Physics 221A
Academic Year 2020–21
Notes 11
Rotations in 3-Dimensional Space†

1. Introduction

The theory of rotations is of direct importance in all areas of atomic, molecular, nuclear and particle physics, and in large areas of condensed matter physics as well. The ideas developed in this and subsequent sets of notes on rotations in quantum mechanics will recur many times in this course, continuing through the second semester. The rotation group is the first nontrivial symmetry group we encounter in a study of quantum mechanics, and serves as a paradigm for other symmetry groups one may encounter later, such as the $SU(3)$ symmetry that acts on the color degrees of freedom in quantum chromodynamics. Furthermore, transformations that have the same mathematical form as rotations but which have nothing to do with rotations in the usual physical sense, such as isotopic spin transformations in nuclear physics, are also important. Rotations are also the spatial part of Lorentz transformations, and Lorentz invariance is one of the basic principles of modern physics. We will study Lorentz transformations in quantum mechanics in Notes 48.

These notes will deal with rotations in ordinary 3-dimensional space, such as they would be used in classical physics. We will deal with quantum representations of rotations in Notes 12 and 13.

2. Inertial Frames

The concept of an inertial frame is the same in Newtonian (nonrelativistic) mechanics and in special relativity, both in classical mechanics and in quantum mechanics. It requires modification only in general relativity, that is, when gravitational fields are strong.

We define an inertial frame as a frame in which free particles (particles upon which no forces act) move in straight lines with constant velocity. An inertial frame has three mutually orthogonal axes, with respect to which coordinates of points can be measured. It assumed that the measurements of distance obey the rules of Euclidean geometry in three-dimensional space. This is a physical assumption that can be tested experimentally. In fact, the geometry of space is not exactly Euclidean, due to the effects of gravitational fields; but in weak fields the deviations are very small and we shall ignore them. It is another physical assumption that there exists a frame in which free particles move

† Links to the other sets of notes can be found at:
`http://bohr.physics.berkeley.edu/classes/221/2021/221.html`.

on straight lines with constant velocity, but given the existence of one such frame, a whole family of other inertial frames may be generated by applying translations, rotations and boosts. Rotating and/or accelerating frames are excluded by the requirement regarding the uniform, rectilinear motion of free particles.

3. Rotation Operators in Ordinary Space and Rotation Matrices

Let O be the origin of an inertial frame, and let "points of space" be identified by their coordinates with respect to this frame. Points of space can also be identified by the vectors from O to the points in question, as with vectors $\mathbf{r}$ and $\mathbf{r}'$ in Fig. 1.

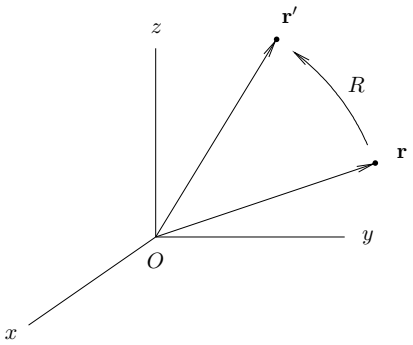

Fig. 1. A rotation R is defined as a map from space to itself that leaves one point O invariant, while preserving all distances between points.

We define a *rotation* R on ordinary space as a map that takes points of space into other points of space in such a way that O is mapped into itself and all distances are preserved. In Fig. 1, map R takes point $\mathbf{r}$ into point $\mathbf{r}'$, so we can write

$$\mathbf{r}' = R\mathbf{r}. \tag{1}$$

Since R preserves distances it maps straight lines into straight lines and it satisfies

$$R(a\mathbf{r}) = a(R\mathbf{r}). \tag{2}$$

Also, since angles can be expressed in terms of distances, R preserves angles, too. This means that R maps parallelograms into congruent parallelograms, and since vector addition can be expressed in terms of parallelograms, it follows that

$$R(\mathbf{r}_1 + \mathbf{r}_2) = (R\mathbf{r}_1) + (R\mathbf{r}_2). \tag{3}$$

Combined with Eq. (2), this means that R is a linear operator on ordinary space, regarded as a vector space. It also follows that R is invertible, and thus R^{-1} exists. This is because a linear operator is invertible if the only vector it maps into the zero vector $\mathbf{0}$ is $\mathbf{0}$ itself, something that is satisfied in this case because $\mathbf{0}$ is the only vector of zero length.

The product of two rotations is denoted $R_1 R_2$, which means, apply R_2 first, then R_1. Rotations do not commute in general, so that $R_1 R_2 \neq R_2 R_1$, in general. It follows from the definition that if R, R_1 and R_2 are rotation operators, then so are R^{-1} and $R_1 R_2$. Also, the identity operator is a rotation. These facts imply that the set of rotation operators R forms a group.

We can describe rotations in coordinate language as follows. We denote the mutually orthogonal unit vectors along the axes of the inertial frame by $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$ or $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$. We assume the frame is right-handed. Vectors such as $\mathbf{r}$ or $\mathbf{r}'$ can be expressed in terms of their components with respect to these unit vectors,

$$\mathbf{r} = \sum_i \hat{\mathbf{e}}_i \, x_i, \qquad \mathbf{r}' = \sum_i \hat{\mathbf{e}}_i \, x'_i. \tag{4}$$

Similarly, we define the *rotation matrix* R associated with the operator R by

$$R_{ij} = \hat{\mathbf{e}}_i \cdot (R\hat{\mathbf{e}}_j), \tag{5}$$

where R_{ij} are the components of the matrix R. (In these notes, we use sans serif fonts for matrices, but ordinary italic fonts for their components.) The definition (5) can be cast into a more familiar form if we use a round bracket notation for the dot product,

$$\mathbf{A} \cdot \mathbf{B} = (\mathbf{A}, \mathbf{B}), \tag{6}$$

where $\mathbf{A}$ and $\mathbf{B}$ are arbitrary vectors. Then the definition (5) can be written,

$$R_{ij} = (\hat{\mathbf{e}}_i, R\hat{\mathbf{e}}_j). \tag{7}$$

This shows that R_{ij} are the matrix elements of R with respect to the basis $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$ in the much same manner in which we define matrix elements of operators in quantum mechanics.

The definition (7) provides an association between geometrical rotation operators R and matrices R which depends on the choice of coordinate axes $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$. If R, R_1 and R_2 are rotation operators corresponding to R, R_1 and R_2, then $R_1 R_2$ corresponds to $\mathsf{R}_1\mathsf{R}_2$ (multiplication in the same order) and R^{-1} corresponds to R^{-1}. As we say, the matrices R form a *representation* of the geometrical rotation operators R.

Assuming that $\mathbf{r}$ and $\mathbf{r}'$ are related by the rotation R as in Fig. 1 or Eq. (1), we can express the components x'_i of $\mathbf{r}'$ in terms of the components x_i of $\mathbf{r}$. We have

$$x'_i = \hat{\mathbf{e}}_i \cdot \mathbf{r}' = \hat{\mathbf{e}}_i \cdot (R\mathbf{r}) = \hat{\mathbf{e}}_i \cdot \Big(R \sum_j \hat{\mathbf{e}}_j x_j \Big) = \sum_j R_{ij} \, x_j, \tag{8}$$

where we use the linearity of R and the definition (7) of R_{ij}. We can write this in matrix-vector notation as

$$\mathbf{r}' = \mathsf{R}\mathbf{r}, \tag{9}$$

which looks just like Eq. (1) except that now $\mathbf{r}$, $\mathbf{r}'$ are seen, not as geometrical vectors as in Fig. 1, but rather as triplets of numbers, that is, the coordinates of the old and new points with respect to the basis $\hat{\mathbf{e}}_i$.

4. Active versus Passive Point of View

When dealing with rotations or any other symmetry group in physics, it is important to keep distinct the active and passive points of view. In this course we will adopt the active point of view unless otherwise noted. The active point of view is used by Messiah, *Quantum Mechanics* and Sakurai, *Modern Quantum Mechanics*, but many other books take the passive point of view, including some standard monographs on rotations, such as Edmonds, *Angular Momentum in Quantum Mechanics*. Books on classical mechanics usually take the passive point of view. The active point of view is usually preferable, because it is more amenable to an abstract or geometrical treatment, whereas the passive point of view is irrevocably chained to coordinate representations.

In the active point of view, we usually imagine one coordinate system, but think of operators that map old points into new points. Then an equation such as $\mathbf{r}' = \mathsf{R}\mathbf{r}$ indicates that $\mathbf{r}$ and $\mathbf{r}'$ are the coordinates of the old and new points with respect to the given coordinate system. This is the interpretation of Eq. (9) above. In the active point of view, we think of rotating our physical system but keeping the coordinate system fixed.

In the passive point of view, we do not rotate our system or the points in it, but we do rotate our coordinate axes. Thus, in the passive point of view, there is only one point, but two coordinate systems. To incorporate the passive point of view into the discussion above, we would introduce the *rotated frame*, defined by

$$\hat{\mathbf{e}}'_i = R\hat{\mathbf{e}}_i, \tag{10}$$

and then consider the coordinates of a given vector with respect to the two coordinate systems and the relations between these components. In a book that adopts the passive point of view, an equation such as $\mathbf{r}' = \mathsf{R}\mathbf{r}$ probably represents the coordinates $\mathbf{r}$ and $\mathbf{r}'$ of a single point with respect to two (the old and new) coordinate systems. With this interpretation, the matrix R has a different meaning than the matrix R used in the active point of view [such as that in Eq. (9) and elsewhere in these notes], being in fact the inverse of the latter. Therefore caution must be exercised in comparing different references. In this course we will make little use of the passive point of view.

5. Properties of Rotation Matrices; the Group $O(3)$

Since the rotation R preserves lengths, we have

$$|\mathbf{r}'|^2 = |\mathbf{r}|^2, \tag{11}$$

when $\mathbf{r}$, $\mathbf{r}'$ are related by Eq. (9). Since this is true for arbitrary $\mathbf{r}$, we have

$$\mathsf{R}^t\mathsf{R} = \mathsf{I}. \tag{12}$$

This is the definition of an orthogonal matrix. The set of all 3×3 real orthogonal matrices is denoted $O(3)$ in standard notation, so our rotation matrices belong to this set. In fact, since every orthogonal matrix in $O(3)$ corresponds to a rotation operator R by our definition, the space of

rotations is precisely the set $O(3)$. The set $O(3)$ forms a group under matrix multiplication that is isomorphic to the group of geometrical rotation operators R introduced above.

Equation (12) implies

$$\mathsf{R}^t = \mathsf{R}^{-1}, \tag{13}$$

so that it is easy to invert an orthogonal matrix. This in turn implies

$$\mathsf{R}\mathsf{R}^t = \mathsf{I} \tag{14}$$

[in the reverse order from Eq. (12)]. Taken together, Eqs. (12) and (14) show that the rows of an orthogonal matrix constitute a set of orthonormal vectors, as do the columns.

6. Proper and Improper Rotations; the Group $SO(3)$

Taking determinants, Eq. (12) also implies $(\det \mathsf{R})^2 = 1$, or

$$\det \mathsf{R} = \pm 1. \tag{15}$$

Orthogonal matrices for which $\det \mathsf{R} = +1$ are said to be *proper* rotations, while those with $\det \mathsf{R} = -1$ are said to be *improper*. Proper rotations have the property that they preserve the sense (right-handed or left-handed) of frames under the transformation (10), while improper rotations reverse this sense.

The set of proper rotations by itself forms a group, that is, the property $\det \mathsf{R} = +1$ is preserved under matrix multiplication and inversion. This group is a subgroup of $O(3)$, denoted $SO(3)$, where the S stands for "special," meaning in this case $\det \mathsf{R} = +1$. The set of improper rotations does not form a group, since it does not contain the identity element. An improper rotation of some importance is

$$\mathsf{P} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = -\mathsf{I}, \tag{16}$$

which inverts vectors through the origin.

For the duration of these notes we will restrict consideration to proper rotations whose matrices belong to the group $SO(3)$, but we will return to improper rotations in Notes 21.

7. Rotations About a Fixed Axis; the Axis-Angle Parameterization

Let us consider now a proper rotation which rotates points of space about a fixed axis, say, $\hat{\mathbf{n}}$, by an angle θ, in which the sense is determined by the right-hand rule. See Fig. 2. We will denote this rotation by $R(\hat{\mathbf{n}}, \theta)$ and the matrix by $\mathsf{R}(\hat{\mathbf{n}}, \theta)$. It is geometrically obvious that rotations about a fixed axis commute,

$$\mathsf{R}(\hat{\mathbf{n}}, \theta_1)\mathsf{R}(\hat{\mathbf{n}}, \theta_2) = \mathsf{R}(\hat{\mathbf{n}}, \theta_2)\mathsf{R}(\hat{\mathbf{n}}, \theta_1) = \mathsf{R}(\hat{\mathbf{n}}, \theta_1 + \theta_2), \tag{17}$$

and the angles add under matrix multiplication as indicated. The rotations about the three coordinate axes are of special interest; these are

$$\mathsf{R}(\hat{\mathbf{x}},\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix},$$
$$\mathsf{R}(\hat{\mathbf{y}},\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \tag{18}$$
$$\mathsf{R}(\hat{\mathbf{z}},\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

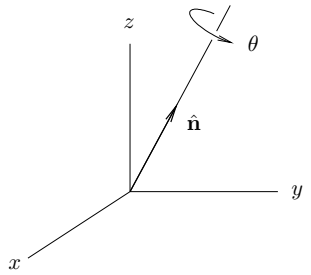

Fig. 2. A proper rotation is specified by a unit vector $\hat{\mathbf{n}}$ that defines an axis, and a rotation of angle θ about that axis. The rotation follows the right-hand rule.

One can show that any proper rotation can be represented as $\mathsf{R}(\hat{\mathbf{n}},\theta)$, for some axis $\hat{\mathbf{n}}$ and some angle θ. Thus, there is no loss of generality in writing a proper rotation in this form. We will call this the *axis-angle parameterization* of the rotations. This theorem is not totally obvious, but the proof is not difficult (see Prob. 4). It is based on the observation that every proper rotation must have an eigenvector with eigenvalue $+1$. This eigenvector or any multiple of it is invariant under the rotation and defines its axis.

8. Infinitesimal Rotations

A rotation that is close to the identity is called *near-identity* or *infinitesimal*. It has the form

$$\mathsf{R} = \mathsf{I} + \epsilon\mathsf{A}, \tag{19}$$

where ϵ is a small scale factor whose purpose is to make the correction term small, and where A is a matrix. By substituting Eq. (19) into the orthogonality condition (12), we find

$$\mathsf{A} + \mathsf{A}^t = 0, \tag{20}$$

that is, the matrix A is antisymmetric.

A convenient parameterization of the antisymmetric matrices is given by

$$\mathsf{A} = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} = \sum_{i=1}^{3} a_i \mathsf{J}_i, \tag{21}$$

where $(\mathsf{J}_1, \mathsf{J}_2, \mathsf{J}_3)$ is a "vector" of matrices, defined by

$$\mathsf{J}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \mathsf{J}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \qquad \mathsf{J}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{22}$$

Equations (22) can be summarized by writing

$$(\mathsf{J}_i)_{jk} = -\epsilon_{ijk}. \tag{23}$$

We will also write the sum in Eq. (21) as $\mathbf{a} \cdot \mathbf{J}$, that is, as a dot product of vectors, but we must remember that $\mathbf{a}$ is a triplet of ordinary numbers, while $\mathbf{J}$ is a triplet of matrices. The notation is similar to that which we use with the Pauli matrices, a "vector" of 2×2 matrices. See Prob. 1.1.

Equation (21) associates an antisymmetric matrix A with a corresponding 3-vector $\mathbf{a}$. In the following we will find it convenient to switch back and forth between these representations.

A useful property of the $\mathbf{J}$ matrices is the following. Let A be an antisymmetric matrix, associated with a vector $\mathbf{a}$ according to Eq. (21), and let $\mathbf{u}$ be another vector. Then

$$\mathsf{A}\mathbf{u} = (\mathbf{a} \cdot \mathbf{J})\mathbf{u}. \tag{24}$$

Now taking the i-th component of the right-hand side and using Eq. (23), we find

$$[(\mathbf{a} \cdot \mathbf{J})\mathbf{u}]_i = (\mathbf{a} \cdot \mathbf{J})_{ij}\, u_j = -a_k\, \epsilon_{kij}\, u_j = +\epsilon_{ikj}\, a_k u_j = (\mathbf{a}\times\mathbf{u})_i, \tag{25}$$

where we use the summation convention. In other words,

$$\mathsf{A}\mathbf{u} = (\mathbf{a} \cdot \mathbf{J})\mathbf{u} = \mathbf{a}\times\mathbf{u}. \tag{26}$$

This gives us an alternative way of writing the cross product (as a matrix multiplication) that is often useful.

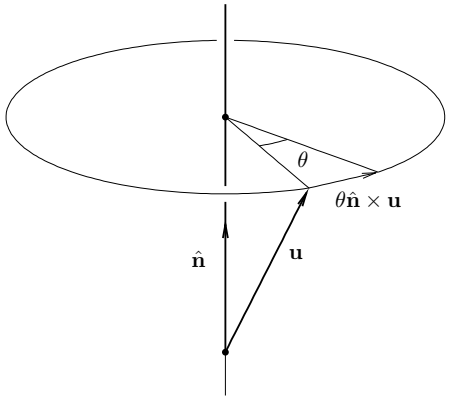

Fig. 3. The effect of a near-identity rotation of small angle θ about axis $\hat{\mathbf{n}}$ on a vector $\mathbf{u}$ is to move it by the amount $\theta\hat{\mathbf{n}}\times\mathbf{u}$.

Since every proper rotation has an axis and an angle, we can ask, what are the axis and angle of the near-identity rotation (19)? To find out, we write

$$\mathsf{R} = \mathsf{I} + \epsilon\mathsf{A} = \mathsf{I} + \epsilon(\mathbf{a}\cdot\mathbf{J}), \tag{27}$$

which implies

$$\mathsf{R}\mathbf{u} = \mathbf{u} + \epsilon\mathbf{a}\times\mathbf{u}, \tag{28}$$

where $\mathbf{u}$ is an arbitrary vector. But another way of writing this equation appears when we examine Fig. 3, which shows the action of a rotation about axis $\hat{\mathbf{n}}$ by small angle θ on the same vector $\mathbf{u}$. The result is

$$\mathsf{R}\mathbf{u} = \mathbf{u} + \theta\hat{\mathbf{n}}\times\mathbf{u}. \tag{29}$$

Comparing this with Eq. (28), we see that the axis and angle of an infinitesimal rotation is given in terms of the correction matrix $\epsilon\mathbf{A}$ in Eq. (19) by

$$\theta\hat{\mathbf{n}} = \epsilon\mathbf{a}, \tag{30}$$

or,

$$\theta = \epsilon|\mathbf{a}|, \qquad \hat{\mathbf{n}} = \frac{\mathbf{a}}{|\mathbf{a}|}. \tag{31}$$

The result is a useful expression for an infinitesimal rotation in axis-angle form,

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = I + \theta\hat{\mathbf{n}}\cdot\mathbf{J} \qquad (\theta \ll 1). \tag{32}$$

We tabulate here some useful properties of the J matrices [see also Eqs. (23) and (26)]. First, the product of two J matrices can be written,

$$\big(\mathsf{J}_i\mathsf{J}_j\big)_{k\ell} = \delta_{i\ell}\delta_{kj} - \delta_{ij}\delta_{k\ell}, \tag{33}$$

as follows from Eqs. (23)and (E.51). From this there follows the commutation relation,

$$\boxed{\big[\mathsf{J}_i, \mathsf{J}_j\big] = \epsilon_{ijk}\,\mathsf{J}_k,} \tag{34}$$

or

$$\boxed{\big[\mathbf{a}\cdot\mathbf{J}, \mathbf{b}\cdot\mathbf{J}\big] = (\mathbf{a}\times\mathbf{b})\cdot\mathbf{J}.} \tag{35}$$

The commutation relations (34) obviously resemble the commutation relations for angular momentum operators in quantum mechanics, but without the $i\hbar$. As we shall see, the quantum angular momentum commutation relations are a reflection of these commutation relations of the J_i matrices at the classical level. Ultimately, these commutation relations represent the geometry of Euclidean space.

9. Exponential Form of Finite Rotations

We will now establish a connection between infinitesimal rotations and the finite rotations $\mathsf{R}(\hat{\mathbf{n}},\theta)$ which take place about a fixed axis. The argument is similar to the one given previously in connection with translation operators [see Eqs. (4.34) and (4.35)].

Our strategy is to set up a differential equation for $(d/d\theta)\mathsf{R}(\hat{\mathbf{n}},\theta)$ and to solve it. By the definition of the derivative, we have

$$\frac{d}{d\theta}\mathsf{R}(\hat{\mathbf{n}},\theta) = \lim_{\epsilon\to 0}\frac{\mathsf{R}(\hat{\mathbf{n}},\theta+\epsilon)-\mathsf{R}(\hat{\mathbf{n}},\theta)}{\epsilon}. \tag{36}$$

But according to Eq. (17), the first factor in the numerator can be written

$$\mathsf{R}(\hat{\mathbf{n}},\theta+\epsilon) = \mathsf{R}(\hat{\mathbf{n}},\epsilon)\mathsf{R}(\hat{\mathbf{n}},\theta), \tag{37}$$

so the derivative becomes

$$\frac{d}{d\theta}\mathsf{R}(\hat{\mathbf{n}},\theta) = \lim_{\epsilon\to 0}\Big(\frac{\mathsf{R}(\hat{\mathbf{n}},\epsilon)-\mathsf{I}}{\epsilon}\Big)\mathsf{R}(\hat{\mathbf{n}},\theta). \tag{38}$$

In the limit, ϵ becomes a small angle, so we can use Eq. (32) to evaluate the limit, obtaining,

$$\frac{d}{d\theta}\mathsf{R}(\hat{\mathbf{n}},\theta) = (\hat{\mathbf{n}}\cdot\mathbf{J})\mathsf{R}(\hat{\mathbf{n}},\theta). \tag{39}$$

Solving this differential equation subject to the initial conditions $\mathsf{R}(\hat{\mathbf{n}},0)=\mathsf{I}$, we obtain

$$\boxed{\mathsf{R}(\hat{\mathbf{n}},\theta) = \exp(\theta\,\hat{\mathbf{n}}\cdot\mathbf{J}).} \tag{40}$$

This exponential can be expanded out in a power series that carries Eq. (32) to higher order terms,

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = \mathsf{I} + \theta\hat{\mathbf{n}}\cdot\mathbf{J} + \frac{\theta^2}{2}(\hat{\mathbf{n}}\cdot\mathbf{J})^2 + \ldots. \tag{41}$$

This series converges for all values of $\hat{\mathbf{n}}$ and θ.

The same result can be obtained in another way that is more pictorial. It is based on the idea that a rotation about a finite angle can be built up out of a sequence of small angle rotations. For example, a rotation of one radian is the product of a million rotations of 10^{-6} radians.

Let us take some angle θ and break it up into a product of rotations of angle θ/N, where N is very large:

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = \mathsf{R}\Big(\hat{\mathbf{n}},\frac{\theta}{N}\Big)^N. \tag{42}$$

We can make θ/N as small as we like by making N large. Therefore we should be able to replace the small angle rotation in Eq. (42) by the small angle formula (32), so that

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = \lim_{N\to\infty}\Big(I + \frac{\theta}{N}\hat{\mathbf{n}}\cdot\mathbf{J}\Big)^N. \tag{43}$$

This is a limit of matrices that is similar to the limit of numbers shown in Eq. (4.42). It turns out that the matrix limit works in the same way, giving Eq. (40).

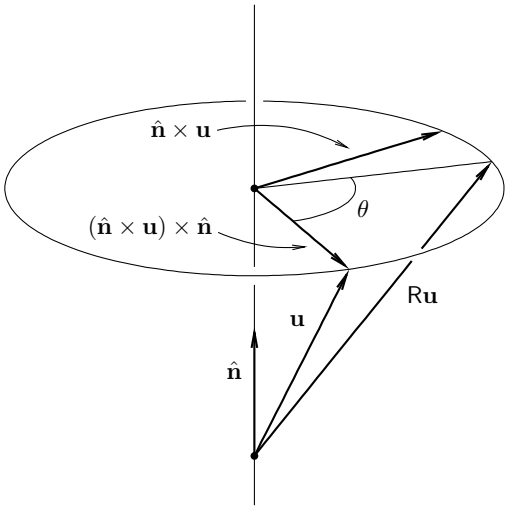

Fig. 4. The result of applying a finite rotation $\mathsf{R}(\hat{\mathbf{n}}, \theta)$ to a vector $\mathbf{u}$ can be expressed as a linear combination of the vectors $\mathbf{u}$, $\hat{\mathbf{n}}\times\mathbf{u}$, and $(\hat{\mathbf{n}}\times\mathbf{u})\times\hat{\mathbf{n}}$, the latter two of which are orthogonal.

10. Explicit Form for Action of Rotation on a Vector

Finite rotations about a fixed axis can be expressed in another way. By using the properties of the $\mathbf{J}$ matrices listed above, one can express higher powers of the matrix $\hat{\mathbf{n}}\cdot\mathbf{J}$ in terms of lower powers, and reexpress the the exponential series (40), acting on an arbitrary vector $\mathbf{u}$, in the following form:

$$\mathsf{R}(\hat{\mathbf{n}}, \theta)\mathbf{u} = \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{u})(1 - \cos\theta) + \mathbf{u}\cos\theta + (\hat{\mathbf{n}}\times\mathbf{u})\sin\theta. \tag{44}$$

The geometrical meaning of this formula is easy to see; the rotation about $\hat{\mathbf{n}}$ leaves the component of $\mathbf{u}$ parallel to $\hat{\mathbf{n}}$ invariant, and rotates the component perpendicular to $\hat{\mathbf{n}}$ by an angle θ in the perpendicular plane. This is illustrated in Fig. 4.

11. Adjoint Formulas

Let $\mathbf{a}$ and $\mathbf{u}$ be arbitrary vectors, and consider how the cross product $\mathbf{a}\times\mathbf{u}$ transforms under proper rotations. For a proper rotation R, the rule is

$$\mathsf{R}(\mathbf{a}\times\mathbf{u}) = (\mathsf{R}\mathbf{a})\times(\mathsf{R}\mathbf{u}), \tag{45}$$

or, as we say in words, the cross product transforms as a vector under proper rotations. The proof of Eq. (45) will be left as an exercise (see Prob. 2), but if you do it, you will find it is equivalent to $\det\mathsf{R} = +1$. In fact, if R is improper, there is a minus sign on the right hand side, an indication that the cross product of two true vectors is not a true vector, but rather a *pseudovector*. More on this in Notes 21, when we return to the improper rotations and discuss parity in quantum mechanics.

Now let us use the notation (26) to express the cross products in Eq. (45) in terms of matrix multiplications. This gives

$$\mathsf{R}(\mathbf{a}\cdot\mathbf{J})\mathbf{u} = [(\mathsf{R}\mathbf{a})\cdot\mathbf{J}]\mathsf{R}\mathbf{u}, \tag{46}$$

or, since $\mathbf{u}$ is arbitrary,

$$\mathsf{R}(\mathbf{a}\cdot\mathbf{J}) = [(\mathsf{R}\mathbf{a})\cdot\mathbf{J}]\mathsf{R}. \tag{47}$$

We now multiply through on the right by R^t, obtaining,

$$\boxed{\mathsf{R}(\mathbf{a}\cdot\mathbf{J})\mathsf{R}^t = (\mathsf{R}\mathbf{a})\cdot\mathbf{J}.} \tag{48}$$

This formula is of such frequent occurrence in applications that we will give it a name. We call it the *adjoint formula*, because of its relation to the adjoint representation of the group $SO(3)$.

Let us now replace R by R_0 to avoid confusion with other rotation matrices to appear momentarily, and let us replace $\mathbf{a}$ by $\theta\hat{\mathbf{n}}$ for the axis and angle of a rotation. Then exponentiating both sides of the adjoint formula (48), we obtain,

$$\exp\big[\mathsf{R}_0(\theta\hat{\mathbf{n}}\cdot\mathbf{J})\mathsf{R}_0^t\big] = \mathsf{R}_0\exp(\theta\hat{\mathbf{n}}\cdot\mathbf{J})\mathsf{R}_0^t = \exp\big[\theta(\mathsf{R}_0\hat{\mathbf{n}})\cdot\mathbf{J}\big], \tag{49}$$

where we have used the rule,

$$e^{ABA^{-1}} = A\,e^B A^{-1}, \tag{50}$$

valid for matrices or operators A and B when A^{-1} exists. This rule can be verified by expanding the exponential series. The result (49) can be written,

$$\boxed{\mathsf{R}_0\mathsf{R}(\hat{\mathbf{n}},\theta)\mathsf{R}_0^t = \mathsf{R}(\mathsf{R}_0\hat{\mathbf{n}},\theta).} \tag{51}$$

We will call this the exponentiated version of the adjoint formula.

The exponentiated version of the adjoint formula (51) says that the axis of a rotation transforms as a vector under proper rotations, while the angle of rotation does not change. As with the original version of the adjoint formula (48), this only holds for proper rotations. Equation (51) might be easier to understand in a passive interpretation, in which R_0 is used to transform the components of $\mathsf{R}(\hat{\mathbf{n}},\theta)$ to a new basis. The new matrix is obtained simply by transforming the axis to the new basis, while keeping the angle fixed.

12. Group Manifolds for $O(3)$ and $SO(3)$

We turn now to the question of parameterizing the rotation matrices or rotation operators. Since an arbitrary, real 3×3 matrix contains 9 real parameters, and since the orthogonality condition $\mathsf{R}\mathsf{R}^t = \mathsf{I}$ constitutes 6 constraints, it follows that it will take 3 parameters to specify a rotation. This is clear already from the axis-angle parameterization of the rotations, since an axis $\hat{\mathbf{n}}$ is equivalent to two parameters (say, the spherical angles specifying the direction of $\hat{\mathbf{n}}$), and the angle of rotation θ is a third parameter.

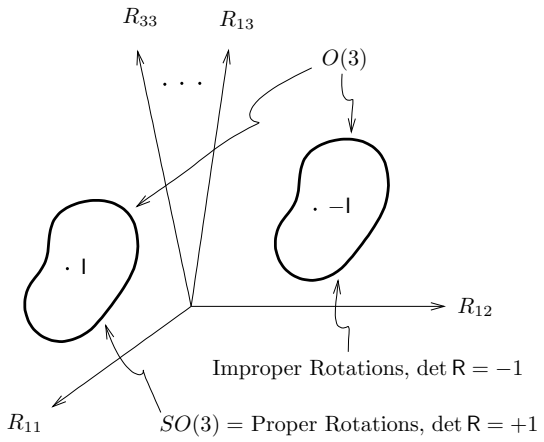

Fig. 5. The rotation group $O(3)$ can be thought of as a 3-dimensional surface imbedded in the 9-dimensional space of all 3×3 matrices. It consists of two disconnected pieces, one of which contains the proper rotations and constitutes the group $SO(3)$, and the other of which contains the improper rotations. The proper rotations include the identity matrix I, and the improper rotations include the parity matrix $-\mathsf{I}$.

It is often useful to think of the rotations and their parameters in geometrical terms. We imagine setting up the 9-dimensional space of all 3×3 real matrices, in which the coordinates are the components of the matrix, R_{ij}. It is difficult to visualize a 9-dimensional space, but we can use our imagination as in Fig. 5. The 6 constraints implicit in $\mathsf{R}\mathsf{R}^t = \mathsf{I}$ imply that the orthogonal matrices lie on a 3-dimensional surface imbedded in this space. This surface is difficult to draw realistically, so it is simply indicated as a nondescript blob in the figure. More exactly, this surface consists of two disconnected pieces, containing the proper and improper matrices. This surface (both pieces) is the *group manifold* for the group $O(3)$, while the piece consisting of the proper rotations alone is the group manifold for $SO(3)$. The identity matrix I occupies a single point in the group manifold $SO(3)$, while the improper matrix $-\mathsf{I}$ lies in the other piece of the group manifold $O(3)$. Any choice of three parameters for the proper rotations can be viewed as a coordinate system on the group manifold $SO(3)$.

The group manifolds $O(3)$ and $SO(3)$ are bounded, that is, they do not run off to infinity. The topological designation for this is that the manifolds are *compact*. This follows from the fact that every row and column of an orthogonal matrix is a unit vector, so all components lie in the range $[-1, 1]$.

The group manifold $SO(3)$ is the configuration space for a rigid body. This is because the orientation of a rigid body is specified relative to a standard or reference orientation by a rotation matrix R that maps the reference orientation into the actual one. In classical rigid body motion, the orientation is a function of time, so the classical trajectory can be seen as a curve $\mathsf{R}(t)$ on the group manifold $SO(3)$. In quantum mechanics, a rigid body is described by a wave function defined on the group manifold $SO(3)$. Many simple molecules behave approximately as a rigid body in their

orientational degrees of freedom.

13. Euler Angles

In addition to the axis-angle parameterization of the rotations, another important parameterization is the *Euler angles*. To construct the Euler angles, let us return to the frame $\hat{\mathbf{e}}_i$ introduced at the beginning of these notes, and recall the rotated frame $\hat{\mathbf{e}}'_i$ defined in Eq. (10). We will take it as geometrically obvious that the rotation operator R or the corresponding rotation matrix R is uniquely specified by the orientation of the rotated frame. Therefore to obtain parameters of the rotation, we can specify the orientation of the rotated frame, that is, the orientation of all three rotated axes.

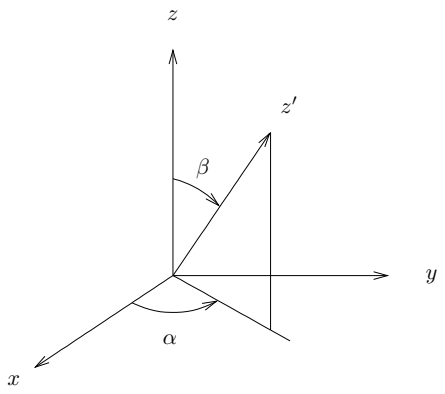

Fig. 6. The Euler angles α and β are the spherical angles of the rotated z'-axis as seen in the unrotated frame.

We begin by specifying the orientation of the rotated z'-axis. This axis points in some direction, which we can indicate by its spherical angles, say, (α, β), with respect to the unrotated frame. This is illustrated in Fig. 6. We have in mind here some rotation R, which corresponds to definite orientations of the three primed axes. The rotation R maps the old (unprimed axes) into the new (primed) ones, and in particular, it satisfies $\hat{\mathbf{z}}' = \mathsf{R}\hat{\mathbf{z}}$. Another rotation matrix that also maps the old z-axis into the new z'-axis is R_1, defined by

$$\mathsf{R}_1 = \mathsf{R}(\hat{\mathbf{z}}, \alpha)\mathsf{R}(\hat{\mathbf{y}}, \beta). \tag{52}$$

By examining Fig. 6, it is easy to see that R_1 satisfies

$$\mathsf{R}_1\hat{\mathbf{z}} = \hat{\mathbf{z}}', \tag{53}$$

since the first rotation by angle β about the y-axis swings the z-axis down in the x-z plane, and then the second rotation by angle α about the z-axis rotates the vector in a cone, bringing it into the final position for the z'-axis. Rotation R_1 is not in general equal to R, for R_1 is only designed to orient the z'-axis correctly, while R puts all three primed axes into their correct final positions. But

R_1 can get the x'- and y'-axes wrong only by some rotation in the x'-y' plane; therefore if we follow the β and α rotations by a third rotation by some new angle, say, γ, about the z'-axis, then we can guarantee that all three axes achieve their desired orientations. That is, we can write an arbitrary rotation R in the form,

$$\mathsf{R} = \mathsf{R}(\hat{\mathbf{z}}', \gamma)\mathsf{R}_1 = \mathsf{R}(\hat{\mathbf{z}}', \gamma)\mathsf{R}(\hat{\mathbf{z}}, \alpha)\mathsf{R}(\hat{\mathbf{y}}, \beta), \tag{54}$$

for some angles (α, β, γ).

But it is not convenient to express a rotation in terms of elementary rotations about a mixture of old and new axes, as in Eq. (54); it is more convenient to express it purely in terms of rotations about the old axes. To do this, we write

$$\mathsf{R}(\hat{\mathbf{z}}', \gamma) = \mathsf{R}(\mathsf{R}_1\hat{\mathbf{z}}, \gamma) = \mathsf{R}_1\mathsf{R}(\hat{\mathbf{z}}, \gamma)\mathsf{R}_1^t, \tag{55}$$

using the exponentiated version of the adjoint formula (51). Therefore

$$\mathsf{R} = \mathsf{R}(\hat{\mathbf{z}}', \gamma)\mathsf{R}_1 = \mathsf{R}_1\mathsf{R}(\hat{\mathbf{z}}, \gamma) = \mathsf{R}(\hat{\mathbf{z}}, \alpha)\mathsf{R}(\hat{\mathbf{y}}, \beta)\mathsf{R}(\hat{\mathbf{z}}, \gamma). \tag{56}$$

This gives us the Euler angle parameterization of proper rotations,

$$\boxed{\mathsf{R}(\alpha, \beta, \gamma) = \mathsf{R}(\hat{\mathbf{z}}, \alpha)\mathsf{R}(\hat{\mathbf{y}}, \beta)\mathsf{R}(\hat{\mathbf{z}}, \gamma).} \tag{57}$$

It is a factorization of an aribtrary, proper rotation into a product of rotations about the coordinate axes, taken in the order (zyz).

Equation (57) constitutes the zyz-convention for the Euler angles, which is particularly appropriate for quantum mechanical applications. Other conventions are possible, and the zxz-convention is common in books on classical mechanics, while the xyz-convention is used in aeronautical engineering (the three rotations are called pitch, yaw and roll, familiar to airplane pilots). Also, you should note that most books on classical mechanics and some books on quantum mechanics adopt the passive point of view, which usually means that the rotation matrices in those books stand for the transposes of the rotation matrices in these notes.

The geometrical meanings of the Euler angles α and β is particularly simple, since these are just the spherical angles of the z'-axis as seen in the unprimed frame. The geometrical meaning of γ is more difficult to see; it is in fact the angle between the y'-axis and the unit vector $\hat{\mathbf{n}}$ lying in the line of nodes. The line of nodes is the line of intersection between the x-y plane and the x'-y' plane. This line is perpendicular to both the z- and z'-axis, and we take $\hat{\mathbf{n}}$ to lie in the direction $\hat{\mathbf{z}}\times\hat{\mathbf{z}}'$.

The allowed ranges on the Euler angles are the following:

$$\begin{aligned} 0 &\le \alpha \le 2\pi, \\ 0 &\le \beta \le \pi, \\ 0 &\le \gamma \le 2\pi. \end{aligned} \tag{58}$$

The ranges on α and β follow from the fact that they are spherical angles, while the range on γ follows from the fact that the γ-rotation is used to bring the x'- and y'-axes into proper alignment in

their plane. If the Euler angles lie within the interior of the ranges indicated, then the representation of the rotations is unique; but if one or more of the Euler angles takes on their limiting values, then the representation may not be unique. For example, if $\beta = 0$, then the rotation is purely about the z-axis, and depends only on the sum of the angles, $\alpha + \gamma$. In other words, apart from exceptional points at the ends of the ranges, the Euler angles form a 1-to-1 coordinate system on the group manifold $SO(3)$.

