The Theory of Fundamental Processes

CRC Press
Taylor & Francis Group

Richard P. Feynman

ARP

Advanced Book Classics

First published 1961 by Westview Press

Published 2018 by CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

CRC Press is an imprint of the Taylor & Francis Group, an informa business

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

ISBN 13: 978-0-201-36077-6 (pbk)

Cover design by Suzanne Heiser

Contents

The Theory of Fundamental Processes

1 Review of the Principles of Quantum Mechanics

These lectures will cover all of physics. Since we believe that the behavior of systems of many particles can be understood in terms of the interactions of a small number of particles, we shall be concerned primarily with the latter. Bearing in mind that the present theories need modifications or revision to account for observed phenomena, we shall want to consider the foundation of quantum mechanics in their most general form. This is so we can get some idea of the minimum assumptions (and their character) which we use to formulate those parts of the theory we use in dealing with the new phenomena of the strange particles.

A rough outline of the book follows: First, we discuss the ideas of quantum mechanics, mainly the concept of amplitudes, emphasizing that other things such as the combination laws of angular momenta are largely consequences of this concept. Next, briefly, relativity and the idea of antiparticles. Following this, we give a complete qualitative description of all the known particles and all that is known about the couplings between them. After that, we return to a detailed quantitative study of the two couplings for which calculations can be carried out today; namely, the -decay coupling and the electromagnetic coupling. The study of the latter is called quantum electrodynamics, and we shall spend most of our time with it.

Accordingly, we begin with a review of the principles of quantum mechanics. It has been found that all processes so far observed can be understood in terms of the following prescription: To every *process* there corresponds an *amplitude*[†]; with proper normalization the probability of the process is equal to the absolute square of this amplitude. The precise meaning of terms will become more clear from the examples that follow. Later we shall find rules for calculating amplitudes.

First, we consider in detail the double–slit experiment for electrons. A uniform beam of electrons of momentum p is incident on the double slit. To be more precise, we consider successive electrons, randomly distributed in the vertical direction (we prepare each electron with $p = p_x$, $p_y = p_z = 0$). (*Feynman*: They should come from a hole, at definite energy.)

When the electron hits the screen we record the position of the hit. The process considered is *thus*: An electron with well–defined momentum some–how goes through the slit system and makes its way to the screen (Fig. 1-1). Now we are not allowed to ask which slit the electron went through unless we actually set up a device to determine whether or not it did. *But then we would be considering a different process*! However we can relate the amplitude of the considered process to the *separate amplitudes* for the electron to have gone through slit (1), (a_1), and through slit (2), (a_2). [For example, when slit (2) is closed the amplitude for the electron to hit the screen is a_1 (prob. $|a_1|^2$) etc.] Nature gives the following simple rule: $a = a_1 + a_2$. This is a special case of the principle of superposition in quantum mechanics (cf. reference 1). Thus the probability of an electron reaching the screen is $P_a = |a|^2 = |a_1 + a_2|^2$. Clearly, in general we have $P_a \neq P_{a_1} + P_{a_2} (P_{a_1} = |a_1|^2, P_{a_2} = |a_2|^2)$, as distinguished from the classical case. We speak of "interference" between the probabilities (see reference 2). The actual form of P_a is familiar from optics.

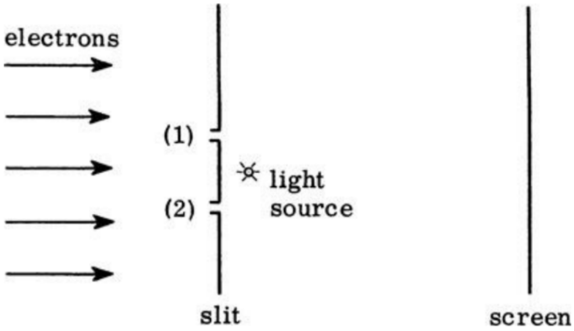


FIG. 1-1

Now suppose we place a light source between slits 1 and 2 (see Fig. 1-1) to find out which slit the electron "really" did go through (we observe the scattered photon). In this case the interference pattern becomes identical to that of the two slits considered independently. One way of interpreting this situation is to say that the act of measurement, of the position of the electron imparts an uncertainty in the momentum (ΔP_y), at the same time changing the phase of the amplitude in an uncontrollable way, so that the average over many electrons yields zero for the "interference" terms, owing to the randomness of the uncontrollable phases (see Bohm[3] for details of this view). However, we prefer the following viewpoint: By looking at the electrons we have actually changed the process under consideration. Now we must consider the photon and its interaction with the electron. So we consider the following amplitudes:

a11= amplitude that electron came through slit 1 and the photon was scattered behind slit 1a21= amplitude that electron came throug

The amplitude that an electron seen at slit 1 arrives at the screen is therefore $a' = a_{11} + a_{21}$; for an electron seen at slit 2, $a' = a_{12} + a_{22}$. Evidently for a properly designed experiment $a_{12} \equiv 0 \equiv a_{21}$ so that $a_{11} \equiv a_1$, $a_{22} \equiv a_2$ of the previous experiment. Now the amplitudes

a' and a' correspond to different processes, so the probability of an electron arriving at the screen is $P'_a = |a'|^2 + |a''|^2 = |a_1|^2 + |a_2|^2$.

