Physics 221A
Fall 1996
Notes 9
Rotations in Ordinary Space

The importance of rotations in quantum mechanics can hardly be overemphasized. The theory of rotations is of direct importance in all areas of atomic, molecular, nuclear and particle physics, and in large areas of condensed matter physics as well. The rotation group is the first nontrivial symmetry group we encounter in a study of quantum mechanics, and serves as a paradigm for other symmetry groups one will encounter perhaps later, such as the $SU(3)$ symmetry which acts on the color degrees of freedom in quantum chromodynamics. Furthermore, transformations which have the same mathematical form as rotations but which have nothing to do with rotations in the usual physical sense, such as isotopic spin transformations in nuclear physics, are also important.

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

Rotations in the abstract are operators which act on points of ordinary, 3-dimensional space, and map them into other points in such a manner that one point, call it O, remains fixed, while all distances between points remain invariant. We will denote such an operator by $\mathcal{R}$. This operator should be viewed in a geometrical sense, and is to be kept distinct from the associated rotation matrix R, which will be introduced momentarily. We note that an operation which preserves lengths must also preserve angles, since angles can be expressed in terms of lengths. If P is a point of space and P' is its image under the action of a rotation $\mathcal{R}$, then we will write

$$P' = \mathcal{R}P. \tag{9.1}$$

If we write $\mathbf{r} = OP$, $\mathbf{r}' = OP'$, then we can also write this as

$$\mathbf{r}' = \mathcal{R}\mathbf{r}. \tag{9.2}$$

In this equation, everything should be understood in a geometrical sense, e.g., $\mathbf{r}$ represents a directed line segment, not a triplet of numbers (which would represent the coordinates of that line segment with respect to some coordinate system).

The product of two rotations is denoted $\mathcal{R}_1\mathcal{R}_2$, which means, apply $\mathcal{R}_2$ first, then $\mathcal{R}_1$. Rotations do not commute in general, so that $\mathcal{R}_1\mathcal{R}_2 \neq \mathcal{R}_2\mathcal{R}_1$, in general. The inverse of a rotation $\mathcal{R}$ is denoted $\mathcal{R}^{-1}$.

We take the point O to be the origin of a set of axes, indicated by unit vectors $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$ or $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$. We can allow the rotation $\mathcal{R}$ to act on these unit vectors to create new unit vectors,

$$\hat{\mathbf{e}}'_i = \mathcal{R}\hat{\mathbf{e}}_i, \tag{9.3}$$

which represents a rotated frame.

The unit vectors $\hat{\mathbf{e}}_i$ allow us to represent rotation operators in terms of matrices. We define the rotation matrix R (with components R_{ij}) corresponding to $\mathcal{R}$ by

$$R_{ij} = (\hat{\mathbf{e}}_i, \mathcal{R}\hat{\mathbf{e}}_j) = \hat{\mathbf{e}}_i \cdot (\mathcal{R}\hat{\mathbf{e}}_j) = \hat{\mathbf{e}}_i \cdot \hat{\mathbf{e}}'_j, \tag{9.4}$$

where, as indicated, the round brackets indicate the ordinary scalar product of vectors. In other words, the rotation matrix R contains the matrix elements of the operator $\mathcal{R}$ with respect to the basis $\hat{\mathbf{e}}_i$. In terms of the rotation matrix, the transformation (9.2) becomes

$$\mathbf{r}' = \mathsf{R}\mathbf{r}, \tag{9.5}$$

where now $\mathbf{r}$, $\mathbf{r}'$ are seen as triplets of numbers, i.e., the coordinates of the old and new points with respect to the basis $\hat{\mathbf{e}}_i$. Equation (9.5) represents ordinary matrix multiplication.

Of course it is a small theorem to show that Eq. (9.5) is really equivalent to Eq. (9.2) under the definition (9.4) of the rotation matrix. The proof runs as follows:

$$r'_i = \hat{\mathbf{e}}_i \cdot \mathbf{r}' = \hat{\mathbf{e}}_i \cdot (\mathcal{R}\mathbf{r}) = \sum_j (\hat{\mathbf{e}}_i, \mathcal{R}\hat{\mathbf{e}}_j) r_j = \sum_j R_{ij}\, r_j. \tag{9.6}$$

In a similar way, we can prove other theorems. One of these says that if operators $\mathcal{R}_1$ and $\mathcal{R}_2$ correspond to matrices R_1, R_2, then the operator $\mathcal{R}_1\mathcal{R}_2$ corresponds to matrix $\mathsf{R}_1\mathsf{R}_2$. Similarly, if $\mathcal{R}$ corresponds to R, then $\mathcal{R}^{-1}$ corresponds to R^{-1}. As we say, the rotation matrices R form a *representation* of the geometrical rotation operators $\mathcal{R}$.

