Chapter 1
Matrices, Vectors, and Vector Calculus

In this chapter, we will focus on the mathematical tools required for the course. The main concepts that will be covered are:

- Coordinate transformations
- Matrix operations
- Scalars and vectors
- Vector calculus
- Differentiation and integration

Coordinate transformation

In order to be able to specify the position of a point P we first must specify the coordinate system that will be used. The coordinates of point P will be a function of the coordinate system being used, and coordinate transformations allow us to define the relation between the coordinates of point P in different coordinate systems.

Different types of coordinate systems are used for different applications. The most commonly used coordinate systems are:

- **Cartesian coordinate systems**. These systems consist out of three perpendicular coordinate axes, called the x, y, and z (or x_1, x_2, and x_3) axes. The coordinates of a point P are usually specified by three coordinates (x, y, z) or (x_1, x_2, x_3). Note: the coordinate axes in a Cartesian coordinate system are usually independent of time.

- **Spherical coordinate systems**. Spherical coordinate systems are most often used when the system under consideration has spherical symmetry. The origin of the coordinate system is chosen to coincide with the point of spherical symmetry. The position of a point P is determined by specifying the distance r from the origin of the coordinate system and the polar and azimuthal angles θ and ϕ. Note: we still need to define a Cartesian coordinate system to define the origin and the two angles. Note 2: the unit vectors associated with the position vector and the angles will be a function of time if the position described is time dependent.

- **Cylindrical coordinate systems**. Cylindrical coordinate systems are most often used when the system under consideration has cylindrical symmetry. The coordinate system is characterized by the axis of cylindrical symmetry, which is usually called the z axis. The position of a point P is determine by specifying the distance r to the z axis, the azimuthal angle ϕ, and the z coordinate. Note: we still need to define a Cartesian coordinate system to define the origin and the two angles. Note 2: the unit vectors associated with the position vector and the azimuthal angle will be a function of time if the position described is time dependent.

A coordinate system is not uniquely defined. For example, the coordinates of a point P in a Cartesian coordinate system depend on the choice of coordinate axes. Different choices will result in different coordinates. **Coordinate transformations** are used to transform the coordinates between coordinate systems.

There are a number of different types of coordinate transformations we will encounter in this course: translation, rotation, and the standard Lorentz transformation. In this chapter we will only focus on the rotational transformation.

Consider two coordinate systems, related to each other via a transformation around the x_3 axis, as shown in Figure 1.

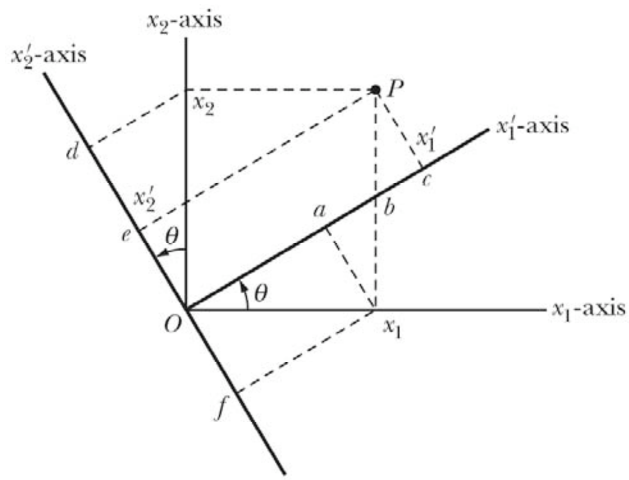

Figure 1. Two different coordinate systems used to represent the position of *P*.

The relation between the coordinates of P in the two coordinate systems can be written as

$$x_i' = \sum_{j=1}^{3} \lambda_{ij} x_i$$

where

$$\lambda_{ij} = \cos\left(x_i', x_j\right)$$

is the cosine of the angle between the x_i'-axis and the x_j-axis (also called the **direction cosine**). The coordinate transformation is most often written in matrix notation:

$$\bar{x}' = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix} \bar{x}$$

Consider the two coordinate systems shown in Figure 1. Since the two coordinate systems are related to each other via a rotation around the x_3-axis we conclude that:

