S. Widnall
16.07 Dynamics
Fall 2009
Lecture notes based on J. Peraire Version 2.0

Lecture L3 - Vectors, Matrices and Coordinate Transformations

By using vectors and defining appropriate operations between them, physical laws can often be written in a simple form. Since we will making extensive use of vectors in Dynamics, we will summarize some of their important properties.

Vectors

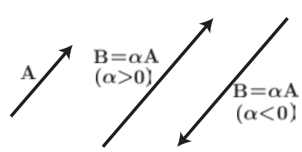
For our purposes we will think of a *vector* as a mathematical representation of a physical entity which has both *magnitude* and *direction* in a 3D space. Examples of physical vectors are forces, moments, and velocities. Geometrically, a vector can be represented as arrows. The length of the arrow represents its magnitude. Unless indicated otherwise, we shall assume that parallel translation does not change a vector, and we shall call the vectors satisfying this property, *free vectors*. Thus, two vectors are equal if and only if they are parallel, point in the same direction, and have equal length.
Vectors are usually typed in boldface and scalar quantities appear in lightface italic type, e.g. the vector quantity $\mathbf{A}$ has magnitude, or modulus, $A = |\mathbf{A}|$. In handwritten text, vectors are often expressed using the arrow, or underbar notation, e.g. $\overrightarrow{A}$, $\underline{A}$.

Vector Algebra

Here, we introduce a few useful operations which are defined for free vectors.

Multiplication by a scalar

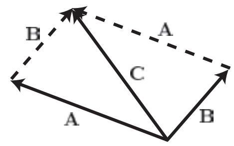
If we multiply a vector $\mathbf{A}$ by a scalar α, the result is a vector $\mathbf{B} = \alpha\mathbf{A}$, which has magnitude $B = |\alpha|A$. The vector $\mathbf{B}$, is parallel to $\mathbf{A}$ and points in the same direction if $\alpha > 0$. For $\alpha < 0$, the vector $\mathbf{B}$ is parallel to $\boldsymbol{A}$ but points in the opposite direction (antiparallel).



If we multiply an arbitrary vector, $\mathbf{A}$, by the inverse of its magnitude, $(1/A)$, we obtain a *unit vector* which is parallel to $\mathbf{A}$. There exist several common notations to denote a unit vector, e.g. $\hat{\mathbf{A}}$, $\mathbf{e}_A$, etc. Thus, we have that $\hat{\mathbf{A}} = \boldsymbol{A}/A = \boldsymbol{A}/|\boldsymbol{A}|$, and $\boldsymbol{A} = A\,\hat{\mathbf{A}}$, $|\hat{\mathbf{A}}| = 1$.

Vector addition

Vector addition has a very simple geometrical interpretation. To add vector $\boldsymbol{B}$ to vector $\boldsymbol{A}$, we simply place the tail of $\boldsymbol{B}$ at the head of $\boldsymbol{A}$. The sum is a vector $\boldsymbol{C}$ from the tail of $\boldsymbol{A}$ to the head of $\boldsymbol{B}$. Thus, we write $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$. The same result is obtained if the roles of $\boldsymbol{A}$ are reversed $\boldsymbol{B}$. That is, $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$. This commutative property is illustrated below with the parallelogram construction.

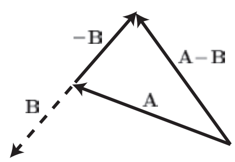


Since the result of adding two vectors is also a vector, we can consider the sum of multiple vectors. It can easily be verified that vector sum has the property of association, that is,

$$(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}).$$

Vector subtraction

Since $\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A} + (-\boldsymbol{B})$, in order to subtract $\boldsymbol{B}$ from $\boldsymbol{A}$, we simply multiply $\boldsymbol{B}$ by -1 and then add.

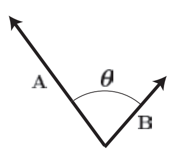


Scalar product ("Dot" product)

This product involves two vectors and results in a scalar quantity. The scalar product between two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$, is denoted by $\boldsymbol{A} \cdot \boldsymbol{B}$, and is defined as

$$\boldsymbol{A} \cdot \boldsymbol{B} = AB \cos\theta \ .$$

Here θ, is the angle between the vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ when they are drawn with a common origin.



We note that, since $\cos\theta = \cos(-\theta)$, it makes no difference which vector is considered first when measuring the angle θ. Hence, $\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{B} \cdot \boldsymbol{A}$. If $\boldsymbol{A} \cdot \boldsymbol{B} = 0$, then either $A = 0$ and/or $B = 0$, or, $\boldsymbol{A}$ and $\boldsymbol{B}$ are orthogonal, that is, $\cos\theta = 0$. We also note that $\boldsymbol{A} \cdot \boldsymbol{A} = A^2$. If one of the vectors is a unit vector, say $B = 1$, then $\boldsymbol{A} \cdot \hat{\boldsymbol{B}} = A\cos\theta$, is the projection of vector $\boldsymbol{A}$ along the direction of $\hat{\boldsymbol{B}}$.

