Vectors & Coordinate Systems

Antoine Lesage-Landry and Francis Dawson

Contents

1 Motivations & Definition

Vectors are defined differently depending on the coordinate system being utilized. In any coordinate system the position vector remains the same although the coordinates used to indicate the endpoint of the vector will be different in a given coordinate system. Typically a coordinate system is chosen in order to make the calculations as simple as possible.

The purpose of this document is for the students to develop the tools that enable them to perform calculations in a different coordinate system and to transfer from one coordinate system to another.

You should read some parts of these notes during the first week of classes (see course description documents). You may wish to be selective in terms of what you read now but you will find this material helpful for the following two reasons: (a) the introductory material is a review of what you learned last year in linear algebra and the second calculus course (b) it provides the background material required to solve problems near the end of this term and for other courses that you may take next term and in the coming years.

There are two geometrical entities which you are familiar with already. A scalar is a entity that only possesses a signed magnitude and describes an amount (a number) or an intensity (representing the magnitude of something associated with a physical entity that can be measured). For example, we use a scalar when we express temperature, $-10°$C, mass, 50kg or voltage, 12 V. Note that a scalar can be positive or negative.

A vector is an entity that possesses both a magnitude and a direction at a position in space. Vectors are an important component of mathematics, physics, and engineering. In a two dimensional (2D) space the vector requires two independent directional components to describe an arbitrary direction in 2D space. In a three dimensional (3D) space, the vector requires three independent directional components to describe an arbitrary direction in 3D space. The vector can be visualized as an arrow pointing in a certain direction where the length gives the magnitude of the vector and the orientation of the arrow the direction. For example, a vector in 3D can be used to express a force. A person jumping up and hitting the ground applies a force of 1000 N downwards when he/she makes impact with the surface.

In defining our vectors we need to first ask in how many dimensions we are making an observation. For example a planar area has a dimension of 2 whereas the world that we live in consists of three dimensions. In representing an arbitrary vector in 2D or 3D we need to create a set of reference vectors from which any vector can be constructed. In 2D, we typically use a vector pointing in the y direction and the x direction to construct an arbitrary vector in 2D. As it turns out using this arrangement is very convenient however there is no reason why vectors with orientations other than x and y cannot be used. In 3D, we use a vector in the x, y and z direction to construct an arbitrary vector. Later in this document we will consider other coordinate systems such as Cartesian coordinates or polar coordinates in two dimensions and cylindrical or spherical coordinates in three dimensions.

Lastly, we define the notion of a scalar field or a vector field by way of an example for each case. We extend the vector definition to multi- variable vector valued functions. In the general case, we refer to these multi-variable vector valued functions as fields.

1.1 Scalar field

A scalar field is a function that for a given set of scalar inputs representing for example a position in space, returns a scalar output. The temperature distribution in a room, e.g. $T(x, y)$ in (1) and represented in Figure 1, is a scalar field where we enter the position in 2D, e.g. $(1, 1)$, and it returns a scalar, the temperature $T(1, 1) = 1.8225$.

$$T(x, y) = \cos\frac{x}{2} + \cos\frac{y}{3} \tag{1}$$

1.2 Vector field

A vector field is a function that returns a vector with respect to a position denoted by scalar coordinate values. For example, the electric field $\mathbf{E}$ in 3D due to a point charge Q given in (2) is a vector field.

$$\mathbf{E}(x, y, z) = \frac{Q}{4\pi\epsilon_0(x^2 + y^2 + z^2)^{3/2}}\left(x\,\hat{\mathbf{\imath}} + y\,\hat{\mathbf{\jmath}} + z\,\hat{\mathbf{k}}\right) \tag{2}$$

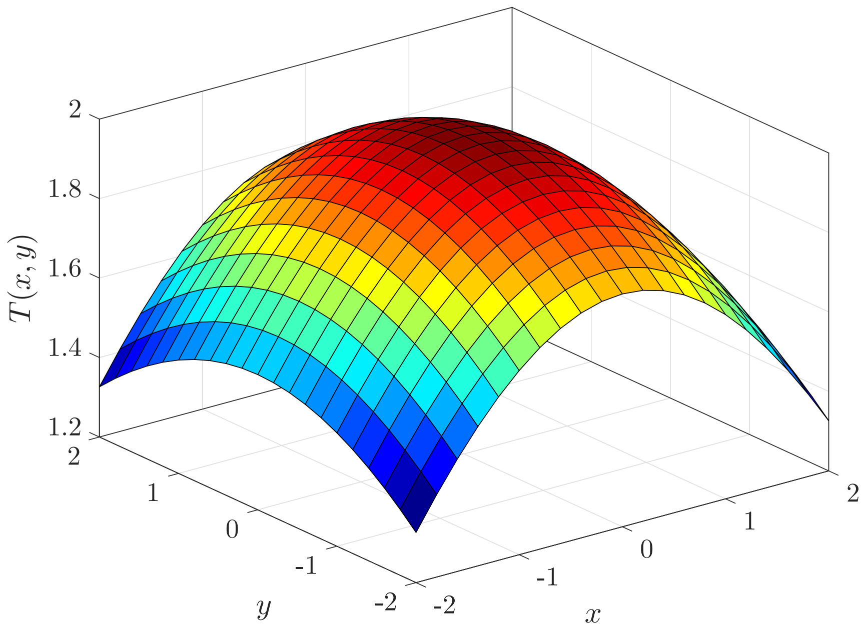


Figure 1: $T(x, y)$ scalar field

For every position, a different vector with its own magnitude and direction is generated. A pictorial representation of the vector field (2) is shown in Figure 2. This vector field returns, when evaluated at position $(1, 1, 1)$, $\mathbf{E}(1, 1, 1) = \frac{Q\left(\hat{\mathbf{\imath}}+\hat{\mathbf{\jmath}}+\hat{\mathbf{k}}\right)}{4\pi\epsilon_0 3^{3/2}}$, a vector, where ϵ_0 is a physical constant.

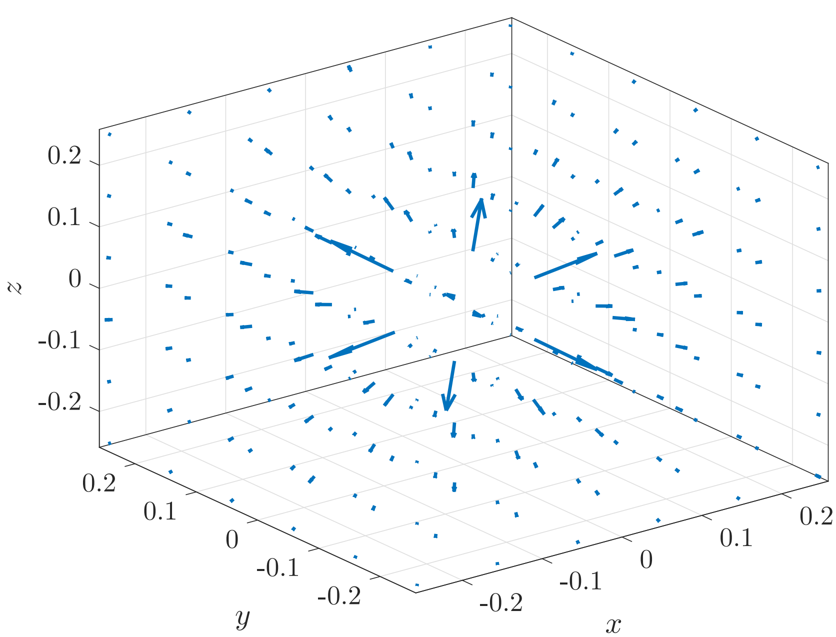


Figure 2: $\mathbf{E}(x, y, z)$ vector field

1.3 Document contents

We start by reviewing some important definitions and axioms associated with vectors, scalar products, vector products and triple products in Section 2. Once the vectors are introduced, we present coordinates systems and revisit the scalar and vector product. Section 3 discusses the two main coordinate systems in 2D: Cartesian and polar coordinates. We then extend the discussion to 3D coordinates systems and discuss Cartesian, cylindrical and spherical coordinates. Section 6 discusses the difference between a position, a position vector, and a general vector. Section 7 presents the velocity and the acceleration vectors in polar, cylindrical and spherical coordinate system. Section 8 gives problems related to the topics covered in this document. Their solutions are given in Section 9. Finally, complementary discussions on the scalar product and on the sum of two vectors described in non-Cartesian coordinate systems is presented in Appendices A and B respectively. These notes are based on the introductory chapters of [1] and [2].

2 Vector operations

Let $\mathbf{v}$ be for example a vector in 2D. $\mathbf{v}$ is visualized as an arrow, and the symbol $\|\mathbf{v}\|$, a scalar, defines its magnitude and represents its length. Alternatively, we may wish to know only the direction of the vector independent of its length in which case we normalize the vector's magnitude to 1. The normalized vector is referred to as a unit vector $\hat{\mathbf{v}}$ and is defined as follows:

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}.$$

Given a unit vector we can always recover the original vector by multiplying it by the magnitude of the original vector that is

$$\mathbf{v} = \|\mathbf{v}\|\,\hat{\mathbf{v}}. \tag{3}$$

We now define and illustrate several vector operations.

2.1 Scalar multiplication

By scalar multiplication, we refer to the case where we are interested in $\mathbf{r} = \alpha\mathbf{v}$ with $\alpha \in \mathbb{R}$, a scalar. The scalar multiplication is a scaling operation: it modifies the magnitude of the vector. When picturing vectors as arrows, the scalar multiplication has the effect of lengthening ($\alpha > 1$) or shortening ($0 < \alpha < 1$) the arrow without changing the direction in which it points. Lastly, the scalar product of α with a vector for $\alpha < 0$ reverses the direction of the vector, i.e. reverses the orientation of the arrow. Scalar multiplication is summarized in Figure 3.

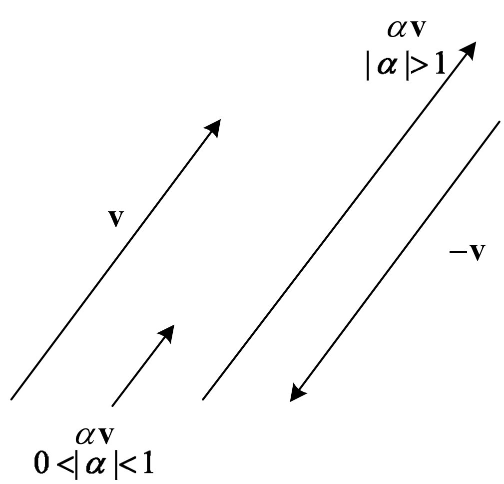
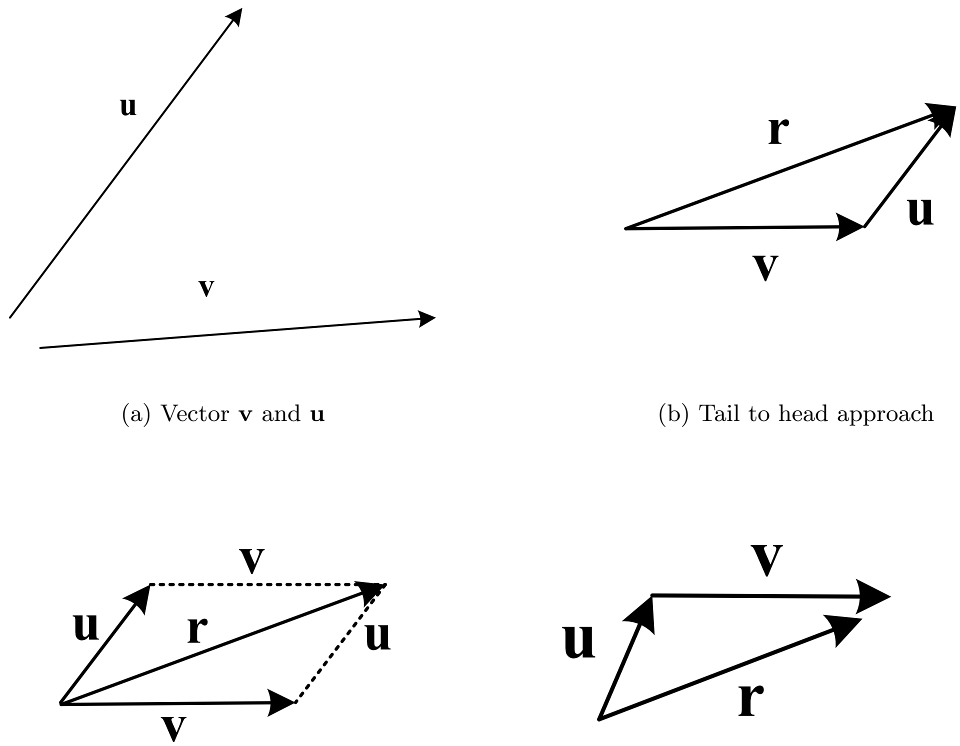


Figure 3: Scalar multiplication of vectors

2.2 Addition

Let $\mathbf{u}$ be a second vector and consider a third vector $\mathbf{r}$ as the addition of the vectors $\mathbf{u}$ and $\mathbf{v}$ that is $\mathbf{r} = \mathbf{u}+\mathbf{v}$. The addition of two vectors can be represented graphically by placing the tail of a vector to the head of the second one. Then, the resulting $\mathbf{r}$ vector is given by the vector between the tail of the initial vector and the head of the second one. Figure 4b summarizes this process.



(a) Vector $\mathbf{v}$ and $\mathbf{u}$ (b) Tail to head approach

(c) Parallelogram approach (d) Commutativity

Figure 4: Addition of two vectors

Another technique for computing the sum of vectors is the parallelogram approach. This is done by first placing both vectors tail to tail. Then, we complete the parallelogram by drawing from the head of each vector a translated version of the two initial vectors. The heads of these two vectors converge on a common point. The resulting vector $\mathbf{r}$ is given by the diagonal of the parallelogram passing through both initial vector tails and the translated vector heads. Figure 4c summarizes this alternative technique.

Properties. The Addition of vectors is commutative: $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$ and associative: $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$ where $\mathbf{w}$ is also a vector.

2.3 Subtraction

The subtraction of two vectors differs from scalar subtraction. Consider $\mathbf{r} = \mathbf{v} - \mathbf{u}$. To evaluate $\mathbf{r}$, we rewrite $\mathbf{r} = \mathbf{v} + (-\mathbf{u})$. In other words, the subtraction of two vectors is the sum of a vector and -1 time the second one. An example for $\mathbf{r} = \mathbf{v} - \mathbf{u}$ is given in Figure 5.

2.4 Linear independence and basis

Knowing the operations we can perform using vectors we can now define the concept of linear independence.

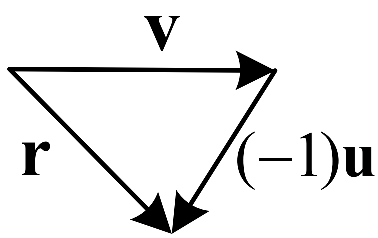


Figure 5: Subtraction of two vectors

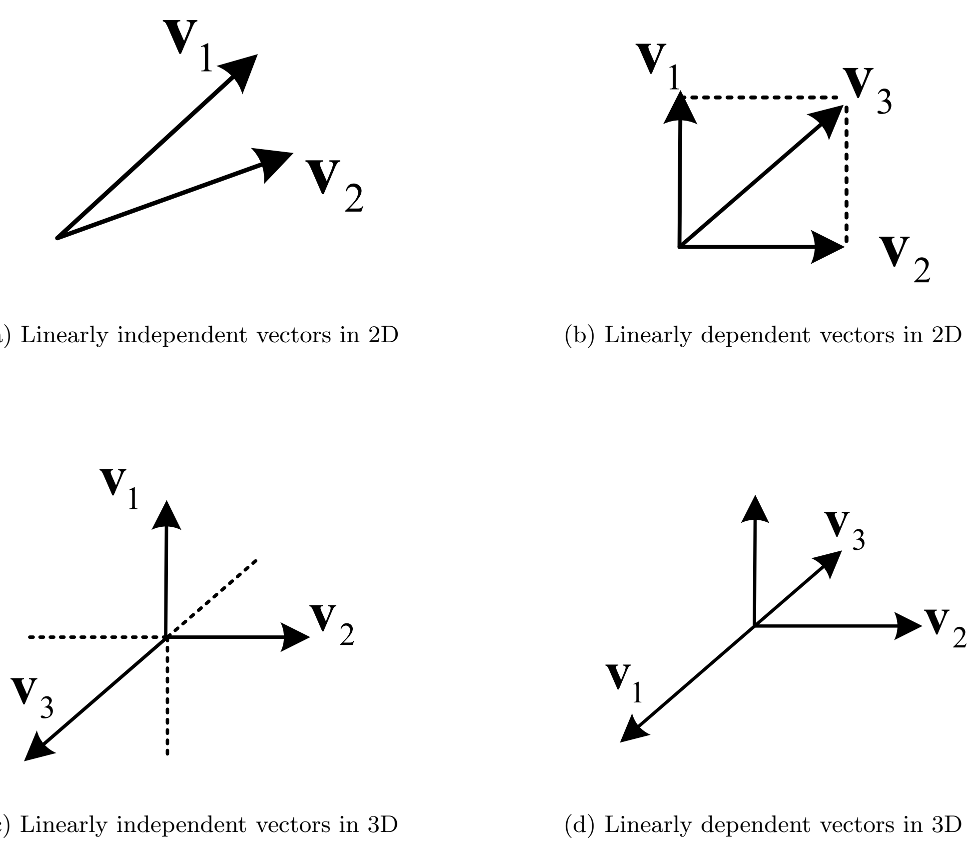
2.4.1 Linear dependence

Consider the set of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Then, this set of vectors is linearly independent if and only if

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots a_n\mathbf{v}_n = \mathbf{0} \tag{4}$$

under the condition that $a_i = 0$ for $i = 1, 2, \ldots, n$ where n represents the dimension of the space. So for example in 2D, $n = 2$ and in 3D, $n = 3$. Alternatively, a set of n dimensional vectors is linearly independent if none of these vectors can not be rewritten as a linear combination of the remaining $n - 1$ vectors.

