5.1 The Mathematical Description of Rotations

We develop the theory of rotations, progressing from infinitesimal rotations to finite rotations, in particular considering the group-theoretic aspects. Some of this material is found in Hand and Finch Chapters 7 and 8, but much is not.

We want to understand how rotation of the coordinate system affects the position vector $\vec{r}$. Such rotations will be of interest to us in two physical situations:

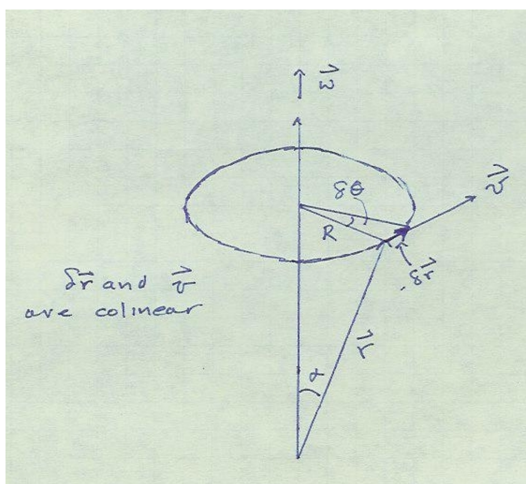
- Apparent motion in a rotating coordinate system of a point that is fixed relative to an inertial coordinate system.
- Rotation of a rigid body about some axis. Recall that a rigid body is a body for which the relative distance between any pair of points is fixed. Such a body may still rotate about any axis. A point P in the body follows rotational motion about the axis of motion relative to the non-rotating system.

We will consider the physical systems in the sections following this one. This section focuses on the mathematical language and general properties of rotational transformations.

5.1.1 Infinitesimal Rotations

Vector Cross-Product Version

Consider the following figure:



$\vec{r}$ points in an arbitrary direction in the non-rotating frame. The vector $\vec{r}$ is rotated by an angle $\delta\theta$ about the z axis. We are allowed to take the rotation axis to coincide with the z axis because $\vec{r}$ is arbitrary. If we define a vector $\delta\vec{\theta}$ that points along the axis of rotation (the z axis) and has magnitude $\delta\theta$, then the change $\delta\vec{r}$ in $\vec{r}$ is related to $\delta\vec{\theta}$ by

$$\delta\vec{r} = \delta\vec{\theta} \times \vec{r}$$

where $\times$ indicates a vector cross-product, and the rotated vector in the non-rotating system is

$$\vec{r}\,' = \vec{r} + \delta\vec{r}$$

The cross-product gives the correct direction for the displacement $\delta\vec{r}$ (perpendicular to the axis and $\vec{r}$) and the correct amplitude ($|\delta\vec{r}| = R\,\delta\theta = r\,\delta\theta \sin\alpha$). If we then divide by the time δt required to make this displacement, we have

$$\frac{\delta\vec{r}}{\delta t} = \frac{\delta\vec{\theta}}{\delta t} \times \vec{r} \Longrightarrow \dot{\vec{r}} = \vec{\omega} \times \vec{r}$$

where ω is the angular frequency of rotation about the axis and $\vec{\omega}$ points along the axis of rotation also.

Matrix Version

A more generic and therefore more useful way to look at a rotation is as a matrix operation on vectors. The infinitesimal rotation can be viewed as a matrix operation:

$$\vec{r}\,' = \vec{r} + \delta\theta\,\hat{z} \times \vec{r} = \begin{pmatrix} x - y\,\delta\theta \\ y + x\,\delta\theta \\ z \end{pmatrix} \equiv \mathbf{R}_{\delta\vec{\theta}}\,\vec{r}$$

with

$$\mathbf{R}_{\delta\vec{\theta}} = \begin{pmatrix} 1 & -\delta\theta & 0 \\ \delta\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{1} + \delta\theta\,\mathbf{M}_z \qquad \mathbf{M}_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

where we have defined the infinitesimal rotation matrix $\mathbf{R}_{\delta\vec{\theta}}$ and the matrix $\mathbf{M}_z$. More generally, one can show that an infinitesimal rotation about an arbitrary axis can be written in matrix form using

$$\mathbf{R}_{\delta\vec{\theta}} = \mathbf{1} + \left(\delta\vec{\theta}\cdot\hat{x}\,\mathbf{M}_x + \delta\vec{\theta}\cdot\hat{y}\,\mathbf{M}_y + \delta\vec{\theta}\cdot\hat{z}\,\mathbf{M}_z\right) \equiv \mathbf{1} + \delta\vec{\theta}\cdot\vec{\mathbf{M}}$$

