Chapter 2
Mathematical Background

2.1 Preamble

First and foremost, the study of motions undergone by robotic mechanical systems or, for that matter, by mechanical systems at large, requires a suitable motion representation. Now, the motion of mechanical systems involves the motion of the particular links comprising those systems, which in this book are supposed to be rigid. The assumption of rigidity, although limited in scope, still covers a wide spectrum of applications, while providing insight into the motion of more complicated systems, such as those involving deformable bodies.

The most general kind of rigid-body motion consists of both translation and rotation. While the study of the former is covered in elementary mechanics courses and is reduced to the mechanics of particles, the latter is more challenging. Indeed, point translation can be studied simply with the aid of three-dimensional vector calculus, while rigid-body rotations require the introduction of *tensors*, i.e., entities mapping vector spaces into vector spaces.

Emphasis is placed on *invariant* concepts, i.e., items that do not change upon a change of coordinate frame. Examples of invariant concepts are geometric quantities such as distances and angles between lines. Although we may resort to a coordinate frame and vector algebra to compute distances and angles, and will represent vectors in that frame, the final result will be independent of how we choose that frame. The same applies to quantities whose evaluation calls for the introduction of tensors. Here, we must distinguish between the physical quantity represented by a vector or a tensor and the representation of that quantity in a coordinate frame using a one-dimensional array of components in the case of vectors, or a two-dimensional array in the case of tensors. It is unfortunate that the same word is used in English to denote a vector and its array representation in a given coordinate frame. Regarding tensors, the associated arrays are called *matrices*. By abuse of terminology, we will refer to both tensors and their arrays as matrices, although keeping in mind the essential conceptual differences involved.

J. Angeles, *Fundamentals of Robotic Mechanical Systems: Theory, Methods, and Algorithms*, Mechanical Engineering Series 124, DOI 10.1007/978-3-319-01851-5_2,

2.2 Linear Transformations

The physical three-dimensional space is a particular case of a *vector space*. A vector space is a set of objects, called *vectors*, that follow certain algebraic rules. Throughout the book, vectors will be denoted by boldface lowercase characters, whereas tensors and their matrix representations will be denoted by boldface uppercase characters. Let $\mathbf{v}$, $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{w}$ be elements of a given vector space $\mathcal{V}$, which is *defined over the real field*, and let α and β be two elements of this field, i.e., α and β are two real numbers. Below we summarize the rules mentioned above:

(i) The sum of $\mathbf{v}_1$ and $\mathbf{v}_2$, denoted by $\mathbf{v}_1 + \mathbf{v}_2$, is itself an element of $\mathcal{V}$ and is *commutative*, i.e., $\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_2 + \mathbf{v}_1$;
(ii) $\mathcal{V}$ contains an element $\mathbf{0}$, called the *zero* vector of $\mathcal{V}$, which, when added to any other element $\mathbf{v}$ of $\mathcal{V}$, leaves it unchanged, i.e., $\mathbf{v} + \mathbf{0} = \mathbf{v}$;
(iii) The sum defined in (i) is *associative*, i.e., $\mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3) = (\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3$;
(iv) For every element $\mathbf{v}$ of $\mathcal{V}$, there exists a corresponding element, $\mathbf{w}$, also of $\mathcal{V}$, which, when added to $\mathbf{v}$, produces the zero vector, i.e., $\mathbf{v} + \mathbf{w} = \mathbf{0}$. Moreover, $\mathbf{w}$ is represented as $-\mathbf{v}$;
(v) The product $\alpha\mathbf{v}$, or $\mathbf{v}\alpha$, is also an element of $\mathcal{V}$, for every $\mathbf{v}$ of $\mathcal{V}$ and every real α. This product is associative, i.e., $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$;
(vi) If α is the real unity, then $\alpha\mathbf{v}$ is identically $\mathbf{v}$;
(vii) The product defined in (v) is *distributive* in the sense that (a) $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$ and (b) $\alpha(\mathbf{v}_1 + \mathbf{v}_2) = \alpha\mathbf{v}_1 + \alpha\mathbf{v}_2$.

Although vector spaces can be defined over other fields, we will deal with vector spaces over the real field, unless explicit reference to another field is made. Moreover, vector spaces can be either finite- or infinite-dimensional, but we will not need the latter. In geometry and elementary mechanics, the dimension of the vector spaces needed is usually three, but when studying multibody systems, an arbitrary finite dimension will be required. The concept of *dimension* of a vector space is discussed in more detail later.

A *linear transformation*, represented as an *operator* $\mathbf{L}$, of a vector space $\mathcal{U}$ into a vector space $\mathcal{V}$, is a rule that assigns to every vector $\mathbf{u}$ of $\mathcal{U}$ at least one vector $\mathbf{v}$ of $\mathcal{V}$, represented as $\mathbf{v} = \mathbf{L}\mathbf{u}$, with $\mathbf{L}$ endowed with two properties:

(i) *homogeneity*: $\mathbf{L}(\alpha\mathbf{u}) = \alpha\mathbf{v}$; and
(ii) *additivity*: $\mathbf{L}(\mathbf{u}_1 + \mathbf{u}_2) = \mathbf{v}_1 + \mathbf{v}_2$.

Note that, in the foregoing definitions, no mention has been made of components, and hence, vectors and their transformations should not be confused with their *array representations*.

Particular types of linear transformations of the three-dimensional Euclidean space that will be encountered frequently in this context are *projections*, *reflections*, and *rotations*. One further type of transformation, which is not linear, but nevertheless appears frequently in kinematics, is the one known as *affine transformation*. The foregoing transformations are defined below. It is necessary, however, to introduce additional concepts pertaining to general linear transformations before expanding into these definitions.

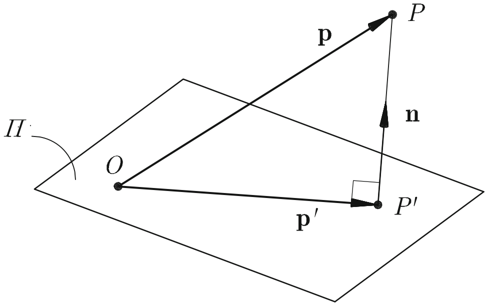

Fig. 2.1 A projection onto a plane Π of unit normal $\mathbf{n}$

The *range* of a linear transformation $\mathbf{L}$ of $\mathcal{U}$ into $\mathcal{V}$ is the set of vectors $\mathbf{v}$ of $\mathcal{V}$ into which some vector $\mathbf{u}$ of $\mathcal{U}$ is mapped, i.e., the range of $\mathbf{L}$ is defined as the set of $\mathbf{v} = \mathbf{Lu}$, for every vector $\mathbf{u}$ of $\mathcal{U}$. The *kernel* of $\mathbf{L}$ is the set of vectors $\mathbf{u}_N$ of $\mathcal{U}$ that are mapped by $\mathbf{L}$ into the zero vector $\mathbf{0} \in \mathcal{V}$. It can be readily proven (see Exercises 2.1–2.3) that the kernel and the range of a linear transformation are both vector subspaces of $\mathcal{U}$ and $\mathcal{V}$, respectively, i.e., they are themselves vector spaces, but of a dimension smaller than or equal to that of their associated vector spaces. Moreover, the kernel of a linear transformation is often called the *null space* of the said transformation.

Henceforth, the three-dimensional Euclidean space is denoted by $\mathcal{E}^3$. Having chosen an origin O for this space, its geometry can be studied in the context of general vector spaces. Hence, points of $\mathcal{E}^3$ will be identified with vectors of the associated three-dimensional vector space. Moreover, lines and planes passing through the origin are subspaces of dimensions 1 and 2, respectively, of $\mathcal{E}^3$. Clearly, lines and planes not passing through the origin of $\mathcal{E}^3$ are not subspaces but can be handled with the algebra of vector spaces, as will be shown here.

An *orthogonal projection* $\mathbf{P}$ of $\mathcal{E}^3$ onto itself is a linear transformation of the said space onto a plane Π passing through the origin and having a unit normal $\mathbf{n}$, with the properties:

$$\mathbf{P}^2 = \mathbf{P}, \quad \mathbf{Pn} = \mathbf{0} \tag{2.1a}$$

Any matrix with the first property above is termed *idempotent*. For $n \times n$ matrices, it is sometimes necessary to indicate the lowest integer l for which an analogous relation follows, i.e., for which $\mathbf{P}^l = \mathbf{P}$. In this case, the matrix is said to be idempotent of degree l.

Clearly, the projection of a position vector $\mathbf{p}$, denoted by $\mathbf{p}'$, onto a plane Π of unit normal $\mathbf{n}$, is $\mathbf{p}$ itself minus the component of $\mathbf{p}$ along $\mathbf{n}$ as shown in Fig. 2.1, i.e.,

$$\mathbf{p}' = \mathbf{p} - \mathbf{n}(\mathbf{n}^T\mathbf{p}) \tag{2.1b}$$

where the superscript T denotes either vector or matrix transposition and $\mathbf{n}^T\mathbf{p}$ is equivalent to the usual *dot product* $\mathbf{n} \cdot \mathbf{p}$.

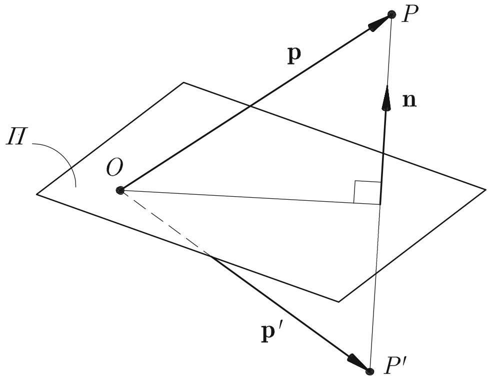

Fig. 2.2 A reflection onto a plane Π of unit normal $\mathbf{n}$

Now, the *identity* matrix $\mathbf{1}$ is defined as the mapping of a vector space $\mathcal{V}$ into itself leaving every vector $\mathbf{v}$ of $\mathcal{V}$ unchanged, i.e.,

$$\mathbf{1v} = \mathbf{v} \tag{2.2}$$

Thus, $\mathbf{p}'$, as given by Eq. (2.1b), can be rewritten as

$$\mathbf{p}' = \mathbf{1p} - \mathbf{nn}^T\mathbf{p} \equiv (\mathbf{1} - \mathbf{nn}^T)\mathbf{p} \tag{2.3}$$

and hence, the *orthogonal projection* $\mathbf{P}$ onto Π can be represented as

$$\mathbf{P} = \mathbf{1} - \mathbf{nn}^T \tag{2.4}$$

where the product $\mathbf{nn}^T$ amounts to a 3×3 matrix.

Now we turn to reflections. Here we have to take into account that reflections occur frequently accompanied by rotations, as yet to be studied. Since reflections are simpler to represent, we first discuss these, rotations being discussed in full detail in Sect. 2.3. What we shall discuss in this section is *pure reflections*, i.e., those occurring without any concomitant rotation. Thus, all reflections studied in this section are pure reflections, but for the sake of brevity, they will be referred to simply as *reflections*.

A *reflection* $\mathbf{R}$ of $\mathcal{E}^3$ onto a plane Π passing through the origin and having a unit normal $\mathbf{n}$ is a linear transformation of the said space into itself, as depicted in Fig. 2.2, such that a vector $\mathbf{p}$ is mapped by $\mathbf{R}$ into a vector $\mathbf{p}'$ given by

$$\mathbf{p}' = \mathbf{p} - 2\mathbf{nn}^T\mathbf{p} \equiv (\mathbf{1} - 2\mathbf{nn}^T)\mathbf{p}$$

Thus, the reflection $\mathbf{R}$ can be expressed as

$$\mathbf{R} = \mathbf{1} - 2\mathbf{nn}^T \tag{2.5}$$

From Eq. (2.5) it is then apparent that a pure reflection is represented by a linear transformation that is symmetric and whose square equals the identity matrix, i.e., $\mathbf{R}^2 = \mathbf{1}$. Indeed, symmetry is apparent from the equation above; the second property is readily proven below:

$$\begin{aligned}\mathbf{R}^2 &= (\mathbf{1} - 2\mathbf{n}\mathbf{n}^T)(\mathbf{1} - 2\mathbf{n}\mathbf{n}^T) \\ &= \mathbf{1} - 2\mathbf{n}\mathbf{n}^T - 2\mathbf{n}\mathbf{n}^T + 4(\mathbf{n}\mathbf{n}^T)(\mathbf{n}\mathbf{n}^T) = \mathbf{1} - 4\mathbf{n}\mathbf{n}^T + 4\mathbf{n}(\mathbf{n}^T\mathbf{n})\mathbf{n}^T\end{aligned}$$

which apparently reduces to $\mathbf{1}$ because $\mathbf{n}$ is a unit vector. Note that from the second property above, we find that pure reflections observe a further interesting property, namely,

$$\mathbf{R}^{-1} = \mathbf{R}$$

i.e., every pure reflection equals its inverse. This result can be understood intuitively by noticing that, upon doubly reflecting an image using two mirrors, the original image is recovered. Any square matrix which equals its inverse will be termed *self-inverse* henceforth.

Further, we take to deriving the *orthogonal decomposition* of a given vector $\mathbf{v}$ into two components, one along and one normal to a unit vector $\mathbf{e}$. The component of $\mathbf{v}$ along $\mathbf{e}$, termed here the *axial component*, $\mathbf{v}_{\|}$—read *v-par*—is simply given as

$$\mathbf{v}_{\|} \equiv \mathbf{e}\mathbf{e}^T\mathbf{v} \tag{2.6a}$$

while the corresponding *normal component*, $\mathbf{v}_{\perp}$—read *v-perp*—is simply the difference $\mathbf{v} - \mathbf{v}_{\|}$, i.e.,

$$\mathbf{v}_{\perp} \equiv \mathbf{v} - \mathbf{v}_{\|} \equiv (\mathbf{1} - \mathbf{e}\mathbf{e}^T)\mathbf{v} \tag{2.6b}$$

the matrix in parentheses in the foregoing equation being rather frequent in kinematics. This matrix will appear when studying rotations.

Further concepts are now recalled: The *basis* of a vector space $\mathcal{V}$ is a set of *linearly independent* vectors of $\mathcal{V}$, $\{\mathbf{v}_i\}_1^n$, in terms of which *any* vector $\mathbf{v}$ of $\mathcal{V}$ can be expressed as

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n \tag{2.7}$$

where the elements of the set $\{\alpha_i\}_1^n$ are all elements of the field over which $\mathcal{V}$ is defined, i.e., they are real numbers in the case at hand. The number n of elements in the set $\mathcal{B} = \{\mathbf{v}_i\}_1^n$ is called *the dimension* of $\mathcal{V}$. Note that *any* set of n linearly independent vectors of $\mathcal{V}$ can play the role of a basis of this space, but once this basis is defined, the set of real coefficients $\{\alpha_i\}_1^n$ representing a given vector $\mathbf{v}$ is *unique*.

Let $\mathcal{U}$ and $\mathcal{V}$ be two vector spaces of dimensions m and n, respectively, and $\mathbf{L}$ a linear transformation of $\mathcal{U}$ into $\mathcal{V}$, and define bases $\mathcal{B}_U$ and $\mathcal{B}_V$ for $\mathcal{U}$ and $\mathcal{V}$ as

$$\mathcal{B}_U = \{\mathbf{u}_j\}_1^m, \quad \mathcal{B}_V = \{\mathbf{v}_i\}_1^n \tag{2.8}$$

Since each $\mathbf{Lu}_j$ is an element of $\mathcal{V}$, it can be represented uniquely in terms of the vectors of $\mathcal{B}_V$, namely, as

$$\mathbf{Lu}_j = l_{1j}\mathbf{v}_1 + l_{2j}\mathbf{v}_2 + \cdots + l_{nj}\mathbf{v}_n, \quad j = 1, \ldots, m \tag{2.9}$$

Consequently, in order to represent the *images* of the m vectors of $\mathcal{B}_U$, namely, the set $\{\mathbf{Lu}_j\}_1^m$, $n \times m$ real numbers l_{ij}, for $i = 1, \ldots, n$ and $j = 1, \ldots, m$, are necessary. These real numbers are now arranged in the $n \times m$ array $[\mathbf{L}]_{\mathcal{B}_U}^{\mathcal{B}_V}$ defined below:

$$[\mathbf{L}]_{\mathcal{B}_U}^{\mathcal{B}_V} \equiv \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1m} \\ l_{21} & l_{22} & \cdots & l_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nm} \end{bmatrix} \tag{2.10}$$

The foregoing array is thus called the *matrix representation of* $\mathbf{L}$ *with respect to* $\mathcal{B}_U$ *and* $\mathcal{B}_V$. We thus have an important definition:

Definition 2.2.1. The jth column of the matrix representation of $\mathbf{L}$ with respect to the bases $\mathcal{B}_U$ and $\mathcal{B}_V$ is composed of the n real coefficients l_{ij} of the representation of the image of the jth vector of $\mathcal{B}_U$ in terms of $\mathcal{B}_V$.

The notation introduced in Eq. (2.10) is rather cumbersome, for it involves one subscript and one superscript. Moreover, each of these is subscripted. In practice, the bases involved are self-evident, which makes an explicit mention of these unnecessary. In particular, when $\mathbf{L}$ is a mapping of $\mathcal{U}$ onto itself, a single basis suffices to represent $\mathbf{L}$ in matrix form. In this case, its bracket will bear only a subscript, and no superscript, namely, $[\mathbf{L}]_{\mathcal{B}}$. Moreover, we will use, henceforth, the concept of basis and coordinate frame interchangeably, since one implies the other.

Two different bases are unavoidable when the two spaces under study are physically distinct, which is the case in velocity analysis of manipulators. As we will see in Chap. 5, in these analyses we distinguish between the velocity of the manipulator in Cartesian space and that in the joint-rate space. While the Cartesian-space velocity—or Cartesian velocity, for brevity—consists, in general, of a six-dimensional vector containing the three-dimensional angular velocity of the end-effector and the translational velocity of one of its points, the latter is an n-dimensional vector. Moreover, if the manipulator is coupled by revolute joints only, the units of the joint-rate vector are all s^{-1}, whereas the Cartesian velocity contains some components with units of s^{-1} and others with units of ms^{-1}.

Further definitions are now recalled. Given a mapping $\mathbf{L}$ of an n-dimensional vector space $\mathcal{U}$ into the n-dimensional vector space $\mathcal{V}$, a nonzero vector $\mathbf{e}$ that is mapped by $\mathbf{L}$ into a multiple of itself, $\lambda\mathbf{e}$, is called an *eigenvector* of $\mathbf{L}$, the scalar λ being called an *eigenvalue* of $\mathbf{L}$. The eigenvalues of $\mathbf{L}$ are determined by the equation

$$\det(\lambda\mathbf{1} - \mathbf{L}) = 0 \tag{2.11}$$

Note that the matrix $\lambda\mathbf{1}-\mathbf{L}$ is *linear* in λ, and since the determinant of a $n\times n$ matrix is a homogeneous nth-order function of its entries, the left-hand side of Eq. (2.11) is a nth-degree polynomial in λ. The foregoing polynomial is termed *the characteristic polynomial of* $\mathbf{L}$. Hence, every $n\times n$ matrix $\mathbf{L}$ has n complex eigenvalues, even if $\mathbf{L}$ is defined over the real field. If it is, then its complex eigenvalues appear in conjugate pairs. Clearly, the eigenvalues of $\mathbf{L}$ are the roots of its characteristic polynomial, while Eq. (2.11) is called the *characteristic equation* of $\mathbf{L}$.

Example 2.2.1. What is the representation of the reflection $\mathbf{R}$ of $\mathcal{E}^3$ into itself, with respect to the x–y plane, in terms of unit vectors parallel to the X, Y, Z axes that form a coordinate frame $\mathcal{F}$?

Solution: Note that in this case, $\mathcal{U} = \mathcal{V} = \mathcal{E}^3$ and, hence, it is not necessary to use two different bases for $\mathcal{U}$ and $\mathcal{V}$. Now, let $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$, be unit vectors parallel to the *X*, *Y*, and *Z* axes. Clearly,

$$\mathbf{Ri} = \mathbf{i}$$

$$\mathbf{Rj} = \mathbf{j}$$

$$\mathbf{Rk} = -\mathbf{k}$$

Thus, the representations of the images of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ under $\mathbf{R}$, in $\mathcal{F}$, are

$$[\,\mathbf{Ri}\,]_{\mathcal{F}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [\,\mathbf{Rj}\,]_{\mathcal{F}} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad [\,\mathbf{Rk}\,]_{\mathcal{F}} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

where subscripted brackets are used to indicate the representation frame. Hence, the matrix representation of $\mathbf{R}$ in $\mathcal{F}$, denoted by $[\,\mathbf{R}\,]_{\mathcal{F}}$, is

$$[\,\mathbf{R}\,]_{\mathcal{F}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

2.3 Rigid-Body Rotations

A *linear isomorphism*, i.e., a one-to-one linear transformation mapping a space $\mathcal{V}$ onto itself, is called an *isometry* if it preserves distances between any two points of $\mathcal{V}$. If $\mathbf{u}$ and $\mathbf{v}$ are regarded as the position vectors of two such points, then the distance d between these two points is defined as

$$d \equiv \sqrt{(\mathbf{u}-\mathbf{v})^T(\mathbf{u}-\mathbf{v})} \tag{2.12}$$

The volume V of the tetrahedron defined by the origin and three points of the three-dimensional Euclidean space of position vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ is obtained as one-sixth of the absolute value of the *double mixed product* of these three vectors,

$$V \equiv \frac{1}{6}\,|\mathbf{u} \times \mathbf{v} \cdot \mathbf{w}| = \frac{1}{6}\left|\det\left[\mathbf{u}\ \mathbf{v}\ \mathbf{w}\right]\right| \tag{2.13}$$

i.e., if a 3×3 array [**A**] is defined in terms of the components of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, in a given basis, then the first column of [**A**] is given by the three components of $\mathbf{u}$, the second and third columns being defined likewise.