- The direction cosine between the x_3'-axis and the x_1- and x_2-axes will be zero (angle is 90°). Thus: $\lambda_{31} = \lambda_{32} = 0$.
- The direction cosine between the x_3'-axis and the x_3-axis will be one (angle is 0°). Thus $\lambda_{33} = 1$.
- The direction cosine between the x_1'- and x_2'-axes and the x_3-axis will be zero (angle is 90°). Thus: $\lambda_{13} = \lambda_{23} = 0$.
- The direction cosines between the x_1'- and the x_2-axis is $\cos(\pi/2-\theta)$ and the direction cosine between the x_2'- and the x_1-axis is $\cos(\pi/2+\theta)$. Thus: $\lambda_{12} = \cos(\pi/2-\theta)$ and $\lambda_{21} = \cos(\pi/2+\theta)$.
- The direction cosines between the x_1'- and the x_1-axes and between the x_2'- and the x_2-axes is $\cos(\theta)$. Thus: $\lambda_{11} = \lambda_{22} = \cos(\theta)$.

The coordinate transformation for this particular transformation is thus given by

$$\bar{x}' = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \bar{x}$$

Although each rotation matrix has 9 parameters, only 3 are truly independent. This can be seen by realizing that in order to specify a rotation we need to specify the rotation axis (which requires the specification of a polar and azimuthal angle) and the rotation angle. Consider some of the relations we can determine between the parameters of the rotation matrix:

- The rotation preserves the length of a vector. Rotating a unit vector will produce another unit vector. This requires that

$$\sqrt{\lambda_{11}^2 + \lambda_{21}^2 + \lambda_{31}^2} = 1$$

$$\sqrt{\lambda_{12}^2 + \lambda_{22}^2 + \lambda_{32}^2} = 1$$

$$\sqrt{\lambda_{13}^2 + \lambda_{23}^2 + \lambda_{33}^2} = 1$$

- The rotation preserves the angle between two vectors. Since the angle between the coordinate axes is 90°, the angle between these axes after transformation must also be 90°. This requires that

$$\lambda_{11}\lambda_{12} + \lambda_{21}\lambda_{22} + \lambda_{31}\lambda_{32} = 0$$

$$\lambda_{11}\lambda_{13} + \lambda_{21}\lambda_{23} + \lambda_{31}\lambda_{33} = 0$$

$$\lambda_{12}\lambda_{13} + \lambda_{22}\lambda_{23} + \lambda_{32}\lambda_{33} = 0$$

These six equations can be combined in the following manner

$$\sum_j \lambda_{ij}\lambda_{kj} = \delta_{ik}$$

This equation is called the orthogonal condition, and is only satisfied if the coordinate axes are mutually perpendicular.

Matrix operations

When we are use the rotation matrix to carry out coordinate transformations we carry out a matrix operation. There are several important facts to remember about matrix operations:

- The unit matrix is defined as a matrix for which the diagonal values are 1 and the non-diagonal values are 0. When the unit vector operates on a vector, the result is the same vector.
- The inverse of a matrix is defined such that when it operates on the original matrix, the result is the unit matrix.
- A transposed matrix is derived from the original matrix by interchanging the rows and columns.
- When we combine coordinate transformations, we can obtain the resulting transformation by multiplying the rotation matrices.
- When we multiply rotation matrices we must realize that the order of the multiplication matters. See for example Figure 2.

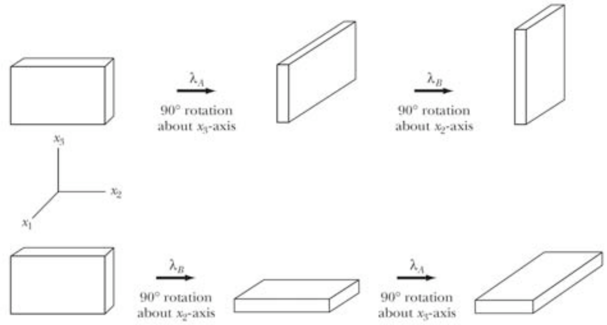

Figure 2. The order of transformation matters.

- A matrix inversion is a transformation that results in the reflection through the origin of all axes. We can not find any series of rotations that result in an inversion. Matrix inversions are examples of the so-called **improper rotations**, and are characterized by a matrix with a determinant equal to -1. **Proper rotations** are those rotations that are characterized by a matrix with a determinant equal to +1.