Another example is neutron scattering from crystals.

(1) Ignore spin: At the observation point the total amplitude equals the sum of the amplitudes for scattering from each atom. One gets the usual Bragg pattern.

(2) Spin effects: Suppose all atoms have spin up, the neutrons spin down (assume the atom spins are localized): (a) no spin flip—as before, (b) spin flip—no diffraction pattern shown even though the energy and wavelengths of the scattered waves are the same as in case a. The reason for this is simply that the atom which did the scattering has its spin flipped down; in principle we can distinguish it from the other atoms. In this case the scattering from atom i is a *different process* from the scattering by atom $j \neq i$.

If instead of (localized) spin flip of the atom we excite (unlocalized) spin waves with wavenumber $k = k_{inc} - k_{scatt}$, we can again expect some partial diffraction effects.

Consider scattering at 90° in the c.m. system [see Fig. 1-2 (a to d)]:

(a) Two identical spinless particles: There are two indistinguishable ways for scatter to occur. Here, total amplitude = 2a and $P = 4|a|^2$, which is twice what we expected classically.

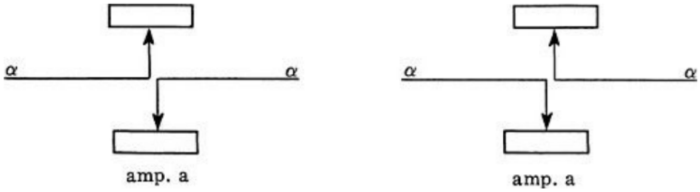


FIG. 1-2a

(b) Two distinguishable spinless particles. Here these processes are distinguishable, so that $P = |a|^2 + |a|^2 = 2|a|^2$.

(c) Two electrons with spin. Here these processes are distinguishable, so that $P = |a|^2 + |a|^2 = 2|a|^2$.

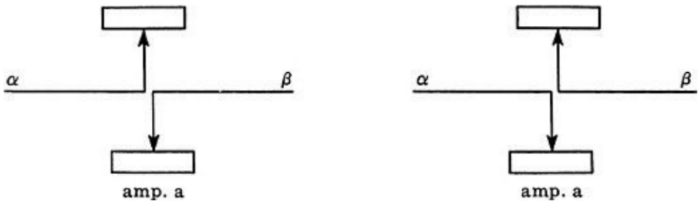


FIG. 1-2b

Problem 1-1: Suppose we have two sources of radio waves (e.g., radio stars) and need to know how far apart they are. We measure this intensity in two receivers at the same time and record the product of the intensities as a function of their relative position. This measurement of the correlation permits the required distance to be computed. With one receiver there is no pattern on the average, because the relative phase of A and B sources is random and fluctuating. For example, in Fig. 1-3 we have put the receivers at a separation corresponding to that of two maxima of the pattern if the relative, phase is 0 (Table 1-1). If L and R are at separation between a maximum and a minimum we have Table 1-2. Thus find the probability of reception of photon coincidence in the counters. Examine the effect of changing the separation between the receivers. Consider the process from the point of view of quantum mechanics.

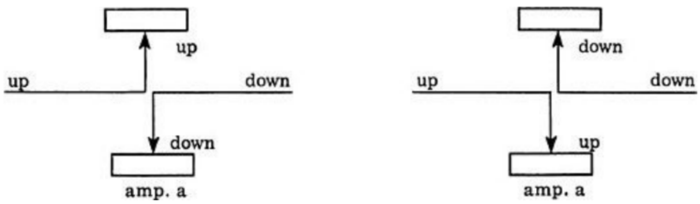


FIG. 1-2c

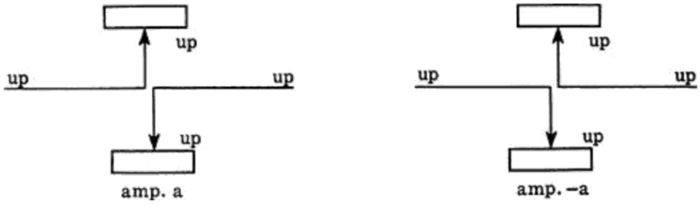


FIG. 1-2d

(d) But if both the incident electrons have spin up, the processes are indistinguishable. The total amplitude = a − a = 0. So here we have a new feature. We discuss this further in the next lecture.

FIG. 1-3

TABLE 1-1

Relative phases of sources	L (common)	R (max)	Product
0°	2	2	4
180°	0	0	0
90°	1	1	1
270°	1	1	1
			Av. = 1.5

TABLE 1-2

Relative phases of sources	L (common)	R (max)	Product
0°	2	0	0
180°	0	2	0
90°	1	1	1
270°	1	1	1
			Av. = 0.5

Discussion of Problem 1-1. There are four ways in which we can have photon coincidences:

(1) Both photons come from A: amp. a_1.

