1
Rotations

This chapter is devoted to rotations in three-dimensional space. Rotations are fundamental to rigid body dynamics because there is a one-one correspondence between orientations of a rigid body and rotations in three-dimensional space. The theory of rotations is a classical subject with a rich history and a variety of modes of expression. We shall begin by expressing the elements of the theory in terms of vectors and linear operators. Next, quaternions will be introduced and additional elements of the theory developed with them. This will lead to some elegant connections between rotations and spherical geometry. This leads, via the stereographic projection and Möbius transformations, to a description of rotations in terms of complex variables.

1.1
Rotations as Linear Operators

One way to approach rotations is to study their effect on spatial objects. The language of vectors and matrices provides a natural calculus. This section reviews some basic algebra of vector spaces and establishes our notation. Then the angle of rotation and axis of rotation as well as the Euler angles are studied as ways to parameterize rotation matrices.

1.1.1
Vector Algebra

Let us first establish some notation. Let $\mathbf{V}$ be a finite-dimensional vector space over the real numbers, $\mathbb{R}$. The elements of $\mathbf{V}$, the vectors, will be denoted by bold, lower case Latin letters, $\boldsymbol{u}$. A basis of $\mathbf{V}$ is a set of vectors $\{\boldsymbol{e}_i\}$ having the property that every vector has a unique representation as a linear combination of basis elements. The basis vectors will be indexed with subscripts. Let $\mathbf{V}^*$ be the dual vector space of $\mathbf{V}$ – the space of all linear, real valued functions on $\mathbf{V}$. The elements of $\mathbf{V}^*$, the covectors, will be denoted by bold, lower case Greek

Rigid Body Mechanics: Mathematics, Physics and Applications. William B. Heard

ISBN: 3-527-40620-4

letters, $\boldsymbol{v}$. For each basis $\{\boldsymbol{e}_i\}$ of $\mathbf{V}$ there is a basis $\{\boldsymbol{\epsilon}^i\}$ of $\mathbf{V}^*$ defined by

$$\boldsymbol{\epsilon}^i(\boldsymbol{e}_j) = \delta_{ij}$$

where δ_{ij} is the Kronecker delta. If $\boldsymbol{u} = u^i \boldsymbol{e}_i$, [1] then $\mathbf{e}$ and u will denote

$$\mathbf{e} = \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{pmatrix} \qquad u = \begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix}$$

and the representation of a vector $\boldsymbol{u}$ in the basis $\mathbf{e}$ is denoted by

$$\boldsymbol{u} = \mathbf{e}u$$

The basis $\mathbf{e}$ will also be referred to as a *frame*. Similarly, if $\boldsymbol{v} = v_i \boldsymbol{\epsilon}^i$, then $\boldsymbol{\epsilon}$ and v will denote

$$\boldsymbol{\epsilon} = \begin{pmatrix} \epsilon^1 \\ \vdots \\ \epsilon^n \end{pmatrix} \qquad v = \begin{pmatrix} v_1 \cdots v_n \end{pmatrix}$$

and the representation of covector $\boldsymbol{v}$ in the basis $\boldsymbol{\epsilon}$ is denoted by

$$\boldsymbol{v} = v\boldsymbol{\epsilon}$$

This notation is also found in [1] which is a good source of material on geometrical aspects of mechanics.

The distinction between a vector $\boldsymbol{u} = \mathbf{e}u$ and its components u relative to a basis $\mathbf{e}$ is of prime importance. The former is invariant under a change of basis and the latter, of course, is not. We shall always reserve the bold-face font for the invariant form and the regular typeface components relative to a basis. A *linear operator* $\mathcal{A}$ on $\boldsymbol{V}$ takes vector $\boldsymbol{x} \in \boldsymbol{V}$ to vector $\mathcal{A}(\boldsymbol{x}) \in \boldsymbol{V}$ and preserves the operations of the vector space

$$\mathcal{A}(a\boldsymbol{x} + b\boldsymbol{y}) = a\mathcal{A}(\boldsymbol{x}) + b\mathcal{A}(\boldsymbol{y}), \text{ for all } a, b \in \mathbb{R} \text{ and } \boldsymbol{x}, \boldsymbol{y} \in \mathbf{V}$$

Given a linear operator $\mathcal{A}$ and a basis $\{\mathbf{e}_i\}$ define the matrix representation $A = [A^i_j]$ (row superscript i and column subscript j) by

$$\mathcal{A}(\mathbf{e}_j) = A^i_j \boldsymbol{e}_i$$

The action of $\mathcal{A}$ on any $\mathbf{u} = u^j \mathbf{e}_j$ is then represented as

$$\mathcal{A}\mathbf{u} = A^i_j u^j \boldsymbol{e}_i$$

1) The summation convention that repeated indices indicate a sum over their range is used here and throughout the text.

The matrix can be regarded as operating on a row of basis vectors from the right according to

$$\mathcal{A}(\boldsymbol{u}) = u^j \boldsymbol{f}_j \qquad \text{with} \quad \boldsymbol{f}_j = A^i_j \boldsymbol{e}_i$$

This can be expressed in a matrix form as

$$\begin{pmatrix} \boldsymbol{f}_1 & \boldsymbol{f}_2 & \boldsymbol{f}_3 \end{pmatrix} = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{pmatrix} \begin{pmatrix} A^1_1 & A^1_2 & A^1_2 \\ A^2_1 & A^2_2 & A^2_3 \\ A^3_1 & A^3_2 & A^3_3 \end{pmatrix}$$

or in the shorthand notation

$$\boldsymbol{f} = \mathbf{e}A$$

Alternatively it can be regarded as operating on a column of components from the left according to

$$\mathcal{A}(\boldsymbol{u}) = v^i \boldsymbol{e}_i \qquad \text{with} \quad v^i = A^i_j u_j$$

This can be expressed as

$$\begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} = \begin{pmatrix} A^1_1 & A^1_2 & A^1_2 \\ A^2_1 & A^2_2 & A^2_3 \\ A^3_1 & A^3_2 & A^3_3 \end{pmatrix} \begin{pmatrix} u^1 \\ u^2 \\ u^3 \end{pmatrix}$$

or

$$v = Au$$

Given a pair of vectors $\boldsymbol{u}, \boldsymbol{v}$, a scalar or inner product on $\boldsymbol{V}$ assigns a non-negative, real number $\langle \mathbf{u}, \mathbf{v} \rangle$ which has the following properties:

$$\text{if } \mathbf{u} \neq 0 \text{ then } \langle \mathbf{u}, \mathbf{u} \rangle > 0$$

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$$

$$\langle a^i \mathbf{e}_i, b^j \mathbf{e}_j \rangle = a^i b^j \langle \mathbf{e}_i, \mathbf{e}_j \rangle$$

An inner product may be expressed in the following equivalent ways:

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u} \cdot \boldsymbol{v} = (u^i \mathbf{e}_i) \cdot (v^j \mathbf{e}_j) = u^i v^i G_{ij}$$

where the real numbers

$$G_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j$$

are components of a symmetric, positive definite matrix G called the metric tensor. The *Euclidean* inner product is distinguished by $G_{ij} = \delta_{ij}$. In the Euclidean case

$$\boldsymbol{u} \cdot \boldsymbol{v} = u^t v = v^t u$$

Given an inner product we can define the length or norm of a vector and the angle between two vectors. The norm of $\mathbf{u}$ is

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$$

The angle between $\mathbf{u}$ and $\mathbf{v}$ is

$$\arccos\left(\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|}\right)$$

Tensors are important objects in rigid body mechanics and we now set down the basics of tensor algebra. We start with the algebraic definition of tensors of rank 2.[2] A tensor of rank 2 assigns a real number to a pair of vectors or covectors and is linear in each argument. A covariant tensor $\boldsymbol{T}$ of rank 2 assigns a real number to pairs of vectors $\boldsymbol{T}(\mathbf{u}, \mathbf{v}) \in \mathbb{R}$. The metric tensor is an example of a covariant tensor. A contravariant tensor $\boldsymbol{T}$ of rank 2 assigns a real number to pairs of covectors $\boldsymbol{T}(\boldsymbol{v}, \boldsymbol{\nu}) \in \mathbb{R}$. A mixed tensor $\boldsymbol{T}$ of rank 2 assigns a real number to a vector–covector pair $\boldsymbol{T}(\mathbf{u}, \boldsymbol{\nu}) \in \mathbb{R}$. The components of a tensor are its values on basis vectors. Thus, a covariant tensor $\boldsymbol{A}$ has components $a_{ij} = \boldsymbol{A}(\mathbf{e}_i, \mathbf{e}_j)$ and a mixed tensor $\boldsymbol{B}$ has components $b^i_j = \boldsymbol{B}(\mathbf{e}_i, \boldsymbol{\epsilon}^j)$.

Tensors can be formed from tensor products of vectors and covectors. The tensor products are denoted with the symbol $\otimes$ and are defined by their action on their arguments. Thus we define a covariant tensor $\boldsymbol{v} \otimes \boldsymbol{\nu}$ by $\boldsymbol{v} \otimes \boldsymbol{\nu}(\mathbf{u}, \mathbf{v}) = \boldsymbol{v}(\mathbf{u})\boldsymbol{\nu}(\mathbf{v})$ and define the mixed tensor $\mathbf{u} \otimes \boldsymbol{\nu}$ by $\mathbf{u} \otimes \boldsymbol{\nu}(\boldsymbol{v}, \mathbf{v}) = \boldsymbol{v}(\mathbf{u})\boldsymbol{\nu}(\mathbf{v})$. It is not the case that every tensor is a tensor product but every tensor is a linear combination of tensor products of basis vectors. For example,

$$\boldsymbol{T}(\mu_i \boldsymbol{\epsilon}^i, v^j \mathbf{e}_j) = \mu_i v^j \boldsymbol{T}(\boldsymbol{\epsilon}^i, \mathbf{e}_j)$$

Addition and scalar multiplication of the tensors can be defined by their action on vectors

$$(a\boldsymbol{T} + b\boldsymbol{S})(\mathbf{u}, \mathbf{v}) = a\boldsymbol{T}(\mathbf{u}, \mathbf{v}) + b\boldsymbol{S}(\mathbf{u}, \mathbf{v})$$

In the case of mixed tensors, the result of this construction is a new vector space $\boldsymbol{V} \otimes \boldsymbol{V}^*$. When $\boldsymbol{V}$ has dimension n, $\boldsymbol{V} \otimes \boldsymbol{V}^*$ has dimension n^2 . The same construction can be carried through for pairs of covectors and the resulting vector space $\boldsymbol{V}^* \otimes \boldsymbol{V}^*$ again has dimension n^2.

The subspace $\bigwedge^2 \boldsymbol{V}^*$ of $\boldsymbol{V}^* \otimes \boldsymbol{V}^*$ consists of skew-symmetric covariant tensors, that is, of covariant tensors $\boldsymbol{\Omega}$ which satisfy

$$\boldsymbol{\Omega}(\mathbf{u}, \mathbf{v}) = -\boldsymbol{\Omega}(\mathbf{v}, \mathbf{u})$$

2) The development extends to any rank but we will need only rank 2.

Example 1.1 *The tensor* $\boldsymbol{T} = \boldsymbol{\nu} \otimes \boldsymbol{v} - \boldsymbol{v} \otimes \boldsymbol{\nu}$ *is skew-symmetric because*

$$\boldsymbol{T}(\mathbf{u}, \mathbf{v}) = \boldsymbol{\nu}(\mathbf{u})\boldsymbol{v}(\mathbf{v}) - \boldsymbol{v}(\mathbf{v})\boldsymbol{\nu}(\mathbf{u})$$

and

$$\boldsymbol{T}(\mathbf{v}, \mathbf{u}) = \boldsymbol{\nu}(\mathbf{v})\boldsymbol{v}(\mathbf{u}) - \boldsymbol{v}(\mathbf{u})\boldsymbol{\nu}(\mathbf{v}) = -\boldsymbol{T}(\mathbf{u}, \mathbf{v}). \quad \diamondsuit$$

The members of $\bigwedge^2 \boldsymbol{V}^*$ are called 2-forms over $\boldsymbol{V}$. Covectors, members of $\boldsymbol{V}^*$, are also called 1-forms. There is a product, the *wedge product*, which produces a 2-form $\boldsymbol{\omega} \wedge \boldsymbol{\mu}$ from two 1-forms $\boldsymbol{\omega}$ and $\boldsymbol{\mu}$. The wedge product is defined by its action on its arguments

$$\boldsymbol{\omega} \wedge \boldsymbol{\mu}(\mathbf{u}, \mathbf{v}) = \boldsymbol{\omega}(\mathbf{u})\boldsymbol{\mu}(\mathbf{v}) - \boldsymbol{\omega}(\mathbf{v})\boldsymbol{\mu}(\mathbf{u})$$

The wedge product is basic to the geometric treatment of Hamiltonian mechanics.