Exercise

Using the definition of scalar product, derive the Law of Cosines which says that, for an arbitrary triangle with sides of length A, B, and C, we have

$$C^2 = A^2 + B^2 - 2AB\cos\theta \ .$$

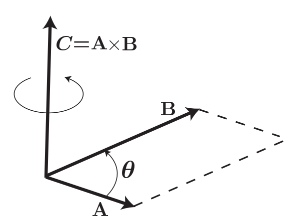
Here, θ is the angle opposite side C. *Hint : associate to each side of the triangle a vector such that* $\mathbf{C} = \mathbf{A} - \mathbf{B}$, *and expand* $C^2 = \mathbf{C} \cdot \mathbf{C}$.

Vector product ("Cross" product)

This product operation involves two vectors $\mathbf{A}$ and $\mathbf{B}$, and results in a new vector $\mathbf{C} = \mathbf{A} \times \mathbf{B}$. The magnitude of $\mathbf{C}$ is given by,

$$C = AB\sin\theta \ ,$$

where θ is the angle between the vectors $\mathbf{A}$ and $\mathbf{B}$ when drawn with a common origin. To eliminate ambiguity, between the two possible choices, θ is always taken as the angle smaller than π. We can easily show that C is equal to the area enclosed by the parallelogram defined by $\boldsymbol{A}$ and $\boldsymbol{B}$.
The vector $\mathbf{C}$ is orthogonal to both $\mathbf{A}$ and $\mathbf{B}$, i.e. it is orthogonal to the plane defined by $\mathbf{A}$ and $\mathbf{B}$. The direction of $\mathbf{C}$ is determined by the *right-hand* rule as shown.



From this definition, it follows that

$$\mathbf{B} \times \mathbf{A} = -\mathbf{A} \times \mathbf{B} \ ,$$

which indicates that vector multiplication is *not* commutative (but anticommutative). We also note that if $\boldsymbol{A} \times \boldsymbol{B} = \mathbf{0}$, then, either $\boldsymbol{A}$ and/or $\boldsymbol{B}$ are zero, or, $\boldsymbol{A}$ and $\boldsymbol{B}$ are parallel, although not necessarily pointing in the same direction. Thus, we also have $\boldsymbol{A} \times \boldsymbol{A} = \mathbf{0}$.
Having defined vector multiplication, it would appear natural to define vector division. In particular, we could say that "$\mathbf{A}$ *divided by* $\mathbf{B}$", is a vector $\mathbf{C}$ such that $\mathbf{A} = \mathbf{B} \times \mathbf{C}$. We see immediately that there are a number of difficulties with this definition. In particular, if $\mathbf{A}$ is not perpendicular to $\mathbf{B}$, the vector $\mathbf{C}$ does not exist. Moreover, if $\mathbf{A}$ is perpendicular to $\mathbf{B}$ then, there are an infinite number of vectors that satisfy $\mathbf{A} = \mathbf{B} \times \mathbf{C}$. To see that, let us assume that $\mathbf{C}$ satisfies, $\mathbf{A} = \mathbf{B} \times \mathbf{C}$. Then, any vector $\mathbf{D} = \mathbf{C} + \beta\mathbf{B}$, for

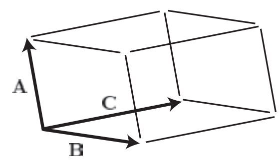
any scalar β, also satisfies $\mathbf{A} = \mathbf{B} \times \mathbf{D}$, since $\mathbf{B} \times \mathbf{D} = \mathbf{B} \times (\mathbf{C} + \beta\mathbf{B}) = \mathbf{B} \times \mathbf{C} = \mathbf{A}$. We conclude therefore, that vector division is not a well defined operation.

Exercise

Show that $|\boldsymbol{A} \times \boldsymbol{B}|$ is the area of the parallelogram defined by the vectors $\boldsymbol{A}$ and $\boldsymbol{B}$, when drawn with a common origin.

Triple product

Given three vectors $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$, the triple product is a scalar given by $\boldsymbol{A} \cdot (\boldsymbol{B} \times \boldsymbol{C})$. Geometrically, the triple product can be interpreted as the volume of the three dimensional parallelepiped defined by the three vectors $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$.



It can be easily verified that $\boldsymbol{A} \cdot (\boldsymbol{B} \times \boldsymbol{C}) = \boldsymbol{B} \cdot (\boldsymbol{C} \times \boldsymbol{A}) = \boldsymbol{C} \cdot (\boldsymbol{A} \times \boldsymbol{B})$.

Exercise

Show that $\boldsymbol{A} \cdot (\boldsymbol{B} \times \boldsymbol{C})$ is the volume of the parallelepiped defined by the vectors $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$, when drawn with a common origin.

Double vector product

The double vector product results from repetition of the cross product operation. A useful identity here is,

$$\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = (\boldsymbol{A} \cdot \boldsymbol{C})\boldsymbol{B} - (\boldsymbol{A} \cdot \boldsymbol{B})\boldsymbol{C} \ .$$

Using this identity we can easily verify that the double cross product is not associative, that is,

$$\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) \neq (\boldsymbol{A} \times \boldsymbol{B}) \times \boldsymbol{C} \ .$$

Vector Calculus

Vector differentiation and integration follow standard rules. Thus if a vector is a function of, say time, then its derivative with respect to time is also a vector. Similarly the integral of a vector is also a vector.