Now, let $\mathbf{Q}$ be an isometry mapping the triad $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ into $\{\mathbf{u}', \mathbf{v}', \mathbf{w}'\}$. Moreover, the distance from the origin to the points of position vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ is given simply as $\|\mathbf{u}\|$, $\|\mathbf{v}\|$, and $\|\mathbf{w}\|$, which are defined as

$$\|\mathbf{u}\| \equiv \sqrt{\mathbf{u}^T\mathbf{u}}, \quad \|\mathbf{v}\| \equiv \sqrt{\mathbf{v}^T\mathbf{v}}, \quad \|\mathbf{w}\| \equiv \sqrt{\mathbf{w}^T\mathbf{w}} \tag{2.14}$$

Clearly,

$$\|\mathbf{u}'\| = \|\mathbf{u}\|, \quad \|\mathbf{v}'\| = \|\mathbf{v}\|, \quad \|\mathbf{w}'\| = \|\mathbf{w}\| \tag{2.15a}$$

and

$$\det\left[\mathbf{u}'\ \mathbf{v}'\ \mathbf{w}'\right] = \pm\det\left[\mathbf{u}\ \mathbf{v}\ \mathbf{w}\right] \tag{2.15b}$$

If, in the foregoing relations, the sign of the determinant is preserved, the isometry represents a *rotation*; otherwise, it represents a reflection. Now, let $\mathbf{p}$ be the position vector of any point of $\mathcal{E}^3$, its image under a rotation $\mathbf{Q}$ being $\mathbf{p}'$. Hence, distance preservation requires that

$$\mathbf{p}^T\mathbf{p} = \mathbf{p}'^T\mathbf{p}' \tag{2.16}$$

where

$$\mathbf{p}' = \mathbf{Q}\mathbf{p} \tag{2.17}$$

condition (2.16) thus leading to

$$\mathbf{Q}^T\mathbf{Q} = \mathbf{1} \tag{2.18}$$

where $\mathbf{1}$ was defined in Sect. 2.2 as the 3×3 *identity matrix*, and hence, Eq. (2.18) states that $\mathbf{Q}$ is an *orthogonal matrix*. Moreover, let $\mathbf{T}$ and $\mathbf{T}'$ denote the two matrices defined below:

$$\mathbf{T} = \left[\mathbf{u}\ \mathbf{v}\ \mathbf{w}\right], \quad \mathbf{T}' = \left[\mathbf{u}'\ \mathbf{v}'\ \mathbf{w}'\right] \tag{2.19}$$

from which it is clear that

$$\mathbf{T}' = \mathbf{Q}\mathbf{T} \tag{2.20}$$

Now, for a rigid-body rotation, Eq. (2.15b) should hold with the positive sign, and hence,

$$\det(\mathbf{T}) = \det(\mathbf{T}') \tag{2.21a}$$

and, by virtue of Eq. (2.20), we conclude that

$$\det(\mathbf{Q}) = +1 \tag{2.21b}$$

Therefore, $\mathbf{Q}$ is a *proper orthogonal matrix*, i.e., it is a proper isometry. Now we have

Theorem 2.3.1. *The eigenvalues of a proper orthogonal matrix* $\mathbf{Q}$ *lie on the unit circle centered at the origin of the complex plane.*

Proof. Let λ be one of the eigenvalues of $\mathbf{Q}$ and $\mathbf{e}$ the corresponding eigenvector, so that

$$\mathbf{Q}\mathbf{e} = \lambda\mathbf{e} \tag{2.22}$$

In general, $\mathbf{Q}$ is not expected to be symmetric, and hence, λ is not necessarily real. Thus, λ is considered complex, in general. In this light, when transposing both sides of the foregoing equation, we will need to take the complex conjugates as well. Henceforth, the complex conjugate of a vector or a matrix will be indicated with an asterisk as a superscript. As well, the conjugate of a complex variable will be indicated with a bar over the said variable. Thus, the transpose conjugate of the above equation takes on the form

$$\mathbf{e}^*\mathbf{Q}^* = \bar{\lambda}\mathbf{e}^* \tag{2.23}$$

Multiplying the corresponding sides of the two previous equations yields

$$\mathbf{e}^*\mathbf{Q}^*\mathbf{Q}\mathbf{e} = \bar{\lambda}\lambda\mathbf{e}^*\mathbf{e} \tag{2.24}$$

However, $\mathbf{Q}$ has been assumed real, and hence, $\mathbf{Q}^*$ reduces to $\mathbf{Q}^T$, the foregoing equation thus reducing to

$$\mathbf{e}^*\mathbf{Q}^T\mathbf{Q}\mathbf{e} = \bar{\lambda}\lambda\mathbf{e}^*\mathbf{e} \tag{2.25}$$

But $\mathbf{Q}$ is orthogonal by assumption, and hence, it obeys Eq. (2.18), which means that Eq. (2.25) reduces to

$$\mathbf{e}^*\mathbf{e} = |\lambda|^2\mathbf{e}^*\mathbf{e} \tag{2.26}$$

where $|\cdot|$ denotes the *module* of the complex variable within it. Thus, the foregoing equation leads to

$$|\lambda|^2 = 1 \tag{2.27}$$

thereby completing the intended proof. As a direct consequence of Theorem 2.3.1, we have

Corollary 2.3.1. *A proper orthogonal* 3×3 *matrix has at least one eigenvalue that is* $+1$.

Now, let $\mathbf{e}$ be the eigenvector of $\mathbf{Q}$ associated with the eigenvalue $+1$. Thus,

$$\mathbf{Qe} = \mathbf{e} \tag{2.28}$$

What Eq. (2.28) states is summarized as a theorem below:

Theorem 2.3.2 (Euler 1776). *A rigid-body motion about a point* O *leaves fixed a set of points lying on a line* $\mathcal{L}$ *that passes through* O *and is parallel to the eigenvector* $\mathbf{e}$ *of* $\mathbf{Q}$ *associated with the eigenvalue* $+1$.

A further result, that finds many applications in robotics and, in general, in system theory, is given below:

Theorem 2.3.3 (Cayley–Hamilton). *Let* $P(\lambda)$ *be the characteristic polynomial of a* $n \times n$ *matrix* $\mathbf{A}$*, i.e.,*

$$P(\lambda) = \det(\lambda\mathbf{1} - \mathbf{A}) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 \tag{2.29}$$

Then $\mathbf{A}$ *satisfies its characteristic equation, i.e.,*

$$\mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \cdots + a_1\mathbf{A} + a_0\mathbf{1} = \mathbf{O} \tag{2.30}$$

where $\mathbf{O}$ *is the* $n \times n$ *zero matrix.*

Proof. See Kaye and Wilson (1998).

What the Cayley–Hamilton Theorem states is that any power $p \geq n$ of the $n \times n$ matrix $\mathbf{A}$ can be expressed as a linear combination of the first n powers of $\mathbf{A}$—the 0th power of $\mathbf{A}$ is, of course, the $n \times n$ identity matrix $\mathbf{1}$. An important consequence of this result is that any *analytic* matrix function of $\mathbf{A}$ can be expressed not as an infinite series, but as a sum, namely, a linear combination of the first n powers of $\mathbf{A}$: $\mathbf{1}$, $\mathbf{A}$, $\ldots$, $\mathbf{A}^{n-1}$. An *analytic* function $f(x)$ of a real variable x is, in turn, a function with a series expansion. Moreover, an analytic matrix function of a matrix argument $\mathbf{A}$ is defined likewise, an example of which is the exponential function. From the previous discussion, then, the exponential of $\mathbf{A}$ can be written as a linear combination of the first n powers of $\mathbf{A}$. It will be shown later that any proper orthogonal matrix $\mathbf{Q}$ can be represented as the exponential of a skew-symmetric matrix derived from the unit vector $\mathbf{e}$ of $\mathbf{Q}$, of eigenvalue $+1$, and the associated angle of rotation, as yet to be defined.

2.3.1 The Cross-Product Matrix

Prior to introducing the matrix representation of a rotation, we will need a few definitions. We will start by defining the partial derivative of a vector with respect to another vector. This is a matrix, as described below: In general, let $\mathbf{u}$ and $\mathbf{v}$ be vectors of spaces $\mathcal{U}$ and $\mathcal{V}$, of dimensions m and n, respectively. Furthermore, let t be a real variable and f be real-valued function of t, $\mathbf{u} = \mathbf{u}(t)$ and $\mathbf{v} = \mathbf{v}(\mathbf{u}(t))$ being m- and n-dimensional vector functions of t as well, with $f = f(\mathbf{u}, \mathbf{v})$. The derivative of $\mathbf{u}$ with respect to t, denoted by $\dot{\mathbf{u}}(t)$, is a m-dimensional vector whose ith component is the derivative of the ith component u_i of $\mathbf{u}$, in a given basis, with respect to t. A similar definition follows for $\dot{\mathbf{v}}(t)$. The partial derivative of f with respect to $\mathbf{u}$ is a m-dimensional vector whose ith component is the partial derivative of f with respect to u_i, with a corresponding definition for the partial derivative of f with respect to $\mathbf{v}$. The foregoing derivatives, as all other vectors, will be assumed, henceforth, to be *column* arrays. Thus,

$$\frac{\partial f}{\partial \mathbf{u}} \equiv \begin{bmatrix} \partial f/\partial u_1 \\ \partial f/\partial u_2 \\ \vdots \\ \partial f/\partial u_m \end{bmatrix}, \quad \frac{\partial f}{\partial \mathbf{v}} \equiv \begin{bmatrix} \partial f/\partial v_1 \\ \partial f/\partial v_2 \\ \vdots \\ \partial f/\partial v_n \end{bmatrix} \tag{2.31}$$

Furthermore, let $\mathbf{v} = \mathbf{v}(\mathbf{u})$. In order to derive $\partial \mathbf{u}/\partial \mathbf{v}$, first the differential $\mathrm{d}\mathbf{v}$ upon a differential $\mathrm{d}\mathbf{u}$ is computed:

$$\mathrm{d}\mathbf{v} = \frac{\partial \mathbf{v}}{\partial u_1}\mathrm{d}u_1 + \frac{\partial \mathbf{v}}{\partial u_2}\mathrm{d}u_2 + \ldots + \frac{\partial \mathbf{v}}{\partial u_m}\mathrm{d}u_m \tag{2.32a}$$

or, in array form,

$$\mathrm{d}\mathbf{v} = \underbrace{\begin{bmatrix} \dfrac{\partial \mathbf{v}}{\partial u_1} & \dfrac{\partial \mathbf{v}}{\partial u_2} & \cdots & \dfrac{\partial \mathbf{v}}{\partial u_m} \end{bmatrix}}_{\dfrac{\partial \mathbf{v}}{\partial \mathbf{u}}} \underbrace{\begin{bmatrix} \mathrm{d}u_1 \\ \mathrm{d}u_2 \\ \vdots \\ \mathrm{d}u_m \end{bmatrix}}_{\mathrm{d}\mathbf{u}} \tag{2.32b}$$

That is, the partial derivative of $\mathbf{v}$ with respect to $\mathbf{u}$ is a $n \times m$ array whose (i, j) entry is defined as $\partial v_i/\partial u_j$, i.e.,

$$\frac{\partial \mathbf{v}}{\partial \mathbf{u}} \equiv \begin{bmatrix} \partial v_1/\partial u_1 & \partial v_1/\partial u_2 & \cdots & \partial v_1/\partial u_m \\ \partial v_2/\partial u_1 & \partial v_2/\partial u_2 & \cdots & \partial v_2/\partial u_m \\ \vdots & \vdots & \ddots & \vdots \\ \partial v_n/\partial u_1 & \partial v_n/\partial u_2 & \cdots & \partial v_n/\partial u_m \end{bmatrix} \tag{2.33}$$

Hence, the total derivative of f with respect to $\mathbf{u}$ can be written as

$$\frac{df}{d\mathbf{u}} = \frac{\partial f}{\partial \mathbf{u}} + \left(\frac{\partial \mathbf{v}}{\partial \mathbf{u}}\right)^T \frac{\partial f}{\partial \mathbf{v}} \tag{2.34}$$

If, moreover, f is an explicit function of t, i.e., if $f = f(\mathbf{u}, \mathbf{v}, t)$ and $\mathbf{v} = \mathbf{v}(\mathbf{u}, t)$, then, one can write the total derivative of f with respect to t as

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \left(\frac{\partial f}{\partial \mathbf{u}}\right)^T \frac{d\mathbf{u}}{dt} + \left(\frac{\partial f}{\partial \mathbf{v}}\right)^T \frac{\partial \mathbf{v}}{\partial t} + \left(\frac{\partial f}{\partial \mathbf{v}}\right)^T \frac{\partial \mathbf{v}}{\partial \mathbf{u}} \frac{d\mathbf{u}}{dt} \tag{2.35}$$

The total derivative of $\mathbf{v}$ with respect to t can be written, likewise, as

$$\frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{v}}{\partial t} + \frac{\partial \mathbf{v}}{\partial \mathbf{u}} \frac{d\mathbf{u}}{dt} \tag{2.36}$$

Example 2.3.1. Let the components of $\mathbf{v}$ and $\mathbf{x}$ in a certain reference frame $\mathcal{F}$ be given as

$$[\mathbf{v}]_{\mathcal{F}} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \quad [\mathbf{x}]_{\mathcal{F}} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{2.37a}$$

Then

$$[\mathbf{v} \times \mathbf{x}]_{\mathcal{F}} = \begin{bmatrix} v_2 x_3 - v_3 x_2 \\ v_3 x_1 - v_1 x_3 \\ v_1 x_2 - v_2 x_1 \end{bmatrix} \tag{2.37b}$$

Hence,

$$\left[\frac{\partial(\mathbf{v} \times \mathbf{x})}{\partial \mathbf{x}}\right]_{\mathcal{F}} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} \tag{2.37c}$$

Henceforth, the partial derivative of $\mathbf{v} \times \mathbf{x}$ with respect to $\mathbf{x}$ will be denoted by the 3×3 matrix $\mathbf{V}$. For obvious reasons, $\mathbf{V}$ is termed the *cross-product matrix* of vector $\mathbf{v}$. Sometimes the cross-product matrix of a vector $\mathbf{v}$ is represented as $\tilde{\mathbf{v}}$, but we do not follow this notation for the sake of consistency, since we decided at the outset to represent matrices with boldface uppercases. Thus, the foregoing cross product admits the alternative representations

$$\mathbf{v} \times \mathbf{x} = \mathbf{V}\mathbf{x} \tag{2.38}$$

Now, it should be apparent that:

Theorem 2.3.4. *The cross-product matrix* $\mathbf{A}$ *of any three-dimensional vector* $\mathbf{a}$ *is skew-symmetric, i.e.,*

$$\mathbf{A}^T = -\mathbf{A}$$

and, as a consequence,

$$\mathbf{a} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{A}^2\mathbf{b} \tag{2.39}$$

where $\mathbf{A}^2$ *can be readily proven to be*

$$\mathbf{A}^2 = -\|\mathbf{a}\|^2\mathbf{1} + \mathbf{a}\mathbf{a}^T \tag{2.40}$$

with $\|\cdot\|$ *denoting the* Euclidean norm *of the vector inside it.*

Note that given *any* three-dimensional vector $\mathbf{a}$, its cross-product matrix $\mathbf{A}$ is *uniquely* defined. Moreover, this matrix is skew-symmetric. The converse also holds, i.e., given any 3×3 skew-symmetric matrix $\mathbf{A}$, its associated *vector* is uniquely defined as well. This result is made apparent from Example 2.3.1 and will be discussed further when we define the *axial vector* of an arbitrary 3×3 matrix below.

Example 2.3.2. Let $\mathbf{a}$ be an arbitrary three-dimensional vector and $\mathbf{A}$ its cross-product matrix. Further, let $\mathbf{B} \equiv \mathbf{1} + \mathbf{A}$, with $\mathbf{1}$ defined as the 3×3 identity matrix. **Without resorting to components**,

(a) prove that

$$\det(\mathbf{B}) = 1 + \|\mathbf{a}\|^2 > 1$$

and hence, $\mathbf{B}$ is nonsingular.

(b) Find $\mathbf{B}^{-1}$ in terms of $\mathbf{A}$ or, equivalently, of $\mathbf{a}$.

Solution:

(a) Let $\{\alpha_i\}_1^3$ be the set of eigenvalues and $\{\mathbf{a}_i\}_1^3$ the set of *corresponding* eigenvectors of $\mathbf{A}$. Likewise, let $\{\beta_i\}_1^3$ be the set of eigenvalues and $\{\mathbf{b}_i\}_1^3$ the set of *corresponding* eigenvectors of $\mathbf{B}$. That is,

$$\mathbf{A}\mathbf{a}_i = \alpha_i\mathbf{a}_i, \quad \mathbf{B}\mathbf{b}_i = \beta_i\mathbf{b}_i, \quad i = 1, 2, 3$$

Next, add $\mathbf{a}_i$ to both sides of the first of the above equations:

$$\mathbf{a}_i + \mathbf{A}\mathbf{a}_i = \mathbf{a}_i + \alpha_i\mathbf{a}_i \quad \Rightarrow \quad (\mathbf{1} + \mathbf{A})\mathbf{a}_i = (1 + \alpha_i)\mathbf{a}_i, \quad i = 1, 2, 3$$

But, in light of the definition of $\mathbf{B}$, the second of the above equations leads to

$$\mathbf{B}\mathbf{a}_i = (1 + \alpha_i)\mathbf{a}_i, \quad i = 1, 2, 3$$

which means that $\mathbf{b}_i = \mathbf{a}_i$ and $\beta_i = 1+\alpha_i$, $i = 1, 2, 3$. Now, $\det(\mathbf{B}) = \beta_1\beta_2\beta_3$, and hence,

$$\det(\mathbf{B}) = (1+\alpha_1)(1+\alpha_2)(1+\alpha_3)$$

Therefore, to find the eigenvalues of $\mathbf{A} \equiv \mathrm{CPM}(\mathbf{a})$, notice that $\mathbf{A}$ can be written as

$$\mathbf{A} = \mathrm{CPM}(\mathbf{a}) \equiv \mathrm{CPM}(\|\mathbf{a}\|\mathbf{e}), \quad \mathbf{e} \equiv \frac{\mathbf{a}}{\|\mathbf{a}\|}$$

i.e., $\mathbf{e}$ is a unit vector obtained upon dividing $\mathbf{a}$ by its norm. By virtue of the definition of the cross-product matrix, moreover, the factor $\|\mathbf{a}\|$ can be taken outside of the $\mathrm{CPM}(\cdot)$ operator, and hence,

$$\mathbf{A} = \|\mathbf{a}\|\mathrm{CPM}(\mathbf{e}) \equiv \|\mathbf{a}\|\mathbf{E}, \quad \mathbf{E} \equiv \mathrm{CPM}(\mathbf{e})$$

But the eigenvalues of the CPM of a unit vector are proven in Exercise 2.11 to be 0, j and $-j$, with $j \equiv \sqrt{-1}$. Therefore, $\alpha_1 = 0$, $\alpha_2 = j\|\mathbf{a}\|$, $\alpha_3 = -j\|\mathbf{a}\|$, and hence, $\beta_1 = 1$, $\beta_2 = 1 + j\|\mathbf{a}\|$, $\beta_3 = 1 - j\|\mathbf{a}\|$. Thus,

$$\det(\mathbf{B}) = 1(1 + j\|\mathbf{a}\|)(1 - j\|\mathbf{a}\|) = 1 + \|\mathbf{a}\|^2$$

thereby completing the intended proof.

(b) Now, to find $\mathbf{B}^{-1}$, the characteristic equations of $\mathbf{A}$ and $\mathbf{B}$ will be needed. These are readily derived below:

$$\mathbf{A}\colon\ (\lambda-\alpha_1)(\lambda-\alpha_2)(\lambda-\alpha_3) = 0 \quad \Rightarrow \quad \lambda^3 + \|\mathbf{a}\|^2\lambda = 0$$

and

$$\mathbf{B}\colon\ (\lambda-\beta_1)(\lambda-\beta_2)(\lambda-\beta_3) = 0 \quad \Rightarrow \quad \lambda^3 - 3\lambda^2 + (3+\|\mathbf{a}\|^2)\lambda - (1+\|\mathbf{a}\|^2) = 0$$

If now the Cayley–Hamilton Theorem is invoked, the foregoing scalar characteristic equations lead to matrix polynomials in $\mathbf{A}$ and $\mathbf{B}$, namely,

$$\mathbf{A}^3 + \|\mathbf{a}\|^2\mathbf{A} = \mathbf{O}, \quad \mathbf{B}^3 - 3\mathbf{B}^2 + (3+\|\mathbf{a}\|^2)\mathbf{B} - (1+\|\mathbf{a}\|^2)\mathbf{1} = \mathbf{O}$$

with $\mathbf{O}$ denoting the 3×3 zero matrix. Since $\mathbf{B}$ is known to be non-singular from (a), the two sides of the polynomial in $\mathbf{B}$ can be multiplied by $\mathbf{B}^{-1}$; the last term of the right-hand side of the matrix polynomial thus resulting involves $\mathbf{B}^{-1}$. Upon solving for this term,

$$\mathbf{B}^{-1} = \frac{1}{1+\|\mathbf{a}\|^2}[\mathbf{B}^2 - 3\mathbf{B} + (3+\|\mathbf{a}\|^2)\mathbf{1}]$$

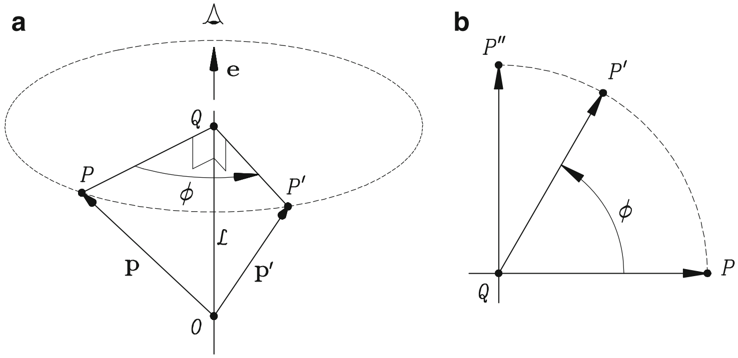

Fig. 2.3 Rotation of a rigid body about a line

which, upon expansion, yields

$$\mathbf{B}^{-1} = \frac{1}{1 + \|\mathbf{a}\|^2}[(1 + \|\mathbf{a}\|^2)\mathbf{1} - \mathbf{A} + \mathbf{A}^2]$$

The reader should be able to verify that the foregoing expression is indeed the inverse of $\mathbf{1} + \mathbf{A}$.

2.3.2 The Rotation Matrix

In deriving the matrix representation of a rotation, we should recall Theorem 2.3.2, which suggests that an explicit representation of $\mathbf{Q}$ in terms of its eigenvector $\mathbf{e}$ is possible. Moreover, this representation must contain information on the amount of the rotation under study, which is nothing but the *angle of rotation*. Furthermore, line $\mathcal{L}$, mentioned in *Euler's Theorem*, is termed the *axis of rotation* of the motion of interest. In order to derive the representation mentioned above, consider the rotation depicted in Fig. 2.3 of angle ϕ about line $\mathcal{L}$.

From Fig. 2.3a, one can apparently write

$$\mathbf{p}' = \overrightarrow{OQ} + \overrightarrow{QP'} \tag{2.41}$$

where $\overrightarrow{OQ}$ is the axial component of $\mathbf{p}$ along vector $\mathbf{e}$, which is derived as in Eq. (2.6a), namely,

$$\overrightarrow{OQ} = \mathbf{e}\mathbf{e}^T\mathbf{p} \tag{2.42}$$

Furthermore, from Fig. 2.3b,

$$\overrightarrow{QP'} = (\cos\phi)\ \overrightarrow{QP} + (\sin\phi)\ \overrightarrow{QP''} \tag{2.43}$$

with $\overrightarrow{QP}$ being nothing but the *normal component* of $\mathbf{p}$ with respect to $\mathbf{e}$, as introduced in Eq. (2.6b), i.e.,

$$\overrightarrow{QP} = (\mathbf{1} - \mathbf{e}\mathbf{e}^T)\mathbf{p} \tag{2.44}$$

and $\overrightarrow{QP''}$ given as

$$\overrightarrow{QP''} = \mathbf{e} \times \mathbf{p} \equiv \mathbf{E}\mathbf{p} \tag{2.45}$$

Substitution of Eqs. (2.44) and (2.45) into Eq. (2.43) leads to

$$\overrightarrow{QP'} = \cos\phi(\mathbf{1} - \mathbf{e}\mathbf{e}^T)\mathbf{p} + \sin\phi\mathbf{E}\mathbf{p} \tag{2.46}$$

If now Eqs. (2.42) and (2.46) are substituted into Eq. (2.41), one obtains

$$\mathbf{p}' = \mathbf{e}\mathbf{e}^T\mathbf{p} + \cos\phi(\mathbf{1} - \mathbf{e}\mathbf{e}^T)\mathbf{p} + \sin\phi\mathbf{E}\mathbf{p} \tag{2.47}$$

Thus, Eq. (2.41) reduces to

$$\mathbf{p}' = [\mathbf{e}\mathbf{e}^T + \cos\phi(\mathbf{1} - \mathbf{e}\mathbf{e}^T) + \sin\phi\mathbf{E}]\mathbf{p} \tag{2.48}$$

From Eq. (2.48) it is apparent that $\mathbf{p}'$ is a linear transformation of $\mathbf{p}$, the transformation being given by the expression inside the brackets, which is the rotation matrix $\mathbf{Q}$ sought, i.e.,

$$\mathbf{Q} = \mathbf{e}\mathbf{e}^T + \cos\phi(\mathbf{1} - \mathbf{e}\mathbf{e}^T) + \sin\phi\mathbf{E} \tag{2.49}$$

A special case arises when $\phi = \pi$,

$$\mathbf{Q} = -\mathbf{1} + 2\mathbf{e}\mathbf{e}^T, \quad \text{for } \phi = \pi \tag{2.50}$$

whence it is apparent that $\mathbf{Q}$ is symmetric if $\phi = \pi$. Of course, $\mathbf{Q}$ becomes symmetric also when $\phi = 0$, but this is a rather obvious case, leading to $\mathbf{Q} = \mathbf{1}$. Except for these two cases, the rotation matrix is not symmetric. However, under no circumstance does the rotation matrix become skew-symmetric, for a 3×3 skew-symmetric matrix is by necessity singular, which contradicts the property of proper orthogonal matrices of Eq. (2.21b).

