Physics 221A
Fall 1996
Notes 10

Rotations in Quantum Mechanics, and Rotations of Spin-$\frac{1}{2}$ Systems

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 begin with a discussion of the physical meaning of rotations in quantum mechanics.

In classical mechanics, we can rotate the state of a dynamical system, i.e., we can rotate all the position and momentum vectors $(\mathbf{r}, \mathbf{p})$ 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 (either pure or mixed) with the apparatus which 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, but without specifying the phase of that state, which is unobservable anyway as long as we can only make measurements on the rotated ensemble. In order to see the physical effects of the phase of the rotated state, it would be necessary to rotate only part of a system, and to observe interference effects with the unrotated part.

There are physical situations which do this for us, that is, situations in which the time evolution generated by some Hamiltonian is equivalent to a time-dependent rotation. A notable example is the evolution of spins in magnetic fields; another example is Thomas precession (a relativistic but purely inertial effect which rotates all the dynamical variables of an accelerated system). In other cases, the time evolution of a subsystem of a given system can be described in terms of rotations; for example, a molecule simultaneously undergoes rotations and vibrations, and the rotational (i.e., orientational) degrees of freedom evolve according to a time-dependent rotation. Similarly, the spins of the electrons in an atom can be seen as precessing in the magnetic field produced by the nucleus (which is a moving charge in the electron rest frame, and which may have an intrinsic magnetic field of its own). By using physical effects which rotate a quantum system (possibly only a subsystem of a larger system), and by studying interference effects, it is possible to observe the phases associated with rotations. 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 which occurs on rotating a spin-$\frac{1}{2}$ system by 2π can be observed.

Let us therefore make a series of postulates or demands which we might reasonably expect rotation operators acting on the Hilbert space of some quantum mechanical system to satisfy. We will assume that these are unitary operators defined or parameterized by the classical rotation operators $\mathcal{R}$ or matrices R, so that given R, we have an operator $U(\mathsf{R})$. For the time being we will work only with proper rotations, so that $\mathsf{R} \in SO(3)$. We will denote the postulated association by

$$\mathsf{R} \mapsto U(\mathsf{R}), \tag{10.1}$$

which means that U is a function of the classical rotation R. We will assume that when the classical rotation is the identity, so is the unitary operator,

$$U(\mathsf{I}) = 1. \tag{10.2}$$

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

$$U(\mathsf{R}_1)U(\mathsf{R}_2) = U(\mathsf{R}_1\mathsf{R}_2). \tag{10.3}$$

These requirements then imply that inverse rotations are mapped into inverse unitary operators,

$$U(\mathsf{R}^{-1}) = U(\mathsf{R})^{-1} = U(\mathsf{R})^\dagger. \tag{10.4}$$

If the requirements (10.2) to (10.4) are satisfied, then we say that $U(\mathsf{R})$ [more precisely, the mapping (10.1)] 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. (10.2) to (10.4), the operators $U(\mathsf{R})$ must also satisfy physical properties we expect of rotations. For example, as mentioned above, the expectation values of vector operators should transform as classical vectors.

The key to finding a unitary representation of the rotations is to begin with infinitesimal rotations. As a special case of Eq. (10.1), if R is near-identity, then so must be $U(\mathsf{R})$, so that when θ is small, we must have

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = I + \theta\hat{\mathbf{n}}\cdot\boldsymbol{\mathsf{J}} \mapsto U(\mathsf{R}) = 1 - \frac{i}{\hbar}K, \tag{10.5}$$

where K is a Hermitian operator, depending on θ and $\hat{\mathbf{n}}$. (We split off a factor of i in the correction term, to make K Hermitian; the factor $-1/\hbar$ is conventional.) In fact, it is easy to use the representation law (10.3) for infinitesimal rotations to show that K must have the form,

$$K = \theta\hat{\mathbf{n}}\cdot\mathbf{J}, \tag{10.6}$$

where $\mathbf{J}$ is a vector of Hermitian operators (not to be confused with the vector of matrices $\boldsymbol{\mathsf{J}}$). We define $\mathbf{J}$ to be the *angular momentum* of the quantum system; the $\hbar$ in the denominator of Eq. (10.5) guarantees that $\mathbf{J}$ has the dimensions of angular momentum as in classical mechanics.

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 which 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 motion of a charged particle in the field of a magnetic monopole). We take the generator of rotations to be the most fundamental role of angular momentum, because these generators (the components of $\mathbf{J}$) are conserved in systems with rotational symmetry (among other reasons).