Derivative of a vector

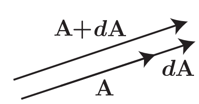
Consider a vector $\boldsymbol{A}(t)$ which is a function of, say, time. The derivative of $\boldsymbol{A}$ with respect to time is defined as,

$$\frac{d\boldsymbol{A}}{dt} = \lim_{\Delta t \to 0} \frac{\boldsymbol{A}(t + \Delta t) - \boldsymbol{A}(t)}{\Delta t}. \tag{1}$$

A vector has magnitude and direction, and it changes whenever either of them changes. Therefore the rate of change of a vector will be equal to the sum of the changes due to magnitude and direction.

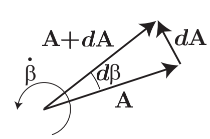
Rate of change due to magnitude changes

When a vector only changes in magnitude from $\boldsymbol{A}$ to $\boldsymbol{A}+d\boldsymbol{A}$, the rate of change vector $d\boldsymbol{A}$ is clearly parallel to the original vector $\boldsymbol{A}$.



Rate of change due to direction changes

Let us look at the situation where only the direction of the vector changes, while the magnitude stays constant. This is illustrated in the figure where a vector $\boldsymbol{A}$ undergoes a small rotation. From the sketch, it is clear that if the magnitude of the vector does not change, $d\boldsymbol{A}$ is perpendicular to $\boldsymbol{A}$ and as a consequence, the derivative of $\boldsymbol{A}$, must be perpendicular to $\boldsymbol{A}$. (*Note that in the picture $d\boldsymbol{A}$ has a finite magnitude and therefore, $\boldsymbol{A}$ and $d\boldsymbol{A}$ are not exactly perpendicular. In reality, $d\boldsymbol{A}$ has infinitesimal length and we can see that when the magnitude of $d\boldsymbol{A}$ tends to zero, $\boldsymbol{A}$ and $d\boldsymbol{A}$ are indeed perpendicular*).



An alternative, more mathematical, explanation can be derived by realizing that even if $\boldsymbol{A}$ changes but its modulus stays constant, then the dot product of $\boldsymbol{A}$ with itself is a constant and its derivative is therefore zero. $\boldsymbol{A} \cdot \boldsymbol{A} = constant$. Differentiating, we have that,

$$d\boldsymbol{A} \cdot \boldsymbol{A} + \boldsymbol{A} \cdot d\boldsymbol{A} = 2\boldsymbol{A} \cdot d\boldsymbol{A} = 0 \ ,$$

which shows that $\boldsymbol{A}$, and $d\boldsymbol{A}$, must be orthogonal.

Suppose that $\boldsymbol{A}$ is instantaneously rotating in the plane of the paper at a rate $\dot{\beta} = d\beta/dt$, with no change in magnitude. In an instant dt, $\boldsymbol{A}$, will rotate an amount $d\beta = \dot{\beta}dt$ and the magnitude of $d\boldsymbol{A}$, will be

$$dA = |d\boldsymbol{A}| = Ad\beta = A\dot{\beta}dt \ .$$

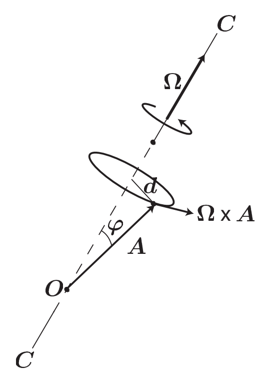
Hence, the magnitude of the vector derivative is

$$\left|\frac{d\boldsymbol{A}}{dt}\right| = A\dot{\beta} \ .$$

In the general three dimensional case, the situation is a little bit more complicated because the rotation of the vector may occur around a general axis. If we express the instantaneous rotation of $\boldsymbol{A}$ in terms of an angular velocity $\boldsymbol{\Omega}$ (recall that the angular velocity vector is aligned with the axis of rotation and the direction of the rotation is determined by the right hand rule), then the derivative of $\boldsymbol{A}$ with respect to time is simply,

$$\left(\frac{d\boldsymbol{A}}{dt}\right)_{\text{constant magnitude}} = \boldsymbol{\Omega} \times \boldsymbol{A} \ . \tag{2}$$

To see that, consider a vector $\boldsymbol{A}$ rotating about the axis $C - C$ with an angular velocity $\boldsymbol{\Omega}$. The derivative will be the velocity of the tip of $\boldsymbol{A}$. Its magnitude is given by $l\Omega$, and its direction is both perpendicular to $\boldsymbol{A}$ and to the axis of rotation. We note that $\boldsymbol{\Omega} \times \boldsymbol{A}$ has the right direction, and the right magnitude since $l = A\sin\varphi$.

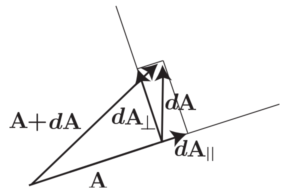


Expression (2) is also valid in the more general case where $\boldsymbol{A}$ is rotating about an axis which does not pass through the origin of $\boldsymbol{A}$. We will see in the course, that a rotation about an arbitrary axis can always be written as a rotation about a parallel axis plus a translation, and translations do not affect the magnitude not the direction of a vector.

