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. (The "apparatus" might be something we create in a laboratory, or it might be provided for us by some natural, physical situation.) 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.

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. 11.11). 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,

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

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 $U(\boldsymbol{\theta})$ for it. 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), (4.60), (5.6) and (5.7). 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

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

Notice that this only defines $\mathbf{J}$ relative to some definition for the unitary rotation operators $U(\mathsf{R})$, so we will have to fill in some things to make the definition complete. But it follows that a near-identity rotation operator has an expansion that begins 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 use the fact that large angle rotations can be built up as products of large numbers of small angle rotations to 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. 11.8. 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,

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

This solution can also be obtained as the limit of the product of a large number of small angle rotations, as in Sec. 11.8 (see Eqs. (11.38)–(11.40)).

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. (11.61), 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

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

The calculation is similar to that in Eq. (11.62) leading to Eq. (11.63). 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. (11.64), 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. (11.65). 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. (11.65). 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 (11.30) 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 a later set of notes.

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 the next set of notes; 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 $\mathsf{R} = \mathsf{R}(\hat{\mathbf{n}}, \theta)$. To be a little more explicit, Eq. (29) means

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

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

$$U^\dagger \boldsymbol{\sigma} U = \left[\cos\frac{\theta}{2} + i(\hat{\mathbf{n}} \cdot \boldsymbol{\sigma})\sin\frac{\theta}{2}\right]\boldsymbol{\sigma}\left[\cos\frac{\theta}{2} - i(\hat{\mathbf{n}} \cdot \boldsymbol{\sigma})\sin\frac{\theta}{2}\right]$$

$$= \cos^2\frac{\theta}{2}\boldsymbol{\sigma} + i\cos\frac{\theta}{2}\sin\frac{\theta}{2}\left[\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}, \boldsymbol{\sigma}\right] + (\hat{\mathbf{n}} \cdot \boldsymbol{\sigma})\boldsymbol{\sigma}(\hat{\mathbf{n}} \cdot \boldsymbol{\sigma})\sin^2\frac{\theta}{2}. \tag{31}$$

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

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

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. (11.41), 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 (11.45), 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 the identity,

$$(\mathsf{R}\mathbf{A}) \cdot \mathbf{B} = \mathbf{A} \cdot (\mathsf{R}^t \mathbf{B}), \tag{35}$$

valid for any rotation R and any vectors $\mathbf{A}$ and $\mathbf{B}$, to obtain

$$U(\mathbf{a} \cdot \boldsymbol{\sigma}) U^\dagger = (\mathsf{R}\mathbf{a}) \cdot \boldsymbol{\sigma}. \tag{36}$$

This may be compared to the adjoint formula (11.45) for classical rotations, which we reproduce here with a slight change of notation:

$$\mathsf{R}(\mathbf{a} \cdot \mathbf{J}) \mathsf{R}^t = (\mathsf{R}\mathbf{a}) \cdot \mathbf{J}. \tag{37}$$

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

$$U(\hat{\mathbf{n}}, 0) = 1, \qquad U(\hat{\mathbf{n}}, 2\pi) = -1, \tag{38}$$

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, 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

$$\mathsf{R}(\hat{\mathbf{n}}, 0) = \mathsf{R}(\hat{\mathbf{n}}, 2\pi) = \mathsf{I}, \tag{39}$$

a single classical rotation for which there are two unitary operators, shown by Eq. (38). More generally, it can be shown that corresponding to every classical rotation R there are two unitary spinor rotations,

$$\mathsf{R}(\hat{\mathbf{n}}, \theta) \mapsto \begin{cases} U(\hat{\mathbf{n}}, \theta), \\ U(\hat{\mathbf{n}}, \theta + 2\pi) = -U(\hat{\mathbf{n}}, \theta), \end{cases} \tag{40}$$

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,

$$U(\mathsf{R}_1)U(\mathsf{R}_2) = \pm U(\mathsf{R}_1\mathsf{R}_2), \tag{41}$$

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. (41) is correct for spinor rotations.

