of Eq. 21 which, according to the boundary condition (24), vanish as $\rho \to \infty$: ϵ is expected to be negative. As we have already remarked, solutions exist only for special values of ϵ. Before describing a systematic search for the solutions, we exhibit the simplest solution by trial and error. Notice that without the potential term $2/\rho$, the solutions for negative ϵ would be

$$u(\rho) = e^{\pm|\epsilon|^{1/2}} \tag{34}$$

The negative exponential is of course the one that satisfies the boundary condition. Since the potential term goes to zero for large ρ (albeit quite slowly), if a solution exists for some ϵ it should resemble (34) for large ρ. The following function can be verified directly to be a solution for $\epsilon = -1$.

$$u_1(\rho) = \rho e^{-\rho}, \qquad \epsilon = -1 \tag{35}$$

The more general analysis to follow indicates that no solution exists for any value of ϵ less than -1. Hence the state described by (35) is the state of lowest energy, called the ground state of the hydrogen atom. Its energy is minus one Rydberg, -13.6 e.V. In terms of usual units of length, the function $\psi(r) = u(r)/r$ is

$$\psi_1(r) = e^{-r/a_o} \qquad \text{(not normalized)} \tag{36}$$

The spatial probability distribution for the ground state of the hydrogen atom is given by the square of (36). Specifically, the probability of detecting the relative coordinate within a spherical shell of inner radius r and outer radius r + dr is

$$P(r) = |\psi(r)|^2 (4\pi r^2 dr) \propto r^2 e^{-2r/a_o} \tag{37}$$

Notice that because of the factor r^2 that comes from the volume element in spherical coordinates, the square of the function u directly measures the radial probability distribution. (That is, $u^2 dr$ is proportional to the probability of finding the electron at a radius between r and r + dr.)

b. General solution. A general analytic solution of Eq. 31 can be obtained by the method of power series:* the analysis shows that solutions that satisfy the boundary condition (24) exist for values of ϵ given by the so-called Rydberg formula

$$\epsilon = -1/n^2 \qquad n = 1, 2, 3, \ldots . \tag{38}$$

The binding energy of the nth state is E_R/n^2 and approaches zero for large n. The corresponding radial distributions are expressible in terms of functions called Laguerre functions. The first few solutions are:

$$\begin{aligned} n = 2:& \quad u_2(\rho) = e^{-\frac{1}{2}\rho}\left(1 - \frac{\rho}{2}\right)\rho & \epsilon = -\frac{1}{4} \\ n = 3:& \quad u_3(\rho) = e^{-\frac{1}{3}\rho}\left(1 - \frac{2}{3}\rho + \frac{2}{27}\rho^2\right)\rho & \epsilon = -\frac{1}{9} \\ n = 4:& \quad u_4(\rho) = e^{-\frac{1}{4}\rho}\left(1 - \frac{3}{4}\rho + \frac{1}{8}\rho^2 - \frac{1}{192}\rho^3\right)\rho & \epsilon = -\frac{1}{16} \end{aligned} \tag{39}$$

The radial probability distribution (u^2) for these states are plotted in Figure 1.

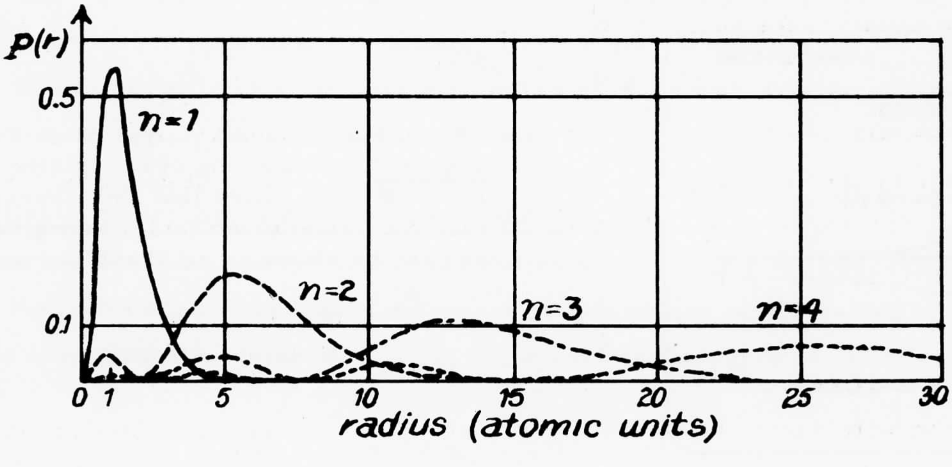
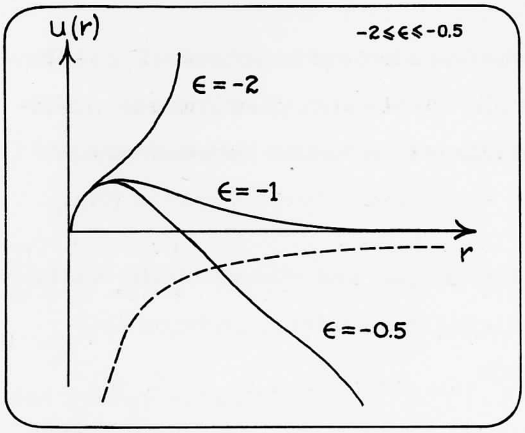


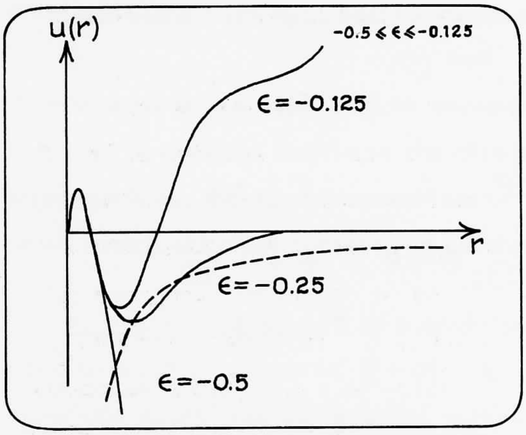
Figure 1. Relative probability of finding an electron at different distances from the proton in the first four spherically symmetric stationary states of hydrogen.

---

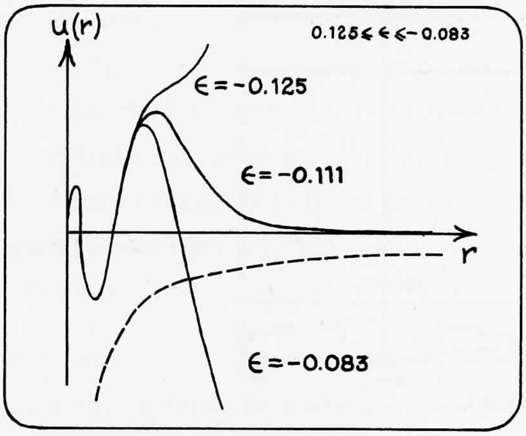
* See, for example, The Feynman Lectures on Physics, Addison-Wesley Publishing Company, Volume 3, Section 2 of Chapter 19.



largest radius shown: $8a_o$



largest radius shown: $32a_o$



largest radius shown: $60a_o$

Figure 2. Search for physically acceptable spherically symmetric solutions to the hydrogen atom. Acceptable solutions are found for $\epsilon = -1, -1/4, -1/9$. Dashed line in each figure represents the Coulomb potential. Notice the changes in horizontal scale. Computer outputs shown on this page courtesy of Dr. Judah Schwartz.