Now one more representation of $\mathbf{Q}$ in terms of $\mathbf{e}$ and ϕ is given. For a fixed axis of rotation, i.e., for a fixed value of $\mathbf{e}$, the rotation matrix $\mathbf{Q}$ is a function of the angle of rotation ϕ, only. Thus, the series expansion of $\mathbf{Q}$ in terms of ϕ is

$$\mathbf{Q}(\phi) = \mathbf{Q}(0) + \mathbf{Q}'(0)\phi + \frac{1}{2!}\mathbf{Q}''(0)\phi^2 + \cdots + \frac{1}{k!}\mathbf{Q}^{(k)}(0)\phi^k + \cdots \tag{2.51}$$

where the superscript (k) stands for the kth derivative of $\mathbf{Q}$ with respect to ϕ. Now, from the definition of $\mathbf{E}$, one can readily prove the relations below:

$$\mathbf{E}^{(2k+1)} = (-1)^k \mathbf{E}, \quad \mathbf{E}^{2k} = (-1)^k (\mathbf{1} - \mathbf{e}\mathbf{e}^T) \tag{2.52}$$

Furthermore, using Eqs. (2.49) and (2.52), one can readily show that

$$\mathbf{Q}^{(k)}(0) = \mathbf{E}^k \tag{2.53}$$

with $\mathbf{E}$ defined already as the cross-product matrix of $\mathbf{e}$. Moreover, from Eqs. (2.51) and (2.53), $\mathbf{Q}(\phi)$ can be expressed as

$$\mathbf{Q}(\phi) = \mathbf{1} + \mathbf{E}\phi + \frac{1}{2!}\mathbf{E}^2\phi^2 + \cdots + \frac{1}{k!}\mathbf{E}^k\phi^k + \cdots$$

whose right-hand side is nothing but the exponential of $\mathbf{E}\phi$, i.e.,

$$\mathbf{Q}(\phi) = e^{\mathbf{E}\phi} \tag{2.54}$$

Equation (2.54) is the exponential representation of the rotation matrix in terms of its *natural invariants*, $\mathbf{e}$ and ϕ. The foregoing parameters are termed *invariants* because they are independent of the coordinate axes chosen to represent the rotation under study. The adjective *natural* is necessary to distinguish them from other invariants that will be introduced presently. This adjective seems suitable because the said invariants stem naturally from Euler's Theorem.

Now, in view of Eqs. (2.52), the above series can be written as

$$\mathbf{Q}(\phi) = \mathbf{1} + \left[-\frac{1}{2!}\phi^2 + \frac{1}{4!}\phi^4 - \cdots + \frac{1}{(2k)!}(-1)^k\phi^{2k} + \cdots \right](\mathbf{1} - \mathbf{e}\mathbf{e}^T)$$
$$+ \left[\phi - \frac{1}{3!}\phi^3 + \cdots + \frac{1}{(2k+1)!}(-1)^k\phi^{2k+1} + \cdots \right]\mathbf{E}$$

The series inside the first pair of brackets is apparently $\cos\phi - 1$, while that in the second pair is $\sin\phi$. We have, therefore, an alternative representation of $\mathbf{Q}$:

$$\mathbf{Q} = \mathbf{1} + \sin\phi\mathbf{E} + (1 - \cos\phi)\mathbf{E}^2 \tag{2.55}$$

which is an expected result in view of the Cayley–Hamilton Theorem.

The Canonical Forms of the Rotation Matrix

The rotation matrix takes on an especially simple form if the axis of rotation coincides with one of the coordinate axes. For example, if the X axis is parallel to the axis of rotation, i.e., parallel to vector $\mathbf{e}$, in a frame that we will label $\mathcal{X}$, then, we will have

$$[\mathbf{e}]_{\mathcal{X}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [\mathbf{E}]_{\mathcal{X}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad [\mathbf{E}^2]_{\mathcal{X}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

In the $\mathcal{X}$-frame, then,

$$[\mathbf{Q}]_{\mathcal{X}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \tag{2.56a}$$

Likewise, if we define the coordinate frames $\mathcal{Y}$ and $\mathcal{Z}$ so that their Y and Z axes, respectively, coincide with the axis of rotation, then

$$[\mathbf{Q}]_{\mathcal{Y}} = \begin{bmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix} \tag{2.56b}$$

and

$$[\mathbf{Q}]_{\mathcal{Z}} = \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.56c}$$

The representations of Eqs. (2.56a–c) can be called the X-, Y-, and Z-*canonical forms* of the rotation matrix. In many instances, a rotation matrix cannot be derived directly from information on the original and the final orientations of a rigid body, but the overall motion can be readily decomposed into a sequence of simple rotations taking the above canonical forms. An application of canonical forms lies in the parameterization of rotations by means of *Euler angles*, consisting of three successive rotations, ϕ, θ and ψ, about one axis of a coordinate frame. Euler angles are introduced in Exercise 2.19, and applications thereof are given in Exercises 2.37, 2.38 and 3.10.

2.3.3 The Linear Invariants of a 3 × 3 *Matrix*

Now we introduce two *linear invariants* of 3 × 3 matrices. Given *any* 3 × 3 matrix **A**, its *Cartesian decomposition*, the counterpart of the Cartesian representation of complex numbers, consists of the sum of its symmetric part, $\mathbf{A}_S$, and its skew-symmetric part, $\mathbf{A}_{SS}$, defined as

$$\mathbf{A}_S \equiv \frac{1}{2}(\mathbf{A} + \mathbf{A}^T), \quad \mathbf{A}_{SS} \equiv \frac{1}{2}(\mathbf{A} - \mathbf{A}^T) \tag{2.57}$$

The *axial vector* or for brevity, the *vector* of **A**, is the vector **a** with the property

$$\mathbf{a} \times \mathbf{v} \equiv \mathbf{A}_{SS}\mathbf{v} \tag{2.58}$$

for any three-dimensional vector **v**. The *trace* of **A** is the sum of the eigenvalues of $\mathbf{A}_S$, which are real. Since no coordinate frame is involved in the above definitions, these are invariant. When calculating these invariants, of course, a particular coordinate frame must be used. Let us assume that the entries of matrix **A** in a certain coordinate frame are given by the array of real numbers a_{ij}, for $i, j = 1, 2, 3$. Moreover, let **a** have components a_i, for $i = 1, 2, 3$, in the same frame. The above-defined invariants are thus calculated as

$$\text{vect}(\mathbf{A}) \equiv \mathbf{a} \equiv \frac{1}{2}\begin{bmatrix} a_{32} - a_{23} \\ a_{13} - a_{31} \\ a_{21} - a_{12} \end{bmatrix}, \quad \text{tr}(\mathbf{A}) \equiv a_{11} + a_{22} + a_{33} \tag{2.59}$$

From the foregoing definitions, we have now

Theorem 2.3.5. *The vector of a* 3×3 *matrix vanishes if and only if it is symmetric, whereas the trace of an* $n \times n$ *matrix vanishes if the matrix is skew symmetric.*

Other useful relations are given below. For any three-dimensional vectors **a** and **b**,

$$\text{vect}(\mathbf{a}\mathbf{b}^T) = -\frac{1}{2}\mathbf{a} \times \mathbf{b} \tag{2.60}$$

and

$$\text{tr}(\mathbf{a}\mathbf{b}^T) = \mathbf{a}^T\mathbf{b} \tag{2.61}$$

The second relation is quite straightforward, but the first one is less so; a proof of the first relation follows: Let **w** denote vect($\mathbf{a}\mathbf{b}^T$). From Definition (2.58), for any three-dimensional vector **v**,

$$\mathbf{w} \times \mathbf{v} = \mathbf{W}\mathbf{v} \tag{2.62}$$

where **W** is the skew-symmetric component of $\mathbf{a}\mathbf{b}^T$, namely,

$$\mathbf{W} \equiv \frac{1}{2}(\mathbf{a}\mathbf{b}^T - \mathbf{b}\mathbf{a}^T) \tag{2.63}$$

and hence,

$$\mathbf{W}\mathbf{v} = \mathbf{w} \times \mathbf{v} = \frac{1}{2}[(\mathbf{b}^T\mathbf{v})\mathbf{a} - (\mathbf{a}^T\mathbf{v})\mathbf{b}] \tag{2.64}$$

Now, let us compare the last expression with the double cross product[1] $(\mathbf{b} \times \mathbf{a}) \times \mathbf{v}$, namely,

$$(\mathbf{b} \times \mathbf{a}) \times \mathbf{v} = (\mathbf{b}^T \mathbf{v})\mathbf{a} - (\mathbf{a}^T \mathbf{v})\mathbf{b} \tag{2.65}$$

from which it becomes apparent that

$$\mathbf{w} = \frac{1}{2}\mathbf{b} \times \mathbf{a} \tag{2.66}$$

thereby proving the aforementioned relation.

Note that Theorem 2.3.5 states a *necessary and sufficient* condition for the vanishing of the vector of a 3×3 matrix, but only a sufficient condition for the vanishing of the trace of a $n \times n$ matrix. What this implies is that the trace of a $n \times n$ matrix can vanish without the matrix being necessarily skew symmetric, but the trace of a skew-symmetric matrix necessarily vanishes. Also note that whereas the vector of a matrix is defined *only* for 3×3 matrices, the trace can be defined more generally for $n \times n$ matrices.

In some applications, the cross-product matrix of the product $\mathbf{Ab}$ of a 3×3 matrix $\mathbf{A}$ by a vector $\mathbf{b}$ is needed[2]:

$$\text{CPM}(\mathbf{Ab}) = (\mathbf{BA})^T - \mathbf{BA} + \text{tr}(\mathbf{A})\mathbf{B} = \left[\text{tr}(\mathbf{A})\mathbf{1} - \mathbf{A}^T\right]\mathbf{B} - \mathbf{BA} \tag{2.67}$$

where $\mathbf{B} = \text{CPM}(\mathbf{b})$. The reader is encouraged to verify the correctness of the above relation using components. *Caveat: a component-free proof of the above relation is particularly challenging.*

2.3.4 The Linear Invariants of a Rotation

From the invariant representations of the rotation matrix, Eqs. (2.49) and (2.55), it is clear that the first two terms of $\mathbf{Q}$, $\mathbf{ee}^T$ and $\cos\phi(\mathbf{1} - \mathbf{ee}^T)$, are symmetric, whereas the third one, $\sin\phi\mathbf{E}$, is skew-symmetric. Hence,

$$\text{vect}(\mathbf{Q}) = \text{vect}(\sin\phi\,\mathbf{E}) = \sin\phi\,\mathbf{e} \tag{2.68}$$

whereas

$$\text{tr}(\mathbf{Q}) = \text{tr}[\mathbf{ee}^T + \cos\phi(\mathbf{1} - \mathbf{ee}^T)] \equiv \mathbf{e}^T\mathbf{e} + \cos\phi(3 - \mathbf{e}^T\mathbf{e}) = 1 + 2\cos\phi \tag{2.69}$$

[1] Popularly known as the *triple cross product*.

[2] This relation was derived by Ph.D. candidate Philippe Cardou.

from which one can readily solve for $\cos\phi$, namely,

$$\cos\phi = \frac{\text{tr}(\mathbf{Q}) - 1}{2} \tag{2.70}$$

Henceforth, the vector of $\mathbf{Q}$ will be denoted by $\mathbf{q}$ and its components in a given coordinate frame by q_1, q_2, and q_3. Moreover, rather than using tr($\mathbf{Q}$) as the other linear invariant, $q_0 \equiv \cos\phi$ will be introduced to refer to the *linear invariants of the rotation matrix*. Hence, the rotation matrix is fully defined by *four scalar parameters*, namely $\{q_i\}_0^3$, which will be conveniently stored in the four-dimensional array $\boldsymbol{\lambda}$, defined as

$$\boldsymbol{\lambda} \equiv [q_1, q_2, q_3, q_0]^T \tag{2.71}$$

Note, however, that the four components of $\boldsymbol{\lambda}$ are not independent, for they obey the relation

$$\|\mathbf{q}\|^2 + q_0^2 \equiv \sin^2\phi + \cos^2\phi = 1 \tag{2.72}$$

Thus, Eq. (2.72) can be written in a more compact form as

$$\|\boldsymbol{\lambda}\|^2 \equiv q_1^2 + q_2^2 + q_3^2 + q_0^2 = 1 \tag{2.73}$$

What Eq. (2.72) states has a straightforward geometric interpretation: As a body rotates about a fixed point, its motion can be described in a four-dimensional space by the motion of a point of position vector $\boldsymbol{\lambda}$ that moves on the surface of the unit sphere centered at the origin of the said space. Alternatively, one can conclude that, as a rigid body rotates about a fixed point, its motion can be described in a three-dimensional space by the motion of position vector $\mathbf{q}$, which moves within the unit solid sphere centered at the origin of the said space. Given the dependence of the four components of vector $\boldsymbol{\lambda}$, one might be tempted to solve for, say, q_0 from Eq. (2.72) in terms of the remaining components, namely, as

$$q_0 = \pm\sqrt{1 - (q_1^2 + q_2^2 + q_3^2)} \tag{2.74}$$

This, however, is not a good idea because the sign ambiguity of Eq. (2.74) leaves angle ϕ undefined, for q_0 is nothing but $\cos\phi$. Moreover, the three components of vector $\mathbf{q}$ alone, i.e., $\sin\phi\,\mathbf{e}$, do not suffice to define the rotation represented by $\mathbf{Q}$. Indeed, from the definition of $\mathbf{q}$, one has

$$\sin\phi = \pm\|\mathbf{q}\|, \quad \mathbf{e} = \mathbf{q}/\sin\phi \tag{2.75}$$

from which it is clear that $\mathbf{q}$ alone does not suffice to define the rotation under study, since it leaves angle ϕ undefined. Indeed, the vector of the rotation matrix provides no information about $\cos\phi$. Yet another representation of the rotation matrix is displayed below, in terms of its linear invariants, that is readily derived from representations (2.49) and (2.55), namely,

$$\mathbf{Q} = \frac{\mathbf{q}\mathbf{q}^T}{\|\mathbf{q}\|^2} + q_0\left(\mathbf{1} - \frac{\mathbf{q}\mathbf{q}^T}{\|\mathbf{q}\|^2}\right) + \overline{\mathbf{Q}} \tag{2.76a}$$

in which $\overline{\mathbf{Q}}$ is the cross-product matrix of vector $\mathbf{q}$, i.e.,

$$\overline{\mathbf{Q}} \equiv \frac{\partial(\mathbf{q}\times\mathbf{x})}{\partial\mathbf{x}}$$

for any vector $\mathbf{x}$.

Note that by virtue of Eq. (2.72), the representation of $\mathbf{Q}$ given in Eq. (2.76a) can be expressed alternatively as

$$\mathbf{Q} = q_0\mathbf{1} + \overline{\mathbf{Q}} + \frac{\mathbf{q}\mathbf{q}^T}{1+q_0} \tag{2.76b}$$

From either Eq. (2.76a) or (2.76b) it is apparent that linear invariants are not suitable to represent a rotation when the associated angle is either π or close to it. Note that a rotation through an angle ϕ about an axis given by vector $\mathbf{e}$ is identical to a rotation through an angle $-\phi$ about an axis given by vector $-\mathbf{e}$. Hence, changing the sign of $\mathbf{e}$ does not change the rotation matrix, provided that the sign of ϕ is also changed. Henceforth, we will choose the sign of the components of $\mathbf{e}$ so that $\sin\phi \geq 0$, which is equivalent to assuming that $0 \leq \phi \leq \pi$. Thus, $\sin\phi$ is calculated as $\|\mathbf{q}\|$, while $\cos\phi$ as indicated in Eq. (2.70). Obviously, $\mathbf{e}$ is simply $\mathbf{q}$ normalized, i.e., $\mathbf{q}$ divided by its Euclidean norm.

2.3.5 Examples

The examples below are meant to stress the foregoing ideas on rotation invariants.

Example 2.3.3. If $[\mathbf{e}]_{\mathcal{F}} = [\sqrt{3}/3,\ -\sqrt{3}/3,\ \sqrt{3}/3\,]^T$ in a given coordinate frame $\mathcal{F}$ and $\phi = 120°$, what is $\mathbf{Q}$ in $\mathcal{F}$?

Solution: From the data,

$$\cos\phi = -\frac{1}{2}, \quad \sin\phi = \frac{\sqrt{3}}{2}$$

Moreover, in the $\mathcal{F}$ frame,

$$[\mathbf{e}\mathbf{e}^T]_{\mathcal{F}} = \frac{1}{3}\begin{bmatrix}1\\-1\\1\end{bmatrix}\begin{bmatrix}1 & -1 & 1\end{bmatrix} = \frac{1}{3}\begin{bmatrix}1 & -1 & 1\\-1 & 1 & -1\\1 & -1 & 1\end{bmatrix}$$

and hence,

$$[\mathbf{1} - \mathbf{e}\mathbf{e}^T]_{\mathcal{F}} = \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}, \quad [\mathbf{E}]_{\mathcal{F}} \equiv \frac{\sqrt{3}}{3}\begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}$$

Thus, from Eq. (2.49),

$$[\mathbf{Q}]_{\mathcal{F}} = \frac{1}{3}\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} - \frac{1}{6}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}$$

i.e.,

$$[\mathbf{Q}]_{\mathcal{F}} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 0 & 0 \end{bmatrix}$$

Example 2.3.4. The matrix representation of a linear transformation $\mathbf{Q}$ in a certain reference frame $\mathcal{F}$ is given below. Find out whether the said transformation is a rigid-body rotation. If it is, find its natural invariants.

$$[\mathbf{Q}]_{\mathcal{F}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Solution: First the given array is tested for orthogonality:

$$[\mathbf{Q}]_{\mathcal{F}}[\mathbf{Q}^T]_{\mathcal{F}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

thereby showing that the said array is indeed orthogonal. Thus, the linear transformation could represent a reflection or a rotation. In order to decide which one this represents, the determinant of the foregoing array is computed:

$$\det(\mathbf{Q}) = +1$$

which makes apparent that $\mathbf{Q}$ indeed represents a rigid-body rotation. Now, its natural invariants are computed. The unit vector $\mathbf{e}$ can be computed as the eigenvector of $\mathbf{Q}$ associated with the eigenvalue $+1$. This requires, however, finding a nontrivial solution of a homogeneous linear system of three equations in three unknowns. This is not difficult to do, but it is cumbersome and is not necessary. In order to find $\mathbf{e}$ and ϕ, it is recalled that $\text{vect}(\mathbf{Q}) = \sin\phi\,\mathbf{e}$, which is readily computed with differences only, as indicated in Eq. (2.59), namely,

$$[\mathbf{q}]_{\mathcal{F}} \equiv \sin\phi\,[\mathbf{e}]_{\mathcal{F}} = -\frac{1}{2}\begin{bmatrix}1\\1\\1\end{bmatrix}$$

Under the assumption that $\sin\phi \geq 0$, then,

$$\sin\phi \equiv \|\mathbf{q}\| = \frac{\sqrt{3}}{2}$$

and hence,

$$[\mathbf{e}]_{\mathcal{F}} = \frac{[\mathbf{q}]_{\mathcal{F}}}{\|\mathbf{q}\|} = -\frac{\sqrt{3}}{3}\begin{bmatrix}1\\1\\1\end{bmatrix}$$

and

$$\phi = 60^\circ \quad \text{or} \quad 120^\circ$$

The foregoing ambiguity is resolved by the trace of $\mathbf{Q}$, which yields

$$1 + 2\cos\phi \equiv \text{tr}(\mathbf{Q}) = 0, \quad \cos\phi = -\frac{1}{2}$$

The negative sign of $\cos\phi$ indicates that ϕ lies in the second quadrant—it cannot lie in the third quadrant because of our assumption about the sign of $\sin\phi$—and hence

$$\phi = 120^\circ$$

Example 2.3.5. A coordinate frame X_1, Y_1, Z_1 is rotated into a configuration X_2, Y_2, Z_2 in such a way that

$$X_2 = -Y_1, \quad Y_2 = Z_1, \quad Z_2 = -X_1$$

Find the matrix representation of the rotation in X_1, Y_1, Z_1 coordinates. From this representation, compute the direction of the axis and the angle of rotation.

Solution: Let $\mathbf{i}_1, \mathbf{j}_1, \mathbf{k}_1$ be unit vectors parallel to X_1, Y_1, Z_1, respectively, $\mathbf{i}_2, \mathbf{j}_2, \mathbf{k}_2$ being defined correspondingly. One has

$$\mathbf{i}_2 = -\mathbf{j}_1, \quad \mathbf{j}_2 = \mathbf{k}_1, \quad \mathbf{k}_2 = -\mathbf{i}_1$$

and hence, from Definition 2.2.1, the matrix representation $[\mathbf{Q}]_1$ of the rotation under study in the X_1, Y_1, Z_1 coordinate frame is readily derived:

$$[\mathbf{Q}]_1 = \begin{bmatrix}0 & 0 & -1\\-1 & 0 & 0\\0 & 1 & 0\end{bmatrix}$$

from which the linear invariants follow, namely,

$$[\mathbf{q}]_1 \equiv [\,\mathrm{vect}(\mathbf{Q})\,]_1 = \sin\phi\,[\,\mathbf{e}\,]_1 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \quad \cos\phi = \frac{1}{2}[\,\mathrm{tr}(\mathbf{Q}) - 1\,] = -\frac{1}{2}$$

Under our assumption that $\sin\phi \geq 0$, we obtain

$$\sin\phi = \|\mathbf{q}\| = \frac{\sqrt{3}}{2}, \quad [\,\mathbf{e}\,]_1 = \frac{[\mathbf{q}]_1}{\sin\phi} = \frac{\sqrt{3}}{3}\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

From the foregoing values for $\sin\phi$ and $\cos\phi$, angle ϕ is computed uniquely as

$$\phi = 120^\circ$$

Example 2.3.6. Show that the matrix **P** given in Eq. (2.4) satisfies properties (2.1a).

Solution: First, we prove idempotency, i.e.,

$$\begin{aligned}\mathbf{P}^2 &= (\mathbf{1} - \mathbf{n}\mathbf{n}^T)(\mathbf{1} - \mathbf{n}\mathbf{n}^T) \\ &= \mathbf{1} - 2\mathbf{n}\mathbf{n}^T + \mathbf{n}\mathbf{n}^T\mathbf{n}\mathbf{n}^T = \mathbf{1} - \mathbf{n}\mathbf{n}^T = \mathbf{P}\end{aligned}$$

thereby showing that **P** is, indeed, idempotent. Now we prove that **n** is an eigenvector of **P** with eigenvalue, 0 and hence, **n** spans the null space of **P**. In fact,

$$\mathbf{P}\mathbf{n} = (\mathbf{1} - \mathbf{n}\mathbf{n}^T)\mathbf{n} = \mathbf{n} - \mathbf{n}\mathbf{n}^T\mathbf{n} = \mathbf{n} - \mathbf{n} = \mathbf{0}$$

thereby completing the proof.

Example 2.3.7. The representations of three linear transformations in a given coordinate frame $\mathcal{F}$ are given below:

$$[\,\mathbf{A}\,]_{\mathcal{F}} = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ -1 & -2 & 2 \end{bmatrix}$$

$$[\,\mathbf{B}\,]_{\mathcal{F}} = \frac{1}{3}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

$$[\,\mathbf{C}\,]_{\mathcal{F}} = \frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}$$

One of the foregoing matrices is an orthogonal projection, one is a reflection, and one is a rotation. Identify each of these and give its invariants.