with

$$\mathbf{M}_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \mathbf{M}_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \qquad \mathbf{M}_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$\vec{\mathbf{M}}$ is called the infinitesimal rotation **generator** because, obviously, it can be used to generate any infinitesimal rotation matrix $\mathbf{R}_{\delta\vec{\theta}}$ when combined with the rotation vector $\delta\vec{\theta}$. A simple way to write the generators is

$$\left(\mathbf{M}_i\right)_{jk} = -\epsilon_{ijk}$$

where ϵ_{ijk} is the completely antisymmetric Levi-Civita symbol (see Appendix A). It is useful to know that the above matrices satisfy the following relations:

$$\mathbf{M}_x^2 = -\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{M}_y^2 = -\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \mathbf{M}_z^2 = -\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$\mathbf{M}_i\mathbf{M}_j - \mathbf{M}_j\mathbf{M}_i \equiv [\mathbf{M}_i, \mathbf{M}_j] = \epsilon_{ijk}\,\mathbf{M}_k$$

Combining Infinitesimal Rotations

It is straightforward to see how to combine two infinitesimal rotations, even if they are not about the same axis. Clearly, our formula $\delta\vec{r} = \delta\vec{\theta} \times \vec{r}$ holds regardless of the orientation of $\delta\vec{\theta}$. So we obtain the result of two successive *infinitesimal* rotations $\delta\vec{\theta}_1$ and $\delta\vec{\theta}_2$ by

$$\begin{aligned}
\vec{r}_2' &= \vec{r}_1' + \delta\vec{\theta}_2 \times \vec{r}_1' \\
&= \vec{r} + \delta\vec{\theta}_1 \times \vec{r} + \delta\vec{\theta}_2 \times \left(\vec{r} + \delta\vec{\theta}_1 \times \vec{r}\right) \\
&= \vec{r} + \left(\delta\vec{\theta}_1 + \delta\vec{\theta}_2\right) \times \vec{r}
\end{aligned}$$

where in the last line we have dropped terms quadratic in the infinitesimal rotation angles. Thus, we see that the effect of two infinitesimal rotations $\delta\vec{\theta}_1$ and $\delta\vec{\theta}_2$ is found simply by obtaining the result of a rotation by the sum rotation vector $\delta\vec{\theta}_1 + \delta\vec{\theta}_2$. Obviously, if $\delta\vec{\theta}_1$ and $\delta\vec{\theta}_2$ are aligned, then the angles sum simply. But if the two rotation axes are not aligned, addition of the angle vectors describes the correct way to combine the rotations. In terms of matrices, we have

$$\begin{aligned}
\vec{r}_2' &= \left(\mathbf{1} + \delta\vec{\theta}_2 \cdot \vec{\mathbf{M}}\right) \vec{r}_1' \\
&= \left(\mathbf{1} + \delta\vec{\theta}_2 \cdot \vec{\mathbf{M}}\right)\left(\mathbf{1} + \delta\vec{\theta}_1 \cdot \vec{\mathbf{M}}\right) \vec{r} \\
&= \left(\mathbf{1} + \left[\delta\vec{\theta}_1 + \delta\vec{\theta}_2\right] \cdot \vec{\mathbf{M}}\right) \vec{r}
\end{aligned}$$

The implication of the addition rule for rotations is that angular velocities add in simple fashion also:

$$\vec{\omega}_{tot} = \vec{\omega}_1 + \vec{\omega}_2$$

NOTE: The simple addition rules for infinitesimal rotations and angular velocities do **not** in general hold for finite rotations, discussed in the next section.

5.1.2 Finite Rotations

There are two different ways of determining the appropriate form for finite rotations – integration of infinitesimal transformations and direct evaluation via direction cosines.

Integration of Infinitesimal Transformations

First, consider a finite rotation about a single axis, $\vec{\theta}$. That rotation can be built up as a product of infinitesimal rotations:

$$\mathbf{R}_{\vec{\theta}} = \lim_{N\to\infty} \left(\mathbf{1} + \frac{1}{N}\,\vec{\theta} \cdot \vec{\mathbf{M}}\right)^N$$

where $\delta\vec{\theta} = \frac{1}{N}\,\vec{\theta}$, which is infinitesimal in the limit $N \to \infty$. This can be rewritten (expand out the infinite product, or just realize that the above is one of the definitions of the exponential function):

$$\mathbf{R}_{\vec{\theta}} = \sum_{n=0}^{\infty} \frac{1}{n!}\left(\vec{\theta} \cdot \vec{\mathbf{M}}\right)^n = \exp\left(\vec{\theta} \cdot \vec{\mathbf{M}}\right)$$