When dealing with rotations or any other symmetry group in physics, it is important to keep distinct the active and passive points of view. In this course we will adopt the active point of view unless otherwise noted (as does Sakurai's book). Many other books, however, take the passive point of view, including some standard monographs on rotations, such as Edmonds. The active point of view is usually preferable, because it is more amenable to an abstract treatment; but the passive point of view is also necessary sometimes.

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

In the passive point of view, we do not rotate our system or the points in it, but we do rotate our coordinate axes. Thus, in the passive point of view, there is only one point, but two coordinate systems. In a book which adopts the passive point of view, an equation such as $\mathbf{r}' = \mathsf{R}\mathbf{r}$ probably represents the coordinates $\mathbf{r}$ and $\mathbf{r}'$ of a single point with respect to two (the old and new) coordinate systems. With this interpretation, the matrix R has a different meaning than the matrix R used in the active point of view [such as that in Eq. (9.5) and elsewhere in these notes], being in fact the inverse of the latter. Therefore caution must be exercised in comparing different references.

Since the rotation $\mathcal{R}$ preserves lengths, we have

$$|\mathbf{r}'|^2 = |\mathbf{r}|^2, \tag{9.7}$$

when $\mathbf{r}$, $\mathbf{r}'$ are related by Eq. (9.5). From this it easily follows that R is orthogonal,

$$\mathsf{R}\mathsf{R}^t = \mathsf{R}^t\mathsf{R} = \mathsf{I}, \tag{9.8}$$

where the t denotes the transpose, or

$$\mathsf{R}^{-1} = \mathsf{R}^t. \tag{9.9}$$

Since R is orthogonal, it belongs to the group $O(3)$; here we invoke the standard notation for the group of $n \times n$, real, orthogonal matrices, which is $O(n)$.

In addition to the requirements placed on $\mathcal{R}$ above, we wish now to add a further one, namely that $\mathcal{R}$ preserve the sense (right-handed or left-handed) of sets of basis vectors. That is, if $\hat{\mathbf{e}}_i$ represents a right-handed system of axes, as we will assume, then we will require that the rotated axes $\hat{\mathbf{e}}'_i$, defined by Eq. (9.3), also be right-handed. This is equivalent to the requirement,

$$\det \mathsf{R} = 1. \tag{9.10}$$

In general, the determinant of an orthogonal matrix is either $+1$ or -1; in the former case, we speak of *proper* rotations, in the latter, of *improper* rotations. As an example of an improper rotation, we have the matrix,

$$\mathsf{P} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{9.11}$$

which inverts vectors through the origin (it is closely related to the parity operator, which, however, acts on the Hilbert space of a quantum system). One can show that every improper rotation is the product of P times a proper rotation, and vice versa. Thus, the group $O(3)$

of orthogonal matrices breaks up into two subsets, the proper and improper rotations. Of these, the proper rotations form a subgroup of $O(3)$, denoted $SO(3)$, whereas the improper rotations do not form a group (since they do not contain the identity element). In the notation $SO(3)$, the S stands for "special," which means det $= +1$; this S has the same meaning in other contexts, like $SU(2)$. [More generally, the group $SO(n)$ is the group of $n \times n$, real orthogonal matrices with determinant $+1$.] For the duration of these notes we will restrict consideration to proper rotations whose matrices belong to the group $SO(3)$, but we will have to return to improper rotations at a later date.

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

$$\mathsf{R}(\hat{\mathbf{n}}, \theta_1)\mathsf{R}(\hat{\mathbf{n}}, \theta_2) = \mathsf{R}(\hat{\mathbf{n}}, \theta_2)\mathsf{R}(\hat{\mathbf{n}}, \theta_1) = \mathsf{R}(\hat{\mathbf{n}}, \theta_1 + \theta_2). \tag{9.12}$$

The rotations about the three coordinate axes are of special interest; these are

$$\begin{aligned} \mathsf{R}_{\hat{\mathbf{x}}}(\theta) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \\ \mathsf{R}_{\hat{\mathbf{y}}}(\theta) &= \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \\ \mathsf{R}_{\hat{\mathbf{z}}}(\theta) &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{9.13}$$

In fact, one can show that any rotation can be represented as $\mathsf{R}(\hat{\mathbf{n}}, \theta)$, for some axis $\hat{\mathbf{n}}$ and some angle θ. Thus, there is no loss of generality in writing a rotation in this form. We will call this form the *axis-angle parameterization* of the rotations. This theorem is not totally obvious, but the proof is not difficult (it can be found in Goldstein). It will not be repeated here.

Next we consider *near-identity* or *infinitesimal* rotations. These are rotations whose matrices are close to the identity, so that

$$\mathsf{R} = \mathsf{I} + \epsilon\mathsf{A}, \tag{9.14}$$

where A is a matrix and the ϵ is a reminder that the correction term is small. Such matrices represent rotations by an infinitesimal angle about some axis. By substituting Eq. (9.14) into the orthogonality condition (9.8), we easily find

$$\mathsf{A} + \mathsf{A}^t = 0, \tag{9.15}$$

i.e., the matrix A is antisymmetric.