Solution: From representations (2.49) and (2.55), it is clear that a rotation matrix is symmetric if and only if $\sin\phi = 0$. This means that a rotation matrix cannot be symmetric unless its angle of rotation is either 0 or π, i.e., unless its trace is either 3 or -1. Since $[\mathbf{B}]_{\mathcal{F}}$ and $[\mathbf{C}]_{\mathcal{F}}$ are symmetric, they cannot be rotations, unless their traces take the foregoing values. Their traces are thus evaluated below:

$$\mathrm{tr}(\mathbf{B}) = 2, \quad \mathrm{tr}(\mathbf{C}) = 1$$

which thus rules out the foregoing matrices as suitable candidates for rotations. Thus, $\mathbf{A}$ is the only candidate left for proper orthogonality, its suitability being tested below:

$$[\mathbf{A}\mathbf{A}^T]_{\mathcal{F}} = \frac{1}{9}\begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}, \quad \det(\mathbf{A}) = +1$$

and hence, $\mathbf{A}$ indeed represents a rotation. Its natural invariants are next computed:

$$\sin\phi\,[\mathbf{e}]_{\mathcal{F}} = [\mathrm{vect}(\mathbf{A})]_{\mathcal{F}} = \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}, \quad \cos\phi = \frac{1}{2}[\mathrm{tr}(\mathbf{A}) - 1] = \frac{1}{2}(2-1) = \frac{1}{2}$$

We assume, as usual, that $\sin\phi \geq 0$. Then,

$$\sin\phi = \|\mathrm{vect}(\mathbf{A})\| = \frac{\sqrt{3}}{2}, \quad \text{i.e.,} \quad \phi = 60^\circ$$

Moreover,

$$[\mathbf{e}]_{\mathcal{F}} = \frac{[\mathrm{vect}(\mathbf{A})]_{\mathcal{F}}}{\|\mathrm{vect}(\mathbf{A})\|} = \frac{\sqrt{3}}{3}\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

Now, one matrix of $\mathbf{B}$ and $\mathbf{C}$ is an orthogonal projection and the other is a reflection. To be a reflection, a matrix has to be orthogonal. Hence, each matrix is tested for orthogonality:

$$[\mathbf{B}\mathbf{B}^T]_{\mathcal{F}} = \frac{1}{9}\begin{bmatrix} 6 & 3 & 3 \\ 3 & 6 & -3 \\ 3 & -3 & 6 \end{bmatrix} = [\mathbf{B}^2]_{\mathcal{F}} = [\mathbf{B}]_{\mathcal{F}}, \;\; [\mathbf{C}\mathbf{C}^T]_{\mathcal{F}} = \frac{1}{9}\begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

thereby showing that $\mathbf{C}$ is orthogonal and $\mathbf{B}$ is not. Furthermore, $\det(\mathbf{C}) = -1$, which confirms that $\mathbf{C}$ is a reflection. Now, if $\mathbf{B}$ is a projection, it is bound to be singular and idempotent. From the orthogonality test it is clear that it is idempotent.

Moreover, one can readily verify that $\det(\mathbf{B}) = 0$, and hence $\mathbf{B}$ is singular, the unit vector $[\,\mathbf{n}\,]_{\mathcal{F}} = [\,n_1,\, n_2,\, n_3\,]^T$ that spans its null space being determined from the general form of projections, Eq. (2.1a), whence,

$$\mathbf{n}\mathbf{n}^T = \mathbf{1} - \mathbf{B}$$

Therefore, if a solution $\mathbf{n}$ has been found, then $-\mathbf{n}$ is also a solution, i.e., *the problem admits two solutions*, one being the negative of the other. These two solutions are found below, by first rewriting the above system of equations in component form:

$$\begin{bmatrix} n_1^2 & n_1n_2 & n_1n_3 \\ n_1n_2 & n_2^2 & n_2n_3 \\ n_1n_3 & n_2n_3 & n_3^2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

Now, from the diagonal entries of the above matrices, it is apparent that the three components of $\mathbf{n}$ have identical absolute values, i.e., $\sqrt{3}/3$. Moreover, from the off-diagonal entries of the same matrices, the second and third components of $\mathbf{n}$ bear equal signs, but we cannot tell whether positive or negative, because of the quadratic nature of the problem at hand. The two solutions are thus obtained as

$$\mathbf{n} = \pm\frac{\sqrt{3}}{3}\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

which is the only invariant of $\mathbf{B}$.

We now look at $\mathbf{C}$, which is a reflection, and apparently, bears the form

$$\mathbf{C} = \mathbf{1} - 2\mathbf{n}\mathbf{n}^T$$

In order to determine $\mathbf{n}$, note that

$$\mathbf{n}\mathbf{n}^T = \frac{1}{2}(\mathbf{1} - \mathbf{C})$$

or in component form,

$$\begin{bmatrix} n_1^2 & n_1n_2 & n_1n_3 \\ n_1n_2 & n_2^2 & n_2n_3 \\ n_1n_3 & n_2n_3 & n_3^2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

which is identical to the matrix equation derived in the case of matrix $\mathbf{B}$. Hence, the solution is the same, i.e.,

$$\mathbf{n} = \pm\frac{\sqrt{3}}{3}\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

thereby finding the invariant sought.

Example 2.3.8. The vector and the trace of a rotation matrix $\mathbf{Q}$, in a certain reference frame $\mathcal{F}$, are given as

$$[\text{vect}(\mathbf{Q})]_{\mathcal{F}} = \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}, \quad \text{tr}(\mathbf{Q}) = 2$$

Find the matrix representation of $\mathbf{Q}$ in the given coordinate frame and in a frame having its Z-axis parallel to vect($\mathbf{Q}$).

Solution: We shall resort to Eq. (2.76a) to determine the rotation matrix $\mathbf{Q}$. The quantities involved in the representation of $\mathbf{Q}$ in $\mathcal{F}$ are readily computed:

$$[\mathbf{q}\mathbf{q}^T]_{\mathcal{F}} = \frac{1}{4}\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}, \quad \|\mathbf{q}\|^2 = \frac{3}{4}, \quad [\overline{\mathbf{Q}}]_{\mathcal{F}} = \frac{1}{2}\begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

from which $\mathbf{Q}$ follows:

$$[\mathbf{Q}]_{\mathcal{F}} = \frac{1}{3}\begin{bmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ -1 & -2 & 2 \end{bmatrix}$$

in the given coordinate frame. Now, let $\mathcal{Z}$ denote a coordinate frame whose Z-axis is parallel to $\mathbf{q}$. Hence,

$$[\mathbf{q}]_{\mathcal{Z}} = \frac{\sqrt{3}}{2}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad [\mathbf{q}\mathbf{q}^T]_{\mathcal{Z}} = \frac{3}{4}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad [\overline{\mathbf{Q}}]_{\mathcal{Z}} = \frac{\sqrt{3}}{2}\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

which readily leads to

$$[\mathbf{Q}]_{\mathcal{Z}} = \begin{bmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and is in the Z-canonical form.

Example 2.3.9. A procedure for trajectory planning produced a matrix representing a rotation for a certain *pick-and-place operation*, as shown below:

$$[\mathbf{Q}] = \begin{bmatrix} 0.433 & -0.500 & z \\ x & 0.866 & -0.433 \\ 0.866 & y & 0.500 \end{bmatrix}$$

where x, y, and z are entries that are unrecognizable due to failures in the printing hardware. Knowing that **Q** is in fact a rotation matrix, find the missing entries.

Solution: Since **Q** is a rotation matrix, the product $\mathbf{P} \equiv \mathbf{Q}^T\mathbf{Q}$ should equal the 3×3 identity matrix, and det(**Q**) should be $+1$. The foregoing product is computed first:

$$[\mathbf{P}]_{\mathcal{F}} = \begin{bmatrix} 0.437 + z^2 & 0.433(x - z - 1) & 0.5(-y + z) + 0.375 \\ * & 0.937 + x^2 & 0.866(x + y) - 0.216 \\ * & * & 1 + y^2 \end{bmatrix}$$

where the entries below the diagonal need not be printed because the matrix is symmetric. Upon equating the diagonal entries of the foregoing array to unity, we obtain

$$x = \pm 0.250, \quad y = 0, \quad z = \pm 0.750$$

while the vanishing of the off-diagonal entries leads to

$$x = 0.250, \quad y = 0, \quad z = -0.750$$

which can be readily verified to produce $\det(\mathbf{Q}) = +1$.

2.3.6 The Euler–Rodrigues Parameters

The invariants defined so far, namely, the natural and the linear invariants of a rotation matrix, are not the only ones that are used in kinematics. Additionally, one has the *Euler parameters*, or *Euler–Rodrigues parameters*, as Cheng and Gupta (1989) propose that they should be called, represented here as **r** and r_0. The Euler–Rodrigues parameters are defined as

$$\mathbf{r} \equiv \sin\left(\frac{\phi}{2}\right)\mathbf{e}, \quad r_0 = \cos\left(\frac{\phi}{2}\right) \tag{2.77}$$

One can readily show that **Q** takes on a quite simple form in terms of the Euler–Rodrigues parameters, namely,

$$\mathbf{Q} = ({r_0}^2 - \mathbf{r} \cdot \mathbf{r})\mathbf{1} + 2\mathbf{r}\mathbf{r}^T + 2r_0\mathbf{R} \tag{2.78}$$

in which **R** is the cross-product matrix of **r**, i.e.,

$$\mathbf{R} \equiv \frac{\partial(\mathbf{r} \times \mathbf{x})}{\partial \mathbf{x}}$$

for arbitrary **x**.

Note that the Euler–Rodrigues parameters appear quadratically in the rotation matrix. Hence, these parameters cannot be computed with simple sums and differences. A closer inspection of Eq. (2.76b) reveals that the linear invariants appear *almost linearly* in the rotation matrix. This means that the rotation matrix, as given by Eq. (2.76b), is composed of two types of terms, namely, linear and rational. Moreover, the rational term is composed of a quadratic expression in the numerator and a linear expression in the denominator, the ratio thus being linear, which explains why the linear invariants can be obtained by sums and differences from the rotation matrix.

The relationship between the linear invariants and the Euler–Rodrigues parameters can be readily derived, namely,

$$r_0 = \pm\sqrt{\frac{1+q_0}{2}}, \quad \mathbf{r} = \frac{\mathbf{q}}{2r_0}, \quad \phi \neq \pi \tag{2.79}$$

Furthermore, note that, if $\phi = \pi$, then $r_0 = 0$, and formulae (2.79) fail to produce $\mathbf{r}$. However, from Eq. (2.77),

$$\text{For } \phi = \pi: \quad \mathbf{r} = \mathbf{e}, \quad r_0 = 0 \tag{2.80}$$

We now derive invariant relations between the rotation matrix and the Euler–Rodrigues parameters. To do this, we resort to the concept of *matrix square root*. As a matter of fact, the square root of a square matrix is nothing but a particular case of an *analytic function* of a square matrix, discussed in connection with Theorem 2.3.3 and the exponential representation of the rotation matrix. Indeed, the square root of a square matrix is an analytic function of that matrix, and hence, admits a series expansion in powers of the matrix. Moreover, by virtue of the Cayley–Hamilton Theorem (Theorem 2.3.3) the said square root should be, for a 3×3 matrix, a linear combination of the identity matrix $\mathbf{1}$, the matrix itself, and its square, the coefficients being found using the eigenvalues of the matrix.

Furthermore, from the geometric meaning of a rotation through the angle ϕ about an axis parallel to the unit vector $\mathbf{e}$, it is apparent that the square of the matrix representing the foregoing rotation is itself a rotation about the same axis, but through the angle 2ϕ. By the same token, the square root of the rotation matrix is again a rotation matrix about the same axis, but through an angle $\phi/2$. Now, while the square of a matrix is unique, its square root is not. This fact is apparent for diagonalizable matrices, whose diagonal entries are their eigenvalues. Each eigenvalue, whether positive or negative, admits two square roots, and hence, a diagonalizable $n \times n$ matrix admits as many square roots as there are combinations of the two possible roots of individual eigenvalues, disregarding rearrangements of the latter. Such a number is 2^n, and hence, a 3×3 matrix admits eight square roots. For example, the eight square roots of the identity 3×3 matrix are displayed below:

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}, \quad \begin{bmatrix}1&0&0\\0&1&0\\0&0&-1\end{bmatrix}, \quad \begin{bmatrix}1&0&0\\0&-1&0\\0&0&1\end{bmatrix}, \quad \begin{bmatrix}-1&0&0\\0&1&0\\0&0&1\end{bmatrix},$$

$$\begin{bmatrix}1&0&0\\0&-1&0\\0&0&-1\end{bmatrix}, \quad \begin{bmatrix}-1&0&0\\0&1&0\\0&0&-1\end{bmatrix}, \quad \begin{bmatrix}-1&0&0\\0&-1&0\\0&0&1\end{bmatrix}, \quad \begin{bmatrix}-1&0&0\\0&-1&0\\0&0&-1\end{bmatrix}$$

In fact, the foregoing result can be extended to orthogonal matrices as well and, for that matter, to any square matrix with n linearly independent eigenvectors. That is, an $n \times n$ orthogonal matrix admits 2^n square roots. However, not all eight square roots of a 3×3 orthogonal matrix are orthogonal. In fact, not all eight square roots of a 3×3 proper orthogonal matrix are proper orthogonal either. Of these square roots, nevertheless, there is one that is proper orthogonal, the one representing a rotation of $\phi/2$. We will denote this particular square root of $\mathbf{Q}$ by $\sqrt{\mathbf{Q}}$. The Euler–Rodrigues parameters of $\mathbf{Q}$ can thus be expressed as the linear invariants of $\sqrt{\mathbf{Q}}$, namely,

$$\mathbf{r} = \text{vect}(\sqrt{\mathbf{Q}}), \quad r_0 = \frac{\text{tr}(\sqrt{\mathbf{Q}}) - 1}{2} \tag{2.81}$$

It is important to recognize the basic differences between the linear invariants and the Euler–Rodrigues parameters. Whereas the former can be readily derived from the matrix representation of the rotation involved by simple additions and subtractions, the latter require square roots and entail sign ambiguities. However, the former fail to produce information on the axis of rotation whenever the angle of rotation is π, whereas the latter produce that information *for any value of the angle of rotation*.

The Euler–Rodrigues parameters are nothing but the *quaternions* invented by Sir William Rowan Hamilton (1844) in an extraordinary moment of creativity on Monday, October 16, 1843, as "Hamilton, accompanied by Lady Hamilton, was walking along the Royal Canal in Dublin towards the Royal Irish Academy, where Hamilton was to preside a meeting" (Altmann 1989).

Moreover, the Euler–Rodrigues parameters should not be confused with the *Euler angles*, which are not invariant and hence, admit multiple definitions. The foregoing means that no single set of Euler angles exists for a given rotation matrix, the said angles depending on how the rotation is decomposed into three simpler rotations. For this reason, Euler angles will not be stressed here. The reader is referred to Exercise 2.19 for a short discussion of Euler angles; Synge (1960) includes a classical treatment, while Kane et al. (1983) provide an extensive discussion of the same.

Example 2.3.10. Find the Euler–Rodrigues parameters of the proper orthogonal matrix $\mathbf{Q}$ given as

$$\mathbf{Q} = \frac{1}{3}\begin{bmatrix}-1&2&2\\2&-1&2\\2&2&-1\end{bmatrix}$$

Solution: Since the given matrix is symmetric, its angle of rotation is π and its vector linear invariant vanishes, which prevents us from finding the direction of the axis of rotation from the linear invariants; moreover, expressions (2.79) do not apply. However, we can use Eq. (2.50) to find the unit vector $\mathbf{e}$ parallel to the axis of rotation, i.e.,

$$\mathbf{e}\mathbf{e}^T = \frac{1}{2}(\mathbf{1} + \mathbf{Q})$$

or in component form,

$$\begin{bmatrix} e_1^2 & e_1e_2 & e_1e_3 \\ e_1e_2 & e_2^2 & e_2e_3 \\ e_1e_3 & e_2e_3 & e_3^2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

A simple inspection of the components of the two sides of the above equation reveals that all three components of $\mathbf{e}$ are identical and moreover, of the same sign, but we cannot tell which sign this is. Therefore,

$$\mathbf{e} = \pm\frac{\sqrt{3}}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Moreover, from the symmetry of $\mathbf{Q}$, we know that $\phi = \pi$, and hence,

$$\mathbf{r} = \mathbf{e}\sin\left(\frac{\phi}{2}\right) = \pm\frac{\sqrt{3}}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad r_0 = \cos\left(\frac{\phi}{2}\right) = 0$$

2.4 Composition of Reflections and Rotations

As pointed out in Sect. 2.2, reflections occur often accompanied by rotations. The effect of this combination is that the rotation destroys the two properties of pure reflections, symmetry and self-inversion, as defined in Sect. 2.2. Indeed, let $\mathbf{R}$ be a pure reflection, taking on the form appearing in Eq. (2.5), and $\mathbf{Q}$ an arbitrary rotation, taking on the form of Eq. (2.49). The product of these two transformations, $\mathbf{QR}$, denoted by $\mathbf{T}$, is apparently neither symmetric nor self-inverse, as the reader can readily verify. Likewise, the product of these two transformations in the reverse order is neither symmetric nor self-inverse.

As a consequence of the foregoing discussion, an improper orthogonal transformation that is not symmetric can always be decomposed into the product of a rotation and a pure reflection, the latter being symmetric and self-inverse. Moreover, this decomposition can take on the form of any of the two possible orderings of the rotation and the reflection. Note, however, that once the order has been selected, the decomposition is not unique. Indeed, if we want to decompose $\mathbf{T}$ in the above paragraph into the product $\mathbf{QR}$, then we can freely choose the unit normal $\mathbf{n}$ of the plane of reflection and write

$$\mathbf{R} \equiv \mathbf{1} - 2\mathbf{n}\mathbf{n}^T$$

vector $\mathbf{n}$ then being found from

$$\mathbf{n}\mathbf{n}^T = \frac{1}{2}(\mathbf{1} - \mathbf{R})$$

Hence, the factor $\mathbf{Q}$ of that decomposition is obtained as

$$\mathbf{Q} = \mathbf{T}\mathbf{R}^{-1} \equiv \mathbf{T}\mathbf{R} = \mathbf{T} - 2(\mathbf{T}\mathbf{n})\mathbf{n}^T$$

where use has been made of the self-inverse property of $\mathbf{R}$. Any other selection of vector $\mathbf{n}$ will lead to a different decomposition of $\mathbf{T}$.

Example 2.4.1. Join the palms of your two hands in the position adopted by swimmers when preparing for plunging, while holding a sheet of paper between them. The sheet defines a plane in each hand that we will call the *hand plane*, its unit normal, pointing outside of the hand, being called the *hand normal* and represented as vectors $\mathbf{n}_R$ and $\mathbf{n}_L$ for the right and left hand, respectively. Moreover, let $\mathbf{o}_R$ and $\mathbf{o}_L$ denote unit vectors pointing in the direction of the finger axes of each of the two hands. Thus, in the swimmer position described above, $\mathbf{n}_L = -\mathbf{n}_R$ and $\mathbf{o}_L = \mathbf{o}_R$. Now, without moving your right hand, let the left hand attain a position whereby the left-hand normal lies at right angles with the right-hand normal, the palm pointing downwards and the finger axes of the two hands remaining parallel. Find the representation of the transformation carrying the right hand to the final configuration of the left hand, in terms of the unit vectors $\mathbf{n}_R$ and $\mathbf{o}_R$.

Solution: Let us regard the desired transformation $\mathbf{T}$ as the product of a rotation $\mathbf{Q}$ by a pure reflection $\mathbf{R}$, in the form $\mathbf{T} = \mathbf{QR}$. Thus, the transformation occurs so that the reflection takes place first, then the rotation. The reflection is simply that mapping the right hand into the left hand, and hence, the reflection plane is simply the hand plane, i.e.,

$$\mathbf{R} = \mathbf{1} - 2\mathbf{n}_R\mathbf{n}_R^T$$

Moreover, the left hand rotates from the swimmer position about an axis parallel to the finger axes through an angle of 90° clockwise from your viewpoint, i.e., in the positive direction of vector $\mathbf{o}_R$. Hence, the form of the rotation involved can be derived readily from Eq. (2.49) and the above information, namely,

$$\mathbf{Q} = \mathbf{o}_R\mathbf{o}_R^T + \mathbf{O}_R$$

where $\mathbf{O}_R$ is the cross-product matrix of $\mathbf{o}_R$. Hence, upon performing the product $\mathbf{QR}$, we have

$$\mathbf{T} = \mathbf{o}_R\mathbf{o}_R^T + 2\mathbf{O}_R - 2(\mathbf{o}_R \times \mathbf{n}_R)\mathbf{n}_R^T$$

which is the transformation sought.

2.5 Coordinate Transformations and Homogeneous Coordinates

Crucial to robotics is the unambiguous description of the geometric relations among the various bodies in the environment surrounding a robot. These relations are established by means of *coordinate frames*, or *frames*, for brevity, attached to each rigid body in the scene, including the robot links. The origins of these frames, moreover, are set at landmark points and orientations defined by key geometric entities like lines and planes. For example, in Chap. 4 we attach two frames to every moving link of a serial robot, with origin at a point on each of the axis of the two joints coupling this link with its two neighbors. Moreover, the Z-axis of each frame is defined, according to the Denavit–Hartenberg notation, introduced in that chapter, along each joint axis, while the X-axis of the frame closer to the base—termed the fore frame—is defined along the common perpendicular to the two joint axes. The origin of the same frame is thus defined as the intersection of the fore axis with the common perpendicular to the two axes. This section is devoted to the study of the coordinate transformations of vectors when these are represented in various frames.

2.5.1 Coordinate Transformations Between Frames with a Common Origin

We will refer to two coordinate frames in this section, namely, $\mathcal{A} = \{X, Y, Z\}$ and $\mathcal{B} = \{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$. Moreover, let $\mathbf{Q}$ be the rotation carrying $\mathcal{A}$ into $\mathcal{B}$, i.e.,

$$\mathbf{Q}\colon \quad \mathcal{A} \quad \to \quad \mathcal{B} \tag{2.82}$$

The purpose of this subsection is to establish the relation between the representations of the position vector of a point P in $\mathcal{A}$ and in $\mathcal{B}$, denoted by $[\,\mathbf{p}\,]_{\mathcal{A}}$ and $[\,\mathbf{p}\,]_{\mathcal{B}}$, respectively. Let

$$[\,\mathbf{p}\,]_{\mathcal{A}} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{2.83}$$

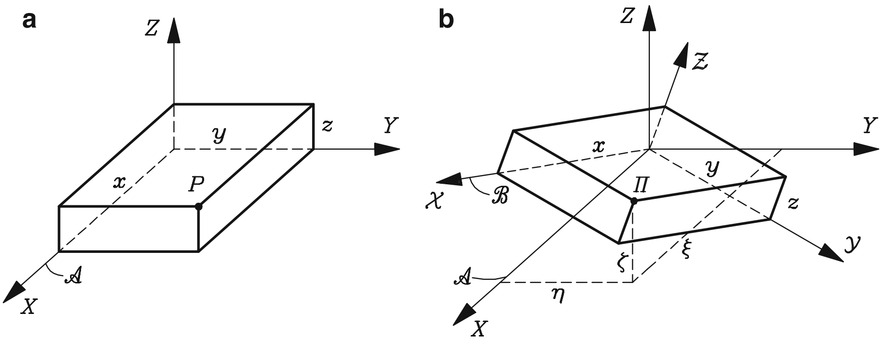

Fig. 2.4 Coordinate transformation: (**a**) coordinates of point P in the $\mathcal{A}$-frame; and (**b**) relative orientation of frame $\mathcal{B}$ with respect to $\mathcal{A}$

We want to find $[\,\mathbf{p}\,]_{\mathcal{B}}$ in terms of $[\,\mathbf{p}\,]_{\mathcal{A}}$ and $\mathbf{Q}$, when the latter is represented in either frame. The coordinate transformation can best be understood if we regard point P as attached to frame $\mathcal{A}$, as if it were a point of a box with sides of lengths x, y, and z, as indicated in Fig. 2.4a. Now, frame $\mathcal{A}$ undergoes a rotation $\mathbf{Q}$ about its origin that carries it into a new attitude, that of frame $\mathcal{B}$, as illustrated in Fig. 2.4b. Point P in its rotated position is labeled Π, of position vector $\boldsymbol{\pi}$, i.e.,

$$\boldsymbol{\pi} = \mathbf{Q}\mathbf{p} \tag{2.84}$$

It is apparent that the relative position of point P with respect to its box does not change under the foregoing rotation, and hence,

$$[\,\boldsymbol{\pi}\,]_{\mathcal{B}} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{2.85}$$

Moreover, let

$$[\,\boldsymbol{\pi}\,]_{\mathcal{A}} = \begin{bmatrix} \xi \\ \eta \\ \zeta \end{bmatrix} \tag{2.86}$$

The relation between the two representations of the position vector of any point of the three-dimensional Euclidean space is given by

Theorem 2.5.1. *The representations of the position vector* $\boldsymbol{\pi}$ *of any point in two frames* $\mathcal{A}$ *and* $\mathcal{B}$*, denoted by* $[\,\boldsymbol{\pi}\,]_{\mathcal{A}}$ *and* $[\,\boldsymbol{\pi}\,]_{\mathcal{B}}$*, respectively, are related by*

$$[\,\boldsymbol{\pi}\,]_{\mathcal{A}} = [\,\mathbf{Q}\,]_{\mathcal{A}}[\,\boldsymbol{\pi}\,]_{\mathcal{B}} \tag{2.87}$$

Proof. Let us write Eq. (2.84) in $\mathcal{A}$:

$$[\boldsymbol{\pi}]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{A}}[\mathbf{p}]_{\mathcal{A}} \tag{2.88}$$

Now, from Fig. 2.4b and Eqs. (2.83) and (2.85) it is apparent that

$$[\boldsymbol{\pi}]_{\mathcal{B}} = [\mathbf{p}]_{\mathcal{A}} \tag{2.89}$$

Upon substituting Eq. (2.89) into Eq. (2.88), we obtain

$$[\boldsymbol{\pi}]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{A}}[\boldsymbol{\pi}]_{\mathcal{B}} \tag{2.90}$$

q.e.d. Moreover, we have

Theorem 2.5.2. *The representations of* $\mathbf{Q}$ *carrying* $\mathcal{A}$ *into* $\mathcal{B}$ *in these two frames are identical, i.e.,*

$$[\mathbf{Q}]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{B}} \tag{2.91}$$

Proof. Upon substitution of Eq. (2.84) into Eq. (2.87), we obtain

$$[\mathbf{Qp}]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{A}}[\mathbf{Qp}]_{\mathcal{B}}$$

or

$$[\mathbf{Q}]_{\mathcal{A}}[\mathbf{p}]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{A}}[\mathbf{Qp}]_{\mathcal{B}}$$

Now, since $\mathbf{Q}$ is orthogonal, it is nonsingular, and hence, $[\mathbf{Q}]_{\mathcal{A}}$ can be deleted from the foregoing equation, thus leading to

$$[\mathbf{p}]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{B}}[\mathbf{p}]_{\mathcal{B}} \tag{2.92}$$

However, by virtue of Theorem 2.5.1, the two representations of $\mathbf{p}$ observe the relation

$$[\mathbf{p}]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{A}}[\mathbf{p}]_{\mathcal{B}} \tag{2.93}$$

the theorem being proved upon equating the right-hand sides of Eqs. (2.92) and (2.93).