The second equality is not much use since an exponential with a matrix argument only has meaning as a power series expansion. If we now specialize to a rotation about the z axis, we get $\vec{\theta} \cdot \vec{\mathbf{M}} = \theta\, \mathbf{M}_z$. Using the relation above for $\mathbf{M}_z^2$, we can rewrite the sum as

$$\begin{aligned}
\mathbf{R} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \sum_{n=0}^{\infty} \left[\frac{\theta^{2n}\,(-1)^n}{(2n)!} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{\theta^{2n+1}\,(-1)^n}{(2n+1)!} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{M}_z \right] \\
&= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + (\cos\theta + \sin\theta\, \mathbf{M}_z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

Of course, similar forms hold if the rotation is aligned with the x or y axis. For more general rotations, the parameterization in terms of the components of $\vec{\theta}$ is not particularly clear; we will develop a simpler parameterization, **Euler angles**, later.

Direction Cosines

Consider our rotation of the vector $\vec{r}$ to be a rotation of a frame F in which $\vec{r}$ is fixed relative to a nonrotating frame F'. In Hand and Finch, these are termed the "body" and "space" coordinate systems, respectively. The original vector has components (x, y, z) in the nonrotating frame. The transformation $\vec{r}\,' = \mathbf{R}\,\vec{r}$ gives the components of the vector in the nonrotating F' frame after the rotation; denote these components by (x', y', z'). In the F frame, the vector retains its original components (x, y, z) after rotation because the F frame rotates with the vector. Thus, the rotation matrix $\mathbf{R}$ provides a linear relationship between the components (x, y, z) and (x', y', z'). There are 9 obvious coefficients of this relationship, the 9 components of $\mathbf{R}$. One way to parameterize these 9 components is to consider how they relate to the angles between the F' coordinate axes $\hat{x}\,'$, $\hat{y}\,'$, and $\hat{z}\,'$ and the F coordinate axes $\hat{x}$, $\hat{y}$, and $\hat{z}$. It turns out that, in fact, $\mathbf{R}$ is the matrix of these dot products:

$$\mathbf{R} = \begin{pmatrix} \hat{x}\,' \cdot \hat{x} & \hat{x}\,' \cdot \hat{y} & \hat{x}\,' \cdot \hat{z} \\ \hat{y}\,' \cdot \hat{x} & \hat{y}\,' \cdot \hat{y} & \hat{y}\,' \cdot \hat{z} \\ \hat{z}\,' \cdot \hat{x} & \hat{z}\,' \cdot \hat{y} & \hat{z}\,' \cdot \hat{z} \end{pmatrix}$$

That this form works can be seen by applying it to the unit vectors of the F frame; we get back the unit vectors of the F frame written in terms of the unit vectors of the F' frame. Since rotation is a linear operation,

$$\mathbf{R}\left(\alpha\,\vec{a} + \beta\,\vec{b}\right) = \alpha\,\mathbf{R}\,\vec{a} + \beta\,\mathbf{R}\,\vec{b}$$

the same matrix is valid for rotating an arbitrary vector $\vec{r}$. In the case of a simple rotation about the z axis, the above direction cosine form is obviously consistent with the matrix operator form given above. The two forms must be consistent in general since the linearity of the relationship between $\vec{r}$ and $\vec{r}\,'$ allows there to be only one such rotation matrix.

Orthogonality of Finite Rotation Matrices

Coordinate rotation matrices are very much like the rotation matrices we obtained for transforming between generalized and normal coordinates in the coupled oscillation problem (Section 3.2). In particular, rotation matrices must be **orthogonal** matrices ($\mathbf{R}^T = \mathbf{R}^{-1}$) because they must be **norm-preserving**. It is intuitively obvious that rotation of a vector about an axis must preserve its length, but it can be seen to be explicitly true by considering an infinitesimal rotation:

$$\left|\vec{r}\,'\right|^2 = (\vec{r} + \delta\vec{r}) \cdot (\vec{r} + \delta\vec{r}) = |\vec{r}|^2 + 2\,\delta\vec{r} \cdot \vec{r} + |\delta\vec{r}|^2 = |\vec{r}|^2$$

where the second term vanishes because $\delta\vec{r}$ is normal to $\vec{r}$ and the third term has been dropped because it is of second order in $\delta\vec{r}$. If infinitesimal rotations preserve length, then so do finite rotations. If we require this condition to be true, we obtain

$$\vec{r}^{\,T}\vec{r} = \vec{r}\,'^{\,T}\vec{r}\,' = (\mathbf{R}\,\vec{r})^T\,(\mathbf{R}\,\vec{r}) = \vec{r}^{\,T}\,\mathbf{R}^T\,\mathbf{R}\,\vec{r} \quad\Longrightarrow\quad \mathbf{R}^T\mathbf{R} = \mathbf{1} \quad\Longrightarrow\quad \mathbf{R}^T = \mathbf{R}^{-1}$$

(We have used the ^{T} notation for displaying dot products of vectors.) Hence, we obtain the orthogonality condition.