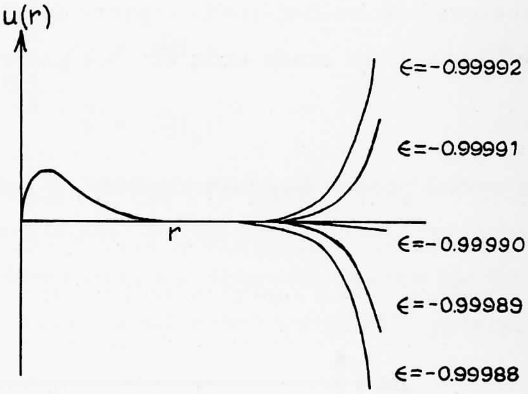


Figure 3. Straddling of acceptable solution near $\epsilon = -1$. Note that this approximation method yields approximate energies. Largest radius shown in all these curves: 16 a_o.

We shall not describe here the power series solution which leads to the results (38), (39). Instead, we indicate a numerical approach which can be used to obtain energy levels and amplitudes even when no analytic solution exists. (The hydrogen atom is one of the few quantum mechanical problems which has a solution in closed form.) The standard method for solving numerically a differential equation such as (31) is to approximate it by a difference equation, in which the function is calculated on a finite set of "mesh points." The difference equation amounts to a set of coupled linear equations, which express the value of the function at the (j + 1)st meshpoint in terms of its value at the jth and (j-1)st points. For any assigned value of ϵ the difference equation may be solved. To start the solution, the values of u at two adjacent netpoints are specified; this corresponds to specifying u(r) and du/dr at a single point in the differential equation. The procedure lends itself readily to machine computation. For details of the process see the appendix to this chapter.

For the present problem, since $u = r\psi$ the requirement that the amplitude $\psi(r)$ be everywhere finite implies that $u(0) = 0$.* Hence it is convenient to start the solution at $r = 0$. The value of u at the first net point is arbitrary; it merely fixes the scale of the solution. Thus the solution for a given ϵ is essentially unique. For most values of ϵ, the solution blows up as $r \to \infty$; such solutions must be rejected because they do not satisfy the physical boundary condition (24). As one systematically changes the value of ϵ, one finds every so often a solution which goes to zero for large r. These are the physically acceptable solutions; the amplitudes and energies are, within the accuracy of the numerical method, just those given by Eqs. 38 and 39. Of course, the numerical method does not yield analytic forms for the solutions, but rather direct plots of the functions u(r). Figure 2 exhibits typical outputs from the computer. Shown are solutions for "wrong" values of ϵ, which diverge at large r, and solutions which do satisfy the boundary conditions.

5. Properties of the spherically symmetric solutions.

We examine here some properties of the solutions obtained in the preceding section. First, notice that the probability distribution moves out to larger and larger values of r with increasing n. The most probable value of r (i.e., the peak in the curve of $u^2(r)$ vs r (Fig. 1):

$$\begin{aligned} n = 1: &\quad r_m = a_0 \\ n = 2: &\quad r_m = 5.2\, a_0 \\ n = 3: &\quad r_m \approx 13\, a_0 \\ n = 4: &\quad r_m \approx 26\, a_0 \end{aligned} \tag{40}$$

---

* The condition that $\psi(r)$ be everywhere finite is a bit stronger than is demanded by the probability interpretation. The latter requires only that $\int |\psi(\vec{r})|^2 d^3r$ be finite, which is satisfied for example, if $\psi(\vec{r})$ behaves like $1/r$ near the origin. However, it can be shown that no functions of this type satisfy the differential equation at the origin. Hence one is justified in imposing the boundary condition $u(0) = 0$ from the beginning.

Notice however that in each case subsidiary peaks are found close to the origin.

Another quantity of interest is the mean value of r, defined by

$$\bar{r}_n = \frac{\int <n|r> r <r|n> r^2 dr}{\int <n|r> <r|n> r^2 dr} = \frac{\int_0^\infty u_n \, r \, u_n dr}{\int_0^\infty u_n^2 dr} \tag{41}$$

Evaluation of the integral leads to the simple result

$$\bar{r}_n = 1.5 n^2 a_0 \tag{42}$$

$\bar{r}$ is not the same as r_m, although the two behave similarly. In like manner, one can calculate the mean value of other powers of r. The mean value of 1/r measures the potential energy. This turns out to be exactly twice the total energy of the state:

$$\bar{V}_n = \frac{e^2 \int_0^\infty (-\frac{1}{r}) u_n^2 dr}{\int_0^\infty u_n^2 dr} = \frac{-2E_R}{n^2} \tag{43}$$

where E_r is the Rydberg energy. The result (43) is called the virial theorem. An analogous relation holds also in classical mechanics: the mean value of potential energy for an orbit in an inverse-square force field is twice the total energy.

The result (43) implies that the mean value of the kinetic energy in each state must be numerically equal to the total energy. Verification of this is left as an exercise.

As we have mentioned earlier, by a peculiarity of the Coulomb potential the set of energy levels (38), which belong to only an incomplete set of hydrogen atom states, already provides the complete set of energy levels. Differences between these energies account for all the lines in the hydrogen spectrum (except for fine and hyperfine structure which is due to small terms not included in the Schrodinger equation used in this chapter).

The correct set of energy levels (38) is given also by the Bohr theory of the atom (exercise 32). It is of interest to compare the picture of the states as given by the Bohr theory with the one obtained from the Schrodinger equation. In the Bohr theory each state is described by a definite orbit. The radius of the nth circular orbit, (for which the energy is given by Eq. 38) is

$$r_n(\text{Bohr}) = n^2 a_0 \tag{44}$$

Except for the factor 1.5, this is the same as the mean radius for the Schrodinger state, Eq. 42. The magnitude of the linear momentum for the nth circular Bohr orbit is

$$p_n(\text{Bohr}) = m v_n = \frac{n\hbar}{r_n} = \frac{\hbar}{n a_0} \tag{45}$$

For the states whose position amplitudes are given by Eq. 39 we can find the momentum amplitudes by the standard methods of Chapter 9. The results for the first two states are

$$<\vec{p}|n = 1> \propto \frac{1}{(p^2 a_0^2 + 1)^2} \tag{46}$$

$$\langle \vec{p} | n = 2 \rangle \propto \frac{4p^2 a_0^2 - 1}{(4p^2 a_0^2 + 1)^3} \tag{47}$$

The mean values again correspond roughly to those of the Bohr theory. Notice , however, that

$$\bar{p}_n \bar{r}_n \sim n\hbar, \tag{48}$$

This shows, in the light of the uncertainty relations discussed in Chapter 10, that the concept of an orbit cannot apply to the states of low n. We shall see in the next chapter that in the limit of high n a correspondence can be exhibited between the solutions of the Schrodinger equation and the Bohr orbits. The correspondence exists, however, not for the spherically symmetric states considered here, but rather for states whose amplitudes have a particular angular dependence. The reason is that the spherically symmetric states carry no angular momentum, whereas an electron in the nth Bohr orbit has angular momentum $n\hbar$. The association of angular momentum with the angular dependence of spatial amplitudes is not an obvious one; it is discussed in the following chapter.