14. The Noncommutativity of Rotations

As pointed out earlier, rotations do not commute, so $\mathsf{R}_1\mathsf{R}_2 \neq \mathsf{R}_2\mathsf{R}_1$, in general. An exception, also noted above, is the case that R_1 and R_2 are about the same axis, but when rotations are taken about different axes they generally do not commute. Because of this, the rotation group $SO(3)$ is said to be *non-Abelian*. This means that the rotation group is noncommutative.

Let us write $\mathsf{R}_1 = \mathsf{R}(\hat{\mathbf{n}}_1, \theta_1)$ and $\mathsf{R}_2 = \mathsf{R}(\hat{\mathbf{n}}_2, \theta_2)$ for two rotations in axis-angle form. If the axes $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ are not identical, then it is not entirely simple to find the axis and angle of the product, that is, $\hat{\mathbf{n}}_3$ and θ_3 in

$$\mathsf{R}(\hat{\mathbf{n}}_1, \theta_1)\mathsf{R}(\hat{\mathbf{n}}_2, \theta_2) = \mathsf{R}(\hat{\mathbf{n}}_3, \theta_3). \tag{59}$$

A formula for $\hat{\mathbf{n}}_3$ and θ_3 in terms of the axes and angles of R_1 and R_2 exists, but it is not trivial. If we write

$$\mathsf{A}_i = \theta_i(\hat{\mathbf{n}}_i \cdot \mathbf{J}), \qquad i = 1, 2, 3, \tag{60}$$

for the three antisymmetric matrices in the exponential expressions for the three rotations in Eq. (59), then that equation can be written as

$$e^{\mathsf{A}_1}\, e^{\mathsf{A}_2} = e^{\mathsf{A}_3}, \tag{61}$$

and the problem is to find A_3 in terms of A_1 and A_2. We see once again the problem of combining exponentials of noncommuting objects (matrices or operators), one that has appeared more than once previously in the course. We comment further on this problem below.

The situation is worse in the Euler angle parameterization; it is quite a messy task to find $(\alpha_3, \beta_3, \gamma_3)$ in terms of $(\alpha_1, \beta_1, \gamma_1)$ and $(\alpha_2, \beta_2, \gamma_2)$ in the product,

$$\mathsf{R}(\alpha_3, \beta_3, \gamma_3) = \mathsf{R}(\alpha_1, \beta_1, \gamma_1)\mathsf{R}(\alpha_2, \beta_2, \gamma_2). \tag{62}$$

The difficulties of these calculations are due to the noncommutativity of the rotation matrices.

A measure of the commutativity of two rotations R_1, R_2 is the matrix

$$\mathsf{C} = \mathsf{R}_1\mathsf{R}_2\mathsf{R}_1^{-1}\mathsf{R}_2^{-1}, \tag{63}$$

which itself is a rotation, and which becomes the identity matrix if R_1 and R_2 should commute. The matrix C is more interesting than the ordinary commutator, $[\mathsf{R}_1, \mathsf{R}_2]$, which is not a rotation. You can think of C as taking a step in the 1-direction, then a step in the 2-direction, then a backwards step in the 1-direction, and finally a backwards step in the 2-direction, all in the space of rotations,

and asking if we return to the identity. The answer is that in general we do not, although if the steps are small then we trace out a path that looks like a small parallelogram that usually does not quite close.

Let us examine C when the angles θ_1 and θ_2 are small. Let us write A_i for the antisymmetric matrices in the exponents of the exponential form for the rotations, as in Eq. (60), and let us expand the exponentials in power series. Then we have

$$\begin{aligned}\mathsf{C} &= e^{\mathsf{A}_1}\, e^{\mathsf{A}_2}\, e^{-\mathsf{A}_1}\, e^{-\mathsf{A}_2} \\ &= \big(\mathsf{I} + \mathsf{A}_1 + \tfrac{1}{2}\mathsf{A}_1^2 + \ldots\big)\big(\mathsf{I} + \mathsf{A}_2 + \tfrac{1}{2}\mathsf{A}_2^2 + \ldots\big) \\ &\qquad \times \big(\mathsf{I} - \mathsf{A}_1 + \tfrac{1}{2}\mathsf{A}_1^2 - \ldots\big)\big(\mathsf{I} - \mathsf{A}_2 + \tfrac{1}{2}\mathsf{A}_2^2 - \ldots\big) \\ &= \mathsf{I} + [\mathsf{A}_1, \mathsf{A}_2] + \ldots .\end{aligned} \tag{64}$$

The first order term vanishes, and at second order we find the commutator of the A matrices, as indicated. Writing these in terms of axes and angles as in Eq. (60), the matrix C becomes

$$\begin{aligned}\mathsf{C} &= \mathsf{I} + \theta_1\theta_2[\hat{\mathbf{n}}_1 \cdot \mathbf{J}, \hat{\mathbf{n}}_2 \cdot \mathbf{J}] + \ldots \\ &= \mathsf{I} + \theta_1\theta_2(\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2) \cdot \mathbf{J} + \ldots ,\end{aligned} \tag{65}$$

where we have used the commutation relations (35). We see that if R_1 and R_2 are near-identity rotations, then so is C, which is about the axis $\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2$.

Matrix C is a near-identity rotation, so we can write it in axis-angle form, say, with axis $\hat{\mathbf{m}}$ and angle ϕ,

$$\mathsf{C} = \mathsf{I} + \phi(\hat{\mathbf{m}} \cdot \mathbf{J}). \tag{66}$$

Comparing this with Eq. (65), we have

$$\phi\, \hat{\mathbf{m}} = \theta_1\theta_2\, \hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2. \tag{67}$$

This result will be of use in Notes 12.

15. The Baker-Campbell-Hausdorff Theorem

It is no surprise to see the commutator of near identity rotations expressed in terms of the commutators of the $\mathbf{J}$ matrices, since the latter are the correction terms for near identity rotations. Equation (65) just gives the details. But there is an important point to be made about the commutation relations of near identity transformations, related to the Baker-Campbell-Hausdorff theorem, which we now discuss.

The theorem concerns products of exponentials of operators or matrices that need not commute, as in Eq. (61). We rewrite that equation in a different notation,

$$e^X\, e^Y = e^Z, \tag{68}$$

where X and Y are given and it is desired to find Z. First, if X and Y commute, then the product of exponentials follows the rules for ordinary numbers, and $Z = X + Y$. Next, if X and Y do not commute but they do commute with their commutator, then the conditions of Glauber's theorem hold (see Sec. 8.9) and we have

$$Z = X + Y + \tfrac{1}{2}[X, Y]. \tag{69}$$

If X and Y do not commute with their commutator, then no simple answer can be given, but we can expand Z in a power series in X and Y. Carried through third order, this gives

$$Z = X + Y + \tfrac{1}{2}[X, Y] + \tfrac{1}{12}[X, [X, Y]] + \tfrac{1}{12}[Y, [Y, X]] + \ldots. \tag{70}$$

The Baker-Campbell-Hausdorff theorem concerns this series, and asserts that the general term of this Taylor series can be written in terms of commutators and iterated commutators of X and Y. This fact can be seen through third order in Eq. (70). Thus, if we know how to compute commutators, we can in principle compute all the terms of this Taylor series and obtain the multiplication rule for products of exponentials.

For example, in the case of the rotations, knowledge of the commutation relations (34) or (35) allows us to compute all the terms of the Taylor series (70) and thus obtain the multiplication law for finite rotations in terms of axis-angle parameters, at least in principle. We do not intend to do this in practice, and in fact we shall use the Baker-Campbell-Hausdorff theorem only for its suggestive and intuitive value. But the point is that all the complexity of the noncommutative multiplication law for finite rotations is implicitly contained in the commutation relations of infinitesimal rotations, that is, of the $\mathbf{J}$ matrices. This explains the emphasis placed on commutation relations in these notes and in physics in general.

Problems

1. Prove the commutation relations (34), using Eq. (23) and the properties of the Levi-Civita symbol ϵ_{ijk}. See Sec. E.16.

2. Some useful identities.

(a) Show that if $\mathsf{R} \in O(3)$, then

$$(\mathsf{R}\mathbf{a}) \cdot (\mathsf{R}\mathbf{b}) = \mathbf{a} \cdot \mathbf{b}, \tag{71}$$

Also show that

$$\mathbf{a} \cdot (\mathsf{R}\mathbf{b}) = (\mathsf{R}^{-1}\mathbf{a}) \cdot \mathbf{b}. \tag{72}$$

These formulas hold for all vectors $\mathbf{a}$ and $\mathbf{b}$.

(b) Show that if $\mathsf{R} \in SO(3)$, then

$$\mathsf{R}(\mathbf{a} \times \mathbf{b}) = (\mathsf{R}\mathbf{a}) \times (\mathsf{R}\mathbf{b}). \tag{73}$$

Hint: Use the fact that if M is any 3×3 matrix, then

$$\epsilon_{ijk} \det \mathsf{M} = \epsilon_{\ell mn} M_{i\ell} M_{jm} M_{kn}. \tag{74}$$

This is essentially the definition of the determinant. This proves Eq. (45), and hence the adjoint formulas (48) and (51). Note that when Eq. (74) is combined with Eq. (71), we obtain

$$(\mathsf{R}\mathbf{a}) \cdot [(\mathsf{R}\mathbf{b}) \times (\mathsf{R}\mathbf{c})] = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}), \tag{75}$$

valid for all $\mathsf{R} \in SO(3)$ and all vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$.

3. The geometrical meaning of Eq. (44) is illustrated in Fig. 4. The rotation leaves the component of $\mathbf{u}$ along the axis $\hat{\mathbf{n}}$ invariant, while rotating the orthogonal component in the plane perpendicular to $\hat{\mathbf{n}}$.

By expressing powers of the $\mathbf{J}$ matrices in terms of lower powers, sum the exponential series (40) and obtain another proof of Eq. (44).

4. The axis $\hat{\mathbf{n}}$ of a proper rotation R is invariant under the action of R, that is, $\mathsf{R}\hat{\mathbf{n}} = \hat{\mathbf{n}}$. Therefore $\hat{\mathbf{n}}$ is a real, normalized eigenvector of R with eigenvalue $+1$.

Prove that every proper rotation has an axis. Show that the axis is unique (apart from the change in sign, $\hat{\mathbf{n}} \to -\hat{\mathbf{n}}$) as long as the angle of rotation is nonzero. This proof is the essential step in showing that every rotation can be expressed in axis-angle form. Do proper rotations in 4 dimensions have an axis?

5. It is claimed that every proper rotation can be written in Euler angle form. Find the Euler angles (α, β, γ) for the rotation $\mathsf{R}(\hat{\mathbf{x}}, \pi/2)$.

Physics 221A
Academic Year 2020–21
Notes 12
Rotations in Quantum Mechanics, and
Rotations of Spin-$\frac{1}{2}$ Systems†

1. Introduction

In these notes we develop a general strategy for finding unitary operators to represent rotations in quantum mechanics, and we work through the specific case of rotations in spin-$\frac{1}{2}$ systems. We find that the relation between spin-$\frac{1}{2}$ rotations and classical rotations is two-to-one, due to the appearance of non-classical phase factors. In a certain sense spin-$\frac{1}{2}$ rotations constitute a basic building block in the theory of rotations in quantum mechanics, out of which rotations for an arbitrary system can be constructed. The general theory of rotations in quantum mechanics will be developed in later sets of notes.

2. Physical Meaning of Rotations in Quantum Mechanics

In classical mechanics, we can rotate the state of a dynamical system, that is, we can rotate all the position and velocity vectors $(\mathbf{r}, \mathbf{v})$ for each particle, to create a new or rotated state. Similarly, in quantum mechanics, given a state $|\psi\rangle$ (taken for simplicity to be pure), it is possible to define a rotated state such that the expectation values of all vector operators in the rotated state are rotated relative to the expectation values in the original state, exactly as classical vectors would transform under rotations in a classical system. The transformation between the original quantum state $|\psi\rangle$ and the rotated quantum state is brought about by means of a certain rotation operator, which is unitary because probabilities must be preserved under rotations.

What does it mean physically to rotate a quantum state? Certainly we cannot go in with a wrench, and rotate the wave function of the electrons in an atom, or rotate the spins of those electrons. In any case, as we have emphasized, the wave function represents the properties of an ensemble of systems, not an individual system, so rotating a quantum state must be equivalent to rotating the properties of the ensemble. One point of view is to identify a quantum state with the apparatus that prepares the ensemble of systems about which the quantum state makes statistical predictions. We can certainly rotate a preparation apparatus, and it is logical to regard the state prepared by the rotated apparatus as the rotated state. This point of view gives us a definition of a rotated state, without however specifying the phase of that state.

† Links to the other sets of notes can be found at:
`http://bohr.physics.berkeley.edu/classes/221/2021/221.html`.

Alternatively, quantum systems may experience rotations as a result of their time evolution. For example, many small molecules are approximately rigid bodies, and their time evolution is approximately given by a time-dependent rotation operator, just as the time evolution of a classical rigid body is specified by a time-dependent rotation $\mathsf{R}(t)$ (see Sec. 11.12). For another example, spins in magnetic fields evolve by means of a time-dependent rotation operator. The magnetic field can either be external or internal; for example, the electrons in an atom precess in the magnetic field of the nucleus (which is a moving charge in the electron rest frame, and which may have an intrinsic magnetic field of its own). Or the magnetic field may be external, permitting experimental control over the orientation of spins. Another example of rotations in quantum mechanics is produced by Thomas precession, a relativistic but purely inertial effect that rotates all the dynamical variables of an accelerated system, for example, the spin of an electron in an orbit in an atom. Different subsystems of a given system may be rotated by different amounts, for example, it is possible to change the orientation of a molecule without rotating its spins (hit it with another molecule), or vice versa (apply a magnetic field). Thus we speak of spatial rotation operators, spin rotation operators, etc.

The phases associated with rotations are observable. For example, in a neutron interferometer, it is possible to split a beam of neutrons into two, and to subject one of the two resulting beams to a magnetic field, which will rotate the neutron spins. By recombining the beams and observing the interference pattern, phase shifts such as the -1 multiplicative factor that occurs on rotating a spin-$\frac{1}{2}$ system by 2π can be observed. This is a rather direct observation of a phase shift, but phases also enter into quantization conditions (recall the Bohr-Sommerfeld rules), and theoretical predictions of energy levels would not agree with experiment if all phases were not accounted for correctly.

3. Postulates for Rotation Operators

Suppose we have some quantum mechanical system, and an associated Hilbert space of states. We will be interested in finding operators that act on these states that represent the action of rotations at the classical level. If a classical rotation is specified by a rotation matrix R (relative to some inertial frame), then we will denote the associated operator by $U(\mathsf{R})$. For the time being we will work only with proper rotations, so that $\mathsf{R} \in SO(3)$. We will denote the association itself by

$$\mathsf{R} \mapsto U(\mathsf{R}), \tag{1}$$

which simply means that U is a function of the classical rotation R.

As you might imagine, the specific form of the operators $U(\mathsf{R})$ depends on the system, but it turns out that there is a great deal that can be said about these operators without going into the specific nature of the system. These are the properties of rotation operators that follow from the properties of classical rotations, that is, ultimately from the Euclidean geometry of three-dimensional

space. We will concentrate on these properties first, and then deal with specific physical systems (spin systems, central force problems, atoms, etc) which are treated in later sets of notes.

We make a series of reasonable assumptions or postulates that the rotation operators $U(\mathsf{R})$ should satisfy. First, these operators should be unitary, because a symmetry operation should preserve probabilities:

$$U(\mathsf{R})^{-1} = U(\mathsf{R})^\dagger. \tag{2}$$

Next, we assume that when the classical rotation is the identity, so is the unitary operator,

$$U(\mathsf{I}) = 1. \tag{3}$$

Finally, we assume that the unitary operator corresponding to the product of two rotations is the product of the unitary operators, so that

$$U(\mathsf{R}_1)U(\mathsf{R}_2) = U(\mathsf{R}_1\mathsf{R}_2). \tag{4}$$

This means that the unitary operators $U(\mathsf{R})$ reproduce the multiplication law of classical rotations. These requirements imply that inverse rotations are mapped into inverse unitary operators,

$$U(\mathsf{R}^{-1}) = U(\mathsf{R})^{-1} = U(\mathsf{R})^\dagger. \tag{5}$$

If the requirements (2) to (5) are satisfied, then we say that $U(\mathsf{R})$ (more precisely, the mapping (1)) forms a *representation* of $SO(3)$ by means of unitary operators. As we will see, these requirements are actually too strong, and in the case of systems of half-integral spin, they cannot be met; for such systems we can almost find a representation, but we ultimately fail because of phase factors. However, the search for a unitary representation of the classical rotations is educational, and the phase factors are not so much a difficulty as an opportunity for obtaining a deeper understanding of rotations, and for finding new physics at the quantum level.

In addition to the mathematical requirements given by Eqs. (3) to (5), the operators $U(\mathsf{R})$ must also satisfy physical properties we expect of rotations. For example, the expectation values of vector operators should transform as classical vectors.

4. Representations of Near-Identity Rotations

The key to finding a unitary representation of the rotations is to begin with near-identity rotations. First, some notation. We notice that a rotation in axis-angle form actually depends only on the product $\theta\hat{\mathbf{n}}$ of the axis and the angle,

$$\mathsf{R}(\hat{\mathbf{n}}, \theta) = e^{\theta\hat{\mathbf{n}}\cdot\mathbf{J}}, \tag{6}$$

so we introduce a vector of angles $\boldsymbol{\theta}$ defined by

$$\boldsymbol{\theta} = \theta\hat{\mathbf{n}}, \tag{7}$$

and write $\mathsf{R}(\boldsymbol{\theta})$ for the rotation. Now if the rotation is near-identity, we can approximate it by the leading terms of the exponential series,

$$\mathsf{R}(\boldsymbol{\theta}) = \mathsf{I} + \boldsymbol{\theta} \cdot \mathbf{J} + \ldots, \tag{8}$$

which is Eq. (11.32). This implies that

$$\mathsf{J}_k = \left. \frac{\partial \mathsf{R}(\boldsymbol{\theta})}{\partial \theta_k} \right|_{\boldsymbol{\theta}=0}, \tag{9}$$

for $k = 1, 2, 3$.

The unitary operator that corresponds to the rotation $\mathsf{R}(\boldsymbol{\theta})$ can also be parameterized by $\boldsymbol{\theta}$, so we will write $U(\boldsymbol{\theta}) = U\big(\mathsf{R}(\boldsymbol{\theta})\big)$. In the case of small angles, we can approximate $U(\boldsymbol{\theta})$ by the leading terms of its Taylor series,

$$U(\boldsymbol{\theta}) = 1 + \sum_k \left. \frac{\partial U(\boldsymbol{\theta})}{\partial \theta_k} \right|_{\boldsymbol{\theta}=0} \theta_k + \ldots, \tag{10}$$

where the first term is 1 (the identity operator) because of Eq. (3). At this point you may wish to review the manner in which we defined linear momentum $\mathbf{p}$ as the generator of translations in Notes 4, or the manner in which we defined the Hamiltonian in Notes 5. See especially Eqs. (4.30), (4.65), (5.6) and (5.8). Following the pattern of those definitions, we now define the *angular momentum* of the quantum system as the vector operator $\mathbf{J}$ with components J_k given by

$$J_k = i\hbar \left. \frac{\partial U(\boldsymbol{\theta})}{\partial \theta_k} \right|_{\boldsymbol{\theta}=0}. \tag{11}$$

This defines $\mathbf{J}$ relative to some definition for the unitary rotation operators $U(\mathsf{R})$, so different systems, with different definitions of rotation operators, will have different definitions of the angular momentum (for some it will be orbital angular momentum, for others,spin, etc). This definition of angular momentum allows us to write a near-identity rotation as

$$U(\boldsymbol{\theta}) = 1 - \frac{i}{\hbar} \boldsymbol{\theta} \cdot \mathbf{J} + \ldots. \tag{12}$$

We split off a factor of i in the definition of J_k in Eq. (11) so that the operators J_k will be Hermitian, because in quantum mechanics we think of Hermitian operators as the generators of unitary operators. The Hermiticity of J_k follows by substituting Eq. (12) into $U(\boldsymbol{\theta})U(\boldsymbol{\theta})^\dagger = 1$ (exactly as we proved the Hermiticity of the operator $\hat{k}$ in Sec. 4.5). As for the factor of $\hbar$, it makes J_k have dimensions of angular momentum.

Reverting now to our original axis-angle notation, we can write the near-identity rotation operator as

$$U(\hat{\mathbf{n}}, \theta) = 1 - \frac{i}{\hbar} \theta \hat{\mathbf{n}} \cdot \mathbf{J} + \ldots. \tag{13}$$

Since the operators $\mathbf{J}$ were defined in terms of $U(\mathsf{R})$, this is not an explicit solution for the rotation operators U, even for infinitesimal rotations. But at least is shows how those rotation operators

depend on the axis and angle, which is progress. Note that $\mathbf{J}$ is a fixed set of three Hermitian operators that are independent of the axis or angle. In a moment we will extend this relation to arbitrary angles.

When defining the angular momentum in quantum mechanics, we have some of the same issues we faced earlier when trying to define linear momentum in quantum mechanics by taking over some definition from classical mechanics. In the case of linear momentum, we decided that the role that linear momentum plays in classical mechanics as the generator of translations was the most fundamental role; this supersedes definitions such as $\mathbf{p} = m\mathbf{v}$, which are not as general. Similarly, in quantum mechanics, we will define the angular momentum as the generator of rotations, rather than as $\mathbf{r}\times\mathbf{p}$, which is not as general. For not only is $\mathbf{r}\times\mathbf{p}$ meaningless for systems such as spin systems, but even for the spatial degrees of freedom of a spinless particle, it is not always true that the generator of rotations is $\mathbf{r}\times\mathbf{p}$ (for example, in the presence of magnetic fields). We take the generator of rotations to be the most fundamental role of angular momentum, because, among other reasons, these generators (the components of $\mathbf{J}$) are conserved in systems with rotational symmetry. Moreover, the generality of our definition allows us to treat the angular momentum of orbital motion, spin systems, systems in magnetic fields, multiparticle systems, relativisitic systems and quantum fields under the same formalism.

5. Rotation Operators for Any Angle

The relation (13), which gives $U(\hat{\mathbf{n}},\theta)$ when θ is small, can be extended to a closed formula valid for any θ. The pattern follows what we did earlier with translation operators in Sec. 4.5 and with classical rotations in Sec. 11.9. We begin by seeking a differential equation for $U(\hat{\mathbf{n}},\theta)$. By the definition of the derivative,

$$\frac{dU(\hat{\mathbf{n}},\theta)}{d\theta} = \lim_{\epsilon\to 0}\frac{U(\hat{\mathbf{n}},\theta+\epsilon)-U(\hat{\mathbf{n}},\theta)}{\epsilon}. \tag{14}$$

But since the operators $U(\hat{\mathbf{n}},\theta)$ form a representation of the classical rotations $\mathsf{R}(\hat{\mathbf{n}},\theta)$ [see Eq. (4)] and since rotations about a fixed axis commute [see Eq. (11.17)], we have

$$U(\hat{\mathbf{n}},\theta+\epsilon) = U(\hat{\mathbf{n}},\epsilon)U(\hat{\mathbf{n}},\theta), \tag{15}$$

and the numerator in Eq. (14) has a common factor of $U(\hat{\mathbf{n}},\theta)$ that can be taken out to the right. Thus we have

$$\frac{dU(\hat{\mathbf{n}},\theta)}{d\theta} = \lim_{\epsilon\to 0}\left(\frac{U(\hat{\mathbf{n}},\epsilon)-1}{\epsilon}\right)U(\hat{\mathbf{n}},\theta). \tag{16}$$

The remaining limit can be evaluated in terms of the angular momentum $\mathbf{J}$ with the aid of Eq. (13), where the small angle θ of that formula is identified with ϵ here. This gives the differential equation,

$$\frac{dU(\hat{\mathbf{n}},\theta)}{d\theta} = -\frac{i}{\hbar}(\hat{\mathbf{n}}\cdot\mathbf{J})U(\hat{\mathbf{n}},\theta). \tag{17}$$

Subject to the initial condition $U(\hat{\mathbf{n}}, 0) = 1$, this has the unique solution,

$$U(\hat{\mathbf{n}}, \theta) = \exp\Big(-\frac{i}{\hbar}\theta\hat{\mathbf{n}}\cdot\mathbf{J}\Big). \tag{18}$$

This solution can also be obtained as the limit of the product of a large number of small angle rotations, as in Sec. 11.9 [see Eqs. (11.42)–(11.43)].

Equation (18) shows that we can find the rotation operators $U(\hat{\mathbf{n}}, \theta)$ corresponding to the classical rotations $R(\hat{\mathbf{n}}, \theta)$ once we know the angular momentum operators $\mathbf{J}$. This greatly simplifies the problem, because there are only three angular momentum operators, but an infinite family of rotation operators. The angular momentum operators are not arbitrary, however; in addition to being Hermitian, they must satisfy certain commutation relations.

6. Commutation Relations for Angular Momentum

Let us consider the rotation C defined in Eq. (11.63), and its unitary representative $U(\mathsf{C})$. We have

$$U(\mathsf{C}) = U(\mathsf{R}_1)U(\mathsf{R}_2)U(\mathsf{R}_1^{-1})U(\mathsf{R}_2^{-1}). \tag{19}$$

Let us expand both sides of this equation in a Taylor series in the angles θ_1 and θ_2, defined by $\mathsf{R}_1 = \mathsf{R}(\hat{\mathbf{n}}_1, \theta_1)$ and $\mathsf{R}_2 = \mathsf{R}(\hat{\mathbf{n}}_2, \theta_2)$. The answer can be obtained in two ways. In one approach, we expand the exponentials for $U(\mathsf{R}_1)$, $U(\mathsf{R}_2)$, etc., according to Eq. (18), and multiply the series. This gives

$$\begin{aligned} U(\mathsf{C}) = &\Big[1 - \frac{i\theta_1}{\hbar}(\hat{\mathbf{n}}_1\cdot\mathbf{J}) - \frac{\theta_1^2}{2\hbar^2}(\hat{\mathbf{n}}_1\cdot\mathbf{J})^2 + \ldots\Big]\Big[1 - \frac{i\theta_2}{\hbar}(\hat{\mathbf{n}}_2\cdot\mathbf{J}) - \frac{\theta_2^2}{2\hbar^2}(\hat{\mathbf{n}}_2\cdot\mathbf{J})^2 + \ldots\Big] \\ &\times\Big[1 + \frac{i\theta_1}{\hbar}(\hat{\mathbf{n}}_1\cdot\mathbf{J}) - \frac{\theta_1^2}{2\hbar^2}(\hat{\mathbf{n}}_1\cdot\mathbf{J})^2 + \ldots\Big]\Big[1 + \frac{i\theta_2}{\hbar}(\hat{\mathbf{n}}_2\cdot\mathbf{J}) - \frac{\theta_2^2}{2\hbar^2}(\hat{\mathbf{n}}_2\cdot\mathbf{J})^2 + \ldots\Big]. \end{aligned} \tag{20}$$

The calculation is similar to that in Eq. (11.64) leading to Eq. (11.65). When the series are multiplied, first order terms vanish, but at second order the product becomes

$$U(\mathsf{C}) = 1 - \frac{1}{\hbar^2}\theta_1\theta_2[\hat{\mathbf{n}}_1\cdot\mathbf{J}, \hat{\mathbf{n}}_2\cdot\mathbf{J}] + \ldots. \tag{21}$$

On the other hand, according to Eq. (11.66), C is a near-identity rotation with axis $\hat{\mathbf{m}}$ and angle ϕ, $\mathsf{C} = \mathsf{I} + \phi\hat{\mathbf{m}}\cdot\mathbf{J}$, where $\hat{\mathbf{m}}$ and ϕ are given by Eq. (11.67). Therefore

$$U(\mathsf{C}) = 1 - \frac{i\phi}{\hbar}\hat{\mathbf{m}}\cdot\mathbf{J} + \ldots = 1 - \frac{i}{\hbar}\theta_1\theta_2(\hat{\mathbf{n}}_1\times\hat{\mathbf{n}}_2)\cdot\mathbf{J} + \ldots, \tag{22}$$

where we use Eq. (11.67). This is consistent with Eq. (21) only if

$$[\hat{\mathbf{n}}_1\cdot\mathbf{J}, \hat{\mathbf{n}}_2\cdot\mathbf{J}] = i\hbar(\hat{\mathbf{n}}_1\times\hat{\mathbf{n}}_2)\cdot\mathbf{J}. \tag{23}$$

By setting $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ to $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$, we obtain

$$\boxed{[J_i, J_j] = i\hbar\,\epsilon_{ijk}\,J_k.} \tag{24}$$

These are the standard commutation relations for angular momentum in quantum mechanics, here extracted from the properties of rotation operators. These commutation relations are the quantum analogs of the classical commutation relations (11.34) for the J matrices; the two commutation relations are the same, apart from conventional factors of i and $\hbar$.

The usual approach to deriving the angular momentum commutation relations (24) in elementary courses in quantum mechanics is to work with the specific example of orbital angular momentum, whose definition $\mathbf{L} = \mathbf{r}\times\mathbf{p}$ is taken over from classical mechanics. One then just works out the commutators of the components of $\mathbf{L}$, and argues that other forms of angular momentum such as spin ought to have the same commutation relations. Here we have derived the angular momentum commutation relations in all generality, following from the properties of rotations. The specific case of orbital angular momentum $\mathbf{L}$ will be dealt with in Notes 15.

We have shown that the only way the unitary operators U can reproduce the multiplication law for the classical rotations R is if the infinitesimal generators in each case satisfy the same commutation relations, apart from conventional factors of i and $\hbar$. Therefore we adopt the following strategy in developing the general theory of the representations of classical rotations. We begin by seeking the most general form that a vector of Hermitian operators $\mathbf{J}$ can take, given that it satisfies the commutation relations (24). For example, we will be interested in the matrices that represent these operators in some appropriately chosen basis. When we have found specific operators $\mathbf{J}$ that satisfy the commutation relations (24), we will say that we have found a *representation* of those commutation relations. Next, given some such operators $\mathbf{J}$, we exponentiate linear combinations of them as in Eq. (18), to obtain the unitary rotation operators $U(\hat{\mathbf{n}},\theta)$. These can also be represented as matrices in some basis. Finally, we explore the physical implications of these operators, to guarantee that they have the physical properties we expect of rotations.

7. Angular Momentum and Rotation Operators for Spin-$\frac{1}{2}$ Systems

Thus, we must begin by finding a representation of the angular momentum commutation relations (24). We will take up the general problem of doing this in Notes 13; for the remainder of these notes, however, we will restrict consideration to a spin-$\frac{1}{2}$ system, which possesses a 2-dimensional ket space. To find operators $\mathbf{J}$ acting on this space that satisfy the commutation relations (24), we simply notice that

$$\mathbf{J} = \frac{\hbar}{2}\boldsymbol{\sigma} \tag{25}$$

will do the trick. This is because of the commutation relations for the Pauli matrices,

$$[\sigma_i,\sigma_j] = 2i\,\epsilon_{ijk}\,\sigma_k. \tag{26}$$

Therefore we provisionally take the rotation operators to be

$$U(\hat{\mathbf{n}},\theta) = e^{-i\theta\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}/2} = \cos\frac{\theta}{2} - i(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sin\frac{\theta}{2}, \tag{27}$$

where we use the standard properties of the Pauli matrices to reexpress the Taylor series for the exponential in terms of trigonometric functions (see Prob. 1.1).

The identification of the operators $U(\hat{\mathbf{n}},\theta)$ with rotations is provisional because we must show that these operators make physical sense as rotations. For example, let us consider the Stern-Gerlach experiment, in which we measure the components of the magnetic moment vector $\boldsymbol{\mu}$. We showed earlier that the operators corresponding to the components of $\boldsymbol{\mu}$, when represented in an eigenbasis of μ_z with appropriate phase conventions, produce matrices proportional to the Pauli matrices $\boldsymbol{\sigma}$. Therefore, since we expect $\boldsymbol{\mu}$ transform as a vector under rotations, so should $\boldsymbol{\sigma}$. This means that if we have an (old) state $|\psi\rangle$, and a new or rotated state $|\psi'\rangle = U(\hat{\mathbf{n}},\theta)|\psi\rangle$, then we expect that the expectation values of $\boldsymbol{\sigma}$ in the old and new states should be related by the classical rotation $\mathsf{R}(\hat{\mathbf{n}},\theta)$. In other words, we should have

$$\langle\psi'|\boldsymbol{\sigma}|\psi'\rangle = \langle\psi|U^\dagger\boldsymbol{\sigma}U|\psi\rangle = \mathsf{R}\langle\psi|\boldsymbol{\sigma}|\psi\rangle, \tag{28}$$

or, since this must hold for all $|\psi\rangle$,

$$U^\dagger\boldsymbol{\sigma}U = \mathsf{R}\boldsymbol{\sigma}, \tag{29}$$

where it is understood that $U = U(\hat{\mathbf{n}},\theta)$ and $\mathsf{R} = \mathsf{R}(\hat{\mathbf{n}},\theta)$. In case Eq. (29) it not clear, we write it out in components,

$$U^\dagger\sigma_i U = \sum_j R_{ij}\sigma_j. \tag{30}$$

To see if Eq. (29) is true, we simply substitute Eq. (27) to obtain

$$\begin{aligned} U^\dagger\boldsymbol{\sigma}U &= \Big[\cos\frac{\theta}{2} + i(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sin\frac{\theta}{2}\Big]\boldsymbol{\sigma}\Big[\cos\frac{\theta}{2} - i(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sin\frac{\theta}{2}\Big] \\ &= \cos^2\frac{\theta}{2}\boldsymbol{\sigma} + i\cos\frac{\theta}{2}\sin\frac{\theta}{2}\big[\hat{\mathbf{n}}\cdot\boldsymbol{\sigma},\boldsymbol{\sigma}\big] + (\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\boldsymbol{\sigma}(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sin^2\frac{\theta}{2}. \end{aligned} \tag{31}$$

Next we use the properties of the Pauli matrices to derive the identities,

$$\big[\hat{\mathbf{n}}\cdot\boldsymbol{\sigma},\boldsymbol{\sigma}\big] = -2i\,\hat{\mathbf{n}}\times\boldsymbol{\sigma}, \qquad (\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\boldsymbol{\sigma}(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}) = 2\hat{\mathbf{n}}(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}) - \boldsymbol{\sigma}, \tag{32}$$

which we use to rewrite Eq. (31) in the form,

$$U^\dagger\boldsymbol{\sigma}U = \cos\theta\,\boldsymbol{\sigma} + (1-\cos\theta)\hat{\mathbf{n}}(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}) + \sin\theta\,\hat{\mathbf{n}}\times\boldsymbol{\sigma}. \tag{33}$$

Finally, by comparing this with Eq. (11.44), we see that Eq. (29) is verified. Encouraged by this result, we henceforth consider the operators $U(\hat{\mathbf{n}},\theta)$ defined by Eq. (27) to be rotation operators for spin-$\frac{1}{2}$ systems.

8. The Spin-$\frac{1}{2}$ Adjoint Formula

Before leaving the result (29), however, we will comment on it further, since it is useful in its own right. In fact, this result is a spinor version of the adjoint formula (11.48), which we derived

earlier for classical rotations. To see the analogy more clearly, we first replace U by $U^{-1} = U^\dagger$ and R by $\mathsf{R}^{-1} = \mathsf{R}^t$ in Eq. (29), so that

$$\boxed{U\boldsymbol{\sigma}U^\dagger = \mathsf{R}^{-1}\boldsymbol{\sigma}.} \tag{34}$$

Next we dot both sides by some vector $\mathbf{a}$, and use Eq. (11.72) to obtain

$$U(\mathbf{a}\cdot\boldsymbol{\sigma})U^\dagger = (\mathsf{R}\mathbf{a})\cdot\boldsymbol{\sigma}. \tag{35}$$

This may be compared to the adjoint formula (11.48) for classical rotations, which we reproduce here with a slight change of notation:

$$\mathsf{R}(\mathbf{a}\cdot\mathbf{J})\mathsf{R}^t = (\mathsf{R}\mathbf{a})\cdot\mathbf{J}. \tag{36}$$

We will refer to Eq. (34) or its variants as the adjoint formula for spinor rotations.

9. The Double-Valued Representation

We seem to be in good shape for the interpretation of the operators $U(\hat{\mathbf{n}},\theta)$ as spinor rotations. There is, however, one wrinkle, which we find upon looking at specific examples of rotations. In particular, if we rotate a spinor about some axis $\hat{\mathbf{n}}$ by the angles of $\theta = 0$ and $\theta = 2\pi$, we find

$$U(\hat{\mathbf{n}},0) = 1, \qquad U(\hat{\mathbf{n}},2\pi) = -1, \tag{37}$$

according to Eq. (27). We see that the spinor of an electron or other spin-$\frac{1}{2}$ particle rotated by 2π does not return to its original value, but rather undergoes a phase change of -1.

The fact that a 2π rotation is not the identity operation is a nonclassical effect, and we must first ask what the physical significance is (in particular, whether it has any physical consequences). Certainly an overall phase factor of a quantum state has no physical significance, but it is possible to split a beam of spin-$\frac{1}{2}$ particles into two, and to subject one of the resulting beams to a rotation by 2π, whereupon the phase shift becomes observable in the interference pattern that results when the beams are recombined. This experiment has actually been performed with neutron beams, which are split and recombined by means of a neutron interferometer (essentially a large silicon crystal used as a kind of neutron diffraction grating). These experiments are discussed in more detail by Sakurai, *Modern Quantum Mechanics*, and they show that we must take the -1 phase shift for 2π rotations to be real.

The fact that a 2π rotation is not the identity operation on spinors of spin-$\frac{1}{2}$ systems means that we must reconsider our original quest for a representation of the classical rotations by means of unitary operators acting on a ket space, as laid out by Eqs. (1)–(5). In fact, the U operators defined by Eq. (27) do not form a representation of the classical rotations in the strict sense of the word, simply because they are not parameterized by the classical rotations. That is, the U operators are not a function of the R matrices, at least not in the sense of a single-valued function; this is clear from the special case of

$$\mathsf{R}(\hat{\mathbf{n}},0) = \mathsf{R}(\hat{\mathbf{n}},2\pi) = \mathsf{I}, \tag{38}$$

a single classical rotation for which there are two unitary operators, shown by Eq. (37). More generally, it can be shown that corresponding to every classical rotation R there are two unitary spinor rotations,

$$\mathsf{R}(\hat{\mathbf{n}},\theta) \mapsto \begin{cases} U(\hat{\mathbf{n}},\theta), \\ U(\hat{\mathbf{n}},\theta+2\pi) = -U(\hat{\mathbf{n}},\theta), \end{cases} \tag{39}$$

which replaces Eq. (1). The two unitary operators U corresponding to a given R differ by a sign. We see that the association between classical and spinor rotations is not one-to-one, but rather one-to-two.

In view of this, notation such as $U(\mathsf{R})$ is not really proper for spin-$\frac{1}{2}$ rotations, without some understanding as to which of the two unitary operators is meant. [On the other hand, the notation $U(\hat{\mathbf{n}},\theta)$ is unambiguous, as indicated by Eq. (27).] For example, the representation law (4) could be rewritten in the form,

$$U(\mathsf{R}_1)U(\mathsf{R}_2) = \pm U(\mathsf{R}_1\mathsf{R}_2), \tag{40}$$

which would mean that if we take one of the two unitary operators corresponding to R_1 and R_2 and multiply them, we will obtain one of the two unitary operators corresponding to $\mathsf{R}_1\mathsf{R}_2$. With this interpretation, Eq. (40) is correct for spinor rotations (but it is awkward and can lead to confusion).