Orthogonality implies that the columns of $\mathbf{R}$, treated as vectors, are in fact orthonormalized:

$$\delta_{ij} = \left(\mathbf{R}^T\mathbf{R}\right)_{ij} = \sum_k R_{ki}R_{kj}$$

(We refer to the components of $\mathbf{R}$ by R_{ab}, dropping the boldface notation because each component is just a number.) There are 6 conditions implied by the above, leaving only 3 independent components of $\mathbf{R}$. As we will see later, these 3 degrees of freedom can be parameterized in terms of the **Euler angles**.

One can easily see that the "transposed" relation also holds. Norm preservation also implies

$$\vec{r}\,'^{\,T}\vec{r}\,' = \vec{r}^{\,T}\vec{r} = \left(\mathbf{R}^T\vec{r}\,'\right)^T\left(\mathbf{R}^T\vec{r}\,'\right) = \vec{r}\,'^{\,T}\mathbf{R}\,\mathbf{R}^T\vec{r}\,' \quad\Longrightarrow\quad \mathbf{R}\,\mathbf{R}^T = \mathbf{1}$$

That is, $\mathbf{R}^T$ is the "right inverse" of $\mathbf{R}$ also.[1] Written out in components, we thus have

$$\delta_{ij} = \left(\mathbf{R}\mathbf{R}^T\right)_{ij} = \sum_k R_{ik}R_{jk}$$

The rows of $\mathbf{R}$ also are othonormal.

5.1.3 Interpretation of Rotations

We at this point should comment on the two different possible physical interpretations of the rotation matrices $\mathbf{R}$. We made a similar distinction when considering coordinate transformations in Section 2.1.10.

Active transformations

In an active transformation, we think of the transformation as actively rotating the particle whose position is given by the vector $\vec{r}(t)$ relative to the coordinate axes. The rotations of the form $\vec{\omega} \times \vec{r}$ that we began with are really of that form. The coordinate system that the rotation is relative to is inertial, the coordinate system in which the rotating vector is fixed is noninertial.

[1] We could have used the fact that, for square matrices, left inverses and right inverses are always the same. But it's nice to see it directly.

Passive transformation

We think of a passive transformation as simply a relabeling of points in space according to a new coordinate system. In this picture, the coordinate system in which the particle is at rest is inertial and the rotating system is not.

Why there can be confusion

The difficulty arises because the transformations relating the two systems are mathematically identical in the two cases, but the physical interpretation is very different. We use the mathematical equivalence to relate one type of transformation to another to allow us to write down the rules. But, the definition of which system is inertial differs between the cases, so we must be very careful.

5.1.4 Scalars, Vectors, and Tensors

We have so far discussed rotations only in terms of their effect on vectors, objects that we have an intuitive feel for. But the link between rotations and vectors is much deeper, leading us to the generic concept of **tensors**.

Vectors and Vector Coordinate Representations

As hinted at in the discussion of the interpretation of rotations, there are multiple concepts described by the same mathematics. As indicated above, we tend to think of a rotation as actively rotating a vector from one position in physical space to another. The rotating coordinate system is noninertial and thus physics is modified there.

But, regardless of which coordinate system is inertial, we are frequently required to transform vectors from one coordinate system to another. This is the passive sense of a rotation transformation. The vector is not undergoing any dynamics, it is simply being looked at in a different frame. The vector, as a physics entity, does not change by looking at it in a different coordinate system. But its **coordinate representation** – the set of numbers that give the components of the vector along the coordinate axes – does change. Rotation matrices thus not only provide the means to dynamically rotate a vector about some axis, but they also provide a method to obtain the coordinate representation in one frame from that in another frame.

It is important to recognize the difference between a vector and its coordinate representation: a vector is a physical entity, independent of any coordinate system, while its coordinate representations let us "write down" the vector in a particular coordinate system. To make this clear, we will use the standard vector notation $\vec{r}$ to denote the "coordinate-independent" vector and underlined notation $\underline{\vec{r}}$ and $\underline{\vec{r}}'$ to denote its coordinate representations in the F and F' coordinate systems.