Now we follow [1] to establish the connection between rank 2 mixed tensors and linear operators. If $\mathcal{A}$ is a linear operator, let $\boldsymbol{T}_{\mathcal{A}}$ be the tensor defined by the action $\boldsymbol{T}_{\mathcal{A}}(\mathbf{v}, \boldsymbol{\nu}) = \boldsymbol{\nu}(\mathcal{A}(\mathbf{v}))$. The components of $\boldsymbol{T}_{\mathcal{A}}$ are given by

$${\boldsymbol{T}_{\mathcal{A}}}^i_j = \boldsymbol{T}_{\mathcal{A}}(\mathbf{e}_j, \boldsymbol{\epsilon}^i) = \boldsymbol{\epsilon}^i(\mathcal{A}(\mathbf{e}_j)) = A^i_j$$

Thus the components of $\boldsymbol{T}_{\mathcal{A}}$ are the same as those of $\boldsymbol{A}$. We will exploit this correspondence, borrowing from the theory of dyadics, to represent linear transformations by

$$\mathcal{A}(u^k \mathbf{e}_k) = \mathbf{e}_i A^i_j \boldsymbol{\epsilon}^j (u^k \mathbf{e}_k) = \mathbf{e}_i A^i_j u^k \delta_{jk} = \mathbf{e}_i A^i_j u^j \tag{1.1}$$

Thus the action of the second-rank tensor depends on its object. If applied to a vector–covector pair it produces a scalar and if applied to a vector it returns another vector. This is reminiscent of the dot and double-dot products of dyadics [2] which are closely related to second-rank tensors.

The combinations $\boldsymbol{\epsilon}\mathbf{e}$ and $\mathbf{e} \otimes \boldsymbol{\epsilon}$ arise frequently in the manipulation of tensors. The product $\boldsymbol{\epsilon}\mathbf{e}$ is the identity matrix because

$$\begin{bmatrix} \boldsymbol{\epsilon}^1 \\ \boldsymbol{\epsilon}^2 \\ \boldsymbol{\epsilon} \end{bmatrix} \left(\begin{bmatrix} \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3 \end{bmatrix} \right) = [\boldsymbol{\epsilon}^i(\mathbf{e}_j)] = I$$

The product $\mathbf{e} \otimes \boldsymbol{\epsilon}$ is the identity operator because

$$\begin{bmatrix} \mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3 \end{bmatrix} \otimes \begin{bmatrix} \boldsymbol{\epsilon}^1 \\ \boldsymbol{\epsilon}^2 \\ \boldsymbol{\epsilon} \end{bmatrix} = \mathbf{e}_i \otimes \boldsymbol{\epsilon}^i$$

and for any vector $\boldsymbol{v} = v^i \mathbf{e}_i$

$$(\mathbf{e}_i \otimes \boldsymbol{\epsilon}^i)(v^k \mathbf{e}_k) = v^k \mathbf{e}_i \delta_{ik} = \boldsymbol{v}$$

The above representation leads to a succinct representation of linear operators

$$\mathcal{A} = \mathbf{e}A\boldsymbol{\epsilon}$$

which is shorthand for $\mathcal{A} = A^i_j \mathbf{e}_i \otimes \boldsymbol{\epsilon}^i$ so that

$$\mathcal{A}(\mathbf{v}) = \mathbf{e}A\boldsymbol{\epsilon}(\mathbf{e}v) = \mathbf{e}Av$$

Now consider the effect of a change of basis. Suppose $\mathbf{e}$ and $\bar{\mathbf{e}}$ are bases of $\boldsymbol{V}$ linearly related by

$$\mathbf{e} = \bar{\mathbf{e}}B$$

Then a vector $\mathbf{v}$ has the representations

$$\mathbf{v} = \mathbf{e}v = \bar{\mathbf{e}}Bv = \bar{\mathbf{e}}\bar{v}$$

so that the components in the two bases are related by

$$\bar{v} = Bv$$

The covector relations

$$\boldsymbol{I} = \mathbf{e}\boldsymbol{\epsilon} = \bar{\mathbf{e}}B\boldsymbol{\epsilon} = \bar{\mathbf{e}}\bar{\boldsymbol{\epsilon}}$$

give

$$\bar{\boldsymbol{\epsilon}} = B\boldsymbol{\epsilon}$$

Thus a covector $\boldsymbol{\mu}$ has the representations

$$\boldsymbol{\mu} = \mu\boldsymbol{\epsilon} = \mu B^{-1}\bar{\boldsymbol{\epsilon}} = \bar{\mu}\bar{\boldsymbol{\epsilon}}$$

so that the components in the two bases are related by

$$\mu = \bar{\mu}B$$

The effect of change of basis on a linear operator follows immediately

$$\mathcal{A} = \mathbf{e}A\boldsymbol{\epsilon} = \bar{\mathbf{e}}BAB^{-1}\bar{\boldsymbol{\epsilon}} = \bar{\mathbf{e}}\bar{A}\bar{\boldsymbol{\epsilon}}$$

or

$$\bar{A} = BAB^{-1}$$

Given the linear operators $\mathcal{A}$, $\mathcal{B}$ with the matrix representations A, B, the matrix representation of the composite map $\mathcal{B} \circ \mathcal{A}$, $\mathcal{A}$ followed by $\mathcal{B}$, is the matrix product BA. The basis vectors transform as

$$\mathbf{e}' = \mathbf{e}BA$$

and the components transform as

$$v' = BAv$$

For any n the vector space $\mathbb{R}^n$ has the standard basis $\{e_i\}$ where e_i is the column of length n having a single nonzero entry, 1 in row i. The inner product in the standard basis is

$$\boldsymbol{u} \cdot \boldsymbol{v} = u^t = v^u = u_i v_i$$

where no distinction is made between u_i and u^i. In $\mathbb{R}^3$ there is also a *vector product* or cross product

$$(a_i \mathbf{e}_i) \times (b_j \mathbf{e}_j) = \epsilon_{ijk} a_i b_j \mathbf{e}_k = (a_2 b_3 - a_3 b_2)\mathbf{e}_1 + (a_3 b_1 - a_1 b_3)\mathbf{e}_2 + (a_1 b_2 - a_2 b_1)\mathbf{e}_3$$

where ϵ_{kij} is the permutation symbol

$$\epsilon_{kij} = \begin{cases} 1 & \text{if } i\,j\,k \text{ is an even permutation of } 123 \\ -1 & \text{if } i\,j\,k \text{ is an odd permutation of } 123 \\ 0 & \text{otherwise} \end{cases}$$

We will have no further need in this chapter to distinguish between subscripts and superscripts, so subscripts will be used. In this case matrix entries will be denoted by A_{ij} with row index i and column index j. Superscripts will return with a vengeance when generalized coordinates are considered.

1.1.2
Rotation Operators on $\mathbb{R}^3$

A rotation is a linear transformation, $\mathcal{R}$, that fixes the origin, preserves the lengths of vectors, and preserves the orientation of bases. That is,

$$\mathcal{R} : \mathbf{V} \to \mathbf{V} : \boldsymbol{x} \mapsto \mathcal{R}(\boldsymbol{x})$$

$$\mathcal{R}(\mathbf{0}) = \mathbf{0}$$

$$\mathcal{R}(\boldsymbol{x}) \cdot \mathcal{R}(\boldsymbol{x}) = \boldsymbol{x} \cdot \boldsymbol{x}$$

$$\mathcal{R}(\mathbf{e}_1) \cdot (\mathcal{R}(\mathbf{e}_2) \times \mathcal{R}(\mathbf{e}_3)) = \mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3)$$

The length-preserving property becomes, in a matrix form,

$$\mathcal{R}(\boldsymbol{x}) \cdot \mathcal{R}(\boldsymbol{x}) = (x_j R_{ij} \mathbf{e}_i) \cdot (x_l R_{kl} \mathbf{e}_k) = x_j x_l R_{ij} R_{il} = x_i x_i \text{ for all } \{x_i\} \in \mathbb{R}^3$$

which implies that

$$R_{ij} R_{il} = \delta_{jl} \quad \text{or} \quad R^t R = I$$

This defines an *orthogonal matrix*. The orientation preserving condition becomes

$$R_{k1} \epsilon_{kij} R_{i2} R_{j3} = 1 \quad \text{or} \quad \det R = +1$$

When R and S are orthogonal, then $(RS)^t RS = S^t R^t RS = I$. When R is orthogonal $R^{-1} = R^t$ and $(R^{-1})^t R^{-1} = RR^t = I$. Therefore RS and R^{-1} are also orthogonal. Clearly I is orthogonal. The product of orthogonal matrices preserves the determinant, that is, if $\det R = \det S = 1$ then $\det RS = \det R \det S = 1$. If $\det R = 1$ then $\det R^{-1} = 1$ because $\det RR^{-1} = \det I = 1$.

It follows from these facts that the orthogonal matrices of determinant 1 form a group.[3] The group is called $SO(3)$ – the *special orthogonal group* of order 3. This group has the additional structure of a three-dimensional manifold and is therefore a Lie group. In the next section we begin to study parameterizations, or coordinates, of $SO(3)$. The basics of the Lie group theory are outlined in Appendix C.

In addition to preserving lengths, orthogonal matrices preserve angles because they preserve inner products

$$(Ru)^t Rv = u^t R^t Rv = u^t v$$

Members of $SO(3)$ also preserve $\mathbb{R}^3$ vector products in the sense that $A(x \times y) = Ax \times Ay$. To prove this, it is enough to show it for e_1, e_2, e_3, the standard basis of $\mathbb{R}^3$. Let $A \in SO(3)$ be presented in terms of its orthonormal columns, $A = [c_1 c_2 c_3]$ so that $Ae_i = c_i$. Then

$$A(e_i \times e_j) = A(\epsilon_{ijk} e_k) = \epsilon_{ijk} A e_k = \epsilon_{ijk} c_k = c_i \times c_j = Ae_i \times Ae_j$$

Every member of $SO(3)$ fixes not only the origin but actually an entire line. This follows from the structure of eigensystems of rotation matrices. The length preserving property of a rotation requires that eigenvalues have magnitude 1. To show this we must allow for complex eigenvalues and eigenvectors and use the norm $\|x\|^2 = (x^*)^t x$ where x^* is the complex conjugate of x. Then $Rx = \lambda x$ implies that $\|Rx\|^2 = \lambda^* \lambda (x^*)^t x = |\lambda|^2 \|x\|^2$. In other words $\|Rx\| = \|x\|$ implies $|\lambda| = \pm 1$. The characteristic polynomial of a rotation matrix is a cubic and one of its roots must be $+1$ because $\det R = 1$. Thus, any eigenvector corresponding to $\lambda = 1$ is fixed and real. The set of all such eigenvectors forms the *axis of rotation*.

1.1.3
Rotations Specified by Axis and Angle

First consider plane rotations. Represent vectors as complex numbers $(x, y) \leftrightarrow x + \imath y = \rho \exp(\imath\theta)$. Then a counterclockwise rotation by angle ϕ is simply multiplication by $\exp(\imath\phi)$: $z \to \exp(\imath\phi) z = \rho \exp[\imath(\theta + \phi)]$. In rectangular components

$$x + \imath y = z \to z' = e^{\imath\theta}(x + \imath y) = (x \cos\theta - y \sin\theta) + \imath(x \sin\theta + y \cos\theta)$$

3) A group G is a set equipped with a binary operation such that if $a, b \in G$ then $ab \in G$. There is an identity element e such that $ea = a$ for every $a \in G$ and for every $a \in G$ there is an inverse a^{-1} such that $aa^{-1} = e$.