A convenient parameterization of the antisymmetric matrices is given by

$$\mathsf{A} = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} = \sum_{i=1}^{3} a_i \mathsf{J}_i, \tag{9.16}$$

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

$$\mathsf{J}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\mathsf{J}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix},$$

$$\mathsf{J}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{9.17}$$

Equations (9.17) can be summarized by writing

$$(\mathsf{J}_i)_{jk} = -\epsilon_{ijk}. \tag{9.18}$$

We will also write the sum in Eq. (9.16) as $\mathbf{a} \cdot \mathbf{J}$, i.e., as a dot product of vectors, but we must remember that $\mathbf{a}$ is a triplet of ordinary numbers, while $\mathbf{J}$ is a triplet of matrices. Notice that this way of writing antisymmetric matrices gives us an alternative notation for the cross product, i.e.,

$$(\mathbf{a} \cdot \mathbf{J})\mathbf{u} = \mathbf{a} \times \mathbf{u}, \tag{9.19}$$

for all vectors $\mathbf{u}$.

We have just seen that every rotation matrix has an axis-angle parameterization. What then are the axis and angle for a near-identity rotation matrix $\mathsf{I} + \epsilon\mathsf{A}$? The answer is that the axis is in the direction of the vector $\mathbf{a}$, i.e., $\hat{\mathbf{n}} = \mathbf{a}/|\mathbf{a}|$, and that the angle is given by $\theta = \epsilon|\mathbf{a}|$. Succinctly, we have

$$\theta\hat{\mathbf{n}} = \epsilon\mathbf{a}. \tag{9.20}$$

To prove this we let a near-identity rotation matrix act on an arbitrary vector $\mathbf{u}$, and find

$$\mathsf{R}\mathbf{u} = (\mathsf{I} + \epsilon\mathsf{A})\mathbf{u} = [\mathsf{I} + \epsilon(\mathbf{a} \cdot \mathbf{J})]\mathbf{u} = \mathbf{u} + \epsilon\mathbf{a} \times \mathbf{u}. \tag{9.21}$$

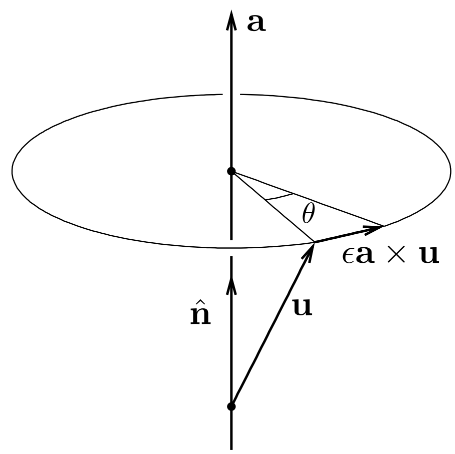

Fig. 9.1. The axis $\hat{\mathbf{n}}$ of a near-identity rotation matrix $\mathsf{R} = \mathsf{I} + \epsilon\mathsf{A}$ is in the direction of the vector $\mathbf{a}$ associated with the antisymmetric matrix A. The small angle of rotation is $\theta = \epsilon|\mathbf{a}|$.

From this it follows that $\mathsf{R}\mathbf{a} = \mathbf{a}$, so that $\mathbf{a}$ is in the direction of the axis unit vector $\hat{\mathbf{n}}$. Then with the help of a picture such as Fig. 9.1, it is easy to see that the small angle of rotation is $\theta = \epsilon|\mathbf{a}|$. Thus, we have

$$\mathsf{R}(\hat{\mathbf{n}}, \theta) = \mathsf{I} + \theta\, \hat{\mathbf{n}} \cdot \mathbf{J}, \tag{9.22}$$

for small angles of rotation θ.

It is convenient to tabulate here some useful properties of the J matrices [see also Eq. (9.19)]. First, the product of two J matrices can be written,

$$\big(\mathsf{J}_i \mathsf{J}_j\big)_{k\ell} = \delta_{i\ell}\delta_{kj} - \delta_{ij}\delta_{k\ell}, \tag{9.23}$$

as follows from Eq. (9.18). From this there follows the commutation relation,

$$\boxed{\big[\mathsf{J}_i, \mathsf{J}_j\big] = \epsilon_{ijk}\, \mathsf{J}_k,} \tag{9.24}$$

or

$$\boxed{\big[\mathbf{a} \cdot \mathbf{J}, \mathbf{b} \cdot \mathbf{J}\big] = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{J}.} \tag{9.25}$$

