Any transformation of a matrix having the form of Eq. (4–41) is known as a *similarity transformation.*

It is appropriate at this point to consider the properties of the determinant formed from the elements of a square matrix. As is customary, we shall denote such a determinant by vertical bars, thus: $|\mathbf{A}|$. It will be noticed that the definition of matrix multiplication is identical with that for the multiplication of determinants (cf. Bôcher, *Introduction to Higher Algebra*, p. 26). Hence

$$|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}|.$$

Since the determinant of the unit matrix is 1, the determinantal form of the orthogonality conditions, Eq. (4–36), can be written

$$|\tilde{\mathbf{A}}| \cdot |\mathbf{A}| = 1.$$

Further, as the value of a determinant is unaffected by interchanging rows and columns, we can write

$$|\mathbf{A}|^2 = 1, \tag{4–42}$$

which implies that the determinant of an orthogonal matrix can only be $+1$ or -1. (The geometrical significance of these two values will be considered in the next section.)

When the matrix is not orthogonal the determinant does not have these simple values, of course. It can be shown however that the value of the determinant is invariant under a similarity transformation. Multiplying the equation (4–41) for the transformed matrix from the right by $\mathbf{B}$, we obtain the relation

$$\mathbf{A}'\mathbf{B} = \mathbf{BA},$$

or in determinantal form

$$|\mathbf{A}'| \cdot |\mathbf{B}| = |\mathbf{B}| \cdot |\mathbf{A}|.$$

Since the determinant of $\mathbf{B}$ is merely a number, and not zero,* we can divide by $|\mathbf{B}|$ on both sides to obtain the desired result:

$$|\mathbf{A}'| = |\mathbf{A}|.$$

In discussing rigid body motion later, all these properties of matrix transformations, especially of orthogonal matrices, will be employed. In addition, other properties are needed, and they will be derived as the occasion requires.

4–4 THE EULER ANGLES

It has already been noted (cf. p. 131) that the nine elements a_{ij} are not suitable as generalized coordinates because they are not independent quantities. The six

* If it were zero there could be no inverse operator $\mathbf{B}^{-1}$ (by Cramer's rule), which is required in order that Eq. (4–41) make sense.

relations that express the orthogonality conditions, Eqs. (4–9) or Eqs. (4–15), of course reduce the number of independent elements to three. But in order to characterize the motion of a rigid body there is an additional requirement the matrix elements must satisfy, beyond those implied by orthogonality. In the previous section it was pointed out that the determinant of a real orthogonal matrix could have the value $+1$ or -1. The following argument shows however that an orthogonal matrix whose determinant is -1 cannot represent a physical displacement of a rigid body.

Consider a simple matrix with the determinant -1:

$$\mathbf{S} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = -\mathbf{1}.$$

The transformation **S** has the effect of changing the sign of each of the components or coordinate axes (cf. Fig. 4–6). Such an operation transforms a right-handed coordinate system into a left-handed one and is known as an *inversion* or *reflection* of the coordinate axes.

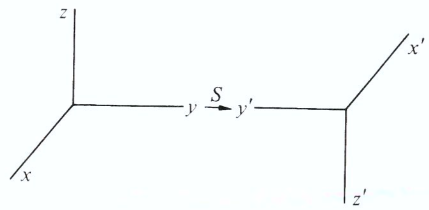


FIGURE 4–6
Inversion of the coordinate axes.

From the nature of this operation it is clear that an inversion of a right-handed system into a left-handed one cannot be accomplished by any *rigid* change in the orientation of the coordinate axes. An inversion therefore never corresponds to a physical displacement of a rigid body. What is true for **S** is equally valid for any matrix whose determinant is -1, for any such matrix can be written as the product of **S** with a matrix whose determinant is $+1$, and thus includes the inversion operation. Consequently it cannot describe a rigid change in orientation. Therefore, the transformations representing rigid body motion must be restricted to matrices having the determinant $+1$. Another method of reaching this conclusion starts from the fact that the matrix of transformation must evolve continuously from the unit matrix, which of course has the determinant $+1$. It would be incompatible with the continuity of the motion to have the matrix determinant change suddenly from its initial value $+1$ to -1 at some given time. Orthogonal transformations with determinant $+1$ are said to be *proper*, so naturally those with the determinant -1 are called *improper*.