This shows that rotations by angle θ about the x, y, z axes are represented by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{pmatrix} \quad \begin{pmatrix} c & 0 & s \\ 0 & 1 & 0 \\ -s & 0 & c \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

respectively, where $c = \cos\theta,\ s = \sin\theta$.

These matrices have alternative expressions as operators which emphasize the role of the axis of rotation and the rotation angle. To state the alternative forms we first introduce the idea of duality between $\mathbb{R}^3$ and $\mathfrak{so}(3)$, which is the set of all skew-symmetric 3×3 matrices, that is matrices which have the property that $A^t = -A$. In fact $\mathfrak{so}(3)$ is the Lie algebra of the Lie group $SO(3)$ (see Appendix C). To each vector v there corresponds a skew-symmetric matrix $\widehat{v}$,

$$v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \longleftrightarrow \begin{pmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{pmatrix} = \widehat{v}$$

and to each skew-symmetric matrix M there is a vector $\vec{M}$,

$$M = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix} \longleftrightarrow \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \vec{M}$$

In the language of Lie algebra $\widehat{v} = ad_v$ (see Section C.3.2).

Using the canonical basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} = \{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$, this is expressed in terms of tensor products as

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \longleftrightarrow u_1\widehat{\mathbf{i}} + u_2\widehat{\mathbf{j}} + u_3\widehat{\mathbf{k}}$$

where

$$\widehat{\mathbf{i}} = \mathbf{k} \otimes \mathbf{j} - \mathbf{j} \otimes \mathbf{k} \qquad \widehat{\mathbf{j}} = \mathbf{i} \otimes \mathbf{k} - \mathbf{k} \otimes \mathbf{i} \qquad \widehat{\mathbf{k}} = \mathbf{j} \otimes \mathbf{i} - \mathbf{i} \otimes \mathbf{j}$$

are the invariant forms of the $\widehat{\ }$ operator. Here we have identified $\mathbb{R}^3$ and $(\mathbb{R}^3)^*$ via

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \longleftrightarrow \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix}$$

or

$$\mathbf{u} \otimes \boldsymbol{v} \longleftrightarrow uv^t$$

With these preliminaries we can write that the rotations about the coordinate axes are

$$\mathbf{i}\otimes\mathbf{i}+c(I-\mathbf{i}\otimes\mathbf{i})+s\widehat{\mathbf{i}}$$
$$\mathbf{j}\otimes\mathbf{j}+c(I-\mathbf{j}\otimes\mathbf{j})+s\widehat{\mathbf{j}}$$
$$\mathbf{k}\otimes\mathbf{k}+c(I-\mathbf{k}\otimes\mathbf{k})+s\widehat{\mathbf{k}}$$

These expressions can be generalized to an arbitrary axis of rotation determined by the unit vector $\mathbf{n}$. The form of the expressions suggests that one might construct a basis containing $\mathbf{n}$ and write immediately that

$$\mathcal{R}_{\mathbf{n}}(\theta)=\mathbf{n}\otimes\mathbf{n}+\cos\theta(I-\mathbf{n}\otimes\mathbf{n})+\sin\theta\widehat{\mathbf{n}}$$

and we now proceed to justify this. Let P represent the operator relative to the standard basis which projects a vector onto the axis of rotation (Fig. 1.1)

$$P=nn^t$$

P has the property that, for any r, Pr is a multiple of n

$$Pr=nn^tr=\lambda n \qquad \lambda=n^tr$$

and P is idempotent

$$P^2=(nn^t)(nn^t)=n(n^tn)n^t=nn^t=P$$

$I-P$ then represents the operator which projects a vector onto the plane normal to $\mathbf{n}$ because for any r

$$n^t(I-P)r=n^tr-n^tnn^tr=0$$

and

$$(I-P)^2=I-P-P+P^2=I-P$$

Let $N=\widehat{n}$. Any rotation of r through an angle θ leaves Pr invariant and rotates $(I-P)r$ by an angle θ in the plane spanned by $(I-P)r$, $n\times r$. Thus the rotation is given by $r'=R_n(\theta)r$ with

$$R_n(\theta)=nn^t+\cos\theta\left(I-nn^t\right)+\sin\theta\,N$$

It is easy to show that the operators $P, I-P, N$ form a closed system defined by Table 1.1. Here we sketch the proof of one entry as an illustration

$$N^2r=n\times(n\times r)=(r\cdot n)n-r=(P-I)r$$

The relations in Table 1.1 can be used to recast the rotation operator entirely in terms of N,

$$R=I+\sin\theta N+(1-\cos\theta)N^2 \tag{1.2}$$

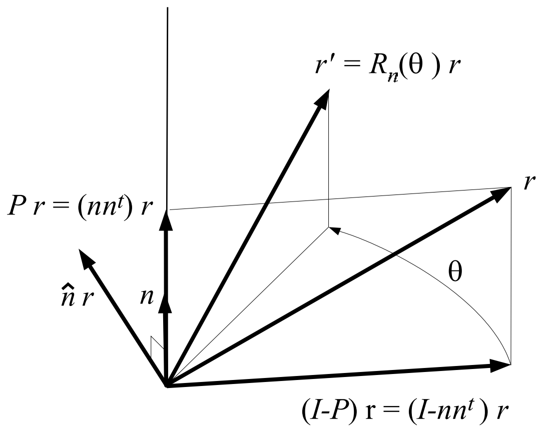

Fig. 1.1 Rotation about the axis $\mathbf{n}$ by an angle θ.

Table 1.1 Composition of projection operators $P = nn^t$, $I - P$, and $N = \hat{n}$.

$\circ$	P	$I-P$	N
P	P	0	0
$I-P$	0	$I-P$	N
N	0	N	$-(I-P)$

$\circ$	P	N	N^2
P	P	0	0
N	0	N^2	$-N$
N^2	0	$-N$	$-N^2$

We close this section with a derivation of the general rotation matrix using general principles of linear algebra. A general maxim in linear algebra is that problems should be worked in the "right" basis, i.e., the basis in which the problem assumes its simplest, most revealing form. Since rotations are given by an axis, $\mathbf{n}$, and an angle, α, the "right" basis in which to study rotations would naturally include $\mathbf{n}$. Accordingly, let $\mathbf{b}_1 = \mathbf{n}$; choose unit vector $\mathbf{b}_2 \perp \mathbf{n}$ and $\mathbf{b}_3 = \mathbf{b}_1 \times \mathbf{b}_2$. In this basis we can write immediately

$$R_n(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{pmatrix}$$

where $c = \cos\alpha$ and $s = \sin\alpha$ and α is reckoned from $\mathbf{b}_2$ in the $\mathbf{b}_2$–$\mathbf{b}_3$ plane. To work in another basis – the standard basis $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, for example – we can use the change of basis relations

$$\mathcal{R} = \mathbf{b} R_n(\alpha) \boldsymbol{\beta} = \mathbf{e} A R_n(\alpha) A^t \boldsymbol{\epsilon} = \mathbf{e} R' \boldsymbol{\epsilon}$$

A is the change of basis matrix such that

$$\mathbf{b} = \mathbf{e} A \qquad A_{ij} = \mathbf{b}_j \cdot \mathbf{e}_i$$

so column j of A contains the components b_j of vector $\mathbf{b}_j$ in the basis $\mathbf{e}$ ($\mathbf{b}_j = \mathbf{e}b_j$). That is

$$A = (b_1\ b_2\ b_2)$$

Thus

$$R' = (b_1\ b_2\ b_3)\begin{pmatrix}1 & 0 & 0\\ 0 & c & -s\\ 0 & s & c\end{pmatrix}\begin{pmatrix}b_1^t\\ b_2^t\\ b_3^t\end{pmatrix} = b_1b_1^t + c(b_2b_2^t + b_3b_3^t) + s(b_3b_2^t - b_2b_3^t)$$

Now we use the identities

$$b_1 = n,$$

$$I = b_1b_1^t + b_2b_2^t + b_3b_3^t$$

and

$$\widehat{b}_1 = b_3b_2^t - b_2b_3^t$$

to recover

$$R = nn^t + c(I - nn^t) + sN \tag{1.3}$$

1.1.4
The Cayley Transform

Equation (1.2) can be written in terms of $t = \tan\frac{1}{2}\theta$ by using the identities

$$\sin\theta = \frac{2t}{1+t^2} \qquad 1 - \cos\theta = \frac{2t^2}{1+t^2}$$

Thus

$$R = I + \frac{2t}{1+t^2}N + \frac{2t^2}{1+t^2}N^2$$

This form of R leads to the Cayley transform of R and yet another form of the rotation matrix which plays an important role in numerical calculations. We need the series expansion

$$\frac{1}{1+t^2} = \sum_{n=0}^{\infty}(-1)^n t^{2n}$$

and the results obtained from Table 1.1,

$$N^{2n} = (-1)^{n+1}N^2$$

$$N^{2n+1} = (-1)^n N$$

These provide

$$\begin{aligned} R &= I + 2\sum_{n=1}^{\infty} t^n N^n \\ &= (I + tN)\sum_{0}^{\infty} t^n N^n \\ &= (I + tN)(I - tN)^{-1} \\ &= (I - tN)^{-1}(I + tN) \end{aligned}$$

and we have arrived at the *Cayley transform*. Given a skew-symmetric matrix S, its Cayley transform is

$$\text{cay}(S) = (I - S)^{-1}(I + S)$$

The importance of $\text{cay}(S)$ is that it is always orthogonal

$$\begin{aligned} \text{cay}(S)\text{cay}(S)^t &= (I - S)^{-1}(I + S)(I + S^t)(I - S^t)^{-1} \\ &= (I - S)^{-1}(I + S)(I - S)(I + S)^{-1} \\ &= I \end{aligned}$$

The inverse of the Cayley transform is expressed as

$$\text{if } S^t = -S \text{ and } R = \text{cay}S \text{ then}$$

$$S = (R - I)(R + I)^{-1}$$

Orthogonality of R insures that S is skew-symmetric.

1.1.5 Reflections

Reflections are more basic than rotations in the sense that every rotation can be obtained by composing two reflections. First consider reflections in a plane through a line containing the complex number $\exp(\imath\phi)$. Then

$$z' = \exp(2\imath\phi)x^* \tag{1.4}$$

is the reflection of z through the line. If we apply two reflections, say through the lines containing $\exp(\imath\phi)$ and $\exp(\imath\psi)$, then

$$z'' = \exp(2\imath\psi)[\exp(2\imath\phi)z^*]^* = \exp[2\imath(\psi - \phi)]z$$

which is a rotation by an angle $2(\psi - \phi)$.