Finally, there are a variety of trace formulas which are useful, of which the following is one:

$$\operatorname{tr}\big[(\mathbf{a} \cdot \mathbf{J})(\mathbf{b} \cdot \mathbf{J})\big] = -2(\mathbf{a} \cdot \mathbf{b}). \tag{9.26}$$

We will now establish a connection between infinitesimal rotations and the finite rotations $\mathsf{R}(\hat{\mathbf{n}},\theta)$ which take place about a fixed axis. When $\Delta\theta$ is a small angle, we invoke Eqs. (9.12) and (9.22) to write,

$$\mathsf{R}(\hat{\mathbf{n}},\theta+\Delta\theta) = \mathsf{R}(\hat{\mathbf{n}},\Delta\theta)\mathsf{R}(\hat{\mathbf{n}},\theta) = [\mathsf{I} + \Delta\theta(\hat{\mathbf{n}}\cdot\mathsf{\mathbf{J}})]\mathsf{R}(\hat{\mathbf{n}},\theta), \tag{9.27}$$

or,

$$\frac{d}{d\theta}\mathsf{R}(\hat{\mathbf{n}},\theta) = (\hat{\mathbf{n}}\cdot\mathsf{\mathbf{J}})\mathsf{R}(\hat{\mathbf{n}},\theta) = \mathsf{R}(\hat{\mathbf{n}},\theta)(\hat{\mathbf{n}}\cdot\mathsf{\mathbf{J}}). \tag{9.28}$$

The matrices $\mathsf{R}(\hat{\mathbf{n}},\theta)$ and $\hat{\mathbf{n}}\cdot\mathsf{\mathbf{J}}$ commute, because they represent finite and infinitesimal rotations about a single axis. Equation (9.28) is a differential equation which is easily solved, subject to the initial condition $\mathsf{R}(\hat{\mathbf{n}},0) = \mathsf{I}$, and it gives

$$\boxed{\mathsf{R}(\hat{\mathbf{n}},\theta) = \exp(\theta\,\hat{\mathbf{n}}\cdot\mathsf{\mathbf{J}}).} \tag{9.29}$$

This exponential can be conceived of as a power series; it converges for all values of $\hat{\mathbf{n}}$ and θ. This result is the generalization of Eq. (9.22) to finite angles.

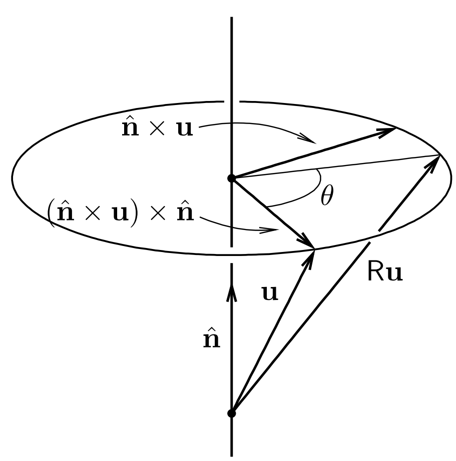

Fig. 9.2. The result of applying a finite rotation $\mathsf{R}(\hat{\mathbf{n}},\theta)$ to a vector $\mathbf{u}$ can be expressed as a linear combination of the vectors $\mathbf{u}$, $\hat{\mathbf{n}}\times\mathbf{u}$, and $(\hat{\mathbf{n}}\times\mathbf{u})\times\hat{\mathbf{n}}$, the latter two of which are orthogonal.

Finite rotations about a fixed axis can be expressed in another way. By using the properties of the J matrices listed above, one can express higher powers of the matrix $\hat{\mathbf{n}}\cdot\mathsf{\mathbf{J}}$

in terms of lower powers, and reexpress the the exponential series (9.29), acting on an arbitrary vector $\mathbf{u}$, in the following form:

$$\mathsf{R}(\hat{\mathbf{n}},\theta)\mathbf{u} = \hat{\mathbf{n}}(\hat{\mathbf{n}}\cdot\mathbf{u})(1-\cos\theta) + \mathbf{u}\cos\theta + (\hat{\mathbf{n}}\times\mathbf{u})\sin\theta. \tag{9.30}$$

The geometrical meaning of this formula is easy to see; the rotation about $\hat{\mathbf{n}}$ leaves the component of $\mathbf{u}$ parallel to $\hat{\mathbf{n}}$ invariant, and rotates the component perpendicular to $\hat{\mathbf{n}}$ by an angle θ in the perpendicular plane. This is illustrated in Fig. 9.2.