To summarize the case for infinitesimal rotations, we have

$$\mathsf{I} + \theta\hat{\mathbf{n}}\cdot\boldsymbol{\mathsf{J}} \mapsto 1 - \frac{i}{\hbar}\theta\hat{\mathbf{n}}\cdot\mathbf{J}, \tag{10.7}$$

which is of course only valid out to first order in the small angle θ. If the angle θ is not small, then the approximation (10.7) cannot be used directly, but we can break the angle θ up into a large number N of small angles θ/N, to which Eq. (10.7) can be applied. This corresponds to the obvious geometrical fact that a finite rotation can be built up out of a large number of infinitesimal ones. Then we have

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = \lim_{N\to\infty}\Bigl(\mathsf{I} + \frac{\theta}{N}\hat{\mathbf{n}}\cdot\boldsymbol{\mathsf{J}}\Bigr)^N = \exp(\theta\hat{\mathbf{n}}\cdot\boldsymbol{\mathsf{J}}), \tag{10.8}$$

where the final equality is a matrix generalization of the familiar limit from elementary calculus,

$$\lim_{N\to\infty}\Bigl(1 + \frac{x}{N}\Bigr)^N = e^x. \tag{10.9}$$

Of course, the exponential expression (10.8) we obtain in this manner agrees with our earlier calculation resulting in Eq. (9.29), which was based on solving a differential equation. In a similar way, we can take the limit of the product of the infinitesimal unitary rotation operators,

$$\lim_{N\to\infty}\Bigl(1 - \frac{i}{\hbar}\frac{\theta}{N}\hat{\mathbf{n}}\cdot\mathbf{J}\Bigr)^N = \exp\Bigl(-\frac{i}{\hbar}\theta\hat{\mathbf{n}}\cdot\mathbf{J}\Bigr). \tag{10.10}$$

In this way we obtain the association between finite rotations and the corresponding unitary operators,

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = \exp(\theta\hat{\mathbf{n}}\cdot\boldsymbol{\mathsf{J}}) \mapsto U(\mathsf{R}) = \exp\Bigl(-\frac{i}{\hbar}\theta\hat{\mathbf{n}}\cdot\mathbf{J}\Bigr) \equiv U(\hat{\mathbf{n}},\theta). \tag{10.11}$$

In effect, we have taken a presumed representation of infinitesimal rotations as in Eq. (10.7) and used it to build up a representation for finite rotations as well.

If we have a unitary representation of the rotations, as above, then the generators $\mathbf{J}$ of the representation must satisfy certain commutation relations. For let us consider the rotation C defined in Eq. (9.46), and its unitary representative $U(\mathsf{C})$. We have

$$U(\mathsf{C}) = U(\mathsf{R}_1)U(\mathsf{R}_2)U(\mathsf{R}_1^{-1})U(\mathsf{R}_2^{-1}). \tag{10.12}$$

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. (10.11), and multiply the series. The calculation is similar to that leading to Eq. (9.47), and it gives

$$U(\mathsf{C}) = 1 - \frac{1}{\hbar^2}\theta_1\theta_2[\hat{\mathbf{n}}_1\cdot\mathbf{J}, \hat{\mathbf{n}}_2\cdot\mathbf{J}] + \ldots. \tag{10.13}$$

On the other hand, we can use the final expression in Eq. (9.47) itself and Eq. (10.7) for infinitesimal rotations to obtain

$$U(\mathsf{C}) = 1 - \frac{i}{\hbar}\theta_1\theta_2(\hat{\mathbf{n}}_1\times\hat{\mathbf{n}}_2)\cdot\mathbf{J} + \ldots. \tag{10.14}$$

These two results are consistent only if

$$[\hat{\mathbf{n}}_1\cdot\mathbf{J}, \hat{\mathbf{n}}_2\cdot\mathbf{J}] = i\hbar(\hat{\mathbf{n}}_1\times\hat{\mathbf{n}}_2)\cdot\mathbf{J}, \tag{10.15}$$

or,

$$\boxed{[J_i, J_j] = i\hbar\,\epsilon_{ijk}\,J_k.} \tag{10.16}$$

These are, of course, 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 (9.24) for the J matrices; the two commutation relations are the same, apart from conventional factors of i and $\hbar$.

Thus, 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 which a vector of Hermitian operators $\mathbf{J}$ can take, given that it satisfies the commutation relations (10.16). For example, we will be interested in the matrices which represent these operators in some appropriately chosen basis. When we have found specific operators $\mathbf{J}$ which satisfy the commutation relations (10.16), 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. (10.11), 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.

