Physics 221A
Fall 2010
Notes 11
Rotations in Ordinary Space

1. Introduction

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 may encounter later, such as the $SU(3)$ symmetry that acts on the color degrees of freedom in quantum chromodynamics. Furthermore, transformations that have the same mathematical form as rotations but which have nothing to do with rotations in the usual physical sense, such as isotopic spin transformations in nuclear physics, are also important. Rotations are also the spatial part of Lorentz transformations, and Lorentz invariance is one of the basic principles of modern physics. We will study Lorentz transformations in quantum mechanics in Physics 221B.

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.

2. Inertial Frames

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

We define an inertial frame as a frame in which free particles (particles upon which no forces act) move in straight lines with constant velocity. The existence of such a frame is based on the assumption that the geometry of three-dimensional, physical space is Euclidean, so that distances between points are given by Euclidean geometry and it is possible to set up three mutually orthogonal axes. The requirement that free particles move in straight lines with constant velocity is a way of saying that the frame is not accelerated or rotating.

3. Rotation Operators in Ordinary Space and Rotation Matrices

Let O be the origin of an inertial frame, and let "points of space" be identified by their coordinates with respect to this frame. We define a *rotation operator* on ordinary space $\mathcal{R}$ as an operator that maps points of space into other points of space in such a way that O is mapped into itself and all distances are preserved. Since angles can be expressed in terms of distances, it follows that all angles are preserved, too. If P is a point mapped into a new point P' under $\mathcal{R}$, we will write

$$P' = \mathcal{R}P. \tag{1}$$

We associate points P and P' with vectors $\mathbf{r} = OP$ and $\mathbf{r}' = OP'$ that are based at O. Then we can also regard $\mathcal{R}$ as an operator acting on such vectors, and write

$$\mathbf{r}' = \mathcal{R}\mathbf{r}. \tag{2}$$

Because $\mathcal{R}$ preserves all distances and angles, it maps straight lines into straight lines, and parallelograms into congruent parallelograms, etc. It follows that $\mathcal{R}$ is a linear operator,

$$\mathcal{R}(a\mathbf{A} + b\mathbf{B}) = a\mathcal{R}\mathbf{A} + b\mathcal{R}\mathbf{B}, \tag{3}$$

for all vectors $\mathbf{A}$ and $\mathbf{B}$ based at O and all real numbers a and b. It also follows that $\mathcal{R}$ is invertible, and thus $\mathcal{R}^{-1}$ exists. This is because a linear operator is invertible if the only vector it maps into the zero vector $\mathbf{0}$ is $\mathbf{0}$ itself, something that is satisfied in this case because $\mathbf{0}$ is the only vector of zero length.

The product of two rotations is denoted $\mathcal{R}_1\mathcal{R}_2$, 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. It follows from the definition that if $\mathcal{R}$, $\mathcal{R}_1$ and $\mathcal{R}_2$ are rotation operators, then so are $\mathcal{R}^{-1}$ and $\mathcal{R}_1\mathcal{R}_2$. Also, the identity operator is a rotation. These facts imply that the set of rotation operators $\mathcal{R}$ forms a group.

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

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

Similarly, we define the *rotation matrix* R associated with the operator $\mathcal{R}$ by

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

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

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

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

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

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

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

Assuming that $\mathbf{r}$ and $\mathbf{r}'$ are related by the rotation $\mathcal{R}$ as in Eq. (2), we can express the components x_i' of $\mathbf{r}'$ in terms of the components x_i of $\mathbf{r}$. We have

$$x_i' = \hat{\mathbf{e}}_i \cdot \mathbf{r}' = \hat{\mathbf{e}}_i \cdot (\mathcal{R}\mathbf{r}) = \hat{\mathbf{e}}_i \cdot \Bigl(\mathcal{R} \sum_j \hat{\mathbf{e}}_j x_j\Bigr) = \sum_j R_{ij}\, x_j, \tag{8}$$

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

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

where now $\mathbf{r}$, $\mathbf{r}'$ are seen, not as geometrical vectors as in Eq. (2), but rather as triplets of numbers, that is, the coordinates of the old and new points with respect to the basis $\hat{\mathbf{e}}_i$.