(2) Both photons come from B: amp. a_2.

(3) Receiver L receives photon from A, R from B: amp. a_3.

(4) Receiver L receives photon from B, R from A: amp. a_4.

Processes (1) and (2) are distinguishable from each other and from (3) and (4). However, (3) and (4) are indistinguishable. [For instance, we could, in principle, measure the energy content of the emitters to find which had emitted the photon in case (1) and (2).]

Thus, $P = |a|^2 + |a_2|^2 + |a_3 + a_4|^2$. The term $|a_3 + a_4|^2$ contains the interference effects. Note that if we were examining electrons instead of photons the latter term would be $|a_3 - a_4|^2$.

† A complex number.

2

Spin and Statistics

We should learn to think directly in terms of quantum mechanics. The only thing mysterious is why we must add the amplitudes, and the rule that P = |total amp.|^2 for a specific process. We return to consider the rules for adding amplitudes when the two alternative processes involve exchange of the two particles.

Consider a process P (amp. a) and the exchange process P_{ex} (amp. a_{ex}) (*indistinguishable from it*). We find the following remarkable rule in nature: For one class of particles (called bosons) the total amplitude is a + a_{ex}; for another class (fermions) the total amplitude is a $-a_{ex}$. It turns out that particles with spin 1/2, 3/2, … are fermions, and particles with spin 0, 1, 2, … are bosons. This is deducible from quantum mechanics plus relativity plus something else. This is discussed in the literature by Pauli[4] and, more recently, by Lüders and Zumino.[5]

It is important to notice that, for this scheme to work, we must know all the states of which the particle (or system) is capable. For example, if we did not know about polarization we would not understand the lack of interference for different polarizations. If we discovered a failure of any of our laws (e.g., for some new particle) we would look for some new degree of freedom to completely specify the state.

Degeneracy. Consider a beam of light polarized in a given direction. Suppose we put the axis of an analyzer (e.g., polaroid, nicol prism) successively in two perpendicular directions, x and y, to measure the number of photons of corresponding polarization in the beam (x and y are of course perpendicular to the direction of the beam). Call the amplitude for the arrival of a photon with polarization in the x direction a_x, in the y direction a_y. Now, if we rotate the analyzer 45°, what is the amplitude $a_{45°}$ for arrival of a photon in that direction? We find that a45° = (1/2)(ax +θ ay); for a general angle (from the x axis) we haθve a( ) = cos a_x + sin a_y. The point is that only two numbers (here a_x and a_y) are required to specify the amplitude for any polarization state. We shall find this result to be connected intimately with the fact that any other choice of axes is equally valid for the description of the photon.

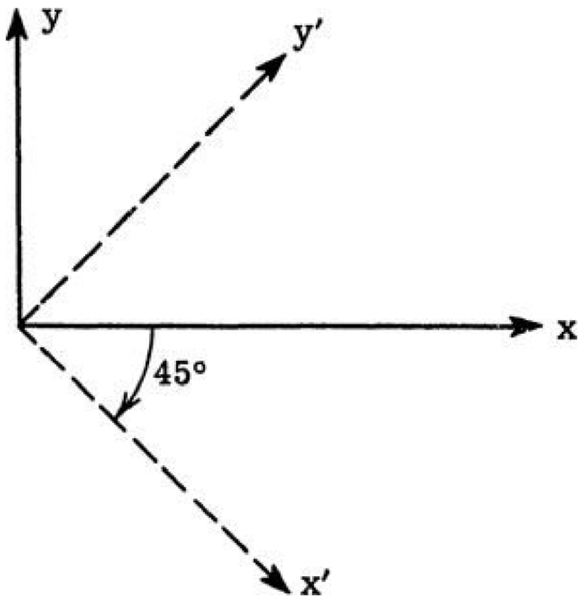


FIG. 2-1

For example (Fig. 2-1) consider the system of coordinates x', y' rotated −45° with respect to (x,y). An observer using this reference frame has

a′ x′ = (ax − ay)/2a′ y′ = (ax + ay)/2a′ 45°(in x′ , y′ = (a′ x′ + a′ y′)/2
= [(ax − ay)/2]= ax (as it should be !) + [(ax + ay)/2]

We could represent the state of the photon by a vector e = a_x**i** + a_y**j** in some two-dimensional space. Then the amplitude for the photon to be found with polarization in direction **v** = **i** cos + **j** sin is **e**·**v**.

The hypothesis that the behavior of a system cannot depend on the orientation in space imposes great restrictions on the properties of the possible states. Consider (Fig. 2-2) a nucleus or an atom which emits a ray preferably along the z axis. Now rotate everything, nucleus plus detecting apparatus. We should expect that the photon is emitted in the corresponding direction.