We can now go back to the general expression for the derivative of a vector (1) and write

$$\frac{d\boldsymbol{A}}{dt} = \left(\frac{d\boldsymbol{A}}{dt}\right)_{\text{constant direction}} + \left(\frac{d\boldsymbol{A}}{dt}\right)_{\text{constant magnitude}} = \left(\frac{d\boldsymbol{A}}{dt}\right)_{\text{constant direction}} + \boldsymbol{\Omega} \times \boldsymbol{A} \ .$$

Note that $(d\boldsymbol{A}/dt)_{\text{constant direction}}$ is parallel to $\boldsymbol{A}$ and $\boldsymbol{\Omega} \times \boldsymbol{A}$ is orthogonal to $\boldsymbol{A}$. The figure below shows the general differential of a vector, which has a component which is parallel to $\boldsymbol{A}$, $d\boldsymbol{A}_{\|}$, and a component which is orthogonal to $\boldsymbol{A}$, $d\boldsymbol{A}_{\perp}$. The magnitude change is given by $d\boldsymbol{A}_{\|}$, and the direction change is given by $d\boldsymbol{A}_{\perp}$.



Rules for Vector Differentiation

Vector differentiation follows similar rules to scalars regarding vector addition, multiplication by a scalar, and products. In particular we have that, for any vectors $\boldsymbol{A}$, $\boldsymbol{B}$, and any scalar α,

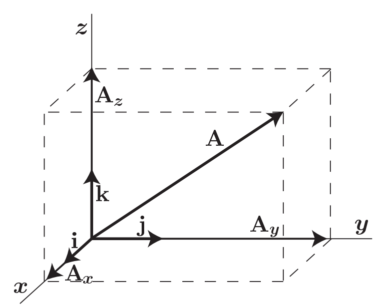
$$\begin{aligned}
d(\alpha\boldsymbol{A}) &= d\alpha\boldsymbol{A} + \alpha d\boldsymbol{A} \\
d(\boldsymbol{A}+\boldsymbol{B}) &= d\boldsymbol{A} + d\boldsymbol{B} \\
d(\boldsymbol{A}\cdot\boldsymbol{B}) &= d\boldsymbol{A}\cdot\boldsymbol{B} + \boldsymbol{A}\cdot d\boldsymbol{B} \\
d(\boldsymbol{A}\times\boldsymbol{B}) &= d\boldsymbol{A}\times\boldsymbol{B} + \boldsymbol{A}\times d\boldsymbol{B} \quad .
\end{aligned}$$

Components of a Vector

We have seen above that it is possible to define several operations involving vectors without ever introducing a reference frame. This is a rather important concept which explains why vectors and vector equations are so useful to express physical laws, since these, must be obviously independent of any particular frame of reference.

In practice however, reference frames need to be introduced at some point in order to express, or measure, the direction and magnitude of vectors, i.e. we can easily measure the direction of a vector by measuring the angle that the vector makes with the local vertical and the geographic north.

Consider a *right-handed* set of axes xyz, defined by three mutually orthogonal unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ ($\boldsymbol{i}\times\boldsymbol{j}=\boldsymbol{k}$) (note that here we are not using the hat ($\hat{}$) notation). Since the vectors $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ are mutually orthogonal they form a basis. The projections of $\boldsymbol{A}$ along the three xyz axes are the *components* of $\boldsymbol{A}$ in the xyz reference frame.



In order to determine the components of $\boldsymbol{A}$, we can use the scalar product and write,

$$A_x = \boldsymbol{A}\cdot\boldsymbol{i}, \quad A_y = \boldsymbol{A}\cdot\boldsymbol{j}, \quad A_z = \boldsymbol{A}\cdot\boldsymbol{k} \ .$$

The vector $\boldsymbol{A}$, can thus be written as a sum of the three vectors along the coordinate axis which have magnitudes A_x, A_y, and A_z and using matrix notation, as a column vector containing the component magnitudes.

$$\boldsymbol{A} = \boldsymbol{A}_x + \boldsymbol{A}_y + \boldsymbol{A}_z = A_x\boldsymbol{i} + A_y\boldsymbol{j} + A_z\boldsymbol{k} = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} .$$

Vector operations in component form

The vector operations introduced above can be expressed in terms of the vector components in a rather straightforward manner. For instance, when we say that $\boldsymbol{A} = \boldsymbol{B}$, this implies that the projections of $\boldsymbol{A}$ and $\boldsymbol{B}$ along the xyz axes are the same, and therefore, this is equivalent to three scalar equations e.g. $A_x = B_x$, $A_y = B_y$, and $A_z = B_z$. Regarding vector summation, subtraction and multiplication by a scalar, we have that, if $\boldsymbol{C} = \alpha\boldsymbol{A} + \beta\boldsymbol{B}$, then,

$$C_x = \alpha A_x + \beta B_x, \qquad C_y = \alpha A_y + \beta B_y, \qquad C_z = \alpha A_z + \beta B_z .$$