Formal Definitions of Scalars, Vectors, and Tensors

We have so far relied on some sort of unstated, intuitive definition of what a vector is. We may, on the other hand, make use of the properties of the coordinate representations of vectors to **define** what is meant by the term "vector." A vector $\vec{v}$ is defined to be an object with coordinate representations in different frames that are related by our orthogonal rotation matrices:

$$\begin{aligned}\underline{\vec{v}}' &= \mathbf{R}\,\underline{\vec{v}} \\ \left(\underline{\vec{v}}'\right)_i &= \sum_j (\mathbf{R})_{ij}\,(\underline{\vec{v}})_j\end{aligned}$$

where in the second line we have written out the relation between coordinate representations in the two frames component-by-component. One might worry that the above definition of a vector is circular because rotation matrices are to some extent defined in terms of vectors. This worry can be eliminated by taking the direction cosine definition of rotation matrices – that definition rests only on the idea of coordinate systems and direction cosines.

We will use as a convenient shorthand the following:

$$v_i' = R_{ij}\,v_j$$

Let us explain the shorthand carefully. First, repeated indices are defined to imply summation (the Einstein summation convention), so $\sum$ symbols are not given explicitly. Second, as noted earlier, the quantity R_{ij} refers to the ij component of the matrix $\mathbf{R}$. Finally, the quantity v_j is the jth component of the coordinate representation $\underline{\vec{v}}$, $v_j = (\underline{\vec{v}})_j$. Similarly, $v_i' = (\underline{\vec{v}}')_i$. That is, the vector object is referred to with an arrow as $\vec{a}$ while the coordinate representations are underlined with an arrow, $\underline{\vec{a}}$ and $\underline{\vec{a}}'$, and the components of the representations are referred to using subscripts but no underlines or arrows, a_i and a_i'.[2]

We have an intuitive definition of a **scalar** as an object whose coordinate representation is independent of frame. Explicitly, the transformation property of a scalar s with coordinate representations $\underline{s}$ and $\underline{s}'$ is

$$\underline{s}' = \underline{s}$$

for any pair of frames. Since the representation of s is independent of frame, there is no distinction between s, $\underline{s}$, and $\underline{s}'$.

We can see, for example, that the norm of the position vector is a scalar by relating its values in two different frames:

$$\underline{\vec{r}}' \cdot \underline{\vec{r}}' = r_i'\,r_i' = R_{ij}\,r_j\,R_{ik}\,r_k = \delta_{jk}\,r_j\,r_k = r_j\,r_j = \underline{\vec{r}} \cdot \underline{\vec{r}}$$

where we have used $\underline{\vec{r}} = \mathbf{R}^T\,\underline{\vec{r}}'$ and the orthonormality property $R_{ij}\,R_{ik} = \delta_{jk}$. In general, the dot product $\vec{a} \cdot \vec{b}$ of two vectors is a scalar by the same argument.

The generalization of scalars and vectors is **tensors**, or, more specifically, **rank** n **tensors**. A rank n tensor $\mathcal{T}$ is an object that has coordinate representation $\underline{\mathcal{T}}$ with N^n components $T_{i_1 \cdots i_n}$ (where N is the dimensionality of the physical space, $N = 3$ for what we have so far considered) with transformation properties

$$T'_{i_1 \cdots i_n} = R_{i_1 j_1} \cdots R_{i_n j_n} T_{j_1 \cdots j_n}$$

A vector is a rank 1 tensor and a scalar is a rank 0 tensor. A rank 2 tensor has coordinate representations that look like square $N \times N$ matrices; what distinguishes a rank 2 tensor

[2] We note that there is no need for underlines for the coordinate representation a_i because a subscript without an arrow implies consideration of a coordinate representation.

from a simple matrix is the relation between the coordinate representations in different frames. It is important to remember the distinction!