Figure 6 gives a few examples of linearly dependent and independent vectors in 2D and 3D. We will



(a) Linearly independent vectors in 2D

(b) Linearly dependent vectors in 2D

(c) Linearly independent vectors in 3D

(d) Linearly dependent vectors in 3D

Figure 6: Linearly dependent and independent vectors

consider the example in Figure 6b for illustration purposes and show that the vectors are linearly dependent.

First, we observe that $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$. Hence, there exists a vector in the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ that can be re-expressed as a linear combination of the remaining two vectors. We can conclude that these vectors are not linearly independent. Another way to show their linear dependence is using (4) and a non-zero a_i coefficient. By setting, $a_1 = -1$, $a_2 = -1$ and $a_3 = 1$, we have,

$$-\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0}. \tag{5}$$

Figure 7 demonstrates this last result using a graphical approach based on the results demonstrated in Sections 2.2 and 2.3.

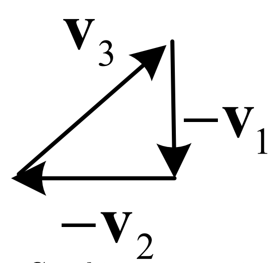


Figure 7: Graphic computation of (5)

2.4.2 Basis

Lastly, we define a basis as a set of linearly independent vectors that can span the whole vector space in which the vectors exist. In others words, a set of vectors forms a basis if using linear combination, we can span any vectors in the space under consideration (e.g. 2D and 3D space).

The choice of basis vectors is not unique and the number of basis vectors required to span a space is equal to the dimension of that space. Later we will show that an orthogonal set of basis vectors is useful from a practical perspective. For example, in the 2D case, we need two vectors, $\mathbf{v}_1$ and $\mathbf{v}_2$, to establish a basis and thus any vector $\mathbf{v}$ can be expressed as a linear combination of basis vectors; e.g. $\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2$. In 3D, we need a set of three vectors to form a basis and thus $\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3$. Note, a,b and c are constants and are real numbers.

2.5 Scalar product (dot product)

The scalar product or dot product between two vectors is denoted by $\mathbf{v} \cdot \mathbf{u}$. Let's first start by considering $\mathbf{v} \cdot \hat{\mathbf{u}}$, where $\hat{\mathbf{u}}$ is the unit vector in the direction of $\mathbf{u}$. Then, the scalar product $\mathbf{v} \cdot \hat{\mathbf{u}}$ represents the projection of the $\mathbf{v}$ vector on $\hat{\mathbf{u}}$. We can interpret projection $\mathbf{v} \cdot \hat{\mathbf{u}}$ as computing how much of $\mathbf{v}$ is directed in the $\hat{\mathbf{u}}$ hat direction. To illustrate a projection, picture a vector $\mathbf{v}$ centered at the origin presented in Figure 8. We are then interested in the projection of this vector onto a vector $\hat{\mathbf{u}}$ along the x axis. The projection of $\mathbf{v}$ represents its x-component given by $\|\mathbf{v}\| \cos\theta_{vu}$. Observe that output of the scalar product and a projection is a scalar.

Now, let's state the formal definition of the scalar product where we put no restriction on the form of the vector $\mathbf{u}$.

$$\boxed{\mathbf{v} \cdot \mathbf{u} = \|\mathbf{v}\|\|\mathbf{u}\| \cos\theta_{vu}} \tag{6}$$

where θ_{vu} is the smallest angle between $\mathbf{v}$ and $\mathbf{u}$ (cf. Figure 8). Hence, for the general case, the scalar product of any two vectors, $\mathbf{v} \cdot \mathbf{u}$, represents a scaled projection of one vector onto the second vector. The scaling factor is given by the magnitude of the second vector. In the definition 6, we retrieve the projection part with $\mathbf{v}\cos\theta_{vu}$ and the scaling factor is the magnitude of $\mathbf{u}$. Note that the other way around is also true: this can represent the projection of $\mathbf{u}$ along $\mathbf{v}$, that is $\mathbf{u}\cos\theta_{vu}$, scaled by the magnitude of $\mathbf{v}$. The scalar product $\mathbf{v} \cdot \mathbf{u}$ is equivalent to the scalar product $\mathbf{u} \cdot \mathbf{v}$ therefore the scalar product is commutative.

The angle θ_{vu} is defined as the smallest angle between the two vectors and hence can only be in the range $[0, \pi]$. This implies that the result of a scalar product can be either positive or negative depending on the angle.

Consider the following three special cases:

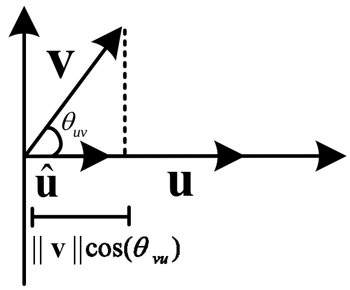


Figure 8: Projection of $\mathbf{v}$ onto $\mathbf{u}$

a) if $\theta_{vu} = \frac{\pi}{2}$, the vectors are perpendicular and using our previous analogy, there is no $\mathbf{u}$-component in $\mathbf{v}$;

b) if $\theta_{vu} = 0$, the scalar product reduces to $\mathbf{v} \cdot \mathbf{u} = \|\mathbf{v}\|\|\mathbf{u}\|$, the product of the magnitude of the two vectors.

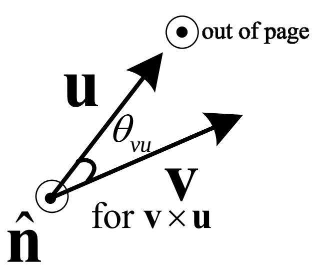
c) if $\mathbf{u} = \mathbf{v}$, we observe the following: $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$, the square of the magnitude of $\mathbf{v}$. This definition of the magnitude in terms of the scalar product holds for all coordinate systems and hence provides a general approach for computing the magnitude of a vector.

Properties. The scalar product is commutative: $\mathbf{v} \cdot \mathbf{u} = \mathbf{u} \cdot \mathbf{v}$. The scalar product is also distributive : $\mathbf{w} \cdot (\mathbf{v} + \mathbf{u}) = \mathbf{w} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{u}$.

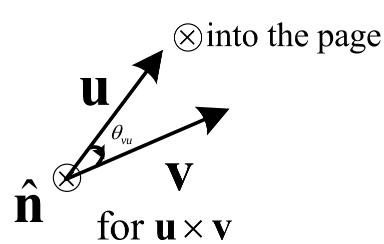
2.6 Vector product (cross product)

The vector product $\mathbf{v} \times \mathbf{u}$ is a vector operation that is only valid in 3D. It takes two vectors as inputs and generates a third vector. This third vector is, by definition, perpendicular to both input vectors and is normal to the plane described by the two vectors. The direction of the resulting vector is given by the right-hand rule that goes as follows: place the edge of your right hand along the first vector, $\mathbf{v}$ and close your fingers in the direction of the smallest angle between the two vectors. Then, the thumb will point in the direction of the resultant vector. We denote $\hat{\mathbf{n}}$ as the unit vector normal to the plane formed by the two vectors in the vector product.

Figure 9 shows graphically the orientation of $\hat{\mathbf{n}}$ for the vector product $\mathbf{v} \times \mathbf{u}$ and $\mathbf{u} \times \mathbf{v}$. Notice that the orientation of $\hat{\mathbf{n}}$ is reversed when the order of the vectors $\mathbf{u}$ and $\mathbf{v}$ are reversed.



(a) $\mathbf{v} \times \mathbf{u}$ and $\hat{\mathbf{n}}$



(b) $\mathbf{v} \times \mathbf{u}$ and $\hat{\mathbf{n}}$

Figure 9: Vector product

The formal definition of the vector product is:

$$\boxed{\mathbf{v} \times \mathbf{u} = \|\mathbf{v}\|\|\mathbf{u}\| \sin \theta_{vu} \hat{\mathbf{n}}} \tag{7}$$

In this case, we note that if $\mathbf{v}$ and $\mathbf{u}$ are parallel, then $\theta_{vu} = 0$ and the cross product is zero.

The magnitude of $\mathbf{v} \times \mathbf{u}$ also has a geometric interpretation. It represents the area of a parallelogram formed by the two vectors placed tail to tail as represented in Figure 10. To see this, we first note that $\|\mathbf{v} \times \mathbf{u}\| = \|\mathbf{v}\|\|\mathbf{u}\| \sin\theta_{vu}\|\hat{\mathbf{n}}\| = \|\mathbf{v}\|\|\mathbf{u}\| \sin\theta_{vu}$, since the magnitude of $\hat{\mathbf{n}}$ is one. We analyze $\|\mathbf{v}\|$ and $\|\mathbf{u}\| \, |\sin\theta_{vu}|$ separately. First, the magnitude of $\mathbf{v}$ represents the base of the parallelogram. Then, trace a right-angled triangle ABC as shown in Figure 10. Using the definition of sin, $\|\mathbf{u}\| \sin\theta_{vu}$ represents the length of the opposite side in the triangle, and is the the height of the parallelogram. The area of a parallelogram is given by the product of its height and base and hence we get our result.

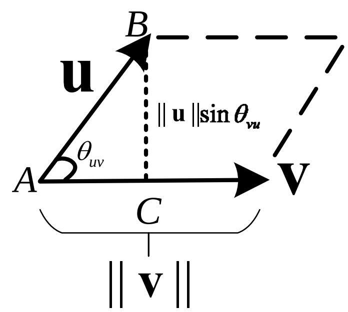


Figure 10: Projection of $\mathbf{v}$ onto $\mathbf{u}$

Properties. The cross product is anti-commutative: $\mathbf{v} \times \mathbf{u} = -\mathbf{u} \times \mathbf{v}$. We can derive this relation using the right-hand rule. The angle θ_{vu} stays the same, but the direction of the thumb and hence of $\hat{\mathbf{n}}$ is reversed and we find the anti-commutative relation. The vector product is distributive over addition: $\mathbf{w} \times (\mathbf{v} + \mathbf{u}) = \mathbf{w} \times \mathbf{v} + \mathbf{w} \times \mathbf{u}$. It is not associative : $\mathbf{w} \times (\mathbf{v} \times \mathbf{u}) \neq (\mathbf{w} \times \mathbf{v}) \times \mathbf{u}$. This last statement can be shown graphically by only considering the direction of each cross product. For example consider two nonorthogonal vectors $\mathbf{u}$ and $\mathbf{v}$ lying in a plane and $\mathbf{w}$ normal to this plane. Then, the left-hand side of the statement is 0 whereas the right-hand side is not zero.

2.7 Scalar triple product

The triple product is a vector operation defined over three input vectors. There are two types of triple product: a scalar triple product and a vector triple product. The vector triple product will be covered in the next section.

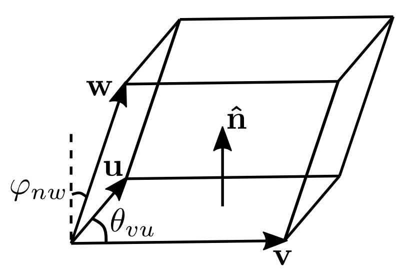


Figure 11: Parallelepiped formed by $\mathbf{w}$, $\mathbf{u}$ and $\mathbf{v}$

The scalar triple product, as its name suggests, generates a scalar output. The scalar triple product takes the form $\mathbf{w} \cdot (\mathbf{v} \times \mathbf{u})$. The scalar triple product, like the vector product, can be geometrically interpreted. The absolute value of the scalar triple product represents the volume of a parallelepiped (3D parallelogram) formed by the three vectors placed tail to tail to tail. This parallelepiped is presented in Figure 11.

To show this, we first apply the vector and scalar product using (6) and (7) respectively:

$$\mathbf{v} \times \mathbf{u} = \hat{\mathbf{n}} \, \|\mathbf{v}\| \, \|\mathbf{u}\| \sin\theta_{vu}, \tag{8}$$

$$\mathbf{w} \cdot (\mathbf{v} \times \mathbf{u}) = \|\mathbf{w}\| \, \|\mathbf{v} \times \mathbf{u}\| \cos\varphi_{nw}. \tag{9}$$

Here $\hat{\mathbf{n}}$ is a unit vector normal to the $\mathbf{u}$,$\mathbf{v}$-plane, θ_{vu} is the angle between $\mathbf{u}$ and $\mathbf{v}$ and φ_{nw} denotes the angle between $\mathbf{w}$ and the normal $\hat{\mathbf{n}}$. The norm of (8) is,

$$\|\mathbf{v} \times \mathbf{u}\| = \|\mathbf{v}\| \, \|\mathbf{u}\| \sin \theta_{vu}$$

since $\hat{\mathbf{n}}$ is a unit vector. Substituting this result in (9) results in,

$$\mathbf{w} \cdot (\mathbf{v} \times \mathbf{u}) = \|\mathbf{w}\| \cos \varphi_{nw} \, \|\mathbf{v}\| \, \|\mathbf{u}\| \sin \theta_{vu}$$

We now have two distinct terms: $\|\mathbf{w}\| \, |\cos \varphi_{nw}|$ and $\|\mathbf{v}\| \, \|\mathbf{u}\| \sin \theta_{vu}$. In the last section, we showed that $|\|\mathbf{v}\| \, \|\mathbf{u}\| \sin \theta_{vu}|$ is equal to the area formed by the parallelogram. In this case, this result gives us the area of the base of the parallelepiped. The scalar product term $\|\mathbf{w}\| \, |\cos \varphi_{nw}|$ is equivalent to the projection of $\mathbf{w}$ onto the direction of $\hat{\mathbf{n}}$, i.e. the magnitude of the $\hat{\mathbf{n}}$ component of $\mathbf{w}$, which represents the base of the parallelepiped. Hence, the product of the two terms gives the volume of the parallelepiped formed by $\mathbf{w}$, $\mathbf{v}$ and $\mathbf{u}$.

Properties. The scalar triple product is invariant under cyclic permutation, that is,

$$\mathbf{w} \cdot (\mathbf{v} \times \mathbf{u}) = \mathbf{v} \cdot (\mathbf{u} \times \mathbf{w}) = \mathbf{u} \cdot (\mathbf{w} \times \mathbf{v}).$$

Since the vector product is anti-commutative, the scalar triple product is also anti-commutative with respect to the two vectors comprising the vector product:

$$\mathbf{w} \cdot (\mathbf{v} \times \mathbf{u}) = -\mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = -\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = -\mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}).$$

The motivation for the use of the scalar triple product aside from its geometric interpretation is described in Appendix A.

2.8 Vector triple product

The second type of triple product is the vector triple product. It won't be used in this course so we only define it briefly and omit the proof of the next identity. The vector triple product outputs a vector and is defined as $\mathbf{w} \times (\mathbf{v} \times \mathbf{u})$. The vector triple product can also be re-expressed according to the following vector identity,

$$\mathbf{w} \times (\mathbf{v} \times \mathbf{u}) = \mathbf{v}(\mathbf{w} \cdot \mathbf{u}) - \mathbf{u}(\mathbf{w} \cdot \mathbf{v}).$$

3 Description and Representation of Coordinate Systems

Thus far, we have looked at vectors from a general perspective by sketching them as directed arrows in space. We can now introduce the tool to represent them according to basis vector components which are defined using a coordinate system. A coordinate system of a n dimensional space is a system at which a point in space is defined by the intersection point of n objects having an $(n-1)$ dimensional form. In 3D $n = 3$ corresponds to a volume, $(n-1 = 2)$-dimensional form is a surface, $(n-2 = 1)$-dimensional form is a line and $(n-3 = 0)$-dimensional form is a point and similarly for 2D. In 2D $(n = 2)$, the intersection of 2 lines (objects of $(n-1)$-dimensional form) produce a point whereas in 3D $(n = 3)$ intersection of 3 surfaces ($(n-1)$-dimensional form) produce a point. The $(n-1)$-dimensional objects also reveal the coordinate direction or vectors that form the basis of a system which we refer as the base vectors. As mentioned previously, base vectors are linearly independent vectors spanning the whole space of interest and hence can span any vectors included in the n dimensional space. There are no restrictions about their relative orientation or norm at this point. It is however convenient to choose the base vectors as unit vectors that are all orthogonal with respect to each other in which case we have an orthonormal basis. Recall that orthogonality implies that the angle between any two base vectors is 90 degrees and hence their scalar dot product is zero.