4. Active versus Passive Point of View

When dealing with rotations or any other symmetry group in physics, it is important to keep distinct the active and passive points of view. In this course we will adopt the active point of view unless otherwise noted (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 or geometrical treatment, whereas the passive point of view is irrevocably chained to coordinate representations. The passive point of view is necessary sometimes, however.

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

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

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

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

5. Properties of Rotation Matrices; the Groups $O(3)$ and $SO(3)$

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

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

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

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

This is the definition of an orthogonal matrix. The set of all 3×3 real orthogonal matrices is denoted $O(3)$ in standard notation, so our rotation matrices belong to this set. In fact, since every orthogonal matrix in $O(3)$ corresponds to a rotation operator $\mathcal{R}$ by our definition, the space of rotations is precisely the set $O(3)$. The set $O(3)$ forms a group under matrix multiplication that is isomorphic to the group of geometrical rotation operators $\mathcal{R}$ introduced above.

Equation (12) implies

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

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

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

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

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

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

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

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

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

which inverts vectors through the origin. This matrix is closely related to the parity operator in quantum mechanics, but the two are not the same since P acts on 3-dimensional space and the parity operator acts on the Hilbert space of a quantum system.

For the duration of these Notes we will mostly restrict consideration to proper rotations whose matrices belong to the group $SO(3)$, but we will return to improper rotations at a later date.

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

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

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

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

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

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

7. Infinitesimal Rotations

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

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

where A is a matrix and ϵ is a reminder that the correction term is small. Such matrices represent rotations by an infinitesimal angle about some axis. By substituting Eq. (19) into the orthogonality condition (12), we easily find

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

that is, the matrix A is antisymmetric.

A convenient parameterization of the antisymmetric matrices is given by

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

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

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

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

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

Equations (22) can be summarized by writing

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

We will also write the sum in Eq. (21) as $\mathbf{a} \cdot \mathbf{J}$, that is, as a dot product of vectors, but we must remember that $\mathbf{a}$ is a triplet of ordinary numbers, while $\mathbf{J}$ is a triplet of matrices.

Now letting $\mathbf{u}$ be an arbitrary vector, we have

$$[(\mathbf{a} \cdot \mathbf{J})\mathbf{u}]_i = (\mathbf{a} \cdot \mathbf{J})_{ij}\, u_j = -a_k\, \epsilon_{kij}\, u_j = +\epsilon_{ikj}\, a_k u_j = (\mathbf{a}\times\mathbf{u})_i, \tag{24}$$

where we use the summation convention. In other words,

$$(\mathbf{a} \cdot \mathbf{J})\mathbf{u} = \mathbf{a}\times\mathbf{u}. \tag{25}$$

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

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 $\mathbf{a}$, that is, $\hat{\mathbf{n}} = \mathbf{a}/|\mathbf{a}|$, and that the angle is given by $\theta = \epsilon|\mathbf{a}|$. Succinctly, we have

$$\theta\hat{\mathbf{n}} = \epsilon\mathbf{a}. \tag{26}$$

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

$$\mathsf{R}\mathbf{u} = (\mathsf{I} + \epsilon\mathsf{A})\mathbf{u} = [\mathsf{I} + \epsilon(\mathbf{a} \cdot \mathbf{J})]\mathbf{u} = \mathbf{u} + \epsilon\mathbf{a}\times\mathbf{u}. \tag{27}$$

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. 1, it is easy to see that the small angle of rotation is $\theta = \epsilon|\mathbf{a}|$. Thus, we have

$$\mathsf{R}(\hat{\mathbf{n}}, \theta) = \mathsf{I} + \theta\, \hat{\mathbf{n}} \cdot \mathbf{J}, \tag{28}$$

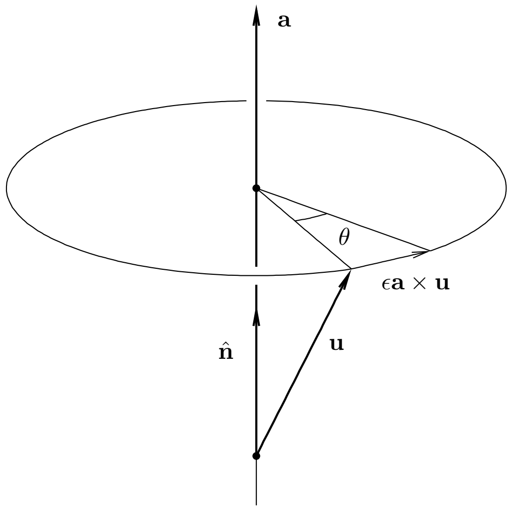

Fig. 1. The axis $\hat{\mathbf{n}}$ of a near-identity rotation matrix $\mathsf{R} = \mathsf{I} + \epsilon\mathsf{A}$ is in the direction of the vector $\mathbf{a}$ associated with the antisymmetric matrix A. The small angle of rotation is $\theta = \epsilon|\mathbf{a}|$.

for small angles of rotation θ.