Next we derive another property of some importance. We have noted that any antisymmetric matrix can be written in the form $\mathbf{a}\cdot\mathbf{J}$ for some vector $\mathbf{a}$, and conversely. If we now conjugate this by a proper rotation matrix R_0, to get $\mathsf{R}_0(\mathbf{a}\cdot\mathbf{J})\mathsf{R}_0^t$, the result is still antisymmetric, and must therefore have the form $\mathbf{b}\cdot\mathbf{J}$ for some vector $\mathbf{b}$. What, then, is the relation between $\mathbf{a}$ and $\mathbf{b}$? The answer is $\mathbf{b} = \mathsf{R}_0\mathbf{a}$, as shown by the formula,

$$\boxed{\mathsf{R}_0(\mathbf{a}\cdot\mathbf{J})\mathsf{R}_0^t = (\mathsf{R}_0\mathbf{a})\cdot\mathbf{J}.} \tag{9.31}$$

This formula is of such frequent occurrance in applications that we will give it a name. We call it the *adjoint formula*, because of its relation to the adjoint representation of the group $SO(3)$. An immediate consequence of this formula is the following. We exponentiate both sides, interpreting $\mathbf{a}$ as $\theta\hat{\mathbf{n}}$. This gives

$$\exp\bigl[\mathsf{R}_0(\theta\hat{\mathbf{n}}\cdot\mathbf{J})\mathsf{R}_0^t\bigr] = \mathsf{R}_0\exp(\theta\hat{\mathbf{n}}\cdot\mathbf{J})\mathsf{R}_0^t = \exp\bigl[(\mathsf{R}_0\mathbf{a})\cdot\mathbf{J}\bigr], \tag{9.32}$$

or,

$$\boxed{\mathsf{R}_0\mathsf{R}(\hat{\mathbf{n}},\theta)\mathsf{R}_0^t = \mathsf{R}(\mathsf{R}_0\hat{\mathbf{n}},\theta).} \tag{9.33}$$

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

An interpretation of the adjoint formula (9.31) is given by allowing both sides to act on the vector $\mathsf{R}_0\mathbf{u}$, where $\mathbf{u}$ is arbitrary. This gives,

$$\mathsf{R}_0(\mathbf{a}\cdot\mathbf{J})\mathbf{u} = \bigl[(\mathsf{R}_0\mathbf{a})\cdot\mathbf{J}\bigr]\mathsf{R}_0\mathbf{u}, \tag{9.34}$$

or

$$\mathsf{R}_0(\mathbf{a}\times\mathbf{u}) = (\mathsf{R}_0\mathbf{a})\times(\mathsf{R}_0\mathbf{u}). \tag{9.35}$$

In other words, the cross product transforms as a vector under the action of the (proper) rotation R_0. (This would not be true if R_0 were improper.) The proof of Eq. (9.35) is straightforward, and provides a proof of the adjoint formula (9.31).

Equation (9.33) has a similar interpretation, which is clearer if we write it in the form

$$\mathsf{R}_0\mathsf{R}(\mathsf{R}_0^t\hat{\mathbf{n}},\theta)\mathsf{R}_0^t = \mathsf{R}(\hat{\mathbf{n}},\theta), \tag{9.36}$$

showing that if we convert a rotation matrix in axis-angle form to a new basis, it is equivalent to transforming the axis to the new basis, and keeping the angle the same. In other words, the axis of a given rotation transforms as a vector under proper rotations. (Again, this would not be true for improper rotations.)

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

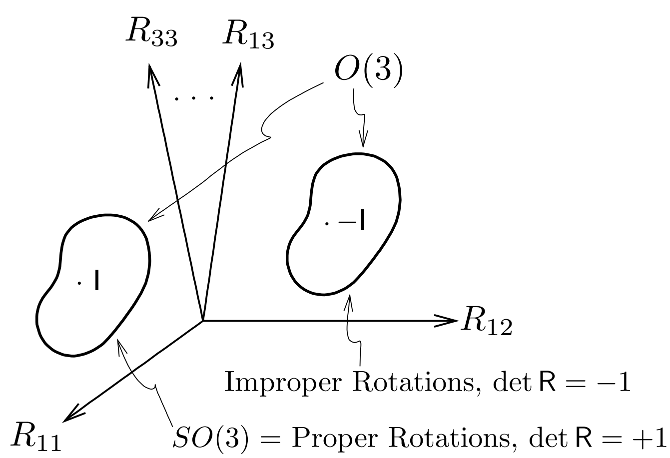

Fig. 9.3. The rotation group $O(3)$ can be thought of as a 3-dimensional surface imbedded in the 9-dimensional space of all 3×3 matrices. It consists of two disconnected pieces, one of which contains the proper rotations and constitutes the group $SO(3)$, and the other of which contains the improper rotations. The proper rotations include the identity matrix I, and the improper rotations include the parity matrix $-\mathsf{I}$.

It is often useful to think of the rotations and their parameters in geometrical terms. We imagine setting up the 9-dimensional space of all 3×3 real matrices, in which the coordinates are the components of the matrix, R_{ij}. Of course it is difficult to visualize a 9-dimensional space, but we can use our imagination as in Fig. 9.3. The 6 constraints implicit in $\mathsf{R}\mathsf{R}^t = \mathsf{I}$ imply that the orthogonal matrices lie on a 3-dimensional surface imbedded in this space.