If the nucleus could be characterized by a single amplitude, say, its energy, then the ray would have to be emitted with equal likelihood in all directions. Why? Because otherwise we could set things up so that the ray comes out in the x direction (for we can always rotate the apparatus, the working system; and the laws of physics do not depend on the direction of the axis). This is a different condition because the subsequent phenomenon (emission) is predicted differently. One amplitude for our state cannot yield two predictions. The system must be described by more amplitudes. If the angular distribution is very sharp we need a large number of amplitudes to characterize the state of the nucleus.

FIG. 2-2

Suppose there are exactly n amplitudes which describe a system

$$(a_1 a_2 \cdots a_n)$$

Now the problem: Suppose we know it is in the state $a_1 = 1$, $a_2 \cdots = a_n = 0$. After rotation what are the amplitudes characterizing the system in the new coordinates ?

We define them as

$$(D_{11}(R) D_{21}(R) \cdots\cdots\cdots D_{n1}(R))$$

Similarly if it starts in the state $a_2 = 1$, $a_1 = a_3 = \cdots = a_n = 0$, we have

$$(D_{12}(R) D_{22}(R) \cdots\cdots\cdots D_{n2}(R))$$

Therefore we need an entire matrix $D_{ij}(R)$.

A more complicated case occurs if initially the system is in a state

$$(a_1 a_2 \cdots a_n)$$

After the rotation the new state is

$$(a'_1 a'_2 \cdots a'_n)$$

whereas $a'_j = \Sigma_j D_{ij}(R) a_j$. Think about why this is so.

3 Rotations and Angular Momentum

In the last lecture we spoke about an apparatus that produced an object in condition a:

$$a = \begin{pmatrix} a_1 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{pmatrix}$$

This requires further explanation; since we have introduced so far only the concept of an amplitude for the complete event: the production and detection of the object. This amplitude can be obtained as follows:

We assume that we have an amplitude b_i that the object produced is in some condition characterized by the index i. If it is in this condition, i, let a_i be the amplitude that it will activate some detector. Then the amplitude for the complete event (production and detection) is $a_i\, b_i$, summed over the possible intermediate conditions i.

Consider again the experiment of an electron passing through two slits (Fig. 3-1). If $a_{1\to 3}$ is the amplitude for an electron to go through one slit and $a_{3\to 2}$ the amplitude for an electron at this slit to reach the screen at *2*, then the amplitude for the complete event is the product $a_{1\to 3} \times a_{3\to 2}$.

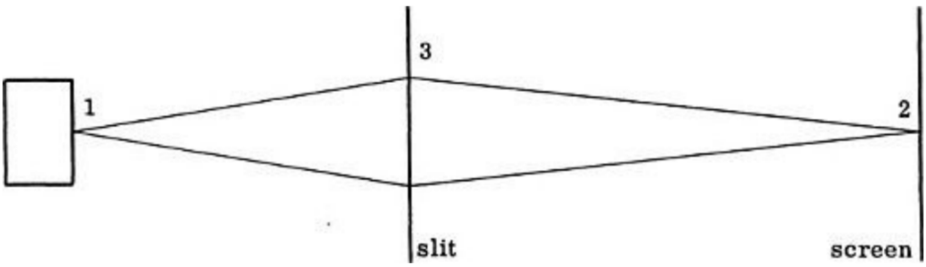
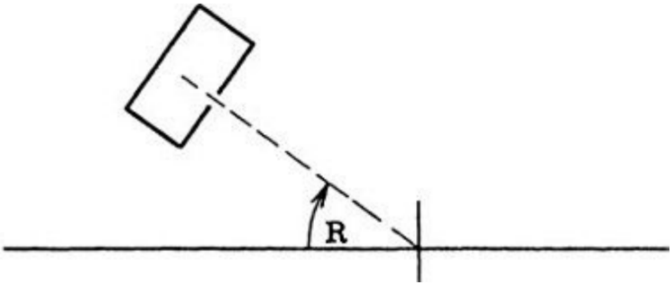


FIG. 3-1

Now rotate the apparatus through **R**(|**R**| = angle of rotation, **R**/|R| = axis of rotation) so that the object is produced in condition a′ with respect to the fixed detector

$$a' = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \cdot \\ \cdot \\ a_n \end{pmatrix}$$



We have pointed out that this must be related to the a by an equation of the form a′ = D(R)a, where the matrix D(R) does not depend on the particular piece of apparatus. In another experiment (Fig. 3-2) we could have the same object produced in some other conditions b and b′ . Then b’ = D(R)b, and the same D(R) is expected. Why must this relation be linear? Because objects can be made to interfere. Suppose we have two pieces of apparatus, one producing an object in condition a, the other producing the same object in condition b, and together producing it in condition a + b. After rotation we would have a′ , + b′ , and also a′ + b′ , and also a′ + b′ , in order that the interference phenomena appear the same way in the rotated system. Then we have

FIG. 3-2

$$a' = D(R)a \quad b' = D(R)b \quad (a + b)' = D(R)(a + b)$$

but (a + b)′ = a′ + b′ , therefore D(R)(a + b) = D(R)a + D(R)b.