10. The Group Manifolds $SU(2)$ and $SO(3)$

We explained in Notes 11 that the Baker-Campbell-Hausdorff theorem guaranteed that the group composition or multiplication law for finite operations was effectively contained in the commutation relations. This was why we focused first on finding a representation of the commutation relations (24), a task that will occupy us at greater length in the next set of notes, and why we were confident that when we exponentiated linear combinations of the angular momentum operators we would obtain operators that would reproduce the composition law of the classical rotations. What, then, has gone wrong, that we should end up with a double-valued representation, so that Eq. (4) must be replaced by Eq. (40)? The answer is that the spinor representation of the rotations is locally one-to-one, but globally one-to-two.

To explain this statement more precisely, we need to discuss the group $SU(2)$. The notation $SU(2)$ is standard in mathematical physics for the group of 2×2 complex unitary matrices with determinant $+1$. As in the notation $SO(3)$, the S stands for "special," which means the determinant is $+1$.

The significance of $SU(2)$ in the present discussion is that every spinor rotation $U(\hat{\mathbf{n}},\theta)$ defined by Eq. (27) for some axis $\hat{\mathbf{n}}$ and some angle θ is a member of the group $SU(2)$; and, conversely, every member of the group $SU(2)$ can be written in the form $U(\hat{\mathbf{n}},\theta)$ for some axis $\hat{\mathbf{n}}$ and some angle θ. We will not prove these facts; the proofs are easy, and are left as exercises. But we note the following consequence. Namely, since any group is closed under multiplication, if we form the product of two spinor rotations of the form $U(\hat{\mathbf{n}},\theta)$, we obtain another spinor rotation of the same form. Thus, every spinor rotation can be written in axis-angle form, and (like the classical rotations)

there is no loss of generality in assuming this form for a spinor rotation. We see that $SU(2)$ *is* the group of spinor rotations.

To understand the one-to-two association between $SO(3)$ and $SU(2)$ more thoroughly, it helps to view things geometrically in terms of the respective group manifolds. As explained in Notes 11, the group manifold for $SO(3)$ can be seen as a 3-dimensional surface living in the 9-dimensional space of all 3×3 real matrices. Similarly, the group manifold for $SU(2)$ can be seen as a surface living in the space of all 2×2 complex matrices. Since a 2×2 complex matrix has 4 complex components, each with a real and imaginary part, it takes 8 real numbers to specify an arbitrary 2×2 complex matrix, and we can say that 2×2 complex matrix space is 8-dimensional. But the condition $U^\dagger U = 1$ constitutes 4 real constraints, and the condition $\det U = +1$ is one more real constraint, for a total of 5 constraints on 8 variables. Therefore the group manifold $SU(2)$ can be seen as a 3-dimensional surface living in the 8-dimensional space of all 2×2 complex matrices. We see that both group manifolds $SO(3)$ and $SU(2)$ are 3-dimensional; this is also evident from the axis-angle parameterization of $SU(2)$ matrices, which involves 3 real parameters.

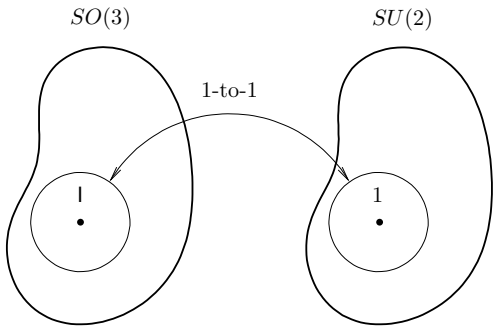

Fig. 1. There exist finite neighborhoods of the identity elements in the two groups, $SO(3)$ and $SU(2)$, which can be placed into one-to-one correspondence in such as way that the composition law is reproduced, according to Eq. (4). But the neighborhoods cannot be expanded to cover the whole group manifold without losing the one-to-one correspondence.

The identity matrix $\mathsf{R} = \mathsf{I}$ is one point of interest on the group manifold $SO(3)$, and the identity $U = 1$ is one point of interest on the group manifold $SU(2)$. These two points are associated with one another by the requirement (3). Next, once we have found a representation of the angular momentum commutation relations by Eq. (25), we have a one-to-one correspondence between infinitesimal neighborhoods of the respective identity elements, as indicated by Eq. (13). Next, the Baker-Campbell-Hausdorff theorem guarantees that by exponentiation we obtain a one-to-one correspondence between finite neighborhoods of the identity elements, which moreover satisfies the representation law (4). This is illustrated in Fig. 1, and it is in this sense that we say that the representation of $SO(3)$ by $SU(2)$ is locally one-to-one. But if we try to expand the two neighborhoods on the two group manifolds, we will find that when the neighborhood in $SO(3)$ has covered the whole group manifold, the corresponding neighborhood in $SU(2)$ has only covered half of that group

manifold. In effect, $SU(2)$ is twice as big as $SO(3)$, corresponding to the fact that the periodicity of spinor rotations about a fixed axis is 4π, not 2π. Thus, we say that globally the representation is one-to-two.

11. Explicit Relationship Between $SU(2)$ and $SO(3)$

The one-to-two character of the spinor represenation of rotations can be seen in another way. Let us return to the spinor adjoint formula (29), which we write in the form

$$U^\dagger \sigma_i U = \sum_j R_{ij}\, \sigma_j. \tag{41}$$

We multiply this equation on the right by σ_k and take traces, using the identity

$$\mathrm{tr}(\sigma_j \sigma_k) = 2\delta_{jk}, \tag{42}$$

to obtain

$$R_{ij} = \frac{1}{2}\,\mathrm{tr}\big(U^\dagger \sigma_i U \sigma_j\big). \tag{43}$$

The significance of this result is that it is an explicit formula giving, not U as a function of R, but rather R as a function of U. We see that since the right hand side is quadratic in U, both U and $-U$ correspond to the same R. Furthermore, it is straightforward to show explicitly from this formula that

$$\mathsf{R}(U_1)\mathsf{R}(U_2) = \mathsf{R}(U_1 U_2). \tag{44}$$

See Prob. 2.

Thus, although we started out looking for representations $U = U(\mathsf{R})$ of the classical rotations by unitary operators, in the case of spin-$\frac{1}{2}$ systems what we have found instead is a representation of unitary spin rotation operators by classical rotations, $\mathsf{R} = \mathsf{R}(U)$. This suggests that the spin rotation group $SU(2)$ is really the more fundamental group, and that the general theory of rotations is best formulated with $SU(2)$ as the starting point. This indeed is the most useful point of view in quantum mechanics, and it has its advantages even in purely classical problems.

12. The Cayley-Klein Parameters

The Cayley-Klein parameters are a set of parameters for representing rotations, either classical or spinor. They were discovered in the nineteenth century before the advent of quantum mechanics, and were originally intended for use in classical problems of rigid body motion. (They are still used for that purpose.) Pauli himself was familiar with the theory of the Cayley-Klein parameters, which apparently helped him to discover the matrices that now bear his name, and to get credit for the theory of electron spin.

To see how the Cayley-Klein parameters come about, we first write an arbitrary 2×2 complex matrix in terms of its four complex components,

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{45}$$

Next, the constraint that U be unitary is equivalent to the demand that the rows of U form a pair of orthonormal unit vectors, or,

$$|a|^2 + |b|^2 = 1, \tag{46a}$$

$$|c|^2 + |d|^2 = 1, \tag{46b}$$

$$a^* c + b^* d = 0. \tag{46c}$$

Also, the requirement $\det U = 1$ is equivalent to

$$\det U = ad - bc = 1. \tag{47}$$

Now if we take Eqs. (46c) and (47) and solve for c and d, assuming a and b are given, we find $c = -b^*$, $d = a^*$, so that an arbitrary element of $SU(2)$ can be written in the form,

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \tag{48}$$

where a and b are complex numbers satisfying the constraint (46a). Finally, if we break a and b into their real and imaginary parts according to

$$a = x_0 + ix_3, \qquad b = x_2 + ix_1, \tag{49}$$

then an arbitrary element of $SU(2)$ has the form

$$U = \begin{pmatrix} x_0 + ix_3 & x_2 + ix_1 \\ -x_2 + ix_1 & x_0 - ix_3 \end{pmatrix} = x_0 + i(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3) = x_0 + i\mathbf{x}\cdot\boldsymbol{\sigma}, \tag{50}$$

where the four real numbers (x_0, x_1, x_2, x_3) satisfy the constraint

$$x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1. \tag{51}$$

The parameters (x_0, x_1, x_2, x_3) are the Cayley-Klein parameters, in terms of which an arbitrary spinor rotation is represented by Eq. (50). An arbitrary classical rotation can also be written in terms of Cayley-Klein parameters, by using Eq. (43) to write R in terms of U. One might ask why we should parameterize rotations by four parameters, subject to one constraint, when we could use three parameters subject to no constraints. The answer is that the various options for three parameters, such as the Euler angles, are unsymmetrical and do not cover the group manifold without introducing coordinate singularities (similar to the singularity in spherical coordinates at the north pole). The lack of symmetry of the Euler angles quickly leads to ugly calculations, and coordinate singularities are inconvenient for many purposes (computer programs, for example).

The constraint (51) is interesting, for it shows that the group manifold $SU(2)$ can be seen as a 3-dimensional surface living, not in 8-dimensional matrix space, as earlier, but in a 4-dimensional space with coordinates (x_1, x_2, x_3, x_4). Furthermore, the surface in question is simply the set of all points in this 4-dimensional space at a unit distance from the origin; the surface is the 3-dimensional surface of a sphere in 4-dimensional space, or the manifold S^3 in standard mathematical terminology.

13. Spinor Rotations about the Coordinate Axes

Let us now tabulate the spinor rotations about the three coordinate axes, much as we did in Eq. (11.18) for classical rotations. We have

$$\begin{aligned} U(\hat{\mathbf{x}},\theta) &= \cos\frac{\theta}{2} - i\sigma_x \sin\frac{\theta}{2} = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \\ U(\hat{\mathbf{y}},\theta) &= \cos\frac{\theta}{2} - i\sigma_y \sin\frac{\theta}{2} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \\ U(\hat{\mathbf{z}},\theta) &= \cos\frac{\theta}{2} - i\sigma_z \sin\frac{\theta}{2} = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}. \end{aligned} \tag{52}$$

The matrix for $U(\hat{\mathbf{z}},\theta)$ is diagonal, because these matrices represent the corresponding operators in the basis of eigenkets of $S_z = (\hbar/2)\sigma_z$. Obviously, the exponential of a diagonal matrix is diagonal.

The elementary rotations in Eq. (52) can be combined to obtain an Euler angle parameterization for spinor rotations. This is a direct transcription of Eq. (11.57),

$$U(\alpha,\beta,\gamma) = U(\hat{\mathbf{z}},\alpha)U(\hat{\mathbf{y}},\beta)U(\hat{\mathbf{z}},\gamma), \tag{53}$$

and is just the spinor representative of the latter. However, the ranges on the Euler angles are different from the classical case; here we have

$$\begin{aligned} 0 &\le \alpha \le 2\pi, \\ 0 &\le \beta \le \pi, \\ 0 &\le \gamma \le 4\pi, \end{aligned} \tag{54}$$

where the final angle γ is allowed to range to 4π to cover the extra spinor rotations.

14. A Spin "Pointing In" a Given Direction

You have no doubt heard the expression, "a spinor pointing in the such-and-such direction." What does this language mean, considering that a spinor for spin-$\frac{1}{2}$ particle is a complex 2-vector, not a real 3-vector? To explain this terminology, we introduce the eigenbasis of the operator S_z, which we denote by $|\pm\rangle$. These kets are of course represented by unit vectors in the S_z basis itself,

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{55}$$

where to be precise the kets are not equal to the column vectors indicated, rather the components of the kets in the S_z-basis are given. The 2-vectors corresponding to $|+\rangle$ and $|-\rangle$ are often denoted in the literature by α and β (the "spin up" and "spin down" spinors, respectively).

To begin, we simply declare that the ket $|+\rangle$ is the spinor "pointing in the $\hat{\mathbf{z}}$ direction." Next, to obtain a spinor pointing in an arbitrary direction $\hat{\mathbf{n}}$, we first consider a classical rotation, say, R_0, which maps the $\hat{\mathbf{z}}$ direction into the $\hat{\mathbf{n}}$ direction,

$$\hat{\mathbf{n}} = \mathsf{R}_0\hat{\mathbf{z}}. \tag{56}$$

A rotation R_0 with this property is easy to write down in Euler angle form; we simply let α be the azimuthal angle of $\hat{\mathbf{n}}$ and β the polar angle, so that R_0 has the form

$$\mathsf{R}_0 = \mathsf{R}(\alpha, \beta, 0) = \mathsf{R}(\hat{\mathbf{z}}, \alpha)\mathsf{R}(\hat{\mathbf{y}}, \beta). \tag{57}$$

The rotation R_0 that satisfies Eq. (56) is not unique; we could allow any value of the Euler angle γ (not just $\gamma = 0$). But R_0 in Eq. (57) will work. Then to obtain the spinor "pointing in" the $\hat{\mathbf{n}}$ direction (call it $|\hat{\mathbf{n}}; +\rangle$), we simply define

$$|\hat{\mathbf{n}}; +\rangle = U_0|+\rangle, \tag{58}$$

where $U_0 = U(\mathsf{R}_0)$. We don't care about the overall phase of this spinor, which is why we can ignore the Euler angle γ, and why we don't care which of the two U_0's is chosen in Eq. (58).

It is easy to work out the 2-vector representing $|\hat{\mathbf{n}}; +\rangle$ in the S_z basis. We simply write out the Euler angle representation of U_0,

$$U_0 = U(\hat{\mathbf{z}}, \alpha)U(\hat{\mathbf{y}}, \beta), \tag{59}$$

and appeal to the matrices (52). We find

$$|\hat{\mathbf{n}}; +\rangle = \begin{pmatrix} e^{-i\alpha/2}\cos\frac{\beta}{2} \\ e^{i\alpha/2}\sin\frac{\beta}{2} \end{pmatrix}. \tag{60}$$

Since the overall phase is immaterial, we can multiply this by $e^{\pm i\alpha/2}$, if we like, to clear one or the other of the two phase factors in the 2-vector.

The spinor $|\hat{\mathbf{n}}; +\rangle$ has several notable properties. First, it is an eigenspinor of $\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}$, that is, of the component of the spin in the $\hat{\mathbf{n}}$ direction, with eigenvalue $+1$. This is easily proved with the help of the spinor adjoint formula (29):

$$\begin{aligned} \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}|\hat{\mathbf{n}}; +\rangle &= U_0 U_0^\dagger (\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}) U_0|+\rangle = U_0 \hat{\mathbf{n}} \cdot (\mathsf{R}_0 \boldsymbol{\sigma})|+\rangle \\ &= U_0\big(\mathsf{R}_0^{-1}\hat{\mathbf{n}}\big) \cdot \boldsymbol{\sigma}|+\rangle = U_0(\hat{\mathbf{z}} \cdot \boldsymbol{\sigma})|+\rangle \\ &= U_0\sigma_z|+\rangle = U_0|+\rangle = |\hat{\mathbf{n}}; +\rangle, \end{aligned} \tag{61}$$

where we use Eq. (56). Next, the expectation value of the spin in any direction orthogonal to $\hat{\mathbf{n}}$, taken with respect to $|\hat{\mathbf{n}}; +\rangle$, vanishes, as indicated by

$$\langle \hat{\mathbf{n}}; +|\boldsymbol{\sigma}|\hat{\mathbf{n}}; +\rangle = \hat{\mathbf{n}}. \tag{62}$$

To prove this we again use the adjoint formula to reexpress the left hand side,

$$\langle \hat{\mathbf{n}}; +|\boldsymbol{\sigma}|\hat{\mathbf{n}}; +\rangle = \langle +|U_0^\dagger \boldsymbol{\sigma} U_0|+\rangle = \mathsf{R}_0\langle +|\boldsymbol{\sigma}|+\rangle. \tag{63}$$

But the final expectation value in this expression is a vector whose x, y and z components are 0, 0, and 1, as a direct appeal to the Pauli matrices will show. That is, this vector is the unit vector $\hat{\mathbf{z}}$, so the right hand side of Eq. (63) becomes $\hat{\mathbf{n}}$ in accordance with Eq. (56). This proves Eq. (62).

It is a fact that for a spin-$\frac{1}{2}$ system, every spinor "points in" some direction, that is, every spinor is an eigenspinor of $\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}$ for some $\hat{\mathbf{n}}$. This is not true for values of the spin greater than $\frac{1}{2}$.

This concludes what we have to say about rotations on spin-$\frac{1}{2}$ systems. In the next notes we consider the general problem of constructing representations of the angular momentum commutation relations and the corresponding representations of rotations.

Problems

1. This is problem 3.8 of Sakurai, *Modern Quantum Mechanics*. Let

$$U(\alpha, \beta, \gamma) = U(\hat{\mathbf{z}}, \alpha) U(\hat{\mathbf{y}}, \beta) U(\hat{\mathbf{z}}, \gamma) \tag{64}$$

be the Euler angle representation of a rotation on a spin-$\frac{1}{2}$ system. Find the axis $\hat{\mathbf{n}}$ and angle θ of this rotation in terms of the Euler angles.

2. The Pauli matrices combined with the 2×2 identity matrix span the space of 2×2 matrices, that is, an arbitrary 2×2 matrix A can be written in the form,

$$A = a_0 I + \mathbf{a} \cdot \boldsymbol{\sigma}, \tag{65}$$

where $(a_0, \mathbf{a})$ are the (generally complex) expansion coefficients. See Prob. 1.3(c). Notice that if A is Hermitian, then $(a_0, \mathbf{a})$ are real. Notice also that

$$\det A = a_0^2 - \mathbf{a} \cdot \mathbf{a}. \tag{66}$$

It is convenient to write $\sigma_0 = I$, and to regard σ_0 as a fourth Pauli matrix. Then Eq. (65) becomes,

$$A = \sum_{\mu=0}^{3} a_\mu \, \sigma_\mu. \tag{67}$$

Notice that we have the orthogonality relation,

$$\operatorname{tr}(\sigma_\mu \sigma_\nu) = 2\delta_{\mu\nu}, \tag{68}$$

which can be used to solve Eq. (67) for the expansion coefficients,

$$a_\mu = \frac{1}{2} \operatorname{tr}(\sigma_\mu A). \tag{69}$$

Use these results to show that

$$\operatorname{tr}(AB) = \frac{1}{2} \sum_{\mu} \operatorname{tr}(\sigma_\mu A) \operatorname{tr}(\sigma_\mu B), \tag{70}$$

where A and B are 2×2 matrices.

Use these results to prove Eq. (44).

3. This problem concerns the polarization states of classical electromagnetic waves, which can be described by the same mathematical formalism used for spinors of spin-$\frac{1}{2}$ particles. It is also provides a good background for the subject of the polarization states of photons, which we will take up later in the course.

(a) The spinor "pointing in" the $\hat{\mathbf{n}}$ direction was defined by Eq. (58). Show that every spinor "points" in some direction. For this, it is sufficient to show that for every normalized spinor $|\chi\rangle$, $\langle\chi|\boldsymbol{\sigma}|\chi\rangle$ is a unit vector. You can prove this directly, or use the formalism based on Eq. (70) above. This property only holds for spin-$\frac{1}{2}$ particles.

(b) Consider now the phenomenon of polarization in classical electromagnetic theory. The most general physical electric field of a plane light wave of frequency ω travelling in the z-direction can be written in the form

$$\mathbf{E}_{\text{phys}}(\mathbf{r},t) = \text{Re}\sum_{\mu=1}^{2}\hat{\boldsymbol{\epsilon}}_\mu E_\mu e^{i(kz-\omega t)}, \tag{71}$$

where $\mu = 1,2$ corresponds to x, y, where $\hat{\boldsymbol{\epsilon}}_1 = \hat{\mathbf{x}}$, $\hat{\boldsymbol{\epsilon}}_2 = \hat{\mathbf{y}}$, and where $E_1 = E_x$ and $E_2 = E_y$ are two complex amplitudes, and where $\omega = ck$. Thus, the wave is parameterized by the two complex numbers, E_x, E_y. Often we are not interested in absolute amplitudes, only relative ones, so we normalize the wave by setting

$$|E_x|^2 + |E_y|^2 = E_0^2, \tag{72}$$

for some suitably chosen reference amplitude E_0. This allows us to associate the wave with a normalized, 2-component complex "spinor," according to

$$\chi = \frac{1}{E_0}\begin{pmatrix} E_x \\ E_y \end{pmatrix}. \tag{73}$$

Furthermore, we are often not interested in any overall phase of this spinor, since such an overall phase corresponds merely to a shift in the origin of time in Eq. (71). This cannot be detected anyway in experiments that average over the rapid oscillations of the wave (a practical necessity at optical frequencies).

If we normalize according to Eqs. (72) and (73) and ignore the overall phase, then the four real parameters originally inherent in (E_x, E_y) are reduced to two, which describe the state of *polarization* of the wave. For example, the spinors

$$\chi_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{74}$$

correspond to linear polarization in the x- and y-directions, respectively. Notice that polarization in the $+x$-direction is the same as that in the $-x$-direction; they differ only by an overall phase. Similarly, the spinors

$$\chi_r = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}, \quad \chi_\ell = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}, \tag{75}$$

correspond to right and left circular polarizations, respectively. Right circular polarization means the electric vector rotates in a clockwise direction in the x-y plane, tracing out a circle, whereas in left circular polarization, the electric vector rotates counterclockwise. Here I am using the conventions of Jackson, *Classical Electrodynamics* and Born and Wolf, *Principles of Optics*, but other authors, such as Sakurai, *Modern Quantum Mechanics*, use the opposite conventions. Right circular polarization corresponds to photons with negative helicity, and vice versa. In more general polarization states, the electric vector traces out an ellipse in the x-y plane; these are called elliptical polarizations. A limiting case of the ellipse is where the electric vector traces a line, back and forth; these are linear polarization states.

In optics it is conventional to introduce the so-called *Stokes' parameters* to describe the state of polarization. These are defined by

$$\begin{aligned} s_0 &= (|E_x|^2 + |E_y|^2)/E_0^2 = 1, \\ s_1 &= (|E_x|^2 - |E_y|^2)/E_0^2, \\ s_2 &= 2\,\mathrm{Re}(E_x E_y^*)/E_0^2, \\ s_3 &= 2\,\mathrm{Im}(E_x E_y^*)/E_0^2. \end{aligned} \tag{76}$$

(See Born and Wolf, *Principles of Optics*, p. 31.) Show that these parameters satisfy

$$s_1^2 + s_2^2 + s_3^2 = 1, \tag{77}$$

so that $\hat{\mathbf{s}} = (s_1, s_2, s_3)$ is a unit vector. The sphere upon which this unit vector lies is called the *Poincaré sphere*; points on this sphere correspond to polarization states. Notice that the Stokes' parameters are independent of the overall phase of the wave, being bilinear in the field amplitudes (E_x, E_y). Indicate which points on the Poincaré sphere correspond to linear x- and y-polarization, and which to right and left circular polarization. What kind of polarization does the positive 2-axis in $\mathbf{s}$-space correspond to? What about the negative 2-axis? (We will refer to directions in $\mathbf{s}$-space by the indices 1,2,3, to avoid confusion with x, y, z in real space).

(c) Compute the expectation value $\langle\chi|\boldsymbol{\sigma}|\chi\rangle = \hat{\mathbf{n}}$ for the spinor (73), and relate the components of $\hat{\mathbf{n}}$ to the Stokes' parameters. You will see that Stokes and Poincaré didn't exactly follow quantum mechanical conventions (since quantum mechanics had not yet been invented in their day), but the basic idea is that the point on the Poincaré sphere indicates the direction in which the spinor (73) is "pointing."

(d) Now suppose we have a quarter-wave plate with its fast and slow axes in the x- and y-direction, respectively. This causes a relative phase shift in the x- and y-components of the spinor (73) by $\pi/2$, that is,

$$\begin{pmatrix} E_x' \\ E_y' \end{pmatrix} = \begin{pmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix}, \tag{78}$$

where the unprimed fields are those entering the quarter wave plate, and the primed ones are those exiting it. Show that the effect of the quarter wave plate on an incoming polarization state, as

represented by a point on the Poincaré sphere, can be represented by a rotation in **s**-space. Find the 3×3 rotation matrix such that $\hat{\mathbf{s}}' = \mathsf{R}\hat{\mathbf{s}}$. Use this picture to determine the effect of the quarter wave plate on linear x- and y- polarizations, and on right and left circular polarizations. What polarization must we feed into the quarter wave plate to get right circular polarization coming out?

Physics 221A
Academic Year 2020–21
Notes 19
Transformation of Operators Under Rotations†

1. Introduction

In both classical and quantum mechanics we classify observables according to how they transform under symmetries, including proper rotations, improper rotations, Lorentz transformations, etc. For example, we speak of scalars, vectors, tensors, pseudoscalars, pseudovectors, etc. In these notes we examine how operators transform under proper rotations, developing a classification scheme that will be important for understanding rotational symmetry in quantum mechanics. We will take up the case of improper rotations in Notes 21, time reversal in Notes 22, and Lorentz transformations in Notes 48. The material of these notes is background on the Wigner-Eckart theorem, which we take up in Notes 20.

2. Definition of a Rotated Operator

We consider a quantum mechanical system with a ket space upon which rotation operators $U(\mathsf{R})$, forming a representation of the classical rotation group $SO(3)$, are defined. The representation will be double-valued if the angular momentum of the system is a half-integer. In these notes we consider only proper rotations R; improper rotations will be taken up later. The operators $U(\mathsf{R})$ map kets into new or rotated kets,

$$|\psi'\rangle = U(\mathsf{R})|\psi\rangle, \tag{1}$$

where $|\psi'\rangle$ is the rotated ket. We will also write this as

$$|\psi\rangle \xrightarrow{\mathsf{R}} U(\mathsf{R})|\psi\rangle. \tag{2}$$

In the case of half-integer angular momenta, the mapping above is only determined to within a sign by the classical rotation R.

Now if A is an operator, we define the *rotated operator* A' by requiring that the expectation value of the original operator with respect to the initial state be equal to the expectation value of the rotated operator with respect to the rotated state, that is,

$$\langle\psi'|A'|\psi'\rangle = \langle\psi|A|\psi\rangle, \tag{3}$$

† Links to the other sets of notes can be found at:
`http://bohr.physics.berkeley.edu/classes/221/2021/221.html`.

which is to hold for all initial states $|\psi\rangle$. But this implies

$$\langle\psi|U(\mathsf{R})^\dagger\, A'\, U(\mathsf{R})|\psi\rangle = \langle\psi|A|\psi\rangle, \tag{4}$$

or, since $|\psi\rangle$ is arbitrary [see Prob. 1.6(b)],

$$U(\mathsf{R})^\dagger\, A'\, U(\mathsf{R}) = A. \tag{5}$$

Solving for A', this becomes

$$A' = U(\mathsf{R})\, A\, U(\mathsf{R})^\dagger, \tag{6}$$

which is our definition of the rotated operator. We will also write this in the form,

$$A \xrightarrow{\mathsf{R}} U(\mathsf{R})\, A\, U(\mathsf{R})^\dagger. \tag{7}$$

Notice that in the case of half-integer angular momenta the rotated operator is specified by the $SO(3)$ rotation matrix R alone, since the sign of $U(\mathsf{R})$ cancels and the answer does not depend on which of the two rotation operators is used on the right hand side. This is unlike the case of rotating kets, where the sign does matter. Equation (7) defines the action of rotations on operators.

3. Scalar Operators

Now we classify operators by how they transform under rotations. First we define a *scalar operator* K to be an operator that is invariant under rotations, that is, that satisfies

$$\boxed{U(\mathsf{R})\, K\, U(\mathsf{R})^\dagger = K,} \tag{8}$$

for all operators $U(\mathsf{R})$. This terminology is obvious. Notice that it is equivalent to the statement that a scalar operator commutes with all rotations,

$$[U(\mathsf{R}), K] = 0. \tag{9}$$

If an operator commutes with all rotations, then it commutes in particular with infinitesimal rotations, and hence with the generators $\mathbf{J}$. See Eq. (12.13). Conversely, if an operator commutes with $\mathbf{J}$ (all three components), then it commutes with any function of $\mathbf{J}$, such as the rotation operators. Thus another equivalent definition of a scalar operator is one that satisfies

$$\boxed{[\mathbf{J}, K] = 0.} \tag{10}$$

The most important example of a scalar operator is the Hamiltonian for an isolated system, not interacting with any external fields. The consequences of this for the eigenvalues and eigenstates of the Hamiltonian are discussed in Secs. 9 and 12 below.

4. Vector Operators

In ordinary vector analysis in three-dimensional Euclidean space, a vector is defined as a collection of three numbers that have certain transformation properties under rotations. It is not sufficient just to have a collection of three numbers; they must in addition transform properly. Similarly, in quantum mechanics, we define a *vector operator* as a vector *of* operators (that is, a set of three operators) with certain transformation properties under rotations.

Our requirement shall be that the expectation value of a vector operator, which is a vector of ordinary or c-numbers, should transform as a vector in ordinary vector analysis. This means that if $|\psi\rangle$ is a state and $|\psi'\rangle = U(\mathsf{R})|\psi\rangle$ is the rotated state as in Eq. (1), then

$$\langle\psi'|\mathbf{V}|\psi'\rangle = \mathsf{R}\langle\psi|\mathbf{V}|\psi\rangle, \tag{11}$$

which must hold if $\mathbf{V}$ is to be considered a genuine vector operator. In case the notation in Eq. (11) is not clear, we write the same equation out in components,

$$\langle\psi'|V_i|\psi'\rangle = \sum_j R_{ij}\langle\psi|V_j|\psi\rangle. \tag{12}$$

Equation (11) or (12) is to hold for all $|\psi\rangle$, so by Eq. (1) they imply

$$U(\mathsf{R})^\dagger\,\mathbf{V}\,U(\mathsf{R}) = \mathsf{R}\mathbf{V}, \tag{13}$$

or, (after swapping R and R^{-1})

$$U(\mathsf{R})\,\mathbf{V}\,U(\mathsf{R})^\dagger = \mathsf{R}^{-1}\mathbf{V}. \tag{14}$$

Inn components this is

$$\boxed{U(\mathsf{R})\,V_i\,U(\mathsf{R})^\dagger = \sum_j V_j\,R_{ji}.} \tag{15}$$

We will take Eq. (13), (14) or (15) as the definition of a vector operator.

In the case of a scalar operator, we had one definition (8) involving its properties under conjugation by rotations, and another (10) involving its commutation relations with the angular momentum $\mathbf{J}$. The latter is in effect a version of the former, when the rotation is infinitesimal. Similarly, for vector operators there is a definition equivalent to Eq. (14) or (15) that involves commutation relations with $\mathbf{J}$. To derive it we let U and R in Eq. (14) have axis-angle form with an angle $\theta \ll 1$, so that

$$U(\mathsf{R}) = 1 - \frac{i}{\hbar}\theta\hat{\mathbf{n}}\cdot\mathbf{J}, \tag{16}$$

and

$$\mathsf{R} = \mathsf{I} + \theta\hat{\mathbf{n}}\cdot\mathbf{J}. \tag{17}$$

See Eqs. (11.22) and (11.32) for the latter. Then the definition (14) becomes

$$\Big(1 - \frac{i}{\hbar}\theta\hat{\mathbf{n}}\cdot\mathbf{J}\Big)\mathbf{V}\Big(1 + \frac{i}{\hbar}\theta\hat{\mathbf{n}}\cdot\mathbf{J}\Big) = (\mathsf{I} - \theta\hat{\mathbf{n}}\cdot\mathbf{J})\mathbf{V}, \tag{18}$$

or

$$[\hat{\mathbf{n}} \cdot \mathbf{J}, \mathbf{V}] = -i\hbar\, \hat{\mathbf{n}} \times \mathbf{V}. \tag{19}$$

Taking the j-th component of this, we have

$$n_i [J_i, V_j] = -i\hbar\, \epsilon_{jik}\, n_i V_k, \tag{20}$$

or, since $\hat{\mathbf{n}}$ is an arbitrary unit vector,

$$\boxed{[J_i, V_j] = i\hbar\, \epsilon_{ijk}\, V_k.} \tag{21}$$

Any vector operator satisfies this commutation relation with the angular momentum of the system.

The converse is also true; if Eq. (21) is satisfied, then $\mathbf{V}$ is a vector operator. This follows since Eq. (21) implies Eq. (19) which implies Eq. (18), that is, it implies that the definition (14) is satisfied for infinitesimal rotations. But it is easy to show that if Eq. (14) is true for two rotations R_1 and R_2, then it is true for the product $\mathsf{R}_1\mathsf{R}_2$. Therefore, since finite rotations can be built up as the product of a large number of infinitesimal rotations (that is, as a limit), Eq. (21) implies Eq. (14) for all rotations. Equations (14) and (21) are equivalent ways of defining a vector operator.

We have now defined scalar and vector operators. Combining them, we can prove various theorems. For example, if $\mathbf{V}$ and $\mathbf{W}$ are vector operators, then $\mathbf{V} \cdot \mathbf{W}$ is a scalar operator, and $\mathbf{V} \times \mathbf{W}$ is a vector operator. This is of course just as in vector algebra, except that we must remember that operators do not commute, in general. For example, it is not generally true that $\mathbf{V} \cdot \mathbf{W} = \mathbf{W} \cdot \mathbf{V}$, or that $\mathbf{V} \times \mathbf{W} = -\mathbf{W} \times \mathbf{V}$.

If we wish to show that an operator is a scalar, we can compute its commutation relations with the angular momentum, as in Eq. (10). However, it may be easier to consider what happens when the operator is conjugated by rotations. For example, the central force Hamiltonian (16.1) is a scalar because it is a function of the dot products $\mathbf{p} \cdot \mathbf{p} = p^2$ and $\mathbf{x} \cdot \mathbf{x} = r^2$. See Sec. 16.2.

5. Examples of Vector Operators

Consider a system consisting of a single spinless particle moving in three-dimensional space, for which the wave functions are $\psi(\mathbf{x})$ and the angular momentum is $\mathbf{L} = \mathbf{x} \times \mathbf{p}$. To see whether $\mathbf{x}$ is a vector operator (we expect it is), we compute the commutation relations with $\mathbf{L}$, finding,

$$[L_i, x_j] = i\hbar\, \epsilon_{ijk}\, x_k. \tag{22}$$

According to Eq. (21), this confirms our expectation. Similarly, we find

$$[L_i, p_j] = i\hbar\, \epsilon_{ijk}\, p_k, \tag{23}$$

so that $\mathbf{p}$ is also a vector operator. Then $\mathbf{x} \times \mathbf{p}$ (see Sec. 4) must also be a vector operator, that is, we must have

$$[L_i, L_j] = i\hbar\, \epsilon_{ijk}\, L_k. \tag{24}$$

This last equation is of course just the angular momentum commutation relations, but here with a new interpretation. More generally, by comparing the adjoint formula (13.89) with the commutation relations (21), we see that the angular momentum $\mathbf{J}$ is always a vector operator.

6. If it Looks Like a Vector ...

Not everything that looks like a vector is a vector operator, according to our definition (15). Consider, for example, the Stark effect, in which an atom interacts with an external electric field $\mathbf{E}_0$. If the atom has a single electron, then the interaction term in the Hamiltonian is the potential

$$V_1(\mathbf{x}) = e\mathbf{E}_0 \cdot \mathbf{x}. \tag{25}$$

The vector $\mathbf{x}$ is a vector operator, as we have seen, but what about $\mathbf{E}_0$? It is in fact just a vector of c-numbers, which, regarded as multiplicative operators on wave functions, commute with everything. Therefore we have

$$U(\mathsf{R})\mathbf{E}_0 U(\mathsf{R})^\dagger = \mathbf{E}_0, \tag{26}$$

which does not satisfy the definition (15) of a vector operator. Since $\mathbf{E}_0$ is not a vector operator, $V_1(\mathbf{x})$ is not a scalar operator.

To understand this situation, we note that the rotation operators $U(\mathsf{R})$ rotate the physical observables of our system, such as the position $\mathbf{x}$ and momentum $\mathbf{p}$ of the electron. They do not rotate external parameters, such as $\mathbf{E}_0$. On the other hand, the external electric field $\mathbf{E}_0$ is created by some charges somewhere. Imagine, for example, that the atom is between the plates of a capacitor, with charges arrayed along the plates to create $\mathbf{E}_0$. If these charges are included in "the system," then when we rotate the system the field $\mathbf{E}_0$ is rotated as well. Then $\mathbf{E}_0$ becomes a vector operator and V_1 becomes a scalar operator.

We will see similar issues regarding the definition of "the system" when we come to conservation of parity and time reversal.

7. Tensor Operators

Finally we define a *tensor operator* as a tensor *of* operators with certain transformation properties that we will illustrate in the case of a rank-2 tensor. In this case we have a set of 9 operators T_{ij}, where $i, j = 1, 2, 3$, which can be thought of as a 3×3 matrix of operators. These are required to transform under rotations according to

$$U(\mathsf{R})\, T_{ij}\, U(\mathsf{R})^\dagger = \sum_{k\ell} T_{k\ell}\, R_{ki} R_{\ell j}, \tag{27}$$

which is a generalization of Eq. (15) for vector operators. As with scalar and vector operators, a definition equivalent to Eq. (27) may be given that involves the commutation relations of T_{ij} with the components of angular momentum.

As an example of a tensor operator, let $\mathbf{V}$ and $\mathbf{W}$ be vector operators, and write

$$T_{ij} = V_i W_j. \tag{28}$$

Then T_{ij} is a tensor operator (it is the tensor product of $\mathbf{V}$ with $\mathbf{W}$). This is just an example; in general, a tensor operator cannot be written as the product of two vector operators as in Eq. (28). In practice, however, many tensor operators appear as linear combinations of such products.

Consider, for example, the quadrupole moment operator. In a system with a collection of particles with positions $\mathbf{x}_\alpha$ and charges q_α, where α indexes the particles, the quadrupole moment operator is

$$Q_{ij} = \sum_\alpha q_\alpha (3 x_{\alpha i}\, x_{\alpha j} - r_\alpha^2\, \delta_{ij}). \tag{29}$$

This is obtained from Eq. (15.88) by setting

$$\rho(\mathbf{x}) = \sum_\alpha q_\alpha\, \delta(\mathbf{x} - \mathbf{x}_\alpha). \tag{30}$$

The quadrupole moment operator is especially important in nuclear physics, in which the particles are the protons in a nucleus with charge $q = e$. Notice that the first term under the sum (29) is an operator of the form (28), with $\mathbf{V} = \mathbf{W} = \mathbf{x}_\alpha$.

Tensor operators of other ranks (besides 2) are possible; a scalar is considered a tensor operator of rank 0, and a vector is considered a tensor of rank 1. In the case of tensors of arbitrary rank, the transformation law involves one copy of the matrix $\mathsf{R}^{-1} = \mathsf{R}^t$ for each index of the tensor.

8. Invariant and Irreducible Subspaces of Operators

Let us return to Eq. (15), the transformation law defining a vector operator. This says that if we rotate any one of the components of a vector operator, we get a linear combination of the three components of that vector operator. The same is true if we rotate any linear combination of the three components.

In this respect it is useful to think of the space of operators as a vector space, and the components of a vector operator as "basis operators" that span a 3-dimensional subspace of the space of operators. (We assume the vector operator is not identically zero.) Operators form a vector space because they can be added to one another and multiplied by scalars, that is, linear combinations of operators are operators.