We tabulate here some useful properties of the J matrices (see also Eqs. (23) and (25)). First, the product of two J matrices can be written,

$$\big(\mathsf{J}_i\mathsf{J}_j\big)_{k\ell} = \delta_{i\ell}\delta_{kj} - \delta_{ij}\delta_{k\ell}, \tag{29}$$

as follows from Eq. (23). From this there follows the commutation relation,

$$\boxed{\big[\mathsf{J}_i, \mathsf{J}_j\big] = \epsilon_{ijk}\,\mathsf{J}_k,} \tag{30}$$

or

$$\boxed{\big[\mathbf{a}\cdot\mathbf{J}, \mathbf{b}\cdot\mathbf{J}\big] = (\mathbf{a}\times\mathbf{b})\cdot\mathbf{J}.} \tag{31}$$

The commutation relations (30) obviously resemble the commutation relations for angular momentum operators in quantum mechanics, but without the $i\hbar$. As we shall see, the quantum angular momentum commutation relations are a reflection of these commutation relations of the J_i matrices at the classical level. Ultimately, these commutation relations represent the geometry of Euclidean space.

8. Exponential Form of Finite Rotations

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

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

$$\frac{d}{d\theta}\mathsf{R}(\hat{\mathbf{n}},\theta) = \lim_{\epsilon\to 0}\frac{\mathsf{R}(\hat{\mathbf{n}},\theta+\epsilon)-\mathsf{R}(\hat{\mathbf{n}},\theta)}{\epsilon}. \tag{32}$$

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

$$\mathsf{R}(\hat{\mathbf{n}},\theta+\epsilon) = \mathsf{R}(\hat{\mathbf{n}},\epsilon)\mathsf{R}(\hat{\mathbf{n}},\theta), \tag{33}$$

so the derivative becomes

$$\frac{d}{d\theta}\mathsf{R}(\hat{\mathbf{n}},\theta) = \lim_{\epsilon\to 0}\Bigl(\frac{\mathsf{R}(\hat{\mathbf{n}},\epsilon)-\mathsf{I}}{\epsilon}\Bigr)\mathsf{R}(\hat{\mathbf{n}},\theta). \tag{34}$$

In the limit, ϵ becomes a small angle, so we can use Eq. (28) to evaluate the limit, obtaining,

$$\frac{d}{d\theta}\mathsf{R}(\hat{\mathbf{n}},\theta) = (\hat{\mathbf{n}}\cdot\mathbf{J})\mathsf{R}(\hat{\mathbf{n}},\theta). \tag{35}$$

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

$$\boxed{\mathsf{R}(\hat{\mathbf{n}},\theta) = \exp(\theta\,\hat{\mathbf{n}}\cdot\mathbf{J}).} \tag{36}$$

This exponential can be expanded out in a power series that carries Eq. (28) to higher order terms,

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = \mathsf{I} + \theta\hat{\mathbf{n}}\cdot\mathbf{J} + \frac{\theta^2}{2}(\hat{\mathbf{n}}\cdot\mathbf{J})^2 + \ldots. \tag{37}$$

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

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

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

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = \mathsf{R}\Bigl(\hat{\mathbf{n}},\frac{\theta}{N}\Bigr)^N. \tag{38}$$

We can make θ/N as small as we like by making N large. Therefore we should be able to replace the small angle rotation in Eq. (38) by the small angle formula (28), so that

$$\mathsf{R}(\hat{\mathbf{n}},\theta) = \lim_{N\to\infty}\Bigl(I + \frac{\theta}{N}\hat{\mathbf{n}}\cdot\mathbf{J}\Bigr)^N. \tag{39}$$

This is a limit of matrices that is similar to a limit of numbers that can be evaluated by elementary calculus,

$$\lim_{N\to\infty}\Bigl(1+\frac{x}{N}\Bigr)^N = e^x. \tag{40}$$

It turns out that the matrix limit works in the same way, giving Eq. (36).