<u>Vectors and Scalars</u>

Consider the following coordinate transformation

$$x_i' = \sum_{j=1}^{3} \lambda_{ij} x_i$$

where

$$\sum_{j=1}^{3} \lambda_{ij} \lambda_{kj} = \delta_{ik}$$

Vectors and scalars are defined based on what happens as a result of such coordinate transformations:

- If a quantity is unaffected by this transformation, it is called a **scalar**.
- If a set of three quantities transforms in the same manner as the coordinates of a point *P*, these quantities are the components of what we call a **vector**.

Note:

- These definitions define scalars and vectors in a very different way from what you may have been used to. For example, the current definition of a vector makes no reference to the geometrical interpretation the vector.
- You will encounter parameters in Physics that look like vectors, but do not transform like vectors under certain operations (a pseudo vector is an example). Just having a magnitude and a direction is not sufficient to define a vector!

Vector and scalar addition and multiplication have a number of properties in common:

- Both satisfy the commutative law: the order of addition does not change the final result.
- Both satisfy the associative law: when we determine the sum of more than two vectors or scalars, the final results will not depend on which pair of vectors or scalars we add first.
- The product of a scalar with a scalar transforms like a scalar (**and is thus a scalar**).
- The product of a vector with a scalar transforms like a vector (**and is thus a vector**).

The following operations are unique to vectors and do not have equivalents for scalars:

- **Scalar product**. The scalar product between two vectors ***A*** and ***B*** is defined as

$$\bar{A} \bullet \bar{B} = \sum_i A_i B_i$$

It can be shown that the scalar product as defined above, is equal to the product of the magnitudes of the two vectors and the direction cosine between them:

$$\bar{A} \bullet \bar{B} = \sum_i A_i B_i = AB\cos(\bar{A}, \bar{B})$$

In order to show that the scalar product behaves like a scalar, we must thus show that the scalar product between ***A*** and ***B*** is the same as the scalar product between ***A'*** and ***B'***.
The scalar product also satisfies the **commutative** and the **distributive** laws:

$$\bar{A} \bullet \bar{B} = \bar{B} \bullet \bar{A}$$

$$\bar{A} \bullet \left(\bar{B} + \bar{C}\right) = \left(\bar{A} \bullet \bar{B}\right) + \left(\bar{A} \bullet \bar{C}\right)$$

- **Vector product**. The vector product between two vectors ***A*** and ***B*** is a third vector ***C***, defined as

$$\bar{C} = \bar{A} \times \bar{B} = \begin{vmatrix} \hat{x}_1 & \hat{x}_2 & \hat{x}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix}$$

In order to show that the vector product behaves like a vector, we must show that the vector product transforms like a vector under a coordinate transformation.
The geometrical interpretation of the vector product is shown in Figure 3. The magnitude of the vector product is the area of the parallelogram defined by the vectors ***A*** and ***B*** and it is directed in a direction perpendicular to the plane defined by the vectors ***A*** and ***B*** (right-hand rule defines the direction).
There are many properties of the vector product. Some of them are listed here (see the text book for a more complete listing):

$$\bar{A} \times \bar{B} = -\bar{B} \times \bar{A}$$

$$\bar{A} \times \left(\bar{B} \times \bar{C}\right) = \left(\bar{A} \bullet \bar{C}\right)\bar{B} - \left(\bar{A} \bullet \bar{B}\right)\bar{C}$$

$$\bar{A} \bullet \left(\bar{B} \times \bar{C}\right) = \bar{B} \bullet \left(\bar{C} \times \bar{A}\right) = \bar{C} \bullet \left(\bar{A} \times \bar{B}\right)$$

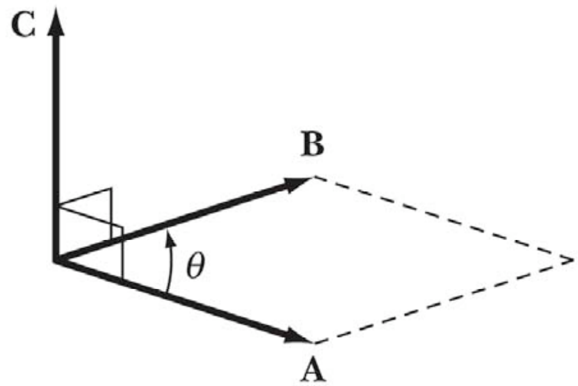

Figure 3. Properties of the vector product between the vectors *A* and *B*.