3.1 Representation of basis vectors

In a coordinate system, the base vectors are defined as the directions perpendicular to the $(n-1)$-dimension objects. With a coordinate system, a position is specified by an n-tuple made of the parameters of the $(n-1)$-dimensional objects specific to the system. Then, arbitrary vectors are rewritten as a linear combination of the base vectors of the coordinate system. Finally, the base vectors point in the increasing direction of an axis and to fix the proper direction, we need to impose the constraint of a right-handed system. A system is called right-handed if it satisfies a property that can be verified by the right hand rule. To check this condition, we first order the base vectors and then we place our index finger along the first direction of the base vector and our middle finger along the direction of the second base vector. Then, if the system is right-handed, the thumb will point in the direction described by the third base vector. The order of the unit base vectors in a Cartesian system is given by convention in terms of an ascending order, i.e. $\hat{\mathbf{x}}$-$\hat{\mathbf{y}}$-$\hat{\mathbf{z}}$. A more formal definition will be given after orthogonal and orthonormal coordinate systems are introduced.

In summary, the base vectors we consider here are unit vectors perpendicular to their respective $(n-1)$-dimensional objects. Since these objects are perpendicular to each other at a given point, the unit vectors are also orthogonal to each other. Hence, a 2D orthonormal coordinate system with base vectors $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ will satisfy the following conditions,

$$\begin{aligned} \hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_1 &= 1 \qquad & \hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_1 &= 0 \\ \hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2 &= 0 \qquad & \hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_2 &= 1, \end{aligned}$$

and a 3D orthonormal coordinate system with base vectors $\hat{\mathbf{n}}_1$, $\hat{\mathbf{n}}_2$ and $\hat{\mathbf{n}}_3$ will satisfy the following conditions:

$$\begin{aligned} \hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_1 &= 1 \qquad & \hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_1 &= 0 \qquad & \hat{\mathbf{n}}_3 \cdot \hat{\mathbf{n}}_1 &= 0 \\ \hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_2 &= 0 \qquad & \hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_2 &= 1 \qquad & \hat{\mathbf{n}}_3 \cdot \hat{\mathbf{n}}_2 &= 0 \\ \hat{\mathbf{n}}_1 \cdot \hat{\mathbf{n}}_3 &= 0 \qquad & \hat{\mathbf{n}}_2 \cdot \hat{\mathbf{n}}_3 &= 0 \qquad & \hat{\mathbf{n}}_3 \cdot \hat{\mathbf{n}}_3 &= 1, \end{aligned}$$

where we used $\hat{\mathbf{n}}_i$ to denote the fact that the base vectors are normalized. However, note that there is nothing which keeps us from using a non-orthogonal coordinate system. For example, a non-orthogonal coordinate system can be used in crystallography when working with a triclinic or monoclinic lattice. The monoclinic system is sketched in Figure 12 . In this case, the angle between the three base vectors associated with the crystallographic axes are $\alpha = \pi/2$, $\gamma = \pi/2$ and $\beta \neq \pi/2$ making this system non-orthogonal.

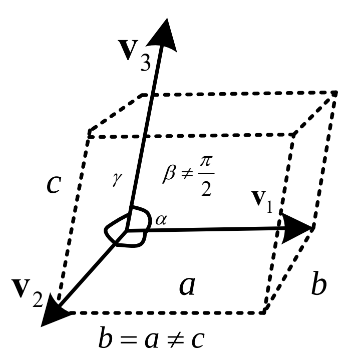


Figure 12: Monoclinic system

For an orthonormal 3D coordinate system, the right-handed system condition can be defined formally using the cross product of its base vectors. The base vectors with the ordered sequence $\hat{\mathbf{n}}_1$, $\hat{\mathbf{n}}_2$ and $\hat{\mathbf{n}}_3$ of an orthonormal right-handed system must satisfy the following condition:

$$\begin{aligned} \hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2 &= \hat{\mathbf{n}}_3 \\ \hat{\mathbf{n}}_2 \times \hat{\mathbf{n}}_3 &= \hat{\mathbf{n}}_1 \\ \hat{\mathbf{n}}_3 \times \hat{\mathbf{n}}_1 &= \hat{\mathbf{n}}_2 \end{aligned} \tag{10}$$

which is equivalent to the right-hand rule.

In the two next sections, we will study the most common orthonormal coordinate systems. Our goal is to express any vector in the form of a linear combination of the base vectors. In the nD case where n represents the number of dimensions, this means that we will express a vector $\mathbf{v}$ as

$$\mathbf{v} = v_1\hat{\mathbf{n}}_1 + v_2\hat{\mathbf{n}}_2 + \ldots + v_n\hat{\mathbf{n}}_n,$$

where $\hat{\mathbf{n}}_i$ $i = 1, 2, \ldots, n$ are the base vectors for an orthonormal basis defined by the coordinate system and v_i, $i = 1, 2, \ldots, n$, are the weights of each base vector. This representation of a vector is referred to the component form. In the 3D case we have,

$$\mathbf{v} = v_1\hat{\mathbf{n}}_1 + v_2\hat{\mathbf{n}}_2 + v_3\hat{\mathbf{n}}_3. \tag{11}$$

Consider a Cartesian coordinate system in which case we can denote the orthonormal basis vectors as,

$$\hat{\mathbf{n}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \hat{\mathbf{n}}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \hat{\mathbf{n}}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

(11) is equivalent to,

$$\begin{aligned} \mathbf{v} &= \begin{bmatrix} | & | & | \\ \hat{\mathbf{n}}_1 & \hat{\mathbf{n}}_2 & \hat{\mathbf{n}}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}. \end{aligned}$$

Alternatively, we can represent the $\mathbf{v}$ vector in a single column form as,

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix},$$

where the base vectors are taken to be implicitly $\hat{\mathbf{n}}_1$, $\hat{\mathbf{n}}_2$ and $\hat{\mathbf{n}}_3$ and the representation of the unit base unit vectors depends on the coordinate system. The goal of the coming section is thus to define the $\hat{\mathbf{n}}_i$ needed to represent a vector in a coordinate system other than Cartesian..

3.2 Scalar product in component form

We now revisit the scalar product using the component form in an orthonormal 3D coordinate system. First, we relate the scalar product formal definition (6) to the component form of the scalar product in the case of an orthonormal system using a geometric interpretation. We demonstrate the relation only in 2D for Cartesian coordinates (cf. Section 4.1) but it can also be extended to higher dimensions and other coordinate systems.

Consider the vectors $\mathbf{v}$ and $\mathbf{u}$ represented in Figure 13.

$$\mathbf{u} \cdot \mathbf{v} = u_x v_x + u_y v_y \tag{12}$$

The dot product $\mathbf{u} \cdot \mathbf{v}$ can be re-expressed as,

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta_{uv} \\ &= \|\mathbf{u}\| \, \|\mathbf{v}\| \cos(\theta_1 - \theta_2) \\ &= \|\mathbf{u}\| \, \|\mathbf{v}\| (\cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2) \end{aligned} \tag{13}$$

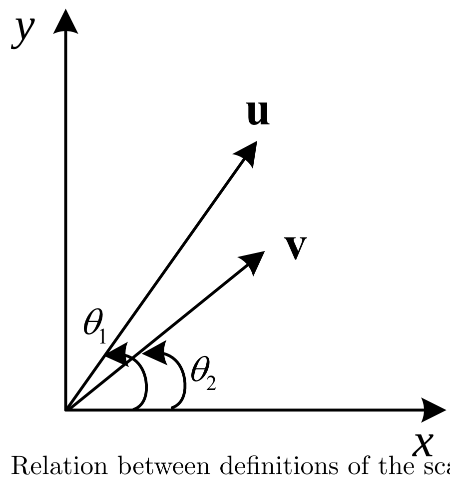


Figure 13: Relation between definitions of the scalar product

Using Figure 13 and denoting u_x, u_y and v_x and v_y as the x-component and y-component of $\mathbf{u}$ and $\mathbf{v}$ respectively we can observe the following relations:

$$\cos\theta_1 = \frac{u_x}{\|\mathbf{u}\|}, \qquad \sin\theta_1 = \frac{u_y}{\|\mathbf{u}\|},$$
$$\cos\theta_2 = \frac{v_x}{\|\mathbf{v}\|}, \qquad \sin\theta_2 = \frac{v_y}{\|\mathbf{v}\|}.$$

We can therefore rewrite (13) as,

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= \|\mathbf{u}\|\,\|\mathbf{v}\|\left(\frac{u_x}{\|\mathbf{u}\|}\frac{v_x}{\|\mathbf{v}\|} + \frac{u_y}{\|\mathbf{u}\|}\frac{v_y}{\|\mathbf{v}\|}\right),\\ &= u_x v_x + u_y v_y,\end{aligned}$$

and we retrieve the component form of the scalar product. In 3D, the angle θ is between vector $\mathbf{u}$ and $\mathbf{v}$ on the plane defined by the vectors $\mathbf{u}$ and $\mathbf{v}$. Performing the same calculations, we get,

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta = u_x v_x + u_y v_y + u_z v_z$$

Finally, we can extend this to spaces with more than three dimensions.

Then, following our introduction of the component form of the scalar product with the geometric interpretation, we can consider the general case of the dot product of two arbitrary 3D vectors expressed in terms of an orthonormal basis where,

$$\begin{aligned}\mathbf{v} &= v_1\hat{\mathbf{n}}_1 + v_2\hat{\mathbf{n}}_2 + v_3\hat{\mathbf{n}}_3,\\ \mathbf{u} &= u_1\hat{\mathbf{n}}_1 + u_2\hat{\mathbf{n}}_2 + u_3\hat{\mathbf{n}}_3,\end{aligned}$$

then,

$$\begin{aligned}\mathbf{v}\cdot\mathbf{u} &= (v_1\hat{\mathbf{n}}_1 + v_2\hat{\mathbf{n}}_2 + v_3\hat{\mathbf{n}}_3)\cdot(u_1\hat{\mathbf{n}}_1 + u_2\hat{\mathbf{n}}_2 + u_3\hat{\mathbf{n}}_3),\\ &= v_1u_1\hat{\mathbf{n}}_1\cdot\hat{\mathbf{n}}_1 + v_1u_2\hat{\mathbf{n}}_1\cdot\hat{\mathbf{n}}_2 + v_1u_3\hat{\mathbf{n}}_1\cdot\hat{\mathbf{n}}_3 + v_2u_1\hat{\mathbf{n}}_2\cdot\hat{\mathbf{n}}_1 + v_2u_2\hat{\mathbf{n}}_2\cdot\hat{\mathbf{n}}_2 + v_2u_3\hat{\mathbf{n}}_2\cdot\hat{\mathbf{n}}_3 + v_3u_1\hat{\mathbf{n}}_3\cdot\hat{\mathbf{n}}_1\\ &\quad + v_3u_2\hat{\mathbf{n}}_3\cdot\hat{\mathbf{n}}_2 + v_3u_3\hat{\mathbf{n}}_3\cdot\hat{\mathbf{n}}_3 \end{aligned}\tag{14}$$

We recall the definition of the dot product, notably that if two vectors are perpendicular, then $\theta = \pi/2$ and the dot product is zero and if they are parallel, $\theta = 0$ and the dot product is equal to the product of the norms. In this case, the norms are all 1 since $\hat{\mathbf{n}}_i$ are unit vectors. Hence, (14) reduces to

$$\mathbf{v}\cdot\mathbf{u} = v_1u_1 + v_2u_2 + v_3u_3. \tag{15}$$

This last form is valid for all orthogonal coordinate systems and can be generalized to a larger number of dimensions in space. This also provides one justification of why orthonormal coordinate systems are easier to work with: there are no cross terms in the dot product expansion.

3.3 Vector product in component form

We can also give an alternate definition of the vector product in component form in 3D. We consider $\mathbf{v}$ and $\mathbf{u}$ as previously defined and compute the cross product between the two vectors,

$$\begin{aligned}\mathbf{v}\times\mathbf{u} &= (v_1\hat{\mathbf{n}}_1 + v_2\hat{\mathbf{n}}_2 + v_3\hat{\mathbf{n}}_3)\times(u_1\hat{\mathbf{n}}_1 + u_2\hat{\mathbf{n}}_2 + u_3\hat{\mathbf{n}}_3),\\ &= v_1u_1\hat{\mathbf{n}}_1\times\hat{\mathbf{n}}_1 + v_1u_2\hat{\mathbf{n}}_1\times\hat{\mathbf{n}}_2 + v_1u_3\hat{\mathbf{n}}_1\times\hat{\mathbf{n}}_3 + v_2u_1\hat{\mathbf{n}}_2\times\hat{\mathbf{n}}_1 + v_2u_2\hat{\mathbf{n}}_2\times\hat{\mathbf{n}}_2 + v_2u_3\hat{\mathbf{n}}_2\times\hat{\mathbf{n}}_3\\ &\quad + v_3u_1\hat{\mathbf{n}}_3\times\hat{\mathbf{n}}_1 + v_3u_2\hat{\mathbf{n}}_3\times\hat{\mathbf{n}}_2 + v_3u_3\hat{\mathbf{n}}_3\times\hat{\mathbf{n}}_3 \end{aligned}\tag{16}$$

In an orthonormal right-handed coordinate set, we can use three properties to compute (16): (i) the cross product of a vector with itself is zero, (ii) the anti-commutative property of the cross product, (iii) the right-handed condition (10). The vector product can be computed in the following way,

$$\mathbf{v}\times\mathbf{u} = v_1u_2\hat{\mathbf{n}}_3 - v_1u_3\hat{\mathbf{n}}_2 - v_2u_1\hat{\mathbf{n}}_3 + v_2u_3\hat{\mathbf{n}}_1 + v_3u_1\hat{\mathbf{n}}_2 - v_3u_2\hat{\mathbf{n}}_1. \tag{17}$$

Rearranging (17), we have,

$$\mathbf{v}\times\mathbf{u} = (v_2u_3 - v_3u_2)\,\hat{\mathbf{n}}_1 - (v_1u_3 - v_3u_1)\,\hat{\mathbf{n}}_2 + (v_1u_2 - v_2u_1)\,\hat{\mathbf{n}}_3. \tag{18}$$

Finally, (18) is equivalent to,

$$\mathbf{u}\times\mathbf{v} = \begin{vmatrix}\hat{\mathbf{n}}_1 & \hat{\mathbf{n}}_2 & \hat{\mathbf{n}}_3\\ u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\end{vmatrix} = |\mathbf{M}|\,,$$

where $|\mathbf{M}|$ represents the determinant of $\mathbf{M}$.

This concludes our introduction of vectors and coordinate system. Let's now explore the common orthogonal coordinate systems.

4 2D coordinate systems

We begin our coordinate system survey by considering the two dimensional case. In 2D, we will cover Cartesian and polar coordinate systems. For all coordinate systems, we will discuss how positions are represented, what the base vectors and vector representation are and how to convert from one coordinate system to another. Finally, we will discuss differential length, surface and volume as we will need these tools later in the course.

4.1 Cartesian or rectangular

4.1.1 Definition

The Cartesian coordinate system (sometimes referred as rectangular in 2D) is the most common system. It is also the easiest to work with, but may on occasion be awkward to use because the symmetry of the problem may suggest the need for an alternative coordinate system to simplify calculations. For example, consider any point at a certain distance from the origin having the same property (e.g. magnitude of potential is constant) . This means that all points lying on a circle of a given radius are equivalent when computations are made and thus we have a symmetry commonly referred to as a polar symmetry.

As described earlier, a position in 2D is described by the intersection of two lines in a Cartesian space. The first line is the line $x = x_0$ and the second line is the line $y = y_0$. We first observe that the two lines are perpendicular and hence the 2D Cartesian system is orthogonal. Both $x_0 \in]-\infty, +\infty[$ and $y_0 \in]-\infty, +\infty[$ are constant and taken together represent the position of a point P ,where x and y are each aligned with one of the two base vectors. This is represented in Figure 14. Please note the following convention for denoting intervals is used throughout this work:

- $]a, b]$ is equivalent to $(a, b]$, that is the interval on the real line from a to b, excluding a and including b;
- $]a, b[$ is equivalent to (a, b), that is from a to b excluding both a and b;

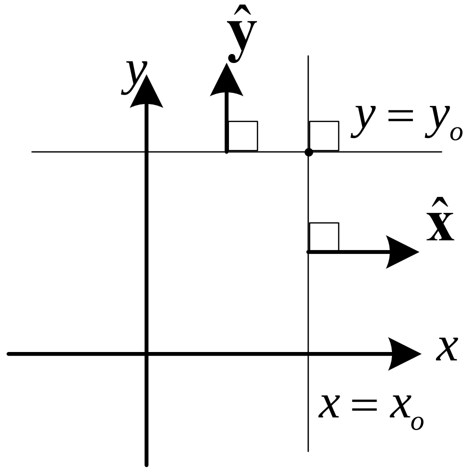


Figure 14: Cartesian coordinate system

- and so on.

With these two lines for varying x_0 and y_0, we can represent any point in a 2D space.

As introduced in the previous section, the unit direction and base vectors of a 2D system are given by a unit vector perpendicular to the $(n-1)$ dimensional object. In a 2D Cartesian, this means that one base vector is perpendicular to the line $x = x_0$ and the second base vector is perpendicular to the line $y = y_0$. However, that leaves us with two possible orientation for each base vector.

In order to fix the orientation, we recall the definition of a right-handed system. By convention, the ordered base vectors are $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}\}$. We then align the hand with the $\hat{\mathbf{x}}$ direction and then the fingers point along the direction of $\hat{\mathbf{y}}$ so it looks like a counterclockwise rotation with reference to the x axis in the xy plane.

The two base vectors denoted as $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are represented in Figure 14 and indicate the increasing dimension of each axis.