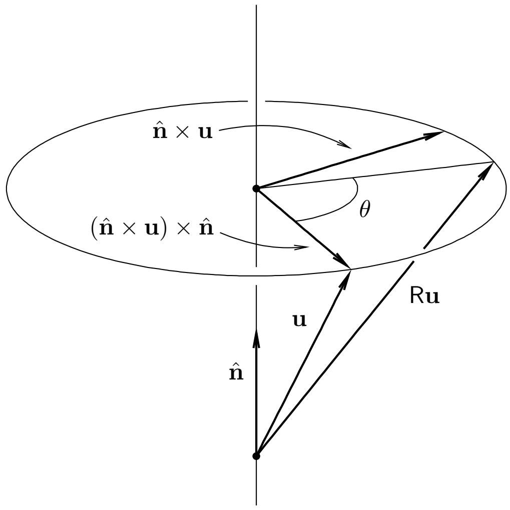

Fig. 2. The result of applying a finite rotation $\mathsf{R}(\hat{\mathbf{n}},\theta)$ to a vector $\mathbf{u}$ can be expressed as a linear combination of the vectors $\mathbf{u}$, $\hat{\mathbf{n}}\times\mathbf{u}$, and $(\hat{\mathbf{n}}\times\mathbf{u})\times\hat{\mathbf{n}}$, the latter two of which are orthogonal.

9. Explicit Form for Action of Rotation on a Vector

Finite rotations about a fixed axis can be expressed in another way. By using the properties of the $\mathbf{J}$ matrices listed above, one can express higher powers of the matrix $\hat{\mathbf{n}}\cdot\mathbf{J}$ in terms of lower powers, and reexpress the the exponential series (36), acting on an arbitrary vector $\mathbf{u}$, in the following form:

$$\mathsf{R}(\hat{\mathbf{n}},\theta)\mathbf{u} = \hat{\mathbf{n}}(\hat{\mathbf{n}}\cdot\mathbf{u})(1-\cos\theta) + \mathbf{u}\cos\theta + (\hat{\mathbf{n}}\times\mathbf{u})\sin\theta. \tag{41}$$

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. 2.

10. Adjoint Formulas

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

$$\mathsf{R}(\mathbf{a}\times\mathbf{u}) = (\mathsf{R}\mathbf{a})\times(\mathsf{R}\mathbf{u}), \tag{42}$$

or, as we say in words, the cross product transforms as a vector under proper rotations. The proof of Eq. (42) will be left as an exercise, but if you do it, you will find it is equivalent to $\det\mathsf{R} = +1$. In fact, if R is improper, there is a minus sign on the right hand side, an indication that the cross product of two true vectors is not a true vector, but rather a *pseudovector*. More on this later, when we return to the improper rotations and discuss parity in quantum mechanics.

Now let us use the notation (25) to express the cross products in Eq. (42) in terms of matrix multiplications,

$$\mathsf{R}(\mathbf{a}\cdot\mathbf{J})\mathbf{u} = [(\mathsf{R}\mathbf{a})\cdot\mathbf{J}]\mathsf{R}\mathbf{u}, \tag{43}$$

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

$$\mathsf{R}(\mathbf{a}\cdot\mathbf{J}) = [(\mathsf{R}\mathbf{a})\cdot\mathbf{J}]\mathsf{R}. \tag{44}$$

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

$$\boxed{\mathsf{R}(\mathbf{a}\cdot\mathbf{J})\mathsf{R}^t = (\mathsf{R}\mathbf{a})\cdot\mathbf{J}.} \tag{45}$$

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

Let us now replace R by R_0 to avoid confusion with other rotation matrices to appear momentarily, and let us replace $\mathbf{a}$ by $\theta\hat{\mathbf{n}}$ for the axis and angle of a rotation. Then exponentiating both sides of the adjoint formula (45), we obtain,

$$\exp\bigl[\mathsf{R}_0(\theta\hat{\mathbf{n}}\cdot\mathbf{J})\mathsf{R}_0^t\bigr] = \mathsf{R}_0\exp(\theta\hat{\mathbf{n}}\cdot\mathbf{J})\mathsf{R}_0^t = \exp\bigl[\theta(\mathsf{R}_0\hat{\mathbf{n}})\cdot\mathbf{J}\bigr], \tag{46}$$

where we have used the rule,

$$e^{ABA^{-1}} = A\,e^B A^{-1}, \tag{47}$$

valid for matrices or operators A and B when A^{-1} exists. This rule can be verified by expanding the exponential series. The result (46) can be written,

$$\boxed{\mathsf{R}_0\mathsf{R}(\hat{\mathbf{n}},\theta)\mathsf{R}_0^t = \mathsf{R}(\mathsf{R}_0\hat{\mathbf{n}},\theta).} \tag{48}$$