An alternate, but equivalent, definition of a rank n tensor is an object whose product with n vectors is a scalar for any n vectors. That is, if we claim $\mathcal{T}$ is a rank n tensor with representation $T_{i_1\cdots i_n}$ in a particular frame, then, for n arbitrary vectors $\{\vec{a}_i\}$ with components $\{a_{i,j_i}\}$ in that frame, if we calculate in each frame the quantity

$$\underline{s} = T_{i_1\cdots i_n}\, a_{1,i_1} \cdots a_{n,i_n}$$

it is required that s be the same in every frame regardless of the choice of the $\{\vec{a}_i\}$. We can see the equivalence of the two definitions by equating the component representations of s calculated in different frames:

$$\begin{aligned} \underline{s}' &= \underline{s} \\ T'_{i_1\cdots i_n}\, a'_{1,i_1} \cdots a'_{n,i_n} &= T_{p_1\cdots p_n}\, a_{1,p_1} \cdots a_{n,p_n} \\ T'_{i_1\cdots i_n}\, R_{i_1 j_1}\, a_{1,j_1} \cdots R_{i_n j_n}\, a_{n,j_n} &= T_{p_1\cdots p_n}\, a_{1,p_1} \cdots a_{n,p_n} \end{aligned}$$

Since the above must hold for all possible choices of the $\{\vec{a}_i\}$, we may conclude

$$\begin{aligned} T'_{i_1\cdots i_n}\, R_{i_1 j_1} \cdots R_{i_n j_n} &= T_{j_1\cdots j_n} \\ R_{k_1 j_1} \cdots R_{k_n j_n}\, T'_{i_1\cdots i_n}\, R_{i_1 j_1} \cdots R_{i_n j_n} &= R_{k_1 j_1} \cdots R_{k_n j_n}\, T_{j_1\cdots j_n} \\ T'_{i_1\cdots i_n}\, \delta_{k_1 i_1} \cdots \delta_{k_n i_n} &= R_{k_1 j_1} \cdots R_{k_n j_n}\, T_{j_1\cdots j_n} \\ T'_{k_1\cdots k_n} &= R_{k_1 j_1} \cdots R_{k_n j_n}\, T_{j_1\cdots j_n} \end{aligned}$$

where we have used the "transposed" orthonormality condition $R_{kj}\, R_{ij} = \delta_{ki}$. Thus, we recover our other definition of a rank n tensor.

Rank 2 tensors are special because their coordinate representations $\underline{\mathcal{T}}$ look like simple $N \times N$ matrices. In particular, the transformation of a rank 2 tensor has a simple matrix form:

$$T'_{ij} = R_{ik}\, R_{jl}\, T_{kl} = R_{ik}\, T_{kl}\, R_{lj}$$
$$\underline{\mathcal{T}}' = \mathbf{R}\, \underline{\mathcal{T}}\, \mathbf{R}^T = \mathbf{R}\, \underline{\mathcal{T}}\, \mathbf{R}^{-1}$$

where $\underline{\mathcal{T}}$ and $\underline{\mathcal{T}}'$ are $N \times N$ matrices. The last expression is the **similarity transformation** of the $N \times N$ matrix $\underline{\mathcal{T}}$ by the orthogonal matrix $\mathbf{R}$.

Examples of Tensors

- One obvious rank 2 tensor is the outer product of two vectors:

  $$T_{ij} = a_i\, b_j \quad \text{or} \quad \mathcal{T} = \vec{a}\, \vec{b}^T$$

  Since each vector transforms as a rank 1 tensor, it is obvious that the above product transforms as a rank 2 tensor.

- More generally, if we take a rank m tensor with coordinate representation components $U_{i_1\cdots i_m}$ and a rank n tensor with coordinate representation components $V_{j_1\cdots j_n}$ and **contract** over – *i.e.*, match up indices and sum, the generalization of a dot product – any p pairs of indices, then the resulting set of quantities is a rank $m + n - 2p$ tensor. Proving it is clearly a tedious exercise in index arithmetic relying on the rotation matrix orthogonality relation $R_{ki}\, R_{kj} = \delta_{ij}$ and its transpose relation $R_{ik}\, R_{jk} = \delta_{ij}$. Taking $p = 0$ as a special case gives us the simple outer product of the two tensors, which gives the previous example when both tensors are rank 1.

- The identity matrix is a rank 2 tensor and, in fact, it is **isotropic**, meaning that its coordinate representation is the same in all frames. Let's just try transforming it to see this:

$$1'_{ij} = R_{ik}\,R_{jl}\,1_{kl} = R_{ik}\,R_{jl}\,\delta_{kl} = R_{ik}\,R_{jk} = \delta_{ij}$$

(We used the "transposed" orthonormality condition $R_{ik}\,R_{jk} = \delta_{ij}$.) So, we see that the identity matrix has representation δ_{ij} in any frame and that the representations in different frames are related by the appropriate transformation relations.