Scalar product

Since $\boldsymbol{i}\cdot\boldsymbol{i} = \boldsymbol{j}\cdot\boldsymbol{j} = \boldsymbol{k}\cdot\boldsymbol{k} = 1$ and that $\boldsymbol{i}\cdot\boldsymbol{j} = \boldsymbol{j}\cdot\boldsymbol{k} = \boldsymbol{k}\cdot\boldsymbol{i} = 0$, the scalar product of two vectors can be written as,

$$\boldsymbol{A}\cdot\boldsymbol{B} = (A_x\boldsymbol{i} + A_y\boldsymbol{j} + A_z\boldsymbol{k})\cdot(B_x\boldsymbol{i} + B_y\boldsymbol{j} + B_z\boldsymbol{k}) = A_xB_x + A_yB_y + A_zB_z .$$

Note that, $\boldsymbol{A}\cdot\boldsymbol{A} = A^2 = A_x^2 + A_y^2 + A_z^2$, which is consistent with Pythagoras' theorem.

Vector product

Here, $\boldsymbol{i}\times\boldsymbol{i} = \boldsymbol{j}\times\boldsymbol{j} = \boldsymbol{k}\times\boldsymbol{k} = \boldsymbol{0}$ and $\boldsymbol{i}\times\boldsymbol{j} = \boldsymbol{k}$, $\boldsymbol{j}\times\boldsymbol{k} = \boldsymbol{i}$, and $\boldsymbol{k}\times\boldsymbol{i} = \boldsymbol{j}$. Thus,

$$\begin{aligned} \boldsymbol{A}\times\boldsymbol{B} &= (A_x\boldsymbol{i} + A_y\boldsymbol{j} + A_z\boldsymbol{k})\times(B_x\boldsymbol{i} + B_y\boldsymbol{j} + B_z\boldsymbol{k}) \\ &= (A_yB_z - A_zB_y)\boldsymbol{i} + (A_zB_x - A_xB_z)\boldsymbol{j} + (A_xB_y - A_yB_x)\boldsymbol{k} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} . \end{aligned}$$

Triple product

The triple product $\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{C})$ can be expressed as the following determinant

$$\boldsymbol{A}\cdot(\boldsymbol{B}\times\boldsymbol{C}) = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix} ,$$

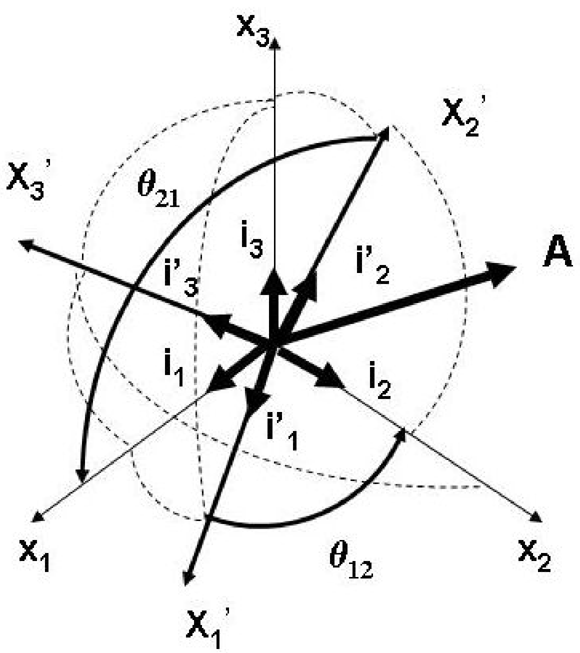
which clearly is equal to zero whenever the vectors are linearly dependent (if the three vectors are linearly dependent they must be co-planar and therefore the parallelepiped defined by the three vectors has zero volume).

Vector Transformations

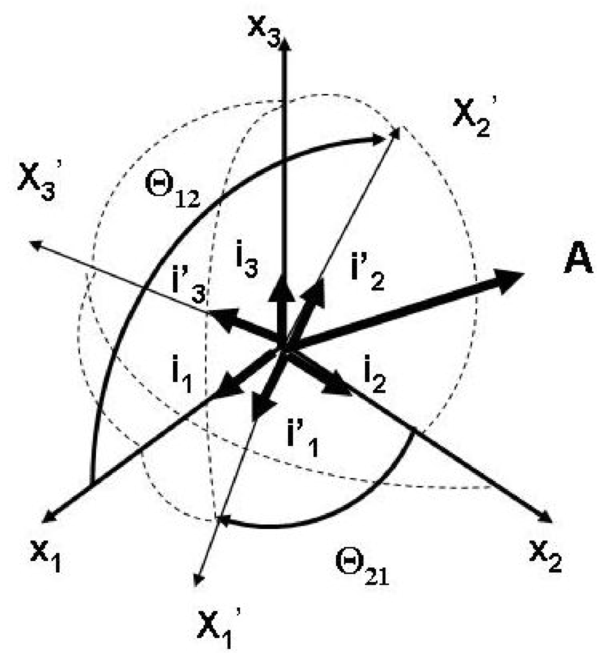
In many problems we will need to use different coordinate systems in order to describe different vector quantities. The above operations, written in component form, only make sense once all the vectors involved are described with respect to the same frame. In this section, we will see how the components of a vector are transformed when we change the reference frame.