The representation of any rotation as a composition of reflections generalizes to $SO(3)$. We first need the matrix representation for reflections in planes in $\mathbb{R}^3$ and the approach used to derive (1.3) also yields the matrix representation for a reflection. The appropriate basis includes a normal $\mathbf{n}$ to the plane of reflection

$$\mathbf{b}_1 = \mathbf{n} \quad \mathbf{b}_2 \perp \mathbf{n} \quad \text{and} \quad \mathbf{b}_3 = \mathbf{b}_1 \times \mathbf{b}_2$$

Then the reflection is represented by (M for mirror)

$$M = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and use of the change of basis matrix to the standard basis gives

$$M_n = (b_1 \; b_2 \; b_3) \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1^t \\ b_2^t \\ b_3^t \end{pmatrix} = -b_1 b_1^t + b_2 b_2^t + b_3 b_3^t$$

and using $b_1 = n$ and $I = b_1 b_1^t + b_2 b_2^t + b_3 b_3^t$ we obtain

$$M_n = I - 2nn^t$$

The invariant form of the reflection operator is

$$\mathcal{M}_{\mathbf{n}} = I - 2\mathbf{n} \otimes \mathbf{n}$$

Now one can show that every rotation is the product of two reflections. Let the rotation be $\mathcal{R}_{\mathbf{n}}(\alpha)$. Choose unit vectors $\mathbf{q}$, $\mathbf{p}$ such that

$$\mathbf{p} \cdot \mathbf{n} = 0 \quad \mathbf{q} \cdot \mathbf{n} = 0 \quad \text{and} \quad \mathbf{p} \cdot \mathbf{q} = \cos \frac{1}{2}\alpha$$

The product of reflections in $\mathbf{q}$ and $\mathbf{p}$ is

$$\mathcal{A} = \mathcal{M}_p \mathcal{M}_q = I - 2\mathbf{p} \otimes \mathbf{p} - 2\mathbf{q} \otimes \mathbf{q} + 4 \cos \frac{1}{2}\alpha \mathbf{p} \otimes \mathbf{q}$$

It is easily verified that

$$\mathcal{A}\mathbf{n} = \mathbf{n}$$

$$\mathbf{n} \cdot \mathcal{A}\mathbf{p} = \mathbf{n} \cdot \mathcal{A}\mathbf{q} = 0,$$

$$\mathbf{p} \cdot \mathcal{A}\mathbf{p} = -1 + 2\cos^2 \frac{1}{2}\alpha = \cos \alpha$$

and

$$\mathbf{q} \cdot \mathcal{A}\mathbf{q} = \cos \alpha$$

which shows that

$$\mathcal{M}_p \mathcal{M}_q = \mathcal{R}_{\mathbf{n}}(\alpha)$$

1.1.6 Euler Angles

The specification of a rotation by axis and angle is convenient and subject to direct geometrical interpretation. It is, however, inconvenient as a basis for rigid body dynamics because of redundancy. Any rotation matrix can be specified by three parameters (there are nine entries in the orthogonal matrix and six constraints imposed on the inner products of the columns). The redundancy arises because all three components of the axis are used as parameters even though only two are independent. The correspondence is also double-valued in that $\boldsymbol{n}$, α yield the same rotation as $-\boldsymbol{n}$, $-\alpha$. We now consider the *Euler angles* which provide a three-parameter specification of a rotation.

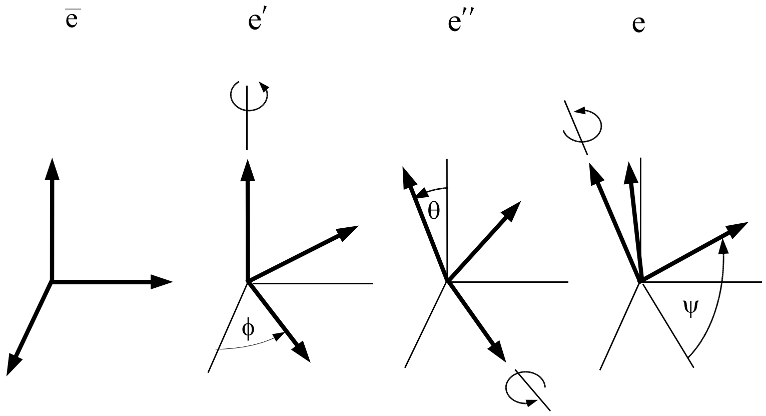

Fig. 1.2 Definition of the Euler angles ϕ, θ, ψ.

The Euler angle parameterization of a rotation is the composite of three intermediate rotations. The rotations are defined relative to the standard basis $\bar{e}$, and the effect of the intermediate rotations on this basis will be denoted by primes (Fig. 1.2).

1. Rotate about $\bar{e}_3$ by an angle ϕ taking $\bar{e}_1$,$\bar{e}_2$ to e'_1,e'_2 and leaving $e'_3 = \bar{e}_3$.
2. Rotate about e'_1 by an angle θ taking e'_2,e'_3 to e''_2,e''_3 and leaving $e''_1 = e'_1$.
3. Rotate about e''_3 by an angle ψ taking e''_1,e''_2 to e_1,e_2 and leaving $\boldsymbol{e} = e''_3$.

The angles ϕ, θ, ψ are sometimes called the angles of *precession*, *nutation*, and *spin*, respectively.

The intermediate rotations have the following matrix representations:

$$\mathbf{e}' = \bar{e} \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} = \bar{e} R_{e_3}(\phi)$$

$$\mathbf{e}'' = \mathbf{e}' \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} = \mathbf{e}' R_{e_1}(\theta)$$

and

$$\mathbf{e} = \mathbf{e}'' \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{e}'' R_{e_3}(\psi)$$

The composite rotation is

$$\mathbf{e} = \bar{\mathbf{e}} R_{e_3}(\phi) R_{e_1}(\theta) R_{e_3}(\psi) = \bar{\mathbf{e}} R(\phi, \theta, \psi)$$

The operator is represented by the fixed basis $\bar{\mathbf{e}}$ by using

$$\mathcal{R} = \mathbf{e}\bar{\boldsymbol{\epsilon}} = \bar{\mathbf{e}} R(\phi, \theta, \psi) \bar{\boldsymbol{\epsilon}}.$$

The matrix products yield

$$R(\phi, \theta, \psi) = \tag{1.5}$$
$$\begin{pmatrix} \cos\phi\cos\psi - \sin\phi\cos\theta\sin\psi & -\cos\phi\sin\psi - \sin\phi\cos\theta\cos\psi & \sin\phi\sin\theta \\ \sin\phi\cos\psi + \cos\phi\cos\theta\sin\psi & -\sin\phi\sin\psi + \cos\phi\cos\theta\cos\psi & -\cos\phi\sin\theta \\ \sin\theta\sin\psi & \sin\theta\cos\psi & \cos\theta \end{pmatrix}$$

Given a rotation parameterized by the rotation angle α and rotation axis $\mathbf{n}$, the Euler angle parameterization can be found as follows:

$$\begin{aligned} \cos\theta &= \mathcal{R}_{\mathbf{n}}(\alpha)(\bar{e}_3) \cdot \bar{e}_3 \\ \cos\phi &= (\bar{e}_3 \times \mathcal{R}_{\mathbf{n}}(\alpha)(\bar{e}_3)) \cdot \bar{e}_1 / \sin\theta \\ \cos\psi &= (\bar{e}_3 \times \mathcal{R}_{\mathbf{n}}(\alpha)(\bar{e}_3)) \cdot \mathcal{R}_{\mathbf{n}}(\alpha)(\bar{e}_2) / \sin\theta \end{aligned}$$

This conversion is not possible if $\sin\theta = 0$. In that case the Euler angle parameterizations is not unique – ϕ and ψ cannot be separated.

There are many ways to represent a rotation as successive rotations about intermediate axes. For example, in aeronautics and astronautics one uses a parameterization dubbed *yaw*, *pitch*, and *roll* defined by the following:

1. Rotate about $\bar{e}_3$ by the yaw angle ψ.
2. Rotate about e'_2 by the pitch angle θ.
3. Rotate about e''_1 by the roll angle ϕ.

In fact [3] there are 12 possible rotation sequences each defining a variant of the Euler angles. These are labeled by $n_1 - n_2 - n_3$ corresponding to the rotation

$$\mathbf{e} = \bar{\mathbf{e}} R_{\bar{e}_{n_1}} R_{e'_{n_2}} R_{e''_{n_3}}$$

The Euler angle parameterization discussed here would be labeled $3-1-3$ while yaw, pitch, and roll would be $3-2-1$. There are 27 possible $n_1 - n_2 - n_3$ combinations but it is necessary that $n_1 \neq n_2$ and $n_2 \neq n_3$ and 12 combinations survive

$$\begin{array}{ccc} 1-2-1 & 2-1-2 & 3-1-3 \\ 1-3-1 & 2-3-2 & 3-2-3 \\ 1-2-3 & 2-3-1 & 3-1-2 \\ 1-3-2 & 2-1-3 & 3-2-1 \end{array}$$

1.2 Quaternions

This section introduces a new number systems called *quaternions* which are intimately connected to three-dimensional rotations (we have already seen that complex numbers are intimately connected to plane rotations). As a prelude, the construction of more familiar number systems will be reviewed. Quaternions will then seem quite natural. A good reference for these matters is [4].

We begin with the integers, $\mathbb{Z}$. In $\mathbb{Z}$ we can add, multiply, and subtract but there is no division. In other words, there is no multiplicative inverse in $\mathbb{Z}$ of any integers other than ± 1. To obtain multiplicative inverses we need rational numbers, traditionally denoted by $\mathbb{Q}$ (for quotients). They are constructed as equivalence classes of ordered pairs of integers (n, m), $m \neq 0$, based on the equivalence relation

$$(n, m) \sim (p, q), \text{ if and only if } nq = mp$$

The equivalence classes are the sets

$$[(n, m)] = \{(p, q) | (p, q) \sim (n, m)\}$$

which is merely an abstract way to declare that rational numbers should be reduced to lowest terms. Having made the distinction between the pairs and the equivalence classes we will now dispense with the square brackets for notational simplicity. The operations of addition and multiplication of rational numbers are defined by

$$(n, m) + (p, q) = (nq + mp, mq)$$
$$(n, m)(p, q) = (np, mq)$$

Now every pair (n, m) has an additive inverse $(-n, m)$

$$(n, m) + (-n, m) = (0, m)$$

and, if $n \neq 0$ a multiplicative inverse (m, n)

$$(n, m)(m, n) = (1, 1)$$

This is quite transparent when translated to grade school notation

$$(n, m) \leftrightarrow \frac{n}{m}$$

The rational numbers are incomplete in the sense that there is no rational number which represents the hypotenuse of a right triangle having unit legs. The real numbers $\mathbb{R}$ are constructed from pairs of sets of rational numbers, called *Dedekind cuts*,

$$(D_1, D_2),\ D_1 \text{ and } D_2 \subset \mathbb{Q}$$

which have the following four properties: each $x \in D_1$ is less than every $y \in D_2$, D_1 has no maximum element,

$$x \in D_1 \text{ and } y < x \text{ then } y \in D_1$$

and

$$x \in D_2 \text{ and } y > x \text{ then } y \in D_2$$

For example, let

$$D_1 = \{r \in \mathbb{Q} | r^2 < 2\}$$

and let D_2 be the set complement of D_1 in $\mathbb{Q}$. The Dedekind cut (D_1, D_2) represents the real number $\sqrt{2}$.

The real numbers are incomplete in the sense that there is no solution in $\mathbb{R}$ to the polynomial equation

$$x^2 + 1 = 0$$

We obtain an algebraically complete number system, the complex numbers $\mathbb{C}$, from ordered pairs of real numbers

$$z = (x, y) \qquad x \text{ and } y \in \mathbb{R}$$

with addition and multiplication defined as

$$(x, y) + (u, v) = (x + u, y + v)$$

and

$$(x, y)(u, v) = (xu - yv, xv + yu)$$

There is also a new operation, multiplication of a complex number by a real number:

$$a(x, y) = (a, 0)(x, y) = (ax, ay)$$

This appears more familiar in high school notation

$$(x, y) \leftrightarrow x + \imath y$$

The Irish mathematician Rowan Hamilton struggled in vain to extend complex numbers to three dimensions. Eventually he realized that it is necessary to go to four dimensions and he invented a new number system called the quaternions. We will introduce quaternions, $\mathbb{H}$ for Hamilton, as ordered pairs of complex numbers. Although Hamilton did not use the ordered pair construction for quaternions, he was the inventor of the pair construction for complex numbers. A quaternion, then, is an ordered pair of complex numbers

$$\mathsf{q} = (z_1, z_2) \qquad z_1 \text{ and } z_2 \in \mathbb{C}$$

with addition and multiplication defined by

$$(z_1, z_2) + (w_1, w_2) = (z_1 + w_1, z_2 + w_2)$$
$$(z_1, z_2)(w_1, w_2) = (z_1 w_1 - z_2 w_2^*, z_1 w_2 + z_2 w_1^*)$$

and

$$a(z, w) = (a, 0)(z, w) = (az, aw) \qquad a \in \mathbb{R}$$

It turns out that multiplication is not commutative. That is, in general

$$\mathsf{rq} \neq \mathsf{qr}$$

We shall denote quaternions with *sans serif* typeface as in the last equation.