This surface is difficult to draw realistically, so it is simply indicated as a nondescript blob in the figure. More exactly, this surface consists of two disconnected pieces, containing the proper and improper matrices. This surface (both pieces) is the *group manifold* for the group $O(3)$, while the piece consisting of the proper rotations alone is the group manifold for $SO(3)$. The identity matrix I occupies a single point in the group manifold $SO(3)$, while the improper matrix $-\mathsf{I}$ lies in the other piece of the group manifold $O(3)$. Finally, we see that parameters of the rotations are nothing but coordinates on the group manifold. Naturally, a 3-dimensional manifold requires 3 coordinates.

In addition to the axis-angle parameterization of the rotations, another important parameterization is the *Euler angles*. To construct the Euler angles, let us return to the frame $\hat{\mathbf{e}}_i$ introduced at the beginning of these notes, and recall the rotated frame $\hat{\mathbf{e}}'_i$ defined in Eq. (9.3). We will take it as geometrically obvious that the rotation operator $\mathcal{R}$ or the corresponding rotation matrix R is uniquely specified by the orientation of the rotated frame. Therefore to obtain parameters of the rotation, we can specify the orientation of the rotated frame, that is, the orientation of all three rotated axes.

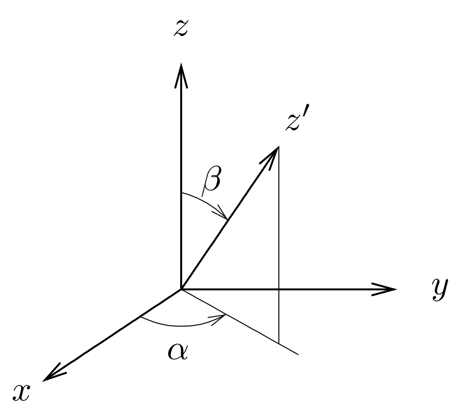

Fig. 9.4. The Euler angles α and β are the spherical angles of the rotated z'-axis as seen in the unrotated frame.

We begin by specifying the orientation of the rotated z'-axis. This axis points in some direction, which we can indicate by its spherical angles, say, (α, β), with respect to the unrotated frame. This is illustrated in Fig. 9.4. We have in mind here some rotation R, which corresponds to definite orientations of the three primed axes. Of course, R maps the old (unprimed axes) into the new (primed) ones, and in particular, it satisfies $\hat{\mathbf{z}}' = \mathsf{R}\hat{\mathbf{z}}$. Another rotation matrix which also maps the old z-axis into the new z'-axis is R_1, defined by

$$\mathsf{R}_1 = \mathsf{R}(\hat{\mathbf{z}}, \alpha)\mathsf{R}(\hat{\mathbf{y}}, \beta). \tag{9.37}$$

By examining Fig. 9.4, it is easy to see that R_1 satisfies

$$\mathsf{R}_1\hat{\mathbf{z}} = \hat{\mathbf{z}}', \tag{9.38}$$

since the first rotation by angle β about the y-axis swings the z-axis down in the x-z plane, and then the second rotation by angle α about the z-axis rotates the vector in a cone, bringing it into the final position for the z'-axis. Of course, R_1 is not in general equal to R, for R_1 is only designed to orient the z'-axis correctly, while R puts all three primed axes into their correct final positions. But R_1 can get the $\hat{\mathbf{x}}'$- and $\hat{\mathbf{y}}'$-axes wrong only by some rotation in the x'-y' plane; therefore if we follow the β and α rotations by a third rotation by some new angle, say, γ, about the z'-axis, then we can guarantee that all three axes achieve their desired orientations. That is, we can write an arbitrary rotation R in the form,

$$\mathsf{R} = \mathsf{R}(\hat{\mathbf{z}}', \gamma)\mathsf{R}_1 = \mathsf{R}(\hat{\mathbf{z}}', \gamma)\mathsf{R}(\hat{\mathbf{z}}, \alpha)\mathsf{R}(\hat{\mathbf{y}}, \beta), \tag{9.39}$$

for some angles (α, β, γ).