Consider two different orthogonal, right-hand sided, reference frames x_1, x_2, x_3 and X_1', X_2', X_3'. A vector $\mathbf{A}$ in coordinate system $\mathbf{x}$ can be transformed to coordinate system $\mathbf{X'}$ by considering the 9 angles that define the relationships between the two systems. (Only three of these angles are independent, a point we shall return to later.)

Referring to a) in the figure we see the vector $\mathbf{A}$, the $\mathbf{x}$ and $\mathbf{X'}$ coordinate systems, the unit vectors $\mathbf{i_1}, \mathbf{i_2}, \mathbf{i_3}$ of the $\mathbf{x}$ system and the unit vectors $\mathbf{i_1'}, \mathbf{i_2'}, \mathbf{i_3'}$ of the $\mathbf{X'}$ system; a) focuses on the transformation of coordinates from $\mathbf{x}$ to $\mathbf{X'}$ while b) focuses on the "reverse" transformation from $\mathbf{X'}$ to $\mathbf{x}$.



a) x transformed into X'



b) X' transformed into x

In the **x** coordinate system, the vector $\boldsymbol{A}$, can be written as

$$\boldsymbol{A} = A_1\boldsymbol{i}_1 + A_2\boldsymbol{i}_2 + A_3\boldsymbol{i}_3, \tag{3}$$

or, when referred to the frame **X'**, as

$$\boldsymbol{A} = A'_1\boldsymbol{i}'_1 + A'_2\boldsymbol{i}'_2 + A'_3\boldsymbol{i}'_3. \tag{4}$$

Since the vector **A** remains the same regardless of our coordinate transformation

$$\boldsymbol{A} = A_1\boldsymbol{i}_1 + A_2\boldsymbol{i}_2 + A_3\boldsymbol{i}_3 = A'_1\boldsymbol{i}'_1 + A'_2\boldsymbol{i}'_2 + A'_3\boldsymbol{i}'_3, \tag{5}$$

We can find the components of the vector **A** in the transformed system in term of the components of **A** in the original system by simply taking the dot product of this equation with the desired unit vector $\boldsymbol{i}'_j$ in the **X'** system so that

$$A'_j = A_1\boldsymbol{i}'_j \cdot \boldsymbol{i}_1 + A_2\boldsymbol{i}'_j \cdot \boldsymbol{i}_2 + A_3\boldsymbol{i}'_j \cdot \boldsymbol{i}_3 \tag{6}$$

where A'_j is the jth component of **A** in the **X'** system. Repeating this operation for each component of **A** in the **X'** system results in the matrix form for **A**

$$\begin{pmatrix} A'_1 \\ A'_2 \\ A'_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{i}'_1 \cdot \boldsymbol{i}_1 & \boldsymbol{i}'_1 \cdot \boldsymbol{i}_2 & \boldsymbol{i}'_1 \cdot \boldsymbol{i}_3 \\ \boldsymbol{i}'_2 \cdot \boldsymbol{i}_1 & \boldsymbol{i}'_2 \cdot \boldsymbol{i}_2 & \boldsymbol{i}'_2 \cdot \boldsymbol{i}_3 \\ \boldsymbol{i}'_3 \cdot \boldsymbol{i}_1 & \boldsymbol{i}'_3 \cdot \boldsymbol{i}_2 & \boldsymbol{i}'_3 \cdot \boldsymbol{i}_3 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix}.$$

The above expression is the relationship that expresses how the components of a vector in one coordinate system relate to the components of the ***same*** vector in a different coordinate system.

Referring to the figure, we see that $\boldsymbol{i}'_j{\cdot}\boldsymbol{i}_i$ is equal to the cosine of the angle between $\boldsymbol{i'_j}$ and $\boldsymbol{i_i}$ which is $\theta_{j'i}$; in particular we see that $\boldsymbol{i}'_2 \cdot \boldsymbol{i}_1 = cos\theta_{21}$ while $\boldsymbol{i}'_1 \cdot \boldsymbol{i}_2 = cos\theta_{12}$; these angles are in general not equal. Therefore, the components of the vector **A** are transformed from the **x** coordinate system to the **X'** system through the transformation

$$A'_j = A_1 \cos\theta_{j1} \ + A_2 \cos\theta_{j2} \ + A_3 \cos\theta_{j3} \quad . \tag{7}$$

where the coefficients relating the components of **A** in the two coordinate systems are the various direction cosines of the angles between the coordinate directions.

The above relations for the transformation of **A** from the **x** to the **X'** system can be written in matrix form as

$$\mathbf{A}' = \begin{pmatrix} A'_1 \\ A'_2 \\ A'_3 \end{pmatrix} = \begin{pmatrix} \cos(\theta_{11}) & \cos(\theta_{12}) & \cos(\theta_{13}) \\ \cos(\theta_{21}) & \cos(\theta_{22}) & \cos(\theta_{23}) \\ \cos(\theta_{31}) & \cos(\theta_{32}) & \cos(\theta_{33}) \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix}. \tag{8}$$

We use the symbol **A'** to denote the components of the vector **A** in the ' system. Of course the vector **A** is unchanged by the transformation. We introduce the symbol $[T]$ for the transformation matrix from **x** to **X'**.