Note that the foregoing theorem states a relation valid only for the conditions stated therein. The reader should not conclude from this result that rotation matrices have the same representations in every frame. This point is stressed in Example 2.5.1. Furthermore, we have

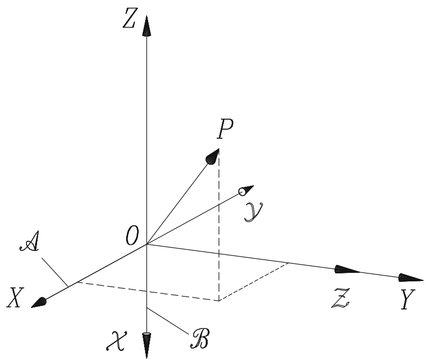

Fig. 2.5 Coordinate frames $\mathcal{A}$ and $\mathcal{B}$ with a common origin

Theorem 2.5.3. *The inverse relation of Theorem 2.5.1 is given by*

$$[\,\boldsymbol{\pi}\,]_{\mathcal{B}} = [\,\mathbf{Q}^T\,]_{\mathcal{B}}[\,\boldsymbol{\pi}\,]_{\mathcal{A}} \tag{2.94}$$

Proof. This is straightforward in light of the two foregoing theorems, and is left to the reader as an exercise.

Example 2.5.1. Coordinate frames $\mathcal{A}$ and $\mathcal{B}$ are shown in Fig. 2.5. Find the representations of $\mathbf{Q}$ rotating $\mathcal{A}$ into $\mathcal{B}$ in these two frames and show that they are identical. Moreover, if $[\,\mathbf{p}\,]_{\mathcal{A}} = [\,1,\ 1,\ 1\,]^T$, find $[\,\mathbf{p}\,]_{\mathcal{B}}$.

Solution: Let $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ be unit vectors in the directions of the X-, Y-, and Z-axes, respectively; unit vectors $\boldsymbol{\iota}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\kappa}$ are defined likewise as parallel to the $\mathcal{X}$-, $\mathcal{Y}$-, and $\mathcal{Z}$-axes of Fig. 2.5. Therefore,

$$\mathbf{Q}\,\mathbf{i} \equiv \boldsymbol{\iota} = -\mathbf{k}, \quad \mathbf{Q}\,\mathbf{j} \equiv \boldsymbol{\gamma} = -\mathbf{i}, \quad \mathbf{Q}\,\mathbf{k} \equiv \boldsymbol{\kappa} = \mathbf{j}$$

Therefore, using Definition 2.2.1, the matrix representation of $\mathbf{Q}$ carrying $\mathcal{A}$ into $\mathcal{B}$, in $\mathcal{A}$, is given by

$$[\,\mathbf{Q}\,]_{\mathcal{A}} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}$$

Now, in order to find $[\,\mathbf{Q}\,]_{\mathcal{B}}$, we apply $\mathbf{Q}$ to the three unit vectors of $\mathcal{B}$, $\boldsymbol{\iota}$, $\boldsymbol{\gamma}$, and $\boldsymbol{\kappa}$. Thus, for $\boldsymbol{\iota}$, we have

$$\mathbf{Q}\boldsymbol{\iota} = \mathbf{Q}(-\mathbf{k}) = -\mathbf{Q}\mathbf{k} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = -\mathbf{j} = -\boldsymbol{\kappa}$$

Likewise,

$$\mathbf{Q}\boldsymbol{\gamma} = -\boldsymbol{\iota}, \quad \mathbf{Q}\boldsymbol{\kappa} = \boldsymbol{\gamma}$$

again, from Definition 2.2.1, we have

$$[\mathbf{Q}]_{\mathcal{B}} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix} = [\mathbf{Q}]_{\mathcal{A}}$$

thereby confirming Theorem 2.5.2. Note that the representation of this matrix in any other coordinate frame would be different. For example, if we represent this matrix in a frame whose X-axis is directed along the axis of rotation of $\mathbf{Q}$, then we end up with the X-canonical representation of $\mathbf{Q}$, namely,

$$[\mathbf{Q}]_{\mathcal{X}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix}$$

with the angle of rotation ϕ being readily computed as $\phi = 120°$, which thus yields

$$[\mathbf{Q}]_{\mathcal{X}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/2 & -\sqrt{3}/2 \\ 0 & \sqrt{3}/2 & -1/2 \end{bmatrix}$$

Apparently, the entries of $[\mathbf{Q}]_{\mathcal{X}}$ are different from those of $[\mathbf{Q}]_{\mathcal{A}}$ and $[\mathbf{Q}]_{\mathcal{B}}$ found above.

Now, from Eq. (2.94),

$$[\mathbf{p}]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

a result that can be readily verified by inspection.

2.5.2 Coordinate Transformation with Origin Shift

Now, if the coordinate origins do not coincide, let $\mathbf{b}$ be the position vector, in $\mathcal{A}$, of $\mathcal{O}$, the origin of $\mathcal{B}$, as shown in Fig. 2.6. The corresponding coordinate transformation from $\mathcal{A}$ to $\mathcal{B}$, the counterpart of Theorem 2.5.1, is given below.

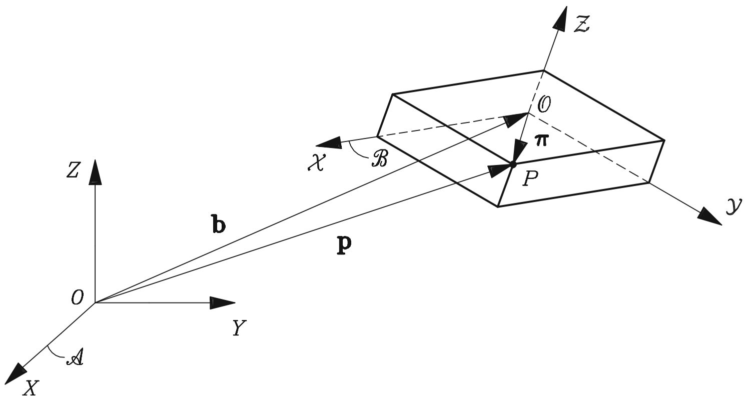

Fig. 2.6 Coordinate frames with different origins

Theorem 2.5.4. *The representations of the position vector* $\mathbf{p}$ *of a point* P *of the Euclidean three-dimensional space in two frames* $\mathcal{A}$ *and* $\mathcal{B}$ *are related by*

$$[\mathbf{p}]_{\mathcal{A}} = [\mathbf{b}]_{\mathcal{A}} + [\mathbf{Q}]_{\mathcal{A}}[\boldsymbol{\pi}]_{\mathcal{B}} \tag{2.95a}$$

$$[\boldsymbol{\pi}]_{\mathcal{B}} = [\mathbf{Q}^T]_{\mathcal{B}}([-\mathbf{b}]_{\mathcal{A}} + [\mathbf{p}]_{\mathcal{A}}) \tag{2.95b}$$

with $\mathbf{b}$ *defined as the vector directed from the origin of* $\mathcal{A}$ *to that of* $\mathcal{B}$*, and* $\boldsymbol{\pi}$ *the vector directed from the origin of* $\mathcal{B}$ *to* P*, as depicted in Fig. 2.6.*

Proof. We have, from Fig. 2.6, in any coordinate frame,

$$\mathbf{p} = \mathbf{b} + \boldsymbol{\pi} \tag{2.96}$$

If we express the above equation in the $\mathcal{A}$-frame, we obtain

$$[\mathbf{p}]_{\mathcal{A}} = [\mathbf{b}]_{\mathcal{A}} + [\boldsymbol{\pi}]_{\mathcal{A}}$$

where $\boldsymbol{\pi}$ is assumed to be readily available in $\mathcal{B}$, and so the foregoing equation must be expressed as

$$[\mathbf{p}]_{\mathcal{A}} = [\mathbf{b}]_{\mathcal{A}} + [\mathbf{Q}]_{\mathcal{A}}[\boldsymbol{\pi}]_{\mathcal{B}} \tag{2.97}$$

which thus proves Eq. (2.95a). To prove Eq. (2.95b), we simply solve Eq. (2.96) for $\boldsymbol{\pi}$ and apply Eq. (2.94) to the equation thus resulting, which readily leads to the desired relation.

Notice the geometric interpretation of the second term in the right-hand side of Eq. (2.97): this term represents, in frame $\mathcal{A}$, the position vector of a point P', whose image under $\mathbf{Q}$ is $\boldsymbol{\pi}$.

Example 2.5.2. If $[\,\mathbf{b}\,]_{\mathcal{A}} = [-1, -1, -1\,]^T$ and $\mathcal{A}$ and $\mathcal{B}$ have the relative orientations given in Example 2.5.1, find the position vector, in $\mathcal{B}$, of a point P of position vector $[\,\mathbf{p}\,]_{\mathcal{A}}$ given as in the same example.

Solution: What we obviously need is $[\,\boldsymbol{\pi}\,]_{\mathcal{B}}$, which is given in Eq. (2.95b). We thus compute first the sum inside the parentheses of that equation, i.e.,

$$[-\mathbf{b}\,]_{\mathcal{A}} + [\,\mathbf{p}\,]_{\mathcal{A}} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

We need further $[\,\mathbf{Q}^T\,]_{\mathcal{B}}$, which can be readily derived from $[\,\mathbf{Q}\,]_{\mathcal{B}}$. We do not have as yet this matrix, but we have $[\,\mathbf{Q}^T\,]_{\mathcal{A}}$, which is identical to $[\,\mathbf{Q}^T\,]_{\mathcal{B}}$ by virtue of Theorem 2.5.2. Therefore,

$$[\,\boldsymbol{\pi}\,]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & -1 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \\ 2 \end{bmatrix}$$

a result that the reader is invited to verify by inspection.

2.5.3 Homogeneous Coordinates

The general coordinate transformation, involving a shift of the origin, is not linear in the sense of the definition given in Subsection 2.2, as can be readily realized by virtue of the *nonhomogeneous* term involved, i.e., the first term of the right-hand side of Eq. (2.95a), which is independent of $\mathbf{p}$. Such a transformation, nevertheless, can be represented in homogeneous form if *homogeneous coordinates* are introduced. These are defined below: Let $[\,\mathbf{p}\,]_{\mathcal{M}}$ be the coordinate array of a *finite* point P in reference frame $\mathcal{M}$. What we mean by a finite point is one whose coordinates are all finite. We are thus assuming that the point P at hand is not *at infinity*, points at infinity being introduced presently. The homogeneous coordinates of P are those in the four-dimensional array $\{\mathbf{p}\}_{\mathcal{M}}$, defined as

$$\{\mathbf{p}\}_{\mathcal{M}} \equiv \begin{bmatrix} [\,\mathbf{p}\,]_{\mathcal{M}} \\ 1 \end{bmatrix} \tag{2.98}$$

The *affine transformation* of Eq. (2.95a) can now be rewritten in homogeneous-coordinate form as

$$\{\mathbf{p}\}_{\mathcal{A}} = \{\mathbf{T}\}_{\mathcal{A}}\{\boldsymbol{\pi}\}_{\mathcal{B}} \tag{2.99}$$

where $\{\mathbf{T}\}_{\mathcal{A}}$ is defined as a 4×4 array, namely,

$$\{\mathbf{T}\}_{\mathcal{A}} \equiv \begin{bmatrix} [\,\mathbf{Q}\,]_{\mathcal{A}} & [\,\mathbf{b}\,]_{\mathcal{A}} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{2.100}$$

The inverse transformation of that defined in Eq. (2.100) is derived from Eqs. (2.95a and b), i.e.,

$$\{\mathbf{T}^{-1}\}_{\mathcal{B}} = \begin{bmatrix} [\mathbf{Q}^T]_{\mathcal{B}} & [-\mathbf{b}]_{\mathcal{B}} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{2.101}$$

Now, $\{\mathbf{T}^{-1}\}_{\mathcal{A}}$ can be readily derived from the above expression, upon application of Theorems 2.5.1 and 2.5.2, which leads to

$$\{\mathbf{T}^{-1}\}_{\mathcal{A}} = \begin{bmatrix} [\mathbf{Q}^T]_{\mathcal{A}} & -[\mathbf{Q}]_{\mathcal{A}}[\mathbf{b}]_{\mathcal{B}} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{2.102}$$

Furthermore, homogeneous transformations can be concatenated. Indeed, let $\mathcal{F}_k$, for $k = i-1,\ i,\ i+1$, denote three coordinate frames, with origins at O_k. Moreover, let $\mathbf{Q}_{i-1}$ be the rotation carrying $\mathcal{F}_{i-1}$ into an orientation coinciding with that of $\mathcal{F}_i$. If a similar definition for $\mathbf{Q}_i$ is adopted, then $\mathbf{Q}_i$ denotes the rotation carrying $\mathcal{F}_i$ into an orientation coinciding with that of $\mathcal{F}_{i+1}$. First, the case in which all three origins coincide is considered. Clearly,

$$[\mathbf{p}]_i = [\mathbf{Q}_{i-1}^T]_{i-1}[\mathbf{p}]_{i-1} \tag{2.103a}$$

$$[\mathbf{p}]_{i+1} = [\mathbf{Q}_i^T]_i[\mathbf{p}]_i = [\mathbf{Q}_i^T]_i[\mathbf{Q}_{i-1}^T]_{i-1}[\mathbf{p}]_{i-1} \tag{2.103b}$$

the inverse relations of those appearing in Eqs. (2.103a and b) being

$$[\mathbf{p}]_{i-1} = [\mathbf{Q}_{i-1}]_{i-1}[\mathbf{p}]_i \tag{2.104a}$$

$$[\mathbf{p}]_{i-1} = [\mathbf{Q}_{i-1}]_{i-1}[\mathbf{Q}_i]_i[\mathbf{p}]_{i+1} \tag{2.104b}$$

If now the origins do not coincide, let $\mathbf{a}_{i-1}$ and $\mathbf{a}_i$ denote the vectors $\overrightarrow{O_{i-1}O_i}$ and $\overrightarrow{O_i O_{i+1}}$, respectively. The transformations $\{\mathbf{T}_{i-1}\}_{i-1}$ and $\{\mathbf{T}_i\}_i$ thus arising are obviously

$$\{\mathbf{T}_{i-1}\}_{i-1} = \begin{bmatrix} [\mathbf{Q}_{i-1}]_{i-1} & [\mathbf{a}_{i-1}]_{i-1} \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad \{\mathbf{T}_i\}_i = \begin{bmatrix} [\mathbf{Q}_i]_i & [\mathbf{a}_i]_i \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{2.105}$$

whereas their inverse transformations are

$$\{\mathbf{T}_{i-1}^{-1}\}_i = \begin{bmatrix} [\mathbf{Q}_{i-1}^T]_i & [\mathbf{Q}_{i-1}^T]_i[-\mathbf{a}_{i-1}]_{i-1} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{2.106a}$$

$$\{\mathbf{T}_i^{-1}\}_{i+1} = \begin{bmatrix} [\mathbf{Q}_i^T]_{i+1} & [\mathbf{Q}_i^T]_{i+1}[-\mathbf{a}_i]_i \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{2.106b}$$

The coordinate transformations involved are derived upon simply replacing $\mathbf{Q}_{i-1}$ and $\mathbf{Q}_i$ with $\mathbf{T}_{i-1}$ and $\mathbf{T}_i$, respectively, in Eqs. (2.104a and b), i.e.,

$$\{\mathbf{p}\}_{i-1} = \{\mathbf{T}_{i-1}\}_{i-1}\{\mathbf{p}\}_i \tag{2.107a}$$

$$\{\mathbf{p}\}_{i-1} = \{\mathbf{T}_{i-1}\}_{i-1}\{\mathbf{T}_i\}_i\{\mathbf{p}\}_{i+1} \tag{2.107b}$$

the corresponding inverse transformations being

$$\{\mathbf{p}\}_i = \{\mathbf{T}_{i-1}^{-1}\}_{i-1}\{\mathbf{p}\}_{i-1} \tag{2.108a}$$

$$\{\mathbf{p}\}_{i+1} = \{\mathbf{T}_i^{-1}\}_i\{\mathbf{p}\}_i = \{\mathbf{T}_i^{-1}\}_i\{\mathbf{T}_{i-1}^{-1}\}_{i-1}\{\mathbf{p}\}_{i-1} \tag{2.108b}$$

which are the counterpart transformations of Eqs. (2.103a and b) for the case of no rigin shift.

Now, if P lies at infinity, we can express its homogeneous coordinates in a simpler form. To this end, we rewrite expression (2.98) in the form

$$\{\mathbf{p}\}_{\mathcal{M}} \equiv \|\mathbf{p}\| \begin{bmatrix} [\mathbf{e}]_{\mathcal{M}} \\ 1/\|\mathbf{p}\| \end{bmatrix}$$

and hence,

$$\lim_{\|\mathbf{p}\|\to\infty} \{\mathbf{p}\}_{\mathcal{M}} = \left(\lim_{\|\mathbf{p}\|\to\infty} \|\mathbf{p}\|\right) \left(\lim_{\|\mathbf{p}\|\to\infty} \begin{bmatrix} [\mathbf{e}]_{\mathcal{M}} \\ 1/\|\mathbf{p}\| \end{bmatrix}\right)$$

or

$$\lim_{\|\mathbf{p}\|\to\infty} \{\mathbf{p}\}_{\mathcal{M}} = \left(\lim_{\|\mathbf{p}\|\to\infty} \|\mathbf{p}\|\right) \begin{bmatrix} [\mathbf{e}]_{\mathcal{M}} \\ 0 \end{bmatrix}$$

We now define the *homogeneous coordinates of a point P lying at infinity* as the four-dimensional array appearing in the foregoing expression, i.e.,

$$\{\mathbf{p}_\infty\}_{\mathcal{M}} \equiv \begin{bmatrix} [\mathbf{e}]_{\mathcal{M}} \\ 0 \end{bmatrix} \tag{2.109}$$

which means that a point at infinity, in homogeneous coordinates, has only a direction, given by the unit vector **e**, but an undefined location. When working with objects within the atmosphere of the Earth, for example, stars can be regarded as lying at infinity, and hence, their location is completely specified simply by their longitude and latitude, which suffice to define the direction cosines of a unit vector in spherical coordinates.

On the other hand, a rotation matrix can be regarded as composed of three columns, each representing a unit vector, e.g.,

$$\mathbf{Q} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{bmatrix}$$

where the triad $\{\mathbf{e}_k\}_1^3$ is orthonormal. We can thus represent $\{\mathbf{T}\}_{\mathcal{A}}$ of Eq. (2.100) in the form

$$\{\mathbf{T}\}_{\mathcal{A}} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \mathbf{b} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.110}$$

thereby concluding that the columns of the 4×4 matrix $\mathbf{T}$ represent the homogeneous coordinates of a set of corresponding points, the first three of which lie at infinity.

Example 2.5.3. An ellipsoid is centered at a point $O_{\mathcal{B}}$ of position vector $\mathbf{b}$, its three axes $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ defining a coordinate frame $\mathcal{B}$. Moreover, its semiaxes have lengths $a = 1$, $b = 2$, and $c = 3$, the coordinates of $O_{\mathcal{B}}$ in a coordinate frame $\mathcal{A}$ being $[\mathbf{b}]_{\mathcal{A}} = [1, 2, 3]^T$. Additionally, the direction cosines of $\mathcal{X}$ are (0.933, 0.067, −0.354), whereas $\mathcal{Y}$ is perpendicular to $\mathbf{b}$ and to the unit vector $\mathbf{u}$ that is parallel to the $\mathcal{X}$-axis. Find the equation of the ellipsoid in $\mathcal{A}$. (This example has relevance in collision-avoidance algorithms, some of which approximate manipulator links as ellipsoids, thereby easing tremendously the computational requirements.)

Solution: Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be unit vectors parallel to the $\mathcal{X}$-, $\mathcal{Y}$-, and $\mathcal{Z}$-axes, respectively. Then,

$$[\mathbf{u}]_{\mathcal{A}} = \begin{bmatrix} 0.933 \\ 0.067 \\ -0.354 \end{bmatrix}, \quad \mathbf{v} = \frac{\mathbf{u} \times \mathbf{b}}{\|\mathbf{u} \times \mathbf{b}\|}, \quad \mathbf{w} = \mathbf{u} \times \mathbf{v}$$

and hence,

$$[\mathbf{v}]_{\mathcal{A}} = \begin{bmatrix} 0.243 \\ -0.843 \\ 0.481 \end{bmatrix}, \quad [\mathbf{w}]_{\mathcal{A}} = \begin{bmatrix} -0.266 \\ -0.535 \\ -0.803 \end{bmatrix}$$

from which the rotation matrix $\mathbf{Q}$, rotating the axes of $\mathcal{A}$ into orientations coinciding with those of $\mathcal{B}$, can be readily represented in $\mathcal{A}$, or in $\mathcal{B}$ for that matter, as

$$[\mathbf{Q}]_{\mathcal{A}} = [\mathbf{u}, \mathbf{v}, \mathbf{w}]_{\mathcal{A}} = \begin{bmatrix} 0.933 & 0.243 & -0.266 \\ 0.067 & -0.843 & -0.535 \\ -0.354 & 0.481 & -0.803 \end{bmatrix}$$

On the other hand, if the coordinates of a point P in $\mathcal{A}$ and $\mathcal{B}$ are $[\mathbf{p}]_{\mathcal{A}} = [p_1, p_2, p_3]^T$ and $[\boldsymbol{\pi}]_{\mathcal{B}} = [\pi_1, \pi_2, \pi_3]^T$, respectively, then the equation of the ellipsoid in $\mathcal{B}$ is, apparently,

$$\mathcal{B}: \quad \frac{\pi_1^2}{1^2} + \frac{\pi_2^2}{2^2} + \frac{\pi_3^2}{3^2} = 1$$

Now, what is needed in order to derive the equation of the ellipsoid in $\mathcal{A}$ is simply a relation between the coordinates of P in $\mathcal{B}$ and those in $\mathcal{A}$. These coordinates are related by Eq. (2.95b), which requires $[\mathbf{Q}^T]_{\mathcal{B}}$, while we have $[\mathbf{Q}]_{\mathcal{A}}$. Nevertheless, by virtue of Theorem 2.5.2

$$[\mathbf{Q}^T]_{\mathcal{B}} = [\mathbf{Q}^T]_{\mathcal{A}} = \begin{bmatrix} 0.933 & 0.067 & -0.354 \\ 0.243 & -0.843 & 0.481 \\ -0.266 & -0.535 & -0.803 \end{bmatrix}$$

Hence,

$$[\boldsymbol{\pi}]_{\mathcal{B}} = \begin{bmatrix} 0.933 & 0.067 & -0.354 \\ 0.243 & -0.843 & 0.481 \\ -0.266 & -0.535 & -0.803 \end{bmatrix} \left(\begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix} + \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} \right)$$

Therefore,

$$\begin{aligned} \pi_1 &= 0.933p_1 + 0.067p_2 - 0.354p_3 - 0.005 \\ \pi_2 &= 0.243p_1 - 0.843p_2 + 0.481p_3 \\ \pi_3 &= -0.266p_1 - 0.535p_2 - 0.803p_3 + 3.745 \end{aligned}$$

Substitution of the foregoing relations into the ellipsoid equation in $\mathcal{B}$ leads to

$$\begin{aligned} \mathcal{A}: \quad & 32.1521{p_1}^2 + 7.70235{p_2}^2 + 9.17286{p_3}^2 - 8.30524p_1 - 16.0527p_2 \\ & -23.9304p_3 + 9.32655p_1p_2 + 9.02784p_2p_3 - 19.9676p_1p_3 + 20.101 = 0 \end{aligned}$$

which is the equation sought, as obtained using computer algebra.