But it is not convenient to express a rotation in terms of elementary rotations about a mixture of old and new axes, as in Eq. (9.39); it is more convenient to express it purely in terms of rotations about the old axes. To do this, we can call on some geometrical intuition, which will convince you that the effect of following R_1 by the rotation $\mathsf{R}(\hat{\mathbf{z}}', \gamma)$ can also be achieved by preceding R_1 by the rotation $\mathsf{R}(\hat{\mathbf{z}}, \gamma)$ (which is about the unprimed z-axis.) That is, we have the identity,

$$\mathsf{R}(\hat{\mathbf{z}}', \gamma)\mathsf{R}_1 = \mathsf{R}_1\mathsf{R}(\hat{\mathbf{z}}, \gamma). \tag{9.40}$$

Sakurai attempts to present geometrical arguments for this identity, but an analytic proof is also easily given, by invoking the second version of the adjoint formula, Eq. (9.36). For in view of Eq. (9.38), we have

$$\mathsf{R}_1^t\mathsf{R}(\hat{\mathbf{z}}', \gamma)\mathsf{R}_1 = \mathsf{R}(\hat{\mathbf{z}}, \gamma), \tag{9.41}$$

which is equivalent to Eq. (9.40). In this manner we obtain the Euler-angle parameterization of the rotations,

$$\boxed{\mathsf{R}(\alpha, \beta, \gamma) = \mathsf{R}_{\hat{\mathbf{z}}}(\alpha)\mathsf{R}_{\hat{\mathbf{y}}}(\beta)\mathsf{R}_{\hat{\mathbf{z}}}(\gamma).} \tag{9.42}$$

Equation (9.42) constitutes the zyz-convention for the Euler angles, which is particularly appropriate for quantum mechanical applications. Other conventions are possible, and the zxz-convention is common in books on classical mechanics. Also, you should note that most books on classical mechanics and some books on quantum mechanics adopt the passive point of view, which usually means that the rotation matrices in those books stand for the transposes of the rotation matrices in these notes.

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

The allowed ranges on the Euler angles are the following:

$$\begin{aligned} &0 \le \alpha \le 2\pi, \\ &0 \le \beta \le \pi, \\ &0 \le \gamma \le 2\pi. \end{aligned} \tag{9.43}$$

The ranges on α and β follow from the fact that they are spherical angles, while the range on γ follows from the fact that the γ-rotation is used to bring the x'- and y'-axes into proper alignment in their plane. If the Euler angles lie within the interior of the ranges indicated, then the represenation of the rotations is unique; but if one or more of the Euler angles takes on their limiting values, then the representation may not be unique. For example, if $\beta = 0$, then the rotation is purely about the z-axis, and depends only on the sum of the angles, $\alpha+\gamma$. In other words, apart from exceptional points at the ends of the ranges, the Euler angles form a 1-to-1 coordinate system on the group manifold $SO(3)$.

We now consider some issues regarding the noncommutativity of the rotations. As pointed out earlier, rotations do not commute, so $\mathsf{R}_1\mathsf{R}_2 \ne \mathsf{R}_2\mathsf{R}_1$, in general. An exception, also noted above, is the case that R_1 and R_2 are about the same axis, but when rotations are taken about different axes they generally do not commute. Because of this, the rotation group $SO(3)$ is said to be *non-Abelian*. (Recall that the translation group is commutative, or *Abelian*.)

Let us write $\mathsf{R}_1 = \mathsf{R}(\hat{\mathbf{n}}_1, \theta_1)$ and $\mathsf{R}_2 = \mathsf{R}(\hat{\mathbf{n}}_2, \theta_2)$ for two rotations written in axis-angle form, so that

$$\mathsf{R}_1\mathsf{R}_2 = \exp\bigl[\theta_1(\hat{\mathbf{n}}_1\cdot\mathbf{J})\bigr]\exp\bigl[\theta_2(\hat{\mathbf{n}}_2\cdot\mathbf{J}\bigr]. \tag{9.44}$$

We are particularly interested in the case $\hat{\mathbf{n}}_1 \ne \hat{\mathbf{n}}_2$. Because of the commutation relations (9.24) or (9.25), the matrices in the two exponentials in Eq. (9.44) do not commute, so the product of exponentials cannot simply be written as the exponential of the sum. Nevertheless, the product of the rotations R_1 and R_2 is certainly some new rotation, which can be written in axis-angle form, say, $\mathsf{R}(\hat{\mathbf{n}}_3, \theta_3)$. But it is somewhat complicated to find $\hat{\mathbf{n}}_3$ and θ_3 in terms of $(\hat{\mathbf{n}}_1, \theta_1)$ and $(\hat{\mathbf{n}}_2, \theta_2)$. The situation is worse in the Euler angle parameterization; it is quite a messy task to find $(\alpha_3, \beta_3, \gamma_3)$ in terms of $(\alpha_1, \beta_1, \gamma_1)$ and $(\alpha_2, \beta_2, \gamma_2)$ in the product,

$$\mathsf{R}(\alpha_3, \beta_3, \gamma_3) = \mathsf{R}(\alpha_1, \beta_1, \gamma_1)\mathsf{R}(\alpha_2, \beta_2, \gamma_2). \tag{9.45}$$