What else can we deduce?

Suppose we consider the apparatus that we rotated through R as a new apparatus, which produces the object in condition a′ . Now we rotate it through S, as shown in Fig. 3-3. According to our rule, the object is now produced in a condition a′ , where a′ = D(S)a′ . Since a′ = D(R)a, we have a′ = D(S)D(R)a, which means D(SR) =D(S)D(R).[†]

Rotations form a group, and the D’s are matrix representations of this group. It is by no means self-evident how to find them.

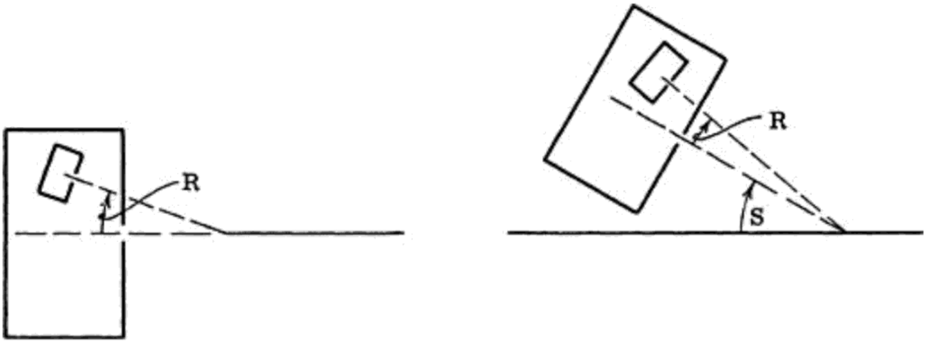


FIG. 3-3

Examples:

(1) An object represented by a single complex number. The D's are 1 × 1 matrices, i.e., a complex number can be chosen to be 1.

(2) An object represented by a vector, hence by three amplitudes, the x,y,z components of the vector. The D's are the familiar matrices relating rotated coordinates.

Let us now go to the general analysis. Suppose we know a matrix for an infinitesimal rotation. Say, the rotation of 1° about the z axis. Then the rotation n° about the z axis is represented by

$$D(n^\circ \text{ around } z) = [\, D(1^\circ \text{ around } z) \,]n$$

More generally, if we know D(ε ° around z), then

$$D(\theta \text{ around } z) = [\, D(\varepsilon \text{ around } z) \,]\, \theta/\epsilon$$

Now, if we rotate just a little we have approximately the identity, so to first order in ε, D(ε around z) =1 + i ε M_Z. Also,

$$D(\varepsilon \text{ around } x) = 1 + i\varepsilon \; MxD(\varepsilon \text{ around } y) = 1 + i\varepsilon \; My$$

Now, we have D(around z) = $(1 + i\varepsilon \; M_Z)^{\theta/\epsilon}$ and using the binomial expansion, one obtains, when $\varepsilon \to 0$,

$$D(\theta \text{ around } z) = 1 + i\theta \; Mz - \theta 22! \; Mz2 - i \; \theta 33! \; Mz3 + \cdots$$

which is often written $e^{i\theta M}$z. The binomial expansion works, since M_Z behaves like ordinary numbers under addition and multiplication.

If we want to rotate ε about an axis along the unit vector **v**, we find

$$D(\varepsilon \text{ around } v) = 1 + i\varepsilon\,(vxMx + vyMy + vzMz)$$

and for a finite about **v**,

$$D(\theta \text{ around } v) = \exp[\, i\theta \; (vxMx + vyMy + vzMz)\,]$$

But now we must be careful about the relative order of M_x, M_y, and M_z in the matrix products that appear in the series; these matrices do not commute, This follows from the fact that finite rotations do not commute. Consider the rotation of an eraser, Fig. 3-4 (a and b). (1) Rotate it 90° about the z axis and then 90° about the x axis (Fig. 3-4a); (2) rotate it 90° about the x axis, and then 90° about the z axis (Fig. 3-4b); and we get two entirely different results.