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. (41)? 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.
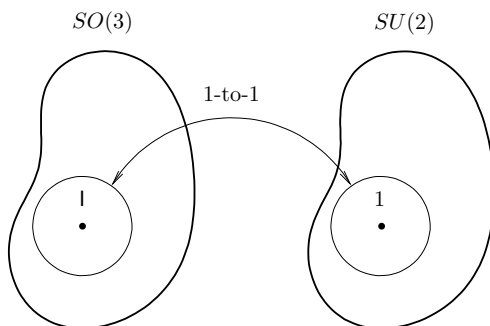

Fig. 1. There exist finite neighborhoods of the identity elements in the two groups, $SO(3)$ and $SU(2)$, which can be placed into one-to-one correspondence in such as way that the composition law is reproduced, according to Eq. (4). But the neighborhoods cannot be expanded to cover the whole group manifold 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

$$U^\dagger \sigma_i U = \sum_j R_{ij}\, \sigma_j. \tag{42}$$

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

$$\text{tr}(\sigma_j \sigma_k) = 2\delta_{jk}, \tag{43}$$

to obtain

$$R_{ij} = \frac{1}{2}\,\text{tr}\big(U^\dagger \sigma_i U \sigma_j\big). \tag{44}$$

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

$$\mathsf{R}(U_1)\mathsf{R}(U_2) = \mathsf{R}(U_1 U_2). \tag{45}$$

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

We now consider the matter of the Cayley-Klein parameters, which are discussed briefly by Sakurai. 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,

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{46}$$

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,

$$|a|^2 + |b|^2 = 1, \tag{47a}$$

$$|c|^2 + |d|^2 = 1, \tag{47b}$$

$$a^*c + b^*d = 0. \tag{47c}$$

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

$$\det U = ad - bc = 1. \tag{48}$$

Now if we take Eqs. (47c) and (48) 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,

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \tag{49}$$

where a and b are complex numbers satisfying the constraint (47a). Finally, if we break a and b into their real and imaginary parts according to

$$a = x_0 + ix_3, \qquad b = x_2 + ix_1, \tag{50}$$

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

$$U = \begin{pmatrix} x_0 + ix_3 & x_2 + ix_1 \\ -x_2 + ix_1 & x_0 - ix_3 \end{pmatrix} = x_0 + i(x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3) = x_0 + i\mathbf{x} \cdot \boldsymbol{\sigma}, \tag{51}$$

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

$$x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1. \tag{52}$$

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. (51). An arbitrary classical rotation can also be written in terms of Cayley-Klein parameters, by using Eq. (44) 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 (52) 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

$$U(\hat{\mathbf{x}}, \theta) = \cos\frac{\theta}{2} - i\sigma_x \sin\frac{\theta}{2} = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix},$$

$$U(\hat{\mathbf{y}}, \theta) = \cos\frac{\theta}{2} - i\sigma_y \sin\frac{\theta}{2} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix},$$

$$U(\hat{\mathbf{z}}, \theta) = \cos\frac{\theta}{2} - i\sigma_z \sin\frac{\theta}{2} = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}. \tag{53}$$

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. (53) can be combined to obtain an Euler angle parameterization for spinor rotations. This is a direct transcription of Eq. (11.55),

$$U(\alpha, \beta, \gamma) = U(\hat{\mathbf{z}}, \alpha)U(\hat{\mathbf{y}}, \beta)U(\hat{\mathbf{z}}, \gamma), \tag{54}$$

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

$$0 \le \alpha \le 2\pi,$$

$$0 \le \beta \le \pi,$$

$$0 \le \gamma \le 4\pi, \tag{55}$$

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,

$$|+\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |-\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{56}$$

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,

$$\hat{\mathbf{n}} = \mathsf{R}_0\hat{\mathbf{z}}. \tag{57}$$

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

$$\mathsf{R}_0 = \mathsf{R}(\alpha, \beta, 0) = \mathsf{R}(\hat{\mathbf{z}}, \alpha)\mathsf{R}(\hat{\mathbf{y}}, \beta). \tag{58}$$

The rotation R_0 that satisfies Eq. (57) is not unique; we could allow any value of the Euler angle γ (not just $\gamma = 0$). But R_0 in Eq. (58) will work. Then to obtain the spinor "pointing in" the $\hat{\mathbf{n}}$ direction (call it $|\hat{\mathbf{n}}; +\rangle$), we simply define

$$|\hat{\mathbf{n}}; +\rangle = U_0|+\rangle, \tag{59}$$

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. (59).

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,

$$U_0 = U(\hat{\mathbf{z}}, \alpha)U(\hat{\mathbf{y}}, \beta), \tag{60}$$

and appeal to the matrices (53). We find

$$|\hat{\mathbf{n}}; +\rangle = \begin{pmatrix} e^{-i\alpha/2} \cos\frac{\beta}{2} \\ e^{i\alpha/2} \sin\frac{\beta}{2} \end{pmatrix}. \tag{61}$$