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

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

$$\mathsf{C} = \mathsf{R}_1\mathsf{R}_2\mathsf{R}_1^{-1}\mathsf{R}_2^{-1}, \tag{9.46}$$

which itself is a rotation, and which becomes the identity matrix if R_1 and R_2 should commute. The matrix C is more interesting than the ordinary commutator, $[\mathsf{R}_1, \mathsf{R}_2]$, which is not a rotation. In particular, the matrix C is of special interest in the case of near-identity rotations. Let us therefore assume both θ_1 and θ_2 are small, and expand C in Taylor series in the angles. We do this by expanding out the exponentials for R_1, R_2, etc., and then multiplying the four series together. We carry this expansion out to second order, which is where the first interesting term occurs. After some algebra, we find

$$\begin{aligned}\mathsf{C} = \mathsf{R}_1\mathsf{R}_2\mathsf{R}_1^{-1}\mathsf{R}_2^{-1} &= \mathsf{I} + \theta_1\theta_2[\hat{\mathbf{n}}_1 \cdot \mathbf{J}, \hat{\mathbf{n}}_2 \cdot \mathbf{J}] + O[(\theta_1, \theta_2)^3] \\ &= \mathsf{I} + \theta_1\theta_2(\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2) \cdot \mathbf{J} + O[(\theta_1, \theta_2)^3],\end{aligned} \tag{9.47}$$

where we have used the commutation relations (9.25). We see that if R_1 and R_2 are near-identity rotations, then so is C, which is about the axis $\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2$. Sakurai does a calculation of this form when R_1 and R_2 are about the x- and y-axes, so that C is about the z-axis.

It should be no surprise to see that the commutation relations of near-identity rotations can be expressed in terms of the commutation relations of the J matrices, since the former can be expressed in terms of the latter. But there is a more profound idea at work here, which is that all of the complexity inherent in the law of composition of two arbitrary rotations such as in Eq. (9.45) is contained in the commutation relations (9.24) or (9.25) for the J matrices.

To be more precise about this, we will now briefly discuss the *Baker-Campbell-Hausdorff theorem*, which is an important result in the theory of Lie groups. This theorem concerns products of exponentials such as we see in Eq. (9.44). As noted above, this product can be written in the form $\exp[\theta_3(\hat{\mathbf{n}}_3 \cdot \mathbf{J})]$ for some axis $\hat{\mathbf{n}}_3$ and angle θ_3. Although the new exponent is not simply a sum of the two old exponents seen in Eq. (9.44), nevertheless the new exponent can be expanded in a Taylor series in powers of the angles θ_1 and θ_2. When we do this, the leading term is, in fact, the sum of the old exponents,

$$\theta_3(\hat{\mathbf{n}}_3 \cdot \mathbf{J}) = \theta_1(\hat{\mathbf{n}}_1 \cdot \mathbf{J}) + \theta_2(\hat{\mathbf{n}}_2 \cdot \mathbf{J}) + \ldots, \tag{9.48}$$

and the higher order terms contain all the complications due to the noncommutativity of the rotations. The Baker-Campbell-Hausdorff theorem concerns this series, and asserts that all the higher order terms can be expressed in terms of commutators or iterated commutators of the two leading terms, shown in Eq. (9.48). Therefore, since the commutation relations of the J matrices are known, we can use iterated commutators to construct the multiplication law for finite rotations.

The point of all this is that the commutation relations of infinitesimal rotations, expressed in terms of the commutation relations of the J matrices in Eq. (9.24), contain all the information necessary to construct the multiplication law for finite rotations. It is for this reason that the commutation relations (9.24) will play such a central role in the theory of rotations in quantum mechanics, which we will develop in subsequent notes.

The commutation relations (9.24) are referred to in common physics parlance as the *Lie algebra* of the rotation group. More generally, when you hear the words "Lie algebra," you should think of the infinitesimal symmetry operations of some symmetry group and their commutation relations, which can in principle be used to build up finite symmetry operations.

To be more precise about this terminology (at least once), we will give here the technical definition of a Lie algebra. A Lie algebra is a real vector space upon which a bracket or commutator operation [,] is defined, which is closed under the action of the bracket. The bracket is required to be linear in the two operands, antisymmetric under exchange of operands, and it must satisfy the Jacobi identity. To say that the vector space is real means that we only contemplate real coefficients when forming linear combinations of the basis vectors; those basis vectors themselves, however, may stand for objects with complex numbers in them (such as the Pauli matrices). In the case of the rotation group, the vector space which constitutes the Lie algebra is the space of all antisymmetric, 3×3 matrices, of which the J matrices form a basis. On this space, the bracket operation is the ordinary matrix commutator, which satisfies all the requirements listed above. The commutation relations (9.24) give the commutators of the basis matrices of this space, so that by forming linear combinations, the commutator of any two antisymmetric matrices can be computed. This is indicated explicitly in Eq. (9.25), which contains the commutator of an arbitrary pair of antisymmetric matrices.