This construction by pairs ties in nicely with the constructions of the rational, real, and complex numbers but is not the traditional approach. If we single out three special pairs and attach Hamilton's notation to them

$$\mathsf{i} = (\imath, 0) \qquad \mathsf{j} = (0, 1) \qquad \mathsf{k} = (0, \imath)$$

and identify

$$(a, 0) \leftrightarrow \mathsf{a} \qquad a \in \mathbb{R}$$

then we find

$$(q_0 + \imath q_1, q_2 + \imath q_3) = q_0 + q_1\mathsf{i} + q_2\mathsf{j} + q_3\mathsf{k} \qquad q_i \in \mathbb{R}$$

which is the form Hamilton used to express quaternions. This form makes it quite clear that quaternions are a four-dimensional generalization of complex numbers.

The quaternions i, j,k satisfy the following relations:

$$\mathsf{i}^2 = \mathsf{j}^2 = \mathsf{k}^2 = \mathsf{ijk} = -1 \tag{1.6}$$

These are the rules Hamilton inscribed on Broome Bridge in Dublin on October 16, 1843 [5].

In the language of abstract algebra, the quaternions form a noncommutative, normed division algebra over $\mathbb{R}$. The eight-dimensional octonians $\mathbb{O}$ [5] can be constructed from pairs of quaternions but there the chain ends. The only normed division algebras over $\mathbb{R}$ are $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathbb{O}$.

1.2.1 Quaternion Algebra

This section summarizes the essentials of the algebra of quaternions. First we express the basic operations in the i, j,k form. Addition becomes

$$\begin{aligned}(q_0 + q_1\mathsf{i} + q_2\mathsf{j} + q_3\mathsf{k}) + (p_0 + p_1\mathsf{i} + p_2\mathsf{j} + p_3\mathsf{k})\\ = (q_0 + p_0) + (q_1 + p_1)\mathsf{i} + (q_2 + p_2)\mathsf{j} + (q_3 + p_3)\mathsf{k}\end{aligned}$$

and multiplication becomes

$$\begin{aligned}(q_0 + q_1\mathsf{i} + q_2\mathsf{j} + q_3\mathsf{k})(p_0 + p_1\mathsf{i} + p_2\mathsf{j} + p_3\mathsf{k}) = {} & q_0p_0 - q_1p_1 - q_2p_2 - q_3p_3\\ & + (q_0p_1 + q_1p_0 + q_2p_3 - q_3p_2)\mathsf{i}\\ & + (q_0p_2 + q_2p_0 + q_3p_1 - q_1p_3)\mathsf{j}\\ & + (q_0p_3 + q_3p_0 + q_1p_2 - q_2p_1)\mathsf{k}\end{aligned}$$

There are many patterns in this last expression and they will become clear by the end of the next section.

Given $\mathsf{q} = q_0 + q_1\mathsf{i} + q_2\mathsf{j} + q_3\mathsf{k}$, q_0 is called the *scalar part* of q

$$S(\mathsf{q}) = q_0$$

and $q_1\mathsf{i} + q_2\mathsf{j} + q_3\mathsf{k}$ is called the *vector part* of q

$$V(\mathsf{q}) = q_1\mathsf{i} + q_2\mathsf{j} + q_3\mathsf{k}$$

The *conjugate* of q is

$$\bar{\mathsf{q}} = q_0 - q_1\mathsf{i} - q_2\mathsf{j} - q_3\mathsf{k}$$

The square of the *norm* of q is

$$|\mathsf{q}|^2 = \mathsf{q}\bar{\mathsf{q}} = q_0^2 + q_1^2 + q_2^2 + q_3^2 \in \mathbb{R}$$

where we have taken the liberty of identifying a real number a and the quaternion q with $S(\mathsf{q}) = a$, $V(\mathsf{q}) = \mathbf{0}$. The *inverse* of q is, when $|\mathsf{q}| \neq 0$,

$$\mathsf{q}^{-1} = \frac{1}{|\mathsf{q}|^2}\bar{\mathsf{q}}$$

Clearly $\mathsf{q}\mathsf{q}^{-1} = 1 + 0\mathsf{i} + 0\mathsf{j} + 0\mathsf{k}$. Note also that $\overline{\mathsf{q}\mathsf{p}} = \bar{\mathsf{p}}\bar{\mathsf{q}}$ and $(\mathsf{q}\mathsf{p})^{-1} = \mathsf{p}^{-1}\mathsf{q}^{-1}$.

1.2.2 Quaternions as Scalar–Vector Pairs

At this point we have two ways to express quaternions – the complex number pair form and the i, j, k form. There is a third way which is convenient for anyone accustomed to vector algebra. This third way represents quaternions as scalar–vector pairs

$$q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k} \leftrightarrow (q_0, \mathbf{v}) \qquad \text{with} \quad \mathbf{v} = q_1\mathbf{e}_1 + q_2\mathbf{e}_2 + q_3\mathbf{e}_3$$

In this way, the quaternion algebra is expressed as

$$\begin{aligned} (q_0, \mathbf{u}) + (p_0, \mathbf{v}) &= (q_0 + p_0, \mathbf{u} + \mathbf{v}) \\ (q_0, \mathbf{u})(p_0, \mathbf{v}) &= (q_0 p_0 - \mathbf{u} \cdot \mathbf{v}, q_0\mathbf{v} + p_0\mathbf{u} + \mathbf{u} \times \mathbf{v}) \\ \overline{(q_0, \mathbf{u})} &= (q_0, -\mathbf{u}) \\ |(q_0, \mathbf{u})|^2 &= q_0^2 + \mathbf{u} \cdot \mathbf{u} \\ (q_0, \mathbf{u})^{-1} &= \frac{1}{q_0^2 + \mathbf{u} \cdot \mathbf{u}}(q_0, -\mathbf{u}) \end{aligned}$$

A perusal of the multiplication formula now reveals the patterns to be elements of the scalar and vector products.

A *pure quaternion* $(0, \mathbf{u})$ is one with zero scalar part and a *unit quaternion* has unit norm $\mathrm{q}\bar{\mathrm{q}} = 1$. The product of two pure quaternions encapsulates the scalar and vector products

$$(0, \mathbf{u})(0, \mathbf{v}) = (-\mathbf{u} \cdot \mathbf{v}, \mathbf{u} \times \mathbf{v})$$

1.2.3 Quaternions as Matrices

Complex numbers can be represented as 2×2 matrices with real entries. Let $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and note that $J^2 = -I$. Then $Z = aI + bJ$ and $W = cI + dJ$ satisfy the rules for addition and multiplication of complex numbers

$$Z + W = (a + c)I + (b + d)J$$

$$ZW = (aI + bJ)(cI + dJ) = (ac - bd)I + (ad + bc)J$$

Similarly, quaternions can be represented as 2×2 matrices of complex numbers. If

$$z_1 = q_0 + \imath q_1 \qquad z_2 = q_2 + \imath q_3$$

$$w_1 = r_0 + \imath r_1 \qquad w_2 = r_2 + \imath r_3$$

then the correspondences

$$Q_1 = \begin{pmatrix} z_1 & z_2 \\ -z_2^* & z_1^* \end{pmatrix} \leftrightarrow q_0 + q_1\mathsf{i} + q_2\mathsf{j} + q_3\mathsf{k}$$

and

$$Q_2 = \begin{pmatrix} w_1 & w_2 \\ -w_2^* & w_1^* \end{pmatrix} \leftrightarrow r_0 + r_1\mathsf{i} + r_2\mathsf{j} + r_3\mathsf{k}$$

are preserved by the matrix product

$$Q_1 Q_2 \leftrightarrow (q_0 + q_1\mathsf{i} + q_2\mathsf{j} + q_3\mathsf{k})(r_0 + r_1\mathsf{i} + r_2\mathsf{j} + r_3\mathsf{k})$$

The quaternion product is linear in the components of each factor. This allows us to express the quaternion operations in terms of linear operations on $\mathbb{R}^4$. If we associate

$$\mathsf{q} = q_0 + q_1\mathsf{i} + q_2\mathsf{j} + q_3\mathsf{k} \leftrightarrow q = \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix}$$

and

$$\mathsf{p} = p_0 + p_1\mathsf{i} + p_2\mathsf{j} + p_3\mathsf{k} \leftrightarrow p = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

then left multiplication by q is represented by

$$L(\mathsf{q}) = \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix} \tag{1.7}$$

that is,

$$\mathsf{qp} \leftrightarrow L(\mathsf{q})p$$

This is expressed explicitly in a component form

$$\begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix} \cdot \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} q_0p_0 - q_1p_1 - q_2p_2 - q_3p_3 \\ q_0p_1 + q_1p_0 + q_2p_3 - q_3p_2 \\ q_0p_2 + q_2p_0 + q_3p_1 - q_1p_3 \\ q_0p_3 + q_3p_0 + q_1p_2 - q_2p_1 \end{pmatrix}$$

Right multiplication by q is represented by

$$R(\mathsf{q}) = \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & q_3 & -q_2 \\ q_2 & -q_3 & q_0 & q_1 \\ q_3 & q_2 & -q_1 & q_0 \end{pmatrix} \tag{1.8}$$

and

$$\mathsf{pq} \leftrightarrow R(\mathsf{q})p$$

These are the representations of the left and right translations L_{q} and R_{q} on the Lie group $S^3 \cong \mathbb{H}_1$ (Appendix C).

Since $L_{\mathsf{p}} L_{\mathsf{p}} = L_{\mathsf{pq}}$, it follows that

$$L(\mathsf{p})L(\mathsf{q}) = L(\mathsf{pq})$$

There is the analogous result for right multiplication

$$R(\mathsf{p})R(\mathsf{q}) = R(\mathsf{qp})$$

$L(\mathsf{p})$ and $R(\mathsf{p})$ have the following properties:

$$\begin{aligned} L(\mathsf{q})p &= -L(\mathsf{p})q \text{ if } q^t p = 0 \\ R(\mathsf{q})p &= -R(\mathsf{p})q \text{ if } q^t p = 0 \\ L(\bar{\mathsf{q}}) &= L(\mathsf{q})^t \\ L(\bar{\mathsf{q}})L(\mathsf{q}) &= (q^t q)I \\ R(\bar{\mathsf{q}})R(\mathsf{q}) &= (q^t q)I \end{aligned}$$

If $\mathsf{q} \in \mathbb{H}_1$ then $R(q)$ and $L(q)$ are orthogonal matrices. These relations are discussed and applied in [6,7]. This correspondence is a generalization of the $\widehat{v} \leftrightarrow \overrightarrow{M}$ correspondence of $\mathbb{R}^3$ and $\mathfrak{so}(3)$.

1.2.4 Rotations via Unit Quaternions

Quaternions are useful in rigid body dynamics because they provide succinct expressions for quantities which can be rather unwieldy in matrix form. A good example is the composition of rotations which is a simple quaternion product. Rotations are effected in the quaternions by means of unit quaternions. We denote the set of all unit quaternions by $\mathbb{H}_1$.

Let $\mathsf{r} = (0, \mathbf{v})$ be a pure quaternion and $\mathsf{q} = (q_0, \mathbf{q})$ be a unit quaternion. Let

$$\mathsf{r}' = \mathsf{qrq}^{-1} = \mathsf{qr}\bar{\mathsf{q}}$$

Now we compute

$$\mathsf{r}' = (0, [q_0^2 - \mathbf{q} \cdot \mathbf{q}]\mathbf{v} + 2\mathbf{v} \cdot \mathbf{qq} + 2q_0 \mathbf{q} \times \mathbf{v})$$

and observe that if an angle θ and a unit vector $\mathbf{n}$ are defined by

$$q_0 = \cos\frac{1}{2}\theta \quad \text{and} \quad \mathbf{q} \equiv (q_1, q_2, q_3) = \sin\frac{1}{2}\theta\,\mathbf{n}$$

then

$$r' = (0, [1 - \cos\theta]\mathbf{n}\mathbf{n}\cdot\mathbf{v} + \cos\theta\mathbf{v} + \sin\theta\mathbf{n}\times\mathbf{v})$$

which reproduces the effect of $\mathcal{R}_{\mathbf{n}}(\theta)$ on the vector part of r. The q_i arise most naturally in the context of quaternions. They were, however, introduced much earlier by Euler and Rodrigues. For that reason they are sometimes called the *Euler parameters* or *Euler–Rodrigues parameters*.