Appendix. Numerical Approximation: Spherically Symmetric Solutions to the Radial Equation

In this appendix we discuss a method of finding approximate solutions to Eq. 31

$$\frac{d^2u}{d\rho^2} + \left(\epsilon + \frac{2}{\rho}\right) u = 0 \tag{A1}$$

The approximation procedure is as follows: Divide the radial coordinate ρ into small segments ("mesh points") separated by the distance $\Delta\rho$. Let j be the number of the segment. Then make the following substitutions

$$\rho \rightarrow \rho_j = j\Delta\rho \qquad j = 0, 1, 2, \ldots$$

$$u(\rho) \longrightarrow u(\rho_j) \equiv u_j$$

$$\left.\frac{d^2u}{d\rho^2}\right|_{\text{at } \rho_j} = \frac{\dfrac{u_{j+1} - u_j}{\Delta\rho} - \dfrac{u_j - u_{j-1}}{\Delta\rho}}{\Delta\rho} = \frac{u_{j+1} - 2u_j + u_{j-1}}{(\Delta\rho)^2}$$

With these substitutions, Eq. A1 is converted from a differential equation to a difference equation.

$$u_{j+1} = \left[2 - \epsilon(\Delta\rho)^2 - 2\Delta\rho/j\right] u_j - u_{j-1} \tag{A2}$$

Equation A2 is used to approximate trial solutions to the Schrodinger equation according to the following procedure. First a mesh size $\Delta\rho$ is chosen. The smaller the mesh size is, the longer is the calculation and the more nearly the resulting approximate solution conforms to the analytic solution. Figures 2 and 3 use a mesh size of 50 points per Bohr radius a_0. After choosing a mesh size, we assign a trial value for the energy ϵ. Integration then begins at $\rho = 0$ with $u(0) = u_o = 0$, as required by the boundary conditions presented on page 12-11. The value of u_1 at the next step $\rho = 0 + \Delta\rho$ determines the initial slope of the function and thus the vertical scale of the diagram. Therefore we choose u_1 by trial and error to yield a convenient scale. The value of u_2 at the second mesh point is then calculated using Eq. A2 with $j = 1$. Then the value of u_3 is calculated using u_2 and u_1, and so forth. In this manner the entire approximate function is drawn for the chosen trial value of energy ϵ.

The trial value of energy is physically acceptable only if the resultant radial function u approaches zero for large radii. It is easy to show that once the function begins to diverge from the $u = 0$ axis at a large radius, then it will continue to diverge with increasingly diverging slope as ρ increases. For large values of ρ, Eq. A1 becomes

$$\frac{d^2u}{d\rho^2} \approx -\epsilon u \qquad \text{large } \rho \tag{A3}$$

For negative ϵ (bound states) the slope increases in a positive sense for positive u and increases in a negative sense for negative u. This means that for large ρ once the function begins to diverge from the $u = 0$ axis it will continue to diverge as ρ increases. This qualitative conclusion is verified in the diverging functions of Figures 2 and also assures us that these curves do not approach the $u = 0$ axis for greater values of ρ than we have used in the figures.

Chapter 13. The Hydrogen Atom: States with Angular Dependence

1. Introduction

In this chapter we complete our study of the hydrogen atom by investigating those solutions of the Schrodinger equation which are not spherically symmetric. The equation can be separated into two angular equations and a radial equation. The angular equations do not involve the potential; their solutions therefore describe the angular dependence of the stationary states of a two-body system with any central force. These angular solutions are characterized by two quantum numbers, ℓ and m, which are restricted to integral values and which are interpreted as measures of the angular momentum and the z-component of angular momentum respectively. All the solutions of the radial equation for hydrogen contain the same exponential factors as do the spherically symmetric states discussed in the preceding chapter. Furthermore, the allowed energy values are not different from those of the spherically symmetric states. The stationary states of the hydrogen atom are thus characterized by the three quantum numbers n, ℓ, and m; but the energy depends only upon n.

2. Separation of the Schrodinger equation

In the preceding chapter we wrote the amplitude $\langle \vec{r}_e, \vec{r}_p | \psi \rangle$ for a stationary state of the hydrogen atom in the form

$$\langle \vec{r}_e, \vec{r}_p | \psi \rangle = \phi(\vec{R})\psi(\vec{r}) \tag{1}$$

and showed that the "internal" function $\psi(\vec{r})$ satisfies the differential equation

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2 + V(r) \right] \psi(\vec{r}) = E\psi(\vec{r}) \tag{2}$$

where $V(r) = -e^2/r$; for the present we keep the general form $V(r)$. The solutions of Eq. 2 described in Chapter 12 are those in which $\psi(\vec{r})$ depends only on the radial coordinate r. We now drop this restriction and let $\psi(\vec{r})$ be a function of all three spherical coordinates r, θ, and ϕ. The transformation of the Laplacian to spherical coordinates is a straightforward (though tedious) procedure. The result is

$$\nabla^2\psi(\vec{r}) = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\psi) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2} \tag{3}$$

The method of separation of variables is commonly used to solve an equation of the form (2), in which the potential is independent of at least one coordinate. One looks for solutions that are products of functions, each of which depends upon only one coordinate. In the present case, this amounts to looking for solutions of the form

---

$$\psi(\vec{r}) = \frac{u(r)}{r} P(\theta)\,\Phi(\phi) \tag{4}$$

where we have taken advantage of the experience with the radial equation in Chapter 12 to write $u(r)/r$ as the radial dependence. Substituting this form for $\psi(\vec{r})$, and expression (3) for the Laplacian, into Eq. 2, one gets

$$P(\theta)\,\Phi(\phi)\left[\frac{1}{r}\,\frac{d^2u(r)}{dr^2}\right] + \frac{u(r)\Phi(\phi)}{r^3}\left[\frac{1}{\sin\theta}\,\frac{d}{d\theta}\left(\sin\theta\,\frac{dP(\theta)}{d\theta}\right)\right]$$
$$+ \frac{u(r)P(\theta)}{r^3\sin^2\theta}\,\frac{d^2\,\Phi(\phi)}{d\phi^2} + \frac{2\mu}{r\hbar^2}\,u(r)P(\theta)\Phi(\phi)\,[E - V(r)] = 0 \tag{5}$$

This equation is simplified by dividing through by the product $\frac{u(r)}{r} P(\theta)\Phi(\phi)$ and multiplying by $r^2\sin^2\theta$. One then obtains

$$\left[\frac{r\sin^2\theta}{u}\,\frac{d^2u}{dr^2} + \frac{\sin\theta}{P}\,\frac{d}{d\theta}\left(\sin\theta\,\frac{dP}{d\theta}\right) + \frac{2\mu r^2}{\hbar^2}\sin^2\theta(E - V)\right] + \left[\frac{1}{\Phi}\,\frac{d^2\Phi}{d\phi^2}\right] = 0 \tag{6}$$

In Eq. 6 the first square bracket is a function of r and θ only; the second is a function of ϕ only; their sum is zero. By the same argument already employed in the preceding chapter (p. 12-5), we conclude that each "function" must be a constant. Accordingly we put

$$\frac{1}{\Phi}\,\frac{d^2\Phi}{d\phi^2} = -m^2 \tag{7}$$

At the present stage the so-called "separation constant" $-m^2$ can be arbitrary; its values are restricted by a physical argument below. Substitute (7) into (6), divide through by $\sin^2\theta$, and rearrange to obtain

$$\left[\frac{r}{u}\,\frac{d^2u}{dr^2} + \frac{2\mu r^2}{\hbar^2}\Big(E - V(r)\Big)\right] + \left[\frac{1}{P\sin\theta}\,\frac{d}{d\theta}\left(\sin\theta\,\frac{dP}{d\theta}\right) - \frac{m^2}{\sin^2\theta}\right] = 0 \tag{8}$$

Equation 8 again consists of two terms, one a function of one variable and the other of a different variable, whose sum is zero. Each of these must be a constant. Call these constants $+\lambda$ and $-\lambda$. Then

$$\frac{1}{P\sin\theta}\,\frac{d}{d\theta}\left(\sin\theta\,\frac{dP}{d\theta}\right) - \frac{m^2}{\sin^2\theta} = -\lambda \tag{9}$$

and

$$\frac{r}{u}\,\frac{d^2u}{dr^2} + \frac{2\mu r^2}{\hbar^2}\Big(E - V(r)\Big) = +\lambda \tag{10}$$

The new separation constant λ is again arbitrary as far as one can tell at the present stage of the argument; its values will likewise be restricted on the basis of a physical requirements.