Then the 3-dimensional subspace of operators spanned by the components of a vector operator is invariant under the action of rotations, which is defined by Eq. (7). This means that if we rotate any operator belonging to this subspace, we obtain another operator belonging to the same subspace, a fact that is a simple consequence of Eq. (15).

Similarly, the nine components T_{ij} of a tensor operator, if linearly independent, span a 9-dimensional space of operators that according to the definition (27) is invariant under the action of rotations.

Another example of an invariant subspace of operators is the 1-dimensional subspace spanned by any nonzero scalar operator K, as follows from the definition (8) of a scalar operator.

Thus, we have three examples of subspaces of operators that are invariant under rotations, those spanned by the components of a scalar operator, a vector operator, or a second-rank tensor operator, which are of dimensionality 1, 3 and 9 respectively.

Now in general when we have a subspace (of any vector space) that is invariant under rotations, we can ask if it possesses a smaller subspace that is also invariant. If it does, we say it is *reducible*; it not, it is *irreducible*. Invariant, reducible subspaces can always be broken up into smaller invariant, irreducible subspaces, so that the irreducible subspaces are the building blocks of reducible subspaces. It turns out that the invariant, irreducible subspaces are orthogonal to one another.

By our definition, every invariant, 1-dimensional subspace is automatically irreducible, because it contains no smaller subspace. Therefore the 1-dimensional subspace of operators spanned by a scalar operator is automatically irreducible.

What about the 3-dimensional subspace spanned by the components of any nonzero vector operator? It turns out that this, too, is irreducible. See Prob. 2.

What about the 9-dimensional space spanned by the components T_{ij} of a tensor operator (we assume the nine components are linearly independent)? It turns out this subspace is reducible. To see how, let us work with the example (28) of a tensor operator, $T_{ij} = V_i W_j$. We draw attention to the trace of this tensor,

$$\operatorname{tr} T = \sum_i T_{ii} = \sum_i V_i W_i = \mathbf{V} \cdot \mathbf{W}, \tag{31}$$

which as shown is the dot product of the two vector operators $\mathbf{V}$ and $\mathbf{W}$. This is a scalar, which is invariant under rotations, so the 9-dimensional space of operators spanned by the T_{ij} has a 1-dimensional, invariant subspace, spanned by $T_{11} + T_{22} + T_{33}$. This subspace, being 1-dimensional, is irreducible.

This leaves a complementary, 8-dimensional subspace of operators. Is this reducible? Yes, it possesses a 3-dimensional invariant subspace spanned by the operators

$$\begin{aligned} X_3 &= T_{12} - T_{21} = V_1 W_2 - V_2 W_1, \\ X_1 &= T_{23} - T_{32} = V_2 W_3 - V_3 W_2, \\ X_2 &= T_{31} - T_{13} = V_3 W_1 - V_1 W_3, \end{aligned} \tag{32}$$

or, in other words,

$$\mathbf{X} = \mathbf{V} \times \mathbf{W}. \tag{33}$$

The components of $\mathbf{X}$ form a vector operator, so by themselves they span an irreducible invariant subspace under rotations. As we see, the components of $\mathbf{X}$ contain the antisymmetric part of the original tensor T_{ij}.

The remaining 5-dimensional subspace is irreducible. It is spanned by operators containing the symmetric part of the tensor T_{ij}, with the trace removed (or, as we say, the symmetric, traceless

part of T_{ij}). The following five operators form a basis in this subspace:

$$\begin{aligned} S_1 &= T_{12} + T_{21}, \\ S_2 &= T_{23} + T_{32}, \\ S_3 &= T_{31} + T_{13}, \\ S_4 &= T_{11} - T_{22}, \\ S_5 &= T_{11} + T_{22} - 2T_{33}. \end{aligned} \tag{34}$$

The original tensor T_{ij} breaks up in three irreducible subspaces, a 1-dimensional scalar (the trace), a 3-dimensional vector (the antisymmetric part), and the 5-dimensional symmetric, traceless part. Notice that these dimensionalities are in accordance with the Clebsch-Gordan decomposition,

$$1 \otimes 1 = 0 \oplus 1 \oplus 2, \tag{35}$$

which corresponds to the count of dimensionalities,

$$3 \times 3 = 1 + 3 + 5 = 9. \tag{36}$$

This Clebsch-Gordan series arises because the vector operators $\mathbf{V}$ and $\mathbf{W}$ form two $\ell = 1$ irreducible subspaces of operators, and when we form T according to $T_{ij} = V_i W_j$, we are effectively combining angular momenta as indicated by Eq. (35). The only difference from our usual practice is that we are forming products of vector spaces of operators, instead of tensor products of ket spaces.

We have examined this decomposition in the special case $T_{ij} = V_i W_j$, but the decomposition itself applies to any second rank tensor T_{ij}. More generally, Cartesian tensors of any rank ≥ 2 are reducible.

It is possible that a given tensor T_{ij} may have one or more of the three irreducible components that vanish. The quadrupole moment tensor (29), for example, is already symmetric and traceless, so its nine components are actually linear combinations of just five linearly independent operators. For another example, an antisymmetric tensor satisfying $T_{ij} = -T_{ji}$ contains only the three-dimensional (vector) subspace.

The decomposition of a tensor operator T_{ij} into its irreducible components gives us a preview of how angular momentum quantum numbers are attached to operators. This topic will be developed in more detail in Notes 20.

9. Energy Eigenstates in Isolated Systems

In this section we explore the consequences of rotational invariance for the eigenstates, eigenvalues and degeneracies of a scalar operator. The most important scalar operator in practice is the Hamiltonian for an isolated system, so for concreteness we will speak of such a Hamiltonian, but the following analysis applies to any scalar operator.

Let H be the Hamiltonian for an isolated system, and let $\mathcal{E}$ be the Hilbert space upon which it acts. Since H is a scalar it commutes with $\mathbf{J}$, and therefore with the commuting operators J^2 and J_3. Let us denote the simultaneous eigenspaces of J^2 and J_3 with quantum numbers j and m by $\mathcal{S}_{jm}$, as illustrated in Fig. 13.5. It was shown in Notes 13 that for a given system, j takes on certain values that must be either integers or half-integers. For example, in central force motion we have only integer values of j (which is called ℓ in that context), while for the ^{57}Fe nucleus, which is discussed in more detail in the next section, we have only half-integer values. For each value of j that occurs there is a collection of $2j+1$ eigenspaces $\mathcal{S}_{jm}$ of J^2 and J_3, for $m = -j, \ldots, +j$. These spaces are mapped invertibly into one another by J_+ and J_-, as illustrated in Fig. 13.5, and if they are finite-dimensional, then they all have the same dimension. As in Notes 13, we write $N_j = \dim \mathcal{S}_{jj}$, which we call the multiplicity of the given j value.

In Notes 13 we constructed a standard angular momentum basis by picking an arbitrary orthonormal basis in each stretched space $\mathcal{S}_{jj}$, with the basis vectors labeled by γ as in Fig. 13.5, where $\gamma = 1, \ldots, N_j$. We denote these basis vectors in $\mathcal{S}_{jj}$ by $|\gamma jj\rangle$. Then by applying lowering operators, we construct an orthonormal basis in each of the other $\mathcal{S}_{jm}$, for m running down to $-j$. In this way we construct a standard angular momentum basis $|\gamma jm\rangle$ on the whole Hilbert space $\mathcal{E}$. In this construction, it does not matter how the basis $|\gamma jj\rangle$ is chosen in $\mathcal{S}_{jj}$, as long as it is orthonormal.

Now, however, we have a Hamiltonian, and we would like a simultaneous eigenbasis of H, J^2 and J_3. To construct this we restrict H to $\mathcal{S}_{jj}$ for some j (see Sec. 1.23 for the concept of the restriction of an operator to a subspace, and how it is used in proving that commuting operators possess a simultaneous eigenbasis). This restricted H is a Hermitian operator on $\mathcal{S}_{jj}$ so it possesses an eigenbasis on that space.

The spectrum of H on $\mathcal{S}_{jj}$ can be either discrete, continuous, or mixed (in most problems we will consider in this course it has a continuous spectrum above a threshold energy, and may have discrete bound states below that). Let us focus on the bound states and assume that H possesses at least one bound eigenstate $|\psi\rangle$ on $\mathcal{S}_{jj}$ with corresponding eigenvalue E. Then $|\psi\rangle$ satisfies

$$J^2|\psi\rangle = j(j+1)\hbar^2\,|\psi\rangle, \qquad J_3|\psi\rangle = j\hbar\,|\psi\rangle, \qquad H|\psi\rangle = E|\psi\rangle. \tag{37}$$

Now by applying a lowering operator J_- we find

$$J_-H|\psi\rangle = HJ_-|\psi\rangle = EJ_-|\psi\rangle, \tag{38}$$

so that $J_-|\psi\rangle$ is an eigenstate of H, lying in the space $\mathcal{S}_{j,j-1}$, with the same eigenvalue E as $|\psi\rangle \in \mathcal{S}_{jj}$. Continuing to apply lowering operators, we generate a set of $2j+1$ linearly independent eigenstates of H with the same eigenvalue E, that is, E is independent of the quantum number m. These states span an irreducible, invariant subspace of $\mathcal{E}$.

There may be other irreducible subspaces with the same energy. This can occur in two ways. It could happen that there is another energy eigenstate in $\mathcal{S}_{jj}$, linearly independent of $|\psi\rangle$, with the same energy E. That is, it is possible that E is a degenerate eigenvalue of H restricted to $\mathcal{S}_{jj}$. In

general, every discrete eigenvalue E of H restricted to $\mathcal{S}_{jj}$ corresponds to an eigenspace, a subspace of $\mathcal{S}_{jj}$ that may be multidimensional. Choosing an orthonormal basis in this subspace and applying lowering operators, we obtain a set of orthogonal, irreducible subspaces of the same value of j, each with $2j+1$ dimensions and all having the same energy.

It could also happen that there is another bound energy eigenstate, in a different space $\mathcal{S}_{j'j'}$ for $j' \neq j$, with the same energy E as $|\psi\rangle \in \mathcal{S}_{jj}$. This would be a degeneracy of H that crosses j values. If such a degeneracy exists, then we have at least two irreducible subspaces of the same energy, one of dimension $2j+1$ and the other of dimension $2j'+1$. In other words, degeneracies can occur either within a given j value or across j values.

These facts that we have accumulated can be summarized by a theorem:

Theorem 1. *The discrete energy eigenspaces of an isolated system consist of one or more invariant, irreducible subspaces under rotations, each associated with a definite j value. The different irreducible subspaces can be chosen to be orthogonal.*

Let us look at two examples of how this theorem works out in practice, the first a simple one with a small number of degrees of freedom that is exactly solvable, and the other a complicated one with a large number of degrees of freedom, in which all we know about the Hamiltonian is that it is invariant under rotations.

10. Example: Central Force Motion

For the simple example we take the case of central force motion, for which we use the notation $\mathbf{L}$, ℓ etc. instead of $\mathbf{J}$, j etc.

In central force motion the stretched subspace $\mathcal{S}_{\ell\ell}$ consists of wave functions $R(r)Y_{\ell\ell}(\theta,\phi)$, where $R(r)$ is any radial wave function. To find the energy eigenstates in this stretched subspace we solve the radial Schrödinger equation for the given ℓ value, which produces in general a continuous and a discrete spectrum. We assume there is a discrete spectrum for the given ℓ value and denote the energy eigenvalues and corresponding radial wave functions by $E_{n\ell}$ and $R_{n\ell}(r)$, as in Notes 16. By applying lowering operators to the wave function $R_{n\ell}(r)\,Y_{\ell\ell}(\theta,\phi)$, we obtain an irreducible subspace of degenerate energy eigenfunctions, spanned by $\{R_{n\ell}(r)\,Y_{\ell m}(\theta,\phi), m = +\ell,\ldots,-\ell\}$.

Now we consider degeneracies. Is it possible, for a given value of ℓ in a central force problem, that a bound energy eigenvalue can be degenerate? That is, can there be more than one linearly independent bound energy eigenstate of a given energy in $\mathcal{S}_{\ell\ell}$? As discussed in Sec. 16.4, the answer is no, the boundary conditions on the radial wave functions guarantee that there can be no degeneracy of this type. In central force problems, we do not have degeneracies within a given ℓ value.

Then is it possible that there is a degeneracy between different values of ℓ? Again, as discussed in Notes 16, the answer is that in general it is not very likely, since the different radial equations for different values of ℓ are effectively different Schrödinger equations whose centrifugal potentials are different.

The fact is that systematic degeneracies require a symmetry. We are already taking into account the $SO(3)$ symmetry of proper rotations, which explains the degeneracy in the magnetic quantum number m, so any additional degeneracy will require a larger symmetry group than $SO(3)$. In the absence of such extra symmetry, degeneracies between different ℓ values can occur only by "accident," that is, by fine tuning parameters in a Hamiltonian to force a degeneracy to happen. This is not likely in most practical situations. Therefore in central force problems we do not normally expect degeneracies that cross subspaces of different values of ℓ.

As explained in Notes 17, however, the electrostatic model of hydrogen is a notable exception, due to the symmetry group $SO(4)$ possessed by this model, which is larger than the rotation group $SO(3)$. The extra symmetry in this model of hydrogen explains why the energy levels $E_n = -1/2n^2$ (in the right units) are the same across the angular momentum values $\ell = 0, \ldots, n-1$. The isotropic harmonic oscillator in two or more dimensions is another example of a system with extra degeneracy; such oscillators are approximate models for certain types of molecular vibrations.

For a more complicated example of how Theorem 1 works out in practice we examine some energy levels of the nucleus ^{57}Fe, which is important in the Mössbauer effect. We use the opportunity to digress into some of the interesting physics connected with this effect. We begin with a general discussion of aspects of the emission and absorption of photons by quantum systems.

11. Emission and Absorption of Photons

When an atom, nucleus or other quantum system is in an excited state B and emits a photon while dropping into the ground state A,

$$B \to A + \gamma, \tag{39}$$

then a simple description of the process says that the energy of the photon is given by

$$E_\gamma = E_B - E_A, \tag{40}$$

where E_B and E_A are the energies of the states B and A. If now there is another atom, nucleus or other system of the same type nearby in its ground state A, then it would appear that that photon has exactly the right energy to induce the inverse reaction,

$$A + \gamma \to B, \tag{41}$$

thereby lifting the second system into the excited state B.

Does this mean, for example, that when an atom in a gas emits a photon that the photon only travels as far as the nearest neighboring atom before being absorbed again? No, because there are several effects that complicate the basic picture just presented, modifying the energy E_γ of the photon so that is it not exactly given by Eq. (40). These include the natural line width of the transition $B \to A$; Doppler shifts; recoil; and, in the case of nuclei, what are called chemical shifts.

In quantum mechanics the energy of a system is only precisely defined in a process that takes place over an infinite amount of time. The excited state B of our system is unstable with some lifetime τ, so its energy E_B is only defined to within an uncertainty of order $\Delta E_B = \hbar/\tau$. Assuming A is the ground state, it is stable and can exist over an infinite amount of time, so there is no uncertainty in its energy. Overall, the uncertainty in the energy E_B creates uncertainty of order $\hbar/\tau$ to the energy of the photon E_γ emitted in the process (39). This can be seen experimentally; if all other sources of broadening of the spectral line are eliminated, then the energy of photons emitted in an atomic or nuclear transition does not have a definite value, but rather there is a spread of order $\Delta E = \hbar/\tau$ about the nominal value $E_B - E_A$. This spread is called the *natural line width* of the spectral line. The natural line width of spectral lines is examined in some detail in Notes 44.

Similarly, if a photon of energy E_γ encounters a quantum system of the same type at rest in its ground state A, then if E_γ is roughly within the range $\Delta E = \hbar/\tau$ about the nominal energy $E_B - E_A$ it will be able to lift the second system into the excited state B, that is, the inverse reaction (41) will take place.

On the other hand, if the emitting atom, nucleus or other quantum system is in a state of motion, then the frequency $\omega_\gamma = E_\gamma/\hbar$ of the emitted photon will be Doppler shifted and may no longer be within the resonance needed to raise another such system into its excited state. Writing simply E for the nominal energy $E_B - E_A$ of the photon, the velocity v needed to shift the photon out of resonance is given by

$$\frac{v}{c} = \frac{\Delta E}{E} = \frac{\hbar}{E\tau}. \tag{42}$$

Whether or not the photon is shifted out of resonance depends on the velocity and other parameters, but in many practical circumstances one will find that thermal velocities do exceed the value given by Eq. (42). A similar logic applies in case the receiving system is in a state of motion (or both, as would be the case of a gas).

Even if the emitting atom or nucleus or other system is at rest, the energy E_γ is not given exactly by Eq. (40) because of the recoil of the emitting system when the photon is emitted. The photon has energy $E = \hbar\omega$ and momentum $\mathbf{p} = \hbar\mathbf{k}$ where $\omega = c|\mathbf{k}|$, so by conservation of momentum the emitting system suffers a recoil and has momentum $m\mathbf{v} = -\hbar\mathbf{k}$ after the photon of frequency ω is emitted, where m is the mass of the emitting system. Thus Eq. (40) should be replaced by

$$E_\gamma + \frac{1}{2}mv^2 = E_B - E_A, \tag{43}$$

where $\mathbf{v}$ is the recoil velocity. Some of the available energy goes into kinetic energy of the recoiling system, and the energy E_γ of the emitted photon is actually less than the nominal value (40). Again, whether this recoil shift is greater or less than the natural line width depends on the parameters of the problem.

12. The ^{57}Fe Nucleus and the Mössbauer Effect

The Mössbauer effect involves a transition $^{57}\text{Fe}^* \to {}^{57}\text{Fe} + \gamma$ between two energy levels of the ^{57}Fe nucleus, where the simple notation ^{57}Fe (or A) refers to the ground state and $^{57}\text{Fe}^*$ (or B) refers to an excited state. These states are illustrated in the energy level diagram for the nucleus given in Fig. 1. The photon emitted has energy 14.4 KeV, and the lifetime of the excited state $^{57}\text{Fe}^*$ is $\tau = 9.8 \times 10^{-8}$ sec. From these figures we find $\Delta\omega/\omega = \Delta E/E = 4.7 \times 10^{-13}$, where E and ω are the energy and frequency of the emitted photon. The spread in the energy is very small compared to the energy. For example, according to Eq. (42), to Doppler shift the photon out of resonance it would require a velocity of $v/c = 4.7 \times 10^{-13}$, or $v = 0.014\,\text{cm/sec}$.

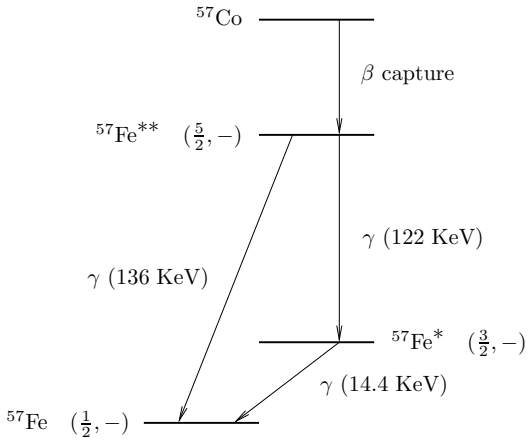

Fig. 1. Energy level diagram relevant for the Mössbauer effect in ^{57}Fe. ^{57}Fe is the ground state, $^{57}\text{Fe}^*$ is an excited state, and $^{57}\text{Fe}^{**}$ is a more highly excited state. Principal transitions via photon emission are shown.

The Mössbauer effect makes use of a source containing iron nuclei in the excited state $^{57}\text{Fe}^*$, which emits photons, and a receiver containing iron nuclei in the ground state which may absorb them by being lifted into the excited state $^{57}\text{Fe}^*$ via the reverse reaction. The receiver can be a block of natural iron, which contains the isotope ^{57}Fe at the 2% level, behind which a gamma-ray detector is placed. If the incident photons are within the narrow resonant range of energies, then they will be absorbed by the block of iron, and the detector will detect nothing. But if they are shifted out of resonance, the gamma rays will pass through the block of iron and the detector will detect them. If there is some effect that shifts the frequency of the gamma rays from their nominal energy (for example, the gravitational red shift in the Pound-Rebka experiment), then a compensating Doppler shift can be introduced by giving the source some velocity. By measuring the velocity of the source needed to shift the gamma rays back into resonance, one can measure the shift caused by the effect in question.

However, plugging in the numbers shows that the recoil shift described by Eq. (43), where m is the mass of an iron nucleus, is much greater than the natural line width, so the recoil would seem

to spoil the whole idea. But the iron nucleus is not free, rather it is part of a crystal lattice, whose vibrations are described by a large number of harmonic oscillators, the normal modes of the lattice. (See the discussion in Sec. 8.2.) Thus the recoil kinetic energy $E = (1/2)mv^2$ in Eq. (43) is not free to take on any value, rather it must be some multiple of $\hbar\omega_\ell$, where ω_ℓ is the frequency of a normal mode of the lattice. That is, when the photon is emitted by the nucleus, some number of phonons are also emitted into the lattice, representing the recoil energy. Note that ω_ℓ is the frequency of a sound wave in the lattice, which is much less than the frequency of the emitted gamma ray.

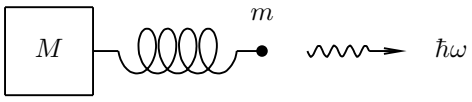

Fig. 2. A one dimensional model of an iron atom coupled to a normal mode of the lattice. The mass of the iron atom is m, the mass of the crystal is M. The iron atom emits a photon of energy $\hbar\omega$ and suffers a recoil (an impulse) as a result.

To model this situation in the simplest possible way, let us imagine an iron atom connected to a spring, forming a one-dimensional harmonic oscillator, as illustrated in Fig. 2. In the figure m is the mass of the iron atom while M is the mass of the crystal lattice to which it is coupled. When the atom emits a photon it suffers an an impulse, that is, a change Δp in its momentum. The emission takes place over a short time compared to the frequency of the harmonic oscillator (a normal mode of the lattice), so the position of the iron atom does not change much during the emission process. Classically we can model the impulse by the map,

$$x \mapsto x, \qquad p \mapsto p + \Delta p. \tag{44}$$

To model the effect of the recoil in quantum mechanics, we use the momentum displacement operator $S(b)$ introduced in Notes 8 [see Eqs. (8.64)–(8.66)]. That is, if $|\psi\rangle$ is the state of the oscillator before the photon is emitted, then the state after the emission is

$$|\psi'\rangle = e^{i\Delta p x/\hbar}|\psi\rangle. \tag{45}$$

The final state can be expanded in energy eigenstates,

$$|\psi'\rangle = \sum_n c_n|n\rangle, \tag{46}$$

so that the probability of finding the oscillator in state n after the photon has been emitted is $|c_n|^2$. In particular, if the initial state of the oscillator $|\psi\rangle = |n_i\rangle$ is an energy eigenstate, then the probability $|c_n|^2$ is the probability to make a transition $n_i \to n$ as a result of the recoil. As we say, $n - n_i$ phonons are emitted. See Prob. 8.4.

In fact, there is a certain probability that no phonons are emitted at all, that is $n = n_i$ and the oscillator remains in the initial state. Because of quantum mechanics, the recoil energy cannot take

on any value, but rather is quantized, and the value zero is allowed. What makes the ^{57}Fe nucleus attractive for the Mössbauer effect is that it has a reasonable probability for this *recoilless emission*.

Of course there is not only a recoil energy but also a recoil momentum. In the Mössbauer effect, recoilless emission does not violate conservation of momentum because the entire crystal lattice, with an effectively infinite mass M (as in the figure) takes up the recoil momentum.

The excited state ^{57}Fe* has only a short lifetime but in practice a population of these excited states is maintained in the source as a part of the decay chain of ^{57}Co, as shown in Fig. 1. ^{57}Co has a lifetime of 271 days, which is long enough to make it practical to use it as a source of ^{57}Fe* in a real experiment. As shown in the figure, ^{57}Co transforms into an excited state ^{57}Fe** by electron capture, after which ^{57}Fe** decays by the emission of a photon into ^{57}Fe*, which is the source of the photons of interest in the Mössbauer effect.

Mössbauer was awarded the Nobel Prize in 1961 for his discovery of recoilless emission of gamma ray photons and some of its applications. A notable early application was the Pound-Rebka experiment, carried out in 1959, in which the Mössbauer effect was used to make the first measurement of the gravitational red shift. This is the red shift photons experience when climbing in a gravitational field, in accordance with the 1911 prediction of Einstein. The gravitational red shift is one of the physical cornerstones of general relativity.

13. Energy levels in ^{57}Fe

To return to the subject of Hamiltonians and their energy levels in isolated systems, let us draw attention to the three levels ^{57}Fe, ^{57}Fe* and ^{57}Fe**, in Fig. 1. These are energy levels of the Hamiltonian for the ^{57}Fe nucleus, and, according to Theorem 1, each must consist of one or more irreducible subspaces under rotations. In fact, each of them consists of precisely one such irreducible subspace, with a definite j value, which is indicated in the figure ($\frac{1}{2}$ for the ground state ^{57}Fe, and $\frac{3}{2}$ and $\frac{5}{2}$ for the two excited states ^{57}Fe* and ^{57}Fe**, respectively). Also indicated are the parities of these states (all three have odd parity). The parity of energy eigenstates of isolated systems is discussed in Sec. 21.8.

A model for the Hamiltonian of the ^{57}Fe nucleus views it as a 57-particle system, that is, with 26 protons and 31 neutrons. The Hamiltonian is some function of the positions, momenta and spins of the particles,

$$H = H(\mathbf{x}_\alpha, \mathbf{p}_\alpha, \mathbf{S}_\alpha), \tag{47}$$

where $\alpha = 1, \ldots, 57$. In this model the total angular momentum of the system is the sum of the orbital and spin angular momenta of the nucleons,

$$\mathbf{J} = \sum_{\alpha=1}^{57} \mathbf{x}_\alpha \times \mathbf{p}_\alpha + \mathbf{S}_\alpha, \tag{48}$$

and the "spins" of the various nuclear states shown in Fig. 1 are actually the quantum numbers of J^2 (that is, we call $\mathbf{J}$ the "spin" and use the notation $\mathbf{S}$ etc. for it). For example, we say that the

ground state ^{57}Fe has spin $s = \frac{1}{2}$.

This model is more or less crude, due to the fact that protons and neutrons are composite particles, each made up of three quarks, which interact with the quark and gluon fields via the strong interactions. For our purposes the only thing that matters is that rotations act upon the state space of the system by means of unitary operators, and that these commute with the Hamiltonian. The model (47) at least gives us something concrete to think about.

Each nuclear energy eigenstate consists of a single irreducible subspace under rotations for the same reasons discussed in connection with central force motion in Sec. 9. That is, extra degeneracy requires extra symmetry or else an unlikely accident, and neither of these is to be expected in nuclei. Therefore each energy level is characterized by a unique angular momentum value, as indicated in the figure.

We can summarize these accumulated facts by stating an addendum to Theorem 1.

Addendum to Theorem 1. *With a few exceptions, notably the electrostatic model of hydrogen, the bound state energy eigenspaces of isolated systems consist of a single invariant, irreducible subspace under rotations. Thus, the energy eigenvalues are characterized by an angular momentum quantum number, which is variously denoted ℓ, s, j, etc, depending on the system.*

We can now understand why the Hilbert space for spins in magnetic fields consists of a single irreducible subspace under rotations for a large class of particles, a question that was raised in Notes 14. For example, if we place the ^{57}Fe nucleus in a magnetic field that is strong by laboratory standards, say, 10T, then the energy splitting between the two magnetic substates $m = \pm\frac{1}{2}$ will be of the order of 100 MHz in frequency units, or about 4×10^{-7}eV, or roughly 3×10^{-11} times smaller than the energy separation from the first excited state ^{57}Fe*. Therefore it is an excellent approximation to ignore the state ^{57}Fe* and all other excited states of the ^{57}Fe nucleus, and to treat the Hilbert space of the nucleus as if it were a single irreducible subspace with $s = \frac{1}{2}$, that is, the ground eigenspace. In other words, in the case of nuclei, the $2s + 1$-dimensional Hilbert space used in our study of spins in magnetic fields in Notes 14 is actually a subspace of a larger Hilbert space. It is, in fact, an energy eigenspace of an isolated system. This in turn explains why the magnetic moment is proportional to the spin [see Prob. 20.2(a)].

Problems

1. Some questions on the Mössbauer effect and the Pound-Rebka experiment.

The gravitational red shift is a prediction of general relativity, but the basic physics behind it can be understood in elementary terms. In the Pound-Rebka experiment, photons were launched upward from the ground, just outside a building at Harvard University. These photons were then received by a detector just outside the window of an upper floor of the building, approximately 20 meters above the ground.

Suppose at the very moment a photon is released at the ground, a physicist with a detector in hand jumps out of the upper storey window. As the photon is climbing upward in the earth's gravitational field, suffering a red shift in the process, the physicist is falling, accelerating downward. At a certain point the detector captures the photon and the physicist measures its frequency. This frequency is blue shifted compared to what it would be if the physicist had just held the detector out the window, instead of jumping, because of the Doppler shift due to the downward motion of the detector.

It turns out that the gravitational red shift and the blue shift due to the falling detector (cum physicist) exactly cancel. Thus, the physicist finds a measured frequency of the photon exactly equal to the frequency it had when emitted at ground level. You can use this fact to calculate the gravitational red shift as seen by a detector that is just held out the window, not falling.

This situation is similar to that illustrated in the "shoot the monkey" demonstration used in elementary physics classes, except that instead of an arrow shot at the monkey particles of light are used. The basic physical reasoning used here is close to that employed by Einstein in his 1911 paper which first predicted the gravitational red shift.

(a) If an atom is free (not a part of a crystal lattice or otherwise bound to anything else), then it suffers some recoil on emitting a photon, which produces a shift $\Delta\omega$ in the frequency of the emitted photon. In the case of the 14.4 KeV photon emitted by the ^{57}Fe nucleus, calculate the fractional shift $\Delta\omega/\omega$ due to this recoil and compare to the natural line width.

(b) If the ^{57}Fe atom is free-floating in a gas at 300 K, calculate the average $\Delta\omega/\omega$ due to the Doppler shift due to the thermal motion of the atom.

(c) Calculate the $\Delta\omega/\omega$ for the gravitational red shift of the same photon climbing (as in the Pound-Rebka experiment) about 20 meters in the earth's gravitational field, and compare to the $\Delta\omega/\omega$ due to the natural line width. You will see that the experiment was a delicate one that required careful measurement and attention to systematic errors.

2. Consider a vector operator $\mathbf{V}$ and a tensor operator T_{ij}.

(a) Show that if one component of $\mathbf{V}$ vanishes, then they all do.

(b) Is this true for a tensor operator T_{ij}?

Physics 221A
Academic Year 2020–21
Notes 20
The Wigner-Eckart Theorem†

1. Introduction

The Wigner-Eckart theorem concerns matrix elements of a type that is of frequent occurrence in all areas of quantum physics, especially in perturbation theory and in the theory of the emission and absorption of radiation. This theorem allows one to determine quickly the selection rules for the matrix element that follow from rotational invariance. In addition, if matrix elements must be calculated, the Wigner-Eckart theorem frequently offers a way of significantly reducing the computational effort. We will make quite a few applications of the Wigner-Eckart theorem in this course.

The Wigner-Eckart theorem is based on an analysis of how operators transform under rotations, a study of which was initiated in Notes 19. It turns out that operators of a certain type, the irreducible tensor operators, are associated with angular momentum quantum numbers and have transformation properties similar to those of kets with the same quantum numbers. An exploitation of these properties leads to the Wigner-Eckart theorem.

2. The Spherical Basis

We return to our development of the properties of operators under rotations. We take up the subject of the *spherical basis*, which is a basis of unit vectors in ordinary three-dimensional space that is alternative to the usual Cartesian basis. Initially we just present the definition of the spherical basis without motivation, and then we show how it can lead to some dramatic simplifications in certain problems. Then we explain its deeper significance.

We denote the usual Cartesian basis by $\hat{\mathbf{c}}_i$, $i = 1, 2, 3$, so that

$$\hat{\mathbf{c}}_1 = \hat{\mathbf{x}}, \qquad \hat{\mathbf{c}}_2 = \hat{\mathbf{y}}, \qquad \hat{\mathbf{c}}_3 = \hat{\mathbf{z}}. \tag{1}$$

We have previously denoted this basis by $\hat{\mathbf{e}}_i$, but in these notes we reserve the symbol $\hat{\mathbf{e}}$ for the spherical basis.

The spherical basis is defined by

$$\hat{\mathbf{e}}_1 = -\frac{\hat{\mathbf{x}} + i\hat{\mathbf{y}}}{\sqrt{2}},$$

† Links to the other sets of notes can be found at:
`http://bohr.physics.berkeley.edu/classes/221/2021/221.html`.

$$\hat{\mathbf{e}}_0 = \hat{\mathbf{z}},$$
$$\hat{\mathbf{e}}_{-1} = \frac{\hat{\mathbf{x}} - i\hat{\mathbf{y}}}{\sqrt{2}}. \tag{2}$$

This is a complex basis, so vectors with real components with respect to the Cartesian basis have complex components with respect to the spherical basis. We denote the spherical basis vectors collectively by $\hat{\mathbf{e}}_q$, $q = 1, 0, -1$.

The spherical basis vectors have the following properties. First, they are orthonormal, in the sense that

$$\hat{\mathbf{e}}_q^* \cdot \hat{\mathbf{e}}_{q'} = \delta_{qq'}. \tag{3}$$

Next, an arbitrary vector $\mathbf{X}$ can be expanded as a linear combination of the vectors $\hat{\mathbf{e}}_q^*$,

$$\mathbf{X} = \sum_q \hat{\mathbf{e}}_q^* X_q, \tag{4}$$

where the expansion coefficients are

$$X_q = \hat{\mathbf{e}}_q \cdot \mathbf{X}. \tag{5}$$

These equations are equivalent to a resolution of the identity in 3-dimensional space,

$$\mathsf{I} = \sum_q \hat{\mathbf{e}}_q^* \hat{\mathbf{e}}_q, \tag{6}$$

in which the juxtaposition of the two vectors is dyad notation for representing the tensor product.

You may wonder why we expand $\mathbf{X}$ as a linear combination of $\hat{\mathbf{e}}_q^*$, instead of $\hat{\mathbf{e}}_q$. The latter type of expansion is possible too, that is, any vector $\mathbf{Y}$ can be written

$$\mathbf{Y} = \sum_q \hat{\mathbf{e}}_q Y_q, \tag{7}$$

where

$$Y_q = \hat{\mathbf{e}}_q^* \cdot \mathbf{Y}. \tag{8}$$

These relations correspond to a different resolution of the identity,

$$\mathsf{I} = \sum_q \hat{\mathbf{e}}_q \hat{\mathbf{e}}_q^*. \tag{9}$$

The two types of expansion give the contravariant and covariant components of a vector with respect to the spherical basis; in this course, however, we will only need the expansion indicated by Eq. (4).

3. An Application of the Spherical Basis

To show some of the utility of the spherical basis, we consider the problem of dipole radiative transitions in a single-electron atom such as hydrogen or an alkali. It is shown in Notes 42 that the transition amplitude for the emission of a photon is proportional to matrix elements of the dipole

operator between the initial and final states. We use an electrostatic, spinless model for the atom, as in Notes 16, and we consider the transition from initial energy level $E_{n\ell}$ to final level $E_{n'\ell'}$. These levels are degenerate, since the energy does not depend on the magnetic quantum number m or m'. The wave functions have the form,

$$\psi_{n\ell m}(r,\theta,\phi) = R_{n\ell}(r)Y_{\ell m}(\Omega), \tag{10}$$

as in Eq. (16.15).

The dipole operator is proportional to the position operator of the electron, so we must evaluate matrix elements of the form,

$$\langle n\ell m|\mathbf{x}|n'\ell'm'\rangle, \tag{11}$$

where the initial state is on the left and the final one on the right. The position operator $\mathbf{x}$ has three components, and the initial and final levels consist of $2\ell+1$ and $2\ell'+1$ degenerate states, respectively. Therefore if we wish to evaluate the intensity of a spectral line as it would be observed, we really have to evaluate $3(2\ell'+1)(2\ell+1)$ matrix elements, for example, $3\times3\times5 = 45$ in a $3d \to 2p$ transition. This is because the energy of the photon is $E_\gamma = E_{n\ell} - E_{n'\ell'}$, which is independent of the initial and final m and m' quantum numbers. This count is actually an exaggeration, as we shall see, because many of the matrix elements vanish, but there are still many nonvanishing matrix elements to be calculated.

A great simplification can be achieved by expressing the components of $\mathbf{x}$, not with respect to the Cartesian basis, but with respect to the spherical basis. First we define

$$x_q = \hat{\mathbf{e}}_q \cdot \mathbf{x}, \tag{12}$$

exactly as in Eq.(5). Next, by inspecting a table of the $Y_{\ell m}$'s (see Sec. 15.7), we find that for $\ell = 1$ we have

$$\begin{aligned}
rY_{11}(\theta,\phi) &= -r\sqrt{\frac{3}{8\pi}}\,\sin\theta e^{i\phi} = \sqrt{\frac{3}{4\pi}}\Big(-\frac{x+iy}{\sqrt{2}}\Big),\\
rY_{10}(\theta,\phi) &= r\sqrt{\frac{3}{4\pi}}\,\cos\theta = \sqrt{\frac{3}{4\pi}}(z),\\
rY_{1,-1}(\theta,\phi) &= r\sqrt{\frac{3}{8\pi}}\,\sin\theta e^{-i\phi} = \sqrt{\frac{3}{4\pi}}\Big(\frac{x-iy}{\sqrt{2}}\Big),
\end{aligned} \tag{13}$$

where we have multiplied each Y_{1m} by the radius r. On the right hand side we see the spherical components x_q of the position vector $\mathbf{x}$, as follows from the definitions (2). The results can be summarized by

$$rY_{1q}(\theta,\phi) = \sqrt{\frac{3}{4\pi}}\,x_q, \tag{14}$$

for $q = 1, 0, -1$, where q appears explicitly as a magnetic quantum number. This equation reveals a relationship between vector operators and the angular momentum value $\ell = 1$, something we will have more to say about presently.

Now the matrix elements (11) become a product of a radial integral times an angular integral,

$$\begin{aligned}\langle n\ell m|x_q|n'\ell'm'\rangle = &\int_0^\infty r^2\,dr\,R^*_{n\ell}(r)rR_{n'\ell'}(r)\\ &\times\sqrt{\frac{4\pi}{3}}\int d\Omega\,Y^*_{\ell m}(\theta,\phi)Y_{1q}(\theta,\phi)Y_{\ell'm'}(\theta,\phi).\end{aligned}\tag{15}$$

We see that all the dependence on the three magnetic quantum numbers (m, q, m') is contained in the angular part of the integral. Moreover, the angular integral can be evaluated by the three-$Y_{\ell m}$ formula, Eq. (18.67), whereupon it becomes proportional to the Clebsch-Gordan coefficient,

$$\langle \ell m|\ell' 1 m' q\rangle.\tag{16}$$

The radial integral is independent of the three magnetic quantum numbers (m, q, m'), and the trick we have just used does not help us to evaluate it. But it is only one integral, and after it has been done, all the other integrals can be evaluated just by computing or looking up Clebsch-Gordan coefficients.