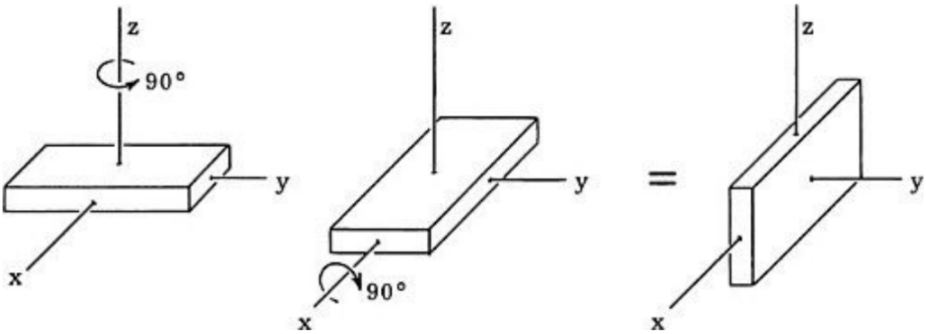


FIG. 3-4a

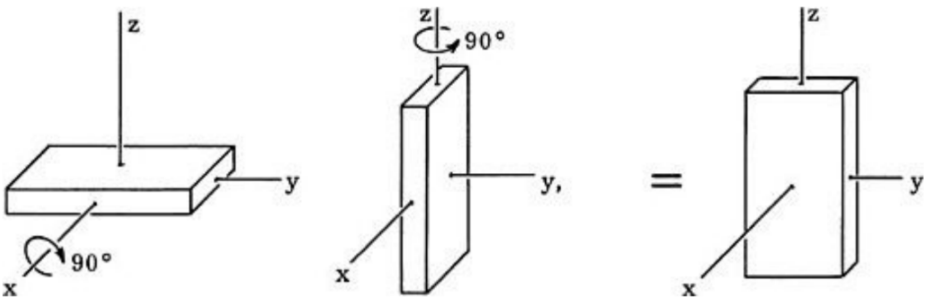


FIG. 3-4b

Let us discover the commutation relations between M_x and M_y. We consider a rotation ε about the x axis, followed by about the y axis, then $-\varepsilon$ about the x axis and $-$ about the y axis as in Fig. 3-5.

We follow the motion of a point starting on the y axis. Clearly the result is a second-order effect. It ends up just displaced by about ε toward the x axis. We note also that a point which starts on the z axis returns to the origin, and therefore the net displacement of the point on the sphere is just a rotation by an angle ε about the z axis. Keeping terms up to the second order, we have

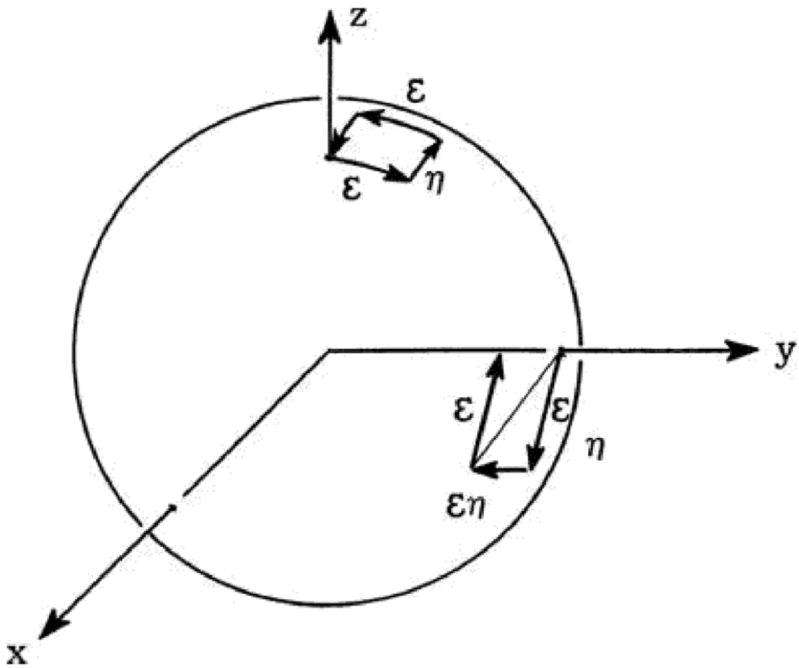


FIG. 3-5

$$[1 - i\eta\, My - \eta 2\,(1/2)My2]\,[1 - i\varepsilon\, Mx - \varepsilon 2(1/2)Mx2]\,[1 + i\eta\, My - \eta 2(1/2)My2] \times [1 + i\varepsilon\, Mx - \varepsilon 2(1/2)Mx2] = 1 + i\varepsilon\,\eta\, Mz$$

Collecting coefficients of ε we find

$$MxMy - MxMy = iMz$$

Similarly,

$$MyMz - MzMy = iMxMzMx - MxMz = iMy$$

These are the rules of commutation for the matrices M_x, M_y, and M_z. Everything else can be derived from these rules. How this is done is given in detail in many books (e.g., Schiff). We give only a bare outline here. First we prove that $Mx2 + My2 + Mz2 = M2$ commutes with all M's. Then we can choose our a's so that they satisfy $M^2a = ka$, where k is some number. Construct

$$M- = Mx - iMy$$

and note

$$Mz\ M- = M-(Mz - 1)$$

Now, suppose $a^{(m)}$ satisfies

$$M_z a^{(m)} = m a^{(m)}$$

where m is another number; then

$$M_z b = M_z M_- a^{(m)} = M_-(M_z - 1)a^{(m)} = (m - 1)M_- a^{(m)} = (m - 1)b$$

Therefore,

$$b = c a^{(m-1)}$$

We normalize $a^{(m)}$ to unity; i.e.,

$$\sum_{j=1}^{n} a_j^{(m)*} a_j^{(m)} = 1 \quad \text{for all } m$$

Therefore,

$$1 = (1/c^*c) \sum_n (M_- a^{(m)})_n^* (M_- a^{(m)})_n = (1/c^*c) \sum_n a_n^{(m)*} (M_+ M_-) a^{(m)}$$

where $M_+ = M_x + iM_y$. Now

$$M_+ M_- = M_x^2 + M_y^2 + M_z = M^2 - M_z^2 + M_z$$

and

$$M^2 a^{(m)} = k a^{(m)}$$

Therefore

$$c = [k - m(m - 1)]^{1/2}$$

Let m = –j be the “last” state. How can we fail to get another if we operate by M_- ? Only if $M_- a^{(-j)} = 0$ or c = 0 for m = –j, so k = –j(–j – 1) =j(j + 1).