Finally, we can re-write any vector $\mathbf{v} \in \mathbb{R}^2$ as

$$\boxed{\mathbf{v} = v_x \hat{\mathbf{x}} + v_y \hat{\mathbf{y}}}$$

where v_x, v_y are two constants. These constants represent the amount of $\hat{\mathbf{x}}$-vector and $\hat{\mathbf{y}}$-vector needed to construct $\mathbf{v}$. They hence represent the projection of $\mathbf{v}$ onto $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ respectively. Figure 15 gives an example of this representation.

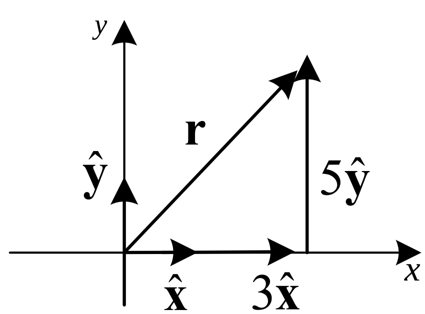


Figure 15: Vector decomposition in Cartesian

4.1.2 Differential elements

We now investigate the representation of the differential geometric elements in a Cartesian coordinate system. The differential geometric elements; length and surface area, can be found by looking at the change due to an infinitesimal variation in every direction associated with the base vectors. We start by computing the differential length $\mathrm{d}\ell_i$, $i = x, y$. From that, we can obtain two things, (i) the vector representing differential length for an arbitrary orientation and (ii) the differential surface area.

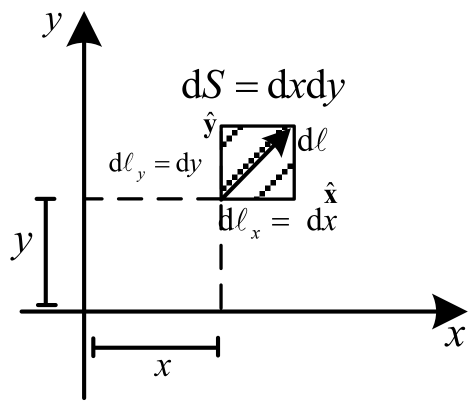


Figure 16: Differential elements in Cartesian

For the Cartesian case, Figure 16 gives a representation of the infinitesimal variation along the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ direction. From the figure, we deduce that $\mathrm{d}\ell_x = \mathrm{d}x$ and $\mathrm{d}\ell_y = \mathrm{d}y$. Hence, the vector representing the differential length is

$$\boxed{\mathrm{d}\boldsymbol{\ell} = \mathrm{d}\ell_x \hat{\mathbf{x}} + \mathrm{d}\ell_y \hat{\mathbf{y}}}$$

Next, the differential surface area $\mathrm{d}S$ is given by the area of the rectangle created by the variations along the direction of the basis vectors. Hence, we have

$$\mathrm{d}S = \mathrm{d}\ell_x \mathrm{d}\ell_y$$

or,

$$\boxed{\mathrm{d}S = \mathrm{d}x\mathrm{d}y}$$

4.2 Polar

4.2.1 Definition

Like all 2D coordinate systems, the polar coordinate system is described by two lines. In a polar coordinate system, the first line is the circle defined by $r = r_0$. The variable r is taken to be greater than or equal to zero. This circle represents all the points located on a line which is at a distance r_0 from the origin, in other words all (x, y) points that satisfy the equation $\sqrt{x^2 + y^2} = r_0$. We can also see this circular line as the union of two half circular lines: $f(x) = \sqrt{r_0^2 - x^2}$ and $f(x) = -\sqrt{r_0^2 - x^2}$ for $x \in [-r_0, r_0]$. The second line is the straight line going through the origin and making an angle $\theta = \theta_0$ with respect to the x axis measured in a counterclockwise direction. Figure 17 shows the two lines that define the polar coordinate system. Hence, any point in 2D can be expressed as a pair (r, θ) where $r \in [0, +\infty[$ and $\theta \in [0, 2\pi]$ as shown in Figure 17. We observe that for any θ_0 and r_0, the line $\theta = \theta_0$ and the circle $r = r_0$ are perpendicular, thus making this system an orthogonal system. Then, the base vectors of this system are defined by unit vectors oriented perpendicularly to the two lines. Once again, both unit vectors could have two orientations: inward and outward for $\hat{\mathbf{r}}$ and clockwise and counterclockwise for $\hat{\boldsymbol{\theta}}$. In this case, we first order the two base vectors as $\{\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}\}$ and set $\hat{\mathbf{r}}$ to point outward so that the r component increases when going further away from the origin.

This has a geometric meaning since r represents the distance of a point from the origin which increases as we go further away from the origin. Then placing our hand along the direction of $\hat{\mathbf{r}}$ and closing it, our fingers point in the counterclockwise direction which gives the direction of $\hat{\boldsymbol{\theta}}$. Setting $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ this way makes the 2D coordinate system right-handed. Finally, the base vectors $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ for the polar coordinate system are represented in Figure 17.

In conclusion, in a right-handed polar coordinate system, the base vector or unit vector $\hat{\mathbf{r}}$ is radial, in other words, from the origin it points outwards and the base vector $\hat{\boldsymbol{\theta}}$ is oriented counterclockwise and is always oriented tangential to a circle that intersects the point (r_0, θ_0). We can write any vector $\mathbf{v} \in \mathbb{R}^2$ in a polar coordinate system in the form of

$$\boxed{\mathbf{v} = v_r \hat{\mathbf{r}} + v_\theta \hat{\boldsymbol{\theta}},}$$

where the coefficients v_r, v_θ are given by the projection along the $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ direction.

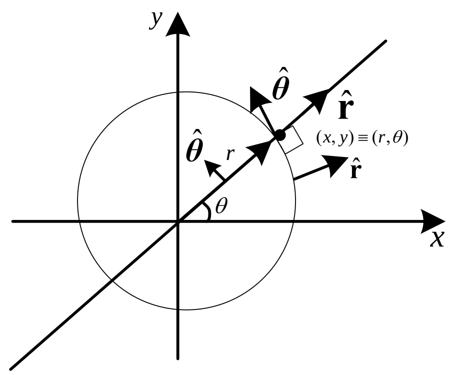


Figure 17: Polar coordinate system

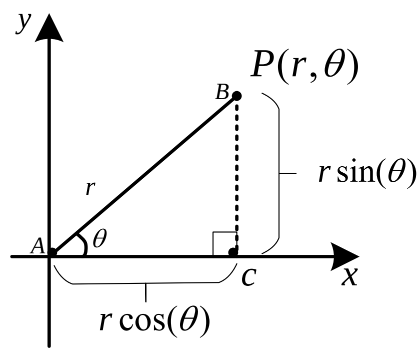


Figure 18: Relation between position in Cartesian and polar coordinates

4.2.2 Transformation

We now want to convert a position $P(r, \theta)$ and a vector $\mathbf{v} = v_r \hat{\mathbf{r}} + v_\theta \hat{\boldsymbol{\theta}}$ from polar coordinates to Cartesian coordinates. The former is straightforward and only uses trigonometry. The latter, however, needs a bit more work and we detail our approach later on and re-use it in the sections to follow.

Given a point $P(r, \theta)$ as shown in Figure 18, we can trace the right-angled triangle ABC and use trigonometry to find the base x and height y of this vector. Recalling the definition of sine and cosine, we obtain the following relation to convert from polar coordinates to Cartesian coordinates:

$$\boxed{\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned}} \tag{19}$$

Or inversely, from polar coordinates to Cartesian coordinates:

$$\boxed{\begin{aligned} r &= \sqrt{x^2 + y^2} \\ \theta &= \arctan \frac{y}{x} \end{aligned}} \tag{20}$$

Note that arctan can have two possible values for the same ratio y/x. Using (20) and (19) we convert from one position in one coordinate system to an equivalent position in the other coordinate system.

We now want to convert a vector representation in one coordinate system to a new vector in a different coordinate system. We need to define a mapping between a vector expressed in one coordinate system to a vector in the other coordinate system. For vectors, this mapping is done using a transformation matrix. Hence, we need to determine the proper transformation matrix which we will refer to as $\mathbf{M}$. The objective will be to determine the elements of the transformation matrix.

Denote $\mathbf{v}$ as a 2D vector and denote $\mathbf{v}_{\text{cart}} = v_x\hat{\mathbf{x}} + v_y\hat{\mathbf{y}}$ as the vector expressed in Cartesian coordinates and $\mathbf{v}_{\text{pol}} = v_r\hat{\mathbf{r}} + v_\theta\hat{\boldsymbol{\theta}}$ as the vector expressed in polar coordinates. Note that $\mathbf{v} = \mathbf{v}_{\text{cart}} = \mathbf{v}_{pol}$ even though they are expressed using different basis vectors associated with the different coordinate systems. Let's first start by converting $\mathbf{v}_{\text{pol}}$ into $\mathbf{v}_{\text{cart}}$. The form of the desired equation is

$$\mathbf{v}_{\text{cart}} = \mathbf{M}\mathbf{v}_{\text{pol}} \tag{21}$$

Or equivalently,

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} v_r \\ v_\theta \end{bmatrix} \tag{22}$$

Now, recall how we obtained v_x in the previous section. It is the projection of $\mathbf{v}_{\text{cart}}$ onto the base vector $\hat{\mathbf{x}}$. Equivalently since $\mathbf{v}_{\text{cart}} = \mathbf{v}_{\text{pol}}$, $v_x = \mathbf{v}_{pol} \cdot \hat{\mathbf{x}}$

Similarly for v_y, we have $v_y = \mathbf{v}_{pol} \cdot \hat{\mathbf{y}}$. In summary,, we have,

$$\begin{aligned} v_x &= \mathbf{v}_{\text{pol}} \cdot \hat{\mathbf{x}} \\ v_y &= \mathbf{v}_{\text{pol}} \cdot \hat{\mathbf{y}} \end{aligned}$$

We replace $\mathbf{v}_{pol}$ in the two previous equations above by their equivalent equations in polar coordinates. The set of equations becomes:

$$\begin{aligned} v_x &= \left(v_r\hat{\mathbf{r}} + v_\theta\hat{\boldsymbol{\theta}}\right) \cdot \hat{\mathbf{x}} \\ v_y &= \left(v_r\hat{\mathbf{r}} + v_\theta\hat{\boldsymbol{\theta}}\right) \cdot \hat{\mathbf{y}} \end{aligned}$$

or,

$$\begin{aligned} v_x &= \hat{\mathbf{r}} \cdot \hat{\mathbf{x}} v_r + \hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{x}} v_\theta \\ v_y &= \hat{\mathbf{r}} \cdot \hat{\mathbf{y}} v_r + \hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{y}} v_\theta \end{aligned}$$

We see that we can rewrite this equation in the form of (22) and we obtain

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{r}} \cdot \hat{\mathbf{x}} & \hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{x}} \\ \hat{\mathbf{r}} \cdot \hat{\mathbf{y}} & \hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{y}} \end{bmatrix} \begin{bmatrix} v_r \\ v_\theta \end{bmatrix}$$

where each matrix element is the projection of a unit vector of one coordinate system onto the other coordinate system. In a 2D system there are 4 possible combinations representing the 2 by 2 matrix, given by the permutation the arguments of the scalar product and their rows. We will see later that this can be generalized to 3D to obtain a similar form of transformation matrix from cylindrical to Cartesian or spherical to Cartesian but in the form of a 3 by 3 matrix.

The next step is to evaluate the dot product of each combination of unit vector. To do so, we refer to Figure 19. We find,

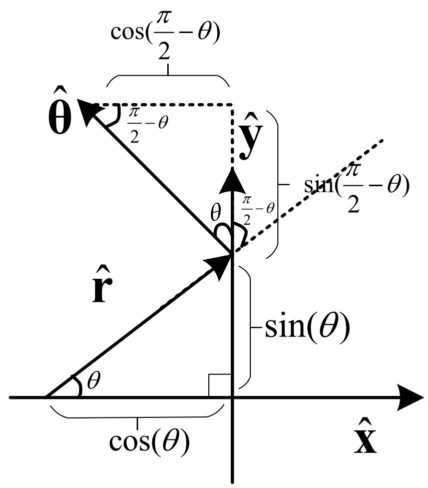


Figure 19: Relation between Cartesian and polar vectors

$$
\begin{aligned}
\hat{\mathbf{r}} \cdot \hat{\mathbf{x}} &= \cos\theta \\
\hat{\mathbf{r}} \cdot \hat{\mathbf{y}} &= \sin\theta \\
\hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{x}} &= -\cos\left(\frac{\pi}{2} - \theta\right) = -\sin\theta \\
\hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{y}} &= \sin\left(\frac{\pi}{2} - \theta\right) = \cos\theta
\end{aligned}
$$

Finally, we have,

$$
\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} v_r \\ v_\theta \end{bmatrix} \tag{23}
$$

Therefore, the transformation matrix from polar coordinates to Cartesian coordinates is

$$
\boxed{\mathbf{M} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}
$$

Finally, from (23), we deduce that

$$
\begin{aligned}
v_x &= v_r \cos\theta - v_\theta \sin\theta \\
v_y &= v_r \sin\theta + v_\theta \cos\theta
\end{aligned}
$$

There exists an alternative to this approach where we focus our effort on re-expressing the base vector of one system in terms of the base vector for the new system. For this example, we continue working on converting a polar vector into a Cartesian vector. We will discuss how this is almost equivalent to converting a Cartesian vector into a polar vector shortly after this demonstration. Our starting point is the vector in polar form which we re-state here

$$
\mathbf{v} = \mathbf{v}_{\text{pol}} = v_r \hat{\mathbf{r}} + v_\theta \hat{\boldsymbol{\theta}} \tag{24}
$$

Our approach consists of finding an expression for both $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ in terms of $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$. We first define the position vector describing the vector going from the origin to any point (x, y):

$$
\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}},
$$

and we use the polar relations (20) to re-express the coefficient of the equation in terms of (r, θ). To relate the polar base vector, we want to keep this equation in terms of the Cartesian base vectors. We now have,

$$
\mathbf{r} = r\cos\theta\hat{\mathbf{x}} + r\sin\theta\hat{\mathbf{y}}
$$

The base vectors for the polar coordinate system can be computed by taking the derivatives of the position vector with respect to a new set of variables. See Chapter 12, section 12.6 of your textbook [3] for more details. The base vectors represent the direction of an infinitesimal change in one variable as shown in Figure 21. Finally, since we work with an orthonormal system, we normalize each direction to get a pair of unit vectors.

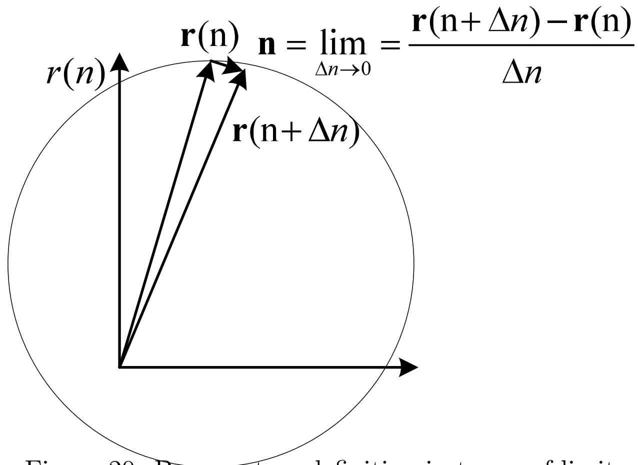


Figure 20: Base vectors definition in terms of limits

Formally, the definition of a base vector in an orthonormal system is

$$\boxed{\hat{\mathbf{n}}_i = \frac{\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}n_i}}{\left\|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}n_i}\right\|}} \tag{25}$$

where $\mathbf{r}$ is the position vector expressed in terms of variables n_i, where $i = 1, \ldots, n$ represent the indices for the individual basis vectors of the final coordinate system and n represents the number of dimensions.

For a polar coordinate system, we obtain the following expressions for the base vectors $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$:

$$\hat{\mathbf{r}} = \frac{\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}r}}{\left\|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}r}\right\|}, \tag{26}$$

$$\hat{\boldsymbol{\theta}} = \frac{\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\theta}}{\left\|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\theta}\right\|}. \tag{27}$$

These expressions are represented in Figure 21 where the derivatives are expressed in terms of limits.