2.6 Similarity Transformations

Transformations of the position vector of points under a change of coordinate frame involving both a translation of the origin and a rotation of the coordinate axes was the main subject of Sect. 2.5. In this section, we study the transformations of components of vectors other than the position vector, while extending the concept to the transformation of matrix entries. How these transformations take place is the subject of this section.

What is involved in the present discussion is a *change of basis* of the associated vector spaces, and hence, this is not limited to three-dimensional vector spaces. That is, n-dimensional vector spaces will be studied in this section. Moreover, only isomorphisms, i.e., transformations $\mathbf{L}$ of the n-dimensional vector space $\mathcal{V}$ onto itself will be considered. Let $\mathcal{A} = \{\mathbf{a}_i\}_1^n$ and $\mathcal{B} = \{\mathbf{b}_i\}_1^n$ be two *different* bases of the same space $\mathcal{V}$. Hence, any vector $\mathbf{v}$ of $\mathcal{V}$ can be expressed in either of two ways, namely,

$$\mathbf{v} = \alpha_1\mathbf{a}_1 + \alpha_2\mathbf{a}_2 + \cdots + \alpha_n\mathbf{a}_n \tag{2.111}$$

$$\mathbf{v} = \beta_1\mathbf{b}_1 + \beta_2\mathbf{b}_2 + \cdots + \beta_n\mathbf{b}_n \tag{2.112}$$

from which two representations of **v** are readily derived, namely,

$$[\mathbf{v}]_{\mathcal{A}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad [\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \tag{2.113}$$

In order to find the representation of the *linear transformation* that relates the two bases, let **A** be the linear transformation that carries the set $\mathcal{A}$ into the set $\mathcal{B}$, i.e., **A**: $\mathcal{A} \to \mathcal{B}$.

Furthermore, let the two foregoing bases be related by

$$\mathbf{b}_j = a_{1j}\mathbf{a}_1 + a_{2j}\mathbf{a}_2 + \cdots + a_{nj}\mathbf{a}_n, \quad j = 1, \ldots, n \tag{2.114}$$

Hence, the representation sought, $[\mathbf{A}]_{\mathcal{A}}$—in the basis $\mathcal{A}$— is, according to Definition 2.2.1,

$$[\mathbf{A}]_{\mathcal{A}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{2.115}$$

Now, in order to find the relationship between the two representations of Eq. (2.113), Eq. (2.114) is substituted into Eq. (2.112), which yields

$$\begin{aligned} \mathbf{v} = {} & \beta_1(a_{11}\mathbf{a}_1 + a_{21}\mathbf{a}_2 + \cdots + a_{n1}\mathbf{a}_n) \\ & + \beta_2(a_{12}\mathbf{a}_1 + a_{22}\mathbf{a}_2 + \cdots + a_{n2}\mathbf{a}_n) \\ & \vdots \\ & + \beta_n(a_{1n}\mathbf{a}_1 + a_{2n}\mathbf{a}_2 + \cdots + a_{nn}\mathbf{a}_n) \end{aligned} \tag{2.116}$$

This can be rearranged to yield

$$\begin{aligned} \mathbf{v} = {} & (a_{11}\beta_1 + a_{12}\beta_2 + \cdots + a_{1n}\beta_n)\mathbf{a}_1 \\ & + (a_{21}\beta_1 + a_{22}\beta_2 + \cdots + a_{2n}\beta_n)\mathbf{a}_2 \\ & \vdots \\ & + (a_{n1}\beta_1 + a_{n2}\beta_2 + \cdots + a_{nn}\beta_n)\mathbf{a}_n \end{aligned} \tag{2.117}$$

Comparing Eq. (2.117) with Eq. (2.111), one readily derives

$$[\mathbf{v}]_{\mathcal{A}} = [\mathbf{A}]_{\mathcal{A}}[\mathbf{v}]_{\mathcal{B}} \tag{2.118}$$

with $[\mathbf{A}]_{\mathcal{A}}$ introduced in Eq. (2.115). Equation (2.118) is the relation, sought, its inverse being

$$[\mathbf{v}]_{\mathcal{B}} = [\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{v}]_{\mathcal{A}} \tag{2.119}$$

Next, let $\mathbf{L}$ have the representation in $\mathcal{A}$ given below:

$$[\mathbf{L}]_{\mathcal{A}} = \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1n} \\ l_{21} & l_{22} & \cdots & l_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \tag{2.120}$$

Now we aim to find the relationship between $[\mathbf{L}]_{\mathcal{A}}$ and $[\mathbf{L}]_{\mathcal{B}}$. To this end, let $\mathbf{w}$ be the image of $\mathbf{v}$ under $\mathbf{L}$, i.e.,

$$\mathbf{Lv} = \mathbf{w} \tag{2.121}$$

which can be expressed in terms of either $\mathcal{A}$ or $\mathcal{B}$ as

$$[\mathbf{L}]_{\mathcal{A}}[\mathbf{v}]_{\mathcal{A}} = [\mathbf{w}]_{\mathcal{A}} \tag{2.122}$$

$$[\mathbf{L}]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [\mathbf{w}]_{\mathcal{B}} \tag{2.123}$$

Further, since $\mathbf{L}$ is an isomorphism by hypothesis, $\mathbf{w}$ of Eq. (2.121) lies in the same space $\mathcal{V}$ as $\mathbf{v}$. Hence, similar to Eq. (2.118),

$$[\mathbf{w}]_{\mathcal{A}} = [\mathbf{A}]_{\mathcal{A}}[\mathbf{w}]_{\mathcal{B}} \tag{2.124}$$

Now, substitution of Eqs. (2.118) and (2.124) into Eq. (2.122) yields

$$[\mathbf{A}]_{\mathcal{A}}[\mathbf{w}]_{\mathcal{B}} = [\mathbf{L}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}}[\mathbf{v}]_{\mathcal{B}} \tag{2.125}$$

which can be readily rearranged in the form

$$[\mathbf{w}]_{\mathcal{B}} = [\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}}[\mathbf{v}]_{\mathcal{B}} \tag{2.126}$$

Comparing Eq. (2.123) with Eq. (2.126) readily leads to

$$[\mathbf{L}]_{\mathcal{B}} = [\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}} \tag{2.127}$$

which upon rearrangement, becomes

$$[\mathbf{L}]_{\mathcal{A}} = [\mathbf{A}]_{\mathcal{A}}[\mathbf{L}]_{\mathcal{B}}[\mathbf{A}^{-1}]_{\mathcal{A}} \tag{2.128}$$

Now, paraphrasing Theorems 2.5.2 and 2.5.4, we can state

Theorem 2.6.1. *The representations of* $\mathbf{A}$ *carrying* $\mathcal{A}$ *into* $\mathcal{B}$ *in these two frames are identical, i.e.,*

$$[\mathbf{A}]_{\mathcal{A}} = [\mathbf{A}]_{\mathcal{B}} \tag{2.129}$$

Proof. Substitute $\mathbf{L}$ for $\mathbf{A}$ in Eq. (2.127) to obtain the above relation, q.e.d.

Relations (2.118), (2.119), (2.127), and (2.128) constitute what are called *similarity transformations*. These are important because they preserve *invariant* quantities such as the eigenvalues and eigenvectors of matrices, the magnitudes of vectors, the angles between vectors, and so on. Indeed, one has:

Theorem 2.6.2. *The characteristic polynomial of a given* $n \times n$ *matrix remains unchanged under a similarity transformation. Moreover, the eigenvalues of two matrix representations of the same* $n \times n$ *linear transformation are identical, and if* $[\mathbf{e}]_{\mathcal{B}}$ *is an eigenvector of* $[\mathbf{L}]_{\mathcal{B}}$*, then under the similarity transformation (2.128), the corresponding eigenvector of* $[\mathbf{L}]_{\mathcal{A}}$ *is* $[\mathbf{e}]_{\mathcal{A}} = [\mathbf{A}]_{\mathcal{A}}[\mathbf{e}]_{\mathcal{B}}$.

Proof. From Eq. (2.11), the characteristic polynomial of $[\mathbf{L}]_{\mathcal{B}}$ is

$$P(\lambda) = \det(\lambda[\mathbf{1}]_{\mathcal{B}} - [\mathbf{L}]_{\mathcal{B}}) \tag{2.130}$$

which can be rewritten as

$$\begin{aligned} P(\lambda) &\equiv \det(\lambda[\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{1}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}} - [\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}}) \\ &= \det([\mathbf{A}^{-1}]_{\mathcal{A}}(\lambda[\mathbf{1}]_{\mathcal{A}} - [\mathbf{L}]_{\mathcal{A}})[\mathbf{A}]_{\mathcal{A}}) \\ &= \det([\mathbf{A}^{-1}]_{\mathcal{A}})\det(\lambda[\mathbf{1}]_{\mathcal{A}} - [\mathbf{L}]_{\mathcal{A}})\det([\mathbf{A}]_{\mathcal{A}}) \end{aligned}$$

But

$$\det([\mathbf{A}^{-1}]_{\mathcal{A}})\det([\mathbf{A}]_{\mathcal{A}}) = 1$$

and hence, the characteristic polynomial of $[\mathbf{L}]_{\mathcal{A}}$ is identical to that of $[\mathbf{L}]_{\mathcal{B}}$. Since both representations have the same characteristic polynomial, they have the same eigenvalues. Now, if $[\mathbf{e}]_{\mathcal{B}}$ is an eigenvector of $[\mathbf{L}]_{\mathcal{B}}$ associated with the eigenvalue λ, then

$$[\mathbf{L}]_{\mathcal{B}}[\mathbf{e}]_{\mathcal{B}} = \lambda[\mathbf{e}]_{\mathcal{B}}$$

Next, Eq. (2.127) is substituted into the foregoing equation, which thus leads to

$$[\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}}[\mathbf{e}]_{\mathcal{B}} = \lambda[\mathbf{e}]_{\mathcal{B}}$$

Upon rearrangement, this equation becomes

$$[\mathbf{L}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}}[\mathbf{e}]_{\mathcal{B}} = \lambda[\mathbf{A}]_{\mathcal{A}}[\mathbf{e}]_{\mathcal{B}} \tag{2.131}$$

whence it is apparent that $[\mathbf{A}]_{\mathcal{A}}[\mathbf{e}]_{\mathcal{B}}$ is an eigenvector of $[\mathbf{L}]_{\mathcal{A}}$ associated with the eigenvalue λ, q.e.d.

Theorem 2.6.3. *If $[\mathbf{L}]_{\mathcal{A}}$ and $[\mathbf{L}]_{\mathcal{B}}$ are related by the similarity transformation (2.127), then*

$$[\mathbf{L}^k]_{\mathcal{B}} = [\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}^k]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}} \tag{2.132}$$

for any integer k.

Proof. This is done by induction. For $k = 2$, one has

$$\begin{aligned}[\mathbf{L}^2]_{\mathcal{B}} &\equiv [\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}}[\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}} \\ &= [\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}^2]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}}\end{aligned}$$

Now, assume that the proposed relation holds for $k = n$. Then,

$$\begin{aligned}[\mathbf{L}^{n+1}]_{\mathcal{B}} &\equiv [\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}^n]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}}[\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}} \\ &= [\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}^{n+1}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}}\end{aligned}$$

i.e., the relation holds for $k = n + 1$ as well, thereby completing the proof.

Theorem 2.6.4. *The trace of a $n \times n$ matrix does not change under a similarity transformation.*

Proof. A preliminary relation will be needed: Let $[\mathbf{A}]$, $[\mathbf{B}]$ and $[\mathbf{C}]$ be three different $n \times n$ matrix arrays, in a given reference frame, that need not be indicated with any subscript. Moreover, let a_{ij}, b_{ij}, and c_{ij} be the components of the said arrays, with indices ranging from 1 to n. Hence, using standard index notation,

$$\mathrm{tr}([\mathbf{A}][\mathbf{B}][\mathbf{C}]) \equiv a_{ij}b_{jk}c_{ki} = b_{jk}c_{ki}a_{ij} \equiv \mathrm{tr}([\mathbf{B}][\mathbf{C}][\mathbf{A}]) \tag{2.133}$$

Taking the trace of both sides of Eq. (2.127) and applying the foregoing result produces

$$\mathrm{tr}([\mathbf{L}]_{\mathcal{B}}) = \mathrm{tr}([\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}]_{\mathcal{A}}[\mathbf{A}]_{\mathcal{A}}) = \mathrm{tr}([\mathbf{A}]_{\mathcal{A}}[\mathbf{A}^{-1}]_{\mathcal{A}}[\mathbf{L}]_{\mathcal{A}}) = \mathrm{tr}([\mathbf{L}]_{\mathcal{A}}) \tag{2.134}$$

thereby proving that the trace remains unchanged under a similarity transformation.

Example 2.6.1. We consider the equilateral triangle sketched in Fig. 2.7, of side length equal to 2, with vertices P_1, P_2, and P_3, and coordinate frames $\mathcal{A}$ and $\mathcal{B}$ of

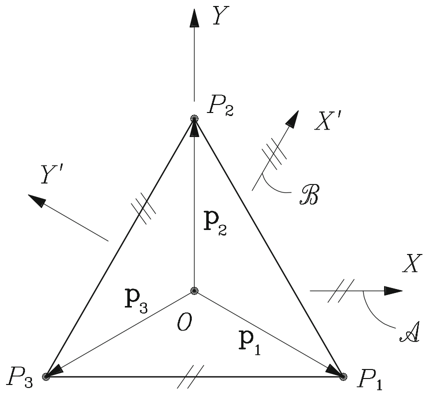

Fig. 2.7 Two coordinate frames used to represent the position vectors of the corners of an equilateral triangle

axes X, Y and X', Y', respectively, both with origin at the centroid of the triangle. Let $\mathbf{P}$ be a 2×2 matrix defined by

$$\mathbf{P} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 \end{bmatrix}$$

with $\mathbf{p}_i$ denoting the position vector of P_i in a given coordinate frame. Show that matrix $\mathbf{P}$ does not obey a similarity transformation upon a change of frame, and compute its trace in frames $\mathcal{A}$ and $\mathcal{B}$ to make it apparent that this matrix does not comply with the conditions of Theorem 2.6.4.

Solution: From the figure it is apparent that

$$[\mathbf{P}]_{\mathcal{A}} = \begin{bmatrix} 1 & 0 \\ -\sqrt{3}/3 & 2\sqrt{3}/3 \end{bmatrix}, \quad [\mathbf{P}]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 \\ -2\sqrt{3}/3 & \sqrt{3}/3 \end{bmatrix}$$

Apparently,

$$\operatorname{tr}([\mathbf{P}]_{\mathcal{A}}) = 1 + \frac{2\sqrt{3}}{3} \neq \operatorname{tr}([\mathbf{P}]_{\mathcal{B}}) = \frac{\sqrt{3}}{3}$$

The reason why the trace of this matrix did not remain unchanged under a coordinate transformation is that the matrix does not obey a similarity transformation under a change of coordinates. Indeed, vectors $\mathbf{p}_i$ change as

$$[\mathbf{p}_i]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{A}}[\mathbf{p}_i]_{\mathcal{B}}$$

under a change of coordinates from $\mathcal{B}$ to $\mathcal{A}$, with $\mathbf{Q}$ denoting the rotation carrying $\mathcal{A}$ into $\mathcal{B}$. Hence,

$$[\mathbf{P}]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{A}}[\mathbf{P}]_{\mathcal{B}}$$

which is different from the similarity transformation of Eq. (2.128). However, if we now define

$$\mathbf{R} \equiv \mathbf{P}\mathbf{P}^T$$

then

$$[\mathbf{R}]_{\mathcal{A}} = \begin{bmatrix} 1 & -\sqrt{3}/3 \\ -\sqrt{3}/3 & 5/3 \end{bmatrix}, \quad [\mathbf{R}]_{\mathcal{B}} = \begin{bmatrix} 1 & \sqrt{3}/3 \\ \sqrt{3}/3 & 5/3 \end{bmatrix}$$

and hence,

$$\mathrm{tr}([\mathbf{R}]_{\mathcal{A}}) = \mathrm{tr}([\mathbf{R}]_{\mathcal{B}}) = \frac{8}{3}$$

thereby showing that the trace of **R** does not change under a change of frame. In order to verify whether matrix **R** complies with the conditions of Theorem 2.6.4, we notice that, under a change of frame, matrix **R** changes as

$$[\mathbf{R}]_{\mathcal{A}} = [\mathbf{P}\mathbf{P}^T]_{\mathcal{A}} = [[\mathbf{Q}]_{\mathcal{A}}[\mathbf{P}]_{\mathcal{B}}([\mathbf{Q}]_{\mathcal{A}}[\mathbf{P}]_{\mathcal{B}})^{\mathbf{T}} = [\mathbf{Q}]_{\mathcal{A}}[\mathbf{P}\mathbf{P}^{\mathbf{T}}]_{\mathcal{B}}[\mathbf{Q}^{\mathbf{T}}]_{\mathcal{A}}]$$

which is indeed a similarity transformation.

2.7 Invariance Concepts

From Example 2.6.1 it is apparent that certain properties, like the trace of certain square matrices, do not change under a coordinate transformation. For this reason, a matrix like **R** of that example is said to be *frame-invariant*, or simply *invariant*, whereas matrix **P** of the same example is not. In this section, we formally define the concept of *invariance* and highlight its applications and its role in robotics. Let a scalar, a vector, and a matrix function of the position vector **p** be denoted by $f(\mathbf{p})$, $\mathbf{f}(\mathbf{p})$ and $\mathbf{F}(\mathbf{p})$, respectively. The representations of $\mathbf{f}(\mathbf{p})$ in two different coordinate frames, labelled $\mathcal{A}$ and $\mathcal{B}$, will be indicated as $[\mathbf{f}(\mathbf{p})]_{\mathcal{A}}$ and $[\mathbf{f}(\mathbf{p})]_{\mathcal{B}}$, respectively, with a similar notation for the representations of $\mathbf{F}(\mathbf{p})$. Moreover, let the two frames differ both in the location of their origins and in their orientations. Additionally, let the *proper orthogonal* matrix $[\mathbf{Q}]_{\mathcal{A}}$ denote the rotation of coordinate frame $\mathcal{A}$ into $\mathcal{B}$. Then, the scalar function $f(\mathbf{p})$ is said to be frame invariant, or invariant for brevity, if

$$f([\mathbf{p}]_{\mathcal{B}}) = f([\mathbf{p}]_{\mathcal{A}}) \tag{2.135}$$

Moreover, the vector quantity **f** is said to be invariant if

$$[\mathbf{f}]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{A}}[\mathbf{f}]_{\mathcal{B}} \tag{2.136}$$

and finally, the matrix quantity $\mathbf{F}$ is said to be invariant if

$$[\mathbf{F}]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{A}}[\mathbf{F}]_{\mathcal{B}}[\mathbf{Q}^T]_{\mathcal{A}} \tag{2.137}$$

Thus, the difference in origin location becomes irrelevant in this context, and hence, will no longer be considered. From the foregoing discussion, it is clear that the same vector quantity has different components in different coordinate frames; moreover, the same matrix quantity has different entries in different coordinate frames. However, certain scalar quantities associated with vectors, e.g., the inner product, and matrices, e.g., the matrix *moments*, to be defined presently, remain unchanged under a change of frame. Additionally, such vector operations as the cross product of two vectors are invariant. In fact, the scalar product of two vectors $\mathbf{a}$ and $\mathbf{b}$ remains unchanged under a change of frame, i.e.,

$$[\,\mathbf{a}\,]_{\mathcal{A}}^T\,[\,\mathbf{b}\,]_{\mathcal{A}} = [\,\mathbf{a}\,]_{\mathcal{B}}^T\,[\,\mathbf{b}\,]_{\mathcal{B}} \tag{2.138}$$

Additionally,

$$[\,\mathbf{a}\times\mathbf{b}\,]_{\mathcal{A}} = [\,\mathbf{Q}\,]_{\mathcal{A}}\,[\,\mathbf{a}\times\mathbf{b}\,]_{\mathcal{B}} \tag{2.139}$$

The kth *moment* of a $n \times n$ matrix $\mathbf{T}$, denoted by $\mathcal{I}_k$, is defined as (Leigh 1968)

$$\mathcal{I}_k \equiv \mathrm{tr}(\mathbf{T}^k), \quad k = 0, 1, \ldots \tag{2.140}$$

where $\mathcal{I}_0 = \mathrm{tr}(\mathbf{1}) = n$. Now we have

Theorem 2.7.1. *The moments of a $n \times n$ matrix are invariant under a similarity transformation.*

Proof. This is straightforward. Indeed, from Theorem 2.6.3, we have

$$[\,\mathbf{T}^k\,]_{\mathcal{B}} = [\,\mathbf{A}^{-1}\,]_{\mathcal{A}}[\,\mathbf{T}^k\,]_{\mathcal{A}}[\,\mathbf{A}\,]_{\mathcal{A}} \tag{2.141}$$

Now, let $[\,\mathcal{I}_k\,]_{\mathcal{A}}$ and $[\,\mathcal{I}_k\,]_{\mathcal{B}}$ denote the kth moment of $[\,\mathbf{T}\,]_{\mathcal{A}}$ and $[\,\mathbf{T}\,]_{\mathcal{B}}$, respectively. Thus,

$$\begin{aligned}[\,\mathcal{I}_k\,]_{\mathcal{B}} &= \mathrm{tr}([\,\mathbf{A}^{-1}\,]_{\mathcal{A}}[\,\mathbf{T}^k\,]_{\mathcal{A}}[\,\mathbf{A}\,]_{\mathcal{A}}) \equiv \mathrm{tr}([\,\mathbf{A}\,]_{\mathcal{A}}[\,\mathbf{A}^{-1}\,]_{\mathcal{A}}[\,\mathbf{T}^k\,]_{\mathcal{A}}) \\ &= \mathrm{tr}([\,\mathbf{T}^k\,]_{\mathcal{A}}) \equiv [\,\mathcal{I}_k\,]_{\mathcal{A}}\end{aligned}$$

thereby completing the proof.

Furthermore,

Theorem 2.7.2. *A $n \times n$ matrix has only n linearly independent moments.*

Proof. Let the characteristic polynomial of $\mathbf{T}$ be

$$P(\lambda) = a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n = 0 \tag{2.142}$$

Upon application of the Cayley–Hamilton Theorem, Eq. (2.142) leads to

$$a_0\mathbf{1} + a_1\mathbf{T} + \cdots + a_{n-1}\mathbf{T}^{n-1} + \mathbf{T}^n = \mathbf{0} \tag{2.143}$$

where $\mathbf{1}$ denotes the $n \times n$ identity matrix.

Now, if we take the trace of both sides of Eq. (2.143), and Definition (2.140) is recalled, one has

$$a_0\mathcal{I}_0 + a_1\mathcal{I}_1 + \cdots + a_{n-1}\mathcal{I}_{n-1} + \mathcal{I}_n = 0 \tag{2.144}$$

from which it is apparent that $\mathcal{I}_n$ can be expressed as a linear combination of the first n moments of $\mathbf{T}$, $\{\mathcal{I}_k\}_0^{n-1}$. By simple induction, one can likewise prove that the mth moment is dependent upon the first n moments if $m \geq n$, thereby completing the proof. Also notice that $\mathcal{I}_0 = n$, and hence, *all $n \times n$ matrices share the same zeroth moment $\mathcal{I}_0$*.