Equations 7, 9, and 10 are separate ordinary differential equations whose solutions combine to give the separable solutions of the original Equation 2. Since the potential appears only in the radial equation, the solutions of Eqs. 7 and 9 apply to any central field problem: the space amplitudes of the stationary states for <u>any</u> central field problem have the <u>same</u> angular dependence. These angular solutions, called spherical harmonics, have therefore a wide applicability.

3. The ϕ equation

The simplest of the separated equations is the ϕ equation, number 7. Its general solution is a linear combination of the functions

$$\Phi_m(\phi) = e^{im\phi} \tag{11a}$$

and

$$\Phi_{-m}(\phi) = e^{-im\phi} \tag{11b}$$

(An alternative form would be a linear combination of sin $m\phi$ and cos $m\phi$.) In the most general case of complex m,

$$m = m_1 + i\,m_2 \tag{12}$$

the solution (11a) reads

$$\Phi_m(\phi) = e^{-m_2\phi + im_1\phi} \tag{13}$$

with a similar expression for $\Phi_{-m}(\theta)$. However, the physical requirement that $\psi(\vec{r})$ be a single-valued function or $\vec{r}$ leads to the condition

$$\Phi(2\pi) = \Phi(0), \tag{14}$$

since $\phi = 2\pi$ and $\phi = 0$ represent the same point in space. The solutions (11) satisfy (14) only if m is a positive or negative integer. The requirement of a single-valued ψ seems stronger than can be justified on purely physical grounds. One may argue that only $|\psi|^2$, which represents a measurable probability, has to be single-valued. Indeed, we have already encountered a case (spin 1/2 amplitudes, Chapter 7) in which going from $\phi = 0$ to $\phi = 2\pi$ changes the sign of amplitudes. The weaker requirement that $|\psi|^2$ be single-valued allows half-integer as well as integer values of m. However, if $\Phi(2\pi)$ differs from $\Phi(0)$, then the spatial derivatives of $\psi(\vec{r})$ (for example the gradient and Laplacian) are undefined. Hence even though $\Phi(\phi)$ satisfes the ϕ equation for half-integer m, the complete function cannot be a solution of the full equation 6 near $\phi = 0$. Hence we must exclude the half-integer values of m.

Interpretation of the constant m. The function $e^{im\phi}$ satisfies the following simple relation

$$-i\hbar \frac{\partial}{\partial\phi}(e^{im\phi}) = m\hbar(e^{im\phi}) \tag{15}$$

We use this relation to infer that m measures the z-component of angular momentum. For this purpose, transform the ϕ derivative to rectangular coordinates:

$$\begin{aligned}\frac{\partial}{\partial\phi} &= \frac{\partial x}{\partial\phi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\phi}\frac{\partial}{\partial y} + \frac{\partial z}{\partial\phi}\frac{\partial}{\partial z}\\ &= -r\sin\theta\sin\phi\frac{\partial}{\partial x} + r\sin\theta\cos\phi\frac{\partial}{\partial y}\\ &= -y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}\end{aligned} \tag{16}$$

The last form can be written as the z-component of $\vec{r} \times \vec{\nabla}$:

$$\frac{\partial}{\partial \phi} = (\vec{r} \times \vec{\nabla})_z \tag{17}$$

We have seen earlier (Chapter 11) that $-i\hbar$ times the gradient can be associated with linear momentum. Recall the relation

$$-i\hbar\vec{\nabla} e^{i\vec{k}\cdot\vec{r}} = (\hbar\vec{k}) e^{i\vec{k}\cdot\vec{r}} = \vec{p} e^{i\vec{k}\cdot\vec{r}} \tag{18}$$

which gives the linear momentum of the state $|\vec{k}\rangle$ with space amplitude $e^{i\vec{k}\cdot\vec{r}}$. Now, the classical expression for angular momentum of a particle is $\vec{L} = \vec{r} \times \vec{p}$. This suggests that $-i\hbar(\vec{r} \times \vec{\nabla})$ should similarly be associated with the angular momentum L, while (17) leads us to associate $-i\hbar \frac{\partial}{\partial \phi}$ with the z-component of angular momentum. If such an interpretation is valid, then according to Eq. 15 a state whose ϕ dependence is $e^{im\phi}$ has a z-component of angular momentum $m\hbar$. The restriction of m to integer values then means that *the z-component of orbital angular momentum is quantized.*

The states under discussion in the present chapter, characterized by a quantized value of the z-component of orbital angular momentum, have many properties in common with the spin states studied in Chapter 7. In particular, they are both separated into discrete beams in a Stern-Gerlach experiment. In general, atomic states involve both electron spin and electron orbital angular momentum. In the present discussion the spin has been ignored. Complete analysis shows that the spin and orbital angular momenta combine to form a total angular momentum J, whose z-component, written m_J, is quantized. The values of m_J are sometimes integer, sometimes half-integer. The number of separated beams into which an incident atomic beam is (ideally) split in a Stern-Gerlach experiment corresponds to the number of different values of m_J for the state under study. Because of the intimate connection between the quantum number m (or m_J) and the magnetic properties of an atom, m is often called the *magnetic quantum number.*

4. The θ equation

We turn next to the equation for $P(\theta)$, Eq. 9. When the derivatives are written out in full, the equation reads

$$\frac{d^2P}{d\theta^2} + \cot\theta \, \frac{dP}{d\theta} + \left(\lambda - \frac{m^2}{\sin^2\theta}\right)P = 0 \tag{19}$$

Equation 19, called Legendre's equation, appears in many branches of theoretical physics.

We impose as a physical boundary condition the requirement that $P(\theta)$ be everywhere finite.*

---

* The condition that $P(\theta)$ be everywhere finite is a bit stronger than is demanded by the probability interpretation. The latter requires only that $\int |\psi(\vec{r})|^2 d^3r \propto \int |P(\theta)|^2 \sin\theta d\theta$ be finite. This weaker condition can be satisfied by a function $P(\theta)$ that becomes infinite at $\theta = 0$ or π, as long as the divergence is no stronger than $(\sin\theta)^{-1/2}$. However, no solutions with such singularities exist. Therefore no harm is done by requiring from the beginning that $P(\theta)$ itself be everywhere finite. (See similar footnote in connection with the radial equation, Chapter 12, page 11.)

Because of the factors $\cot\theta$ and $m^2/\sin^2\theta$, both of which are infinite at $\theta = 0$ and $\theta = \pi$, the boundary condition cannot be satisfied for most values of λ. A solution which is finite at one endpoint will in general diverge at the other; if one carries out a numerical integration beginning at some point other than 0 or π, the solution will in general diverge at both endpoints (see Figure 1). However, for certain discrete values of λ, a solution that is everywhere finite does exist. These are the solutions in which we are interested. The allowed values of λ and corresponding solutions $P(\theta)$ can be determined approximately by numerical integration (Figure 2); however, exact solutions exist and have been known for a long time. The method of solution is described in many standard texts.*

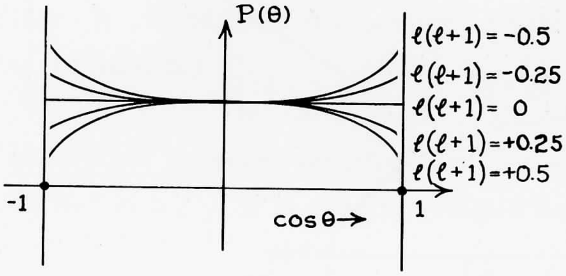


Figure 1. Computer-generated trial solutions to Eq. 19 for $m = 0$ and $\lambda = \ell(\ell+1)$ near the value zero. Horizontal dimensions are proportional to $\cos\theta$ and cover the range -1 to 1, corresponding to the full range of the angle θ: $0 \leqslant \theta \leqslant \pi$. Integration proceeds in both directions starting at the midpoint $\theta = \pi/2$, $\cos\theta = 0$. The acceptable solution, for which $\lambda = 0 = \ell$, is identified by the fact that it does not diverge near the endpoints, $\cos\theta = \pm 1$. Divergence of the other trial solutions occur very close to these endpoints. Only the beginnings of these divergences are shown in the figure. Computer output courtesy Dr. Judah Schwartz.