In order to describe the motion of rigid bodies in the Lagrangian formulation of mechanics, it will therefore be necessary to seek three independent parameters specifying the orientation of a rigid body in such a manner that the corresponding orthogonal matrix of transformation has the determinant $+1$. Only when such generalized coordinates have been found can one write a Lagrangian for the system and obtain the Lagrangian equations of motion. A number of such sets of parameters have been described in the literature, but the most common and useful are the *Euler angles*.* We shall therefore define these angles at this point, and show how the elements of the orthogonal transformation matrix can be expressed in terms of them.

One can carry out the transformation from a given cartesian coordinate system to another by means of three successive rotations performed in a specific sequence. The Euler angles are then defined as the three successive angles of rotation. Within limits, the choice of rotation angles is arbitrary. The main convention that will be followed here is used widely in celestial mechanics, applied mechanics, and frequently in molecular and solid state physics. Other conventions will be described below.

The sequence employed here is started by rotating the initial system of axes, xyz, by an angle ϕ counterclockwise about the z axis, and the resultant coordinate system is labeled the $\xi\eta\zeta$ axes. In the second stage the intermediate axes, $\xi\eta\zeta$, are rotated about the ξ axis counterclockwise by an angle θ to produce another intermediate set, the $\xi'\eta'\zeta'$ axes. The ξ' axis is at the intersection of the xy and $\xi'\eta'$ planes and is known as the *line of nodes*. Finally the $\xi'\eta'\zeta'$ axes are rotated system of axes. Figure 4–7 illustrates the various stages of the sequence. The counterclockwise by an angle ψ about the ζ' axis to produce the desired $x'y'z'$ Euler angles θ, ϕ, and ψ thus completely specify the orientation of the $x'y'z'$ system relative to the xyz and can therefore act as the three needed generalized coordinates.†

The elements of the complete transformation **A** can be obtained by writing the matrix as the triple product of the separate rotations, each of which has a relatively simple matrix form. Thus, the initial rotation about z can be described by a matrix **D**:

$$\boldsymbol{\xi} = \mathbf{D}\mathbf{x},$$

where $\boldsymbol{\xi}$ and $\mathbf{x}$ stand for column matrices. Similarly the transformation from $\xi\eta\zeta$ to $\xi'\eta'\zeta'$ can be described by a matrix **C**,

$$\boldsymbol{\xi}' = \mathbf{C}\boldsymbol{\xi},$$

and the last rotation to $x'y'z'$ by a matrix **B**

$$\mathbf{x}' = \mathbf{B}\boldsymbol{\xi}'.$$

* Also denoted, interchangeably, as Euler's angles, or Eulerian angles.

† A number of minor variations will be found in the older literature even within this convention. The differences are not very great, but they are often sufficient to frustrate easy comparison of the end formulae, such as the matrix elements. Greatest confusion, perhaps, arises from the occasional use of left-handed coordinate systems (as by Osgood and by Margenau and Murphy). Some European authors agree with the practice given here except that the meanings of ϕ and ψ are interchanged.