Now it is natural to ask how the rotation matrix and the quaternion are related. First, let us express the rotation matrix $R = R_n(\theta)$ explicitly in terms of the components of $\mathbf{n}$ and the rotation angle via $c = \cos\theta$, $s = \sin\theta$,

$$R = \begin{pmatrix} (1-c)n_1^2 + c & (1-c)n_1 n_2 - s n_3 & (1-c)n_1 n_3 + s n_2 \\ (1-c)n_2 n_1 + s n_3 & (1-c)n_2^2 + c & (1-c)n_2 n_3 - s n_1 \\ (1-c)n_3 n_1 - s n_2 & (1-c)n_3 n_2 + s n_1 & (1-c)n_3^2 + c \end{pmatrix} \quad (1.9)$$

Then, to express R in terms of the quaternion components q_i, we use the identities

$$c = q_0^2 - q_1^2 - q_2^2 - q_3^2 \qquad 1 - c = 2(1 - q_0^2) \qquad (1-c)n_i n_j = 2q_i q_j$$

and

$$s\mathbf{n} = 2q_0\boldsymbol{q}.$$

Thus

$$R = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1 q_2 - q_0 q_3) & 2(q_1 q_3 + q_0 q_2) \\ 2(q_2 q_1 + q_0 q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2 q_3 - q_0 q_1) \\ 2(q_3 q_1 - q_0 q_2) & 2(q_3 q_2 + q_0 q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix} \quad (1.10)$$

Note that quaternions q and $-q$ correspond the same rotation matrix.

Apart from its intrinsic interest, this expression is of practical importance because it is a purely algebraic version of the rotation matrix. No transcendental functions are involved. This is relevant when computing requirements are an issue.

The rotation defined by quaternion $(q_0, \boldsymbol{q})$ can also be expressed in the axis-angle form as

$$\mathcal{R}_{\mathbf{n}}(\theta) = 2\boldsymbol{q}\otimes\boldsymbol{q} + (q_0^2 - \boldsymbol{q}\cdot\boldsymbol{q})I + 2q_0\widehat{\boldsymbol{q}} \quad \boldsymbol{q} = \sin\frac{1}{2}\theta\,\mathbf{n}$$

The process can be reversed. That is, given a rotation matrix, R, one can construct the components of the associated quaternion. Let $\tau = \mathrm{Trace}(R)$ – the sum of the diagonal elements of R. From (1.9) we find

$$\tau = 1 + 2\cos\theta \quad \text{or} \quad \cos\theta = \frac{1}{2}(\tau - 1) \quad (1.11)$$

and

$$q_0 = \cos\frac{1}{2}\theta = \frac{1}{2}\sqrt{\tau + 1}. \tag{1.12}$$

The vector part of the quaternion is found from the off-diagonal elements of R

$$q_1 = \frac{1}{4q_0}(R_{32} - R_{23}) \tag{1.13}$$

$$q_2 = \frac{1}{4q_0}(R_{13} - R_{31}) \tag{1.14}$$

$$q_3 = \frac{1}{4q_0}(R_{21} - R_{12}) \tag{1.15}$$

This may be expressed in matrix notation as

$$R - R^t = 4q_0\widehat{V(\mathsf{q})}$$

Quaternions also handle reflections concisely. To reflect through the plane with unit normal $\mathbf{n}$ we have

$$(0, \mathbf{n})(0, -\boldsymbol{r})(0, -\mathbf{n}) = (0, \boldsymbol{r} - 2\mathbf{nn} \cdot \boldsymbol{r})$$

Conjugation applied to a pure quaternion is the same as negation. Therefore the reflection operation is expressed in a slightly simpler form as

$$(0, \mathbf{n})(0, \boldsymbol{r})(0, \mathbf{n}) = (0, \boldsymbol{r} - 2\mathbf{nn} \cdot \boldsymbol{r})$$

Let two reflections be defined by the unit vectors $\mathbf{n}$ and $\boldsymbol{m}$. The rotation determined by the composition of the two reflections is simply

$$(0, \mathbf{n})(0, \boldsymbol{m}) = (-\mathbf{n} \cdot \boldsymbol{m}, -\mathbf{n} \times \boldsymbol{m}) = -(\cos\theta, \sin\theta\,\frac{\mathbf{n} \times \boldsymbol{m}}{\|\mathbf{n} \times \boldsymbol{m}\|})$$

where θ is the angle between $\mathbf{n}$ and $\boldsymbol{m}$. Therefore the angle of rotation is twice the angle between the unit vectors and the axis of rotation is parallel to $\mathbf{n} \times \boldsymbol{m}$ (recall that rotations defined by quaternions q and $-q$ are the same).

1.2.5 Composition of Rotations

The composition of two rotations is itself a rotation because the product of two elements of $SO(3)$ is itself in $SO(3)$. Now we wish to inquire about the rotation axis and rotation angle for the composite of two rotations.

Consider two rotations with axes specified by unit vectors $\mathbf{a}$, $\mathbf{b}$ and angles by α, β. We wish to find the axis and angle for the rotation composed of $\mathcal{R}_{\mathbf{a}}(\alpha)$ followed by $\mathcal{R}_{\mathbf{b}}(\beta)$. Construct the spherical triangle $\mathbf{abc}$ on the unit sphere

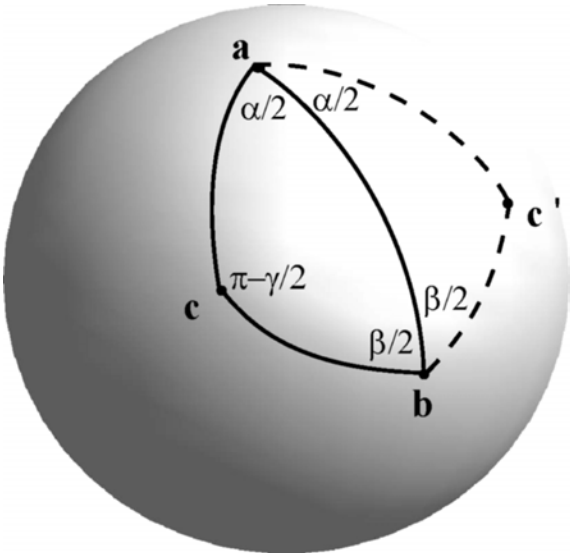

Fig. 1.3 Defining figure for the Rodrigues spherical triangle.

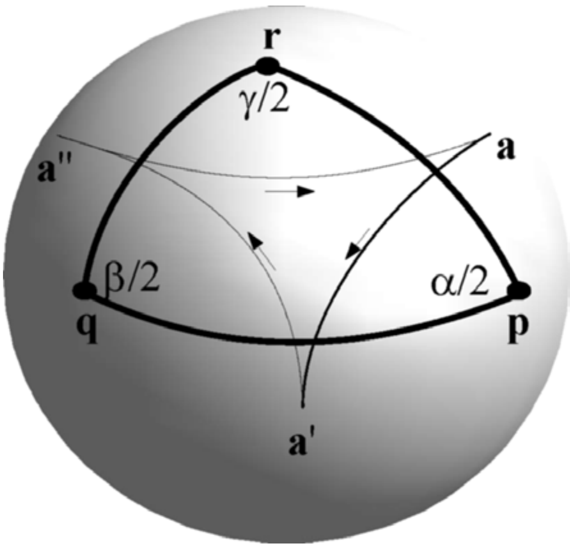

Fig. 1.4 A sequence of finite rotations taking point a to a', then to a'' and then back to a illustrating the Rodrigues triangle.

shown in Fig. 1.3. The angles **a** and **b** are half the rotation angles and the orientation is such that rotation about **a** would bring arc **ac** toward arc **ab**. We will refer to this triangle as the *Rodrigues triangle* after Olinde Rodrigues who discovered the construction [8].

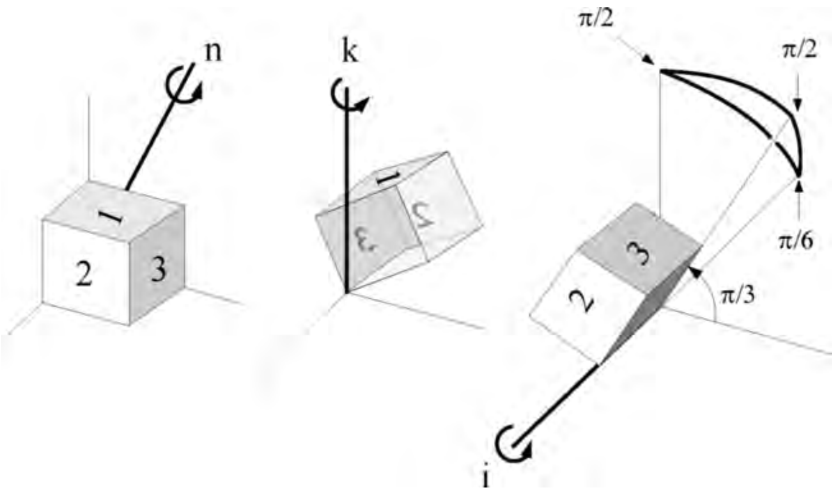

Fig. 1.5 A sequence of finite rotations of a cube: 180° about $\mathbf{n} = [0\ 1/2\ \sqrt{3}/2]$ followed by 180° about $\boldsymbol{k} = [0\ 0\ 1]$. The composite rotation is about $\boldsymbol{i} = [1\ 0\ 0]$ by 60° as the shown by the Rodrigues spherical triangle displayed in the last panel.

The rotation axis and angle for the composite rotation can be obtained by an elegant geometrical argument [9–11]. Rotate arc **ac** about **a** by angle α taking point **c** to $\mathbf{c}'$. Then rotate arc $\mathbf{bc}'$ about axis **b** by angle β. This returns $\mathbf{c}'$ to **c** thus showing that **c** is fixed under the composite rotation and must lie on the axis of rotation. In more formal terms, consider the reflections $\mathcal{M}_{\mathbf{a}'}$, $\mathcal{M}_{\mathbf{b}'}$, and $\mathcal{M}_{\mathbf{c}'}$ (using the polar triangle notation established in Appendix A). We know that reflections preserve lengths of vectors and therefore reflections map the unit sphere into itself. The relation between reflections and rotations provides

$$\mathcal{R}_{\mathbf{a}}(\alpha) = \mathcal{M}_{\mathbf{c}'}\mathcal{M}_{\mathbf{b}'}$$

$$\mathcal{R}_{\mathbf{b}}(\beta) = \mathcal{M}_{\mathbf{a}'}\mathcal{M}_{\mathbf{c}'}$$

and therefore

$$\begin{aligned} \mathcal{R}_{\mathbf{b}}(\beta)\mathcal{R}_{\mathbf{a}}(\alpha) &= \mathcal{M}_{\mathbf{a}'}\mathcal{M}_{\mathbf{c}'}\mathcal{M}_{\mathbf{c}'}\mathcal{M}_{\mathbf{b}'} \\ &= \mathcal{M}_{\mathbf{a}'}\mathcal{M}_{\mathbf{b}'} \\ &= \mathcal{R}_{-\mathbf{c}}(2\pi - \gamma) \\ &= \mathcal{R}_{\mathbf{c}}(\gamma) \end{aligned}$$

It is instructive to use quaternions to calculate the composite rotation angle. Let the rotations be represented by the unit quaternions $\mathsf{p} = (\cos\frac{1}{2}\alpha, \sin\frac{1}{2}\alpha\mathbf{a})$ and $\mathsf{q} = (\cos\frac{1}{2}\beta, \sin\frac{1}{2}\beta\mathbf{b})$ and let $\mathbf{a}\cdot\mathbf{b} = \cos\phi$. The effect of rotation 1 followed by rotation 2 is

$$\mathsf{r}' = \mathsf{q}(\mathsf{prp}^{-1})\mathsf{q}^{-1} = (\mathsf{qp})\mathsf{r}(\mathsf{qp})^{-1}$$

so that the composition is obtained from the product $\mathsf{r} = \mathsf{qp}$

$$\begin{aligned} \mathsf{qp} &= \Big(\cos\frac{1}{2}\alpha\cos\frac{1}{2}\beta - \sin\frac{1}{2}\alpha\sin\frac{1}{2}\beta\cos\phi, \\ &\cos\frac{1}{2}\alpha\sin\frac{1}{2}\beta\,\mathbf{b} + \cos\frac{1}{2}\beta\sin\frac{1}{2}\alpha\,\mathbf{a} + \sin\frac{1}{2}\alpha\sin\frac{1}{2}\beta\,\mathbf{a}\times\mathbf{b}\Big) \end{aligned} \tag{1.16}$$

Thus

$$\cos\frac{1}{2}\gamma = S(\mathsf{qp}) = \cos\frac{1}{2}\alpha\cos\frac{1}{2}\beta - \sin\frac{1}{2}\alpha\sin\frac{1}{2}\beta\cos\phi$$

Now, referring to Appendix A, Eq. (A.15), we find that

$$\cos C = -\cos\frac{1}{2}\alpha\cos\frac{1}{2}\beta + \sin\frac{1}{2}\alpha\sin\frac{1}{2}\beta\cos\phi$$

Therefore

$$\cos\frac{1}{2}\gamma = -\cos C$$

or

$$C = \pi - \frac{1}{2}\gamma$$

and the construction is validated. With γ in hand, the axis of the composite rotation is simply

$$\mathbf{n} = V(\mathsf{qp}) / \sin\frac{1}{2}\gamma$$

The Rodrigues triangle construction may also be given by the following equivalent description. Refer to the spherical triangle constructed in Fig. 1.4.