Thus, we must begin by finding a representation of the angular momentum commutation relations (10.16). 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 of course possesses a 2-dimensional ket space. To find operators $\mathbf{J}$ acting on this space which satisfy the commutation relations (10.16), we simply notice that

$$\mathbf{J} = \frac{\hbar}{2}\boldsymbol{\sigma} \tag{10.17}$$

will do the trick. Therefore we provisionally take the rotation operators to be

$$U(\hat{\mathbf{n}}, \theta) = e^{-i\theta\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}/2} = \cos\frac{\theta}{2} - i(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sin\frac{\theta}{2}, \tag{10.18}$$

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

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

$$\langle\psi'|\boldsymbol{\sigma}|\psi'\rangle = \langle\psi|U^\dagger\boldsymbol{\sigma}U|\psi\rangle = \mathsf{R}\langle\psi|\boldsymbol{\sigma}|\psi\rangle, \tag{10.19}$$

or,

$$U^\dagger\boldsymbol{\sigma}U = \mathsf{R}\boldsymbol{\sigma}, \tag{10.20}$$

where it is understood that $U = U(\hat{\mathbf{n}}, \theta)$ and $\mathsf{R} = \mathsf{R}(\hat{\mathbf{n}}, \theta)$.

To see if Eq. (10.20) is true, we simply substitute Eq. (10.18) to obtain

$$\begin{aligned} U^\dagger\boldsymbol{\sigma}U &= \Big[\cos\frac{\theta}{2} + i(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sin\frac{\theta}{2}\Big]\boldsymbol{\sigma}\Big[\cos\frac{\theta}{2} - i(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sin\frac{\theta}{2}\Big] \\ &= \cos^2\frac{\theta}{2}\boldsymbol{\sigma} + i\cos\frac{\theta}{2}\sin\frac{\theta}{2}\big[\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}, \boldsymbol{\sigma}\big] + (\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\boldsymbol{\sigma}(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\sin^2\frac{\theta}{2}. \end{aligned} \tag{10.21}$$

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

$$\big[\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}, \boldsymbol{\sigma}\big] = -2i\,\hat{\mathbf{n}}\times\boldsymbol{\sigma}, \qquad (\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})\boldsymbol{\sigma}(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}) = 2\hat{\mathbf{n}}(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}) - \boldsymbol{\sigma}, \tag{10.22}$$

which we use to rewrite Eq. (10.21) in the form,

$$U^\dagger\boldsymbol{\sigma}U = \cos\theta\,\boldsymbol{\sigma} + (1-\cos\theta)\hat{\mathbf{n}}(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}) + \sin\theta\,\hat{\mathbf{n}}\times\boldsymbol{\sigma}. \tag{10.23}$$

Finally, by comparing this with Eq. (9.30), we see that Eq. (10.20) is verified. Encouraged by this result, we henceforth consider the operators $U(\hat{\mathbf{n}},\theta)$ defined by Eq. (10.18) to be rotation operators for spin-$\frac{1}{2}$ systems.

Before leaving the result (10.20), 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 (9.31), 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. (10.20), so that

$$\boxed{U\boldsymbol{\sigma}U^\dagger = \mathsf{R}^{-1}\boldsymbol{\sigma}.} \tag{10.24}$$

Next we dot both sides by some vector $\mathbf{a}$, and use the identity,

$$(\mathsf{R}\mathbf{A})\cdot\mathbf{B} = \mathbf{A}\cdot(\mathsf{R}^t\mathbf{B}), \tag{10.25}$$

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

$$U(\mathbf{a}\cdot\boldsymbol{\sigma})U^\dagger = (\mathsf{R}\mathbf{a})\cdot\boldsymbol{\sigma}. \tag{10.26}$$

This may be compared to the adjoint formula (9.31) for classical rotations, which we reproduce here with a slight change of notation:

$$\mathsf{R}(\mathbf{a}\cdot\mathbf{J})\mathsf{R}^t = (\mathsf{R}\mathbf{a})\cdot\mathbf{J}. \tag{10.27}$$

We will refer to Eq. (10.24) or its variants as the adjoint formula for spinor rotations.