This relationship, which expresses how the components of a vector in one coordinate system relate to the components of the **same** vector in a different coordinate system, is then written

$$\mathbf{A'} = [T]\mathbf{A}. \tag{9}$$

where $[T]$ is the transformation matrix.

We now consider the process that transforms the vector **A'** from the **X'** system to the **x** system.

$$\mathbf{A} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \end{pmatrix} = \begin{pmatrix} \cos(\Theta_{11}) & \cos(\Theta_{12}) & \cos(\Theta_{13}) \\ \cos(\Theta_{21}) & \cos(\Theta_{22}) & \cos(\Theta_{23}) \\ \cos(\Theta_{31}) & \cos(\Theta_{32}) & \cos(\Theta_{33}) \end{pmatrix} \begin{pmatrix} A'_1 \\ A'_2 \\ A'_3 \end{pmatrix}. \tag{10}$$

By comparing the two coordinate transformations shown in a) and b), we see that $\cos(\theta_{12})$=$\cos(\Theta_{21})$, and that therefore the matrix element of magnitude $\cos(\theta_{12})$ which appears in the 12 position in the transformation matrix from **x** to **X'** now appears in the 21 position in the matrix which transforms **A** from **X'** to **x**. This pattern is repeated for all off-diagonal elements. The diagonal elements remain unchanged since $\cos(\theta_{ii})$=$\cos(\Theta_{ii})$. Thus the matrix which transposes the vector **A** in the **X'**system back to the **x** system is the transpose of the original transformation matrix,

$$\mathbf{A} = [T]^T\mathbf{A'}. \tag{11}$$

where $[T]^T$ is the transpose of $[T]$. (A transpose matrix has the rows and columns reversed.)

Since transforming **A** from **x** to **X'** and back to **x** results in no change, the matrix $[T]^T$ is also $[T]^{-1}$ the inverse of $[T]$ since

$$\mathbf{A} = [T]^T[T]\mathbf{A} = [T]^{-1}[T]\mathbf{A} = [\mathbf{I}]\mathbf{A} = \mathbf{A}. \tag{12}$$

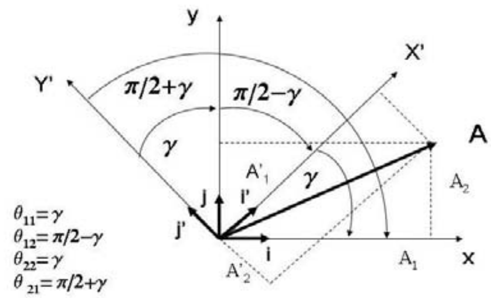
where $[I]$ is the identity matrix

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{13}$$

This is a remarkable and useful property of the transformation matrix, which is not true in general for any matrix.

Example ***Coordinate transformation in two dimensions***

Here, we apply for illustration purposes, the above expressions to a two-dimensional example. Consider the change of coordinates between two reference frames xy, and $x'y'$, as shown in the diagram.



The angle between $\boldsymbol{i}$ and $\boldsymbol{i}'$ is γ. Therefore, $\boldsymbol{i}\cdot\boldsymbol{i}' = \cos\gamma$. Similarly, $\boldsymbol{j}\cdot\boldsymbol{i}' = \cos(\pi/2-\gamma) = \sin\gamma$, $\boldsymbol{i}\cdot\boldsymbol{j}' = \cos(\pi/2+\gamma) = -\sin\gamma$, and $\boldsymbol{j}\cdot\boldsymbol{j}' = \cos\gamma$. Finally, the transformation matrix $[T]$ is

$$[T] = \begin{pmatrix} cos\,(\theta_{11}) & \cos\,(\theta_{12}) \\ \cos\,(\theta_{21}) & \cos\,(\theta_{22}) \end{pmatrix} = \begin{pmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{pmatrix},$$

and we can write,

$$\begin{pmatrix} A_1' \\ A_2' \end{pmatrix} = [T] \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}.$$

and

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = [T]^T \begin{pmatrix} A_1' \\ A_2' \end{pmatrix}.$$

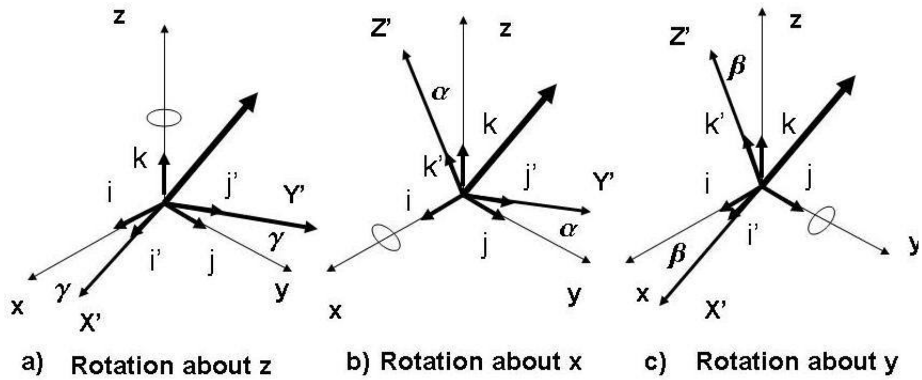
Therefore,

$$A_1' = A_1\cos\gamma + A_2\sin\gamma \tag{14}$$

$$A_2' = -A_1\sin\gamma + A_2\cos\gamma. \tag{15}$$

For instance, we can easily check that when $\gamma = \pi/2$, the above expressions give $A_1' = A_2$, and $A_2' = -A_1$, as expected.