Given a spherical triangle with vertices **p**, **q**, and **r** and vertex angles $\alpha/2$, $\beta/2$ and $\gamma/2$ successive rotations about **p** by angle α, **q** by β, and **r** by γ yield the identity transformation.

$$\begin{aligned} \mathcal{R}_{\mathbf{r}}(\gamma)\mathcal{R}_{\mathbf{q}}(\beta)\mathcal{R}_{\mathbf{p}}(\alpha) &= \mathcal{M}_{\mathbf{q}'}\mathcal{M}_{\mathbf{p}'}\mathcal{M}_{\mathbf{p}'}\mathcal{M}_{\mathbf{r}'}\mathcal{M}_{\mathbf{r}'}\mathcal{M}_{\mathbf{q}'} \\ &= I \end{aligned}$$

Example 1.2 *An illustration of the composition of rotations of a cube and the associated Rodrigues spherical triangle is shown in Fig. 1.5. The analytic expression of that composition is*

$$\mathcal{R}_{\mathbf{k}}(\pi)\mathcal{R}_{\mathbf{n}}(\pi) = \mathcal{R}_{-\mathbf{i}}(2\pi - \pi/3) = \mathcal{R}_{\mathbf{i}}(\pi/3)$$

Note the orientation of the Rodrigues triangle. $\diamondsuit$

We close this section with an application of the Rodrigues triangle. Let **a**, **b**, and $\mathbf{b}' = \mathcal{R}_{\mathbf{a}}(\alpha)\mathbf{b}$ be points on the unit sphere. Construct the spherical triangles shown in Fig. 1.6. If each triangle is interpreted as a Rodrigues triangle for composition of rotations, we obtain the relation

$$\mathcal{R}_{\mathbf{a}}(\alpha)\mathcal{R}_{\mathbf{b}}(\beta) = \mathcal{R}_{\mathbf{b}'}(\beta)\mathcal{R}_{\mathbf{a}}(\alpha) = \mathcal{R}_{\mathbf{c}}(\gamma)$$

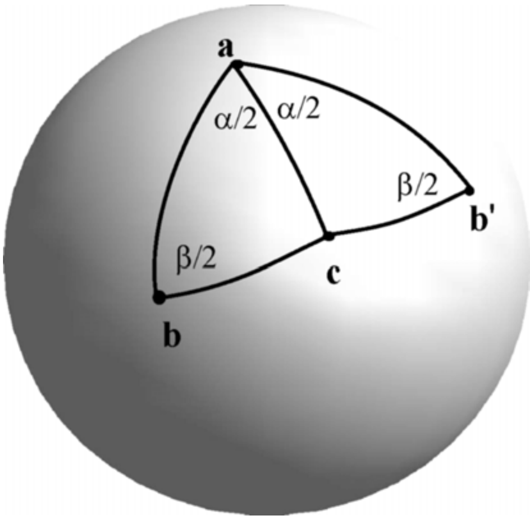

Fig. 1.6 Transposition of rotations about $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{b}' = \mathcal{R}_{\mathbf{a}}(\alpha)\mathbf{b}$.

This is referred to as *Rodrigues' transposition of rotations* [3, 12] or *rotation reversal* [13].

Example 1.3 *This relation can be used to clarify the order of factors occurring in the operator and matrix versions of the Euler angle parameterization. In operator notation,*

$$\mathcal{R} = \mathcal{R}_{\mathbf{e}_3''}(\psi)\mathcal{R}_{\mathbf{e}_1'}(\theta)\mathcal{R}_{\bar{\mathbf{e}}_3}(\phi)$$

which can be subjected to transpositions to obtain

$$\mathcal{R} = \mathcal{R}_{\mathbf{e}_1'}(\theta)\mathcal{R}_{\mathbf{e}_3'}(\psi)\mathcal{R}_{\bar{\mathbf{e}}_3}(\phi) = \mathcal{R}_{\mathbf{e}_1'}(\theta)\mathcal{R}_{\bar{\mathbf{e}}_3}(\phi)\mathcal{R}_{\bar{\mathbf{e}}_3}(\psi) = \mathcal{R}_{\bar{\mathbf{e}}_3}(\phi)\mathcal{R}_{\bar{\mathbf{e}}_1}(\theta)\mathcal{R}_{\bar{\mathbf{e}}_3}(\psi)$$

This leads immediately, in the notation of Section 1.1.6, to the representation

$$\mathcal{R} = \bar{\mathbf{e}}R_{e_3}(\phi)R_{e_1}(\theta)R_{e_3}(\psi)\bar{\mathbf{e}}. \quad \diamondsuit$$

1.3 Complex Numbers

In this section we describe rotations parameterized by Möbius transformations of the complex plane. The development begins by recasting the representation of quaternions as 2×2 matrices with complex entries. Section 1.2.3 examined a mapping from the quaternion components to the complex entries in the matrices. Here we explore a different mapping which will lead to interesting connections between rotations and complex functions.[4]

4) There are actually 24 such mappings which are isomorphisms.

1.3.1 Cayley–Klein Parameters

Define the following matrices.[5]

$$J_1 = \begin{pmatrix} 0 & \imath \\ \imath & 0 \end{pmatrix} \qquad J_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad J_3 = \begin{pmatrix} \imath & 0 \\ 0 & -\imath \end{pmatrix}$$

These matrices obey the quaternion multiplication rules (1.6), namely $J_1^2 = J_2^2 = J_3^2 = J_1 J_2 J_3 = -I$, and the quaternions are mapped to the set of real, linear combinations of I, J_1, J_2, J_3,

$$q_0 + q_1\mathsf{i} + q_2\mathsf{j} + q_3\mathsf{k} \quad \leftrightarrow \quad q_0 I + q_1 J_1 + q_2 J_2 + q_3 J_3 = \begin{pmatrix} q_0 + \imath q_3 & -q_2 + \imath q_1 \\ q_2 + \imath q_1 & q_0 - \imath q_3 \end{pmatrix}$$

It is easy to show that if $\mathsf{p}, \mathsf{q} \in \mathbb{H}$ are mapped to P, Q, then $\mathsf{p} + \mathsf{q}$ and pq are mapped to $P + Q$ and PQ. Conjugation in the quaternions corresponds to taking the adjoint of the associated matrix

$$\bar{\mathsf{q}} \leftrightarrow Q^\dagger = (Q^t)^*$$

In other words, $\mathbb{H}$ is isomorphic to the algebra over $\mathbb{R}$ of complex 2×2 matrices of the form

$$\begin{pmatrix} u & v \\ -v^* & u^* \end{pmatrix}$$

Now if

$$Q = \begin{pmatrix} q_0 + \imath q_3 & -q_2 + \imath q_1 \\ q_2 + \imath q_1 & q_0 - \imath q_3 \end{pmatrix}, \qquad \mathsf{q} = q_0 + q_1\mathsf{i} + q_2\mathsf{j} + q_3\mathsf{k}$$

then

$$\det Q = q_0^2 + q_1^2 + q_2^2 + q_3^2 = |\mathsf{q}|^2$$

which shows that the subalgebra of unit quaternions, $\mathbb{H}_1$, is isomorphic to the subgroup of $GL(2, \mathbb{C})$ in which each element satisfies $QQ^\dagger = I$ and $\det Q = 1$. This group is called $SU(2)$, the group of *special unitary matrices* of order 2.

Given the correspondence between $SU(2)$ and $\mathbb{H}_1$ it follows that rotations can be represented in $SU(2)$ as follows. Let X be the matrix corresponding to the pure quaternion $\mathsf{r} = (0, \boldsymbol{r}), \boldsymbol{r} = (x\ y\ x)$ and let $\mathsf{q} \in \mathbb{H}_1$. Then

$$\mathsf{r} \leftrightarrow \begin{pmatrix} \imath z & -y + \imath x \\ y + \imath x & -\imath z \end{pmatrix} = X$$

and

$$\mathsf{q}\mathsf{r}\bar{\mathsf{q}} \leftrightarrow QXQ^\dagger$$

5) These matrices are related to the Pauli matrices σ_i of quantum physics by $J_1 = \imath\sigma_1$, $J_2 = -\imath\sigma_2$, and $J_3 = \imath\sigma_3$.

In the context of rigid body mechanics the elements of the matrices in $SU(2)$ are called the *Cayley–Klein parameters*. These parameters can be expressed in terms of the Euler angles. From Eqs. (1.5) and (1.12–1.15) we obtain

$$4q_0^2 = \tau + 1 = (1+\cos\theta)(1+\cos(\phi+\psi)) = 4\cos^2\frac{1}{2}\theta\,\cos^2\frac{1}{2}(\phi+\psi)$$

and

$$\begin{aligned}4q_0q_1 &= \sin\theta(\cos\phi+\cos\psi)\\ &= \sin\theta\left(\cos\left[\frac{1}{2}(\phi+\psi)+\frac{1}{2}(\phi-\psi)\right]+\cos\left[\frac{1}{2}(\phi+\psi)-\frac{1}{2}(\phi-\psi)\right]\right)\\ &= 4\cos\frac{1}{2}\theta\,\sin\frac{1}{2}\theta\,\cos\frac{1}{2}(\phi+\psi)\,\cos\frac{1}{2}(\phi-\psi)\end{aligned}$$

Thus

$$q_0 = \cos\frac{1}{2}\theta\cos\frac{1}{2}(\phi+\psi) \tag{1.17}$$

$$q_1 = \sin\frac{1}{2}\theta\cos\frac{1}{2}(\phi-\psi) \tag{1.18}$$

and similarly

$$q_2 = \sin\frac{1}{2}\theta\sin\frac{1}{2}(\phi-\psi) \tag{1.19}$$

$$q_3 = \cos\frac{1}{2}\theta\sin\frac{1}{2}(\phi+\psi) \tag{1.20}$$

With the identifications

$$\alpha = q_0 + \imath q_3,\;\; \beta = -q_2 + \imath q_1,\;\; \gamma = -\beta^*,\;\; \delta = \alpha^*$$

the $SU(2)$ representation of the rotation matrix becomes

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix} = \begin{pmatrix}\cos\frac{1}{2}\theta e^{\frac{1}{2}\imath(\phi+\psi)} & \imath\sin\frac{1}{2}\theta e^{\frac{1}{2}\imath(\phi-\psi)}\\ \imath\sin\frac{1}{2}\theta e^{-\frac{1}{2}\imath(\phi-\psi)} & \cos\frac{1}{2}\theta e^{-\frac{1}{2}\imath(\phi+\psi)}\end{pmatrix}$$