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

$$U(\hat{\mathbf{n}},0) = 1, \qquad U(\hat{\mathbf{n}},2\pi) = -1, \tag{10.28}$$

according to Eq. (10.18). 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 which 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. (10.1)–(10.4). In fact, the U operators defined by Eq. (10.18) 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

$$\mathsf{R}(\hat{\mathbf{n}}, 0) = \mathsf{R}(\hat{\mathbf{n}}, 2\pi) = \mathsf{I}, \tag{10.29}$$

a single classical rotation for which there are two unitary operators, shown by Eq. (10.28). More generally, it can be shown that corresponding to every classical rotation R there are two unitary spinor rotations,

$$\mathsf{R}(\hat{\mathbf{n}}, \theta) \mapsto \begin{cases} U(\hat{\mathbf{n}}, \theta), \\ U(\hat{\mathbf{n}}, \theta + 2\pi) = -U(\hat{\mathbf{n}}, \theta), \end{cases} \tag{10.30}$$

which replaces Eq. (10.1). 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. (10.18).] For example, the representation law (10.3) could be rewritten in the form,

$$U(\mathsf{R}_1)U(\mathsf{R}_2) = \pm U(\mathsf{R}_1\mathsf{R}_2), \tag{10.31}$$

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. (10.31) is correct for spinor rotations.

We explained in Notes 9 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 (10.16), a task which 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 which 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. (10.3) must be replaced by Eq. (10.31)? 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 $+1$. For example, the notation $U(2)$ represents the group of 2×2 complex unitary matrices without additional restriction.

The significance of $SU(2)$ in the present discussion is that every spinor rotation $U(\hat{\mathbf{n}}, \theta)$ defined by Eq. (10.18) 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 is 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 9, 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.

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 (10.2). Next, once we have found a

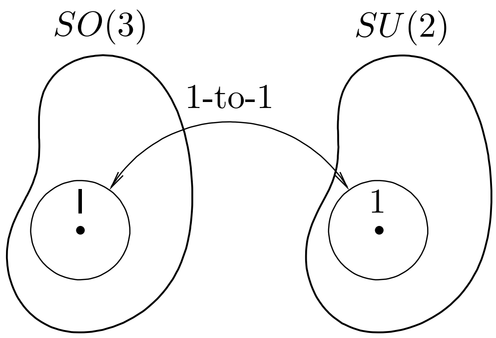

Fig. 10.1. There exist finite neighborhoods of the identity elements in the two groups, $SO(3)$ and $SU(2)$, which can be placed into 1-to-1 correspondence in such as way that the composition law is reproduced, according to Eq. (10.3). But the neighborhoods cannot be expanded to cover the whole group manifold without losing the 1-to-1 correspondence.

representation of the angular momentum commutation relations by Eq. (10.17), we have a one-to-one correspondence between infinitesimal neighborhoods of the respective identity elements, as indicated by Eq. (10.7). 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 (10.3). This is illustrated in Fig. 10.1, and it is in this sense that we say that the representation of $SO(3)$ by $SU(2)$ is locally 1-to-1. 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 1-to-2.

The 1-to-2 character of the spinor represenation of rotations can be seen in another way. Let us return to the spinor adjoint formula (10.20), which we write in the form

$$U^\dagger \sigma_i U = \sum_j R_{ij}\, \sigma_j. \tag{10.32}$$

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

$$\mathrm{tr}(\sigma_j \sigma_k) = 2\delta_{jk}, \tag{10.33}$$

to obtain

$$R_{ij} = \frac{1}{2}\,\mathrm{tr}\big(U^\dagger \sigma_i U \sigma_j\big). \tag{10.34}$$

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

$$\mathsf{R}(U_1)\mathsf{R}(U_2) = \mathsf{R}(U_1 U_2). \tag{10.35}$$

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.

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 which 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,

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{10.36}$$

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,

$$|a|^2 + |b|^2 = 1, \tag{10.37a}$$

$$|c|^2 + |d|^2 = 1, \tag{10.37b}$$

$$a^* c + b^* d = 0. \tag{10.37c}$$

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

$$\det U = ad - bc = 1. \tag{10.38}$$

Now if we take Eqs. (10.37c) and (10.38) 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,

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \tag{10.39}$$

where a and b are complex numbers satisfying the constraint (10.37a). Finally, if we break a and b into their real and imaginary parts according to

$$a = x_1 + ix_2, \qquad b = x_3 + ix_4, \tag{10.40}$$

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

$$U = \begin{pmatrix} x_1 + ix_2 & x_3 + ix_4 \\ -x_3 + ix_4 & x_1 - ix_2 \end{pmatrix}, \tag{10.41}$$

where the four real numbers (x_1, x_2, x_3, x_4) satisfy the constraint

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 1. \tag{10.42}$$

The parameters (x_1, x_2, x_3, x_4) are the Cayley-Klein parameters, in terms of which an arbitrary spinor rotation is represented by Eq. (10.41). An arbitrary classical rotation can also be written in terms of Cayley-Klein parameters, by using Eq. (10.34) 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 (10.42) 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.