- We can demonstrate that the ϵ_{ijk} Levi-Civita symbol is an isotropic rank 3 tensor. Let's explicitly calculate the effect of the transformation rule on it:

$$\epsilon'_{ijk} = R_{il}\,R_{jm}\,R_{kn}\,\epsilon_{lmn}$$

We may evaluate the above by recognizing that the "transposed" orthonormality condition on $\mathbf{R}$ implies that the rows of $\mathbf{R}$ look like N mutually orthonormal vectors in N-dimensional space. (Here we use the term vector more loosely – we have no need to prove that these rows behave like vectors in rotated frames, we only need the fact that their component representations in a given frame looks like that of N orthonormal vectors.) Denote these "vectors" by $\vec{\mathbf{R}}^r_i$, where $\left(\vec{\mathbf{R}}^r_i\right)_j = R_{ij}$. (The ^{r} superscript indicates we are treating the rows, rather than the columns, of $\mathbf{R}$ as vectors.) With this notation, the above product looks like $\vec{\mathbf{R}}^r_i \cdot \left(\vec{\mathbf{R}}^r_j \times \vec{\mathbf{R}}^r_k\right)$. In $N = 3$ dimensions, the expression will only be nonvanishing when the triplet ijk is a cyclic or anticyclic combination; and the expression will have magnitude 1 and take the sign of the permutation (cyclic or anticyclic). These are exactly the properties of ϵ_{ijk}, so we have

$$\epsilon'_{ijk} = \epsilon_{ijk}$$

So the Levi-Civita symbol is an isotropic rank 3 tensor for $N = 3$ (and for arbitrary N, though we will not prove it here). Note that this implies some properties of $\vec{\mathbf{M}}$:

1. When treated as a single rank 3 tensor $\mathcal{M}$ with coordinate representation components $M_{ijk} = \left(\vec{\mathbf{M}}_i\right)_{jk} = -\epsilon_{ijk}$, $\mathcal{M}$ is clearly an isotropic rank 3 tensor. For this particularly interesting case, we will take the symbol $\vec{\mathbf{M}}$ to stand for the rank 3 tensor $\mathcal{M}$. Since $\vec{\mathbf{M}}$ is isotropic, there is no distinction between $\underline{\vec{\mathbf{M}}}$ and $\vec{\mathbf{M}}$.
2. Given a vector $\vec{a}$, the quantity $\vec{a} \cdot \vec{\mathbf{M}}$ has in frames F and F' coordinate representations $\underline{\vec{a}} \cdot \underline{\vec{\mathbf{M}}} = \underline{\vec{a}} \cdot \vec{\mathbf{M}}$ and $\underline{\left(\vec{a} \cdot \vec{\mathbf{M}}\right)}' = \underline{\vec{a}}' \cdot \underline{\vec{\mathbf{M}}}' = \underline{\vec{a}}' \cdot \vec{\mathbf{M}}$, where the last step in each case is possible because $\vec{\mathbf{M}}$ is isotropic. Thus, only the coordinate representation of the vector $\vec{a}$ need be changed to write $\vec{a} \cdot \vec{\mathbf{M}}$ in different frames.

- With the above, we may show that the operator $\vec{a}\,\times$ is a rank 2 tensor. We are referring to the operator, not just the vector $\vec{a}$. The operation $\vec{a} \times \vec{b}$ is written as

$$\left(\vec{a} \times \vec{b}\right)_i = \epsilon_{ijk}\,a_j\,b_k = (\epsilon_{ijk}\,a_j)\,b_k = \left(\vec{a} \cdot \vec{\mathbf{M}}\right)_{ik}\,b_k$$

which shows that $\vec{a}\,\times$ looks like the operation of a matrix on a vector. Since we know $\vec{a}$ is a vector and $\vec{\mathbf{M}}$ is a rank 3 tensor, the contraction over one index must yield a rank 2 tensor as discussed above. But, since we did not explicitly prove the general relation, let's prove this specific case explicitly using the transformation properties of $\vec{a}$ and $\vec{\mathbf{M}}$:

$$\begin{aligned}
\underline{\left(\vec{a}\cdot\vec{\mathbf{M}}\right)'_{ij}} = -a'_k\epsilon'_{kij} &= -R_{kl}\,a_l\,R_{km}\,R_{in}\,R_{jp}\,\epsilon_{mnp} \\
&= -\delta_{lm}\,a_l\,R_{in}\,R_{jp}\,\epsilon_{mnp} \\
&= -R_{in}\,R_{jp}\,a_m\,\epsilon_{mnp} \\
&= R_{in}\,R_{jp}\underline{\left(\vec{\mathbf{a}}\cdot\vec{\mathbf{M}}\right)_{np}}
\end{aligned}$$

So, indeed, $\vec{a}\cdot\vec{\mathbf{M}}$, and thus $\vec{a}\,\times$, transforms like a rank 2 tensor. This last fact will be useful in the next section on dynamics in rotating systems.