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

The exponentiated version of the adjoint formula (48) says that the axis of a rotation transforms as a vector under proper rotations, while the angle of rotation does not change. As with the original version of the adjoint formula (45), this only holds for proper rotations. Equation (48) might be easier to understand in a passive interpretation, in which R_0 is used to transform the components of $\mathsf{R}(\hat{\mathbf{n}},\theta)$ to a new basis. The new matrix is obtained simply by transforming the axis to the new basis, while keeping the angle fixed.

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

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

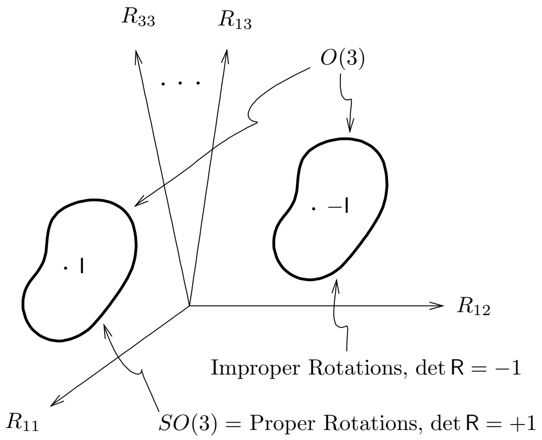

Fig. 3. The rotation group $O(3)$ can be thought of as a 3-dimensional surface imbedded in the 9-dimensional space of all 3×3 matrices. It consists of two disconnected pieces, one of which contains the proper rotations and constitutes the group $SO(3)$, and the other of which contains the improper rotations. 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. 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 $O(3)$. Any choice of three parameters for the proper rotations can be viewed as a coordinate system on the group manifold $SO(3)$.

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

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

12. Euler Angles

In addition to the axis-angle parameterization of the rotations, another important parameterization is the *Euler angles*. To construct the Euler angles, let us return to the frame $\hat{\mathbf{e}}_i$ introduced at the beginning of these notes, and recall the rotated frame $\hat{\mathbf{e}}'_i$ defined in Eq. (10). We will take it as geometrically obvious that the rotation operator $\mathcal{R}$ 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.

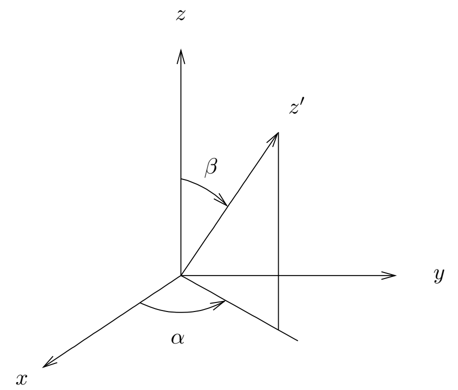

Fig. 4. The Euler angles α and β are the spherical angles of the rotated z'-axis as seen in the unrotated frame.

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. 4. We have in mind here some rotation R, which corresponds to definite orientations of the three primed axes. Matrix R maps the old (unprimed axes) into the new (primed) ones, and in particular, it satisfies $\hat{\mathbf{z}}' = \mathsf{R}\hat{\mathbf{z}}$. Another rotation matrix that also maps the old z-axis into the new z'-axis is R_1, defined by

$$\mathsf{R}_1 = \mathsf{R}(\hat{\mathbf{z}}, \alpha)\mathsf{R}(\hat{\mathbf{y}}, \beta). \tag{49}$$