The same kind of steps (using M_+, which raises m by one, just like M_- lowers it) prove that if the largest value of m is +j′ , then k = j′ (j′ + 1), so that j = j′ . Hence 2j′ is an integer. The total number of states is 2j + 1.

Examples:

(1) 1 state: j = 0

(2) 3 states: j = 1

m	Transforms like
1	12 (x + iy)
0	z
–1	12 (x − iy)

(3) 2 states: $j = 1/2$. This is a very interesting case. Let

$$a(1/2) = (10) a(-1/2) = (01)$$

Using our general results we obtain

$$M_- \ (10) = (01)$$

since

$$[\, j(j+1) - m(m-1) \,]1/2 = [\, (1/2)(3/2) - (1/2)(1/2) \,]1/2 = 1 M_-(01) = 0$$

Therefore,

$$M_- = (0010)$$

Likewise,

$$M_z \ (10) = 1/2 \ (10) M_z \ (01) = -1/2 \ (01)$$

Therefore,

$$M_z = 1/2(10 - 01)$$

Similarly we can show that

$$M_+ = (0100)$$

so that we can write

$$M_x = 1/2 \ (0110) = (1/2)\sigma x M_y = 1/2 \ (0-ii0) = (1/2)\sigma y M_z = 1/2 \ (100-1) = (1/2)\sigma z$$

The above expressions also serve as the definition of the three important 2 × 2 matrices, the Pauli matrices σ_x, σ_y, σ_z. Check also that $\sigma x2 = \sigma y2 = \sigma z2 = 1$, $\sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z$. The main point of this is, that it all came out of nothing: that nature has no preferred axis and the nature of the principle of superposition were the only assumptions invoked.

However, we have made a very important hypothesis: We have assumed that the processes of production and detection are *well separated* and that in between one can talk of an amplitude that characterizes the object. This hypothesis has always been made (particularly in field theory) no matter how small the distance between the apparatus and the detector. It may turn out that it is not valid if these are too close together.

Another important assumption was to disregard any dynamic interference: There are no forces between our producing and measuring apparatus at least that are not describable by transfer of our object between them. An amplitude for two independent events is then also the product of the amplitude for each separate event.

Look at the example of the two stars A, B and the counters X, Y (Fig. 3-6). If $a_{B \to x}$ is the amplitude for the photon emitted at B to reach counter X and $a_{A \to y}$ is the corresponding amplitude for the photon emitted at A to reach counter Y, then $a = a_{B \to x} \times a_{A \to y}$ is the amplitude for occurrence of both events.

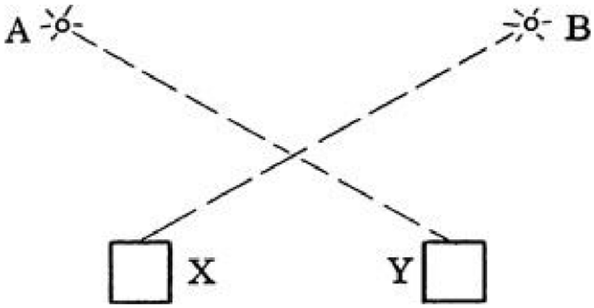


FIG. 3-6

† Strictly speaking, we cannot prove that the amplitudes after rotation must be the same in both cases; only the squares must be the same. The amplitudes could differ by a phase factor. However, Wigner has shown that it could always be eliminated by redefining the D's.

4 Rules of Composition of Angular Momentum

A spin 1/2 state is characterized by two amplitudes. In general a = a_+(1/2) + a_-(–1/2) where (1/2) stands for (10), (−1/2) for (01), and a for (a+a−).

For instance, the solution of

$$Mxa = (1/2)a$$

corresponding to spin up along the x axis is

$$a = (1/2)(1/2) + (12)(-1/2)$$

Also, down in x, (1/2)(1/2) − (12)(−1/2); up in y, (1/2)(1/2) + (i2) × (−1/2); down in y, (1/2)(1/2) − (i2) (−1/2). In fact, it can be shown that every state represents spin in some direction.

Any system that has two complex numbers has an analogy in spin 1/2. For instance let us consider the polarization of light. Let x polarization be spin up and y polarization be spin down along an axis ζ in a “crazy” three-dimensional space. The other two axes we label ξ and η. Then spin up along ξ = 45° polarization; down, ξ = −45° polarization; up, η = RHC (right–hand circular polarization); down, η = LHC (left-hand circular polarization). If we draw a unit sphere centered at the origin of this space (Fig. 4-1), every state of polarization is represented by a point on it.