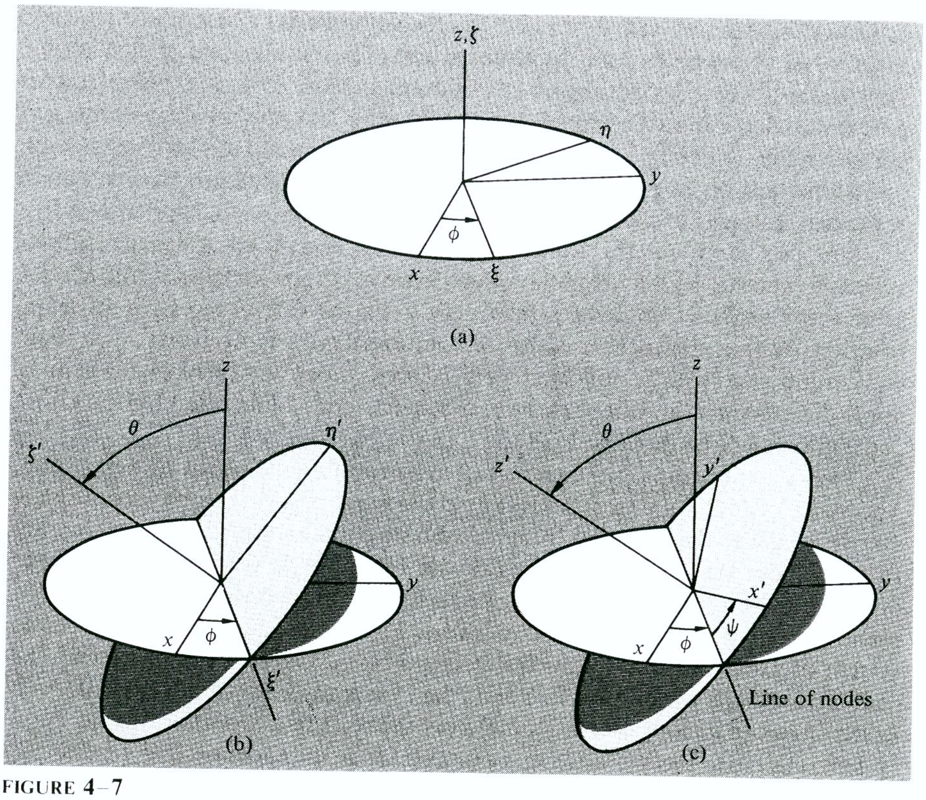


FIGURE 4–7
The rotations defining the Eulerian angles.

Hence the matrix of the complete transformation

$$\mathbf{x}' = \mathbf{A}\mathbf{x}$$

is the product of the successive matrices,

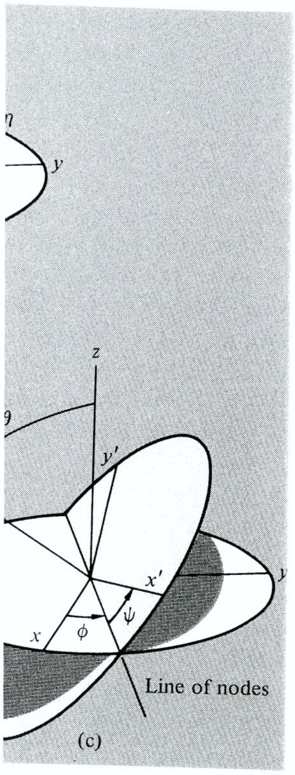
$$\mathbf{A} = \mathbf{BCD}.$$

Now the **D** transformation is a rotation about z, and hence has a matrix of the form (cf. Eq. (4–17))

$$\mathbf{D} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad (4\text{–}43)$$

The **C** transformation corresponds to a rotation about ξ, with the matrix

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \qquad (4\text{–}44)$$



(c)

and finally **B** is a rotation about ζ' and therefore has the same form as **D**:

$$\mathbf{B} = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{4-45}$$

The product matrix $\mathbf{A} = \mathbf{BCD}$ then follows as

$$\mathbf{A} = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix}. \tag{4-46}$$

The inverse transformation from body coordinates to space axes

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{x}'$$

is then given immediately by the transposed matrix $\tilde{\mathbf{A}}$:

$$\mathbf{A}^{-1} = \tilde{\mathbf{A}} = \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & \sin\theta\sin\phi \\ \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & -\sin\theta\cos\phi \\ \sin\theta\sin\psi & \sin\theta\cos\psi & \cos\theta \end{pmatrix}. \tag{4-47}$$

Verification of the multiplication, and demonstration that **A** represents a proper, orthogonal matrix will be left to the exercises.