An additional observation can be made. If in three dimensions, we rotate the x, y, z coordinate system about the z axis, as shown in **a)** leaving the z component unchanged,



the transformation matrix becomes

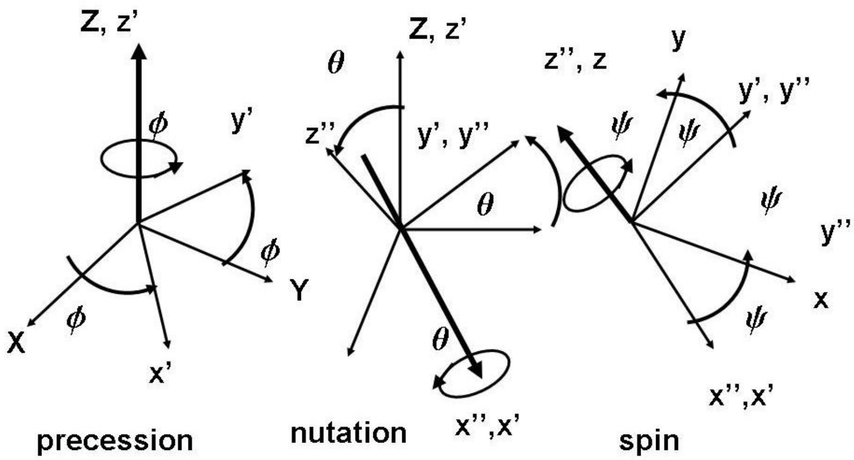
$$[T] = \begin{pmatrix} cos\,(\theta_{11}) & \cos(\theta_{12}) & 0 \\ \cos(\theta_{21}) & \cos(\theta_{22}) & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Analogous results can be obtained for rotation about the x axis or rotation about the y axis as shown in **b)** and **c)**.

Sequential Transformations; Euler Angles

The general orientation of a coordinate system can be described by a sequence of rotations about coordinate axis. One particular set of such rotations leads to a description particularly convenient for describing the motion of a three-dimensional rigid body in general spinning motion, call Euler angles. We shall treat this topic in Lecture 28. For now, we examine how this rotation fits into our general study of coordinate transformations. A coordinate description in terms of Euler angles is obtained by the sequential rotation of axis as shown in the figure; the order of transformation makes a difference.

To develop the description of this motion, we use a series of transformations of coordinates. The final result is shown below. This is the coordinate system used for the description of motion of a general three-dimensional rigid body such as a top described in body-fixed axis. To identify the new position of the coordinate axes as a result of angular displacement through the three Euler angles, we go through a series of coordinate rotations.



First, we rotate from an initial X, Y, Z system into an x', y', z' system through a rotation ϕ about the Z, z' axis.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} cos\phi & sin\phi & 0 \\ -sin\phi & cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = [T_1] \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}.$$

The resulting x', y' coordinates remain in the X, Y plane. Then, we rotate about the x' axis into the x'', y'', z'' system through an angle θ. The x'' axis remains coincident with the x' axis. The axis of rotation for this transformation is called the "line of nodes". The plane containing the x'', y'' coordinate is now tipped through an angle θ relative to the original X, Y plane. coordinates

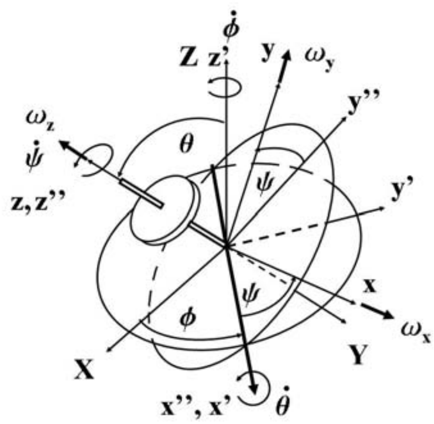
$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\theta & sin\theta \\ 0 & -sin\theta & cos\theta \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = [T_2] \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}.$$

And finally, we rotate about the z'', z system through an angle ψ into the x, y, z system. The z'' axis is called the spin axis. It is coincident with the z axis.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} cos\psi & sin\psi & 0 \\ -sin\psi & cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = [T_3] \begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix}.$$

The final coordinate system used to describe the position of the body is shown below. The angle ψ is called the spin; the angle ϕ is called the precession; the angle θ is called the nutation. The total transformation is given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = [T_3][T_2][T_1] \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}.$$



Euler angles are not always defined in exactly this manner, either the notation or the order of rotations can differ. The particular transformation used in any example should be clearly described.

References

[1] J.B. Marion and S.T. Thornton, *Classical Dynamics of Particles and Systems*, Harcourt Brace, 1995.

[2] D. Kleppner and R.J. Kolenkow, *An Introduction to Mechnics*, McGraw Hill, 1973.