<u>Differentiation and Integration</u>

Two important operations on both scalars and vectors are differentiation and integration. These operations are used to define important mechanical quantities (such as velocity and acceleration), and a thorough understanding of operations involving differentiation and integration is required in order to succeed in this course.

- **Scalar Differentiation**. We can differentiate vectors and scalars with respect to a scalar variable *s*.
  - The result of the differentiation of a scalar with respect to another scalar variable will be another scalar. The result of the differentiation will be independent of the coordinate system.
  - The result of the differentiation of a vector function with respect to a scalar variable will be another vector. The resulting vector will be directed tangential to the curve that represents the function.

- **Scalar Differentiation in different coordinate systems**. An important scalar variable used in differentiations is the time *t*. Based on the position vector, we can obtain the velocity and acceleration vectors by differentiating the position vectors once and twice, respectively, with respect to time. If the Cartesian coordinates are being used, the axes are independent of time, and differentiation the position vector with respect to time is equivalent to differentiating the individual components with respect to time:

$$\bar{v} = \dot{\bar{r}} = \sum_i \frac{dx_i}{dt}\hat{x}_i$$

$$\bar{a} = \dot{\bar{v}} = \ddot{\bar{r}} = \sum_i \frac{d^2 x_i}{dt^2} \hat{x}_i$$

The situation is more complicated if we are using spherical or cylindrical coordinates. Consider for example the motion of an object shown in Figure 4.

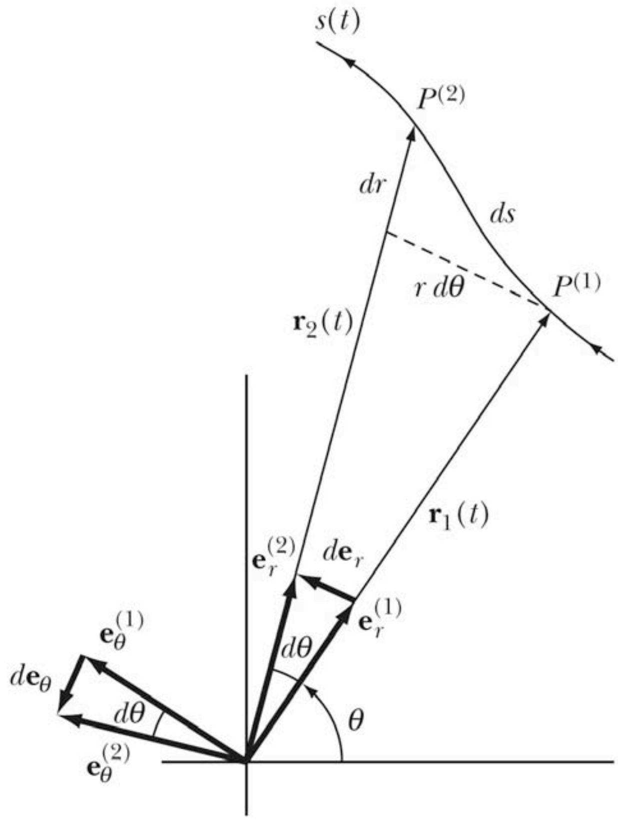

Figure 4. Motion of an object described in terms of spherical coordinates.

When the object moves during a time *dt* from *P*(1) to *P*(2), the spherical unit vectors change too, as shown in Figure 4:

$$d\hat{r} = \hat{r}(t + dt) - \hat{r}(t)$$

$$d\hat{\theta} = \hat{\theta}(t + dt) - \hat{\theta}(t)$$

Based on the definition of the unit vectors in the spherical coordinate systems we can conclude:

$$d\hat{r} = (d\theta)\hat{\theta}$$

$$d\hat{\theta} = -(d\theta)\hat{r}$$

By dividing each side by *dt* we obtain the following relations:

$$\dot{\hat{r}} = \frac{d\hat{r}}{dt} = \left(\frac{d\theta}{dt}\right)\hat{\theta} = \dot{\theta}\hat{\theta}$$

$$\dot{\hat{\theta}} = \frac{d\hat{\theta}}{dt} = -\dot{\theta}\hat{r}$$

Using these relations we can calculate the velocity and acceleration:

$$\bar{v} = \frac{d}{dt}(r\hat{r}) = \frac{dr}{dt}\hat{r} + r\frac{d\theta}{dt}\hat{\theta} = \dot{r}\hat{r} + r\dot{\theta}\hat{\theta}$$

$$\bar{a} = \frac{d}{dt}\left(\dot{r}\hat{r} + r\dot{\theta}\hat{\theta}\right) = \left(\ddot{r} - r\dot{\theta}^2\right)\hat{r} + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\hat{\theta}$$

Other relations for spherical and cylindrical coordinates can be found in the textbook.