The vector invariants of a $n \times n$ matrix are its eigenvectors, which have a direct geometric significance in the case of symmetric matrices. The eigenvalues of these matrices are all real, its eigenvectors being also real and mutually orthogonal. Skew-symmetric matrices, in general, need not have either real eigenvalues or real eigenvectors. However, if we limit ourselves to 3×3 skew-symmetric matrices, exactly one of their eigenvalues, and its associated eigenvector, are both real. The eigenvalue of interest is 0, and the associated vector is the axial vector of the matrix under study.

It is now apparent that two $n \times n$ matrices related by a similarity transformation have the same set of moments. Now, by virtue of Theorem 2.7.2, one may be tempted to think that if two $n \times n$ matrices share their first n moments $\{\mathcal{I}_k\}_0^{n-1}$, then the two matrices are related by a similarity transformation. To prove that this is not the case, let two $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ have characteristic polynomials with coefficients $\{a_k\}^{n-1}$ and $\{b_k\}_0^{n-1}$, respectively, the two sets being not necessarily identical. Moreover, let the $n+1$ moments of $\mathbf{A}$ and $\mathbf{B}$ be denoted by $\{\mathcal{I}_k\}_0^n$ and $\{\mathcal{I}'_k\}_0^n$, with

$$\mathcal{I}_0 = \mathcal{I}'_0 = n, \quad \mathcal{I}_k = \mathcal{I}'_k, \quad \text{for} \quad k = 1, \ldots, n-1$$

Hence, from Eq. (2.144),

$$\begin{aligned}
\mathcal{I}_n &= -(a_0 n + a_1\mathcal{I}_1 + \cdots + a_{n-1}\mathcal{I}_{n-1}) \\
\mathcal{I}'_n &= -(b_0 n + b_1\mathcal{I}_1 + \cdots + b_{n-1}\mathcal{I}_{n-1})
\end{aligned}$$

Therefore, in spite of Theorem 2.7.2, two $n \times n$ matrices with identical moments $\mathcal{I}_k = \mathcal{I}'_k$, for $k = 1, \ldots, n-1$ may still have $\mathcal{I}_n \neq \mathcal{I}'_n$ if these matrices are not related by a similarity transformation, and hence, have distinct characteristic polynomials. We thus have

Theorem 2.7.3. *Two $n \times n$ matrices are related by a similarity transformation if and only if their n moments $\{\mathcal{I}_k\}_1^n$ are identical.*

Hence,

Corollary 2.7.1. *If two $n \times n$ matrices share the same n moments $\{\mathcal{I}_k\}_1^n$, then their characteristic polynomials are identical.*

Consider the two matrices **A** and **B** given below:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

The two foregoing matrices cannot possibly be related by a similarity transformation, for the first one is the identity matrix, while the second is not. However, the two matrices share the two moments $\mathcal{I}_0 = 2$ and $\mathcal{I}_1 = 2$. Let us now compute the second moments of these matrices:

$$\mathrm{tr}(\mathbf{A}^2) = 2, \quad \mathrm{tr}(\mathbf{B}^2) = \mathrm{tr}\begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} = 10$$

which are indeed different. Therefore, to test whether two different $n \times n$ matrices represent the same linear transformation, and hence, are related by a similarity transformation, we must verify that they share the same set of $n+1$ moments $\{\mathcal{I}_k\}_0^n$. In fact, since all $n \times n$ matrices share the same zeroth moment, only the n moments $\{\mathcal{I}_k\}_1^n$ need be tested for similarity verification. That is, if two $n \times n$ matrices share the same n moments $\{\mathcal{I}_k\}_1^n$, then they represent the same linear transformation, albeit in different coordinate frames.

The foregoing discussion does not apply, in general, to nonsymmetric matrices, for these matrices are not fully characterized by their eigenvalues. For example, consider the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \Rightarrow \quad \mathbf{A}^2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Its two moments of interest are $\mathcal{I}_1 = \mathrm{tr}(\mathbf{A}) = 2$, $\mathcal{I}_2 = 2$, which happen to be the corresponding moments of the 2×2 identity matrix as well. However, while the identity matrix leaves all two-dimensional vectors unchanged, matrix **A** does not.

Now, if two symmetric matrices, say **A** and **B**, represent the same transformation, they are related by a similarity transformation, i.e., a nonsingular matrix **T** exists such that

$$\mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$$

Given **A** and **T**, then, finding **B** is trivial, a similar statement holding if **B** and **T** are given; however, if **A** and **B** are given, finding **T** is more difficult. The latter problem occurs sometimes in robotics in the context of *calibration*, to be discussed in Sect. 2.7.1.

Example 2.7.1. Two symmetric matrices are displayed below. Find out whether they are related by a similarity transformation.

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

Solution: The traces of the two matrices are apparently identical, namely, 4. Now we have to verify whether their second and third moments are also identical. To do this, we need the square and the cube of the two matrices, from which we then compute their traces. Thus, from

$$\mathbf{A}^2 = \begin{bmatrix} 2 & 0 & 3 \\ 0 & 1 & 0 \\ 3 & 0 & 5 \end{bmatrix}, \quad \mathbf{B}^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & -3 \\ 0 & -3 & 2 \end{bmatrix}$$

we readily obtain

$$\mathrm{tr}(\mathbf{A}^2) = \mathrm{tr}(\mathbf{B}^2) = 8$$

Moreover,

$$\mathbf{A}^3 = \begin{bmatrix} 5 & 0 & 8 \\ 0 & 1 & 0 \\ 8 & 0 & 13 \end{bmatrix}, \quad \mathbf{B}^3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 13 & -8 \\ 0 & -8 & 5 \end{bmatrix}$$

whence

$$\mathrm{tr}(\mathbf{A}^3) = \mathrm{tr}(\mathbf{B}^3) = 19$$

Therefore, the two matrices are related by a similarity transformation. Hence, they represent the same linear transformation.

Example 2.7.2. Same as Example 2.7.1, for the two matrices displayed below:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Solution: As in the previous example, the traces of these matrices are identical, i.e., 2. However, $\mathrm{tr}(\mathbf{A}^2) = 10$, while $\mathrm{tr}(\mathbf{B}^2) = 6$. We thus conclude that the two matrices cannot be related by a similarity transformation.

2.7.1 Applications to Redundant Sensing

A sensor, such as a camera or a range finder, is often mounted on a robotic end-effector to determine the *pose*—i.e., the position and orientation, as defined in Sect. 3.2.3—of an object. If two redundant sensors are introduced, and we attach frames $\mathcal{A}$ and $\mathcal{B}$ to each of these, then each sensor can be used to determine the orientation of the end-effector with respect to a reference pose. This is a simple task, for all that is needed is to measure the rotation $\mathbf{R}$ that each of the foregoing frames underwent from the reference pose, in which these frames are denoted by $\mathcal{A}_0$ and $\mathcal{B}_0$, respectively. Let us assume that these measurements produce the orthogonal matrices $\mathbf{A}$ and $\mathbf{B}$, representing $\mathbf{R}$ in $\mathcal{A}$ and $\mathcal{B}$, respectively. With this information we would like to determine the relative orientation $\mathbf{Q}$ of frame $\mathcal{B}$ with respect to frame $\mathcal{A}$, a problem that is called here *instrument calibration*.

We thus have $\mathbf{A} \equiv [\mathbf{R}]_{\mathcal{A}}$ and $\mathbf{B} \equiv [\mathbf{R}]_{\mathcal{B}}$, and hence, the algebraic problem at hand consists in determining $[\mathbf{Q}]_{\mathcal{A}}$ or equivalently, $[\mathbf{Q}]_{\mathcal{B}}$. The former can be obtained from the similarity transformation of Eq. (2.137), which leads to

$$\mathbf{A} = [\mathbf{Q}]_{\mathcal{A}}\mathbf{B}[\mathbf{Q}^T]_{\mathcal{A}}$$

or

$$\mathbf{A}[\mathbf{Q}]_{\mathcal{A}} = [\mathbf{Q}]_{\mathcal{A}}\mathbf{B}$$

This problem could be solved if we had three invariant vectors associated with each of the two matrices $\mathbf{A}$ and $\mathbf{B}$. Then, each corresponding pair of vectors of these triads would be related by Eq. (2.136), thereby obtaining three such vector equations that should be sufficient to compute the nine components of the matrix $\mathbf{Q}$ rotating frame $\mathcal{A}$ into $\mathcal{B}$. However, since $\mathbf{A}$ and $\mathbf{B}$ are orthogonal matrices, they admit only one real invariant vector, namely, their axial vector, and we are short of two vector equations. We thus need two more invariant vectors, represented in both $\mathcal{A}$ and $\mathcal{B}$, to determine $\mathbf{Q}$. The obvious way of obtaining one additional vector in each frame is to take not one, but two measurements of the orientation of $\mathcal{A}_0$ and $\mathcal{B}_0$ with respect to $\mathcal{A}$ and $\mathcal{B}$, respectively. Let the matrices representing these orientations be given, in each of the two coordinate frames, by $\mathbf{A}_i$ and $\mathbf{B}_i$, for $i = 1, 2$. Moreover, let $\mathbf{a}_i$ and $\mathbf{b}_i$, for $i = 1, 2$, be the axial vectors of matrices $\mathbf{A}_i$ and $\mathbf{B}_i$, respectively.

Now, if none of $\mathbf{a}_1$ and $\mathbf{a}_2$ vanishes, and the two vectors are linearly independent, a third vector can be obtained out of each pair, namely,

$$\mathbf{a}_3 = \mathbf{a}_1 \times \mathbf{a}_2, \quad \mathbf{b}_3 = \mathbf{b}_1 \times \mathbf{b}_2 \tag{2.145}$$

If one of the vectors of the two pairs vanishes, we have two more possibilities, namely, the angle of rotation of that orthogonal matrix, $\mathbf{A}_1$ or $\mathbf{A}_2$, whose axial vector vanishes is either 0 or π. If the foregoing angle vanishes, then $\mathcal{A}$ underwent a pure translation from $\mathcal{A}_0$, the same holding, of course, for $\mathcal{B}$ and $\mathcal{B}_0$. This means that the corresponding measurement becomes useless for our purposes, and a new measurement is needed, involving a rotation. If, on the other hand, the same angle is π, then the associated rotation is symmetric and the unit vector $\mathbf{e}$ parallel to its axis can be determined from Eq. (2.50) in both $\mathcal{A}$ and $\mathcal{B}$. This unit vector, then, would play the role of the vanishing axial vector, and we would thus end up, in any event, with two pairs of nonzero vectors, $\{\mathbf{a}_i\}_1^2$ and $\{\mathbf{b}_i\}_1^2$. Moreover, the pairs can be linearly dependent while none of its two vectors vanishes and the vectors are distinct. This is possible if the two rotations take place about the same axis but through distinct angles. In this case, the second rotation becomes useless, should be rejected, and a new second rotation must be taken. In conclusion, we can always find two triads of nonzero vectors, $\{\mathbf{a}_i\}_1^3$ and $\{\mathbf{b}_i\}_1^3$, that are related by

$$\mathbf{a}_i = [\mathbf{Q}]_{\mathcal{A}}\,\mathbf{b}_i, \quad \text{for } i = 1, 2, 3 \tag{2.146}$$

The problem at hand now reduces to computing $[\mathbf{Q}]_{\mathcal{A}}$ from Eq. (2.146). In order to perform this computation, we write the three foregoing equations in matrix form, namely,

$$\mathbf{E} = [\mathbf{Q}]_{\mathcal{A}}\,\mathbf{F} \tag{2.147}$$

with $\mathbf{E}$ and $\mathbf{F}$ defined as

$$\mathbf{E} \equiv \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3\end{bmatrix}, \quad \mathbf{F} \equiv \begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3\end{bmatrix} \tag{2.148}$$

Now, by virtue of the form in which the two vector triads were defined, none of the two above matrices is singular, and hence, we have

$$[\mathbf{Q}]_{\mathcal{A}} = \mathbf{E}\mathbf{F}^{-1} \tag{2.149}$$

Moreover, note that the inverse of $\mathbf{F}$ can be expressed in terms of its columns explicitly, without introducing components, if the concept of *reciprocal bases* is recalled (Brand 1965). Thus,

$$\mathbf{F}^{-1} = \frac{1}{\Delta}\begin{bmatrix}(\mathbf{b}_2 \times \mathbf{b}_3)^T \\ (\mathbf{b}_3 \times \mathbf{b}_1)^T \\ (\mathbf{b}_1 \times \mathbf{b}_2)^T\end{bmatrix}, \quad \Delta \equiv \mathbf{b}_1 \times \mathbf{b}_2 \cdot \mathbf{b}_3 \tag{2.150}$$

Therefore,

$$[\mathbf{Q}]_{\mathcal{A}} = \frac{1}{\Delta}[\mathbf{a}_1(\mathbf{b}_2 \times \mathbf{b}_3)^T + \mathbf{a}_2(\mathbf{b}_3 \times \mathbf{b}_1)^T + \mathbf{a}_3(\mathbf{b}_1 \times \mathbf{b}_2)^T] \tag{2.151}$$

thereby completing the computation of $[\mathbf{Q}]_{\mathcal{A}}$ *directly and with simple vector operations*.

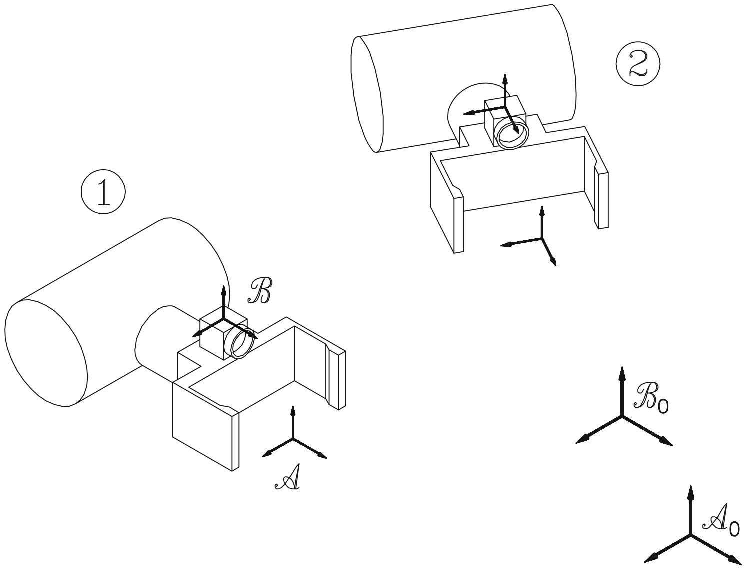

Fig. 2.8 Measuring the orientation of a camera-fixed coordinate frame with respect to a frame fixed to a robotic end-effector

Example 2.7.3 (Hand–Eye Calibration). Determine the relative orientation of a frame $\mathcal{B}$ attached to a camera mounted on a robot end-effector, with respect to a frame $\mathcal{A}$ fixed to the latter, as shown in Fig. 2.8. It is assumed that two measurements of the orientation of the two frames with respect to frames $\mathcal{A}_0$ and $\mathcal{B}_0$ in the reference configuration of the end-effector are available. These measurements produce the orientation matrices $\mathbf{A}_i$ of the frame fixed to the camera and $\mathbf{B}_i$ of the frame fixed to the end-effector, for $i = 1, 2$. The numerical data of this example are given below:

$$\mathbf{A}_1 = \begin{bmatrix} -0.92592593 & -0.37037037 & -0.07407407 \\ 0.28148148 & -0.80740741 & 0.51851852 \\ -0.25185185 & 0.45925926 & 0.85185185 \end{bmatrix}$$

$$\mathbf{A}_2 = \begin{bmatrix} -0.83134406 & 0.02335236 & -0.55526725 \\ -0.52153607 & 0.31240270 & 0.79398028 \\ 0.19200830 & 0.94969269 & -0.24753503 \end{bmatrix}$$

$$\mathbf{B}_1 = \begin{bmatrix} -0.90268482 & 0.10343126 & -0.41768659 \\ 0.38511568 & 0.62720266 & -0.67698060 \\ 0.19195318 & -0.77195777 & -0.60599932 \end{bmatrix}$$

$$\mathbf{B}_2 = \begin{bmatrix} -0.73851280 & -0.54317226 & 0.39945305 \\ -0.45524951 & 0.83872293 & 0.29881721 \\ -0.49733966 & 0.03882952 & -0.86668653 \end{bmatrix}$$

Solution: Shiu and Ahmad (1987) formulated this problem in the form of a matrix linear homogeneous equation, while Chou and Kamel (1988) solved the same problem using quaternions and very cumbersome numerical methods that involve singular-value computations. The latter require an iterative procedure within a Newton–Raphson method, itself iterative, for nonlinear-equation solving. Other attempts to solve the same problem have been reported in the literature, but these also resorted to extremely complicated numerical procedures for nonlinear-equation solving (Chou and Kamel 1991; Horaud and Dornaika 1995). The latter proposed a more concise method based on quaternions—isomorphic to the Euler–Rodrigues parameters—that nevertheless is still computationally expensive.

More recently, Daniilidis (1999) proposed an algorithm based on dual quaternions to simultaneously estimate the relative pose of the two frames of interest. In this book we do not study either quaternions—at least, not by this name—or dual algebra; the former are, in fact, isomorphic to the Euler–Rodrigues parameters of a rotation, which were introduced in Sect. 2.3.6. Dual algebra, in turn, is used to manipulate scalars, vectors and matrices comprising one rotation and one translation, or their statics counterparts, one moment and one force (Angeles 1988). In the above reference, Daniilidis resorts to the singular-value decomposition to find the relative pose in question, but this decomposition slows down the computational procedure. Angeles et al. (2000), in turn, proposed an alternative approach based on the invariance concepts introduced in this section, that leads to an algorithm involving only linear equations. This algorithm, moreover, relies on *recursive least-square* computations, thereby doing away with singular-value computations and allowing for real-time performance. This reference and (Daniilidis 1999) report experimental results.

First, the vector of matrix $\mathbf{A}_i$, represented by $\mathbf{a}_i$, and the vector of matrix $\mathbf{B}_i$, represented by $\mathbf{b}_i$, for $i = 1, 2$, are computed from simple differences of the off-diagonal entries of the foregoing matrices, followed by a division by 2 of all the entries thus resulting, which yields

$$\mathbf{a}_1 = \begin{bmatrix} -0.02962963 \\ 0.08888889 \\ 0.32592593 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 0.07784121 \\ -0.37363778 \\ -0.27244422 \end{bmatrix}$$

$$\mathbf{b}_1 = \begin{bmatrix} -0.04748859 \\ -0.30481989 \\ 0.14084221 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} -0.12999385 \\ 0.44869636 \\ 0.04396138 \end{bmatrix}$$

In the calculations below, 16 digits were used, but only eight are displayed. Furthermore, with the foregoing vectors, we compute $\mathbf{a}_3$ and $\mathbf{b}_3$ from cross products, thus obtaining

$$\mathbf{a}_3 = \begin{bmatrix} 0.09756097 \\ 0.01730293 \\ 0.00415020 \end{bmatrix}$$

$$\mathbf{b}_3 = \begin{bmatrix} -0.07655343 \\ -0.01622096 \\ -0.06091842 \end{bmatrix}$$

Furthermore, Δ is obtained as

$$\Delta = 0.00983460$$

while the individual *rank-one matrices* inside the brackets of Eq. (2.151) are calculated as

$$\mathbf{a}_1(\mathbf{b}_2 \times \mathbf{b}_3)^T = \begin{bmatrix} 0.00078822 & 0.00033435 & -0.00107955 \\ -0.00236467 & -0.00100306 & 0.00323866 \\ -0.00867044 & -0.00367788 & 0.01187508 \end{bmatrix}$$

$$\mathbf{a}_2(\mathbf{b}_3 \times \mathbf{b}_1)^T = \begin{bmatrix} -0.00162359 & 0.00106467 & 0.00175680 \\ 0.00779175 & -0.00510945 & -0.00843102 \\ 0.00568148 & -0.00372564 & -0.00614762 \end{bmatrix}$$

$$\mathbf{a}_3(\mathbf{b}_1 \times \mathbf{b}_2)^T = \begin{bmatrix} -0.00746863 & -0.00158253 & -0.00594326 \\ -0.00132460 & -0.00028067 & -0.00105407 \\ -0.00031771 & -0.00006732 & -0.00025282 \end{bmatrix}$$

whence $\mathbf{Q}$ in the $\mathcal{A}$ frame is readily obtained as

$$[\mathbf{Q}]_{\mathcal{A}} = \begin{bmatrix} -0.84436553 & -0.01865909 & -0.53545750 \\ 0.41714750 & -0.65007032 & -0.63514856 \\ -0.33622873 & -0.75964911 & 0.55667078 \end{bmatrix}$$

thereby completing the desired computation.

2.8 Exercises

N.B.: Unless either a numerical result is required or you are instructed to do otherwise, do not resort to components in the exercises below.

2.1 Prove that the range and the null space of any linear transformation $\mathbf{L}$ of vector space $\mathcal{U}$ into vector space $\mathcal{V}$ are vector spaces as well, the former of $\mathcal{V}$, the latter of $\mathcal{U}$.

2.2 Let $\mathbf{L}$ map $\mathcal{U}$ into $\mathcal{V}$ and $\dim\{\mathcal{U}\} = n$, $\dim\{\mathcal{V}\} = m$. Moreover, let $\mathcal{R}$ and $\mathcal{N}$ be the range and the null space of $\mathbf{L}$, their dimensions being ρ and ν, respectively. Show that $\rho + \nu = n$.

2.3 Given two arbitrary nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{E}^3$, find the matrix $\mathbf{P}$ representing the projection of $\mathcal{E}^3$ onto the subspace spanned by $\mathbf{u}$ and $\mathbf{v}$.

2.4 Verify that $\mathbf{P}$, whose matrix representation in a certain coordinate system is given below, is a projection. Then, describe it geometrically, i.e., identify the plane onto which the projection takes place. Moreover, find the null space of $\mathbf{P}$.

$$[\mathbf{P}] = \frac{1}{3}\begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$$

2.5 If for any three-dimensional vectors $\mathbf{a}$ and $\mathbf{v}$, matrix $\mathbf{A}$ is defined as

$$\mathbf{A} \equiv \frac{\partial(\mathbf{a} \times \mathbf{v})}{\partial \mathbf{v}}$$

then we have

$$\mathbf{A}^T \equiv \frac{\partial(\mathbf{v} \times \mathbf{a})}{\partial \mathbf{v}}$$

Show that $\mathbf{A}$ is skew-symmetric *without introducing components*.

2.6 Let $\mathbf{u}$ and $\mathbf{v}$ be any three-dimensional vectors, and define $\mathbf{T}$ as

$$\mathbf{T} \equiv \mathbf{1} + \mathbf{u}\mathbf{v}^T$$

The (unit) eigenvectors of $\mathbf{T}$ are denoted by $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_3$. Show that, say, $\mathbf{w}_1$ and $\mathbf{w}_2$ are any unit vectors perpendicular to $\mathbf{v}$ and different from each other, whereas $\mathbf{w}_3 = \mathbf{u}/\|\mathbf{u}\|$. Also show that the corresponding eigenvalues, denoted by λ_1, λ_2, and λ_3, associated with $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_3$, respectively, are given as

$$\lambda_1 = \lambda_2 = 1, \quad \lambda_3 = 1 + \mathbf{u} \cdot \mathbf{v}$$

2.7 Show that if $\mathbf{u}$ and $\mathbf{v}$ are any three-dimensional vectors, then

$$\det(\mathbf{1} + \mathbf{u}\mathbf{v}^T) = 1 + \mathbf{u} \cdot \mathbf{v}$$

Hint: Use the results of the Exercise 2.6.

2.8 For the two unit vectors $\mathbf{e}$ and $\mathbf{f}$ in three-dimensional space, define the two reflections

$$\mathbf{R}_1 = \mathbf{1} - 2\mathbf{e}\mathbf{e}^T, \quad \mathbf{R}_2 = \mathbf{1} - 2\mathbf{f}\mathbf{f}^T$$

Now, show that $\mathbf{Q} = \mathbf{R}_1\mathbf{R}_2$ is a rigid-body rotation, and find its axis and its angle of rotation in terms of unit vectors $\mathbf{e}$ and $\mathbf{f}$. Again, no components are permitted in this exercise.