5.1.5 Comments on Lie Algebras and Lie Groups

We make some qualitative comments on Lie algebras and groups. Such concepts are not incredibly useful here, but serve to allow analogies to be made between the physical rotations we have talked about here and transformations in more complex systems. We progress through some definitions of mathematical objects that lead up to Lie algebras and Lie groups.

Groups

A **group** is defined to be a set of elements with a binary operation rule that specifies how combinations of pairs of elements yield other members of the group. The set must be **closed** under the operation – binary combinations of members of the set may only yield other members of the set – to be called a group. In addition, the binary operation must be associative, $a\,(b\,c) = (a\,b)\,c$, there must be an identity element 1 such that $a\,1 = a$, and each element must have an inverse such that $a^{-1}\,a = a\,a^{-1} = 1$. Note that the group operation need not be commutative.

Fields

A **field** is a group that has two kinds of operations, addition and multiplication. It is a group under each of the operations separately, and in addition satisfies distributivity: $a\,(b+c) = a\,b + a\,c$.

Group Algebras

A **group algebra** is a combination of a field F (with addition $+$ and multiplication $\cdot$) and a group G (with multiplication $*$), consisting of all *finite* linear combinations of elements of G with coefficients from F, $a\,g + b\,h$, where a and b belong to F and g and h belong to G. The group operations on F and G continue to work:

$$\begin{aligned}
a\,g + b\,g &= (a+b)\,g \\
a\cdot\sum_i a_i\,g_i &= \sum_i (a\cdot a_i)\,g_i \\
\left(\sum_i a_i\,g_i\right) * \left(\sum_j b_j\,h_j\right) &= \sum_{i,j}(a_i\cdot b_j)\,(g_i * h_j)
\end{aligned}$$

Lie Algebras

A **Lie algebra** is a group algebra with the additional conditions that the group elements $\{\tau_i\}$ belonging to G satisfy the commutation relations

$$[\tau_i, \tau_j] \equiv \tau_i \, \tau_j - \tau_j \, \tau_i = c_{ij}^k \tau_k$$

The $\{c_{ij}^k\}$ are called the **structure constants** of the Lie algebra. They satisfy $c_{ij}^k = -c_{ji}^k$. (Actually, the real definition of a Lie algebra is slightly more generic, allowing for definition of a commutator without G being part of a group algebra.) A Lie algebra may have a finite or infinite number of group members, though in general we will only consider finite ones. The matrices $\mathbf{M}_x$, $\mathbf{M}_y$, and $\mathbf{M}_z$ defined earlied form a Lie algebra with the real numbers as the field F.

Lie Groups

A **Lie group** is the exponential of the Lie algebra, consisting of all possible elements

$$a_m = \exp\left(\sum_k \theta_m^k \tau_k\right)$$

where the θ_m^k are members of the field F and the τ_k are members of the group G. The exponential is defined in terms of its infinite power series expansion, which is well-defined for a group. Note that the members of the Lie group are not members of the Lie algebra because, by dint of the power series expansion, they are *infinite* linear combinations of members of the Lie algebra. The Lie group is entirely separate from the Lie algebra. The Lie group is, as its name indicates, a group. Moreover, thanks to the method of definition, the group is differentiable with respect to the members of the field.

Representations

Our rotation matrices are a particular **representation** of the rotation group in three dimensions, $O(3)$, which is a Lie group. The $\mathbf{M}$ matrices are a representation of the Lie algebra that generate the Lie group. The word representation is used because the matrices are not really necessary; it is only the group and field rules that are fundamental. The representation in terms of matrices provides one way that those rules can be realized, but not the only way.

Significance

These definitions serve only to illustrate that there is a fundamental structure underlying rotation matrices, that they form a Lie group can be generated from a small set of elements forming a Lie algebra. Lie algebras and groups are found everywhere in physics because they are the means by which continuous transformations are performed. As we saw in Section 2.1.10, continuous coordinate transformations under which a Lagrangian is invariant are of great significance because they give rise to conserved canonical momenta. The continuous parameters of such symmetry transformations are the field members θ_m^k, so the conserved momenta are given by the group members τ_k. When rotations are symmetry transformations, we thus naturally see that angular momentum $\vec{l} = \vec{p}^T \vec{\mathbf{M}} \vec{r}$ is conserved. We will see another Lie group, the Lorentz group, in connection with special relativity.