The selection rule $m = q + m'$ in the Clebsch-Gordan coefficient (16) means that many of the integrals vanish, so we have exaggerated the total number of integrals that need to be done. But had we worked with the Cartesian components x_i of $\mathbf{x}$, this selection rule might not have been obvious. In any case, even with the selection rule, there may still be many nonzero integrals to be done (nine, in the case $3d \to 2p$).

The example we have just given of simplifying the calculation of matrix elements for a dipole transition is really an application of the Wigner-Eckart theorem, which is the main topic of these notes.

The process we have just described is not just a computational trick, rather it has a physical interpretation. The initial and final states of the atom are eigenstates of L^2 and L_z, and the photon is a particle of spin 1 (see Notes 41). Conservation of angular momentum requires that the angular momentum of the initial state (the atom, with quantum numbers ℓ and m) should be the same as the angular momentum of the final state (the atom, with quantum numbers ℓ' and m', plus the photon with spin 1). Thus, the selection rule $m = m' + q$ means that q is the z-component of the spin of the emitted photon, so that the z-component of angular momentum is conserved in the emission process. As for the selection rule $\ell \in \{\ell' - 1, \ell', \ell' + 1\}$, it means that the amplitude is zero unless the possible total angular momentum quantum number of the final state, obtained by combining $\ell' \otimes 1$, is the total angular momentum quantum number of the initial state. This example shows the effect of symmetries and conservation laws on the selection rules for matrix elements.

This is only an incomplete accounting of the symmetry principles at work in the matrix element (11) or (15); as we will see in Notes 21, parity also plays an important role.

4. Significance of the Spherical Basis

To understand the deeper significance of the spherical basis we examine Table 1. The first row of this table summarizes the principal results obtained in Notes 13, in which we worked out the matrix representations of angular momentum and rotation operators. To review those results, we start with a ket space (first column) upon which proper rotations act by means of unitary operators $U(\mathsf{R})$ (second column). We refer only to proper rotations $\mathsf{R} \in SO(3)$, and we note that the representation may be double-valued.

The rotation operators have generators, defined by Eq. (12.13), that is, that equation can be taken as the definition of $\mathbf{J}$ when the rotation operators $U(\mathsf{R})$ are given. [Equation (12.11) is equivalent.] The generators appear in the third column of the table. The components of $\mathbf{J}$ satisfy the usual commutation relations (12.24) since the operators $U(\mathsf{R})$ form a representation of the rotation group.

Space	Action	Ang Mom	SAMB	Action on SAMB
Kets	$\lvert\psi\rangle \mapsto U\lvert\psi\rangle$	$\mathbf{J}$	$\lvert\gamma jm\rangle$	$U\lvert\gamma jm\rangle = \sum_{m'} \lvert\gamma jm'\rangle D^j_{m'm}$
3D Space	$\mathbf{x} \mapsto \mathsf{R}\mathbf{x}$	$i\mathbf{J}$	$\hat{\mathbf{e}}_q$	$\mathsf{R}\hat{\mathbf{e}}_q = \sum_{q'} \hat{\mathbf{e}}_{q'} D^1_{q'q}$
Operators	$A \mapsto UAU^\dagger$	$\ldots$	T^k_q	$UT^k_qU^\dagger = \sum_{q'} T^k_{q'} D^k_{q'q}$

Table 1. The rows of the table indicate different vector spaces upon which rotations act by means of unitary operators. The first row refers to a ket space (a Hilbert space of a quantum mechanical system), the second to ordinary three-dimensional space (physical space), and the third to the space of operators. The operators in the third row are the usual linear operators of quantum mechanics that act on the ket space, for example, the Hamiltonian. The first column identifies the vector space. The second column shows how rotations $\mathsf{R} \in SO(3)$ act on the given space. The third column shows the generators of the rotations, that is, the 3-vector of Hermitian operators that specify infinitesimal rotations. The fourth column shows the standard angular momentum basis (SAMB), and the last column, the transformation law of vectors of the standard angular momentum basis under rotations.

Next, since J^2 and J_z commute, we construct their simultaneous eigenbasis, with an extra index γ to resolve degeneracies. Also, we require states with different m but the same γ and j to be related by raising and lowering operators. This creates the standard angular momentum basis (SAMB), indicated in the fourth column.

In the last column, we show how the vectors of the standard angular momentum basis transform under rotations. This table entry is essentially Eq. (13.58). A basis vector $\lvert\gamma jm\rangle$, when rotated, produces a linear combination of other basis vectors for the same values of γ and j but different values of m. This implies that the space spanned by $\lvert\gamma jm\rangle$ for fixed γ and j, but for $m = -j, \ldots, +j$ is invariant under rotations. This space has dimensionality $2j + 1$. It is, in fact, an irreducible invariant space.

Note that the D-matrices that appear in the transformation law in the last column are universal matrices, dependent only on the angular momentum commutation relations and phase conventions, but independent of the nature of the system. See the discussion in Seec. 18.13.

Now we turn to the other rows of the table. At the beginning of Notes 13 we remarked that the analysis of those notes applies to other spaces besides ket spaces. All that is required is that we have a vector space upon which rotations act by means of unitary operators. For other vectors spaces the notation may change (we will not call the vectors kets, for example), but otherwise everything else goes through.

The second row of Table 1 summarizes the case in which the vector space is ordinary three-dimensional (physical) space. Rotations act on this space by means of the matrices R, which, being orthogonal, are also unitary (an orthogonal matrix is a special case of a unitary matrix). The action consists of just rotating vectors in the usual sense, as indicated in the second column.

The generators of rotations in this case must be a vector $\mathbf{J}$ of Hermitian operators, that is, Hermitian matrices, that satisfy

$$U(\hat{\mathbf{n}},\theta) = 1 - \frac{i}{\hbar}\theta\hat{\mathbf{n}}\cdot\mathbf{J}, \tag{17}$$

when θ is small. Here U really means the same thing as R, since we are speaking of the action on three-dimensional space, and 1 means the same as the identity matrix I. We will modify this definition of $\mathbf{J}$ slightly by writing $\mathbf{J}' = \mathbf{J}/\hbar$, thereby absorbing the $\hbar$ into the definition of $\mathbf{J}$ and making $\mathbf{J}'$ dimensionless. This is appropriate when dealing with ordinary physical space, since it has no necessary relation to quantum mechanics. (The spherical basis is also useful in classical mechanics, for example.) Then we will drop the prime, and just remember that in the case of this space, we will use dimensionless generators. Then we have

$$U(\hat{\mathbf{n}},\theta) = 1 - i\theta\hat{\mathbf{n}}\cdot\mathbf{J}. \tag{18}$$

But by a change of notation this is the same as

$$R(\hat{\mathbf{n}},\theta) = \mathsf{I} + \theta\hat{\mathbf{n}}\cdot\mathbf{J}, \tag{19}$$

as in Eq. (11.32), where the vector of matrices $\mathbf{J}$ is defined by Eq. (11.22). These imply

$$\mathbf{J} = i\mathbf{J}, \tag{20}$$

as indicated in the third column of Table 1. Writing out the matrices J_i explicitly, we have

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{21}$$

These matrices are indeed Hermitian, and they satisfy the dimensionless commutation relations,

$$[J_i, J_j] = i\epsilon_{ijk}\, J_k, \tag{22}$$

as follows from Eqs. (20) and (11.34).

We can now construct the standard angular momentum basis on three-dimensional space. In addition to Eq. (21), we need the matrices for J^2 and $J_\pm$. These are

$$J^2 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{23}$$

and

$$J_\pm = \begin{pmatrix} 0 & 0 & \mp 1 \\ 0 & 0 & -i \\ \pm 1 & i & 0 \end{pmatrix}. \tag{24}$$

We see that $J^2 = 2\mathsf{I}$, which means that every vector in ordinary space is an eigenvector of J^2 with eigenvalue $j(j+1) = 2$, that is, with $j = 1$. An irreducible subspace with $j = 1$ in any vector space must be 3-dimensional, but in this case the entire space is 3-dimensional, so the entire space consists of a single irreducible subspace under rotations with $j = 1$.

The fact that physical space carries the angular momentum value $j = 1$ is closely related to the fact that vector operators are irreducible tensor operators of order 1, as explained below. It is also connected with the fact that the photon, which is represented classically by the vector field $\mathbf{A}(\mathbf{x})$ (the vector potential), is a spin-1 particle.

Since every vector in three-dimensional space is an eigenvector of J^2, the standard basis consists of the eigenvectors of J_3, related by raising and lowering operators (this determines the phase conventions of the vectors, relative to that of the stretched vector). But we can easily check that the spherical unit vectors (2) are the eigenvectors of J_3, that is,

$$J_3 \hat{\mathbf{e}}_q = q\, \hat{\mathbf{e}}_q, \qquad q = 0, \pm 1. \tag{25}$$

Furthermore, it is easy to check that these vectors are related by raising and lowering operators, that is,

$$J_\pm \hat{\mathbf{e}}_q = \sqrt{(1 \mp q)(1 \pm q + 1)}\, \hat{\mathbf{e}}_{q\pm 1}, \tag{26}$$

where $J_\pm$ is given by Eq. (24). Only the overall phase of the spherical basis vectors is not determined by these relations. The overall phase chosen in the definitions (2) has the nice feature that $\hat{\mathbf{e}}_0 = \hat{\mathbf{z}}$.

Since the spherical basis is a standard angular momentum basis, its vectors must transform under rotations according to Eq. (13.85), apart from notation. Written in the notation appropriate for three-dimensional space, that transformation law becomes

$$\mathsf{R}\hat{\mathbf{e}}_q = \sum_{q'} \hat{\mathbf{e}}_{q'} D^1_{q'q}(\mathsf{R}). \tag{27}$$

We need not prove this as an independent result; it is just a special case of Eq. (13.85), in a different notation. This transformation law is also shown in the final column of Table 1, in order to emphasize its similarity to related transformation laws on other spaces.

Equation (27) has an interesting consequence, obtained by dotting both sides with $\hat{\mathbf{e}}^*_{q'}$. We use a round bracket notation for the dot product on the left hand side, and we use the orthogonality

relation (3) on the right hand side, which picks out one term from the sum. We find

$$\left(\hat{\mathbf{e}}^*_{q'}, \mathsf{R}\hat{\mathbf{e}}_q\right) = D^1_{q'q}(\mathsf{R}), \tag{28}$$

which shows that $D^1_{q'q}$ is just the matrix representing the rotation operator on three-dimensional space with respect to the spherical basis. The usual rotation matrix contains the matrix elements with respect to the Cartesian basis, that is,

$$\left(\hat{\mathbf{c}}_i, \mathsf{R}\hat{\mathbf{c}}_j\right) = R_{ij}. \tag{29}$$

See Eq. (11.7). For a given rotation, matrices R and $D^1(\mathsf{R})$ are similar (they differ only by a change of basis).

In the third row of the table the vector space upon which rotations act is the space of operators, that is, the usual linear operators in quantum mechanics that map the ket space into itself. The action of rotations on operators is indicated in the second column, which is the same as the definition of the rotated operator presented by Eq. (19.7). The third column, for the generators of rotations, is left blank. It is explored in Prob. 1. For the remaining two columns of the third row of Table 1, we turn the definition of a new class of operators.

5. Irreducible Tensor Operators

We define an *irreducible tensor operator* of order k as a set of $2k+1$ operators T^k_q, for $q = -k, \ldots, +k$, that satisfy

$$\boxed{U\, T^k_q\, U^\dagger = \sum_{q'} T^k_{q'}\, D^k_{q'q}(U),} \tag{30}$$

for all rotation operators U. We denote the irreducible tensor operator itself by T^k, and its $2k+1$ components by T^k_q. This definition is really a version of Eq. (13.85), applied to the space of operators. It means that the components of an irreducible tensor operator are basis operators in a standard angular momentum basis that spans an irreducible subspace of operators. Thus we place T^k_q in the SAMB column of the third row of Table 1, and the transformation law (30) in the last column. The three transformation laws in the last column (for three different kinds of spaces) should be compared. We see that the order k of an irreducible tensor operator behaves like an angular momentum quantum number j, and q behaves like m.

The index k of an irreducible tensor operator of a physical observable is restricted to integer values, $k = 0, 1, \ldots$. This is unlike the case of the standard angular momentum basis states in a ket space, in which the index j can take on either integer or half-integer values. The physical reason for this is that operators that represent physically observable quantities must be invariant under a rotation of 2π. The mathematical reason is that our definition of a rotated operator, given by Eq. (19.6), is quadratic $U(\mathsf{R})$, so that the representation of rotations on the vector space of operators is always a single-valued representation of $SO(3)$.

Let us look at some examples of irreducible tensor operators. A scalar operator K is an irreducible tensor operator of order 0, that is, it is an example of an irreducible tensor operator T^0_0. This follows easily from the fact that K commutes with any rotation operator U, and from the fact that the $j = 0$ rotation matrices are simply given by the 1×1 matrix (1) [see Eq. (13.68)].

Irreducible tensor operators of order 1 are constructed from vector operators by transforming from the Cartesian basis to the spherical basis. If we let $\mathbf{V}$ be a vector operator as defined by Eq. (19.14), and define its spherical components by

$$V_q = T^1_q = \hat{\mathbf{e}}_q \cdot \mathbf{V}, \tag{31}$$

then we have

$$\begin{aligned} U(\mathsf{R})V_q U(\mathsf{R})^\dagger &= \hat{\mathbf{e}}_q \cdot (\mathsf{R}^{-1}\mathbf{V}) = (\mathsf{R}\hat{\mathbf{e}}_q) \cdot \mathbf{V} \\ &= \sum_{q'} V_{q'} D^1_{q'q}(\mathsf{R}), \end{aligned} \tag{32}$$

where we use Eq. (27) and (11.72).

The electric quadrupole operator is given as a Cartesian tensor in Eq. (19.29). This Cartesian tensor is symmetric and traceless, so it contains only 5 independent components, which span an irreducible subspace of operators. In fact, this subspace is associated with angular momentum value $k = 2$. It is possible to introduce a set of operators T^2_q, $q = -2, \ldots, +2$ that form a standard angular momentum basis in this space, that is, that form an order 2 irreducible tensor operator. These can be regarded as the spherical components of the quadrupole moment tensor. This subject is explored in more detail in Prob. 2.

6. Commutation Relations of an Irreducible Tensor Operator with J

Above we presented two equivalent definitions of scalar and vector operators, one involving transformation properties under rotations, and the other involving commutation relations with $\mathbf{J}$. We will now do the same with irreducible tensor operators. To this end, we substitute the infinitesimal form (19.16) of the rotation operator U into both sides of the definition (30).

On the right we will need the D-matrix for an infinitesimal rotation. Since the D-matrix contains just the matrix elements of U with respect to a standard angular momentum basis [this is the definition of the D-matrices, see Eq. (13.56)], we require these matrix elements in the case of an infinitesimal rotation. For $\theta \ll 1$, Eq. (13.56) becomes

$$D^j_{m'm}(\hat{\mathbf{n}}, \theta) = \langle jm'| \Big(1 - \frac{i}{\hbar}\theta \hat{\mathbf{n}} \cdot \mathbf{J}\Big) |jm\rangle = \delta_{m'm} - \frac{i}{\hbar}\theta \langle jm'|\hat{\mathbf{n}} \cdot \mathbf{J}|jm\rangle. \tag{33}$$

Changing notation $(jm'm) \to (kq'q)$ and substituting this and Eq. (19.16) into the definition (30) of an irreducible tensor operator, we obtain

$$\Big(1 - \frac{i}{\hbar}\theta \hat{\mathbf{n}} \cdot \mathbf{J}\Big) T^k_q \Big(1 + \frac{i}{\hbar}\theta \hat{\mathbf{n}} \cdot \mathbf{J}\Big) = \sum_{q'} T^k_{q'} \Big(\delta_{q'q} - \frac{i}{\hbar}\theta \langle kq'|\hat{\mathbf{n}} \cdot \mathbf{J}|kq\rangle\Big), \tag{34}$$

or, since $\hat{\mathbf{n}}$ arbitrary unit vector,

$$[\mathbf{J}, T^k_q] = \sum_{q'} T^k_{q'} \langle kq'|\mathbf{J}|kq\rangle. \tag{35}$$

The operators $\mathbf{J}$ on the left- and right-hand sides of Eqs. (34) and (35) are not the same operators. On the left $\mathbf{J}$ is the angular momentum on the same space upon which the operators T^k_q act; in practice this is usually the state space of a quantum system. The $\mathbf{J}$ on the right is the angular momentum operator on a model space in which the matrices $D^k_{q'q}$ are defined. See the discussion in Sec. 18.13.

Equation (35) specifies a complete set of commutation relations of the components of $\mathbf{J}$ with the components of an irreducible tensor operator, but it is usually transformed into a different form. First we take the z-component of both sides and use $J_z|kq\rangle = \hbar q|kq\rangle$, so that

$$\langle kq'|J_z|kq\rangle = q\hbar\,\delta_{q'q}. \tag{36}$$

This is Eq. (13.47) with a change of notation. Then Eq. (35) becomes Eq. (43a) below. Next dot both sides of Eq. (35) with $\hat{\mathbf{x}} \pm i\hat{\mathbf{y}}$, and use

$$J_\pm|kq\rangle = \sqrt{(k \mp q)(k \pm q + 1)}\hbar\,|k, q \pm 1\rangle, \tag{37}$$

or

$$\langle kq'|J_\pm|kq\rangle = \sqrt{(k \mp q)(k \pm q + 1)}\hbar\,\delta_{q',q\pm1}. \tag{38}$$

This is Eq. (13.48b) with a change of notation. Then we obtain Eq. (43b) below. Finally, take the i-th component of Eq. (35),

$$[J_i, T^k_q] = \sum_{q'} T^k_{q'} \langle kq'|J_i|kq\rangle, \tag{39}$$

and form the commutator of both sides with J_i,

$$\begin{aligned}[J_i, [J_i, T^k_q]] &= \sum_{q'} [J_i, T^k_{q'}]\langle kq'|J_i|kq\rangle = \sum_{q'q''} T^k_{q''}\langle kq''|J_i|kq'\rangle\langle kq'|J_i|kq\rangle \\ &= \sum_{q''} T^k_{q''}\langle kq''|J_i^2|kq\rangle,\end{aligned} \tag{40}$$

where we have used Eq. (35) again to create a double sum. Finally summing both sides over i, we obtain,

$$\sum_i [J_i, [J_i, T^k_q]] = \sum_{q''} T^k_{q''}\langle kq''|J^2|kq\rangle. \tag{41}$$

But

$$\langle kq''|J^2|kq\rangle = k(k+1)\hbar^2\,\delta_{q''q}, \tag{42}$$

a version of Eq. (13.46), so we obtain Eq. (43c) below.

In summary, an irreducible tensor operator satisfies the following commutation relations with the components of angular momentum:

$$[J_z, T^k_q] = \hbar q\, T^k_q, \tag{43a}$$

$$[J_\pm, T^k_q] = \hbar\sqrt{(k \mp q)(k \pm q + 1)}\, T^k_{q\pm 1}, \tag{43b}$$

$$\sum_i [J_i, [J_i, T^k_q]] = \hbar^2 k(k+1)\, T^k_q. \tag{43c}$$

We see that forming the commutator with $J_\pm$ plays the role of a raising or lowering operator for the components of an irreducible tensor operator. As we did with scalar and vector operators, we can show that these angular momentum commutation relations are equivalent to the definition (30) of an irreducible tensor operator. This is done by showing that Eqs. (43) are equivalent to Eq. (30) in the case of infinitesimal rotations, and that if Eq. (30) is true for any two rotations, it is also true for their product. Thus by building up finite rotations as products of infinitesimal ones we show the equivalence of Eqs. (30) and (43). Many books take Eqs. (43) as the definition of an irreducible tensor operator.

7. Statement and Applications of the Wigner-Eckart Theorem

The Wigner-Eckart theorem is not difficult to remember and it is quite easy to use. In this section we discuss the statement of the theorem and ways of thinking about it and its applications, before turning to its proof.

The Wigner-Eckart theorem concerns matrix elements of an irreducible tensor operator with respect to a standard angular momentum basis of kets, something we will write in a general notation as $\langle \gamma' j' m' | T^k_q | \gamma j m \rangle$. As an example of such a matrix element, you may think of the dipole matrix elements $\langle n'\ell' m' | x_q | n\ell m \rangle$ that we examined in Sec. 3. In that case the operator (the position or dipole operator) is an irreducible tensor operator with $k = 1$.

The matrix element $\langle \gamma' j' m' | T^k_q | \gamma j m \rangle$ depends on 8 indices, $(\gamma' j' m'; \gamma j m; k q)$, and in addition it depends on the specific operator T in question. The Wigner-Eckart theorem concerns the dependence of this matrix element on the three magnetic quantum numbers $(m' m q)$, and states that that dependence is captured by a Clebsch-Gordan coefficient. More specifically, the Wigner-Eckart theorem states that $\langle \gamma' j' m' | T^k_q | \gamma j m \rangle$ is proportional to the Clebsch-Gordan coefficient $\langle j' m' | j k m q \rangle$, with a proportionality factor that is independent of the magnetic quantum numbers. That proportionality factor depends in general on everything else besides the magnetic quantum numbers, that is, $(\gamma' j'; \gamma j; k)$ and the operator in question. The standard notation for the proportionality factor is $\langle \gamma' j' \| T^k \| \gamma j \rangle$, something that looks like the original matrix element except the magnetic quantum numbers are omitted and a double bar is used. The quantity $\langle \gamma' j' \| T^k \| \gamma j \rangle$ is called the *reduced matrix element*. With this notation, the Wigner-Eckart theorem states

$$\boxed{\langle \gamma' j' m' | T^k_q | \gamma j m \rangle = \langle \gamma' j' \| T^k \| \gamma j \rangle\, \langle j' m' | j k m q \rangle.} \tag{44}$$

The reduced matrix element can be thought of as depending on the irreducible tensor operator T^k and the two irreducible subspaces $(\gamma' j')$ and (γj) that it links. Some authors (for example, Sakurai) include a factor of $1/\sqrt{2j+1}$ on the right hand side of Eq. (44), but here that factor has been absorbed into the definition of the reduced matrix element. The version (44) is easier to remember and closer to the basic idea of the theorem.

To remember the Clebsch-Gordan coefficient it helps to suppress the bra $\langle \gamma' j' m'|$ from the matrix element and think of the ket $T^k_q|\gamma j m\rangle$, or, more precisely, the $(2j+1)(2k+1)$ kets that are produced by letting m and q vary over their respective ranges. This gives an example of an operator with certain angular momentum indices multiplying a ket with certain angular momentum indices. It turns out that such a product of an operator times a ket has much in common with the product (i.e., the tensor product) of two kets, insofar as the transformation properties of the product under rotations are concerned. That is, suppose we were multiplying a ket $|kq\rangle$ with the given angular momentum quantum numbers times another ket $|jm\rangle$ with different angular momentum quantum numbers. Then we could find the eigenstates of total angular momentum by combining the constituent angular momenta according to $k \otimes j$. Actually, in thinking of kets $T^k_q|jm\rangle$, it is customary to think of the product of the angular momenta in the reverse order, that is, $j \otimes k$. This is an irritating convention because it makes the Wigner-Eckart theorem harder to remember, but I suspect it is done this way because in practice k tends to be small and j large.

In any case, thinking of the product of kets, the product

$$|jm\rangle \otimes |kq\rangle = |jkmq\rangle \tag{45}$$

contains various components of total J^2 and J_z, that is, it can be expanded as a linear combination of eigenstates of total J^2 and J^z, with expansion coefficients that are the Clebsch-Gordan coefficients. The coefficient with total angular momentum j' and z-component m' is the Clebsch-Gordan coefficient $\langle j'm'|jkmq\rangle$, precisely what appears in the Wigner-Eckart theorem (44).

Probably the most useful application of the Wigner-Eckart theorem is that it allows us to easily write down selection rules for the given matrix element, based on the selection rules of the Clebsch-Gordan coefficient occurring in Eq. (44). In general, a *selection rule* is a rule that tells us when a matrix element must vanish on account of some symmetry consideration. The Wigner-Eckart theorem provides us with all the selection rules that follow from rotational symmetry; a given matrix element may have other selection rules based on other symmetries (for example, parity). The selection rules that follow from the Wigner-Eckart theorem are that the matrix element $\langle \gamma j' m'|T^k_q|\gamma j m\rangle$ vanishes unless $m' = m + q$ and j' takes on one of the values, $|j-k|, |j-k|+1, \ldots, j+k$.

Furthermore, suppose we actually have to evaluate the matrix elements $\langle \gamma' j' m'|T^k_q|\gamma j m\rangle$ for all $(2k+1)(2j+1)$ possibilities we get by varying q and m. We must do this, for example, in computing atomic transition rates. (We need not vary m' independently, since the selection rules enforce $m' = m+q$.) Then the Wigner-Eckart theorem tells us that we actually only have to do one of these matrix elements (presumably, whichever is the easiest), because if we know the left hand side of Eq. (44) for one set of magnetic quantum numbers, and if we know the Clebsch-Gordan coefficient on

the right-hand side, then we can determine the proportionality factor, that is, the reduced matrix element. Then all the other matrix elements for other values of the magnetic quantum numbers follow by computing (or looking up) Clebsch-Gordan coefficients. This procedure requires that the first matrix element we calculate be nonzero.

In some other cases, we have analytic formulas for the reduced matrix element. That was the case of the application in Sec. 3, where the three-$Y_{\ell m}$ formula allowed us to compute the proportionality factor explicitly.

8. The Wigner-Eckart Theorem for Scalar Operators

Let us consider a scalar operator for which $k = q = 0$, such as the Hamiltonian H for an isolated system, that is, with $T^0_0 = H$. In this case the Clebsch-Gordan coefficient is

$$\langle j'm'|j0m0\rangle = \delta_{j'j}\,\delta_{m'm}, \tag{46}$$

so the Wigner-Eckart theorem can be written

$$\langle \gamma' j'm'|H|\gamma jm\rangle = C^j_{\gamma'\gamma}\,\delta_{j'j}\,\delta_{m'm}, \tag{47}$$

where

$$C^j_{\gamma'\gamma} = \langle \gamma' j\|H\|\gamma j\rangle. \tag{48}$$

The matrix elements of the Hamiltonian of an isolated system in a standard angular momentum basis are diagonal in both j and m, and moreover are independent of m.

If we wish to find the eigenvalues of the Hamiltonian we can diagonalize the matrices $C^j_{\gamma'\gamma}$, where j labels the matrix and $\gamma, \gamma' = 1, \ldots, N_j$, where N_j is the multiplicity of the given j value.

9. Recursion Relations for Matrix Elements

Many books rationalize the Wigner-Eckart theorem by showing that matrix elements of the form $\langle \gamma' j'm'|T^k_q|\gamma jm\rangle$ satisfy the same recursion relations as the Clebsch-Gordan coefficients $\langle j'm'|jkmq\rangle$, and by arguing that that fact implies that the two are proportional. The conclusion is correct, but the ususual presentations fall far short of actually proving it. Nevertheless, it is worthwhile outlining the usual argument.

We obtain a selection rule for the matrix element $\langle \gamma' j'm'|T^k_q|\gamma jm\rangle$ by sandwiching $\langle \gamma' j'm'|$ and $|\gamma jm\rangle$ around the commutation relation (43a). This gives

$$(m' - m - q)\langle \gamma' j'm'|T^k_q|\gamma jm\rangle = 0, \tag{49}$$

so either $m' = m + q$ or $\langle \gamma' j'm'|T^k_q|\gamma jm\rangle = 0$.

Similarly, we obtain recursion relations by sandwiching $\langle \gamma' j'm'|$ and $|\gamma jm\rangle$ around the commutation relations (43b). Noting that

$$\langle \gamma' j'm'|J_+ = [J_-|\gamma' j'm'\rangle]^\dagger = \sqrt{(j' + m')(j' - m' + 1)}\langle \gamma' j'm' - 1|, \tag{50}$$

we obtain

$$\begin{aligned}\sqrt{(j'+m')(j'-m'+1)}\,\langle\gamma' j', m'-1|T^k_q|\gamma jm\rangle \\ = \sqrt{(j-m)(j+m+1)}\,\langle\gamma' j'm'|T^k_q|\gamma j, m+1\rangle \\ + \sqrt{(k-q)(k+q+1)}\,\langle\gamma' j'm'|T^k_{q+1}|\gamma jm\rangle,\end{aligned} \tag{51}$$

which may be compared to the recursion relation (18.53) for the Clebsch-Gordan coefficients. Similarly, we find

$$\begin{aligned}\sqrt{(j'-m')(j'+m'+1)}\,\langle\gamma' j', m'+1|T^k_q|\gamma jm\rangle \\ = \sqrt{(j+m)(j-m+1)}\,\langle\gamma' j'm'|T^k_q|\gamma j, m-1\rangle \\ + \sqrt{(k+q)(k-q+1)}\,\langle\gamma' j'm'|T^k_{q-1}|\gamma jm\rangle,\end{aligned} \tag{52}$$

which may be compared to Eq. (18.54).

The rest of the proof is based on an argument from linear algebra that proceeds as follows. Let M be an $n \times n$ matrix and x an n-dimensional vector. The equation $\mathsf{M}x = 0$ has a set of solutions $\{x\}$ that constitute a vector space, called the *kernel* of M. If $\det \mathsf{M} \neq 0$, the kernel is the trivial space $\{0\}$, but if $\det \mathsf{M} = 0$ then the dimension of the kernel is ≥ 1.

If the kernel is 1-dimensional, then any solution x is proportional to any nonzero solution x_0, that is, $x = cx_0$. In the application to the proof of the Wigner-Eckart theorem, the matrix M is the set of coefficients of the recursion relations, the unknown x is the collection of matrix elements $\langle\gamma' j'm'|T^k_q|\gamma jm\rangle$, x_0 is the set of Clebsch-Gordan coefficients $\langle j'm'|jkmq\rangle$, and c is the reduced matrix element.

Since filling in the details of this approach is somewhat tedious, we present an alternative proof of the Wigner-Eckart theorem, one based on the observation that if a set of vectors indexed by jm values transforms under rotations as a standard angular momentum basis, then it is a standard angular momentum basis.

10. Proof of the Wigner-Eckart Theorem

Consider the product of kets $|jm\rangle \otimes |kq\rangle = |jkmq\rangle$ with the given angular momentum quantum numbers, and consider the $(2j+1)(2k+1)$-dimensional product space spanned by such kets when we allow the magnetic quantum numbers m and q to vary over their respective ranges. The eigenstates $|JM\rangle$ of total J^2 and J_z in this space are given by the Clebsch-Gordan expansion,

$$|JM\rangle = \sum_{mq} |jkmq\rangle\langle jkmq|JM\rangle. \tag{53}$$

Moreover, the states $|JM\rangle$ for fixed J and $M = -J, \ldots, +J$ form a standard angular momentum basis in an invariant, irreducible subspace of dimension $2J+1$ in the product space. This means that the basis states $|JM\rangle$ are not only eigenstates of total J^2 and J_z, but they are also linked by raising and lowering operators. Equivalently, the states $|JM\rangle$ transform as a standard angular momentum

basis under rotations,

$$U|JM\rangle = \sum_{M'} |JM'\rangle D^J_{M'M}(U). \tag{54}$$

Now consider the $(2j+1)(2k+1)$ kets $T^k_q|\gamma jm\rangle$ obtained by varying m and q. We construct linear combinations of these with the same Clebsch-Gordan coefficients as in Eq. (53),

$$|X;JM\rangle = \sum_{mq} T^k_q|\gamma jm\rangle\langle jkmq|JM\rangle, \tag{55}$$

and define the result to be the ket $|X;JM\rangle$, as indicated. The indices JM in the ket $|X;JM\rangle$ indicate that the left-hand side depends on these indices, because the right hand side does; initially we assume nothing else about this notation. Similarly, X simply stands for everything else the left-hand side depends on, that is, X is an abbreviation for the indices (γkj).

However, in view of the similarity between Eqs. (53) and (55), we can guess that $|X;JM\rangle$ is actually an eigenstate of J^2 and J_z with quantum numbers J and M, and that the states $|X;JM\rangle$ are related by raising and lowering operators. That is, we guess

$$J_z|X;JM\rangle = M\hbar\,|X;JM\rangle, \tag{56a}$$

$$J_\pm|X;JM\rangle = \sqrt{(J\mp M)(J\pm M+1)}\hbar\,|X;J,M\pm1\rangle, \tag{56b}$$

$$J^2|X;JM\rangle = J(J+1)\hbar^2\,|X;JM\rangle. \tag{56c}$$

If true, this is equivalent to the transformation law,

$$U|X;JM\rangle = \sum_{M'} |X;JM'\rangle\, D^J_{M'M}(U), \tag{57}$$

exactly as in Eq. (54). Equations (56) and (57) are equivalent because Eq. (56) can be obtained from Eq. (57) by specializing to infinitesimal rotations, while Eq. (57) can be obtained from Eq. (56) by building up finite rotations out of infinitesimal ones.

In Sec. 11 below we will prove that these guesses are correct. For now we merely explore the consequences. To begin, since $|X;JM\rangle$ is an eigenstate of J^2 and J_z with quantum numbers J and M, it can be expanded as a linear combination of the standard basis kets $|\gamma jm\rangle$ with the same values $j=J$ and $m=M$, but in general all possible values of γ. That is, we have an expansion of the form,

$$|X;JM\rangle = \sum_{\gamma'} |\gamma' JM\rangle\, C^{kJMj}_{\gamma'\gamma}, \tag{58}$$

where the indices on the expansion coefficients $C^{kJMj}_{\gamma'\gamma}$ simply list all the parameters they can depend on. These coefficients, do not, however, depend on M, as we show by applying raising or lowering operators to both sides, and using Eq. (56b). This gives

$$\begin{aligned}&\sqrt{(J\mp M)(J\pm M+1)}\hbar\,|X;J,M\pm1\rangle \\ &\qquad = \sum_{\gamma'} \sqrt{(J\mp M)(J\pm M+1)}\hbar\,|\gamma' J,M\pm1\rangle\, C^{kJMj}_{\gamma'\gamma},\end{aligned} \tag{59}$$

or, after canceling the square roots,

$$|X;J,M\pm1\rangle=\sum_{\gamma'}|\gamma'J,M\pm1\rangle\,C^{kJMj}_{\gamma'\gamma}. \tag{60}$$

Comparing this to Eq. (58), we see that the expansion coefficients are the same for all M values, and thus independent of M. We will henceforth write simply $C^{kJj}_{\gamma'\gamma}$ for them.

Now we return to the definition (55) of the kets $|X;JM\rangle$ and use the orthogonality of the Clebsch-Gordan coefficients (18.50) to solve for the kets $T^k_q|\gamma jm\rangle$. This gives

$$T^k_q|\gamma jm\rangle=\sum_{JM}|X;JM\rangle\langle JM|jkmq\rangle=\sum_{\gamma''JM}|\gamma''JM\rangle\,C^{kJj}_{\gamma''\gamma}\,\langle JM|jkmq\rangle, \tag{61}$$

where we use Eq. (58), replacing γ' with γ''. Now multiplying this by $\langle\gamma'j'm'|$ and using the orthonormality of the basis $|\gamma jm\rangle$, we obtain

$$\langle\gamma'j'm'|T^k_q|\gamma jm\rangle=C^{kj'j}_{\gamma'\gamma}\,\langle j'm'|jkmq\rangle, \tag{62}$$

which is the Wigner-Eckart theorem (44) if we identify

$$C^{kj'j}_{\gamma'\gamma}=\langle\gamma'j'\|T^k\|\gamma j\rangle. \tag{63}$$

11. Proof of Eq. (57)

To complete the proof of the Wigner-Eckart theorem we must prove Eq. (57), that is, we must show that the kets $|X;JM\rangle$ transform under rotations like the vectors of a standard angular momentum basis. To do this we call on the definition of $|X;JM\rangle$, Eq. (55), and apply U to both sides,

$$U|X;JM\rangle=\sum_{mq}UT^k_qU^\dagger\,U|\gamma jm\rangle\langle jkmq|JM\rangle. \tag{64}$$

Next we use the definition of an irreducible tensor operator (30) and the transformation law for standard basis vectors under rotations, Eq. (13.85), to obtain

$$U|X;JM\rangle=\sum_{\substack{mq\\m'q'}}T^k_{q'}|\gamma jm'\rangle\,D^j_{m'm}(U)\,D^k_{q'q}(U)\,\langle jkmq|JM\rangle. \tag{65}$$

We now call on Eq. (18.64) with a change of indices,

$$D^j_{m'm}(U)D^k_{q'q}(U)=\sum_{J'M'M''}\langle jkm'q'|J'M'\rangle\,D^{J'}_{M'M''}(U)\,\langle J'M''|jkmq\rangle, \tag{66}$$

which expresses the product of D-matrices in Eq. (65) in terms of single D-matrices. When we substitute Eq. (66) into Eq. (65), the $m'q'$-sum is doable by the definition (55),

$$\sum_{m'q'}T^k_{q'}|\gamma jm'\rangle\langle jkm'q'|J'M'\rangle=|X;J'M'\rangle, \tag{67}$$

and the mq-sum is doable by the orthogonality of the Clebsch-Gordan coefficients,

$$\sum_{mq} \langle J'M''|jkmq\rangle\langle jkmq|JM\rangle = \delta_{J'J}\,\delta_{M''M}. \tag{68}$$

Altogether, Eq. (65) becomes

$$U|X;JM\rangle = \sum_{J'M'M''} |X;J'M'\rangle D^{J'}_{M'M''}(U)\,\delta_{J'J}\,\delta_{M''M} = \sum_{M'} |X;JM'\rangle D^{J}_{M'M}(U). \tag{69}$$

This proves Eq. (57).

12. Products of Irreducible Tensor Operators

As we have seen, the idea behind the Wigner-Eckart theorem is that a product of an irreducible tensor operator T^k_q times a ket of the standard basis $|\gamma jm\rangle$ transforms under rotations exactly as the tensor product of two kets of standard bases with the same quantum numbers, $|jm\rangle\otimes|kq\rangle$. Similarly, it turns out that the product of two irreducible tensor operators, say, $X^{k_1}_{q_1}Y^{k_2}_{q_2}$, transforms under rotations exactly like the tensor product of kets with the same quantum numbers, $|k_1q_1\rangle\otimes|k_2q_2\rangle$. In particular, such a product of operators can be represented as a linear combination of irreducible tensor operators with order k lying in the range $|k_1-k_2|,\ldots,k_1+k_2$, with coefficients that are Clebsch-Gordan coefficients. That is, we can write

$$X^{k_1}_{q_1}Y^{k_2}_{q_2} = \sum_{kq} T^k_q\,\langle kq|k_1k_2q_1q_2\rangle, \tag{70}$$

where the T^k_q are new irreducible tensor operators.