We first evaluate the derivative with respect to r using (26):

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}r} = \cos\theta\hat{\mathbf{x}} + \sin\theta\hat{\mathbf{y}}$$

with a norm given by,

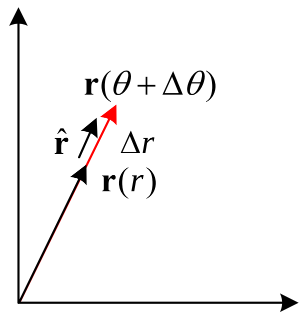
$$\left\|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}r}\right\| = \sqrt{\cos^2\theta + \sin^2\theta} = 1$$

The first orthonormal base vector is therefore,

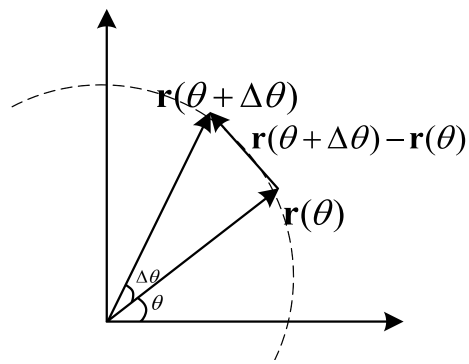
$$\hat{\mathbf{r}} = \cos\theta\hat{\mathbf{x}} + \sin\theta\hat{\mathbf{y}} = r_x\hat{\mathbf{x}} + r_y\hat{\mathbf{y}}.$$

We proceed the same way for $\hat{\boldsymbol{\theta}}$ using (27) and obtain,

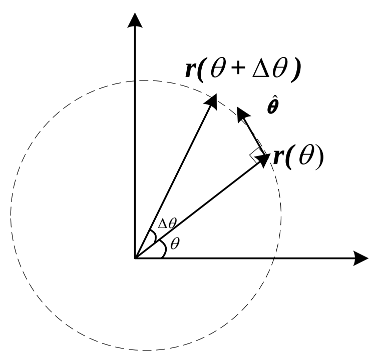
$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\theta} = -r\sin\theta\hat{\mathbf{x}} + r\cos\theta\hat{\mathbf{y}}$$



(a) $\hat{\mathbf{r}}$



(b) $\theta = \mathbf{r}(\theta + \boldsymbol{\Delta}\theta) - \mathbf{r}(\theta)$



(c) $\hat{\boldsymbol{\theta}}$

Figure 21: Definition of the base vectors $\hat{\boldsymbol{\theta}}$ in terms of limits

and with a norm given by

$$\begin{aligned} \left\| \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\theta} \right\| &= \sqrt{r^2 \sin^2\theta + r^2 \cos^2\theta} \\ &= \sqrt{r^2} \\ &= r. \end{aligned}$$

The second orthonormal base vector is therefore,

$$\hat{\boldsymbol{\theta}} = -\sin\theta \hat{\mathbf{x}} + \cos\theta \hat{\mathbf{y}} = \theta_x \hat{\mathbf{x}} + \theta_y \hat{\mathbf{y}}.$$

Now, we substitute these two relations in (24) and we get,

$$\begin{aligned} \mathbf{v} &= v_r \left(\cos\theta \hat{\mathbf{x}} + \sin\theta \hat{\mathbf{y}}\right) + v_\theta \left(-\sin\theta \hat{\mathbf{x}} + \cos\theta \hat{\mathbf{y}}\right) \\ &= \left(v_r \cos\theta - v_\theta \sin\theta\right) \hat{\mathbf{x}} + \left(v_r \sin\theta + v_\theta \cos\theta\right) \hat{\mathbf{y}} \end{aligned}$$

We re-write $\mathbf{v}$ in Cartesian coordinates as,

$$\begin{aligned} \mathbf{v}_{\text{cart}} &= v_x \hat{\mathbf{x}} + v_y \hat{\mathbf{y}} \\ &= \begin{bmatrix} v_x \\ v_y \end{bmatrix}, \end{aligned}$$

and finally obtain the the transformation matrix as given in (23),

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} v_r \\ v_\theta \end{bmatrix}$$

By inspection, we observe that

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{r}} & \hat{\boldsymbol{\theta}} \end{bmatrix} \begin{bmatrix} v_r \\ v_\theta \end{bmatrix}$$

and hence,

$$\boxed{\mathbf{M} = \begin{bmatrix} \hat{\mathbf{r}}_x & \hat{\boldsymbol{\theta}}_x \\ \hat{\mathbf{r}}_y & \hat{\boldsymbol{\theta}}_y \end{bmatrix}}$$

where $\hat{\mathbf{r}}_x$, $\hat{\mathbf{r}}_y$ and $\hat{\boldsymbol{\theta}}_x$, $\hat{\boldsymbol{\theta}}_y$ are the x,y component of $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ respectively and form $\mathbf{M}$, the transformation matrix. We can generalize this approach for any orthonormal coordinate system in any dimension. The general transformation matrix has the following form,

$$\boxed{\mathbf{M} = \begin{bmatrix} | & | & \dots & | \\ \hat{\mathbf{n}}_1 & \hat{\mathbf{n}}_2 & \dots & \hat{\mathbf{n}}_n \\ | & | & \dots & | \end{bmatrix}}$$

In summary, to convert a vector from one set of coordinates to another, we need to find the transformation matrix $\mathbf{M}$. Once this is done, we use matrix multiplication to obtain the vector in the desired form.

This is, however, only one of many transformations that we could think of. One might be required to convert a Cartesian vector to a polar vector. The procedure is to recall (21) and solve for $\mathbf{v}_{\text{cart}}$:

$$\begin{aligned} \mathbf{v}_{\text{cart}} &= \mathbf{M}\mathbf{v}_{\text{pol}} \\ \mathbf{M}^{-1}\mathbf{v}_{\text{cart}} &= \mathbf{M}^{-1}\mathbf{M}\mathbf{v}_{\text{pol}} \\ \mathbf{v}_{\text{pol}} &= \mathbf{M}^{-1}\mathbf{v}_{\text{cart}} \end{aligned} \tag{28}$$

At this point it is useful to note the following property for $\mathbf{M}$. The transformation matrix between two orthonormal coordinate systems is orthogonal. An orthogonal matrix means that each of its columns represent the orthonormal base vectors. From previous courses you may recall that the following condition is satisfied for an orthonormal matrix: $\mathbf{M}\mathbf{M}^{\text{T}} = \mathbf{I}$. This expression implies $\mathbf{M}^{\text{T}} = \mathbf{M}^{-1}$. Hence we can retrieve the polar vector using the transpose of the matrix $\mathbf{M}$ that we developed before,

$$\mathbf{v}_{\text{pol}} = \mathbf{M}^{\text{T}}\mathbf{v}_{\text{cart}}$$

Finally, let us obtain one more useful characterization of the orthogonality of $\mathbf{M}$ by looking at the determinant of $\mathbf{M}$. We recall that for any matrix $\mathbf{A} \in \mathbb{R}^{p\times p}$, $\det\mathbf{A} = \det\mathbf{A}^{\text{T}}$. We note that $\det\mathbf{I} = \det\left(\mathbf{M}^{\text{T}}\mathbf{M}\right) = \det\left(\mathbf{M}^{\text{T}}\right)\det(\mathbf{M}) = (\det\mathbf{M})^2$ where the determinant of the identity matrix is 1. $\mathbf{M}$ being orthogonal also implies that its determinant is ± 1. Then, computing the determinant of the matrix allows one to confirm that there was no computation error. However, if the determinant is ± 1, then it does not imply that we have the correct matrix.

Let's conclude our discussion on the transformation matrix by proving that the determinant of $\mathbf{M}$ is 1. We have,

$$\begin{aligned} \det\mathbf{M} &= \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} \\ &= \cos^2\theta + \sin^2\theta \\ &= 1 \end{aligned}$$

4.2.3 Differential elements

The next important topic to discuss when introducing a new coordinate set is how to define the differential length $\mathrm{d}\boldsymbol{\ell}$, a vector and the surface area $\mathrm{d}S$, a scalar, both in 2D. This is obtained by looking at an infinitesimal change along the directions of each base vector. Formally,

$$\mathrm{d}\boldsymbol{\ell} = \mathrm{d}\ell_r\hat{\mathbf{r}} + \mathrm{d}\ell_\theta\hat{\boldsymbol{\theta}},$$
$$\mathrm{d}S = \mathrm{d}\ell_r\mathrm{d}\ell_\theta.$$

The definition of differential elements for a polar coordinate system is represented in Figure 22. With

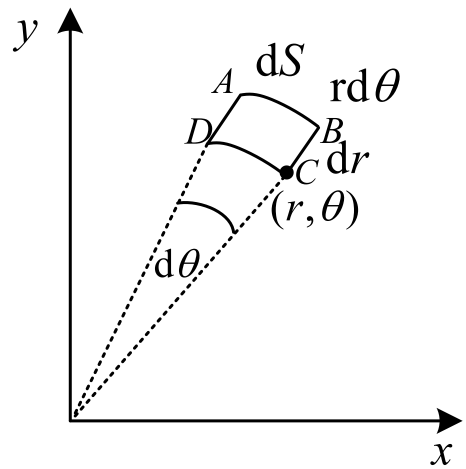


Figure 22: Differential elements in a polar coordinate system

reference to Figure 22, we define the differential surface as the area $ABCD$ and compute the differential length as

$$\mathrm{d}\ell_r = \mathrm{d}r,$$
$$\mathrm{d}\ell_\theta = r\mathrm{d}\theta.$$

We note here one important difference between the Cartesian differential lengths and $\mathrm{d}\ell_\theta$ and also between $\mathrm{d}\ell_r$ and $\mathrm{d}\ell_\theta$. We are interested in length, something that we could measure in meters. However, an infinitesimal change along the $\hat{\boldsymbol{\theta}}$ direction has a magnitude $\mathrm{d}\theta$ radian (or degree) and hence has units that are not in meters as any length. This represents the infinitesimal angle change going from the line defined as $\theta = \theta_0$ to the line defined as $\theta = \theta_0 + \mathrm{d}\theta$. To retrieve the length of the CD segment in Figure 22, we compute the arc length given by $r\mathrm{d}\theta$, that is the radius of the circle created by the arc segment times the angle in radians. Finally, the differential length is

$$\boxed{\mathrm{d}\ell = \mathrm{d}r\hat{\mathbf{r}} + r\mathrm{d}\theta\hat{\boldsymbol{\theta}}.}$$

The differential surface area element is then

$$\boxed{\mathrm{d}S = r\mathrm{d}\theta\mathrm{d}r.}$$

Later in this course we will determine the differential surface area for an arbitrary coordinate system using the concept of a Jacobian.

5 3D coordinate systems

We now expand our space to three dimensions (3D) and present three common coordinate systems. The Cartesian and cylindrical will first be introduced as an extension to the previously described Cartesian and polar coordinate sets in 2D. Finally, we will conclude with the spherical coordinate system. In this section, we build on the previous section and several demonstrations will start from previously stated results.

5.1 Cartesian

5.1.1 Definition

In 3D, a position can be described by the intersection of three planar surfaces rather than by two straight lines. The Cartesian coordinate system is the most common coordinate set and we are used to referring to a position in terms of (x, y, z). In this case, a position $P(x_0, y_0, z_0)$ is described by three planes: $x = x_0$, $y = y_0$ and $z = z_0$ as shown in Figure 23 where $x_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}$ and $z_0 \in \mathbb{R}$. The intersection of three planes defines a point in 3D space. The x and y part are identical to the 2D case and are respectively perpendicular to the yz and xz plane. The base vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are also consistent with the 2D case. Now in 3D, we add a third dimension perpendicular to the plane formed by the xy axis. This is the z dimension.

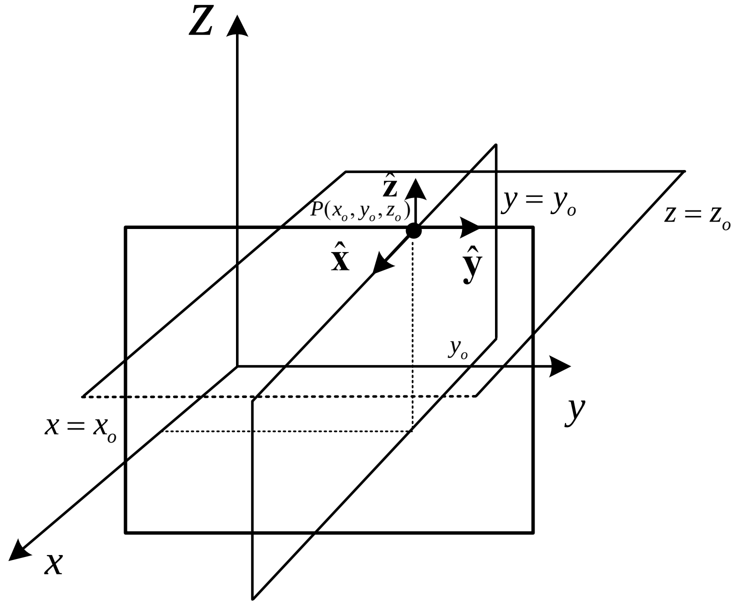


Figure 23: 3D Cartesian coordinate system

The three base vectors are unit vectors perpendicular to their respective surfaces. The base vectors are represented in Figure 23. When compared to the 2D case, we added $\hat{\mathbf{z}}$ perpendicular to both $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$.

One thing to notice here again is the direction of the base vectors. The perpendicularity condition tells us that, according to Figure 23, $\hat{\mathbf{x}}$ and $-\hat{\mathbf{x}}$ are valid base vectors along the x axis and for $\pm\hat{\mathbf{y}}$ and $\pm\hat{\mathbf{z}}$ for the y axis and z axis respectively. There is an ambiguity in choosing a direction for the base vectors so we must impose a rule which we refer to as a right-handed system. There is also a left handed rule but it is infrequently used and must be stated up front. In a right handed system the base vectors satisfy the following rules:

$$\hat{\mathbf{x}} \times \hat{\mathbf{y}} = \hat{\mathbf{z}},$$
$$\hat{\mathbf{y}} \times \hat{\mathbf{z}} = \hat{\mathbf{x}},$$
$$\hat{\mathbf{z}} \times \hat{\mathbf{x}} = \hat{\mathbf{y}}.$$

You can check by yourselves that the base vectors and axes in Figure 23 respect this condition by using the right-hand rule. Note, in sketching a problem you may rotate the axes to make drawing easier and you are free to label the axes as you wish but make sure that the rules above are satisfied.

Then, any vector in 3D can be expressed in a Cartesian coordinate system as,

$$\boxed{\mathbf{v} = v_x\hat{\mathbf{x}} + v_y\hat{\mathbf{y}} + v_z\hat{\mathbf{z}}.}$$

The coefficient v_x, v_y and v_x are the projection of $\mathbf{v}$ along the $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ directions respectively. Figure 24 shows the decomposition of the vector $\mathbf{v} = 4\hat{\mathbf{x}} + 6\hat{\mathbf{y}} + 3\hat{\mathbf{z}}$ into each base vector.

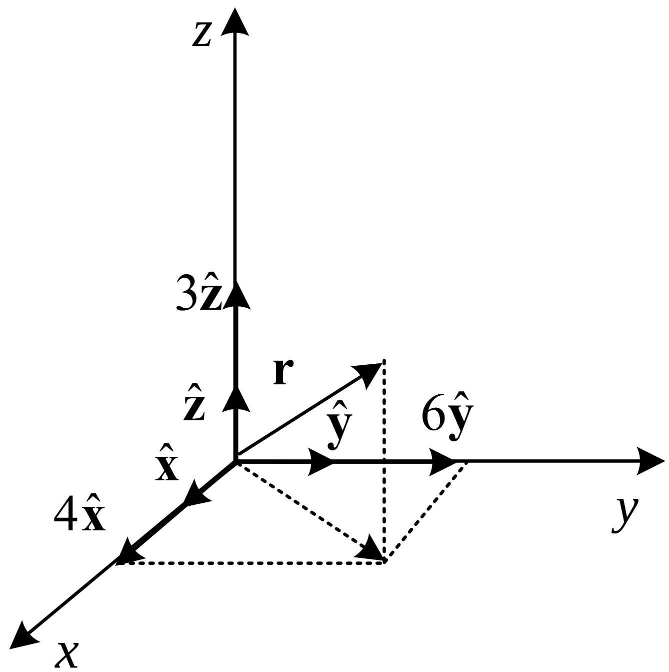


Figure 24: 3D Cartesian coordinate system

5.1.2 Differential elements

The differential elements in 3D Cartesian coordinates are similar to the ones previously stated in the 2D case. We simply add one other infinitesimal change in the z direction. The differential elements for the 3D Cartesian case are given in Figure 25a.