The simplest case is $m = 0$; in this case the values of λ for which physically acceptable solutions exist turn out to be

$$\lambda = 0, 2, 6, 12, \ldots \tag{20}$$

These values are summarized by introducing a quantum number ℓ, such that

$$\lambda = \ell(\ell + 1) \qquad \ell = 0, 1, 2, 3, \ldots \tag{21}$$

The corresponding solutions of (19) are polynomials in $\cos\theta$ called Legendre polynomials and written $P_\ell(\cos\theta)$. The first few polynomials are

---

* See, for example, page 172 of Quantum Mechanics by Eugen Merzbacher, John Wiley and Sons, Inc., New York, 1961.

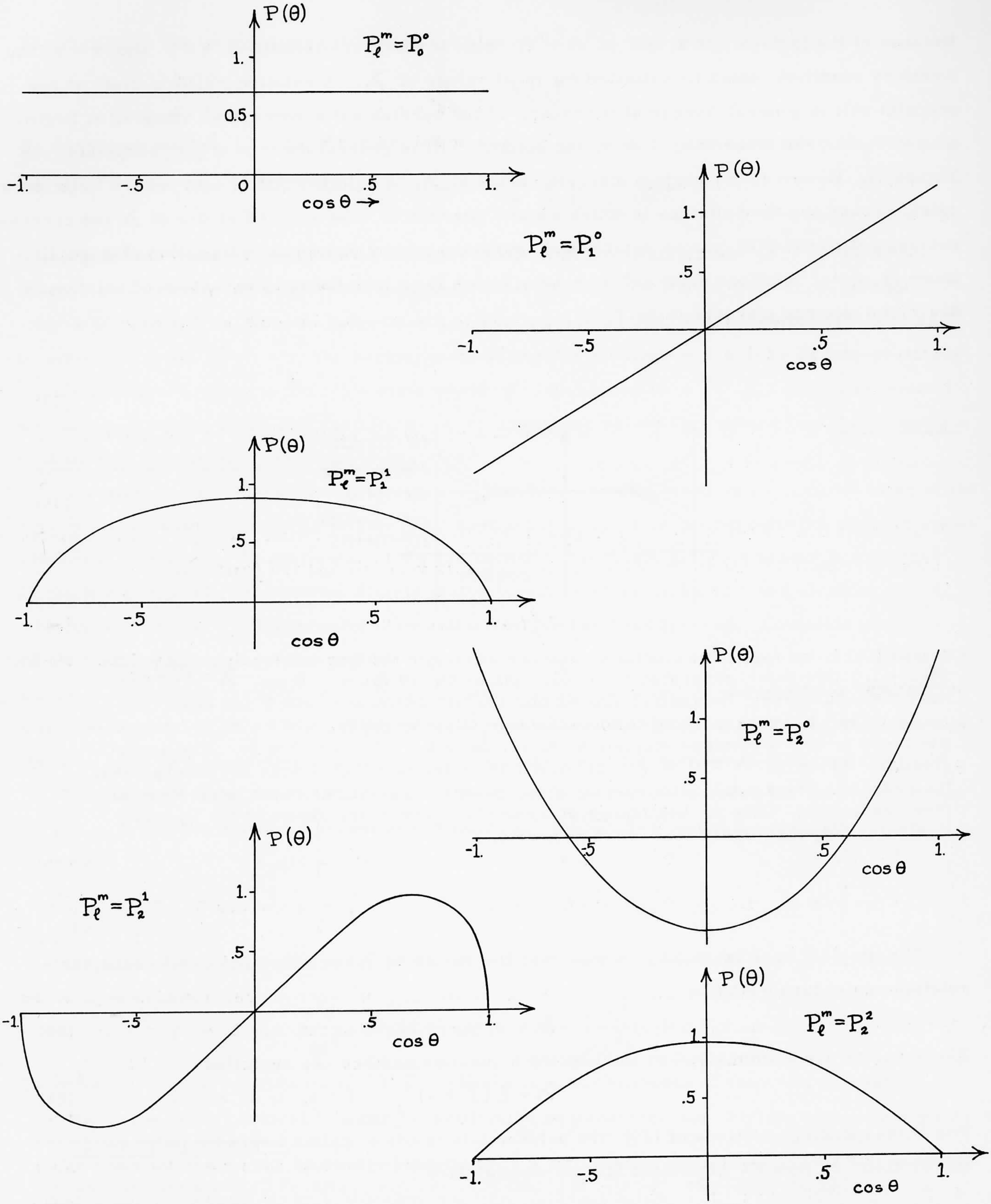


Figure 2. Solutions of the θ equation (19) for $\ell = 0$, 1, and 2, plotted as functions of cos θ. Analytic expressions for these solutions are found in Eq. 26.

$$\begin{aligned} P_0 &= 1 & \lambda &= 0 \\ P_1 &= \cos\theta & \lambda &= 2 \\ P_2 &= 3\cos^2\theta - 1 & \lambda &= 6 \\ P_3 &= 5\cos^3\theta - 3\cos\theta & \lambda &= 12 \end{aligned} \qquad (22)$$

It is readily verified by substitution that these polynomials satisfy Eq. 19, with $m = 0$, for the indicated values of λ.

Consider next the case of nonzero m. According to the results of the preceding section, only integer m need be considered. The values of λ for which physically acceptable solutions exist are again of the form $\ell(\ell + 1)$, just as for $m = 0$. In addition, however, the allowed values of ℓ must now satisfy the condition

$$\ell \geqslant |m| \qquad (23)$$

Alternatively, if we consider λ to be fixed at a particular value $\ell(\ell + 1)$, we find that acceptable solutions of Eq. 19 exist for

$$m^2 = 0, 1, 4, \ldots\ldots, \ell^2 \qquad (24)$$

The solutions are called associated Legendre functions and are written as P_ℓ^m. The simplest of these functions are those for $m = \ell$. It is readily verified that the functions

$$P_\ell^\ell(\theta) = (\sin\theta)^\ell \qquad (25)$$

satisfys (19) if $m = \ell$. The functions P_ℓ^0 are of course the Legendre polynomials (22). We list here all the solutions for $\ell \leqslant 3$:

Since only m^2 enters Eq. 19, it follows that $P_\ell^{-m}(\theta) = P_\ell^m(\theta)$.

$$\begin{aligned} &P_0^0 = 1 \\ &P_1^0 = \cos\theta, && P_1^1 = \sin\theta \\ &P_2^0 = 3\cos^2\theta - 1, && P_2^1 = \cos\theta\sin\theta, && P_2^2 = \sin^2\theta \\ &P_3^0 = 5\cos^3\theta - 3\cos\theta, && P_3^1 = \sin\theta(5\cos^2\theta - 1), && P_3^2 = \sin^2\theta\cos\theta\,, \quad P_3^3 = \sin^3\theta \end{aligned} \qquad (26)$$

The complete angular dependence of the solution of the original Schrodinger equation is given by the function $P_\ell^m(\theta)$ multiplied by the solution of the ϕ equation obtained in the preceding section. The full angular functions are called spherical harmonics and are symbolized by $Y_{\ell m}$:

$$Y_{\ell m}(\theta, \phi) = P_\ell^m(\theta)\, e^{im\phi} \qquad (27)$$

Since m runs from $-\ell$ to $+\ell$, there are $2\ell + 1$ distinct spherical harmonics for each value of ℓ. Note that $|Y_{\ell m}|^2$, which gives the angular dependence of the probability distribution, is independent of ϕ. This is because of our choice of $e^{\pm im\phi}$, instead of $\cos m\phi$ and $\sin m\phi$, as the independent solutions of Eq. 7.