To prove this, we first solve for T^k_q,

$$T^k_q = \sum_{q_1q_2} X^{k_1}_{q_1}Y^{k_2}_{q_2}\,\langle k_1k_2q_1q_2|kq\rangle, \tag{71}$$

which we must show is an irreducible tensor operator. To do this, we conjugate both sides of this with a rotation operator U and use the fact that X and Y are irreducible tensor operators,

$$\begin{aligned} UT^k_qU^\dagger &= \sum_{q_1q_2} UX^{k_1}_{q_1}U^\dagger\,UY^{k_2}_{q_2}U^\dagger\,\langle k_1k_2q_1q_2|kq\rangle \\ &= \sum_{\substack{q_1q_2\\q_1'q_2'}} X^{k_1}_{q_1'}Y^{k_2}_{q_2'}\,D^{k_1}_{q_1'q_1}(U)D^{k_2}_{q_2'q_2}(U)\,\langle k_1k_2q_1q_2|kq\rangle. \end{aligned} \tag{72}$$

Next we use Eq. (18.64) with a change of symbols,

$$D^{k_1}_{q_1'q_1}(U)D^{k_2}_{q_2'q_2}(U) = \sum_{KQQ'} \langle k_1k_2q_1'q_2'|KQ'\rangle D^K_{Q'Q}(U)\langle KQ|k_1k_2q_1q_2\rangle, \tag{73}$$

which we substitute into Eq. (72). Then the $q_1'q_2'$-sum is doable in terms of the expression (71) for T^k_q,

$$\sum_{q_1'q_2'} X^{k_1}_{q_1'}Y^{k_2}_{q_2'}\langle k_1k_2q_1'q_2'|KQ'\rangle = T^K_{Q'}, \tag{74}$$

and the $q_1 q_2$-sum is doable by the orthogonality of the Clebsch-Gordan coefficients,

$$\sum_{q_1 q_2} \langle KQ|k_1 k_2 q_1 q_2\rangle \langle k_1 k_2 q_1 q_2|kq\rangle = \delta_{Kk}\,\delta_{Qq}. \tag{75}$$

Then Eq. (72) becomes

$$UT^k_q U^\dagger = \sum_{KQQ'} T^K_{Q'} D^K_{Q'Q}\,\delta_{Kk}\,\delta_{Qq} = \sum_{q'} T^k_{q'} D^k_{q'q}(U). \tag{76}$$

This shows that T^k_q is an irreducible tensor operator, as claimed.

As an application, two vector operators $\mathbf{V}$ and $\mathbf{W}$, may be converted into $k = 1$ irreducible tensor operators V_q and W_q by going over to the spherical basis. From these we can construct $k = 0, 1, 2$ irreducible tensor operators according to

$$T^k_q = \sum_{q_1 q_2} V_{q_1} W_{q_2} \langle 11 q_1 q_2|kq\rangle. \tag{77}$$

This will yield the same decomposition of a second rank tensor discussed in Sec. 19.8, where we found a scalar ($k = 0$), a vector ($k = 1$), and a symmetric, traceless tensor ($k = 2$).

Problems

1. This will help you understand irreducible tensor operators better. Let $\mathcal{E}$ be a ket space for some system of interest, and let $\mathcal{A}$ be the space of linear operators that act on $\mathcal{E}$. For example, the ordinary Hamiltonian is contained in $\mathcal{A}$, as are the components of the angular momentum $\mathbf{J}$, the rotation operators $U(\mathsf{R})$, etc. The space $\mathcal{A}$ is a vector space in its own right, just like $\mathcal{E}$; operators can be added, multiplied by complex scalars, etc. Furthermore, we may be interested in certain subspaces of $\mathcal{A}$, such as the 3-dimensional space of operators spanned by the components V_x, V_y, V_z of a vector operator $\mathbf{V}$.

Now let $\mathcal{S}$ be the space of linear operators that act on $\mathcal{A}$. We call an element of $\mathcal{S}$ a "super" operator because it acts on ordinary operators; ordinary operators in $\mathcal{A}$ act on kets in $\mathcal{E}$. We will denote super-operators with a hat, to distinguish them from ordinary operators. (This terminology has nothing to do with supersymmetry.)

Given an ordinary operator $A \in \mathcal{A}$, it is possible to associate it in several different ways with a super-operator. For example, we can define a super operator $\hat{A}_L$, which acts by left multiplication:

$$\hat{A}_L X = AX, \tag{78}$$

where X is an arbitrary ordinary operator. This equation obviously defines a linear super-operator, that is, $\hat{A}_L(X + Y) = \hat{A}_L X + \hat{A}_L Y$, etc. Similarly, we can define a super-operator associated with A by means of right multiplication, or by means of the forming of the commutator, as follows:

$$\begin{aligned} \hat{A}_R X &= XA, \\ \hat{A}_C X &= [A, X]. \end{aligned} \tag{79}$$

There are still other ways of associating an ordinary operator with a super-operator. Let R be a classical rotation, and let $U(\mathsf{R})$ be a representation of the rotations acting on the ket space $\mathcal{E}$. Thus, the operators $U(\mathsf{R})$ belong to the space $\mathcal{A}$. Now associate such a rotation operator $U(\mathsf{R})$ in $\mathcal{A}$ with a super-operator $\hat{U}(\mathsf{R})$ in $\mathcal{S}$, defined by

$$\hat{U}(\mathsf{R})X = U(\mathsf{R})\,X\,U(\mathsf{R})^\dagger. \tag{80}$$

Again, $\hat{U}(\mathsf{R})$ is obviously a linear super-operator.

(a) Show that $\hat{U}(\mathsf{R})$ forms a representation of the rotations, that is, that

$$\hat{U}(\mathsf{R}_1)\hat{U}(\mathsf{R}_2) = \hat{U}(\mathsf{R}_1\mathsf{R}_2). \tag{81}$$

This is easy.

Now let $U(\mathsf{R})$ be infinitesimal as in Eq. (19.16), and let

$$\hat{U}(\mathsf{R}) = 1 - \frac{i}{\hbar}\theta\hat{\mathbf{n}}\cdot\hat{\mathbf{J}}. \tag{82}$$

(Here the hat on $\hat{\mathbf{n}}$ denotes a unit vector, while that on $\hat{\mathbf{J}}$ denotes a super-operator.) Express the super-operator $\hat{\mathbf{J}}$ in terms of ordinary operators. Write Eqs. (43) in super-operator notation. Work out the commutation relations of the super-operators $\hat{\mathbf{J}}$.

(b) Now write out nine equations, specifying the action of the three super-operators $\hat{J}_i$ on the the basis operators V_j. Write the answers as linear combinations of the V_j's. Then write out six more equations, specifying the action of the super raising and lowering operators, $\hat{J}_\pm$, on the three V_j.

Now find the operator A that is annihilated by $\hat{J}_+$. Do this by writing out the unknown operator as a linear combination of the V_j's, in the form

$$A = a_xV_x + a_yV_y + a_zV_z, \tag{83}$$

and then solving for the coefficients a_i. Show that this operator is an eigenoperator of $\hat{J}_z$ with eigenvalue $+\hbar$. In view of these facts, the operator A must be a "stretched" operator for $k = 1$; henceforth write T^1_1 for it. This operator will have an arbitrary, complex multiplicative constant, call it c. Now apply $\hat{J}_-$, and generate T^1_0 and T^1_{-1}. Choose the constant c to make T^1_0 look as simple as possible. Then write

$$T^1_q = \hat{\mathbf{e}}_q\cdot\mathbf{V}, \tag{84}$$

and thereby "discover" the spherical basis.

2. This problem concerns quadrupole moments and spins. It provides some background for problem 3.

(a) In the case of a nucleus, the spin Hilbert space $\mathcal{E}_{\text{spin}} = \text{span}\{|sm\rangle, m = -s,\ldots,+s\}$ is actually the ground state of the nucleus. It is customary to denote the angular momentum j of the ground

state by s. This state is $(2s+1)$-fold degenerate. The nuclear spin operator $\mathbf{S}$ is really the restriction of the total angular momentum of the nucleus $\mathbf{J}$ to this subspace of the (much larger) nuclear Hilbert space.

Let A^k_q and B^k_q be two irreducible tensor operators on $\mathcal{E}_{\rm spin}$. As explained in these notes, when we say "irreducible tensor operator" we are really talking about the collection of $2k+1$ operators obtained by setting $q = -k, \ldots, +k$. Use the Wigner-Eckart theorem to explain why any two such operators of the same order k are proportional to one another. This need not be a long answer.

Thus, all scalars are proportional to a standard scalar (1 is convenient), and all vector operators (for example, the magnetic moment $\boldsymbol{\mu}$) are proportional to a standard vector ($\mathbf{S}$ is convenient), etc.

For a given s, what is the maximum value of k? What is the maximum order of an irreducible tensor operator that can exist on space $\mathcal{E}_{\rm spin}$ for a proton (nucleus of ordinary hydrogen)? A deuteron (heavy hydrogen)? An alpha particle (nucleus of helium)? These rules limit the electric and magnetic multipole moments that a nucleus is allowed to have, as is discussed more fully in Notes 27.

(b) Let $\mathbf{A}$ and $\mathbf{B}$ be two vector operators (on any Hilbert space, not necessarily $\mathcal{E}_{\rm spin}$), with spherical components A_q, B_q, as in Eq. (31). As explained in the notes, A_q and B_q are $k=1$ irreducible tensor operators. As explained in Sec. 12, it is possible to construct irreducible tensor operators T^k_q for $k = 0, 1, 2$ out of the nine operators, $\{A_q B_{q'}, q, q' = -1, 0, 1\}$. Write out the three operators T^0_0, T^1_1 and T^2_2 in terms of the Cartesian products $A_i B_j$. Just look up the Clebsch-Gordan coefficients. There are nine operators in T^0_0, T^1_q and T^2_q, but I'm only asking you to compute these three to save you some work.

Show that T^0_0 is proportional to $\mathbf{A}\cdot\mathbf{B}$, that T^1_1 is proportional to a spherical component of $\mathbf{A}\times\mathbf{B}$, and that T^2_2 can be written in terms of the components of the symmetric and traceless part of the Cartesian tensor $A_i B_j$, which is

$$\frac{1}{2}(A_i B_j + A_j B_i) - \frac{1}{3}(\mathbf{A}\cdot\mathbf{B})\delta_{ij}. \tag{85}$$

(c) In classical electrostatics, the quadrupole moment tensor Q_{ij} of a charge distribution $\rho(\mathbf{x})$ is defined by

$$Q_{ij} = \int d^3\mathbf{x}\, \rho(\mathbf{x})[3x_i x_j - r^2\, \delta_{ij}], \tag{86}$$

where $\mathbf{x}$ is measured relative to some origin inside the charge distribution. The quadrupole moment tensor is a symmetric, traceless tensor. The quadrupole energy of interaction of the charge distribution with an external electric field $\mathbf{E} = -\nabla\phi$ is

$$E_{\rm quad} = \frac{1}{6}\sum_{ij} Q_{ij}\frac{\partial^2\phi(0)}{\partial x_i \partial x_j}. \tag{87}$$

This energy must be added to the monopole and dipole energies, plus the higher multipole energies.

In the case of a nucleus, we choose the origin to be the center of mass, whereupon the dipole moment and dipole energy vanish. The monopole energy is just the usual Coulomb energy $q\phi(0)$,

where q is the total charge of the nucleus. Thus, the quadrupole term is the first nonvanishing correction. However, the energy must be understood in the quantum mechanical sense.

Let $\{\mathbf{x}_\alpha, \alpha = 1, \ldots, Z\}$ be the position operators for the protons in a nucleus. The neutrons are neutral, and do not contribute to the electrostatic energy. The electric quadrupole moment operator for the nucleus is defined by

$$Q_{ij} = e \sum_\alpha (3x_{\alpha i} x_{\alpha j} - r_\alpha^2 \, \delta_{ij}), \tag{88}$$

where e is the charge of a single proton. In an external electric field, the nuclear Hamiltonian contains a term H_{quad}, exactly in the form of Eq. (87), but now interpreted as an operator.

The operator Q_{ij}, being symmetric and traceless, constitutes the Cartesian specification of a $k = 2$ irreducible tensor operator, that you could turn into standard form $T_q^2, q = -2, \ldots, +2$ using the method of part (b) if you wanted to. We'll stay with the Cartesian form here, however. When the operator Q_{ij} is restricted to the ground state (really a manifold of $2s + 1$ degenerate states), it remains a $k = 2$ irreducible tensor operator. According to part (a), it must be proportional to some standard irreducible tensor operator, for which $3S_iS_j - S^2\delta_{ij}$ is convenient. That is, we must be able to write

$$Q_{ij} = a(3S_iS_j - S^2\delta_{ij}), \tag{89}$$

for some constant a.

It is customary in nuclear physics to denote the "quadrupole moment" of the nucleus by the real number Q, defined by

$$Q = \langle ss|Q_{33}|ss\rangle, \tag{90}$$

where $|ss\rangle$ is the stretched state. Don't confuse Q_{ij}, a tensor of operators, with Q, a single number.

The book, *Modern Quantum Mechanics* by J. J. Sakurai gives the interaction energy of a nucleus in an external electric field as

$$H_{\text{int}} = \frac{eQ}{2s(s-1)\hbar^2}\Big[\Big(\frac{\partial^2\phi}{\partial x^2}\Big)S_x^2 + \Big(\frac{\partial^2\phi}{\partial y^2}\Big)S_y^2 + \Big(\frac{\partial^2\phi}{\partial z^2}\Big)S_z^2\Big], \tag{91}$$

where ϕ is the electrostatic potential for the external field satisfying the Laplace equation $\nabla^2\phi = 0$ and where the coordinate axes are chosen so that the off-diagonal elements of $\partial^2\phi/\partial x_i \partial x_j$ vanish. Here ϕ and its derivatives are evaluated at the center of mass of the nucleus and ϕ satisfies the Laplace equation rather than the Poisson equation because the sources of the external electric field are outside the nucleus.

Express the quantity a in Eq. (89) in terms of Q, and derive a version of Eq. (91). This equation, copied out of the book, has an error in it; correct it.

3. This is Sakurai, problem 3.29, p. 247; or Sakurai and Napolitano, problem 3.33, p. 261.

A spin-$\frac{3}{2}$ nucleus situated at the origin is subjected to an external inhomogeneous electric field. The basic electric quadrupole interaction is given by Eq. (91) (but corrected), where as above ϕ

satisfies the Laplace equation and the off-diagonal components $\partial^2\phi/\partial x_i \partial x_j$ vanish. Show that the interaction energy can be written

$$A(3S_z^2 - S^2) + B(S_+^2 + S_-^2), \tag{92}$$

and express A and B in terms of the nonvanishing second derivatives of ϕ, evaluated at the origin. Determine the energy eigenkets (in terms of $|m\rangle$, where $m = \pm\frac{3}{2}, \pm\frac{1}{2}$) and the corresponding energy eigenvalues. Is there any degeneracy?

Physics 221A
Academic Year 2020–21
Notes 21
Parity†

1. Introduction

We have now completed our study of proper rotations in quantum mechanics, one of the important space-time symmetries. In these notes we shall examine parity, another space-time symmetry. In a latter set of notes we shall take up the third, which is time reversal. All these symmetries are special cases of Lorentz transformations, which are united in a relativistic treatment.

2. Spatial Inversion

Parity begins life as an operation on three-dimensional, physical space, represented in an inertial frame by the matrix

$$\mathsf{P} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = -\mathsf{I}. \tag{1}$$

See also Eq. (11.16) and the discussion in Sec. 11.5. This operation, which we call *spatial inversion*, inverts vectors through the origin, mapping $\mathbf{x}$ into $-\mathbf{x}$. The most important properties of P for our purposes are

$$\mathsf{P}^2 = \mathsf{I}, \tag{2}$$

and

$$\mathsf{P}\mathsf{R}\mathsf{P}^{-1} = \mathsf{R}, \tag{3}$$

for all proper rotations $\mathsf{R} \in SO(3)$. These properties are trivial in view of Eq. (1); in particular, P, being a multiple of the identity, commutes with all matrices, so $\mathsf{P}\mathsf{R}\mathsf{P}^{-1} = \mathsf{R}\mathsf{P}\mathsf{P}^{-1} = \mathsf{R}$.

The spatial inversion P is an example of an improper rotation, that is, it belongs to $O(3)$ but not $SO(3)$. In fact, every proper rotation is mapped into an improper one by multiplying by P, and vice versa. See Fig. 11.5 to visualize the spaces of proper and improper rotations.

† Links to the other sets of notes can be found at:
`http://bohr.physics.berkeley.edu/classes/221/2021/221.html`.

3. Parity in Quantum Mechanics

In this section we introduce the operator π, called the *parity* operator, which corresponds to the spatial inversion operation P.

Recall that in the case of proper rotations, we sought and found operators $U(\mathsf{R})$, acting on the Hilbert space of various quantum mechanical systems, that implement the effect of classical, proper rotations R. The operators $U(\mathsf{R})$ were required to satisfy certain postulates, including one that they be unitary, and another that they reproduce the multiplication law of the classical, proper rotations. See Sec. 12.3. We represent this mapping between classical rotations and rotation operators by

$$\mathsf{R} \mapsto U(\mathsf{R}), \tag{4}$$

and we say that it specifies a unitary representation of the classical rotation group $SO(3)$.

Similarly, in the case of spatial inversion, we seek an operator π acting on the Hilbert space of various quantum mechanical systems that is associated with the spatial inversion operation P,

$$\mathsf{P} \mapsto \pi, \tag{5}$$

which is required to satisfy certain postulates. These are that π be unitary,

$$\pi^\dagger \pi = 1, \tag{6}$$

and that it should satisfy

$$\pi^2 = 1, \tag{7}$$

and

$$\pi U(\mathsf{R}) \pi^\dagger = U(\mathsf{R}) \tag{8}$$

for all $\mathsf{R} \in SO(3)$. The latter two postulates are representations of the multiplication laws (2) and (3) at the level of spatial transformations. The first two postulates imply that π is both unitary and Hermitian, so that

$$\pi = \pi^\dagger = \pi^{-1}, \tag{9}$$

while the third postulate (8) is equivalent to

$$[\pi, \mathbf{J}] = 0, \tag{10}$$

where $\mathbf{J}$ is the angular momentum (the generator of the rotation operators $U(\mathsf{R})$). In short, π is a scalar operator [see Eqs. (19.8) and (19.10)]. These postulates do not uniquely determine the parity operator for all quantum mechanical systems, but they narrow the possibilities considerably so that it is easy to find a suitable definition of π in specific cases.

Taken together with the properties of $U(\mathsf{R})$, the postulates for π amount to requiring that the set of operators $U(\mathsf{R})$ and $\pi U(\mathsf{R})$ for $\mathsf{R} \in SO(3)$ form a unitary representation of the full rotation group $O(3)$. That is, generalizing the map (4) to include improper rotations, we can say

$$\pi = U(\mathsf{P}). \tag{11}$$

4. Parity for a Spinless Particle in Three Dimensions

Let us begin with the case of a spinless particle moving in three dimensions, for which the basis states can be taken to be the position eigenkets $|\mathbf{x}\rangle$ and the Hilbert space is isomorphic to the space of configuration space wave functions $\psi(\mathbf{x})$. In this case we defined the rotation operators by

$$U(\mathsf{R})|\mathbf{x}\rangle = |\mathsf{R}\mathbf{x}\rangle, \tag{12}$$

which we found satisfactory. See Eq. (15.3) and the discussion in Sec. 15.2. In a similar spirit, we now guess that the definition of parity for such systems should be

$$\pi|\mathbf{x}\rangle = |\mathsf{P}\mathbf{x}\rangle = |-\mathbf{x}\rangle. \tag{13}$$

This transformation law on basis kets implies that wave functions transform according to

$$\psi(\mathbf{x}) \xrightarrow{\pi} \psi(-\mathbf{x}). \tag{14}$$

We must check that this definition satisfies the postulates (6)–(8).

First, by applying π twice, we have two changes of sign,

$$\pi^2|\mathbf{x}\rangle = |\mathbf{x}\rangle, \tag{15}$$

so the postulate (7) is satisfied. Next, since

$$\int d^3\mathbf{x}\,|\psi(\mathbf{x})|^2 = \int d^3\mathbf{x}\,|\psi(-\mathbf{x})|^2 \tag{16}$$

for all wave functions ψ, the operator π is unitary (it preserves the normalization of states), and the postulate (6) is satisfied. See Problem 1.6(c). Finally, since

$$\pi U(\mathsf{R})|\mathbf{x}\rangle = \pi|\mathsf{R}\mathbf{x}\rangle = |-\mathsf{R}\mathbf{x}\rangle = U(\mathsf{R})\pi|\mathbf{x}\rangle \tag{17}$$

we have $\pi U(\mathsf{R}) = U(\mathsf{R})\pi$, which is equivalent to the postulate (8). All postulates are satisfied, so we shall take Eq. (13) as the definition of π for a spinless particle in three-dimensional space.

As for the last postulate (8) and its equivalent form (10), since the angular momentum for a spinless particle in three-dimensional space is the orbital angular momentum $\mathbf{L} = \mathbf{x}\times\mathbf{p}$, we have

$$\pi\mathbf{L}\pi^\dagger = \mathbf{L}. \tag{18}$$

This follows from the postulates, but can also be verified directly.

It is also of interest to compute the conjugation relations of other vector operators with respect to π. We start with the position operator $\mathbf{x}$. For the sake of the following demonstration, we will write $\hat{\mathbf{x}}$ for the operator, and $\mathbf{x}$ for associated vectors of c-numbers. In the configuration representation, the operator $\hat{\mathbf{x}}$ means multiplication of the wave function by the position of the particle, that is,

$$(\hat{\mathbf{x}}\psi)(\mathbf{x}) = \mathbf{x}\psi(\mathbf{x}). \tag{19}$$

Now computing the effect of $\pi\hat{\mathbf{x}}\pi^\dagger$ on an arbitrary wave function, we have

$$\psi(\mathbf{x}) \xrightarrow{\pi^\dagger} \psi(-\mathbf{x}) \xrightarrow{\hat{\mathbf{x}}} \mathbf{x}\psi(-\mathbf{x}) \xrightarrow{\pi} -\mathbf{x}\psi(\mathbf{x}) = -(\hat{\mathbf{x}}\psi)(\mathbf{x}), \tag{20}$$

where we use $\pi^\dagger = \pi$.

Thus we have

$$\pi\mathbf{x}\pi^\dagger = -\mathbf{x}, \tag{21}$$

where we have dropped the hats so that now $\mathbf{x}$ represents the operator again. We see that conjugating the position operator $\mathbf{x}$ by π converts it into $-\mathbf{x}$. This is reasonable, in view of the action of P on points of space. Similarly, we find

$$\pi\mathbf{p}\pi^\dagger = -\mathbf{p}. \tag{22}$$

This is fairly obvious, since $\mathbf{p} = -i\hbar\nabla$ in the position representation, and $\nabla = \partial/\partial\mathbf{x}$ should change sign when $\mathbf{x}$ does. Taking the cross product of Eqs. (21) and (22), we obtain Eq. (18).

5. Parity for Spin Systems

Next let us consider just the spin degrees of freedom for a particle with spin, ignoring the spatial degrees of freedom for the time being. The Hilbert space is

$$\mathcal{E}_{\text{spin}} = \text{span}\{|sm\rangle, m = -s, \ldots, +s\}, \tag{23}$$

with the standard angular momentum basis shown. We would like to find a reasonable definition of π that satisfies the postulates (6)–(8). This means finding the action of π on the basis states $|sm\rangle$. Since the angular momentum in this case is the spin vector $\mathbf{S}$, the form (10) of the third postulate is

$$[\pi, \mathbf{S}] = 0. \tag{24}$$

The states $|sm\rangle$ are nondegenerate eigenstates of S_z on the Hilbert space (23), and Eq. (24) implies $[\pi, S_z] = 0$. Therefore by Theorem 1.5, $|sm\rangle$ must also be an eigenstate of parity,

$$\pi|sm\rangle = \eta_m|sm\rangle, \tag{25}$$

where as indicated the eigenvalue η_m might depend on m, as far as we know at this point. Theorem 1.5 does not tell us the value of this eigenvalue, but it must be real since π is Hermitian. Now by applying raising and lowering operators and using Eq. (24) again, we have

$$\begin{aligned} S_\pm(\pi|sm\rangle) = \pi S_\pm|sm\rangle &= \sqrt{(s \mp m)(s \pm m + 1)}\hbar\,\pi|s, m \pm 1\rangle \\ &= \eta_m\sqrt{(s \mp m)(s \pm m + 1)}\hbar\,|s, m \pm 1\rangle, \end{aligned} \tag{26}$$

or, on canceling the square roots,

$$\pi|s, m \pm 1\rangle = \eta_m|s, m \pm 1\rangle. \tag{27}$$

By comparing this to Eq. (25), we see that $\eta_m = \eta_{m\pm1}$, or, by induction, that η is independent of m for all $2s+1$ values it can take on. This fact could have been seen from another standpoint, for since π is a scalar operator, by the Wigner-Eckart theorem for scalars its eigenvalues must be independent of m. See Secs. 19.7-19.10. Thus we shall write simply

$$\pi|sm\rangle = \eta|sm\rangle. \tag{28}$$

In fact, since the states $|sm\rangle$ form a basis for $\mathcal{E}_{\text{spin}}$, all kets in this space are eigenstates of π with eigenvalue η, that is, π is η times the identity operator.

Now using the fact that $\pi^2 = 1$, we find that $\eta^2 = 1$, or $\eta = \pm1$, and we have

$$\pi|sm\rangle = \pm|sm\rangle. \tag{29}$$

The postulates required of π reduce its definition to a choice of a sign. This sign is the same for all vectors in the Hilbert space $\mathcal{E}_{\text{spin}}$ of Eq. (23), so it is really a characteristic of the particle, that is, a constant like the spin quantum number s.

Can this sign be determined by any further arguments or by appeal to experimental results? Within nonrelativistic quantum mechanics, the answer is no, the final choice of sign is without physical consequences and can be made arbitrarily. The simplest choice is $\eta = 1$, and for as long as we remain with the nonrelativistic theory, we shall assume that the parity operator satisfies

$$\pi|sm\rangle = |sm\rangle. \tag{30}$$

That is, the parity operator has no effect on the spin state of the particle.

In relativistic quantum mechanics, however, it becomes possible to create and destroy particles. In some cases, an initial state of a definite parity decays or otherwise evolves into a state in which a particle has been created or destroyed, and in order to account for conservation of parity it is necessary to attribute a certain parity to a particle because of the kind of particle it is. This is called *intrinsic parity*, and it is equivalent to making one of the two choice $\eta = \pm1$ for each type of particle in the process. Some details of this assignment depend on the type of particle, fermion or boson. (It turns out that for fermions, the intrinsic parity of a particle-antiparticle pair is odd; it is arbitrary whether the particle is interpreted as having even intrinsic parity and the antiparticle odd, or vice versa. See Notes 48.) Not all interactions conserve parity, but for those that do, it is necessary in relativistic interactions to take into account the intrinsic parities of particles. Then the total parity of a state is the product of the intrinsic parities of the particles times the parities of the spatial parts of the wave functions. For example, the photon turns out to have negative intrinsic parity. This is related to Laporte's rule, which we discuss presently, which states that the parity of an atom changes when a photon is emitted in an electric dipole transition. The parity of the atom changes, but the parity of the entire system, including the electromagnetic field, is conserved, as we see when we include the intrinsic parity of the photon. The intrinsic parity of particles cannot be understood very well outside of the relativistic theory, but we will have more to say about it Notes 48.

Returning to nonrelativistic quantum mechanics, if we include the spatial degrees of freedom of a particle, so that the Hilbert space is spanned by the basis kets

$$|\mathbf{x}\rangle \otimes |sm\rangle = |\mathbf{x}, m\rangle, \tag{31}$$

then the parity operator is defined by

$$\pi|\mathbf{x}, m\rangle = |-\mathbf{x}, m\rangle. \tag{32}$$

This is equivalent to the action on the wave function,

$$\psi_m(\mathbf{x}) \xrightarrow{\pi} \psi_m(-\mathbf{x}). \tag{33}$$

As we say, parity is a purely spatial operator, and does not affect the spin (in the nonrelativistic theory).

6. Transformation Properties Under Parity

Returning to the conjugation relations (18), (21) and (22), we see that some vectors (for example, $\mathbf{x}$ and $\mathbf{p}$) change sign when conjugated by parity, while some (for example, $\mathbf{L}$ and $\mathbf{S}$) do not. All these vectors are vector operators according to the definitions (19.15) or (19.21), that is, they all transform in the same way under conjugation by proper rotations, but their transformation laws under improper rotations differ. Thus we can take the set of vector operators and break it down according to how the vectors transform under parity. If $\mathbf{V}$ is a vector operator and

$$\pi\mathbf{V}\pi^\dagger = \pm\mathbf{V}, \tag{34}$$

then we will say that $\mathbf{V}$ is a *true vector* or *polar vector* if the sign is -1, or a *pseudovector* or *axial vector* if the sign is $+1$. Thus, $\mathbf{x}$ and $\mathbf{p}$ are true vectors, while $\mathbf{L}$ and $\mathbf{S}$ (more generally, any angular momentum $\mathbf{J}$) are pseudovectors. Other vectors in physics can be classified similarly, for example, the electric field $\mathbf{E}$ is a true vector, while the magnetic field $\mathbf{B}$ is a pseudovector.

Similarly, we can classify scalars. If K is a scalar operator according to the definitions (19.8) or (19.10), and if

$$\pi K\pi^\dagger = \pm K, \tag{35}$$

then we will say that K is a *true scalar* if the sign is $+1$ or a *pseudoscalar* if the sign is -1 (notice the signs are reversed compared to the terminology for a vector). An example of a true scalar is $\mathbf{x}\cdot\mathbf{p}$, while an example of a pseudoscalar is $\mathbf{p}\cdot\mathbf{S}$.

7. Hamiltonians Invariant Under Parity

We argued previously that the Hamiltonian for an isolated system must be a scalar, that is, invariant under proper rotations, because the energy could not depend on the orientation. See

Sec. 19.7. But are these Hamiltonians scalars or pseudoscalars or a combination of the two? That is, are they also invariant under parity?

As an example, consider a central force Hamiltonian,

$$H = \frac{p^2}{2m} + V(r). \tag{36}$$

This certainly commutes with parity, for although $\mathbf{p}$ is mapped into $-\mathbf{p}$, the kinetic energy contains $\mathbf{p} \cdot \mathbf{p}$, which is invariant. Also, the potential energy is a function of the radius r, the magnitude of the vector $\mathbf{x}$, which does not change when $\mathbf{x}$ changes its sign. Parity is a good quantum number for all central force Hamiltonians.

As we know, central force Hamiltonians usually represent a two-particle system interacting by means of forces derivable from a potential that is a function only of the distance between the two particles. See Sec. 16.9. All these Hamiltonians commute with parity. This generalizes to any number of particles interacting by means of forces derivable from potentials depending only on the distances between the particles, such as the electrostatic Coulomb approximation to the interactions of charged particles. This covers a wide variety of approximate Hamiltonians for atoms, molecules, and other systems. What happens when we include electromagnetic and relativistic effects? As it turns out, such Hamiltonians still commute with parity, as long as the electromagnetic fields are internally generated (that is, the system is still isolated). For example, one of the magnetic and relativistic corrections in a single electron atom is the spin-orbit term, which is proportional to $\mathbf{L} \cdot \mathbf{S}$. But by the rules above, this is a true scalar (the dot product of two pseudovectors). What about when we include the emission and absorption of radiation, that is, of photons? In this case, too, parity is conserved, as long as we include the parity of the electromagnetic field (that is, of the photons) in our account. We summarize these facts by saying that the electromagnetic forces conserve parity.

The strong forces, responsible for the interactions of nucleons inside a nucleus as well as interactions of other hadronic particles such as the π-meson, also conserve parity. This fact was originally established by experiments involving nuclei and particles, which showed that in strong interactions the parity of the initial state is always equal to the parity of the final state. In reckoning the parity of states, however, it is necessary to take into account the intrinsic parities of various particles, for example, the π-meson (all three charge states of it) has negative intrinsic parity (just like the photon).

For these reasons, it used to be believed that parity conservation was a fundamental law of nature, much like the rotational invariance of isolated systems. This belief ended abruptly in 1956, however, when Lee and Yang suggested that parity might be violated in weak interactions. They also suggested experiments to test the idea that were quickly carried out. These showed that, indeed, parity is violated in β-decay, a prime example of a weak interaction.

The weak forces are present in ordinary atomic and nuclear interactions, but, as the name implies, they are very small compared to the electromagnetic and strong interactions, at least at low

energy. The weak forces become relatively stronger as the energy is increased, and it is currently believed that at energies much higher than now available in accelerators the strong, electromagnetic and weak forces all become comparable in strength. Considering only low energy interactions, however, we can say that the weak interactions are extremely small compared to the electromagnetic and nuclear interactions, in fact, they are so small that they are quite difficult to detect even if one is looking for them. There has been some effort in the Berkeley Physics department to detect the effects of weak interactions in atomic physics, and these experiments are difficult. For most ordinary purposes at ordinary energies, the weak interactions can be ignored and all Hamiltonians for isolated systems conserve parity to an extremely high degree of accuracy.

It is, however, easy to create a system that does not conserve parity, simply by placing it in an external field. Such a system is no longer isolated, and parity need not be conserved. For example, suppose a particle of charge q moving in a central force field is placed in an external, uniform electric field $\mathbf{E}_0$. Then the Hamiltonian is

$$H = \frac{p^2}{2m} + V(r) - q\mathbf{x}\cdot\mathbf{E}_0. \tag{37}$$

This does not commute with parity because the last term changes sign upon conjugation by parity.

In such cases, however, parity conservation is restored if we enlarge our definition of "the system" to include the charges creating the electric field. For example, if $\mathbf{E}_0$ is created by positive and negative charges arrayed along the parallel plates of a capacitor, and if these charges are included in "the system," then π will swap the positive and negative charges and reverse the direction of $\mathbf{E}_0$. Since $\mathbf{x}$ also changes sign, the whole Hamiltonian will once again be invariant under parity.

8. Parity and Energy Eigenstates in Isolated Systems

Suppose we have an isolated system with Hamiltonian H. As we have explained, in most applications H commutes with parity to an extremely high degree of accuracy, so let us assume that $\pi H \pi^\dagger = H$, that is, $[\pi, H] = 0$.

Since the system is isolated, H commutes with all proper rotations, and is a scalar operator according to the definition (19.8). Then Theorem 19.1 states that the energy eigenspaces consist of one or more irreducible subspaces under rotations, and the addendum says that, with few exceptions, they consist of precisely one irreducible subspace. So let us consider this generic case, for which each energy eigenvalue is associated with a definite j value and has an energy eigenspace that is $2j+1$-fold degenerate.

As in Notes 13 and 19 let $\mathcal{S}_{jm}$ be the eigenspace of operators J^2 and J_3 with quantum numbers j and m, and let $|\psi\rangle$ be a bound energy eigenstate in the stretched subspace $\mathcal{S}_{jj}$ with eigenvalue E, so that

$$J^2|\psi\rangle = j(j+1)\hbar^2\,|\psi\rangle, \quad J_3|\psi\rangle = j\hbar\,|\psi\rangle, \quad H|\psi\rangle = E|\psi\rangle, \tag{38}$$

Under our assumptions of genericity, H restricted to $\mathcal{S}$ is nondegenerate, that is, $|\psi\rangle$ is the only state in $\mathcal{S}_{jj}$ with the energy E.

Now π commutes with $\mathbf{J}$ and hence with J^2 and J_z, and so it can be restricted to $\mathcal{S}_{jj}$. And since $|\psi\rangle$ is a nondegenerate eigenstate of H and since π commutes with H, according to Theorem 1.5, $|\psi\rangle$ must also be an eigenstate of parity,

$$\pi|\psi\rangle = e|\psi\rangle, \tag{39}$$

where $e = \pm 1$. Theorem 1.5 does not tell us what the value of e is.

Now it is easy to show that all the states obtained by applying lowering operators to $|\psi\rangle$ are also eigenstates of parity with the same eigenvalue e as the stretched state. That is,

$$\pi(J_-)^r|\psi\rangle = (J_-)^r\pi|\psi\rangle = e(J_-)^r|\psi\rangle, \tag{40}$$

for $r = 1, 2, \ldots, 2j$. That is, all $2j+1$ degenerate energy eigenstates in the irreducible subspace have the same parity.

Thus the parity value can be associated with the energy level itself, just like the j value. This is why the three nuclear states shown in Fig. 19.1 are assigned definite values of parity (all three, ^{57}Fe, ^{57}Fe* and ^{57}Fe**, have negative parity).

In the case of nuclei, the energies, angular momenta (or spins) and parities of energy levels are determined experimentally, since theoretical calculations of these quantities are difficult. In atoms and molecules, theory is in better shape, and these quantities can be calculated with greater or lesser difficulty.

In the nongeneric case of the electrostatic model of hydrogen, it is not true that all energy eigenstates are eigenstates of parity. It is true that in central force motion each angular momentum value ℓ is associated with the definite parity $(-1)^\ell$ [see Eq. (52)], but since there are degeneracies of energy that cross angular momentum values, it is possible to form linear combinations of energy eigenstates of the same energy but opposite parity. For example, the $2s$ and $2p$ levels in hydrogen have the same energy but opposite parities. These facts play an important role in understanding the Stark effect, which is discussed in Notes 24.

9. Parity in One-Dimensional Systems

Consider a one-dimensional Hamiltonian,

$$H = \frac{p^2}{2m} + V(x). \tag{41}$$

This commutes with parity if and only if $V(x) = V(-x)$, that is, if the potential is symmetric about $x = 0$. Now in one-dimensional systems with a localized potential, the bound energy eigenstates are nondegenerate because the wave functions must vanish at $x = \pm\infty$. See Sec. 6.2. Therefore by Theorem 1.5, the energy eigenstates must also be eigenstates of parity. That is, the bound energy eigenfunctions satisfy $\psi(x) = \pm\psi(-x)$, for some choice of sign that depends on the eigenfunction. For example, in the harmonic oscillator, which has the symmetric potential $V(x) = m\omega^2x^2/2$, the n-th energy eigenfunction has parity $(-1)^n$. See Sec. 8.6.

10. Considerations of Symmetry Raised by Parity

Parity illustrates some issues that arise whenever a Hamiltonian possesses a symmetry. Some of these have already been discussed in connection with rotations, but parity is a simpler symmetry so some of the ideas are easier to appreciate.

Whenever a Hamiltonian possesses a symmetry and we wish to diagonalize it or otherwise study its properties, it is a good idea to diagonalize the symmetry operators first. This is because symmetry operators are usually easier to diagonalize than Hamiltonians, and their eigenbases, sometimes called *symmetry-adapted* bases, are especially convenient for further study of the Hamiltonian. For example, in the case of rotational invariance, the symmetry adapted basis is the standard angular momentum basis $|\gamma jm\rangle$, an eigenbasis of J^2 and J_z.

Similarly, if a Hamiltonian commutes with parity, it may be convenient to study it in a basis which is an eigenbasis of parity. Since $\pi^2 = 1$, the eigenvalues e satisfy $e^2 = 1$, or $e = \pm 1$. Thus the Hilbert space is decomposed into two subspaces, the eigenspaces of parity with $e = \pm 1$. We can call these the even and odd subspaces,

$$\mathcal{E} = \mathcal{E}_{\text{even}} \oplus \mathcal{E}_{\text{odd}}. \tag{42}$$

Then a symmetry adapted basis consists of an orthonormal basis inside $\mathcal{E}_{\text{even}}$ plus an orthonormal basis inside $\mathcal{E}_{\text{odd}}$. We can write the basis vectors as $|\gamma e\rangle$, where

$$\pi|\gamma e\rangle = e|\gamma e\rangle, \tag{43}$$

so that e indicates which subspace the vectors lie in, and where γ is a label of the orthonormal basis vectors inside $\mathcal{E}_{\text{even}}$ (for $e = +1$) or inside $\mathcal{E}_{\text{odd}}$ (for $e = -1$). This notation is similar to the notation $|\gamma jm\rangle$ which we used for a standard angular momentum basis, where the index γ plays a similar role.