- **Vector Differential Operator**. A very important operator in this course will be the gradient operator. It operates on a scalar function and the result of the operation is a vector. Consider a scalar function ϕ that is a function of the Cartesian coordinates. The value of the scalar function at a point *P* in two different coordinate systems must be the same:

$$\phi'(x_1', x_2', x_3') = \phi(x_1, x_2, x_3)$$

The coordinates in the two different coordinate systems are connected to each other via a rotation matrix:

$$x_i' = \sum_{j=1}^{3} \lambda_{ij} x_i$$

When we differentiate the scalar function we find the following relation:

$$\frac{\partial \phi'}{\partial x_i'} = \sum_{j=1}^{3} \lambda_{ij} \frac{\partial \phi}{\partial x_j}$$

As we can see, the components of the differential of the scalar function transform like a vector, and the components can thus be considered the components of a vector we call the **gradient** of a scalar function:

$$grad = \overline{\nabla} = \sum_{j=1}^{3} \hat{x}_j \frac{\partial}{\partial x_j}$$

Other important operators are defined in terms of the gradient operator:

$$\text{grad}\phi = \overline{\nabla}\phi = \sum_{j=1}^{3} \frac{\partial \phi}{\partial x_j} \hat{x}_j$$

$$\text{div}\overline{A} = \overline{\nabla} \bullet \overline{A} = \sum_{j=1}^{3} \frac{\partial A_j}{\partial x_j}$$

$$\text{curl}\overline{A} = \overline{\nabla} \times \overline{A} = \sum_{i,j,k} \varepsilon_{ijk} \frac{\partial A_k}{\partial x_j} \hat{x}$$

$$\nabla^2 \phi = \sum_{j=1}^{3} \frac{\partial^2 \phi}{\partial x_j{}^2}$$

The gradient of a scalar function has the following properties:

- The gradient of a scalar function at a point P is directed normal to the lines or surfaces for which the scalar function is constant.
- The gradient of a scalar function at a point P is directed in the direction of maximum change in the scalar function.

- **Integration**. The opposite of differentiation is integration. Both scalar and vector functions can be integrated, and we can encounter volume, surface, and line integration:
  - **Volume integration of a vector**. When we integrate a vector over a volume, the result is another vector with components obtained by volume integration of the components of the original vector.

$$\int_V \overline{A} dv = \begin{pmatrix} \int_V A_1 dv \\ \int_V A_2 dv \\ \int_V A_3 dv \end{pmatrix}$$

- **Surface integration of a vector**. The surface integral of a vector function is given by the integral of its component perpendicular to the surface. When we integrate a vector over a surface, the result is a scalar.

$$\int_S \bar{A} \bullet d\bar{a} = \int_S \left(\bar{A} \bullet \hat{n}\right) da$$

- **Line integration of a vector.** The line integration of a vector is given by the integral of the component of the vector along the path (does not need to be a straight line).

$$\int_{Line} \bar{A} \bullet d\bar{s} = \int_{Line} \sum_i A_i dx_i$$

Various theorems relate volume, surface, line integrals of vectors. Some of the most important theorems are:

- **Gauss's theorem for volume integrals**:

$$\int_S \bar{A} \bullet d\bar{a} = \int_V \left(\bar{\nabla} \bullet \bar{A}\right) dv$$

Note: the surface integral of a vector function is replaced by the volume integral of a scalar function.

- **Stoke's theorem for line integrals**:

$$\int_{Line} \bar{A} \bullet d\bar{s} = \int_S \left(\bar{\nabla} \times \bar{A}\right) \bullet d\bar{a}$$

Stoke's theorem is most useful if it reducing a two-dimensional surface integral to a one-dimensional line integral.