The angular equations (7) and (9) do not involve the potential. Therefore the spherical harmonics represent the angular dependence of the stationary states for any central field problem.

Interpretation of the constant ℓ. In the preceding section we remarked that the differential operator $-i\hbar \frac{\partial}{\partial \phi}$, acting on the function $e^{im\phi}$, gives the same function multiplied by the constant $m\hbar$. In mathematical language, the statement is that $e^{im\phi}$ is a characteristic function (or eigenfunction) of the differential operator $-i\hbar\partial/\partial \phi$, which we recognized as being associated with the z-component of the angular momentum. The corresponding characteristic value (eigenvalue) $m\hbar$ was then interpreted as the value of the z-component of angular momentum for the state in question. Here we wish to make a similar argument which leads to the identification of the quantity $\ell(\ell+1)\hbar^2$ as the square of the total (orbital) angular momentum. We saw that with the identification

$$\vec{p} \longleftrightarrow -i\hbar\vec{\nabla}, \tag{28}$$

the differential operation $-i\hbar\partial/\partial\phi$ corresponds to

$$L_z = (\vec{r} \times \vec{p})_z \longleftrightarrow -i\hbar(\vec{r} \times \vec{\nabla})_z \tag{29}$$

Analogously, the other two components of $\vec{L}$ are also associated with differential operators, thus

$$L_x \longleftrightarrow -i\hbar(\vec{r} \times \vec{\nabla})_x = -i\hbar\left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y}\right) \tag{30}$$

$$L_y \longleftrightarrow -i\hbar(\vec{r} \times \vec{\nabla})_y = -i\hbar\left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z}\right) \tag{31}$$

Similar expressions can be written for the squares of the angular momentum components. For example,

$$\begin{aligned} L_x^2 &\longleftrightarrow (-i\hbar)^2 \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\left(y \frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) \\ &= -\hbar^2\left(y^2 \frac{\partial^2}{\partial z^2} - 2yz \frac{\partial^2}{\partial y \partial z} + z^2 \frac{\partial^2}{\partial y^2} - y\frac{\partial}{\partial y} - z\frac{\partial}{\partial z}\right) \end{aligned} \tag{32}$$

The expression for $L^2 = L_x^2 + L_y^2 + L_z^2$ has a most suggestive form when written in spherical coordinates. The transformation to spherical coordinates is a tedious one and leads to the result

$$L^2 \longleftrightarrow -\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2}\right) \tag{33}$$

These are just the derivatives that appear in Eqs. 7 and 9. The spherical harmonics $Y_{\ell m}(\theta, \phi)$ therefore satisfy the equation

$$-\hbar^2\left(\frac{\partial^2}{\partial\theta^2} + \cot\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2}\right) Y_{\ell m}(\theta, \phi) = \ell(\ell+1)\hbar^2 Y_{\ell m}(\theta, \phi) \tag{34}$$

Equation 34 asserts that the $Y_{\ell m}$ are eigenfunctions of a differential operator which, according to Eq. 33, is associated with the square of the angular momentum. The corresponding eigenvalues, $\ell(\ell+1)\hbar^2$, give the value of this quantity for the state with the quantum numbers ℓ and m.

The result that the stationary states for the central field problem are characterized by definite values for the square of the angular momentum is a logical extension of classical mechanics. Central forces imply conservation of angular momentum; hence any classical orbit in a central field is characterized by a definite value of angular momentum. (Not, of course, restricted by classical mechanics to any set of discrete values.) A further correspondence with classical mechanics may be established by writing the radial equation, 10, as

$$-\frac{\hbar^2}{2\mu}\frac{d^2u(r)}{dr^2} + \left[V(r) + \frac{\ell(\ell+1)\hbar^2}{2\mu r^2}\right]u(r) = Eu(r) \qquad (35)$$

Equation 35 looks like a one-dimensional Schrodinger equation, with the potential energy $V(r)$ and an additional repulsive potential $\ell(\ell + 1)\hbar^2/(2\mu r^2)$. Recall that when a three-dimensional central field problem in classical mechanics is reduced to an equivalent one-dimensional problem for the radial motion,* the equation of conservation of energy takes a form very similar to (35), namely

$$\frac{1}{2}m\dot{r}^2 + V(r) + \frac{L^2}{2mr^2} = E \qquad (36)$$

Equation 36, together with the conservation of angular momentum

$$mr(d\theta/dt)^2 = L = \text{constant} \qquad (37)$$

specifies the classical motion. In the radial equation, 36, the term $L^2/(2mr^2)$ acts as an additional "centripetal potential" which comes from the angular momentum. The presence of the analogous term $\ell(\ell + 1)\hbar^2/(2\mu r^2)$ in Eq. 35 strongly supports the interpretation of $\ell(\ell + 1)\hbar^2$ as the square of the angular momentum.

According to the interpretations made in this and the preceding sections, the solutions of the angular parts of the Schrodinger equation, with the separation constants $\ell(\ell + 1)$ and m^2, represent states in which the square of the angular momentum and the z-component of angular momentum respectively have the values $\ell(\ell + 1)\hbar^2$ and $m\hbar$. Note that this says nothing about the x or y components of angular momentum. It can be verified that the solutions are *not* eigenfunctions of the operators (30) and (31) that represent L_x and L_y. This means that the x and y components of angular momentum do not have unique values for the states we have exhibited. The situation is analogous to the one encountered in our analysis of spin states. There also, only the z-component of angular momentum could be specified in a given state. Of course the choice of the z axis is arbitrary. The essential point is that the *square* of the angular momentum and the magnitude of *one* component are specified.

Finally, we remark that the condition (24) implies

$$m^2 \leqslant \ell^2 < \ell(\ell + 1) \qquad (38)$$

---

* See, for example, *Physics, a New Introductory Course*, Part II, page 17-3.

which is also consistent with the interpretation we have made of the constants $\ell(\ell + 1)$ and m. Evidently the square of one component of angular momentum can be no larger than the square of the total angular momentum. The fact that (38) is an inequality rather than an equality implies that even when L_z has its maximum permissible value, the other two components cannot be considered to be identically zero. In fact, the expectation value of $L_x^2 + L_y^2$ must be $\ell\hbar^2$ in the state for which $m = \ell$, in order to make a consistent interpretation. The situation is similar to the one we found for spin states; no quantum state corresponds exactly to the classical notion of an angular momentum oriented in a definite direction in space.

5. The radial equation.

There remains now to solve the radial equation, which we have just rewritten in the form (35). Notice that the radial equation involves the separation constant λ, but not m. This means that whatever the form of the radial functions, the energies of the stationary states are independent of the magnetic quantum number m. The physical interpretation is that since there is no preferred direction in a central field problem, the energy of a state cannot depend on the direction of the z-axis and thus on the z-component of angular momentum.