It is possible to write down projection operators that project onto the even and odd subspaces in terms of the parity operator itself. These are

$$P_\pm = \frac{1 \pm \pi}{2}. \tag{44}$$

It is easy to show that the operators $P_\pm$ satisfy the requirements of a projection operator (see Sec. 1.24), and that when applied to an arbitrary state $|\psi\rangle$ (not necessarily an eigenstate of parity), they project it onto one or the other of the two subspaces, producing an eigenstate of parity of the eigenvalue indicated. For example, in a one-dimensional problem, if $\psi(x)$ is an arbitrary wave function (not necessarily even or odd), then

$$(P_+\psi)(x) = \frac{1}{2}[\psi(x) + \psi(-x)] \tag{45}$$

and

$$(P_-\psi)(x) = \frac{1}{2}[\psi(x) - \psi(-x)] \tag{46}$$

are the even and odd parts of $\psi(x)$, respectively.

The matrix elements of the Hamiltonian in the symmetry adapted basis are noteworthy. Assuming the Hamiltonian commutes with parity, we have

$$\langle\gamma' e'|H|\gamma e\rangle = \langle\gamma' e'|\pi^\dagger H\pi|\gamma e\rangle = e'e\,\langle\gamma' e'|H|\gamma e\rangle, \tag{47}$$

which implies that the matrix element vanishes unless $e'e = 1$. But this is equivalent to $e' = e$, so the Hamiltonian is diagonal in the parity quantum number e, and we have

$$\langle\gamma' e'|H|\gamma e\rangle = \delta_{e'e}\,C^e_{\gamma'\gamma}, \tag{48}$$

where $C^e_{\gamma'\gamma}$ just indicates the indices the nonvanishing matrix elements can depend on. This may be compared to Eq. (20.54), which applies to Hamiltonians that are invariant under proper rotations. The fact that the matrix element vanishes if $e' \neq e$ is sometimes stated in words by saying that the Hamiltonian does not "connect" states of different parity.

In many cases it is necessary to diagonalize a Hamiltonian numerically, by setting up a matrix for it in some basis. There is the question, however, of what basis to use. If the Hamiltonian respects a symmetry then there is great advantage in using a symmetry adapted basis, since it reduces the size of the matrices that must be diagonalized. For example, in the parity-adapted basis $|\gamma e\rangle$, to diagonalize a Hamiltonian that commutes with parity we must diagonalize the two matrices $C^e_{\gamma'\gamma}$ with rows and columns indexed by γ' and γ, for $e = \pm 1$. Effectively we are diagonalizing the Hamiltonian separately in each of the subspaces $\mathcal{E}_{\rm even}$ and $\mathcal{E}_{\rm odd}$. If we arrange the basis vectors so that all the even vectors $|\gamma, +1\rangle$ come first, and all the odd vectors $|\gamma, -1\rangle$ come second, then the Hamiltonian matrix has the form,

$$\langle\gamma' e'|H|\gamma e\rangle = \left(\begin{array}{c|c} X & 0 \\ \hline 0 & X \end{array}\right), \tag{49}$$

where the upper (or even-even) block is the matrix $C^{+1}_{\gamma'\gamma}$, and the lower (or odd-odd) block is the matrix $C^{-1}_{\gamma'\gamma}$, and where X indicates a non-zero block.

There are many numerical algorithms for diagonalizing a matrix, but a common one (the Householder algorithm) involves a computational effort that scales as N^3, where N is the size of the matrix. Thus the effort to diagonalize two matrices of size $N/2$ is $2(N/2)^3 = N^3/4$, or about $1/4$ the effort of diagonalizing a single $N \times N$ matrix. This is the improvement in computational efficiency when using a symmetry adapted basis for a Hamiltonian that commutes with parity, as compared to diagonalizing the Hamiltonian in an arbitrary (nonsymmetric) basis.

11. Parity and Selection Rules

Parity gives rise to important selection rules which we will illustrate for the case of dipole transitions in an electrostatic model of a single-electron atom. The energy eigenfunctions are the usual ones in central force motion,

$$\psi_{n\ell m}(r,\theta,\phi) = R_{n\ell}(r)Y_{\ell m}(\theta,\phi), \tag{50}$$

which we shall write in ket language as $|n\ell m\rangle$. The matrix elements relevant to electric dipole transitions are

$$\langle n'\ell'm'|\mathbf{x}|n\ell m\rangle. \tag{51}$$

We need to find the action of parity on the eigenstates $|n\ell m\rangle$. Since the radius r is invariant under the replacement $\mathbf{x} \to -\mathbf{x}$, only the $Y_{\ell m}$ part of the wave function is affected by parity. The $Y_{\ell m}$'s have the property that if they are multiplied by r^ℓ, where r is the radius, then the result is a homogeneous polynomial in the Cartesian coordinates x, y and z of degree ℓ. See Sec. 15.8. But since x, y and z all change sign under parity, and since the factor r^ℓ is invariant, the $Y_{\ell m}$ changes by $(-1)^\ell$ under parity. We summarize this by writing

$$\pi|n\ell m\rangle = (-1)^\ell|n\ell m\rangle. \tag{52}$$

The energy eigenfunctions (50) are also eigenfunctions of parity, with eigenvalue $(-1)^\ell$. This is often a useful fact. Notice that the eigenvalues do not depend on m. This was to be expected, since π is a scalar operator, and according to the Wigner-Eckart theorem, the eigenvalues of a scalar operator do not depend on m. See Sec. 19.7.

Now using the fact that $\mathbf{x}$ is odd under parity, we can make some statements about the matrix element (51). That is,

$$\langle n'\ell'm'|\mathbf{x}|n\ell m\rangle = -\langle n'\ell'm'|\pi^\dagger\mathbf{x}\pi|n\ell m\rangle = -(-1)^{\ell+\ell'}\langle n'\ell'm'|\mathbf{x}|n\ell m\rangle. \tag{53}$$

Thus, the matrix element (51) vanishes unless $-(-1)^{\ell+\ell'} = 1$, or

$$\Delta\ell = \text{odd}, \tag{54}$$

where $\Delta\ell = \ell - \ell'$.

We see that the parity of the atomic state must change under an electric dipole transition. This selection rule is known as *Laporte's rule*. Initially it was observed experimentally, and later it was explained by Wigner on the basis of parity. As noted above, Laporte's rule is equivalent to the statement that the intrinsic parity of the photon is negative.

In the case of the matrix element (51), the Wigner-Eckart theorem gives the selection rule $\Delta\ell = 0, \pm 1$, that is, these are the values of $\Delta\ell$ allowed by rotational invariance alone. But parity excludes the value $\Delta\ell = 0$, so the final selection rule is $\Delta\ell = \pm 1$.

Problems

1. This is essentially Sakurai's problem 4.4, p. 282 (revised edition), or problem 4.4, p. 300 (second edition), but I've rewritten it to make it more clear.

We did not discuss the spin angular functions in class, but these are spinor functions on the unit sphere that arise when we combine orbital and spin angular momentum for a central force problem

for a spinning particle. Here we will take the case $s = \frac{1}{2}$ (for example, in hydrogen). Let $|n\ell m_\ell\rangle$ be ket language for the wave function $R_{n\ell}(r)Y_{\ell m_\ell}(\theta,\phi)$, the solution of the Schrödinger equation for a spinless particle in a central force field, and let $|sm_s\rangle$ be the usual spin states (here $s = \frac{1}{2}$ and $m_s = \pm\frac{1}{2}$). We distinguish between m_ℓ and m_s, the two types of magnetic quantum numbers. We multiply the wave functions times the spin functions and form linear combinations with the Clebsch-Gordan coefficients to get eigenstates of J^2 and J_z, where $\mathbf{J} = \mathbf{L} + \mathbf{S}$. These are

$$|n\ell j m_j\rangle = \sum_{m_\ell, m_s} |n\ell m_\ell\rangle |sm_s\rangle \langle \ell s m_\ell m_s | j m_j\rangle. \tag{55}$$

The spatial wave functions $\psi_{n\ell m_\ell}(\mathbf{x}) = R_{n\ell}(r)Y_{\ell m_\ell}(\theta,\phi)$ factor into a radial part times an angular part. Let us write this in ket language as $|n\ell m_\ell\rangle = |n\ell\rangle|\ell m_\ell\rangle$. The factor $R_{n\ell}(r)$ or $|n\ell\rangle$ is the same for all terms in the sum above, so it can be taken out and what is left is a two-component spinor that depends only on the angles. This is what in wave function language Sakurai calls $\mathcal{Y}_\ell^{jm_j}$; in ket language we will write

$$|\ell j m_j\rangle = \sum_{m_\ell m_s} |\ell m_\ell\rangle |sm_s\rangle \langle \ell s m_\ell m_s | j m_j\rangle. \tag{56}$$

(a) For the case $\ell = 0$, $j = \frac{1}{2}$, $m_j = \frac{1}{2}$, write out the two-component spinor $\mathcal{Y}_\ell^{jm_j}$ as functions of (θ,ϕ).

(b) Multiply this by $\boldsymbol{\sigma}\cdot\mathbf{x}$, and express the result as a linear combination of other spin angular functions $\mathcal{Y}_\ell^{jm_j}$.

(c) Certain values of j, m_j and ℓ occur in the sum, and certain others do not. Use symmetry principles to explain why the values that occur are allowed and the others are not.

Hint: It may help to think of three kinds of rotations: spin rotations, orbital rotations, and total (spin plus orbital) rotations.

2. This is Sakurai, problem 4.5, p. 282 (revised edition), or problem 4.5, p. 300 (second edition).

Because of weak (neutral-current) interactions, there is a parity-violating potential between the atomic electron and the nucleus as follows:

$$V = \lambda[\delta(\mathbf{x})\,\mathbf{S}\cdot\mathbf{p} + \mathbf{S}\cdot\mathbf{p}\,\delta(\mathbf{x})], \tag{57}$$

where $\mathbf{S}$ and $\mathbf{p}$ are the spin and momentum operators of the electron, and the nucleus is assumed to be situated at the origin. As a result, the ground state of an alkali atom, usually characterized by $|n\ell jm\rangle$, actually contains very tiny contributions from other eigenstates as follows:

$$|n\ell jm\rangle \to |n\ell jm\rangle + \sum_{n'\ell'j'm'} C_{n'\ell'j'm'}\,|n'\ell'j'm'\rangle, \tag{58}$$

On the basis of symmetry considerations alone, what can you say about the values of $(n'\ell'j'm')$ that give rise to nonvanishing contributions? Suppose the radial wave functions and energy levels are all

known. Indicate how you can calculate $C_{n'\ell'j'm'}$. Are there any further restrictions on $(n'\ell'j'm')$ that go beyond those due to symmetry considerations?

3. This problem concerns a system whose symmetry is an improper rotation.

A collection of $N \geq 3$ nuclei is distributed in the x-y plane. They do not lie on a line. Nucleus i has position $\mathbf{R}_i = (R_{ix}, R_{iy}, 0)$. The positions of the nuclei are fixed. An electron moves in the field of these nuclei.

Let A be the reflection in the x-y plane. It is represented by the matrix

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{59}$$

which maps (x, y, z) into $(x, y, -z)$. The electrostatic potential $\Phi(\mathbf{x})$ of the nuclei is invariant under reflection in the x-y plane, that is,

$$\Phi(x, y, z) = \Phi(x, y, -z), \tag{60}$$

or

$$\Phi(\mathsf{A}\mathbf{x}) = \Phi(\mathbf{x}) \tag{61}$$

for short.

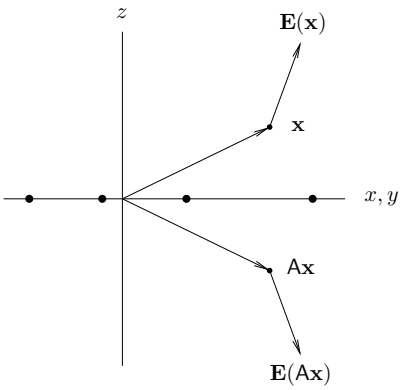

Fig. 1. The x-y plane, viewed from the side. The spots in the x-y plane are the nuclei. $\mathbf{x}$ is a typical point, and $\mathsf{A}\mathbf{x}$ is the reflection of this point through the x-y plane. The figure illustrates Eqs. (62) and (63).

From $\mathbf{E} = -\nabla\Phi$ and Eq. (60) or (61) it follows that

$$\begin{pmatrix} E_x(x, y, -z) \\ E_y(x, y, -z) \\ E_z(x, y, -z) \end{pmatrix} = \begin{pmatrix} E_x(x, y, z) \\ E_y(x, y, z) \\ -E_z(x, y, z) \end{pmatrix}, \tag{62}$$

or

$$\mathbf{E}(\mathsf{A}\mathbf{x}) = \mathsf{A}\mathbf{E}(\mathbf{x}) \tag{63}$$

for short. This also follows by inspection of Fig. 1.

(a) Every improper rotation can be written as $\mathsf{PR}(\hat{\mathbf{n}},\theta)$, where P is given by Eq. (1) and $\mathsf{R}(\hat{\mathbf{n}},\theta)$ is a proper rotation in axis-angle form. Write

$$\mathsf{A} = \mathsf{PR}(\hat{\mathbf{n}},\theta) \tag{64}$$

and identify $\hat{\mathbf{n}}$ and θ.

(b) Is $\Phi(\mathbf{x})$ a scalar operator? Is $\mathbf{E}(\mathbf{x})$ a vector operator?

(c) In the electrostatic approximation the Hamiltonian for the electron is

$$H_0 = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}), \tag{65}$$

where

$$V(\mathbf{x}) = -e\Phi(\mathbf{x}). \tag{66}$$

Define a unitary operator W that acts on the Hilbert space of the electron and that corresponds to A above in some reasonable way. I suggest you denote the usual parity operator by Π, to avoid confusion with the angle π. You may find it convenient to refer to different types of rotation operators,

$$\begin{aligned} U_o(\hat{\mathbf{n}},\theta) &= e^{-i\theta\hat{\mathbf{n}}\cdot\mathbf{L}/\hbar} \qquad \text{(orbital)}, \\ U_s(\hat{\mathbf{n}},\theta) &= e^{-i\theta\hat{\mathbf{n}}\cdot\mathbf{S}/\hbar} \qquad \text{(spin)}, \\ U_t(\hat{\mathbf{n}},\theta) &= e^{-i\theta\hat{\mathbf{n}}\cdot\mathbf{J}/\hbar} \qquad \text{(total)}, \end{aligned} \tag{67}$$

as needed, where $\mathbf{J} = \mathbf{L} + \mathbf{S}$.

Work out the conjugation relations,

$$W^\dagger \begin{pmatrix} x \\ y \\ z \end{pmatrix} W, \qquad \text{and} \qquad W^\dagger \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} W, \tag{68}$$

or simply $W^\dagger \mathbf{x} W$ and $W^\dagger \mathbf{p} W$ for short. Note that for any function $f(\mathbf{x})$,

$$W^\dagger f(\mathbf{x}) W = f\big(W^\dagger \mathbf{x} W\big). \tag{69}$$

Show that

$$W^\dagger H_0 W = H_0. \tag{70}$$

If it does not, go back and change the definition so that it does. Notice that since you defined W to be unitary, this is equivalent to $[W, H_0] = 0$.

(d) So the system is invariant under reflection in the plane, and we have found an operator corresponding to that reflection that commutes with the electrostatic Hamiltonian H_0. Now we will see if the symmetry persists when we include relativistic corrections.

The spin-orbit Hamiltonian worked out in Notes 25 applied only in the case of a radial electric field. The more general expression is

$$H_{SO} = \frac{e}{4m^2c^2}\big[\mathbf{E}(\mathbf{x})\times\mathbf{p} - \mathbf{p}\times\mathbf{E}(\mathbf{x})\big]\cdot\mathbf{S}, \tag{71}$$

where the quantity in the square brackets is the sum of an operator and its Hermitian conjugate, to make H_{SO} Hermitian.

Does the operator W you defined in part (c) commute with H_{SO}? If not, modify it so that it does, and call the new operator X. Make sure it is still unitary. Depending on your choice of W it is possible that no modification is necessary, then $X = W$. If $X \neq W$, reconsider the commutation relations (68). Also work out the commutation relations

$$X^\dagger \begin{pmatrix} E_x(\mathbf{x}) \\ E_y(\mathbf{x}) \\ E_z(\mathbf{x}) \end{pmatrix} X \qquad \text{and} \qquad X^\dagger \begin{pmatrix} S_x \\ S_y \\ S_z \end{pmatrix} X, \tag{72}$$

or $X^\dagger \mathbf{E}(\mathbf{x}) X$ and $X^\dagger \mathbf{S} X$ for short. Hint: notice that $[(\mathsf{R}\mathbf{A})\times(\mathsf{R}\mathbf{B})] \cdot (\mathsf{R}\mathbf{C}) = (\mathbf{A}\times\mathbf{B}) \cdot \mathbf{C}$, for all proper rotations R and all vectors $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$. See Prob. 11.2.

(e) In quantum mechanics we like Hermitian operators that commute with the Hamiltonian, because they possess simultaneous eigenbases, and for other reasons. The operator X you found in part (d) should be unitary. Find a simple expression for X^2.

Is the X you found in part (d) Hermitian? If not, make a simple modification to it to create a new operator, call it K, that is both unitary and Hermitian and that commutes with H. If no modification is necessary, then K will be the same as X obtained in part (d). If K is both unitary and Hermitian, then its its eigenvalues must be ± 1. Show that this is so.

Let the electron wave function be

$$\psi = \begin{pmatrix} \psi_+(x,y,z) \\ \psi_-(x,y,z) \end{pmatrix}. \tag{73}$$

Find the $\pm$ components of the wave function $K\psi$, and find the conditions that ψ_+ and ψ_- must satisfy in order that ψ be an eigenfunction of K with eigenvalues ± 1.

Thus the Hilbert space breaks up into two orthogonal eigenspaces of K, which we denote by $\mathcal{E}_\pm$, where $k = \pm 1$ is the eigenvalue of K. Write this as

$$\mathcal{E} = \mathcal{E}_+ \oplus \mathcal{E}_-. \tag{74}$$

This system and its symmetry become more interesting when time reversal is added to the mix of symmetries. See the continuation of this problem in Prob. 22.6.

Physics 221A
Academic Year 2020–21
Notes 22
Time Reversal†

1. Introduction

We have now considered the space-time symmetries of translations, proper rotations, and spatial inversions (that is, improper rotations) and the operators that implement these symmetries on a quantum mechanical system. We now turn to the last of the space-time symmetries, namely, time reversal. As we shall see, time reversal is different from all the others, in that it is implemented by means of antiunitary transformations.

2. Time Reversal in Classical Mechanics

Consider the classical motion of a single particle in three-dimensional space. Its trajectory $\mathbf{x}(t)$ is a solution of the equations of motion, $\mathbf{F} = m\mathbf{a}$. We define the time-reversed classical motion as $\mathbf{x}(-t)$. It is the motion we would see if we took a movie of the original motion and ran it backwards. Is the time-reversed motion also physically allowed (that is, does it also satisfy the classical equations of motion)?

The answer depends on the nature of the forces. Consider, for example, the motion of a charged particle of charge q in a static electric field $\mathbf{E} = -\nabla\Phi$, for which the equations of motion are

$$m\frac{d^2\mathbf{x}}{dt^2} = q\mathbf{E}(\mathbf{x}). \tag{1}$$

If $\mathbf{x}(t)$ is a solution of these equations, then so is $\mathbf{x}(-t)$, as follows from the fact that the equations are second order in time, so that the two changes of sign coming from $t \to -t$ cancel. However, this property does not hold for magnetic forces, for which the equations of motion include first order time derivatives:

$$m\frac{d^2\mathbf{x}}{dt^2} = \frac{q}{c}\frac{d\mathbf{x}}{dt}\times\mathbf{B}(\mathbf{x}). \tag{2}$$

In this equation, the left-hand side is invariant under $t \to -t$, while the right-hand side changes sign. For example, in a constant magnetic field, the sense of the circular motion of a charged particle (clockwise or counterclockwise) is determined by the charge of the particle, not the initial conditions, and the time-reversed motion $\mathbf{x}(-t)$ has the wrong sense. We see that motion in a given electric field is time-reversal invariant, while in a magnetic field it is not.

† Links to the other sets of notes can be found at:
`http://bohr.physics.berkeley.edu/classes/221/2021/221.html`.

We must add, however, that whether a system is time-reversal invariant depends on the definition of "the system." In the examples above, we were thinking of the system as consisting of a single charged particle, moving in given fields. But if we enlarge "the system" to include the charges that produce the fields (electric and magnetic), then we find that time-reversal invariance is restored, even in the presence of magnetic fields. This is because when we set $t \to -t$, the velocities of all the particles change sign, so the current does also. But this change does nothing to the charges of the particles, so the charge density is left invariant. Thus, the rules for transforming charges and currents under time reversal are

$$\rho \to \rho, \qquad \mathbf{J} \to -\mathbf{J}. \tag{3}$$

But according to Maxwell's equations, this implies the transformation laws

$$\mathbf{E} \to \mathbf{E}, \qquad \mathbf{B} \to -\mathbf{B}, \tag{4}$$

for the electromagnetic field under time reversal. With these rules, we see that time-reversal invariance is restored to Eq. (2), since there are now two changes of sign on the right hand side.

Thus we have worked out the basic transformation properties of the electromagnetic field under time reversal, and we find that electromagnetic effects are overall time-reversal invariant. We have shown this only in classical mechanics, but it is also true in quantum mechanics.

Similarly, in quantum physics we are often interested in the time-reversal invariance of a given system, such as an atom interacting with external fields. The usual point of view is to take the external fields as just given, and not to count them as part of the system. Under these circumstances the atomic system is time-reversal invariant if there are no external magnetic fields, but time-reversal invariance is broken in their presence. On the other hand, the atom generates its own, internal, magnetic fields, such as the dipole fields associated with the magnetic moments of electrons or nuclei, or the magnetic field produced by the moving charges. Since these fields are produced by charges that are a part of "the system," however, they do not break time-reversal invariance. We summarize these facts by saying that electromagnetic effects are time-reversal invariant in isolated systems.

It turns out the same is true for the strong forces, a fact that is established experimentally. The weak forces do, however, violate time-reversal invariance (or at least CP-invariance) at a small level. We shall say more about such violations later in these notes.

3. Time Reversal and the Schrödinger Equation

Let us consider the quantum analog of Eq. (1), that is, the motion of a charged particle in a given electric field. The Schrödinger equation in this case is

$$i\hbar\frac{\partial\psi(\mathbf{x},t)}{\partial t} = \Big[-\frac{\hbar^2}{2m}\nabla^2 + q\Phi(\mathbf{x})\Big]\psi(\mathbf{x},t). \tag{5}$$

Suppose $\psi(\mathbf{x},t)$ is a solution of this equation. Following what we did in the classical case, we ask if $\psi(\mathbf{x},-t)$ is also a solution. The answer is no, for unlike the classical equations of motion (1), the

Schrödinger equation is first order in time, so the left hand side changes sign under $t \to -t$, while the right hand side does not. If, however, we take the complex conjugate of Eq. (5), then we see that $\psi^*(\mathbf{x}, -t)$ is a solution of the Schrödinger equation, since the complex conjugation changes the sign of i on the left hand side, which compensates for the change in sign from $t \to -t$.

Altogether, we see that if we define the time-reversed motion in quantum mechanics by the rule

$$\psi_r(\mathbf{x}, t) = \psi^*(\mathbf{x}, -t), \tag{6}$$

where the r-subscript means "reversed," then charged particle motion in a static electric field is time-reversal invariant. We can see already from this example that time reversal in quantum mechanics is represented by an antilinear operator, since a linear operator is unable to map a wave function into its complex conjugate.

Similarly, the quantum analog of Eq. (2) is the Schrödinger equation for a particle in a magnetic field,

$$i\hbar \frac{\partial \psi(\mathbf{x}, t)}{\partial t} = \frac{1}{2m}\Big[-i\hbar\nabla - \frac{q}{c}\mathbf{A}(\mathbf{x})\Big]^2 \psi(\mathbf{x}, t). \tag{7}$$

In this case if $\psi(\mathbf{x}, t)$ is a solution, it does not follow that $\psi^*(\mathbf{x}, -t)$ is a solution, because of the terms that are linear in $\mathbf{A}$. These terms are purely imaginary, and change sign when we complex conjugate the Schrödinger equation. But $\psi^*(\mathbf{x}, -t)$ is a solution in the reversed magnetic field, that is, after the replacement $\mathbf{A} \to -\mathbf{A}$. This is just as in the classical case.

4. The Time-Reversal Operator Θ

The definition (6) of the time-reversed wave function applies to a spinless particle moving in three-dimensions. We shall be interested in generalizing it to other systems, such as multiparticle systems with spin, as well as preparing the ground for generalizations to relativistic systems and quantum fields. If $|\psi(t)\rangle$ is a time-dependent state vector of a system that satisfies the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle, \tag{8}$$

then on analogy with Eq. (6) we shall write the time-reversed state as

$$|\psi_r(t)\rangle = \Theta|\psi(-t)\rangle, \tag{9}$$

where Θ, the *time-reversal operator*, is to be defined for different systems based on certain postulates that we shall require of it. The operator Θ by itself does not involve time, as we see from Eq. (6), where it has the effect of complex conjugating the wave function; rather, it is a mapping that takes kets into other kets. In particular, setting $t = 0$ in Eq. (9),

$$|\psi_r(0)\rangle = \Theta|\psi(0)\rangle, \tag{10}$$

we see that Θ maps the initial conditions of the original motion into the initial conditions of the time-reversed motion.

We obtain a set of postulates for Θ as follows. First, since probabilities should be conserved under time reversal, we require

$$\Theta^\dagger \Theta = 1. \tag{11}$$

Next, in classical mechanics, the initial conditions of a motion $\mathbf{x}(t)$ transform under time reversal according to $(\mathbf{x}_0, \mathbf{p}_0) \to (\mathbf{x}_0, -\mathbf{p}_0)$, so we postulate that the time-reversal operator in quantum mechanics should satisfy the conjugation relations,

$$\Theta \mathbf{x} \Theta^\dagger = \mathbf{x}, \qquad \Theta \mathbf{p} \Theta^\dagger = -\mathbf{p}. \tag{12}$$

This should hold in systems where the operators $\mathbf{x}$ and $\mathbf{p}$ are meaningful. In such systems, these requirements imply

$$\Theta \mathbf{L} \Theta^\dagger = -\mathbf{L}, \tag{13}$$

where $\mathbf{L} = \mathbf{x} \times \mathbf{p}$ is the orbital angular momentum. As for spin angular momentum, we shall postulate that it transform in the same way as orbital angular momentum,

$$\Theta \mathbf{S} \Theta^\dagger = -\mathbf{S}. \tag{14}$$

This is plausible in view of a simple classical model of a spin, in which a particle like an electron is seen as a small charged sphere spinning on its axis. The rotation produces both an angular momentum and a magnetic moment. This model has flaws and cannot be taken very seriously, but at least it does indicate that if we reverse the motion of the charges on the sphere, both the angular momentum and the magnetic moment should reverse. Accepting both Eqs. (13) and (14), we see that we should have

$$\Theta \mathbf{J} \Theta^\dagger = -\mathbf{J}, \tag{15}$$

for all types of angular momentum.

5. Θ Cannot Be Unitary

It turns out that the conjugation relations (12) cannot be satisfied by any unitary operator. For if we take the canonical commutation relations,

$$[x_i, p_j] = i\hbar\, \delta_{ij}, \tag{16}$$

and conjugate with Θ, we find

$$\Theta [x_i, p_j] \Theta^\dagger = -[x_i, p_j] = -i\hbar\, \delta_{ij} = \Theta (i\hbar\, \delta_{ij}) \Theta^\dagger. \tag{17}$$

The quantity $i\hbar\, \delta_{ij}$ is just a number, so if Θ is unitary it can be brought through to cancel $\Theta^\dagger$, and we obtain a contradiction. Thus, we are forced to conclude that the time-reversal operator Θ must be antilinear, so that the imaginary unit i on the right-hand side of Eq. (17) will change into $-i$ when Θ is pulled through it.

6. Wigner's Theorem

A famous theorem proved by Wigner says that if we have a mapping of a ket space onto itself, taking, say, kets $|\psi\rangle$ and $|\phi\rangle$ into kets $|\psi'\rangle$ and $|\phi'\rangle$, such that the absolute values of all scalar products are preserved, that is, such that

$$|\langle\psi|\phi\rangle| = |\langle\psi'|\phi'\rangle| \tag{18}$$

for all $|\psi\rangle$ and $|\phi\rangle$, then, to within inessential phase factors, the mapping must be either a linear unitary operator or an antilinear unitary operator. The reason Wigner does not demand that the scalar products themselves be preserved (only their absolute values) is that the only quantities that are physically measurable are absolute squares of scalar products. These are the probabilities that are experimentally measurable. This theorem is discussed in more detail by Messiah, *Quantum Mechanics*, in which a proof is given. (See also Steven Weinberg, *The Quantum Theory of Fields I.*) Its relevance for the discussion of symmetries in quantum mechanics is that a symmetry operation must preserve the probabilities of all experimental outcomes, and thus all symmetries must be implemented either by unitary or antiunitary operators. In fact, all symmetries except time reversal (translations, proper rotations, parity, and others as well) are implemented by unitary operators. Time reversal, however, requires antiunitary operators.

7. Properties of Antilinear Operators

Since we have not encountered antilinear operators before, we now make a digression to discuss their mathematical properties. We let $\mathcal{E}$ be the ket space of some quantum mechanical system. In the following general discussion we denote linear operators by L, L_1, etc., and antilinear operators by A, A_1, etc. Both linear and antilinear operators are mappings of the ket space onto itself,

$$\begin{aligned} L &: \mathcal{E} \to \mathcal{E}, \\ A &: \mathcal{E} \to \mathcal{E}, \end{aligned} \tag{19}$$

but they have different distributive properties when acting on linear combinations of kets:

$$L\big(c_1|\psi_1\rangle + c_2|\psi_2\rangle\big) = c_1\, L|\psi_1\rangle + c_2\, L|\psi_2\rangle \tag{20a}$$

$$A\big(c_1|\psi_1\rangle + c_2|\psi_2\rangle\big) = c_1^*\, A|\psi_1\rangle + c_2^*\, A|\psi_2\rangle \tag{20b}$$

(see Eqs. (1.34)). In particular, an antilinear operator does not commute with a constant, when the latter is regarded as a multiplicative operator in its own right. Rather, we have

$$\boxed{Ac = c^* A.} \tag{21}$$

It follows from these definitions that the product of two antilinear operators is linear, and the product of a linear with an antilinear operator is antilinear. More generally, a product of operators is either linear or antilinear, depending on whether the number of antilinear factors is even or odd, respectively.

We now have to rethink the entire Dirac bra-ket formalism, to incorporate antilinear operators. To begin, we define the action of antilinear operators on bras. We recall that a bra, by definition, is a complex-valued, linear operator on kets, that is, a mapping,

$$\text{bra} : \mathcal{E} \to \mathbb{C}, \tag{22}$$

and that the value of a bra acting on a ket is just the usual scalar product. Thus, if $\langle\phi|$ is a bra, then we have

$$\big(\langle\phi|\big)\big(|\psi\rangle\big) = \langle\phi|\psi\rangle. \tag{23}$$

We now suppose that an antilinear operator A is given, that is, its action on kets is known, and we wish to define its action on bras. For example, if $\langle\phi|$ is a bra, we wish to define $\langle\phi|A$. In the case of linear operators, the definition was

$$\big(\langle\phi|L\big)\big(|\psi\rangle\big) = \big(\langle\phi|\big)\big(L|\psi\rangle\big). \tag{24}$$

Since the positioning of the parentheses is irrelevant, it is customary to drop them, and to write simply $\langle\phi|L|\psi\rangle$. In other words, we can think of L as acting either to the right or to the left. However, the analogous definition for antilinear operators does not work, for if we try to write

$$\big(\langle\phi|A\big)\big(|\psi\rangle\big) = \big(\langle\phi|\big)\big(A|\psi\rangle\big), \tag{25}$$

then $\langle\phi|A$ is indeed a complex-valued operator acting on kets, but it is an antilinear operator, not a linear one. Bras are supposed to be linear operators. Therefore we introduce a complex conjugation to make $\langle\phi|A$ a linear operator on kets, that is, we set

$$\boxed{\big(\langle\phi|A\big)|\psi\rangle = \big[\langle\phi|\big(A|\psi\rangle\big)\big]^*.} \tag{26}$$

This rule is easiest to remember in words: we say that in the case of an antilinear operator, it *does* matter whether the operator acts to the right or to the left in a matrix element, and if we change the direction in which the operator acts, we must complex conjugate the matrix element. In the case of antilinear operators, parentheses are necessary to indicate which direction the operator acts. The parentheses are awkward, and the fact is that Dirac's bra-ket notation is not as convenient for antilinear operators as it is for linear ones.

Next we consider the definition of the Hermitian conjugate. We recall that in the case of linear operators, the Hermitian conjugate is defined by

$$L^\dagger|\psi\rangle = \big(\langle\psi|L\big)^\dagger, \tag{27}$$

for all kets $|\psi\rangle$, or equivalently by

$$\langle\phi|L^\dagger|\psi\rangle = \langle\psi|L|\phi\rangle^*, \tag{28}$$

for all kets $|\psi\rangle$ and $|\phi\rangle$. Here the linear operator L is assumed given, and we are defining the new linear operator $L^\dagger$. The definition (27) also works for antilinear operators, that is, we set

$$A^\dagger|\psi\rangle = \big(\langle\psi|A\big)^\dagger. \tag{29}$$

We note that by this definition, $A^\dagger$ is an antilinear operator if A is antilinear. Now, however, when we try to write the analog of (28), we must be careful about the parentheses. Thus, we have

$$\langle\phi|\big(A^\dagger|\psi\rangle\big) = \big[\big(\langle\psi|A\big)|\phi\rangle\big]^*. \tag{30}$$

This rule is also easiest to remember in words. It is really a reconsideration of the rule stated in Sec. 1.13, that the Hermitian conjugate of any product of complex numbers, kets, bras and operators is obtained by reversing the order and taking the Hermitian conjugate of all factors. This rule remains true when antilinear operators are in the mix, but the parentheses indicating the direction in which the antilinear operator acts must also be reversed at the same time that A is changed into $A^\dagger$ or vice versa. That is, the direction in which the antilinear operator acts is reversed.

The boxed equations (21) and (26) summarize the principal rules for antilinear operators that differ from those of linear operators.

8. Antiunitary Operators

We wrote down Eq. (11) thinking that it would require probabilities to be preserved under Θ. This would certainly be true if Θ were unitary, but since we now know Θ must be antilinear, we should think about probability conservation under antilinear transformations.

We define an *antiunitary* operator A as an antilinear operator that satisfies

$$AA^\dagger = A^\dagger A = 1. \tag{31}$$

We note that the product $AA^\dagger$ or $A^\dagger A$ is a linear operator, so this definition is meaningful. Just like unitary operators, antiunitary operators preserve the absolute values of scalar products, as indicated by Wigner's theorem. To see this, we let $|\psi\rangle$ and $|\phi\rangle$ be arbitrary kets, and we set $|\psi'\rangle = A|\psi\rangle$, $|\phi'\rangle = A|\phi\rangle$, where A is antiunitary. Then we have

$$\langle\phi'|\psi'\rangle = \big(\langle\phi|A^\dagger\big)\big(A|\psi\rangle\big) = \big[\langle\phi|\big(A^\dagger A|\psi\rangle\big)\big]^* = \langle\phi|\psi\rangle^*, \tag{32}$$

where we reverse the direction of $A^\dagger$ in the second equality and use $A^\dagger A = 1$ in the third. Antiunitary operators take scalar products into their complex conjugates, and Eq. (18) is satisfied. Thus, we were correct in writing down Eq. (11) for probability conservation under time reversal.

9. The LK Decomposition

Given an antilinear operator A of interest, it is often convenient to factor it into the form

$$A = LK, \tag{33}$$

where L is a linear operator and K is a particular antilinear operator chosen for its simplicity. The idea is that K takes care of the antilinearity of A, while L takes care of the rest. The choices made for K are usually of the following type.

Let Q stand for a complete set of commuting observables (a single symbol Q for all operators in the set). Let n be the collective set of quantum numbers corresponding to Q, so that the basis kets in this representation are $|n\rangle$. The index n can include continuous quantum numbers as well as discrete ones. Then we define a particular antilinear operator K_Q by requiring, first, that K_Q be antilinear, and second, that

$$K_Q|n\rangle = |n\rangle. \tag{34}$$

Notice that the definition of K_Q depends not only on the operators Q that make up the representation, but also the phase conventions for the eigenkets $|n\rangle$. If K_Q were a linear operator, Eq. (34) would imply $K_Q = 1$; but since K_Q is antilinear, the equation $K_Q = 1$ is not only not true, it is meaningless, since it equates an antilinear operator to a linear one. But Eq. (34) does completely specify K_Q, for if $|\psi\rangle$ is an arbitrary ket, expanded according to

$$|\psi\rangle = \sum_n c_n|n\rangle, \tag{35}$$

then

$$K_Q|\psi\rangle = \sum_n c_n^*|n\rangle, \tag{36}$$

where we use Eqs. (20b) and (34). Thus, the action of K_Q on an arbitrary ket is known. The effect of K_Q is to bring about a complex conjugation of the expansion coefficients in the Q representation. These expansion coefficients are the same as the wave function in the Q representation; thus, in wave function language in the Q representation, K_Q just maps the wave function into its complex conjugate.

Consider, for example, the ket space for a spinless particle in three dimensions. Here we can work in the position representation, in which $Q = \mathbf{x}$ and in which the basis kets are $|\mathbf{x}\rangle$. Then we define the antilinear operator $K_\mathbf{x}$ by

$$K_\mathbf{x}|\mathbf{x}\rangle = |\mathbf{x}\rangle, \tag{37}$$

so that if $|\psi\rangle$ is an arbitrary ket and $\psi(\mathbf{x})$ its wave function, then

$$K_\mathbf{x}|\psi\rangle = K_\mathbf{x}\int d^3\mathbf{x}\,|\mathbf{x}\rangle\langle\mathbf{x}|\psi\rangle = K_\mathbf{x}\int d^3\mathbf{x}\,|\mathbf{x}\rangle\psi(\mathbf{x}) = \int d^3\mathbf{x}\,|\mathbf{x}\rangle\psi^*(\mathbf{x}). \tag{38}$$

Thus, $\psi(\mathbf{x})$ is mapped into $\psi(\mathbf{x})^*$.

The operator K_Q looks simple in the Q-representation. It may of course be expressed in other representations, but then it no longer looks so simple. For example, $K_\mathbf{x}$ is not as simple in the momentum representation as in the configuration representation (an explicit expression for $K_\mathbf{x}$ in the momentum representation will be left as an exercise).