We recall the definition of the differential length from the previous section. The differential surface in 3D, contrary to the 2D case, is oriented which means it has a direction associated with it. We define the orientation of a surface (and differential surface) as the direction of its normal: pointing out of a surface or pointing into a surface. This leaves an ambiguity so in any given problem one must understand the problem being solved. We can now summarize the differential elements in 3D considering the general case where the directions of the base vectors are $\hat{\mathbf{n}}_1$, $\hat{\mathbf{n}}_2$ and $\hat{\mathbf{n}}_3$. We then consider the specific form of these expressions for the Cartesian coordinate system.

$$\mathrm{d}\boldsymbol{\ell} = \mathrm{d}\ell_1\hat{\mathbf{n}}_1 + \mathrm{d}\ell_2\hat{\mathbf{n}}_2 + \mathrm{d}\ell_3\hat{\mathbf{n}}_3 \tag{29}$$

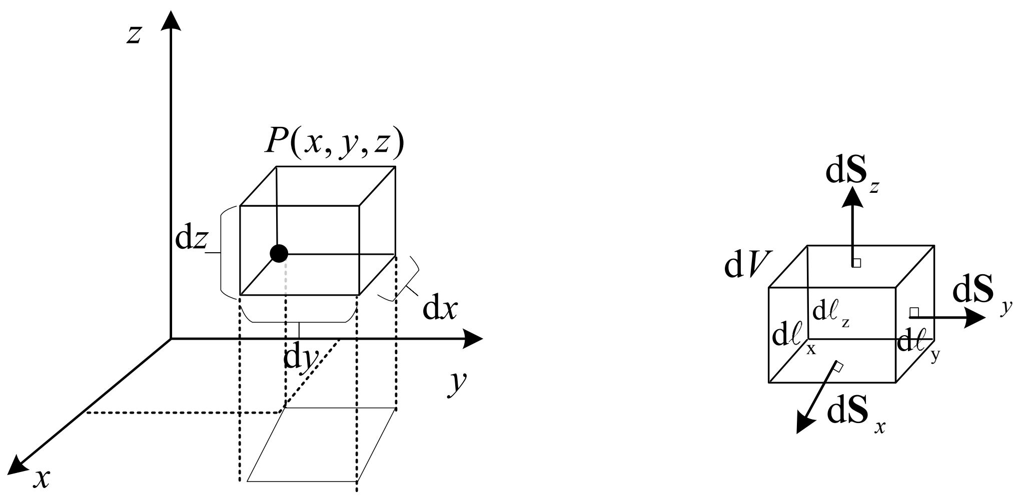
$$\mathrm{d}\mathbf{S}_1 = \mathrm{d}\ell_2\mathrm{d}\ell_3\hat{\mathbf{n}}_1 \tag{30}$$

$$\mathrm{d}\mathbf{S}_2 = \mathrm{d}\ell_1\mathrm{d}\ell_3\hat{\mathbf{n}}_2 \tag{31}$$

$$\mathrm{d}\mathbf{S}_3 = \mathrm{d}\ell_1\mathrm{d}\ell_2\hat{\mathbf{n}}_3 \tag{32}$$

$$\mathrm{d}V = \mathrm{d}\ell_1\mathrm{d}\ell_2\mathrm{d}\ell_3 \tag{33}$$

In Cartesian coordinates we draw an infinitesimal rectangular box. Each face of the box is an oriented surface. The three visible surfaces have the following differential surface areas and orientations and are shown in Figure 25b: $\mathrm{d}\mathbf{S}_x = \mathrm{d}S_x\hat{\mathbf{x}}$, $\mathrm{d}\mathbf{S}_y = \mathrm{d}S_y\hat{\mathbf{y}}$ and $\mathrm{d}\mathbf{S}_z = \mathrm{d}S_z\hat{\mathbf{z}}$.



(a) Differential elements due to infinitesimal change

(b) Zoom on differential elements

Figure 25: Differential elements in 3D Cartesian

The plane representing $\mathrm{d}\mathbf{S}_x$, $\mathrm{d}\mathbf{S}_y$ and $\mathrm{d}\mathbf{S}_z$ can be constructed using the bases vectors $(\hat{\mathbf{y}}, \hat{\mathbf{z}})$, $(\hat{\mathbf{x}}, \hat{\mathbf{z}})$ and $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ respectively. So for example, the plane xy is defined by a vector which has the form of $a\hat{\mathbf{x}} + b\hat{\mathbf{y}}$. Every point in the plane can be reached by choosing a and b an element of $\mathbb{R}$. Similarly, we obtain the plane yz and xz with a vector of the form $c\hat{\mathbf{y}} + d\hat{\mathbf{z}}$ and $e\hat{\mathbf{x}} + f\hat{\mathbf{z}}$ respectively where c, d, e, f are elements of $\mathbb{R}$. The planes represented by $\mathrm{d}\mathbf{S}_x$, $\mathrm{d}\mathbf{S}_y$ and $\mathrm{d}\mathbf{S}_z$ are perpendicular to $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$ respectively.

Lastly, in 3D we now have the differential volume which can be inferred from Figure 25b.

As shown in Figure 25, the differential length along the three directions is $\mathrm{d}\ell_x = \mathrm{d}x$, $\mathrm{d}\ell_y = \mathrm{d}y$ and $\mathrm{d}\ell_z = \mathrm{d}z$. Using definitions (29)-(33), the differential elements in the 3D Cartesian case are:

$$\boxed{\begin{aligned} \mathrm{d}\boldsymbol{\ell} &= \mathrm{d}x\hat{\mathbf{x}} + \mathrm{d}y\hat{\mathbf{y}} + \mathrm{d}z\hat{\mathbf{z}} \\ \mathrm{d}\mathbf{S}_x &= \mathrm{d}y\mathrm{d}z\hat{\mathbf{x}} \\ \mathrm{d}\mathbf{S}_y &= \mathrm{d}x\mathrm{d}z\hat{\mathbf{y}} \\ \mathrm{d}\mathbf{S}_z &= \mathrm{d}x\mathrm{d}y\hat{\mathbf{z}} \\ \mathrm{d}V &= \mathrm{d}x\mathrm{d}y\mathrm{d}z. \end{aligned}}$$

5.2 Cylindrical

5.2.1 Definition

The cylindrical coordinate system is the polar coordinate system to which we add the same z axis as used in the 3D Cartesian coordinate system to describe the added dimension. The three surfaces used to determine the intersection point are first the cylinder of radius $r = r_0$ (we can see this as a circle for every value of z, hence a cylinder), the plane $\theta = \theta_0$ and the plane $z = z_0$. We recall that θ is defined as the angle between the positive x axis and and a line passing through the origin on the xy plane. The intersection of the three surfaces can describe any point $P(r_0, \theta_0, z_0)$ where $r_0 \in [0, +\infty[$, $\theta_0 \in [0, 2\pi]$ and $z_0 \in \mathbb{R}$, as presented in Figure 26. We note that by setting $z = 0$ we retrieve the polar coordinate system.

The base vectors are given by the unit vector perpendicular to each surface at a given point and are also represented in Figure 26. For a given point, each surface is perpendicular to another and hence the system is orthonormal. Once again, the base vector can point in one of two directions and we pick the direction that leads to a right-handed system.

Using the right-hand and the ordering $\{\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{z}}\}$, we then set the $\hat{\mathbf{r}}$ direction as outward similarly to polar coordinates. Then, retaining $\hat{\mathbf{z}}$ as in Cartesian coordinates, the right-hand rule instructs us that the angle

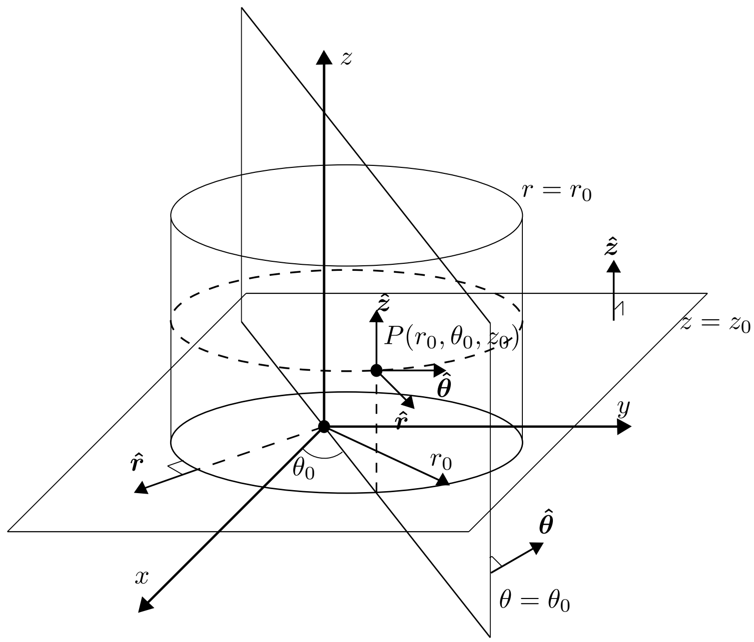


Figure 26: Cylindrical coordinate system

direction $\hat{\boldsymbol{\theta}}$ for the unit vector is counterclockwise. As a last step, we can verify that the base vectors presented in Figure 26 make a right-handed system. Alternatively, we can check if the system is right-handed by using (10). Finally, any vector $\mathbf{v} \in \mathbb{R}^3$ can be expressed in cylindrical coordinates as,

$$\boxed{\mathbf{v} = v_r\hat{\mathbf{r}} + v_\theta\hat{\boldsymbol{\theta}} + v_z\hat{\mathbf{z}}.}$$

5.2.2 Transformation

The next point of interest is how to to convert a position $P(r, \theta, z)$ and a vector $\mathbf{v} = v_r\hat{\mathbf{r}} + v_\theta\hat{\boldsymbol{\theta}} + v_z\hat{\mathbf{z}}$ to Cartesian coordinates. The position in a cylindrical coordinate system can be related to a position in a Cartesian coordinate system according to the following relations:

$$\boxed{\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned}}$$

or inversely, to convert from a cylindrical coordinate system to a Cartesian coordinate system:

$$\boxed{\begin{aligned} r &= \sqrt{x^2 + y^2} \\ \theta &= \arctan\frac{y}{x} \\ z &= z \end{aligned}}$$

These sets of equations are obtained the same way as in polar coordinates to which we then add the z coordinate. This z coordinate is equivalent to the Cartesian z since both axes are defined the same way.

To convert a vector from one coordinate system into another, we need to compute the transformation matrix with one of the two approaches described in Section 4.2.2. In this case, we use the first approach. Recall that we are looking for $\mathbf{M}$ such that

$$\mathbf{v}_{\text{cart}} = \mathbf{M}\mathbf{v}_{\text{cyl}}$$

We expand all terms of the matrix $\mathbf{M}$ and we have,

$$\begin{bmatrix} v_x \\ v_y \\ v_x \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} v_r \\ v_\theta \\ v_z \end{bmatrix}$$

We previously defined M_{ij} as the projection of the j^{th} original system base vector onto the i^{th} base vector of the desired system. Hence, we have

$$\mathbf{M} = \begin{bmatrix} \hat{\mathbf{r}} \cdot \hat{\mathbf{x}} & \hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{x}} & \hat{\mathbf{z}} \cdot \hat{\mathbf{x}} \\ \hat{\mathbf{r}} \cdot \hat{\mathbf{y}} & \hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{y}} & \hat{\mathbf{z}} \cdot \hat{\mathbf{y}} \\ \hat{\mathbf{r}} \cdot \hat{\mathbf{z}} & \hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{z}} & \hat{\mathbf{z}} \cdot \hat{\mathbf{z}} \end{bmatrix}.$$

The matrix elements M_{11}, M_{12}, M_{21} and M_{22} are exactly the same as in polar coordinates since the cylindrical and polar coordinate system share the same $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ vectors. $\hat{\mathbf{z}}$ is perpendicular to the xy-plane and hence to both $\hat{\mathbf{x}}$ and $\hat{\boldsymbol{\theta}}$ and thus the cylindrical coordinate system we defined is orthonormal. We thus have,

$$\begin{aligned} \hat{\mathbf{z}} \cdot \hat{\mathbf{x}} &= 0, \\ \hat{\mathbf{z}} \cdot \hat{\mathbf{y}} &= 0, \\ \hat{\mathbf{z}} \cdot \hat{\mathbf{z}} &= 1, \\ \hat{\mathbf{r}} \cdot \hat{\mathbf{z}} &= 0, \\ \hat{\boldsymbol{\theta}} \cdot \hat{\mathbf{z}} &= 0. \end{aligned}$$

The transformation matrix from a cylindrical coordinate system to Cartesian coordinate system is therefore,

$$\mathbf{M} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Using the transpose of $\mathbf{M}$, we can then convert a Cartesian vector into a cylindrical vector.

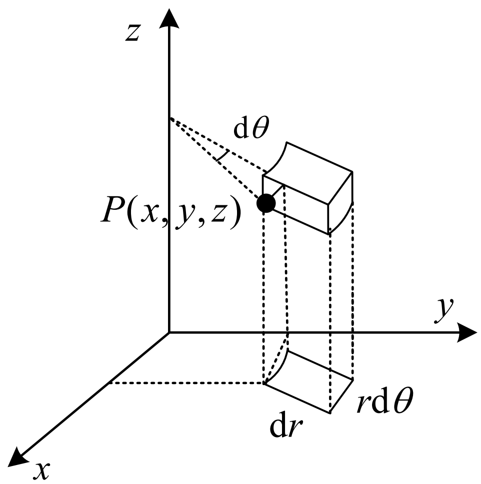
***Problem A*:** Use the second approach of Section 4.2.2 to find the transformation matrix $\mathbf{M}$ from a cylindrical coordinate system to a Cartesian coordinate system.

5.2.3 Differential elements

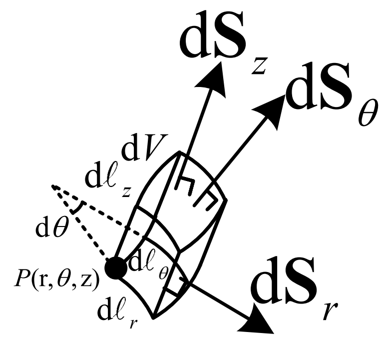
The differential elements for cylindrical coordinates are represented in Figure 27. Using this figure, we find $\mathrm{d}\ell_r = \mathrm{d}r$ and $\mathrm{d}\ell_z = \mathrm{d}z$. For $\mathrm{d}\ell_\theta$, it is once again given by the arc length rather than by the change in angle, hence we have $\mathrm{d}\ell_\theta = r\mathrm{d}\theta$. We use the differential element definitions given in (29)-(33) and substitute the differential length along each direction. We will show a more general approach for computing the differential surface area element and differential volume element using the concept of the Jacobian later in this course.

The differential elements in a cylindrical coordinate system are:

$$\boxed{\begin{aligned} \mathrm{d}\boldsymbol{\ell} &= \mathrm{d}r\hat{\mathbf{r}} + r\mathrm{d}\theta\hat{\boldsymbol{\theta}} + \mathrm{d}z\hat{\mathbf{z}} \\ \mathrm{d}\mathbf{S}_r &= r\mathrm{d}\theta\mathrm{d}z\hat{\mathbf{r}} \\ \mathrm{d}\mathbf{S}_\theta &= \mathrm{d}r\mathrm{d}z\hat{\boldsymbol{\theta}} \\ \mathrm{d}\mathbf{S}_z &= r\mathrm{d}r\mathrm{d}\theta\hat{\mathbf{z}} \\ \mathrm{d}V &= r\mathrm{d}r\mathrm{d}\theta\mathrm{d}z. \end{aligned}}$$



(a) Differential elements due to infinitesimal change



(b) Zoom on differential elements

Figure 27: Differential elements in cylindrical

5.3 Spherical

5.3.1 Definition

We conclude our survey of coordinates systems with spherical coordinates. In a spherical coordinate system, we use the following three surfaces to describe a position. First we generate a spherical shell $\rho = \rho_0$ which represents all points located at a distance equal to ρ_0 from the origin. Second, we generate a cone $\varphi = \varphi_0$ where φ is defined as the direction that takes one from the z axis to the surface of the cone and where the apex of the cone is placed at the origin. Finally, the plane $\theta = \theta_0$ is defined in the same way as for cylindrical coordinates that is an angle in the xy plane and with respect to the x axis. The intersection of these three surfaces represents a position $(\rho_0, \varphi_0, \theta_0)$ in a 3D spherical coordinate system where $\rho_0 \in [0, +\infty[$, $\varphi \in [0, \pi]$ and $\theta_0 \in [0, 2\pi]$. Note that φ is only defined between 0 and π since any value between π and 2π can be retrieved with values of θ greater than π. The spherical coordinate system is presented in Figure 28.

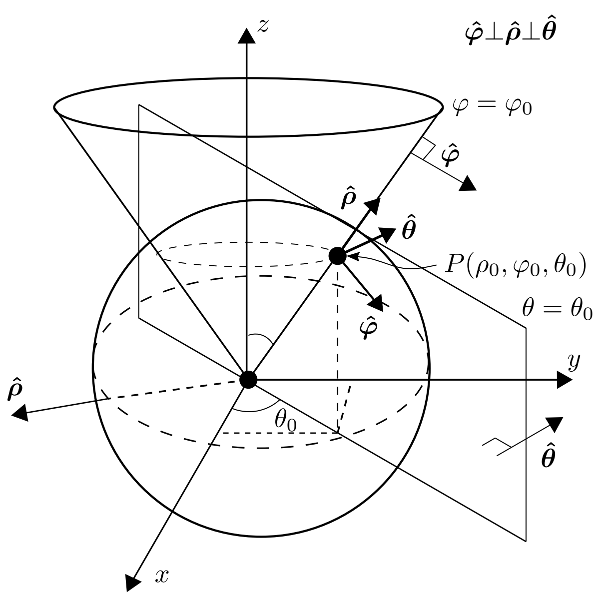


Figure 28: Spherical coordinate system

The vectors perpendicular to the cone, the spherical surface and plane are all perpendicular to each other at a given point and hence the vectors are orthogonal to each other. The base vectors are given by the three unit vectors which are perpendicular to their respective surfaces. We set the orientation of the base vectors to have a right-handed system using the ordered sequence $\{\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\varphi}}, \hat{\boldsymbol{\theta}}\}$. We set $\hat{\boldsymbol{\rho}}$ to be pointing out so that the ρ component increases when going further away from the origin. Once again, ρ is set this way since it reflects the distance from a point to the origin. Then, the azimuthal angle $\hat{\boldsymbol{\theta}}$, in the xy plane, is taken in the counterclockwise direction as in cylindrical coordinates. Finally, we use the right-hand to set the direction of the elevation angle $\hat{\boldsymbol{\varphi}}$. To do so, we place our index finger along the $\hat{\boldsymbol{\rho}}$ direction and our thumb in the $\hat{\boldsymbol{\theta}}$ direction. The middle finger reveals the direction of $\hat{\boldsymbol{\varphi}}$. This direction is given in Figure 28.