It will be noted that the sequence of rotations used to define the final orientation of the coordinate system is to some extent arbitrary. The initial rotation could be taken about any of the three Cartesian axes. In the subsequent two rotations, the only limitation is that no two successive rotations can be about the same axis. A total of twelve conventions is therefore possible in defining the Euler angles (in a right-handed coordinate system). The two most frequently used conventions differ only in the choice of axis for the second rotation. In the Euler's angle definitions described above, and used throughout the book, the second rotation is about the intermediate x axis. We will refer to this choice as the *x-convention.* In quantum mechanics, nuclear physics, and particle physics, the custom has arisen to take the second defining rotation about the intermediate y axis,* and this form will be denoted as the *y-convention.*

A third convention is commonly used in engineering applications relating to the orientation of moving vehicles such as aircraft and satellites. Both the x- and y-conventions have the drawback that when the primed coordinate system is only slightly different from the unprimed system, the angles ϕ and ψ become indistinguishable, as their respective axes of rotation, z and z' are then nearly coincident. To get around this problem all three rotations are taken around different axes. The first rotation is about the vertical axis and gives the *heading* or *yaw* angle. The second is around a perpendicular axis fixed in the vehicle and

* The usage of Wigner in *Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra* and of Rose in *Elementary Theory of Angular Momentum* appears to have been decisive in this regards.

normal to the figure axis; it is measured by the *pitch* or *attitude* angle. Finally the third angle is one of rotation about the figure axis of the vehicle and is the *roll* or *bank* angle. Because all three axes are involved in the rotations it will be designated as the *xyz-convention* (although the order of axes chosen may actually be different). This last convention is sometimes referred to as the *Tait–Bryan* angles.

While only the x-convention will be used in the text, for reference purposes Appendix B lists all the formulas involving Euler's angles, such as rotation matrices, in both the y- and xyz-conventions.

4–5 THE CAYLEY–KLEIN PARAMETERS AND RELATED QUANTITIES

We have seen that only three independent quantities are needed to specify the orientation of a rigid body. Nonetheless, there are occasions when it is desirable to use sets of variables containing more than the minimum number of quantities to describe a rotation even though they are not suitable as generalized coordinates. Thus, Felix Klein introduced the set of four parameters bearing his name to facilitate the integration of complicated gyroscopic problems. The Euler angles are difficult to use in numerical computation because of the large number of trigonometric functions involved, and the four-parameter representations are much better adapted for use on computers. Further, the four-parameter sets are of great theoretical interest in branches of physics beyond the scope of this book, wherever rotations or rotational symmetry are involved. It therefore seems worthwhile to devote some space to describe these enlarged parameter sets. However, none of the results of this section will be directly used in the discussion of rigid body motion in the following chapter.

In the previous sections we employed on occasion a two-dimensional real space with axes x_1 and x_2 to illustrate the properties of orthogonal transformations. We shall now consider a different two-dimensional space, this time having complex axes denoted by u and v. A general linear transformation in such a space appears as

$$\begin{aligned} u' &= \alpha u + \beta v, \\ v' &= \gamma u + \delta v, \end{aligned} \tag{4–48}$$

with the corresponding transformation matrix

$$\mathbf{Q} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}. \tag{4–49}$$

As it stands $\mathbf{Q}$ has eight quantities to be specified, since each of the four elements is complex. To reduce the transformation to three independent quantities, additional conditions must be imposed on $\mathbf{Q}$. For much of the following discussion, it is sufficient to require that the transformation be such that $\mathbf{Q}$ is unitary:

$$\mathbf{Q}^{\dagger}\mathbf{Q} = \mathbf{1} = \mathbf{Q}\mathbf{Q}^{\dagger}. \tag{4–50}$$