By examining Fig. 4, it is easy to see that R_1 satisfies

$$\mathsf{R}_1\hat{\mathbf{z}} = \hat{\mathbf{z}}', \tag{50}$$

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

$$\mathsf{R} = \mathsf{R}(\hat{\mathbf{z}}', \gamma)\mathsf{R}_1 = \mathsf{R}(\hat{\mathbf{z}}', \gamma)\mathsf{R}(\hat{\mathbf{z}}, \alpha)\mathsf{R}(\hat{\mathbf{y}}, \beta), \tag{51}$$

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. (51); it is more convenient to express it purely in terms of rotations about the old axes. To do this, we write Eq. (51) in the form,

$$\mathsf{R} = \mathsf{R}_1\mathsf{R}_1^{-1}\mathsf{R}(\hat{\mathbf{z}}', \gamma)\mathsf{R}_1, \tag{52}$$

in which the product of the last three rotations can be rewritten with the help of the adjoint formula (48),

$$\mathsf{R}_1^{-1}\mathsf{R}(\hat{\mathbf{z}}', \gamma)\mathsf{R}_1 = \mathsf{R}\big(\mathsf{R}_1^{-1}\hat{\mathbf{z}}', \gamma\big) = \mathsf{R}(\hat{\mathbf{z}}, \gamma), \tag{53}$$

where we use Eq. (50). Notice that R_1^{-1} has rotated the axis $\hat{\mathbf{z}}'$ back into $\hat{\mathbf{z}}$. Now we can write

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

which leads to the Euler angle parameterization for the rotations,

$$\boxed{\mathsf{R}(\alpha, \beta, \gamma) = \mathsf{R}_{\hat{\mathbf{z}}}(\alpha)\mathsf{R}_{\hat{\mathbf{y}}}(\beta)\mathsf{R}_{\hat{\mathbf{z}}}(\gamma).} \tag{55}$$

Equation (55) 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:

$$\begin{aligned} &0 \le \alpha \le 2\pi, \\ &0 \le \beta \le \pi, \\ &0 \le \gamma \le 2\pi. \end{aligned} \tag{56}$$

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

13. The Noncommutativity of Rotations

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

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

$$\mathsf{R}(\hat{\mathbf{n}}_1, \theta_1)\mathsf{R}(\hat{\mathbf{n}}_2, \theta_2) = \mathsf{R}(\hat{\mathbf{n}}_3, \theta_3). \tag{57}$$

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

$$\mathsf{A}_i = \theta_i(\hat{\mathbf{n}}_i \cdot \mathbf{J}), \qquad i = 1, 2, 3, \tag{58}$$

for the three antisymmetric matrices in the exponential expressions for the three rotations in Eq. (57), then that equation can be written as

$$\exp(\mathsf{A}_1)\exp(\mathsf{A}_2) = \exp(\mathsf{A}_3), \tag{59}$$

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

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

$$\mathsf{R}(\alpha_3, \beta_3, \gamma_3) = \mathsf{R}(\alpha_1, \beta_1, \gamma_1)\mathsf{R}(\alpha_2, \beta_2, \gamma_2). \tag{60}$$

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

$$\mathsf{C} = \mathsf{R}_1\mathsf{R}_2\mathsf{R}_1^{-1}\mathsf{R}_2^{-1}, \tag{61}$$

which itself is a rotation, and which becomes the identity matrix if R_1 and R_2 should commute. The matrix C is more interesting than the ordinary commutator, $[\mathsf{R}_1, \mathsf{R}_2]$, which is not a rotation. You can think of C as taking a step in the 1-direction, then a step in the 2-direction, then a backwards step in the 1-direction, and finally a backwards step in the 2-direction, all in the space of rotations, and asking if we return to the identity. The answer is that in general we do not, although if the steps are small then we trace out a path that looks like a small parallelogram that usually does not quite close.

Let us examine C when the angles θ_1 and θ_2 are small. Let us write A_i for the antisymmetric matrices in the exponents of the exponential form for the rotations, as in Eq. (58), and let us expand the exponentials in power series. Then we have

$$\begin{aligned}\mathsf{C} &= e^{\mathsf{A}_1}\, e^{\mathsf{A}_2}\, e^{-\mathsf{A}_1}\, e^{-\mathsf{A}_2} \\ &= \big(\mathsf{I} + \mathsf{A}_1 + \tfrac{1}{2}\mathsf{A}_1^2 + \ldots\big)\big(\mathsf{I} + \mathsf{A}_2 + \tfrac{1}{2}\mathsf{A}_2^2 + \ldots\big) \\ &\qquad \times \big(\mathsf{I} - \mathsf{A}_1 + \tfrac{1}{2}\mathsf{A}_1^2 - \ldots\big)\big(\mathsf{I} - \mathsf{A}_2 + \tfrac{1}{2}\mathsf{A}_2^2 - \ldots\big) \\ &= \mathsf{I} + [\mathsf{A}_1, \mathsf{A}_2] + \ldots.\end{aligned} \tag{62}$$