1.3.2
Rotations and the Complex Plane

We have considered rotations in the $SO(3)$, $\mathbb{H}_1$, and $SU(2)$ settings. The last setting we wish to consider is the complex plane with point at infinity. This is the so-called *Riemann sphere*, $\widehat{\mathbb{C}}$, the simplest of the compact Riemann surfaces. There is a one–one map, called the *stereographic projection*, from the unit sphere to $\widehat{\mathbb{C}}$. Imagine the complex plane intersecting a unit sphere in its equator. To map the point $r = (\xi\ \eta\ \zeta)$ on the sphere to the complex plane, one draws a

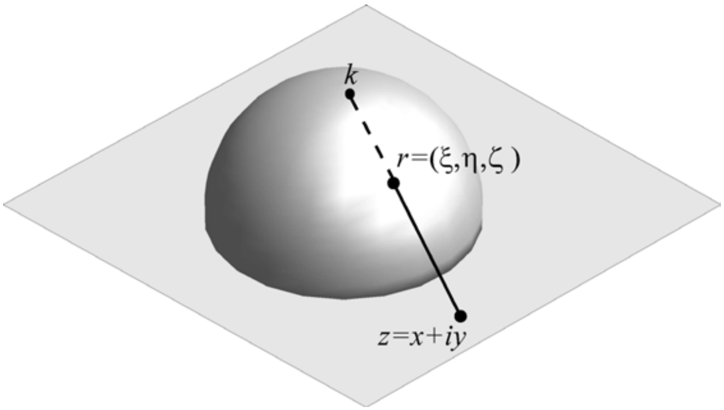

Fig. 1.7 Stereographic projection from the unit sphere to the complex plane.

line from the north pole of the sphere, say k, through the point r to the place z where the line intersects the complex plane (Fig. 1.7)

$$z = \mathrm{St}(\xi, \eta, \zeta) = \frac{\xi + \imath\eta}{1 - \zeta} = x + \imath y \tag{1.21}$$

Note that points on the equator $(\zeta = 0)$ are fixed points of the map, k $(\zeta = 1)$ is mapped to the point at infinity, and $-k$ $(\zeta = -1)$ is mapped to the origin. The upper hemisphere is mapped outside the unit circle and the lower hemisphere is mapped inside the unit circle.

The inverse map is

$$r = \mathrm{St}^{-1}(x + \imath y) = \left(\frac{2x}{x^2 + y^2 + 1}, \frac{2y}{x^2 + y^2 + 1}, \frac{x^2 + y^2 - 1}{x^2 + y^2 + 1} \right) \tag{1.22}$$

A rotation of $\mathbb{R}^3$ induces a single-valued map of the unit sphere into itself and also induces, via the stereographic projection, a map of the complex plane into itself with the same properties. The images of the poles of the rotation will be fixed points of the map. The orbits of the rotation on the sphere are circles in planes normal to the axis of rotation. These circles are mapped to circles in the complex plane (Fig. 1.8). Note that if the circles are traversed clockwise (seen from the exterior of the sphere) their images will be traversed counterclockwise.

We now sketch the argument that the transformation $z \to z'$ induced by a rotation must take the form [14]

$$\frac{z' - n_+}{z' - n_-} = C \frac{z - n_+}{z - n_-}$$

where $z' = f(z)$ is the image of z induced by the rotation and we have denoted by n_+ and n_- the images of the poles of the rotation, i.e., the intersections of

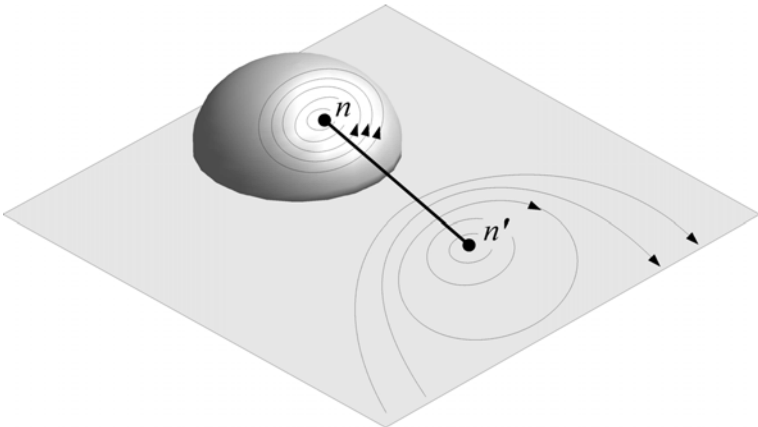

Fig. 1.8 Rotation of the sphere about $\boldsymbol{n}$ induces, via stereographic projection of the orbits, a map of the complex plane with the fixed point $\boldsymbol{n}'$.

the axis of rotation **n** with the sphere. The details of the argument can be found in the beautifully rendered [11]. We begin by examining the transformation induced by a reflection since any rotation is composed of two reflections. The simplest case is reflection in the equatorial plane. If $z = \mathrm{St}(\boldsymbol{x})$, then

$$\mathrm{St}(\mathcal{M}_{\mathsf{k}}(\boldsymbol{x})) = 1/z^*$$

This transformation is inversion in the unit circle. It can be shown that stereographic projection maps circles on the sphere to circles in the plane and that it preserves angles of intersection (angles between tangents at the point of intersection). Inversion in the unit circle generalizes to inversion in any circle, say radius R and center a, and the mapping is $\mathrm{Inv}_{R,a}$,

$$\mathrm{Inv}_{R,a}(z) = a + \frac{R^2}{z^* - a^*}$$

Inv also preserves circles and angles. Now consider a rotation composed of the reflections $\mathcal{M}_{\boldsymbol{a}}$ and $\mathcal{M}_{\boldsymbol{b}}$. The vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are poles of great circles, say C_a and C_b. $\mathrm{St}(C_a)$ and $\mathrm{St}(C_b)$ are circles in the complex plane and they intersect at points n_+ and n_- which are the stereographic images of the axis of rotation and are the fixed points of the induced transformation. Thus the induced transformation is the composition of two inversions and is clearly of the form

$$z' = \frac{az + b}{cz + d}$$

Functions of this form with $ad - bc \neq 0$ are called Möbius transformations. For each pair of complex triplets q, r, s and q', r', s' there is a unique Möbius transformation taking one to the other

$$\frac{z'-q'}{z'-r'}\frac{s'-r'}{s'-q'} = \frac{z-q}{z-r}\frac{s-r}{s-q}$$

If we choose $q = n_+$, $r = n_-$, and $s = \infty$

$$\frac{z'-n_+}{z'-n_-} = \frac{z'_\infty - n_+}{z'_\infty - n_-}\frac{z-n_+}{z-n_-}$$

By (1.21)

$$n_+ = \frac{n_1 + \imath n_2}{1-n_3} \qquad n_- = -\frac{n_1+\imath n_2}{1+n_3}$$

The image of the point at infinity in the complex plane, z'_∞, is the stereographic projection of the image of $\boldsymbol{k}$ under the rotation which is, in the standard basis,

$$\begin{aligned} R_n(\chi)(k) &= \left[(1-\cos\chi)nn^t + \cos\chi I + \sin\chi\widehat{n}\right]k \\ &= \begin{pmatrix} 2(n_1n_3\sin^2\frac{1}{2}\chi + n_2\sin\frac{1}{2}\chi\cos\frac{1}{2}\chi) \\ 2(n_2n_3\sin^2\frac{1}{2}\chi - n_1\sin\frac{1}{2}\chi\cos\frac{1}{2}\chi) \\ 2n_3^2\sin^2\frac{1}{2}\chi - \sin^2\frac{1}{2}\chi + \cos^2\frac{1}{2}\chi \end{pmatrix} \\ &= \begin{pmatrix} 2(q_1q_3+q_0q_2) \\ 2(q_2q_3-q_0q_1) \\ 2(q_0^2+q_3^2)-1 \end{pmatrix} \end{aligned}$$

Stereographic projection yields

$$z'_\infty = \imath\frac{(n_1+\imath n_2)(\cos\frac{1}{2}\chi + \imath\sin\frac{1}{2}\chi)}{(1-n_3^2)\sin^2\frac{1}{2}\chi}$$

from which it follows that

$$C = \frac{z'_\infty - n_+}{z'_\infty - n_-} = e^{-\imath\chi}$$

Now it is just a matter of algebra to show that

$$z' = f(z) = \frac{\alpha z + \beta}{-\beta^* z + \alpha^*}$$

where

$$\alpha = q_0 + \imath q_3 \qquad \beta = -q_2 + \imath q_1 \tag{1.23}$$

are the Cayley–Klein parameters. The form of this function is not unique because the numerator and denominator can be multiplied by a common factor without affecting its value. In other words, there is an equivalence class of functions which correspond to each rotation of the sphere. We have chosen, as is always possible if $\alpha\delta - \beta\gamma \neq 0$, the function satisfying

$$z' = f(z) = \frac{\alpha z + \beta}{-\beta^* z + \alpha^*} \qquad \alpha\alpha^* + \beta\beta^* = 1$$

These equivalence classes of functions are called normalized *Möbius transformations*, $Mo(\mathbb{C})$, and there is a one–one correspondence between them and matrices in $SU(2)$

$$\begin{pmatrix} u & v \\ -v^* & u^* \end{pmatrix} \leftrightarrow \frac{uz+v}{-v^*z+u^*}$$

The correspondence yields an isomorphism. If $A, B \in SU(2)$ and $f, g \in Mo(\mathbb{C})$ with $A \leftrightarrow f$, $B \leftrightarrow g$, then $AB \leftrightarrow f \circ g$ and $A^{-1} \leftrightarrow f^{-1}$.

1.4 Summary

Table 1.2 Summary of rotation parameterizations.

Object	Typical element	Group operation	Representation of rotation
$\mathbb{H}_1$	$(q_0, \boldsymbol{q})$	Quaternion multiply	$\mathsf{q}\mathsf{r}\bar{\mathsf{q}}$
$SU(2)$	$\begin{pmatrix} q_0 + \imath q_3 & -q_2 + \imath q_1 \\ q_2 + \imath q_1 & q_0 - \imath q_3 \end{pmatrix}$	Matrix multiply	$AXA^\dagger$
$Mo(\mathbb{C})$	$\frac{(q_0 + \imath q_3)z - q_2 + \imath q_1}{(q_2 + \imath q_1)z + q_0 - \imath q_3}$	Composition	$f(z)$
$SO(3)$	$2qq^t + (q_0^2 - q \cdot q)I + 2q_0\widehat{q}$	Matrix multiply	$R\,v$

The aspects of the theory of rotations in $\mathbb{R}^3$ covered in this chapter are summarized in Table 1.2 . The first three objects are isomorphic but they each bear a 2–1 relationship with $SO(3)$. This state of affairs is sometimes expressed as "$SU(2)$ is a double cover for $SO(3)$." This has profound implications in particle physics where the duplicity represents spin.

We close this chapter with comment on the topology of the systems used to represent rotations. $\mathbb{H}_1$, and hence $SU(2)$, is homeomorphic to the three-dimensional sphere

$$S^3 = [x \in \mathbb{R}^4 | x \cdot x = 1]$$

On the other hand, $SO(3)$ is homeomorphic to S^3 with antipodal points identified. This is the projective space $\mathbb{R}P^3$ – the space of all lines through the origin in $\mathbb{R}^4$. The former is simply connected (any loop can be contracted to a point) but the latter is not. In $\mathbb{H}_1$ or $SU(2)$ it is possible, to "unwind" the loops which project to noncontractible loops in $SO(3)$. For example, an arc from a point on S^3 to its antipode projects to a closed loop in $SO(3)$. This can actually be visualized with strings attached to a rotated body and is very nicely illustrated in [10]. It is also related to the "waiter's trick" described in [15]. If one holds a tray on one's outstretched, upturned hand one can rotate the tray by 360° using only motion of the wrist and the arm. The "trick" is that in doing so the elbow winds up pointed skyward. A further rotation by 360° using only motions of the wrist and the arm restores the elbow to its original position.