A general direction corresponds to elliptical polarization. Passing light through a 1/4-wave plate is a certain rotation. The connection between the polarization of light and direction in a three-dimensional space was exploited long ago by Stokes. It is very useful to understand certain processes, for example, masers. (The maser is a device using a system, the ammonia molecule, making transitions between two states under the influence of electric fields. Its analysis can be more easily understood by representing the state of the ammonia molecule at any time as a direction in some three-dimensional space, analogous to the ordinary space for a spin-1/2-electron.)

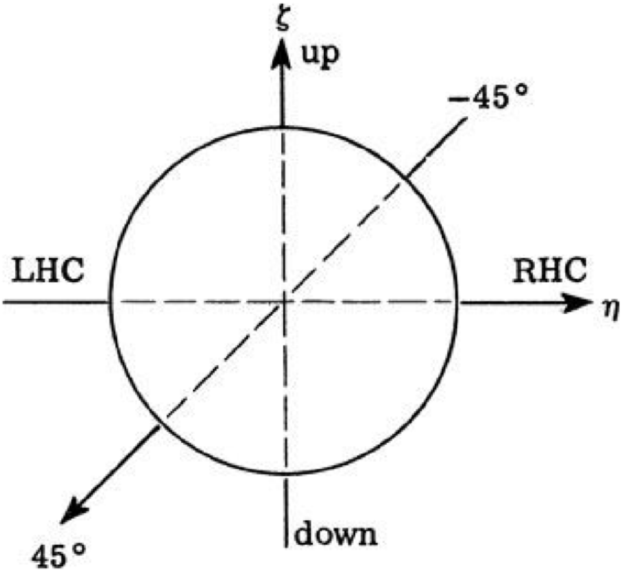


FIG. 4-1

Rules of Composition of Angular Momentum. Consider an apparatus that produces two particles which we label A and B. Suppose particle A has spin 1 and exists in three states with m = +1, 0, -1; and that particle B has spin 1/2 and exists in two states with m = +1/2 and -1/2. For each of A's three states, B can have two, so there are six possible states of the two particles together.

We may be thinking of an electron revolving around a nucleus. How do we characterize the combined system? We have matrices M_A and M_B which operate on the states $_A$ and $_B$. Then

$$(1 + i\varepsilon M_z)\psi_A\psi_B = (1 + i\varepsilon M_{zA})\psi_A(1 + i\varepsilon M_{zB})\psi_B = 1 + i\varepsilon(M_{zA} + M_{zB})\psi_A\psi_B$$

or[†]

$$M_z = M_{zA} + M_{zB}$$

The states of the combined system are given in Table 4-1. There are six states and one could jump to the conclusion that j = 5/2. However, there is no value of m = ±5/2 and also m = ±1/2 appears twice.

Actually $M^2 = (\mathbf{M}_A + \mathbf{M}_B)^2$ has two values for j:

$$j = 3/2 \quad m = 3/2, \ -1/2, \ -3/2$$

and

$$j = 1/2 \quad m = 1/2, \ -1/2$$

TABLE 4-1

m_A	m_B	m
1	1/2	3/2
0	1/2	1/2
1	–1/2	1/2
0	–1/2	–1/2
–1	1/2	–1/2
–1	–1/2	–3/2

Clearly the state j = 3/2, m = 3/2 is (+ 1)(1/2). But which state corresponds to j = 3/2, m = 1/2? Recall

M−(m) = [j(j + 1) − m(m − 1)]1/2 (m − 1)

We have

M− = (M−A + M−B)M−(1/2) = (−1/2)M−(−1/2) = 0M−(1) = 2 (0)M−(0) = 2 (−1)M−(−1) = 0

Then

M−(1)(1/2) = 2 (0) (1/2) + (1)(−1/2)

and

M−(3/2, 3/2) = 3 (3/2, 1/2)

Therefore,

(3/2, 1/2) = (2/3)(0) (1/2) + (1/3)(1) (−1/2)

The state (1/2, 1/2) is obtained by forming the linear combination of (0)(1/2) and (1)(–1/2), which is orthogonal to (3/2, 1/2). We obtain the results given in Table 4-2.

TABLE 4-2

m	j = 3/2	j = 1/2
3/2	(1)(1/2)	
1/2	(2/3)(0)(1/2) + (1/3)(1) (−1/2)	(1/3)(0)(1/2) − (2/3)(1) (−1/2)
–1/2	(2/3)(0)(−1/2) + (−1) (1/2)	−(1/3)(0)(−1/2) + (2/3) (−1)(1/2)
–3/2	(–1)(–1/2)	

More examples: Add two spin = 1/2 states (Table 4-3) under exchange of spins. Now add two spin = 1 states (Table 4-4). For the addition of two equal angular momentum the biggest state is symmetric, the next antisymmetric, and so on.

TABLE 4-3