For the hydrogen atom $V(r) = -e^2/r$. Equation (35) is then identical with the radial equation for the spherically symmetric case (Eq. 28 of Chapter 12), except for the additional centripetal term $\ell(\ell + 1)\hbar^2/(2\mu r^2)$. If we make the same changes of variable as in the Chapter 12, defining

$$\rho = r/a_o \qquad \epsilon = E/E_R \tag{39}$$

then Eq. (35) becomes

$$\frac{d^2u}{d\rho^2} + \left[\epsilon - \frac{\ell(\ell + 1)}{\rho^2} + \frac{2}{\rho}\right] u = 0 \tag{40}$$

For $\ell = 0$, the solutions are of course the functions discussed in Chapter 12. (The angular solution for $\ell = 0$ is a constant independent of θ.) A clue as to the nature of the solutions for nonzero ℓ is obtained by observing that near the origin, the inverse-square centripetal potential dominates the Coulomb potential (which varies as $1/\rho$) as well as the energy term ϵ (which is a constant). Thus the solutions in the neighborhood of the origin must resemble those of the simple equation

$$\frac{d^2u}{d\rho^2} - \frac{\ell(\ell + 1)}{\rho^2} u = 0 \tag{41}$$

Equation 41 is readily solved; the solutions are

$$u \propto \rho^{\ell + 1} \tag{42a}$$

and

$$u \propto \rho^{-\ell} \tag{42b}$$

Since ℓ is positive, (42b) is divergent at the origin and must be excluded. Therefore, any acceptable solution of Eq. 40 must behave as $\rho^{\ell+1}$ near the origin. This is an extension of the condition $u(0) = 0$ which we employed for the spherically symmetric states. For large ρ, the argument of Chapter 12, page 8 still applies and the solution must resemble the form

$$e^{-|\epsilon|^{1/2}\rho} \tag{43}$$

Once again the physical boundary conditions can be satisfied only for certain discrete values of ϵ. It happens that these acceptable values of ϵ are always of the same form $\epsilon = -1/n^2$, no matter what the value of ℓ. However, for a given ℓ, a solution exists only if

$$n \geqslant \ell + 1 \tag{44}$$

Alternatively, for a fixed value of n, solutions exist for values of ℓ given by the relation

$$0 \leqslant \ell < n \tag{45}$$

Each of these solutions has the same energy. Energy is thus independent of angular momentum in the hydrogen atom, a "degeneracy" peculiar to the Coulomb potential. For each n there are n different solutions. Moreover, for each value of ℓ there are $2\ell + 1$ different solutions of the ϕ equation. Thus, altogether, for any given n there are

$$\sum_{\ell=0}^{n-1} (2\ell + 1) = n^2 \tag{46}$$

distinct stationary states, all having the same principal quantum number n and the same energy.

Figure 3 shows the energy level diagram of hydrogen as determined from the solutions of the Schrodinger equation. Despite the degeneracy in the energy values for different values of ℓ, it is useful to use a separate column for each value of ℓ. There are two reasons. First, the effect of small internal magnetic interactions, not included in Eq. 2, is to break the degeneracy: the energies of the different states with the same principal quantum number n in fact differ by small but detectable amounts. This results in the so-called *fine structure and hyperfine structure* of the hydrogen spectrum. A second reason for using a separate column for each value of ℓ in Figure 3 is that when an atom makes a transistion from one state to a state of lower energy, emitting a photon, the most likely transition involves a unit change in ℓ

$$\Delta\ell = \pm 1 \tag{47}$$

Some of these transitions are indicated by arrows in the figure.

The radial function for the state with quantum numbers n and ℓ is written u_n^ℓ. These functions are obtained analytically by series solution, or numerically by the technique already described in Chapter 12. Figure 4 illustrates the numerical determination of the radial functions. We list below the analytic form of the first few functions. Notice that for $\ell = 0$ the functions are just the radially symmetric functions, Eq. 39 of Chapter 12.

$$u_1^0 = \rho e^{-\rho} \tag{48}$$

$$u_2^0 = e^{-\rho/2}(1 - \rho/2)\rho, \qquad u_2^1 = \rho^2 e^{-\rho/2} \tag{49}$$

$$u_3^0 = e^{-\rho/3}(1 - \tfrac{2}{3}\rho + \tfrac{2}{27}\rho^2)\rho, \qquad u_3^1 = e^{-\rho/3}(1 - \rho/6)\rho^2, \qquad u_3^2 = \rho^3 e^{-\rho/3} \tag{50}$$

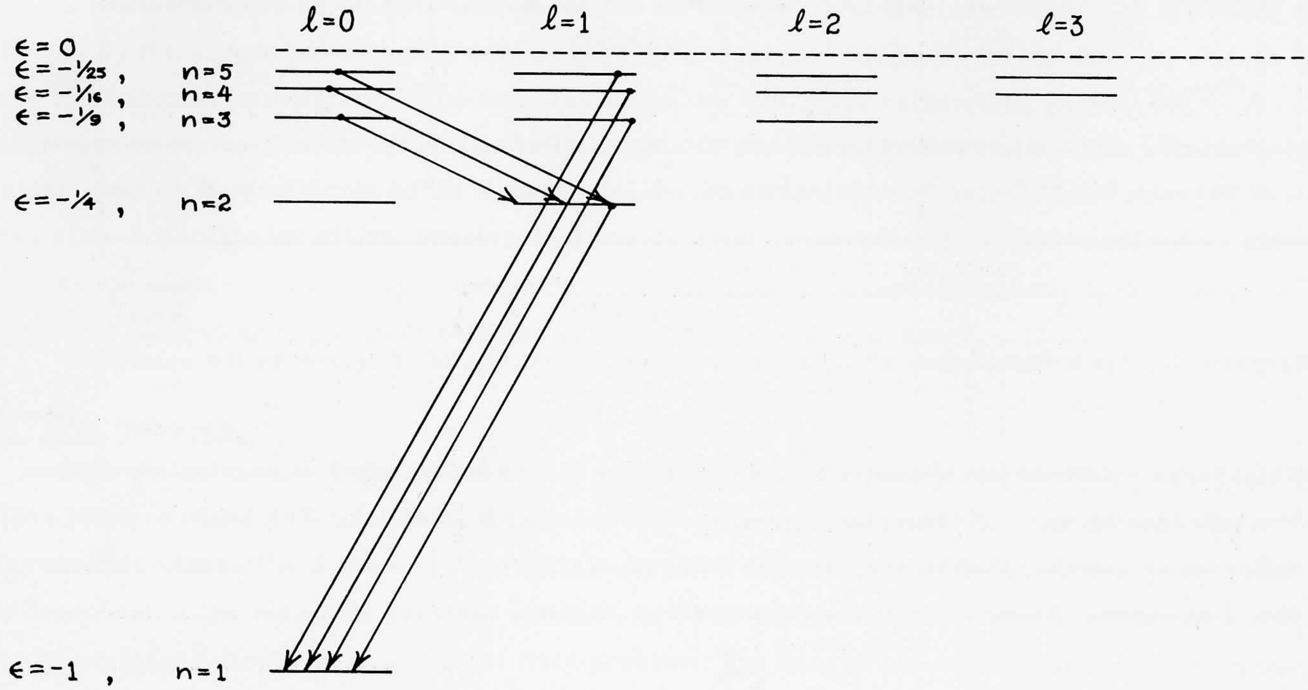


Figure 3. Energy level diagram of hydrogen showing the five lowest energies. States with different orbital angular momenta are separated into different vertical columns. Notice that for each n there are n different values of $\ell = 0, 1, 2, \ldots\ldots, (n-1)$. The arrows indicate a few of the transitions associated with the emission of a photon. These transitions are those for which the angular momentum changes by one unit.