It follows from the definition (34) that K_Q satisfies

$$K_Q^2 = 1. \tag{39}$$

(Just multiply Eq. (34) by K_Q and note that K_Q^2 is a linear operator, so that Eq. (39) is meaningful.) The operator K_Q also satisfies

$$K_Q = K_Q^\dagger, \tag{40}$$

and is therefore antiunitary. To prove this, we let $K_Q^\dagger$ act on a basis ket and then insert a resolution of the identity,

$$K_Q^\dagger|n\rangle = \sum_m |m\rangle\langle m|\big(K_Q^\dagger|n\rangle\big). \tag{41}$$

But

$$\langle m|\big(K_Q^\dagger|n\rangle\big) = \big[\big(\langle m|K_Q^\dagger\big)|n\rangle\big]^* = \langle n|\big(K_Q|m\rangle\big) = \langle n|m\rangle = \delta_{nm}, \tag{42}$$

where in the first step we reverse the direction in which $K_Q^\dagger$ acts, in the second reverse the order and conjugate all the factors to get the complex conjugate, and in the last step we use $K_Q|m\rangle = |m\rangle$. Altogether, this gives

$$K_Q^\dagger|n\rangle = |n\rangle. \tag{43}$$

But since $K_Q^\dagger$ has the same effect on the basis kets as K, and since both are antilinear, they must be equal antilinear operators, $K_Q = K_Q^\dagger$.

10. Time Reversal in Spinless Systems

In the case of a spinless particle, we decided in Sec. 3 that the time-reversal operator has the effect of complex conjugating the wave function $\psi(\mathbf{x})$. Let us reconsider the case of spinless systems from the standpoint of the axioms that Θ is supposed to satisfy, and see if we can rederive this result. Once we have done that, we will turn to particles with spin.

In the case of a spinless particle moving in three dimensions, the ket space is $\mathcal{E} = \text{span}\{|\mathbf{x}\rangle\}$ and the wave functions are $\psi(\mathbf{x})$. The time-reversal operator Θ, whatever it is, can be factored into LK, where $K = K_{\mathbf{x}}$ is the complex conjugation operator in the position representation, and where L is unitary. Let us begin by finding what K does to the operators $\mathbf{x}$ and $\mathbf{p}$ under conjugation. We recall from Sec. 9 that $K = K^\dagger = K^{-1}$. Tracking the effect of $K\mathbf{x}K^\dagger$ on a wave function, we have

$$\psi(\mathbf{x}) \xrightarrow{K^\dagger} \psi^*(\mathbf{x}) \xrightarrow{\mathbf{x}} \mathbf{x}\psi^*(\mathbf{x}) \xrightarrow{K} \mathbf{x}\psi(\mathbf{x}), \tag{44}$$

or,

$$K\mathbf{x}K^\dagger = \mathbf{x}. \tag{45}$$

Similarly, for the momentum operator we have

$$\psi(\mathbf{x}) \xrightarrow{K^\dagger} \psi^*(\mathbf{x}) \xrightarrow{\mathbf{p}} -i\hbar\nabla\psi^*(\mathbf{x}) \xrightarrow{K} +i\hbar\nabla\psi(\mathbf{x}), \tag{46}$$

or,

$$K\mathbf{p}K^\dagger = -\mathbf{p}. \tag{47}$$

These agree with the postulates (12), and produce the right transformation law (13) for orbital angular momentum. Therefore in the LK-decomposition of Θ we take $L = 1$ and define

$$\Theta = K_{\mathbf{x}}, \tag{48}$$

restoring the **x**-subscript to K for clarity. Time reversal is simply complex conjugation of the wave function in the **x**-representation for a spinless particle.

This result can be easily generalized to the case of any number of spinless particles in any number of dimensions. The time-reversal operator Θ is defined as the complex conjugation operation in the configuration representation, so that its action on wave functions is given by

$$\psi(\mathbf{x}_1, \ldots, \mathbf{x}_n) \xrightarrow{\Theta} \psi^*(\mathbf{x}_1, \ldots, \mathbf{x}_n). \tag{49}$$

This is a simple rule that covers many cases occurring in practice.

11. Time Reversal and Spin

Next we consider the spin degrees of freedom of a particle of spin s. For simplicity we will at first ignore the spatial degrees of freedom, so the $(2s+1)$-dimensional ket space is $\mathcal{E} = \text{span}\{|sm\rangle, m = -s, \ldots, s\}$. As usual, the basis kets are eigenstates of S_z. The postulated time-reversal operator must satisfy the conjugation relations,

$$\Theta \mathbf{S} \Theta^\dagger = -\mathbf{S}. \tag{50}$$

As we will show, this condition determines Θ to within a phase factor.

First we consider the operator S_z, which satisfies

$$\Theta S_z \Theta^\dagger = -S_z. \tag{51}$$

From this it easily follows that the ket $\Theta|sm\rangle$ is an eigenket of S_z with eigenvalue $-m\hbar$,

$$S_z \Theta |sm\rangle = -\Theta S_z |sm\rangle = -m\hbar \Theta |sm\rangle. \tag{52}$$

But since the eigenkets of S_z are nondegenerate, we must have

$$\Theta|sm\rangle = c_m |s, -m\rangle, \tag{53}$$

where c_m is a constant that presumably depends on m. In fact, if we square both sides of Eq. (53) and use the fact that Θ is antiunitary, we will that c_m is a phase factor. To find the m-dependence of c_m, we study the commutation relations of Θ with the raising and lowering operators. For example, for S_+, we have

$$\Theta S_+ \Theta^\dagger = \Theta(S_x + iS_y)\Theta^\dagger = -S_x + iS_y = -S_-, \tag{54}$$

where we use Eq. (15) and where a second sign reversal takes place in the S_y-term due to the imaginary unit i. There is a similar equation for S_-; we summarize them both by writing

$$\Theta S_\pm \Theta^\dagger = -S_\mp. \tag{55}$$

Now let us apply S_+ to the ket $\Theta|sm\rangle$, and use the commutation relations (54). We find

$$\begin{aligned} S_+ \Theta |sm\rangle &= -\Theta S_- |sm\rangle = -\hbar\sqrt{(s+m)(s-m+1)}\,\Theta|s, m-1\rangle \\ &= -\hbar\sqrt{(s+m)(s-m+1)}\,c_{m-1}\,|s, -m+1\rangle = c_m S_+ |s, -m\rangle \\ &= \hbar\sqrt{(s+m)(s-m+1)}\,c_m |s, -m+1\rangle, \end{aligned} \tag{56}$$

or, on cancelling the square roots,

$$\Theta|s, m-1\rangle = -c_m|s, -m+1\rangle. \tag{57}$$

But by Eq. (53), this must also equal $c_{m-1}|s, -m+1\rangle$. Thus we find

$$c_{m-1} = -c_m, \tag{58}$$

so c_m changes by a sign every time m increments or decrements by 1.

We can summarize this result by writing

$$\Theta|sm\rangle = \eta(-1)^{s-m}|s, -m\rangle, \tag{59}$$

where η is a phase that is independent of m. We have written the exponent of -1 in Eq. (59) as $s-m$ because if s is half-integer then so is m, but $s-m$ is always an integer. The main point is that the coefficient alternates in sign as m increases or decreases by unit steps. We have proven our earlier assertion, that Θ is determined to within a phase by the conjugation relation (50).

Since the phase η is independent of m, it can be absorbed into the definition of Θ, by writing, say, $\Theta = \eta\Theta_1$, where Θ_1 is a new time-reversal operator. Such an overall phase has no effect on the desired commutation relations (15), as one can easily verify, and in fact is devoid of physical significance. A common choice in practice is to choose $\eta = i^{2s}$ so that

$$\Theta|sm\rangle = i^{2m}|s, -m\rangle. \tag{60}$$

This phase convention is nice because it is the obvious generalization of Eq. (96), which applies to orbital angular momentum.

12. Another Approach to Time Reversal and Spin

Another approach to finding a time-reversal operator that satisfies Eq. (50) is to attempt an LK-decomposition of Θ. Since the usual basis is the S_z basis, we examine the antiunitary operator K_{S_z}, for which we simply write K in the following. We begin by conjugating $\mathbf{S}$ by K, finding,

$$K \begin{pmatrix} S_x \\ S_y \\ S_z \end{pmatrix} K^\dagger = \begin{pmatrix} S_x \\ -S_y \\ S_z \end{pmatrix}. \tag{61}$$

The operators S_x and S_z did not change sign because their matrices in the standard angular momentum basis are real (see Sec. 13.5), while S_y does change sign since its matrix for S_y is purely imaginary. We see that Θ is not equal to K, because the latter only changes the sign of one of the components of spin.

Nevertheless, the result can be fixed up with a unitary operator. Let U_0 be the spin rotation by angle π about the y-axis,

$$U_0 = U(\hat{\mathbf{y}}, \pi) = e^{-i\pi S_y/\hbar}. \tag{62}$$

Such a rotation leaves the y-component of a vector invariant, while flipping the signs of the x- and z-components. This is really an application of the adjoint formula (13.89). Thus we have

$$U_0 \begin{pmatrix} S_x \\ -S_y \\ S_z \end{pmatrix} U_0^\dagger = \begin{pmatrix} -S_x \\ -S_y \\ -S_z \end{pmatrix}. \tag{63}$$

Altogether, we can satisfy the requirement (50) by defining

$$\Theta = e^{-i\pi S_y/\hbar} K = K e^{-i\pi S_y/\hbar}, \tag{64}$$

where $e^{-i\pi S_y/\hbar}$ and K commute because the matrix for $e^{-i\pi S_y/\hbar}$ in the standard basis is real.

Comparing the approach of this section to that in Sec. 11, we see that the operator Θ defined by Eq. (64) must be the same as that defined in Eq. (59), for some choice of η. In fact, with some additional trouble one can show that $\eta = 1$ works (although this is not a very important fact, since η is nonphysical anyway).

13. Spatial and Spin Degrees of Freedom

Let us now include the spatial degrees of freedom, and consider the case of a spinning particle in three-dimensional space. The ket space is $\mathcal{E} = \text{span}\{|\mathbf{x}, m\rangle\}$, following the notation of Eq. (18.10), and the wave functions are $\psi_m(\mathbf{x})$, as in Eq. (18.12). In this case, the obvious definition of the time-reversal operator is the product of the two operators introduced above ($\Theta = K_{\mathbf{x}}$ for the spatial degrees of freedom, and $\Theta = K_{S_z} U_0$ for the spin degrees of freedom). That is, we take

$$\Theta = K e^{-i\pi S_y/\hbar}, \tag{65}$$

where now $K = K_{\mathbf{x}S_z}$ is the complex conjugation operator in the $|\mathbf{x}m\rangle$ basis, and where the rotation operator only rotates the spin (not the spatial degrees of freedom). This is the same as Eq. (64), except for a reinterpretation of the operator K.

For example, in the case of a spin-$\frac{1}{2}$ particle, we have $S_y = (\hbar/2)\sigma_y$, so

$$e^{-i\pi S_y/\hbar} = e^{-i(\pi/2)\sigma_y} = \cos(\pi/2) - i\sigma_y \sin(\pi/2) = -i\sigma_y = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{66}$$

so a two component spinor as in Eq. (18.14) transforms under time reversal according to

$$\begin{pmatrix} \psi_+(\mathbf{x}) \\ \psi_-(\mathbf{x}) \end{pmatrix} \xrightarrow{\Theta} \begin{pmatrix} -\psi_-^*(\mathbf{x}) \\ \psi_+^*(\mathbf{x}) \end{pmatrix}. \tag{67}$$

More generally, for any s the wave function $\psi_m(\mathbf{x})$ transforms under time-reversal according to

$$\psi_m(\mathbf{x}) \xrightarrow{\Theta} \sum_{m'} d^s_{mm'}(\pi)\, \psi^*_{m'}(\mathbf{x}), \tag{68}$$

where we use the reduced rotation matrices defined by Eq. (13.66). Compare this to Eq. (49) for a spinless particle.

Finally, to implement time reversal on a system of many spinning particles, for which the ket space is the tensor product of the ket spaces for the individual particles (both orbital and spin), we simply take Θ to be a product of operators of the form (65), one for each particle. The result is the formula (65) all over again, with K now interpreted as the complex conjugation in the tensor product basis,

$$|\mathbf{x}_1 m_{s1}\rangle \otimes \ldots \otimes |\mathbf{x}_n m_{sn}\rangle, \tag{69}$$

where n is the number of particles, and where the spin rotation is the product of the individual spin rotations,

$$e^{-i\pi S_y/\hbar} = e^{-i\pi S_{1y}/\hbar} \ldots e^{-i\pi S_{ny}/\hbar}. \tag{70}$$

Here S_y is the y-component of the total spin of the system,

$$S_y = S_{1y} + \ldots + S_{ny}. \tag{71}$$

It may seem strange that a rotation about the y-axis should appear in Eq. (64) or (65), since the time-reversal operator should not favor any particular direction in space. Actually, the time-reversal operator does not favor any particular direction, it is just the decomposition into the indicated unitary operator $e^{-i\pi S_y/\hbar}$ and the antiunitary complex conjugation operator K which has treated the three directions in an asymmetrical manner. That is, the complex conjugation antiunitary operator K is tied to the S_z representation and the standard phase conventions used in angular momentum theory; since K does not treat the three directions symmetrically, the remaining unitary operator $e^{-i\pi S_y/\hbar}$ cannot either. However, their product does.

Often in multiparticle systems we are interested to combine spin states of individual particles together to form eigenstates of total S^2 and S_z. This gives us a complete set of commuting observables that include the total S^2 and S_z, rather than the S_z's of individual particles as in Eq. (69). But since the Clebsch-Gordan coefficients are real, the complex conjugation operator K in the new basis is the same as in the old, and Eq. (65) still holds.

14. Examples of Hamiltonians

Let us now consider some examples of Hamiltonians that either do or do not commute with time reversal.

First, any kinetic-plus-potential Hamiltonian in three dimensions,

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}), \tag{72}$$

commutes with time reversal, because the kinetic energy is even in the momentum and the position vector is invariant. We emphasize that the potential need not be a central force potential. The motion of a charged particle in a given electrostatic field, discussed in Sec. 3, is an example of such a system. More generally, kinetic-plus-potential Hamiltonians for any number of particles in any number of dimensions commute with time reversal.

The easiest way to break time-reversal invariance is to introduce a external magnetic field. Then the kinetic energy becomes

$$\frac{1}{2m}\Big[\mathbf{p} - \frac{q}{c}\mathbf{A}(\mathbf{x})\Big]^2, \tag{73}$$

which does not commute with time reversal because $\mathbf{p}$ changes sign under conjugation by Θ, while $\mathbf{A}(\mathbf{x})$ does not.

As mentioned previously, however, if the magnetic field is internally generated, then time-reversal invariance is not broken. For example, in an atom the spin-orbit interaction is the magnetic interaction between the spin of the electron and the magnetic field produced by the motion of the nucleus around the electron, as seen in the electron rest frame. It is described by a term in the Hamiltonian of the form

$$f(r)\mathbf{L}\cdot\mathbf{S}, \tag{74}$$

which according to Eq. (15) is invariant under time reversal (both $\mathbf{L}$ and $\mathbf{S}$ change sign). Similarly, spin-spin interactions such as hyperfine effects in an atom, which are proportional to $\mathbf{I}\cdot\mathbf{S}$ ($\mathbf{I}$ is the nuclear spin, $\mathbf{S}$ the electron spin) are invariant under time reversal. These are all examples of the invariance of electromagnetic effects under time reversal.

Are there any examples of interactions that break time-reversal invariance but that only involve internally generated fields? Yes, suppose for example that the nucleus has an electric dipole moment, call it $\boldsymbol{\mu}_e$. By the Wigner-Eckart theorem, this vector is proportional to the spin $\mathbf{S}$, so we get a term in the Hamiltonian,

$$H_{\rm int} = -\boldsymbol{\mu}_e\cdot\mathbf{E} = k\mathbf{S}\cdot\mathbf{E}, \tag{75}$$

where k is a constant. This term is odd under time reversal, and so breaks time-reversal invariance. Time-reversal invariance is known to be respected to a very high degree of approximation, so terms of the form (75), if they are present in ordinary atoms, are very small.

There is currently considerable experimental interest in the detection of electric dipole moments of nuclei and particles such as the proton, neutron and electron, because the existence of such moments would imply a violation of time-reversal invariance and would give information about physics beyond the standard model.

15. The Time-Reversed Motion

We began our discussion of time reversal by working at the classical level and asking under what conditions the time-reversed motion is also a solution of the equations of motion. Let us now address the same question in quantum mechanics, assuming for simplicity that the Hamiltonian is time-independent. We assume $|\psi(t)\rangle$ satisfies the Schrödinger equation, and we define the time-reversed state $|\psi_r(t)\rangle$ by Eq. (9). Does it also satisfy the Schrödinger equation?

To begin we just compute the time derivative of the time-reversed state,

$$i\hbar\frac{\partial}{\partial t}|\psi_r(t)\rangle = i\hbar\frac{\partial}{\partial t}\Theta|\psi(-t)\rangle = \Theta\Big[-i\hbar\frac{\partial}{\partial t}|\psi(-t)\rangle\Big], \tag{76}$$

where Θ changes i to $-i$ when we pull it to the left. Θ commutes with $\partial/\partial t$ because Θ is a Hilbert space operator and $\partial/\partial t$ is just a derivative with respect to a parameter (time) that the state depends on. (If this is not clear, write out the time derivative as a limit in which $\Delta t \to 0$.) Now let us write $\tau = -t$, so that

$$|\psi_r(t)\rangle = \Theta|\psi(\tau)\rangle, \tag{77}$$

and so that Eq. (76) becomes

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\psi_r(t)\rangle &= \Theta\Big[i\hbar\frac{\partial}{\partial\tau}|\psi(\tau)\rangle\Big] = \Theta H|\psi(\tau)\rangle \\ &= (\Theta H\Theta^\dagger)\Theta|\psi(\tau)\rangle = (\Theta H\Theta^\dagger)|\psi_r(t)\rangle. \end{aligned} \tag{78}$$

We see that the time-reversed state satisfies the Schrödinger equation with the time-reversed Hamiltonian, by which we mean $\Theta H\Theta^\dagger$. We also see that the time-reversed motion satisfies the original Schrödinger equation if the Hamiltonian is invariant under time-reversal, that is, if

$$\Theta H\Theta^\dagger = H \qquad \text{or} \qquad [\Theta, H] = 0. \tag{79}$$

Another approach to the same result is to write a solution of the Schrödinger equation as

$$|\psi(t)\rangle = \exp(-itH/\hbar)|\psi(0)\rangle, \tag{80}$$

to which we apply Θ,

$$\Theta|\psi(t)\rangle = \Theta\exp(-itH/\hbar)\Theta^\dagger\Theta|\psi(0)\rangle. \tag{81}$$

The conjugated time-evolution operator can be written,

$$\Theta\exp(-itH/\hbar)\Theta^\dagger = \exp\Big[\Theta(-itH/\hbar)\Theta^\dagger\Big] = \exp(+itH/\hbar), \tag{82}$$

where we have used the antilinearity of Θ and Eq. (79). Then replacing t by $-t$ and using the definition (9) of $|\psi_r(t)\rangle$, we have

$$|\psi_r(t)\rangle = \exp(-itH/\hbar)|\psi_r(0)\rangle. \tag{83}$$

We see that $|\psi_r(t)\rangle$ is also a solution of the time-dependent Schrödinger equation.

16. Time Reversal and Energy Eigenstates

We have seen the effect of time reversal on the solutions of the time-dependent Schrödinger equation. Let us now examine the effect on the energy eigenstates.

Let $|\psi\rangle$ be an energy eigenstate for any system,

$$H|\psi\rangle = E|\psi\rangle, \tag{84}$$

and suppose that $[\Theta, H] = 0$. Then

$$H(\Theta|\psi\rangle) = \Theta H|\psi\rangle = E(\Theta|\psi\rangle), \tag{85}$$

that is, Θ maps eigenstates of H into eigenstates of H with the same energy. If the original eigenstate is nondegenerate, then $\Theta|\psi\rangle$ must be proportional to $|\psi\rangle$,

$$\Theta|\psi\rangle = c|\psi\rangle, \tag{86}$$

where c is a constant. In fact, the constant is a phase factor, as we see by squaring both sides,

$$(\langle\psi|\Theta^\dagger)(\Theta|\psi\rangle) = [\langle\psi|(\Theta^\dagger\Theta|\psi\rangle)]^* = \langle\psi|\psi\rangle^* = 1 = |c|^2, \tag{87}$$

where we reverse the direction of $\Theta^\dagger$ in the first step and use $\Theta^\dagger\Theta = 1$ in the second. Thus we can write

$$\Theta|\psi\rangle = e^{i\alpha}|\psi\rangle. \tag{88}$$

Now multiplying this by $e^{-i\alpha/2}$, we find

$$e^{-i\alpha/2}\Theta|\psi\rangle = \Theta e^{i\alpha/2}|\psi\rangle = e^{i\alpha/2}|\psi\rangle, \tag{89}$$

or, with $|\phi\rangle = e^{i\alpha/2}|\psi\rangle$,

$$\Theta|\phi\rangle = |\phi\rangle. \tag{90}$$

We see that nondegenerate energy eigenstates of a Hamiltonian that is time-reversal invariant can always be chosen (by changing the overall phase, if necessary), so that the eigenstate is invariant under time reversal.

17. Reality of Energy Eigenfunctions in Spinless Systems

In particular, for a spinless system in three dimensions with a kinetic-plus-potential Hamiltonian, Eq. (88) implies that nondegenerate energy eigenfunctions $\psi(\mathbf{x})$ satisfy

$$\psi^*(\mathbf{x}) = e^{i\alpha}\,\psi(\mathbf{x}). \tag{91}$$

Now define a new wave function,

$$\phi(\mathbf{x}) = e^{i\alpha/2}\,\psi(\mathbf{x}), \tag{92}$$

so that

$$\phi(\mathbf{x}) = \phi^*(\mathbf{x}). \tag{93}$$

We see that a nondegenerate energy eigenfunction in a spinless kinetic-plus-potential system is always proportional to a real eigenfunction; the eigenfunction may be chosen to be real. We invoked the identical argument in Notes 6, when we showed that nondegenerate energy eigenfunctions in simple one-dimensional problems can always be chosen to be real.

In the case of degeneracies, the eigenfunctions are not necessarily proportional to real eigenfunctions, but real eigenfunctions can always be constructed out of linear combinations of the degenerate eigenfunctions. We state this fact without proof, but we offer some examples. First, a free particle in one dimension has the degenerate energy eigenfunctions, $e^{ipx/\hbar}$ and $e^{-ipx/\hbar}$, both of which are intrinsically complex; but real linear combinations are $\cos(px/\hbar)$ and $\sin(px/\hbar)$.

Similarly, a spinless particle moving in a central force field in three dimensions possesses the energy eigenfunctions,

$$\psi_{n\ell m}(\mathbf{x}) = R_{n\ell}(r)Y_{\ell m}(\theta,\phi), \tag{94}$$

which are degenerate since by the Wigner-Eckart theorem the energy $E_{n\ell}$ is independent of the magnetic quantum number m. The radial wave functions $R_{n\ell}$ can be chosen to be real, as we suppose, but the $Y_{\ell m}$'s are complex. However, in view of Eq. (15.57), we have

$$\psi^*_{n\ell m}(\mathbf{x}) = (-1)^m \psi_{n\ell,-m}(\mathbf{x}), \tag{95}$$

or, in ket language,

$$\Theta|n\ell m\rangle = (-1)^m |n\ell,-m\rangle. \tag{96}$$

In this case, real wave functions can be constructed out of linear combinations of the states $|n\ell m\rangle$ and $|n\ell,-m\rangle$. Sometimes it is convenient to ignore the radial variables and think of the Hilbert space of functions on the unit sphere, as we did in Notes 15; then we treat Θ as the complex conjugation operator acting on such functions, and we have

$$\Theta|\ell m\rangle = (-1)^m |\ell,-m\rangle, \tag{97}$$

instead of Eq. (96). This is a ket version of Eq. (15.57).

18. Kramers Degeneracy

Equation (65) allows us to compute the square of Θ, which is used in an important theorem to be discussed momentarily. We find

$$\Theta^2 = Ke^{-i\pi S_y/\hbar}Ke^{-i\pi S_y/\hbar} = e^{-2\pi i S_y/\hbar}, \tag{98}$$

where we commute K past the rotation and use $K^2 = 1$. The result is a total spin rotation of angle 2π about the y-axis. This rotation can be factored into a product of spin rotations, one for each particle, as in Eq. (70). For every boson, that is, for every particle with integer spin, the rotation by 2π is $+1$, while for every fermion, that is, for every particle of half-integer spin, the rotation by 2π is -1, because of the double-valued representation of the classical rotations for the case of half-integer angular momentum. Thus, the product (98) is $+1$ if the system contains an even number of fermions, and -1 if it contains an odd number.

This result has an application. Consider an arbitrarily complex system of possibly many spinning particles, in which the Hamiltonian is invariant under time reversal. One may think, for example, of the electronic motion in a solid or a molecule. There is no assumption that the system be invariant under rotations; this would not usually be the case, for example, in the electronic motion in the molecule. Suppose such a system has a nondegenerate energy eigenstate $|\psi\rangle$ with eigenvalue E, as in Sec. 16, so that $|\psi\rangle$ satisfies Eq. (88). Now multiplying that equation by Θ, we obtain

$$\Theta^2|\psi\rangle = \Theta e^{i\alpha}|\psi\rangle = e^{-i\alpha}\Theta|\psi\rangle = |\psi\rangle. \tag{99}$$

But if the system contains an odd number of fermions, then according to Eq. (98) we must have $\Theta^2 = -1$, which contradicts Eq. (99). Therefore the assumption of a nondegenerate energy level must be incorrect. We conclude that in time-reversal invariant systems with an odd number of fermions, the energy levels are always degenerate. This is called *Kramers degeneracy*. More generally, one can show that in such systems, the energy levels have a degeneracy that is even [see Prob. 5(c)]. Kramers degeneracy is lifted by any effect that breaks the time-reversal invariance, notably external magnetic fields.

One example of Kramers degeneracy has appeared already, in Prob. 20.3, in which it is asked to compute the energy levels of a nucleus of spin $\frac{3}{2}$ in an inhomogeneous electric field. The system consists of the nucleus, and it is not isolated, since it is interacting with the external electric field. But such interactions do not break time-reversal invariance, so the nuclear Hamiltonian (which we do not need to know explicitly) commutes with Θ. But since the system has half-integer spin, all eigenstates must be degenerate. In fact, the solution of the problem shows that the four energy eigenstates fall into two degenerate pairs.

Other examples of Kramers degeneracy will be pointed out as they occur in applications studied later in the course.

19. CPT Invariance and CP Violation

In relativistic quantum mechanics, it is believed that all physical fields must be Lorentz covariant and that interactions must be local. These hypotheses lead to the CPT theorem, which states that the product of charge conjugation (C), parity (P) and time reversal (T) is an exact symmetry of nature. We have discussed parity and time reversal but not charge conjugation, which can only be understood properly in a relativistic context. Nevertheless, the basic idea is that charge conjugation maps particles into their antiparticles, thereby changing their charge.

For a period of time after the discovery of parity violation in 1954, it was believed that although P was not a good symmetry of nature, at least the product CP would be a good symmetry. In 1964, however, an example of CP-violation was discovered in the decay of the neutral K-mesons. This research was carried out by Cronin and Fitch, who later received the Nobel prize for their work. If we accept the validity of the CPT theorem, CP-violation implies violation of time reversal. More recently there has been extensive experimental work on the decays of the B-mesons, which also exhibit CP-violation. This work has been carried out in the BaBar experiments at SLAC, and has resulted in a better understanding of CP-violation in the standard model.

There is much current speculation that the observed asymmetry between matter and antimatter in the universe is due to time-reversal violating effects shortly after the big bang. The idea is that matter and antimatter were created in almost equal measure, the small difference being due to T violation. Later the matter and antimatter mostly annihilated, leaving behind only a small residue of matter, which however makes up the matter we see in the universe today, including ourselves. These are just speculations, but they show the importance of time reversal and other fundamental

symmetries in current physical thinking.

Problems

1. This is a variation on Sakurai *Modern Quantum Mechanics*, problem 4.10, but improved.

(a) Let $U(\mathsf{R})$ be a rotation operator on the state space of any system (however complex). Show that $[\Theta, U(\mathsf{R})] = 0$ for all R.

(b) Denote the basis states in a single irreducible subspace under rotations of any system by $|jm\rangle$. By considering $\Theta U(\mathsf{R})|jm\rangle$, show that

$$[D^j_{m'm}(\mathsf{R})]^* = (-1)^{m-m'} D^j_{-m',-m}(\mathsf{R}). \tag{100}$$

(c) Show that if T^k_q is an irreducible tensor operator, then so is

$$S^k_q = (-1)^q \big(T^k_{-q}\big)^\dagger. \tag{101}$$

Operator S^k is regarded as the Hermitian conjugate of T^k.

2. This is Sakurai, *Modern Quantum Mechanics*, revised edition, problem 11, p. 283, or Sakurai and Napolitano (second edition), problem 11, p. 301.

Suppose a spinless particle is bound to a fixed center by a potential $V(\mathbf{x})$ so asymmetric that no energy level is degenerate. Using time-reversal invariance, prove

$$\langle \mathbf{L} \rangle = 0 \tag{102}$$

for any energy eigenstate. (This is known as *quenching* of orbital angular momentum.) If the wave function of such a nondegenerate energy eigenstate is expanded as

$$\sum_{\ell m} F_{\ell m}(r) Y_{\ell m}(\theta, \phi), \tag{103}$$

what kind of phase restrictions do we obtain on $F_{\ell m}(r)$?

3. Time-reversal symmetry is important in many applications in atomic, molecular, nuclear, and condensed matter physics. The notes above have given just the basic facts. In this this problem and the next several problems we will explore various aspects of time-reversal symmetry more deeply and rigorously. In all cases we assume that we have an n-particle system of some kind, on which Θ is defined as in the notes above. In particular, Θ is an antiunitary operator ($\Theta^\dagger\Theta = \Theta\Theta^\dagger = 1$) that satisfies

$$\Theta^2 = s, \tag{104}$$

where $s = +1$ for a system with an even number of fermions, and $s = -1$ for an odd number.

It was noted in Eq. (32) that antiunitary operators map scalar products into their complex conjugates, so, in particular this applies to time reversal.

From this it follows that if we have an orthonormal set of vectors, $\{|n\rangle, n = 1, 2, \ldots\}$ with $\langle n|m\rangle = \delta_{nm}$, and if we define

$$|n'\rangle = \Theta|n\rangle, \tag{105}$$

then

$$\langle n'|m'\rangle = \delta^*_{nm} = \delta_{nm}. \tag{106}$$

The set $\{|n\rangle\}$ need not span the whole space (it need not be a basis). Notice that this does not say whether the new basis vectors $|n'\rangle$ are linearly independent of the old ones $|n\rangle$. In summary, we can say that time reversal maps orthonormal sets into orthonormal sets.

Let $\mathcal{E}$ be the Hilbert space for a system, and let $\mathcal{S} \subset \mathcal{E}$ be a subspace. We say that $\mathcal{S}$ is *invariant* under Θ if for every $|\psi\rangle \in \mathcal{S}$, $\Theta|\psi\rangle \in \mathcal{S}$ (that is, Θ maps $\mathcal{S}$ into itself).

(a) Let H be a Hamiltonian that is invariant under time reversal, that is, $\Theta^\dagger H \Theta = H$. Show that the bound state energy eigenspaces of H are invariant under Θ.

In the following it is worthwhile keeping in mind two examples of subspaces invariant under Θ, namely, the whole space ($\mathcal{S} = \mathcal{E}$) and the energy eigenspaces of a Θ-invariant Hamiltonian.

(b) If $\mathcal{S}$ is invariant under Θ, show that it is also invariant under $\Theta^\dagger$. (Hint: Use $\Theta^2 = s$.)

(c) Let $\mathcal{S}$ be invariant under Θ, and let $\mathcal{A} \subset \mathcal{S}$ be a subspace of $\mathcal{S}$ that is also invariant under Θ. Let $\mathcal{B}$ be the subspace of $\mathcal{S}$ that is orthogonal to $\mathcal{A}$, so that

$$\mathcal{S} = \mathcal{A} \oplus \mathcal{B}. \tag{107}$$

Show that $\mathcal{B}$ is also invariant under Θ. Hint: a vector $|\psi\rangle \in \mathcal{S}$ lies in $\mathcal{B}$ if it is orthogonal to all vectors $|a\rangle \in \mathcal{A}$.

4. This problem continues Prob. 3. In this problem we consider the case of an even number of fermions, so that $s = +1$. As in Prob. 3, we let $\mathcal{S} \subset \mathcal{E}$ be a Θ-invariant subspace. We assume that $\mathcal{S}$ is not the trivial subspace $\{0\}$, that is, that $\dim \mathcal{S} \geq 1$.

(a) Show that $\mathcal{S}$ contains a one-dimensional, invariant subspace. Hint: Let $|\psi\rangle \in \mathcal{S}$ be any nonzero vector in $\mathcal{S}$, and consider $\Theta|\psi\rangle$. Either $|\psi\rangle$ and $\Theta|\psi\rangle$ are linearly independent, or they are not. Consider the two cases.

(b) Assuming that $\mathcal{S}$ is n-dimensional where $n \geq 1$ is finite, show that

$$\mathcal{S} = \mathcal{S}_1 \oplus \ldots \oplus \mathcal{S}_n, \tag{108}$$

where each $\mathcal{S}_i$ is one-dimensional and the various $\mathcal{S}_i$ are orthogonal to one another.

If $\mathcal{S}$ is infinite-dimensional, we will assume that the decomposition (108) is still valid, with $n \to \infty$. (It is obvious that we can keep on splitting off invariant, one-dimensional subspaces for as long as we want.)

(c) Show that a one-dimensional, Θ-invariant subspace possesses a basis that is invariant under Θ, that is, a vector $|e\rangle$ such that $\langle e|e\rangle = 1$ and $\Theta|e\rangle = |e\rangle$.

Thus, if we take $\mathcal{S} = \mathcal{E}$, in the case $s = +1$ we have shown that the Hilbert space possesses an orthonormal basis of Θ-invariant vectors; these might be used for diagonalizing a Hamiltonian. If we take $\mathcal{S}$ to be an energy eigenspace of a Θ-invariant Hamiltonian, then we have shown that a basis of orthonormal energy eigenvectors inside the eigenspace can be chosen to be invariant under Θ.

(d) If H is Θ-invariant, show that its matrix elements with respect to a Θ-invariant basis are real. In the usual applications this is obvious in the case of spinless particles, because the Schrödinger equation is real and Θ-invariant basis functions are real; but it is not so trivial when the particles have spin. Diagonalizing real matrices is easier than diagonalizing complex ones.

(e) If we have any two orthonormal bases $\{|e_n\rangle\}$ and $\{|f_n\rangle\}$ on a subspace $\mathcal{S}$, then these are related by an unitary transformation,

$$|f_n\rangle = \sum_m |e_m\rangle\, U_{mn}, \tag{109}$$

where

$$U_{mn} = \langle e_m|f_n\rangle. \tag{110}$$

The matrix U_{mn} is unitary because

$$(UU^\dagger)_{mn} = \sum_k U_{mk}U^*_{nk} = \sum_k \langle e_m|f_k\rangle\langle f_k|e_n\rangle = \langle e_m|e_n\rangle = \delta_{mn}. \tag{111}$$

Similarly we show that $(U^\dagger U)_{mn} = \delta_{mn}$.

Suppose $\mathcal{S}$ is Θ-invariant as are the bases $\{|e_n\rangle\}$ and $\{|f_n\rangle\}$. Show that the matrix U_{mn} defined by Eq. (110) is orthogonal (that is, the matrix elements are real). Conversely, show that if $\{|e_n\rangle\}$ is Θ-invariant and U is real orthogonal, then $\{|f_n\rangle\}$ is Θ-invariant. This is trivial, but taken together the statements show that the set of Θ-invariant bases on an N-dimensional, Θ-invariant space $\mathcal{S}$ can be placed in one-to-one correspondence with elements of the group $O(N)$.

We see that in the case $s = +1$, a basis can always be chosen such that the matrix elements of a Θ-invariant Hamiltonian are real; and such a matrix can be diagonalized by means of a real, orthogonal transformation, giving us energy eigenstates that are Θ-invariant. This fact reveals the computational advantages of time-reversal invariance.

5. The problem is a continuation of Probs. 3 and 4. It deals with the case of an odd number of fermions, for which $\Theta^2 = s = -1$.

(a) Let $\mathcal{S}$ be an invariant subspace, that is, invariant under Θ, with $\dim\mathcal{S} \geq 1$. Show tht $\mathcal{S}$ does not contain any one-dimensional, invariant subspace.

Thus, $\dim\mathcal{S} \geq 2$. If $\mathcal{S}$ is identified with an energy eigenspace, the conclusion is the same as that regarding Kramers degeneracy in Sec. 18.

(b) Let $\mathcal{S}$ be a Θ-invariant subspace as in part (a), and let $|\psi\rangle$ be a nonzero vector in $\mathcal{S}$. Also let $|\phi\rangle = \Theta|\psi\rangle$. Show that $|\phi\rangle \neq 0$ and that $\langle\phi|\psi\rangle = 0$. Use these facts to show that $\mathcal{S}$ possesses a two-dimensional, invariant subspace.

(c) Assuming that $\mathcal{S}$ is n-dimensional where $n \geq 1$ is finite, show that

$$\mathcal{S} = \mathcal{S}_1 \oplus \ldots \oplus \mathcal{S}_n, \tag{112}$$

where each $\mathcal{S}_i$ is two-dimensional and the various $\mathcal{S}_i$ are orthogonal to one another. Thus, $n = 2N$ is even.

If $\mathcal{S}$ is infinite-dimensional, we will assume that the decomposition (112) is still valid, with $n \to \infty$. (It is obvious that we can keep on splitting off invariant, two-dimensional subspaces for as long as we want.)

6. This problem continues Prob. 21.3, using the same notation.

(a) Work out $\Theta^\dagger H \Theta$ and $\Theta^\dagger K \Theta$, where

$$H = H_0 + H_{SO}. \tag{113}$$

Note that Θ commutes with Π.

(b) Let $\{|\alpha, +\rangle, \alpha = 1, 2, \ldots\}$ be an orthonormal basis in the subspace $\mathcal{E}_+$,

$$\langle\alpha, +|\beta, +\rangle = \delta_{\alpha\beta}. \tag{114}$$

We do not assume it is an energy eigenbasis; it may be a basis we will use for diagonalizing the Hamiltonian. Define

$$|\alpha, -\rangle = \Theta|\alpha, +\rangle. \tag{115}$$

According to Eq. (106), the set $\{|\alpha, -\rangle, \alpha = 1, 2, \ldots\}$ is orthonormal. Show that this set lies in the subspace $\mathcal{E}_-$. It is easy to show that it is actually a basis in $\mathcal{E}_-$; you may assume this.

(c) To obtain the energy eigenstates we must diagonalize the Hamiltonian. Let us look at the matrix elements of H in a four-dimensional space spanned by $|\alpha, \pm\rangle$ and $|\beta, \pm\rangle$ for fixed values of α and β. Put the basis vector in this order: $|\alpha, +\rangle$, $|\beta, +\rangle$, $|\alpha, -\rangle$ and $|\beta, -\rangle$.

The Hamiltonian matrix is 4×4. Find all restrictions on the 16 matrix elements that come from the properties and relations you have worked out among the operators H, K and Θ.

Write out the matrix of the 16 matrix elements in terms of a set of real parameters a, b, c, etc. For example, you can write a complex number as $a + ib$, a real number as just c, etc.

(d) Find the energy eigenvalues. Does Kramers degeneracy apply? By varying the positions of the nuclei (just the x and y components, since they must lie in the x-y plane) we can vary the parameters a, b, etc. How many parameters must we vary for the energy eigenvalues to be completely (four-fold) degenerate?