The first order term vanishes, and at second order we find the commutator of the A matrices, as indicated. Writing these in terms of axes and angles as in Eq. (58), the matrix C becomes

$$\begin{aligned}\mathsf{C} &= \mathsf{I} + \theta_1\theta_2[\hat{\mathbf{n}}_1 \cdot \mathbf{J}, \hat{\mathbf{n}}_2 \cdot \mathbf{J}] + \ldots \\ &= \mathsf{I} + \theta_1\theta_2(\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2) \cdot \mathbf{J} + \ldots,\end{aligned} \tag{63}$$

where we have used the commutation relations (31). 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.

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

$$\mathsf{C} = \mathsf{I} + \phi(\hat{\mathbf{m}} \cdot \mathbf{J}). \tag{64}$$

Comparing this with Eq. (63), we have

$$\phi\, \hat{\mathbf{m}} = \theta_1\theta_2\, \hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2. \tag{65}$$

14. The Baker-Campbell-Hausdorff Theorem

It is no surprise to see the commutator of near identity rotations expressed in terms of the commutators of the $\mathbf{J}$ matrices, since the latter are the correction terms for near identity rotations. Equation (63) just gives the details. But there is an important point to be made about the commutation relations of near identity transformations, related to the Baker-Campbell-Hausdorff theorem, which we now discuss.

The theorem concerns products of exponentials of operators or matrices that need not commute, as in Eq. (59). We rewrite that equation in a different notation,

$$e^X\, e^Y = e^Z, \tag{66}$$

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

$$Z = X + Y + \frac{1}{2}[X, Y]. \tag{67}$$

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

$$Z = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] + \frac{1}{12}[Y, [Y, X]] + \ldots . \tag{68}$$

The Baker-Campbell-Hausdorff theorem concerns this series, and asserts that the general term of this Taylor series can be written in terms of commutators and iterated commutators of X and Y. This fact can be seen through third order in Eq. (68). Thus, if we know how to compute commutators, we can in principle compute all the terms of this Taylor series and obtain the multiplication rule for products of exponentials.

For example, in the case of the rotations, knowledge of the commutation relations (30) or (31) allows us to compute all the terms of the Taylor series (68) and thus obtain the multiplication law for finite rotations in terms of axis-angle parameters, at least in principle. We do not intend to do this in practice, and in fact we shall use the Baker-Campbell-Hausdorff theorem only for its suggestive and intuitive value. But the point is that all the complexity of the noncommutative multiplication law for finite rotations is implicitly contained in the commutation relations of infinitesimal rotations, that is, of the $\mathbf{J}$ matrices. This explains the emphasis placed on commutation relations in these Notes and in physics in general.

Problems

1. Prove the commutation relations (30), using Eq. (23) and the properties of the Levi-Civita symbol ϵ_{ijk}.

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

$$\mathsf{R}(\mathbf{a}\times\mathbf{b}) = (\mathsf{R}\mathbf{a})\times(\mathsf{R}\mathbf{b}). \tag{69}$$

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

$$\epsilon_{ijk} \det \mathsf{M} = \epsilon_{\ell mn} M_{i\ell} M_{jm} M_{kn}. \tag{70}$$

This is essentially the definition of the determinant. This proves Eq. (42), and hence the adjoint formulas (45) and (48).

3. The geometrical meaning of Eq. (41) is illustrated in Fig. 2. The rotation leaves the component of $\mathbf{u}$ along the axis $\hat{\mathbf{n}}$ invariant, while rotating the orthogonal component in the plane perpendicular to $\hat{\mathbf{n}}$.

By expressing powers of the J matrices in terms of lower powers, sum the exponential series (36) and obtain another proof of Eq. (41).

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

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

5. It is claimed that every rotation can be written

in Euler angle form. Find the Euler angles (α, β, γ) for the rotation $\mathsf{R}(\hat{\mathbf{x}}, \pi/2)$.