Notice also that the solution for $\ell = n-1$ has a particularly simple form

$$u_n^{n-1} = \rho^n e^{-\rho/n} \tag{51}$$

6. Discussion of the stationary states of hydrogen

We use the state vector $|n, \ell, m>$ to denote the stationary state of hydrogen characterized by the quantum numbers n, ℓ, and m. When the results of Sections 3 to 5 are combined, one obtains the full spatial dependence of this state:

$$<\vec{r}|n, \ell, m> = \frac{1}{r} u_n^{\ell}(r)\ P_{\ell}^{m}(\theta) e^{im\phi} \tag{52}$$

The projection amplitude from one state to another is

$$<n', \ell', m'|n, \ell, m>$$
$$= \int_0^{\infty} dr\ u_{n'}^{\ell'}(r) u_n^{\ell}(r) \quad \int_0^{\pi} \sin\theta d\theta P_{\ell'}^{m'}(\theta) P_{\ell}^{m}(\theta) \int_0^{2\pi} d\phi e^{i(m-m')\phi} \tag{53}$$

It can be demonstrated that the angular integrations in (53) give zero unless $m = m'$ and $\ell = \ell'$; that is, the spherical harmonics are orthogonal over a sphere. Likewise, the radial integral vanishes unless $n = n'$. Thus the states satisfy a full orthogonality condition:

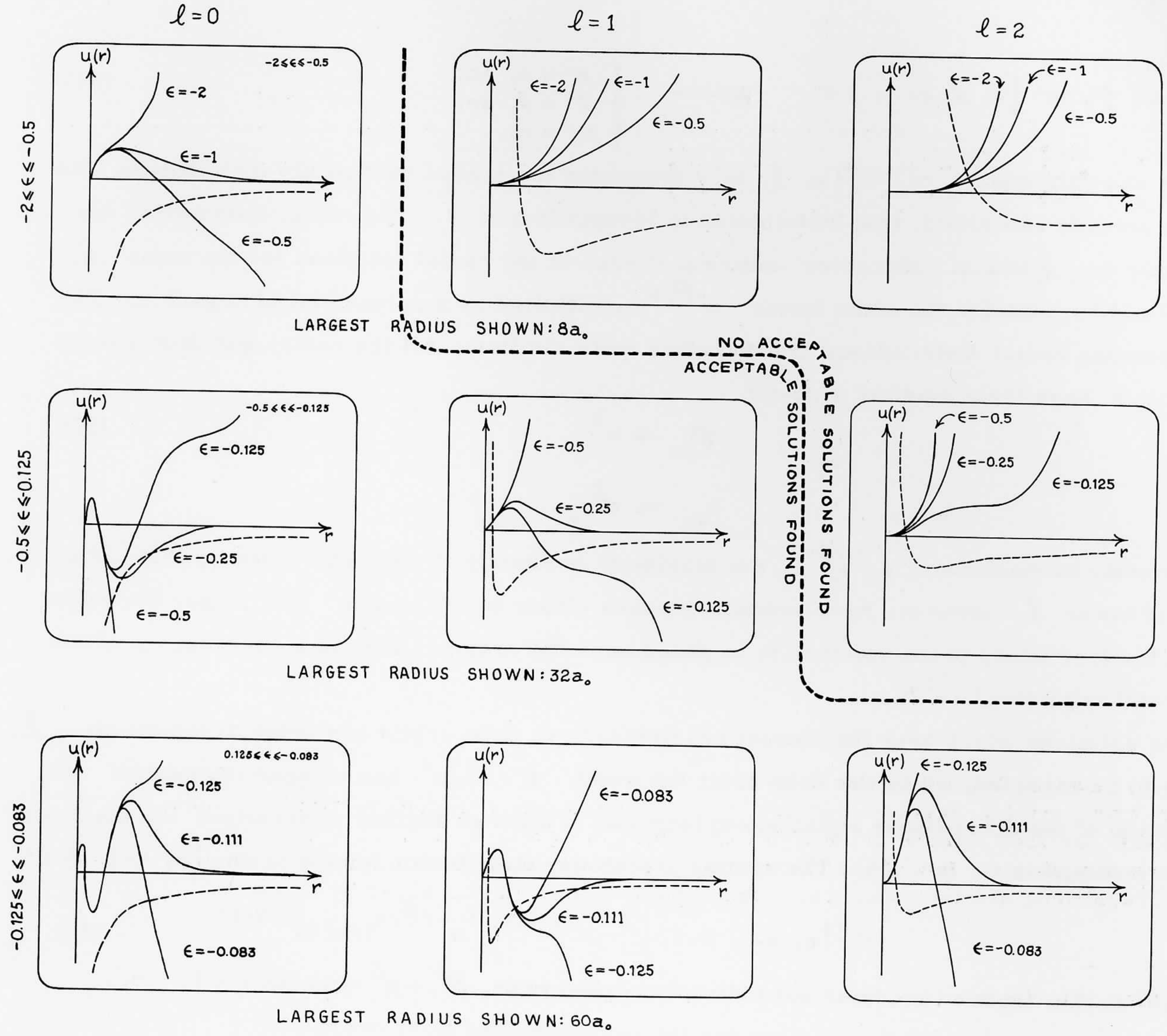


Figure 4. Computer-generated trial solutions for the radial equation (40). The approximation method used is a simple extension of that presented in the appendix to Chapter 12. Physically acceptable solutions are found for energies $\epsilon = -1/n^2$. These energies are independent of the angular momentum quantum number ℓ ; however, acceptable solutions are found only for integer ℓ less than n. Solutions for $\ell = 0$ are the radially-symmetric solutions of Chapter 12, Figure 2.

$$\langle n', \ell', m' | n, \ell, m \rangle = 0 \qquad \text{unless} \quad \begin{cases} n' = n, \text{ and} \\ \ell' = \ell, \text{ and} \\ m' = m. \end{cases} \tag{54}$$

The absolute square of $\langle \vec{r} | n, \ell, m \rangle$ measures the spatial probability distribution. As we have already remarked, this distribution is independent of ϕ. The radial distribution depends only on ℓ and n. Moreover, notice that each of the radial functions for the same n, Eqs. 49 and 50, contains the same factor $e^{-\rho/n}$, multiplied by a polynomial of degree n. The corresponding radial distributions are therefore quite similar. All the radial distributions for the same n have their maxima at about

$$\rho_m \sim n^2 \tag{55}$$

$$\text{i.e.,} \qquad r_m \sim n^2 a_o$$

For the state of maximum ℓ, $u_n^{\ell-1}$, the maximum is exactly at the value given by (55). The states of lower ℓ, however, have secondary peaks closer to the origin. The larger the value of ℓ, the less likely is the electron to be found near the proton. This is a consequence of the centripetal repulsion.

The solutions which bear the closest resemblance to Bohr orbits are those of maximum ℓ, This is to be expected, since the Bohr orbit for which $\epsilon = -1/n^2$ has angular momentum $n\hbar$. No solution of the Schrodinger equation corresponds to such an angular momentum; the maximum angular momentum is $(n - 1)\hbar$. The spatial probability distribution for the state $|n, n-1, n-1\rangle$ is

$$\left|\langle \vec{r} | n, n-1, n-1 \rangle\right|^2 = \rho^{2(n-1)} e^{-2\rho/n} (\sin\theta)^{2(n-1)} \tag{56}$$

For high n, this function becomes very sharply peaked near $\rho^2 = n^2$ and $\sin\theta = 1$, i.e., $\theta = \pi/2$. These are just the parameters for the corresponding Bohr orbit.

Analysis of the line spectrum of hydrogen confirms the correctness of the set of energy levels given by the Schrodinger theory. Each observed line corresponds to a transition between two levels. Of course, the same level scheme is also predicted by the simpler Bohr theory. Thus the overall features of the spectral data cannot be taken as confirming the validity of the Schrodinger equation. Of particular interest is the probability distribution predicted by the Schrodinger theory, which differs markedly, particularly for the states of low n, from that of the Bohr theory. No experiment can measure the probability distribution directly. However, indirect evidence confirms the correctness of the Schrodinger amplitudes. For example, the magnetic interactions between electron and proton spin leads to a small correction to the energy levels. Calculation of this correction involves the average of $1/r^3$ for the state in question. When this is calculated from the Schrodinger probability amplitudes, the resulting energy shifts correctly explain the hyperfine structure of the spectral lines. Another property that depends on the probability amplitudes is the shift in the energy levels when the atom is placed in an external electric field (Stark effect). The magnitude of the Stark shift is likewise correctly calculated from the Schrodinger amplitudes.