2.9 State the conditions on the unit vectors $\mathbf{e}$ and $\mathbf{f}$, of two reflections $\mathbf{R}_1$ and $\mathbf{R}_2$, respectively, under which a given rotation $\mathbf{Q}$ can be factored into the reflections $\mathbf{R}_1$ and $\mathbf{R}_2$ given in the foregoing exercise. In other words, not every rotation matrix $\mathbf{Q}$ can be factored into those two reflections, for fixed $\mathbf{e}$ and $\mathbf{f}$, but special cases can. Identify these cases.

2.10 For given three-dimensional, non-zero $\mathbf{a}$ and $\mathbf{b}$, find $\mathbf{v}$ that verifies

$$\mathbf{v} + \mathbf{a} \times \mathbf{v} = \mathbf{b}$$

When finding an expression for $\mathbf{v}$ in terms of $\mathbf{a}$ and $\mathbf{b}$, answer the questions below:

(a) Is it possible to find $\mathbf{v}$ for arbitrary $\mathbf{a}$ and $\mathbf{b}$? If so, find an expression for $\mathbf{v}$.
(b) Can $\mathbf{v}$ be orthogonal to $\mathbf{a}$? If so, under which conditions?
(c) Can $\mathbf{v}$ be orthogonal to $\mathbf{b}$? If so, under which conditions?

2.11 Prove that the eigenvalues of the cross-product matrix of the unit vector $\mathbf{e}$ are 0, j, and $-j$, where $j = \sqrt{-1}$. Then show that the eigenvectors associated with the complex eigenvalues are both complex *mutually orthogonal vectors*, and find expressions thereof in terms of $\mathbf{e}$. *Note: Given two n-dimensional vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *defined over the complex field, their scalar product is defined as* $\mathbf{u}^*\mathbf{v}$, *where* $\mathbf{u}^*$ *stands for the transpose conjugate of* $\mathbf{u}$.

2.12 Prove that the eigenvalues of a proper orthogonal matrix $\mathbf{Q}$ are 1, $e^{j\phi}$, and $e^{-j\phi}$, with ϕ denoting the angle of rotation. *Hint: Use the result of the foregoing exercise and the Cayley–Hamilton Theorem.*

2.13 Find the axis and the angle of rotation of the proper orthogonal matrix $\mathbf{Q}$ given below in a certain coordinate frame $\mathcal{F}$.

$$[\mathbf{Q}]_{\mathcal{F}} = \frac{1}{3}\begin{bmatrix} -1 & -2 & 2 \\ -2 & -1 & -2 \\ 2 & -2 & -1 \end{bmatrix}$$

2.14 Prove that the vect$(\cdot)$ and the tr$(\cdot)$ operators are linear homogeneous, i.e., that $\text{vect}(\alpha\mathbf{A} + \beta\mathbf{B}) = \alpha\text{vect}(\mathbf{A}) + \beta\text{vect}(\mathbf{B})$, with a similar expression for $\text{tr}(\alpha\mathbf{A} + \beta\mathbf{B})$.

2.15 Cayley's Theorem, which is not to be confused with the Theorem of Cayley–Hamilton, states that every 3×3 proper orthogonal matrix $\mathbf{Q}$ can be *uniquely* factored as

$$\mathbf{Q} = (\mathbf{1} - \mathbf{S})(\mathbf{1} + \mathbf{S})^{-1}$$

where $\mathbf{S}$ is a skew-symmetric matrix. Find a general expression for $\mathbf{S}$ in terms of $\mathbf{Q}$, and state the condition under which this factoring is not possible.

2.16 Find matrix **S** of Cayley's factoring for **Q** as given in Exercise 2.13.

2.17 If **Q** represents a rotation about an axis parallel to the unit vector **e** through an angle ϕ, then the *Rodrigues vector* $\boldsymbol{\rho}$ of this rotation can be defined as

$$\boldsymbol{\rho} \equiv \tan\left(\frac{\phi}{2}\right)\mathbf{e}$$

Note that if $\mathbf{r}$ and r_0 denote the Euler–Rodrigues parameters of the rotation under study, then $\boldsymbol{\rho} = \mathbf{r}/r_0$. Show that

$$\boldsymbol{\rho} = -\text{vect}(\mathbf{S})$$

for **S** given in Exercise 2.15.

2.18 The vertices of a cube, labeled A, B, ..., H, are located so that A, B, C, and D, as well as E, F, G, and H, are in clockwise order when viewed from outside. Moreover, AE, BH, CG, and DF are edges of the cube, which is to be manipulated so that a mapping of vertices takes place as indicated below:

$$A \rightarrow D,\ B \rightarrow C,\ C \rightarrow G,\ D \rightarrow F$$
$$E \rightarrow A,\ F \rightarrow E,\ G \rightarrow H,\ H \rightarrow B$$

Find the angle of rotation and the angles that the axis of rotation makes with edges AB, AD, and AE.

2.19 (Euler angles) A rigid body can attain an arbitrary configuration starting from any reference configuration, 0, by means of the composition of three rotations about coordinate axes, as described below: Attach axes X_0, Y_0, and Z_0 to the body in the reference configuration and rotate the body through an angle ϕ about Z_0, thus carrying the axes into X_1, Y_1, and Z_1 (=Z_0), respectively. Next, rotate the body through an angle θ about axis Y_1, thus carrying the axes into X_2, Y_2, and Z_2, respectively. Finally, rotate the body through an angle ψ about Z_2 so that the axes coincide with their desired final orientation, X_3, Y_3, and Z_3. Angle ψ is chosen so that axis Z_3 lies in the plane of Z_0 and X_1, whereas angle θ is chosen so as to carry axis Z_1 (=Z_0) into Z_3 (=Z_2). Show that the rotation matrix carrying the body from configuration 0 to configuration 3 is:

$$\mathbf{Q} = \begin{bmatrix} c\theta c\phi c\psi - s\phi s\psi & -c\theta c\phi s\psi - s\phi c\psi & s\theta c\phi \\ c\theta s\phi c\psi + c\phi s\psi & -c\theta s\phi s\psi + c\phi c\psi & s\theta s\phi \\ -s\theta c\psi & s\theta s\psi & c\theta \end{bmatrix}$$

where $c(\cdot)$ and $s(\cdot)$ stand for $\cos(\cdot)$ and $\sin(\cdot)$, respectively. Moreover, show that α, the angle of rotation of **Q** given above, obeys the relation

$$\cos\left(\frac{\alpha}{2}\right) = \cos\left(\frac{\psi+\phi}{2}\right)\cos\left(\frac{\theta}{2}\right)$$

Hint: Let $\mathbf{R}_i$ *be the rotation carrying frame*[3] $\mathcal{F}_{i-1}$ *into* $\mathcal{F}_i$, *for* $i = 1, 2, 3$. *Then, the total rotation carrying* $\mathcal{F}_0$ *into* $\mathcal{F}_3$ *can be found to be* $\mathbf{R}_1\mathbf{R}_2\mathbf{R}_3$, *provided that all three rotation matrices are given in the same frame. However, each* $\mathbf{R}_i$ *admits a simple representation, in canonical form, in* $\mathcal{F}_{i-1}$. *Hence, to represent* $\mathbf{R}_i$, *for* $i = 2, 3$, *in* $\mathcal{F}_0$, *a similarity transformation ála Eqs. (2.127) and (2.128) is needed.*

2.20 Given an arbitrary rigid-body rotation about an axis parallel to the unit vector $\mathbf{e}$ through an angle ϕ, it is possible to find both $\mathbf{e}$ and ϕ using the linear invariants of the rotation matrix, as long as the vector invariant does not vanish. The latter happens if and only if $\phi = 0$ or π. Now, if $\phi = 0$, the associated rotation matrix is the identity, and $\mathbf{e}$ is any three-dimensional vector; if $\phi = \pi$, then we have

$$\mathbf{Q}(\pi) \equiv \mathbf{Q}_\pi = -\mathbf{1} + 2\mathbf{e}\mathbf{e}^T$$

whence we can solve for $\mathbf{e}\mathbf{e}^T$ as

$$\mathbf{e}\mathbf{e}^T = \frac{1}{2}(\mathbf{Q}_\pi + \mathbf{1})$$

Now, it is apparent that the three eigenvalues of $\mathbf{Q}_\pi$ are real and the associated eigenvectors are mutually orthogonal. Find these.

2.21 Explain why *all* the off-diagonal entries of a symmetric rotation matrix *cannot* be negative.

2.22 The three entries above the diagonal of a 3×3 matrix $\mathbf{Q}$ that is supposed to represent a rotation are given below:

$$q_{12} = \frac{1}{2}, \quad q_{13} = -\frac{2}{3}, \quad q_{23} = \frac{3}{4}$$

Without knowing the other entries, explain why the above entries are unacceptable.

2.23 Let $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ be the position vectors of three arbitrary points in three-dimensional space. Now, define a matrix $\mathbf{P}$ as

$$\mathbf{P} \equiv \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 \end{bmatrix}$$

Show that $\mathbf{P}$ is not frame-invariant. *Hint: Show, for example, that it is always possible to find a coordinate frame in which* $\mathrm{tr}(\mathbf{P})$ *vanishes. This task will be eased if you represent the position vectors of the three points in a suitable coordinate frame in which a few of their coordinates will vanish.*

2.24 For $\mathbf{P}$ defined as in Exercise 2.23, let

$$q \equiv \mathrm{tr}(\mathbf{P}^2) - \mathrm{tr}^2(\mathbf{P})$$

[3] $\mathcal{F}_i$ is obviously frame X_i, Y_i, Z_i.

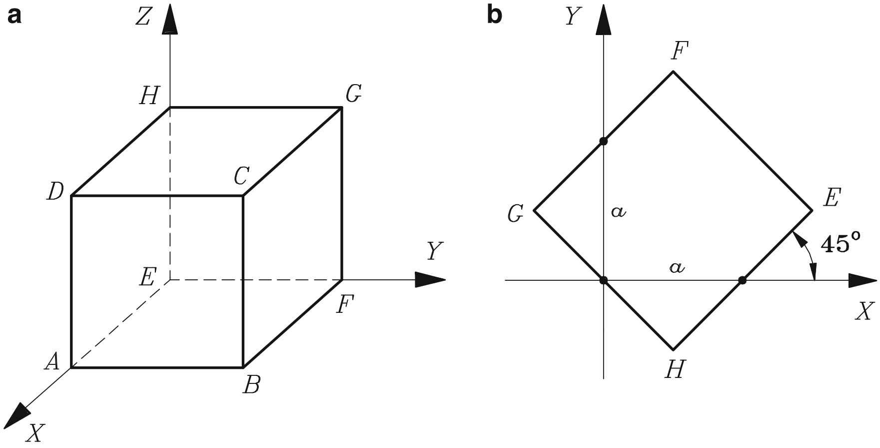

Fig. 2.9 A cube in two different configurations

Show that q vanishes if the three given points and the origin are collinear, for $\mathbf{P}$ represented in any coordinate frame.

2.25 For $\mathbf{P}$ defined, again, as in Exercise 2.23, show that $\mathbf{P}\mathbf{P}^T$ is invariant under frame-rotations about the origin, and becomes singular if and only if either the three given points are collinear or the origin lies in the plane of the three points. Note that this matrix is more singularity-robust than $\mathbf{P}$.

2.26 The diagonal entries of a rotation matrix are known to be -0.5, 0.25, and -0.75. Find the off-diagonal entries.

2.27 As a generalization to the foregoing exercise, discuss how you would go about finding the off-diagonal entries of a rotation matrix whose diagonal entries are known to be a, b, and c. *Hint: This problem can be formulated as finding the intersection of the coupler curve of a four-bar spherical linkage (Chiang 1988), which is a curve on a sphere, with a certain parallel of the same sphere.*

2.28 Shown in Fig. 2.9a is a cube that is to be displaced in an assembly operation to a configuration in which face $EFGH$ lies in the XY plane, as indicated in Fig. 2.9b. Compute the unit vector $\mathbf{e}$ parallel to the axis of the rotation involved and the angle of rotation ϕ, for $0 \le \phi \le \pi$.

2.29 The axes X_1, Y_1, Z_1 of a frame $\mathcal{F}_1$ are attached to the base of a robotic manipulator, whereas the axes X_2, Y_2, Z_2 of a second frame $\mathcal{F}_2$ are attached to its end-effector, as shown in Fig. 2.10. Moreover, the origin P of $\mathcal{F}_2$ has the $\mathcal{F}_1$-coordinates $(1, -1, 1)$. Furthermore, the orientation of the end effector with respect to the base is defined by a rotation $\mathbf{Q}$, whose representation in $\mathcal{F}_1$ is given by

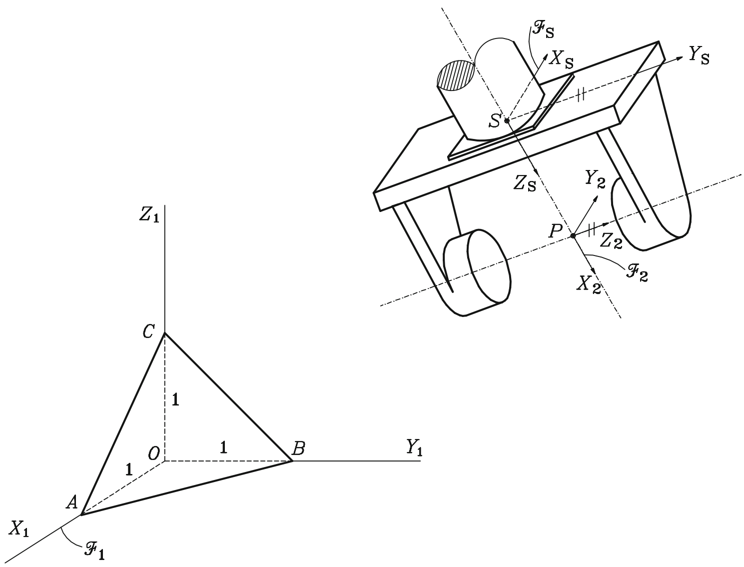

Fig. 2.10 Robotic EE approaching a stationary object ABC

$$[\mathbf{Q}]_1 = \frac{1}{3}\begin{bmatrix} 1 & 1-\sqrt{3} & 1+\sqrt{3} \\ 1+\sqrt{3} & 1 & 1-\sqrt{3} \\ 1-\sqrt{3} & 1+\sqrt{3} & 1 \end{bmatrix}$$

(a) What are the end-effector coordinates of point C of Fig. 2.10?

(b) The end-effector is approaching the ABC plane shown in Fig. 2.10. What is the equation of the plane in end-effector coordinates? Verify your result by substituting the answer to (a) into this equation.

2.30 Shown in Fig. 2.11 is a cube of unit side, which is composed of two parts. Frames (X_0, Y_0, Z_0) and (X_1, Y_1, Z_1) are attached to each of the two parts, as illustrated in the figure. The second part is going to be picked up by a robotic gripper as the part is transported on a belt conveyor and passes close to the stationary first part. Moreover, the robot is to assemble the cube by placing the second part onto the first one in such a way that vertices A_1, B_1, C_1 are coincident with vertices A_0, B_0, C_0. Determine the axis and the angle of rotation that will carry the second part onto the first one as described above.

2.31 A piece of code meant to produce the entries of rotation matrices is being tested. In one run, the code produced a matrix with diagonal entries -0.866, -0.866, -0.866. Explain how without looking at the other entries, you can decide that the code has a bug.

2.32 Shown in Fig. 2.12 is a rigid cube of unit side in three configurations. The second and the third configurations are to be regarded as images of the first

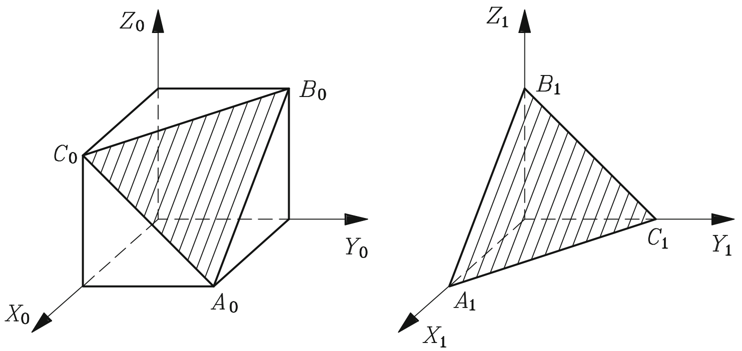

Fig. 2.11 Roboticized assembly operation

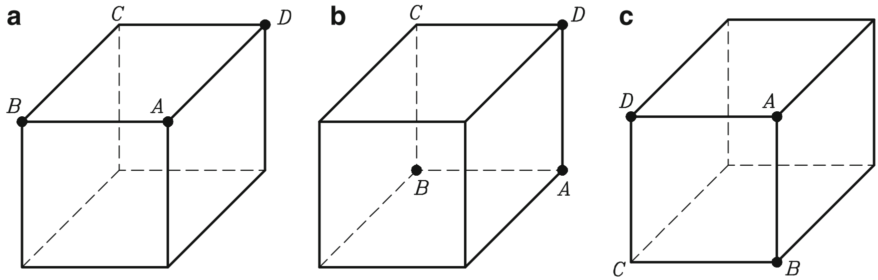

Fig. 2.12 Three configurations of a cube

one. One of the last two configurations is a reflection, and the other is a rotation of the first one. Identify the rotated configuration and find its associated invariants.

2.33 Two frames, $\mathcal{G}$ and $\mathcal{C}$, are attached to a robotic gripper and to a camera mounted on the gripper, respectively. Moreover, the camera is rigidly attached to the gripper, and hence, the orientation of $\mathcal{C}$ with respect to $\mathcal{G}$, denoted by $\mathbf{Q}$, remains constant under gripper motions. The orientation of the gripper with respect to a frame $\mathcal{B}$ fixed to the base of the robot was measured in both $\mathcal{G}$ and $\mathcal{C}$. Note that this orientation is measured in $\mathcal{G}$ simply by reading the joint encoders, which report values of the joint variables, as discussed in detail in Chap. 4. The same orientation is measured in $\mathcal{C}$ from estimations of the coordinates of a set of points fixed to $\mathcal{B}$, as seen by the camera.

Two measurements of the above-mentioned orientation, denoted $\mathbf{R}_1$ and $\mathbf{R}_2$, were taken in $\mathcal{G}$ and $\mathcal{C}$, with the numerical values reported below:

$$[\mathbf{R}_1]_{\mathcal{G}} = \begin{bmatrix} 0.667 & 0.333 & 0.667 \\ -0.667 & 0.667 & 0.333 \\ -0.333 & -0.667 & 0.667 \end{bmatrix},$$

$$[\mathbf{R}_1]_{\mathcal{C}} = \begin{bmatrix} 0.500 & 0 & -0.866 \\ 0 & 1.000 & 0 \\ 0.866 & 0 & 0.500 \end{bmatrix},$$

$$[\mathbf{R}_2]_{\mathcal{G}} = \begin{bmatrix} 0.707 & 0.577 & 0.408 \\ 0 & 0.577 & -0.816 \\ -0.707 & 0.577 & 0.408 \end{bmatrix},$$

$$[\mathbf{R}_2]_{\mathcal{C}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.346 & -0.938 \\ 0 & 0.938 & 0.346 \end{bmatrix}$$

(a) Verify that the foregoing matrices represent rotations.
(b) Verify that the first two matrices represent, in fact, the same rotation $\mathbf{R}_1$, albeit in different coordinate frames.
(c) Repeat item (b) for $\mathbf{R}_2$.
(d) Find $[\mathbf{Q}]_{\mathcal{G}}$. Is your computed $\mathbf{Q}$ orthogonal? If not, can the error be attributed to data-incompatibility? to roundoff-error amplification?

2.34 The orientation of the end-effector of a given robot is to be inferred from joint-encoder readouts, which report an orientation given by a matrix $\mathbf{Q}$ in $\mathcal{F}_1$-coordinates, namely,

$$[\mathbf{Q}]_1 = \frac{1}{3}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$$

(a) Show that the above matrix can indeed represent the orientation of a rigid body.
(b) What is $\mathbf{Q}$ in end-effector coordinates, i.e., in a frame $\mathcal{F}_7$, if Z_7 is chosen parallel to the axis of rotation of $\mathbf{Q}$?

2.35 The rotation $\mathbf{Q}$ taking a coordinate frame $\mathcal{B}$, fixed to the base of a robot, into a coordinate frame $\mathcal{G}$, fixed to its gripper, and the position vector $\mathbf{g}$ of the origin of $\mathcal{G}$ have the representations in $\mathcal{B}$ given below:

$$[\mathbf{Q}]_{\mathcal{B}} = \frac{1}{3}\begin{bmatrix} 1 & 1-\sqrt{3} & 1+\sqrt{3} \\ 1+\sqrt{3} & 1 & 1-\sqrt{3} \\ 1-\sqrt{3} & 1+\sqrt{3} & 1 \end{bmatrix}, \qquad [\mathbf{g}]_{\mathcal{B}} = \begin{bmatrix} 1-\sqrt{3} \\ \sqrt{3} \\ 1+\sqrt{3} \end{bmatrix}$$

Moreover, let $\mathbf{p}$ be the position vector of any point $\mathcal{P}$ of the three-dimensional space, its coordinates in $\mathcal{B}$ being (x, y, z). The robot is supported by a cylindrical column C of circular cross section, bounded by planes Π_1 and Π_2. These are given below:

$$C\colon\ x^2 + y^2 = 4;\ \Pi_1\colon\ z = 0;\ \Pi_2\colon\ z = 10$$

Find the foregoing equations in $\mathcal{G}$ coordinates.

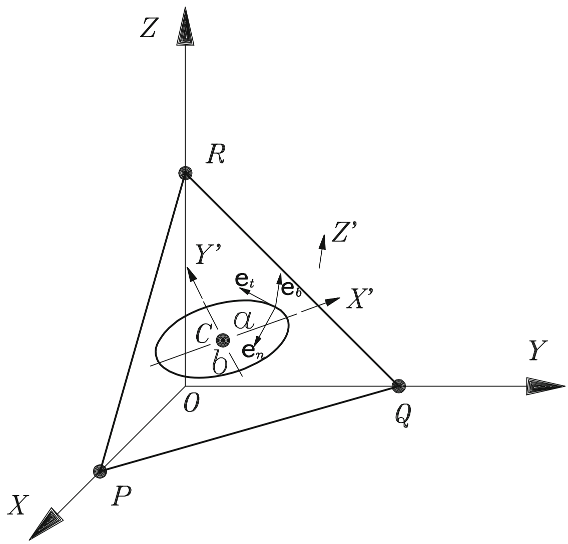

Fig. 2.13 An elliptical path on an inclined plane

2.36 A certain point of the gripper of a robot is to trace an elliptical path of semiaxes a and b, with center at C, the centroid of triangle PQR, as shown in Fig. 2.13. Moreover, the semiaxis of length a is parallel to edge PQ, while the ellipse lies in the plane of the triangle, and all three vertices are located a unit distance away from the origin.

(a) For $b = 2a/3$, the gripper is to keep a fixed orientation with respect to the unit tangent, normal, and binormal vectors of the ellipse, denoted by $\mathbf{e}_t$, $\mathbf{e}_n$, and $\mathbf{e}_b$, respectively.[4] Determine the matrix representing the rotation undergone by the gripper from an orientation in which vector $\mathbf{e}_t$ is parallel to the coordinate axis X, while $\mathbf{e}_n$ is parallel to Y and $\mathbf{e}_b$ to Z. Express this matrix in $X,\ Y,\ Z$ coordinates, if the equation of the ellipse, in parametric form, is given as

$$x' = a\cos\varphi,\ y' = b\sin\varphi,\ z' = 0$$

the orientation of the gripper thus becoming a function of φ.

(b) Find the value of φ for which the angle of rotation of the gripper, with respect to the coordinate axes $X,\ Y,\ Z$, becomes π.

2.37 With reference to Exercise 2.28, find Euler angles ϕ, θ, and ψ that will rotate the cube of Fig. 2.9a into the attitude displayed in Fig. 2.9b. For a definition of Euler angles, see Exercise 2.19

2.38 Find a sequence of Euler angles ϕ, θ, and ψ, as defined in Exercise 2.19, that will carry triangle A_1, B_1, C_1 into triangle A_0, B_0, C_0, of Fig. 2.11.

[4] An account of curve geometry is given in Sect. 11.2.