Let us now tabulate the spinor rotations about the three coordinate axes, much as we did in Eq. (9.13) for classical rotations. We have

$$\begin{aligned}
U(\hat{\mathbf{x}}, \theta) &= \cos\frac{\theta}{2} - i\sigma_x \sin\frac{\theta}{2} = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \\
U(\hat{\mathbf{y}}, \theta) &= \cos\frac{\theta}{2} - i\sigma_y \sin\frac{\theta}{2} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \\
U(\hat{\mathbf{z}}, \theta) &= \cos\frac{\theta}{2} - i\sigma_z \sin\frac{\theta}{2} = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}.
\end{aligned} \tag{10.43}$$

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. (10.43) can be combined to obtain an Euler angle parameterization for spinor rotations. This is a direct transcription of Eq. (9.42),

$$U(\alpha,\beta,\gamma) = U(\hat{\mathbf{z}},\alpha)U(\hat{\mathbf{y}},\beta)U(\hat{\mathbf{z}},\gamma), \tag{10.44}$$

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

$$\begin{aligned} 0 &\le \alpha \le 2\pi, \\ 0 &\le \beta \le \pi, \\ 0 &\le \gamma \le 4\pi, \end{aligned} \tag{10.45}$$

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

Let us now consider a simple application of spinor rotations. 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,

$$[|+\rangle] = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad [|-\rangle] = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{10.46}$$

where again we enclose a ket in brakets to indicate the vector of its components with respect to a chosen basis (here the S_z basis). 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).

In any case, we will 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,

$$\hat{\mathbf{n}} = \mathsf{R}_0\hat{\mathbf{z}}. \tag{10.47}$$

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

$$\mathsf{R}_0 = \mathsf{R}(\alpha,\beta,0) = \mathsf{R}(\hat{\mathbf{z}},\alpha)\mathsf{R}(\hat{\mathbf{y}},\beta). \tag{10.48}$$

The rotation R_0 which satisfies Eq. (10.47) is not unique; we could allow any value of the Euler angle γ (not just $\gamma = 0$). But R_0 in Eq. (10.48) will work. Then to obtain the spinor "pointing in" the $\hat{\mathbf{n}}$ direction (call it $|\hat{\mathbf{n}};+\rangle$), we simply define

$$|\hat{\mathbf{n}};+\rangle = U_0|+\rangle, \tag{10.49}$$

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. (10.49).

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,

$$U_0 = U(\hat{\mathbf{z}},\alpha)U(\hat{\mathbf{y}},\beta), \tag{10.50}$$

and appeal to the matrices (10.43). We find

$$[|\hat{\mathbf{n}};+\rangle] = \begin{pmatrix} e^{-i\alpha/2}\cos\frac{\beta}{2} \\ e^{i\alpha/2}\sin\frac{\beta}{2} \end{pmatrix}. \tag{10.51}$$

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 $\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}$, i.e., 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 (10.20):

$$\begin{aligned} \hat{\mathbf{n}}\cdot\boldsymbol{\sigma}|\hat{\mathbf{n}};+\rangle &= U_0U_0^\dagger(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})U_0|+\rangle = U_0\hat{\mathbf{n}}\cdot(\mathsf{R}_0\boldsymbol{\sigma})|+\rangle \\ &= U_0\big(\mathsf{R}_0^{-1}\hat{\mathbf{n}}\big)\cdot\boldsymbol{\sigma}|+\rangle = U_0(\hat{\mathbf{z}}\cdot\boldsymbol{\sigma})|+\rangle \\ &= U_0\sigma_z|+\rangle = U_0|+\rangle = |\hat{\mathbf{n}};+\rangle, \end{aligned} \tag{10.52}$$

where we use Eq. (10.47). 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

$$\langle\hat{\mathbf{n}};+|\boldsymbol{\sigma}|\hat{\mathbf{n}};+\rangle = \hat{\mathbf{n}}. \tag{10.53}$$

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

$$\langle\hat{\mathbf{n}};+|\boldsymbol{\sigma}|\hat{\mathbf{n}};+\rangle = \langle+|U_0^\dagger\boldsymbol{\sigma}U_0|+\rangle = \mathsf{R}_0\langle+|\boldsymbol{\sigma}|+\rangle. \tag{10.54}$$

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. (10.54) becomes $\hat{\mathbf{n}}$ in accordance with Eq. (10.47). This proves Eq. (10.53).

It is a fact that for a spin-$\frac{1}{2}$ system, every spinor "points in" some direction, i.e., 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.