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):

$$\begin{aligned} \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}|\hat{\mathbf{n}}; +\rangle = U_0 U_0^\dagger(\hat{\mathbf{n}} \cdot \boldsymbol{\sigma})U_0|+\rangle &= U_0\hat{\mathbf{n}} \cdot (\mathsf{R}_0\boldsymbol{\sigma})|+\rangle \\ &= U_0\big(\mathsf{R}_0^{-1}\hat{\mathbf{n}}\big) \cdot \boldsymbol{\sigma}|+\rangle = U_0(\hat{\mathbf{z}} \cdot \boldsymbol{\sigma})|+\rangle \\ &= U_0\sigma_z|+\rangle = U_0|+\rangle = |\hat{\mathbf{n}}; +\rangle, \end{aligned} \tag{62}$$

where we use Eq. (57). 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

$$\langle\hat{\mathbf{n}}; +|\boldsymbol{\sigma}|\hat{\mathbf{n}}; +\rangle = \hat{\mathbf{n}}. \tag{63}$$

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

$$\langle\hat{\mathbf{n}}; +|\boldsymbol{\sigma}|\hat{\mathbf{n}}; +\rangle = \langle+|U_0^\dagger\boldsymbol{\sigma}U_0|+\rangle = \mathsf{R}_0\langle+|\boldsymbol{\sigma}|+\rangle. \tag{64}$$

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. (64) becomes $\hat{\mathbf{n}}$ in accordance with Eq. (57). This proves Eq. (63).

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, revised edition, or 3.9 of Sakurai and Napolitano, second edition. Let

$$U(\alpha, \beta, \gamma) = U(\hat{\mathbf{z}}, \alpha) U(\hat{\mathbf{y}}, \beta) U(\hat{\mathbf{z}}, \gamma) \tag{65}$$

be the Euler angle representation of a rotation on a spin-$\frac{1}{2}$ system. Here we can be sloppy and ignore the distinction between the operator U and its matrix in the $|m\rangle$ basis, where $m = \pm\frac{1}{2}$, but it is traditional to denote the matrix by D, as does the book. Find the axis $\hat{\mathbf{n}}$ and angle θ of this rotation in terms of the Euler angles.

2. You are probably aware that 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,

$$A = a_0 I + \mathbf{a} \cdot \boldsymbol{\sigma}, \tag{66}$$

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

$$\det A = a_0^2 - \mathbf{a} \cdot \mathbf{a}. \tag{67}$$

It is convenient to write $\sigma_0 = I$, and to regard σ_0 as a fourth Pauli matrix. Then Eq. (66) becomes,

$$A = \sum_{\mu=0}^{3} a_\mu \, \sigma_\mu. \tag{68}$$

Notice that we have the orthogonality relation,

$$\mathrm{tr}(\sigma_\mu \sigma_\nu) = 2\delta_{\mu\nu}, \tag{69}$$

which can be used to solve Eq. (68) for the expansion coefficients,

$$a_\mu = \frac{1}{2} \, \mathrm{tr}(\sigma_\mu A). \tag{70}$$

Use these results to show that

$$\mathrm{tr}(AB) = \frac{1}{2} \sum_\mu \mathrm{tr}(\sigma_\mu A) \, \mathrm{tr}(\sigma_\mu B), \tag{71}$$

where A and B are 2×2 matrices.

Use these results to prove Eq. (45).

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. (59). 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. (71) 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

$$\mathbf{E}_{\text{phys}}(\mathbf{r}, t) = \text{Re} \sum_{\mu=1}^{2} \hat{\boldsymbol{\epsilon}}_\mu E_\mu e^{i(kz-\omega t)}, \tag{72}$$

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

$$|E_x|^2 + |E_y|^2 = E_0^2, \tag{73}$$

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

$$\chi = \frac{1}{E_0} \begin{pmatrix} E_x \\ E_y \end{pmatrix}. \tag{74}$$

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. (72). 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. (73) and (74) 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

$$\chi_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \chi_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{75}$$

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

$$\chi_r = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}, \quad \chi_\ell = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \tag{76}$$

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. Sakurai says there is no uniformity in the literature about these conventions, and in fact he uses the opposite conventions for left and right circular polarizations; but the conventions I am quoting here are the ones used by Jackson and Born and Wolf, and I think most physicists use them. 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

$$\begin{aligned} s_0 &= (|E_x|^2 + |E_y|^2)/E_0^2 = 1, \\ s_1 &= (|E_x|^2 - |E_y|^2)/E_0^2, \\ s_2 &= 2\,\mathrm{Re}(E_x E_y^*)/E_0^2, \\ s_3 &= 2\,\mathrm{Im}(E_x E_y^*)/E_0^2. \end{aligned} \tag{77}$$

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

$$s_1^2 + s_2^2 + s_3^2 = 1, \tag{78}$$

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 (74), 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 (74) 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 (74) by

$\pi/2$, that is,

$$\begin{pmatrix} E'_x \\ E'_y \end{pmatrix} = \begin{pmatrix} e^{-i\pi/4} & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \end{pmatrix},\tag{79}$$

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 $\mathbf{s}$-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?