Note that the radius ρ is now the radius in 3D space instead of being the radial distance between a point and the z-axis for a given z in a cylindrical coordinate system. The associated base vector also differs. $\hat{\boldsymbol{\rho}}$ is radial, pointing outward from the origin whereas $\hat{\mathbf{r}}$ for a cylindrical coordinate system is pointing outward from the z axis. Notice also the difference between the radial vectors for the two cases: $\hat{\mathbf{r}}$ is perpendicular to $\hat{\mathbf{z}}$ and $\hat{\boldsymbol{\rho}}$ is not. Finally, $\hat{\boldsymbol{\theta}}$ and the θ component are defined the same way as in cylindrical coordinates. Finally, any vector $\mathbf{v} \in \mathbb{R}^3$ can be expressed as

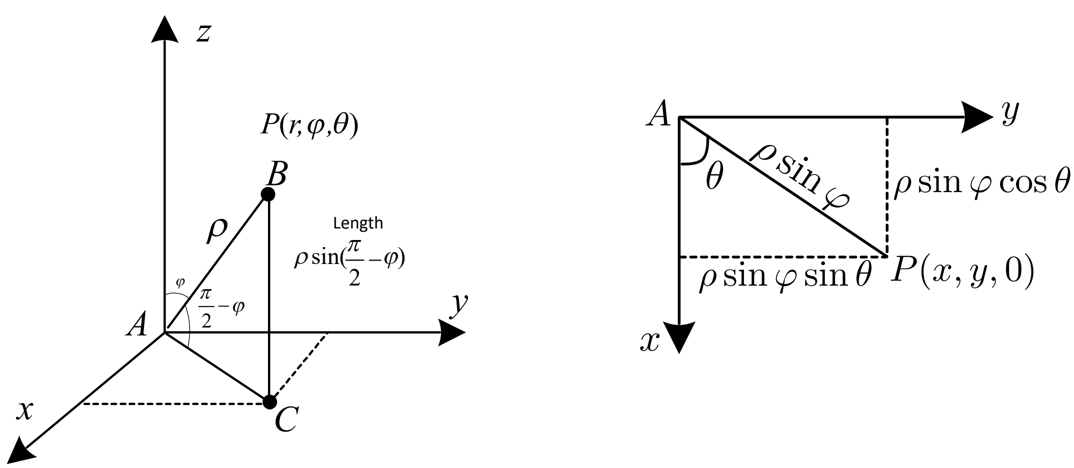
$$\boxed{\mathbf{v} = v_\rho \hat{\boldsymbol{\rho}} + v_\varphi \hat{\boldsymbol{\varphi}} + v_\theta \hat{\boldsymbol{\theta}}.}$$

where v_ρ, v_φ and v_θ are the coefficients given by the projection of $\mathbf{v}$ along their respective base vectors.

5.3.2 Transformation

We now work on a relation between spherical coordinates and Cartesian coordinates. We first look for the relation between a position in the two coordinate systems.

Let's consider a point $P(\rho, \varphi, \theta)$ as presented in Figure 29. Starting with the relation for z, we trace the triangle ABC given in Figure 29a. The z component of the point is the height of the right-angled triangle with angle BAC equal to $\frac{\pi}{2} - \varphi$. Hence, we have $z = \rho \sin\left(\frac{\pi}{2} - \varphi\right) = \rho \cos\varphi$. It also follows that the base of the triangle is $\rho \sin\varphi$. We then look at the point P' to compute the relations for x and y as illustrated in Figure 29b. Notice that the distance from the origin to P' is the base of the triangle used for z. Hence, using basic trigonometry we find the relations between the position in different coordinates.



(a) z component

(b) x and y components

Figure 29: Relation between x, y, z and ρ, φ and θ

$$\boxed{\begin{aligned} x &= \rho\cos\theta\sin\varphi && (34)\\ y &= \rho\sin\theta\sin\varphi && (35)\\ z &= \rho\cos\varphi && (36)\end{aligned}}$$

We can invert these relations and obtain the relationships that allow us to convert from a Cartesian coordinate system to a spherical coordinate system:

$$\boxed{\begin{aligned} \rho &= \sqrt{x^2+y^2+z^2}\\ \varphi &= \arctan\frac{\sqrt{x^2+y^2}}{z}\\ \theta &= \arctan\frac{y}{x}\end{aligned}}$$

To convert a vector from a spherical coordinate system to a Cartesian coordinate system, we proceed the same way as before. We look for the transformation matrix $\mathbf{M}$ such that

$$\mathbf{v}_{\text{cart}} = \mathbf{M}\mathbf{v}_{\text{sph}}$$

or equivalently,

$$\begin{bmatrix} v_x\\ v_y\\ v_x\end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13}\\ M_{21} & M_{22} & M_{23}\\ M_{31} & M_{32} & M_{33}\end{bmatrix}\begin{bmatrix} v_\rho\\ v_\varphi\\ v_\theta\end{bmatrix}$$

We will use the second approach that was derived in Section 4.2.2 to compute $\mathbf{M}$. The $\mathbf{M}$ matrix is therefore given by,

$$\mathbf{M} = \begin{bmatrix} | & | & |\\ \hat{\boldsymbol{\rho}} & \hat{\boldsymbol{\varphi}} & \hat{\boldsymbol{\theta}}\\ | & | & |\end{bmatrix},$$

where we find the unit base vectors found using (25). The base vectors are

$$\boxed{\begin{aligned} \hat{\boldsymbol{\rho}} &= \frac{\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\rho}}{\left\|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\rho}\right\|} && (37)\\ \hat{\boldsymbol{\varphi}} &= \frac{\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\varphi}}{\left\|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\varphi}\right\|} && (38)\\ \hat{\boldsymbol{\theta}} &= \frac{\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\theta}}{\left\|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\theta}\right\|} && (39)\end{aligned}}$$

We recall the position vector in a Cartesian coordinate system as being $\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$ and re-express it in terms of spherical coordinate variables using the relations given in (36). Remember that we want to relate the spherical base vectors to the Cartesian base vectors. The position vector is,

$$\mathbf{r} = \rho\cos\theta\sin\varphi\hat{\mathbf{x}} + \rho\sin\theta\sin\varphi\hat{\mathbf{y}} + \rho\cos\varphi\hat{\mathbf{z}}.$$

We then evaluate (37)-(39) and compute the following derivatives.

$$\begin{aligned} \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\rho} &= \cos\theta\sin\varphi\hat{\mathbf{x}} + \sin\theta\sin\varphi\hat{\mathbf{y}} + \cos\varphi\hat{\mathbf{z}},\\ \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\varphi} &= \rho\cos\theta\cos\varphi\hat{\mathbf{x}} + \rho\sin\theta\cos\varphi\hat{\mathbf{y}} - \rho\sin\varphi\hat{\mathbf{z}},\\ \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\theta} &= -\rho\sin\theta\sin\varphi\hat{\mathbf{x}} + \rho\cos\theta\sin\varphi\hat{\mathbf{y}}.\end{aligned}$$

Their norms are:

$$
\begin{aligned}
\left\|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\rho}\right\| &= \sqrt{\cos^2\theta\sin^2\varphi + \sin^2\theta\sin^2\varphi + \cos^2\varphi} \\
&= \sqrt{\sin^2\varphi\varphi + \cos^2\varphi} \\
&= 1 \\
\left\|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\varphi}\right\| &= \sqrt{\rho^2\cos^2\varphi + \rho^2\sin^2\varphi} \\
&= \rho \\
&= \sqrt{\rho^2\cos^2\theta\cos^2\varphi + \rho^2\sin^2\theta\cos^2\varphi + \rho^2\sin^2\varphi} \\
\left\|\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}\theta}\right\| &= \sqrt{\rho^2\sin^2\theta\sin^2\varphi + \rho^2\cos^2\theta\sin^2\varphi} \\
&= \sqrt{\rho^2\sin^2\varphi} \\
&= \rho\sin\varphi
\end{aligned}
$$

Finally, the spherical unit base vectors in terms of Cartesian unit base vectors are:

$$
\begin{aligned}
\hat{\boldsymbol{\rho}} &= \cos\theta\sin\varphi\hat{\mathbf{x}} + \sin\theta\sin\varphi\hat{\mathbf{y}} + \cos\varphi\hat{\mathbf{z}}, \\
\hat{\boldsymbol{\varphi}} &= \cos\theta\cos\varphi\hat{\mathbf{x}} + \sin\theta\cos\varphi\hat{\mathbf{y}} - \sin\varphi\hat{\mathbf{z}}, \\
\hat{\boldsymbol{\theta}} &= -\sin\theta\hat{\mathbf{x}} + \cos\theta\hat{\mathbf{y}}.
\end{aligned}
$$

If we express the above equation in matrix form we are representing

$$
\mathbf{M}^{\mathrm{T}} = \begin{bmatrix} \cos\theta\sin\varphi & \sin\theta\sin\varphi & \cos\varphi \\ \cos\theta\cos\varphi & \sin\theta\cos\varphi & -\sin\varphi \\ -\sin\theta & \cos\theta & 0 \end{bmatrix},
$$

since the elements of the matrix are orthogonal. To recover $\mathbf{M}$ we need only take the transpose of the matrix $\mathbf{M}^{\mathrm{T}}$ to obtain $\mathbf{M}$. The transformation matrix to convert spherical coordinate vectors to Cartesian coordinate vectors is then,

$$
\boxed{\mathbf{M} = \begin{bmatrix} \cos\theta\sin\varphi & \cos\theta\cos\varphi & -\sin\theta \\ \sin\theta\sin\varphi & \sin\theta\cos\varphi & \cos\theta \\ \cos\varphi & -\sin\varphi & 0 \end{bmatrix},}
$$

The same result will be obtained if one uses the first approach and computes all combinations of dot products. Once again, this matrix is orthogonal and its transpose can be used to convert a Cartesian vector into a spherical vector.

We could now ask ourselves, what about converting a spherical vector into a cylindrical vector? Finding the transformation matrix to convert a vector in spherical coordinates to cylindrical coordinates is tedious since the projection or derivatives used in computing $\mathbf{M}$ are not straightforward to compute. However, we can make use of the transformation matrices we derived to convert spherical coordinate and cylindrical vectors to Cartesian coordinate vectors as a starting point. Denote $\mathbf{M}_s$ and $\mathbf{M}_c$ as the transformation matrix from spherical to Cartesian and from cylindrical to Cartesian respectively. We have,

$$
\mathbf{v}_{\mathrm{cart}} = \mathbf{M}_s\mathbf{v}_{\mathrm{sph}} \tag{40}
$$

$$
\mathbf{v}_{\mathrm{cart}} = \mathbf{M}_c\mathbf{v}_{\mathrm{cyl}} \tag{41}
$$

$$
\tag{42}
$$

We invert (41) and obtain,

$$
\mathbf{v}_{\mathrm{cyl}} = \mathbf{M}_c^{\mathrm{T}}\mathbf{v}_{\mathrm{cart}}
$$

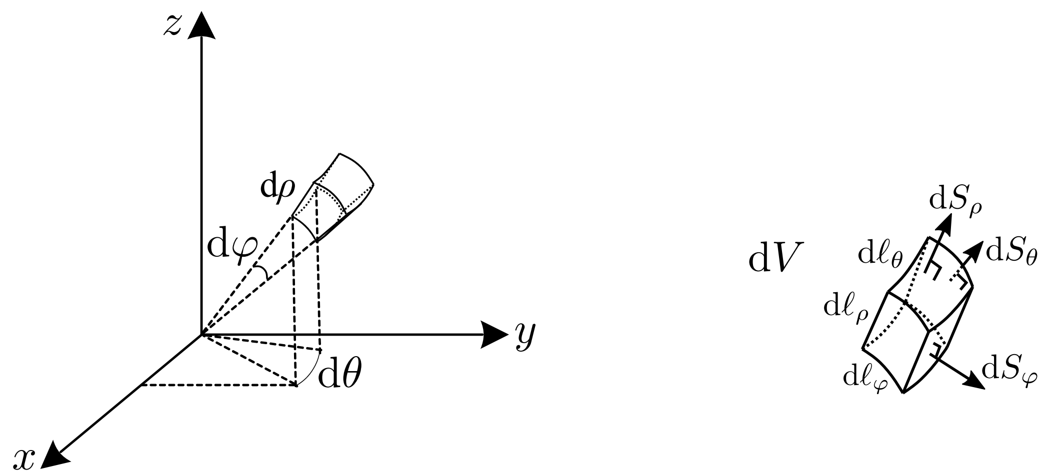
We then substitute the expression $\mathbf{v}_{\mathrm{cart}}$ from (40) into this equation.. Therefore, $\mathbf{v}_{\mathrm{cyl}}$ can be expressed in the form of the following relationship,

$$
\mathbf{v}_{\mathrm{cyl}} = \mathbf{M}_c^{\mathrm{T}}\mathbf{M}_s\mathbf{v}_{\mathrm{sph}}.
$$

***Problem B*:** Use the first approach of Section 4.2.2 to find the transformation matrix $\mathbf{M}$ from a spherical coordinate system to a Cartesian coordinate system.

5.3.3 Differential elements

The volume generated by an infinitesimal change along along the direction of the base vectors is represented in Figure 4. The first step to determine the differential elements is to compute $\mathrm{d}\ell_\rho$, $\mathrm{d}\ell_\varphi$ and $\mathrm{d}\ell_\theta$ as defined in Figure 30b.

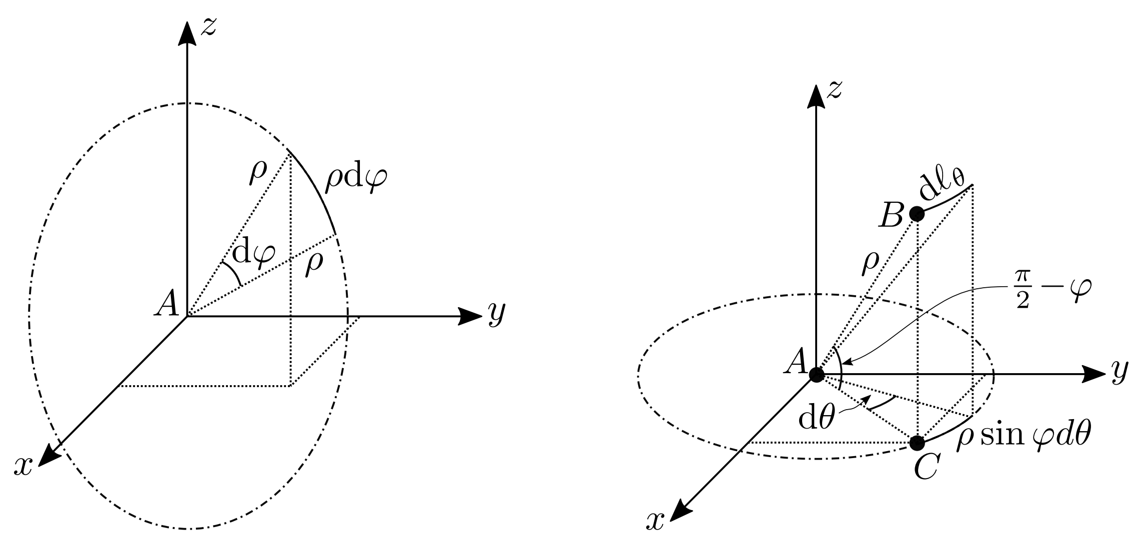


(a) Differential elements due to infinitesimal change

(b) Zoom on differential elements

Figure 30: Differential elements in spherical

We first find $\mathrm{d}\ell_\rho = \mathrm{d}\rho$ using the figure. Then, for the other two differential lengths, we have to use the same approach we did when we looked at cylindrical differential lengths. $\mathrm{d}\varphi$ and $\mathrm{d}\theta$ are angle differences. The lengths we are interested in are the arc lengths defined by the infinitesimal angle changes. Using Figure 31a , we observe that $\mathrm{d}\ell_\varphi$ is the arc length of a circle of radius ρ. Hence, $\mathrm{d}\ell_\varphi = \rho\mathrm{d}\varphi$. Then, with Figure 31b , we see that $\mathrm{d}\ell_\theta$ is the arc length of a circle of radius given by the base of the ABC right-angled triangle. The angle BAC is the complement of φ and thus the base of the triangle is $\rho\cos\left(\frac{\pi}{2}-\varphi\right) = \rho\sin\varphi$ and we have $\mathrm{d}\ell_\theta = \rho\sin\varphi\ \mathrm{d}\theta$.



(a) $\mathrm{d}\ell_\varphi$

(b) $\mathrm{d}\ell_\theta$

Figure 31: Arc length due to infinitesimal change in φ and θ

Finally, we use (29)-(33) to compute the differential elements given in 30b.

$$\boxed{\begin{aligned} \mathrm{d}\boldsymbol{\ell} &= \mathrm{d}\rho\hat{\boldsymbol{\rho}} + \rho\mathrm{d}\varphi\hat{\boldsymbol{\varphi}} + \rho\sin\varphi\mathrm{d}\theta\hat{\boldsymbol{\theta}} \\ \mathrm{d}\mathbf{S}_\rho &= \rho^2\sin\varphi\mathrm{d}\theta\mathrm{d}\varphi\hat{\boldsymbol{\rho}} \\ \mathrm{d}\mathbf{S}_\varphi &= \rho\sin\varphi\mathrm{d}\rho\mathrm{d}\theta\hat{\boldsymbol{\varphi}} \\ \mathrm{d}\mathbf{S}_\theta &= \rho\mathrm{d}\rho\mathrm{d}\varphi\hat{\boldsymbol{\theta}} \\ \mathrm{d}V &= \rho^2\sin\varphi\mathrm{d}\rho\mathrm{d}\varphi\mathrm{d}\theta. \end{aligned}}$$

We will develop a method for calculating a differential surface area element for any type of smooth surface later in this course.

6 Position, position vector and general vector

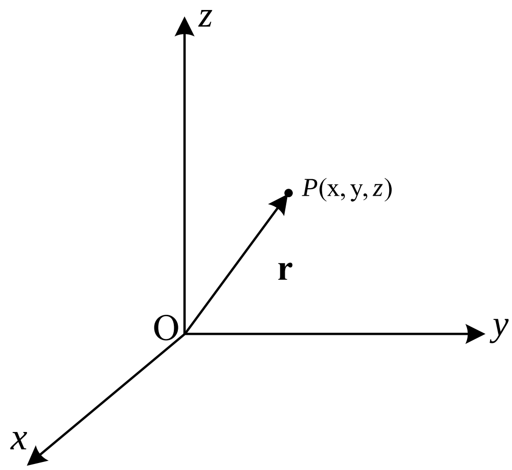


Figure 32: Position P and position vector $\mathbf{r}$

To conclude our review of the coordinate system, we discuss position, position vectors and general vectors. Given any coordinate set in n dimensions, a position P is given by the n-tuple which is characterized by the intersection point of $(n-1)$ objects defined for the system. A position P can be described by (x, y, z) in Cartesian coordinates, (r, θ, z) in cylindrical coordinates or (ρ, φ, θ) in spherical coordinates. Then, we define a position vector $\mathbf{r}$ as the vector going from the origin O to the position P. It is directed and has a magnitude given by the distance between the position and the origin (cf. Figure 32). How can we express this vector? In Cartesian coordinates, we have

$$\mathbf{r} = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}. \tag{43}$$

However, in cylindrical and spherical coordinates, it is not as straightforward. Starting from (43) and recalling the relation between Cartesian and cylindrical coordinates, we have

$$\mathbf{r} = r\cos\theta\hat{\mathbf{x}} + r\sin\theta\hat{\mathbf{y}} + z\hat{\mathbf{z}}$$

and using the vector transformation we derived in Section 5.2.2, the position vector in cylindrical coordinates is given by,

$$\mathbf{r} = r\hat{\mathbf{r}} + z\hat{\mathbf{z}}$$

It is important to observe here that the position vector has no $\hat{\boldsymbol{\theta}}$ term as one could expect. A position vector is only defined by the $\hat{\mathbf{r}}$ and $\hat{\mathbf{z}}$ component. The value of θ for P is taken into consideration inside $r\hat{\mathbf{r}}$ by definition. Another reason why a position vector cannot have a $\hat{\boldsymbol{\theta}}$ component is the following. Consider

$\mathbf{r} = r\hat{\mathbf{r}} + \theta\hat{\boldsymbol{\theta}} + z\hat{\mathbf{z}}$. Then its magnitude is not a distance as defined for a position vector. Why? Since the component $\theta\hat{\boldsymbol{\theta}}$ does not express a length. It expresses an angle.

The same arguments hold for spherical. We first re-express (43) using the relations between Cartesian and spherical coordinates and obtain

$$\mathbf{r} = \rho\cos\theta\sin\varphi\hat{\mathbf{x}} + \rho\sin\theta\sin\varphi\hat{\mathbf{y}} + \rho\cos\varphi\hat{\mathbf{z}}.$$

Then, with the relation between unit vectors of Section 5.3.2 and $\hat{\boldsymbol{\rho}}$, we have

$$\mathbf{r} = \rho\hat{\boldsymbol{\rho}}$$

for a position vector. Once again, it makes no sense to potentially include $\theta\hat{\boldsymbol{\theta}} + \varphi\hat{\boldsymbol{\varphi}}$ in this equation since the magnitude for both vector components represent angles and not lengths.

In summary, position vectors for a Cartesian, cylindrical and spherical coordinate system can only be expressed as

$$\boxed{\begin{aligned} \mathbf{r}_{\text{cart}} &= x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}} \\ \mathbf{r}_{\text{cyl}} &= r\hat{\mathbf{r}} + z\hat{\mathbf{z}} \\ \mathbf{r}_{\text{sph}} &= \rho\hat{\boldsymbol{\rho}} \end{aligned}}$$

respectively. This emphasizes the difference between a position vector and a general vector. A general vector can be expressed with all base vectors: in cylindrical as $\mathbf{v} = a\hat{\mathbf{r}} + b\hat{\boldsymbol{\theta}} + c\hat{\mathbf{z}}$ or in spherical as $\mathbf{v} = e\hat{\boldsymbol{\rho}} + f\hat{\boldsymbol{\varphi}} + g\hat{\boldsymbol{\theta}}$. However, it does not have the same meaning as position which represents the directed distance between the origin and a point in space. The main confusion can come when we consider a vector field $\mathbf{v}(x, y, z)$ which returns a vector at the position (x, y, z). Note that there is no relation with the origin here and $\mathbf{v}(x, y, z)$ is only a vector with its tail at (x, y, z). If we consider a vector field, $\mathbf{v}(r, \theta, z)$, in cylindrical coordinates then the vector at the position (r, θ, z) can have a $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{z}}$ component. Similarly, for a vector field in spherical coordinates, the vector at a position (ρ, φ, θ) can have a $\hat{\boldsymbol{\rho}}$, $\hat{\boldsymbol{\varphi}}$ and $\hat{\boldsymbol{\theta}}$ component.

7 Velocity and acceleration vectors in different coordinate systems

In mechanics we typically formulate the equations of motion in terms of a vector equation. You are all familiar with how to determine the velocity vector and acceleration vector in a Cartesian coordinate system from a previous course on dynamics and from MAT187S last year.

You are given the position vector $\mathbf{r} = \langle x(t), y(t), z(t)\rangle$ where t is time. The velocity vector, $\boldsymbol{\nu}$, and acceleration vector, $\mathbf{a}$, are defined as

$$\begin{aligned} \dot{\mathbf{r}} = \boldsymbol{\nu} &= \langle \dot{x}(t), \dot{y}(t), \dot{z}(t)\rangle \\ \ddot{\mathbf{r}} = \mathbf{a} &= \langle \ddot{x}(t), \ddot{y}(t), \ddot{z}(t)\rangle \end{aligned}$$

respectively.

The question is how do these expressions change when expressed in a different coordinate system.

7.1 2D Polar coordinates

We start by expressing the position vector $\mathbf{r}$ in polar coordinates which is $\mathbf{r} = r\hat{\mathbf{r}}$. Note there is no reference to the unit vector $\hat{\boldsymbol{\theta}}$ since in a polar coordinate system the vector from the origin to a point $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}})$ is defined only in terms of a unit radial vector. We assume that both components of the position $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}})$ are time-dependent. Refer to Section 6 for a more detailed description of position vectors.

First we compute the velocity vector by differentiating the expression for the position vector with respect to the variable t.

$$\boldsymbol{\nu} = \dot{r}\hat{\mathbf{r}} + r\dot{\hat{\mathbf{r}}}$$

Note that we need to differentiate the unit vector $\hat{\mathbf{r}}$ since it changes with respect to θ. Recall that

$$\begin{aligned} \hat{\mathbf{r}} &= \cos\theta\hat{\mathbf{x}} + \sin\theta\hat{\mathbf{y}} \\ \hat{\boldsymbol{\theta}} &= -\sin\theta\hat{\mathbf{x}} + \cos\theta\hat{\mathbf{y}} \end{aligned}$$

Now differentiate $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ with respect to r and θ. This is an exercise in differentiating by parts. The derivatives of the unit vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are zero since they do not depend on r or θ. Hence we are left with:

$$\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}r} = 0, \qquad \frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}\theta} = -\sin\theta\hat{\mathbf{x}} + \cos\theta\hat{\mathbf{y}} = \hat{\boldsymbol{\theta}}, \tag{44}$$

$$\frac{\mathrm{d}\hat{\boldsymbol{\theta}}}{\mathrm{d}r} = 0, \qquad \frac{\mathrm{d}\hat{\boldsymbol{\theta}}}{\mathrm{d}\theta} = -\cos\theta\hat{\mathbf{x}} - \sin\theta\hat{\mathbf{y}} = -\hat{\mathbf{r}}. \tag{45}$$

The only term of relevance in our problem is $\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}\theta}$.

The last step is to recognize that r was differentiated with respect to t and therefore

$$\dot{\hat{\mathbf{r}}} = \frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}t} = \frac{\mathrm{d}\theta}{\mathrm{d}t}\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}\theta} = \dot{\theta}\hat{\boldsymbol{\theta}}$$

Had $\hat{\mathbf{r}}$ been a function of two variables (r, θ) we would have had the following expression, after differentiation with respect to t:

$$\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}t} = \frac{\mathrm{d}\theta}{\mathrm{d}t}\frac{\partial\hat{\mathbf{r}}}{\partial\theta} + \frac{\mathrm{d}r}{\mathrm{d}t}\frac{\partial\hat{\mathbf{r}}}{\partial r}$$

Note we have used a symbol ∂ which refers to partial differentiation and d for a direct derivative. In a few weeks we define what partial differentials are and the distinction between a partial and direct derivative. We will also show how differentiation of multivariable functions for the general case can be computed using the Chain Rule.

The final expression for $\boldsymbol{\nu}$ after substitution becomes

$$\boxed{\boldsymbol{\nu} = \dot{r}\hat{\mathbf{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}}}$$

$\dot{r}$ represents the rate of change in the radial direction, at a given position in space, and $r\dot{\theta}$ represents the rate of change in the circumferential direction, at the same position in space.

Next we compute the acceleration vector $\mathbf{a}$.

$$\mathbf{a} = \dot{\boldsymbol{\nu}} = \ddot{r}\hat{\mathbf{r}} + \dot{r}\dot{\hat{\mathbf{r}}} + \dot{r}\dot{\theta}\hat{\boldsymbol{\theta}} + r\ddot{\theta}\hat{\boldsymbol{\theta}} + r\dot{\theta}\dot{\hat{\boldsymbol{\theta}}} \tag{46}$$

We have expressions already for the unit vectors $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\theta}}$ and $\dot{\hat{\mathbf{r}}}$. We need to now compute the unit vector $\dot{\hat{\boldsymbol{\theta}}}$.

From (45) and (44), we have

$$\frac{\mathrm{d}\hat{\boldsymbol{\theta}}}{\mathrm{d}\theta} = -\cos\theta\hat{\mathbf{x}} - \sin\theta\hat{\mathbf{y}} = -\hat{\mathbf{r}}$$

$$\frac{\mathrm{d}\hat{\boldsymbol{\theta}}}{\mathrm{d}r} = 0.$$

The next step is to backwards substitute for $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\theta}}$, $\dot{\hat{\mathbf{r}}}$ and $\dot{\hat{\boldsymbol{\theta}}}$ in (46). The final result after simplifications is:

$$\boxed{\mathbf{a} = \dot{\boldsymbol{\nu}} = \left(\ddot{r} - r\dot{\theta}^2\right)\hat{\mathbf{r}} + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\hat{\boldsymbol{\theta}}}$$

The term $\ddot{r} - r\dot{\theta}^2$ represents the radial acceleration at a position in space and $r\ddot{\theta} + 2\dot{r}\dot{\theta}$ represents the circumferential acceleration at the same position in space.

7.2 3D Cylindrical coordinates

The detailed procedure for computing the first two time derivatives of a position vector was presented for polar coordinates. For the 3D cylindrical case we follow the same procedure but have more bookkeeping. Only the main steps will be highlighted. In cylindrical coordinates, a position vector is given by,

$$\mathbf{r} = r\hat{\mathbf{r}} + z\hat{\mathbf{z}}.$$

We differentiate this expression with respect to t to obtain,

$$\boldsymbol{\nu} = \dot{r}\hat{\mathbf{r}} + r\dot{\hat{\mathbf{r}}} + \dot{z}\hat{\mathbf{z}} + z\dot{\hat{\mathbf{z}}}. \tag{47}$$

The base vectors $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ in cylindrical coordinates are defined the same way as the base vectors in cylindrical coordinates: $\hat{\mathbf{r}}$ and $\hat{\boldsymbol{\theta}}$ are the base vectors. Their derivative with respect to z is zero since they don't have any z dependency. Hence, the two first terms of (47) are identical to the one in the previous section. Now, we differentiate $\hat{\mathbf{z}}$ for each of the variables and we obtain,

$$\frac{\mathrm{d}\hat{\mathbf{z}}}{\mathrm{d}r} = 0, \qquad \frac{\mathrm{d}\hat{\mathbf{z}}}{\mathrm{d}\theta} = 0, \qquad \frac{\mathrm{d}\hat{\mathbf{z}}}{\mathrm{d}z} = 0.$$

Therefore, the last term of (47) is zero. Finally, the velocity is

$$\boxed{\boldsymbol{\nu} = \dot{r}\hat{\mathbf{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}} + \dot{z}\hat{\mathbf{z}},}$$

where $\dot{r}$ represents the rate of change in the radial direction, $r\dot{\theta}$ represents the rate of change in the circumferential direction and $\dot{z}$ the rate of change in the z direction, all of them at the same position in space.

For the acceleration vector $\mathbf{a}$, we proceed the same way. We first evaluate the time derivative of $\boldsymbol{\nu}$ and we have,

$$\mathbf{a} = \dot{\boldsymbol{\nu}} = \ddot{r}\hat{\mathbf{r}} + \dot{r}\dot{\hat{\mathbf{r}}} + \dot{r}\dot{\theta}\hat{\boldsymbol{\theta}} + r\ddot{\theta}\hat{\boldsymbol{\theta}} + r\dot{\theta}\dot{\hat{\boldsymbol{\theta}}} + \ddot{z}\hat{\mathbf{z}} + \dot{z}\dot{\hat{\mathbf{z}}}.$$

Using the result from Section 7.1 and the time derivative of $\hat{\mathbf{z}}$, we get the following form for the acceleration vector in cylindrical coordinates:

$$\boxed{\mathbf{a} = \dot{\boldsymbol{\nu}} = \left(\ddot{r} - r\dot{\theta}^2\right)\hat{\mathbf{r}} + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\hat{\boldsymbol{\theta}} + \ddot{z}\hat{\mathbf{z}}.}$$

In the acceleration equation, the coefficients from left to right represent the radial or centripetal acceleration, the circumferential or Coriolis acceleration and the translational acceleration in z.

7.3 3D Spherical coordinates

The detailed procedure for computing the first two time derivatives of a position vector was presented for polar coordinates. For the 3D spherical case we follow the same procedure but have more bookkeeping. Only the main steps will be highlighted. The position vector in spherical coordinates is

$$\mathbf{r} = \rho\hat{\boldsymbol{\rho}},$$

where the dependence in φ and θ is included in $\hat{\boldsymbol{\rho}}$. The velocity vector is given by the time derivative of $\mathbf{r}$,

$$\boldsymbol{\nu} = \dot{\mathbf{r}} = \dot{\rho}\hat{\boldsymbol{\rho}} + \rho\dot{\hat{\boldsymbol{\rho}}}.$$

We need to differentiate $\hat{\boldsymbol{\rho}}$ with respect to time. We will have to take into consideration that $\hat{\boldsymbol{\rho}}$ in terms of a Cartesian coordinate system is itself a function of φ and ρ. Recall that,

$$\begin{aligned}
\hat{\boldsymbol{\rho}} &= \cos\theta\sin\varphi\hat{\mathbf{x}} + \sin\theta\sin\varphi\hat{\mathbf{y}} + \cos\varphi\hat{\mathbf{z}}, \\
\hat{\boldsymbol{\varphi}} &= \cos\theta\cos\varphi\hat{\mathbf{x}} + \sin\theta\cos\varphi\hat{\mathbf{y}} - \sin\varphi\hat{\mathbf{z}}, \\
\hat{\boldsymbol{\theta}} &= -\sin\theta\hat{\mathbf{x}} + \cos\theta\hat{\mathbf{y}}.
\end{aligned}$$

We now use the chain rule and compute $\dot{\hat{\boldsymbol{\rho}}}$ given by,

$$\dot{\hat{\boldsymbol{\rho}}} = \frac{\mathrm{d}\hat{\boldsymbol{\rho}}}{\mathrm{d}t} = \frac{\partial\hat{\boldsymbol{\rho}}}{\partial\varphi}\frac{\mathrm{d}\varphi}{\mathrm{d}t} + \frac{\partial\hat{\boldsymbol{\rho}}}{\partial\theta}\frac{\mathrm{d}\theta}{\mathrm{d}t}$$

and similarly for $\dot{\hat{\boldsymbol{\varphi}}}$ and $\dot{\hat{\boldsymbol{\theta}}}$. Notice here the difference between the partial (∂) and direct (d) derivatives. The concept of the chain rule and the distinction between direct and partial differentiation will be studied later in the course. The time derivative $\dot{\hat{\boldsymbol{\varphi}}}$ and $\dot{\hat{\boldsymbol{\theta}}}$ will be needed later to evaluate the acceleration vector $\mathbf{a}$. After evaluating all derivatives, we get,

$$\dot{\hat{\boldsymbol{\rho}}} = \dot{\theta} \sin\varphi \hat{\boldsymbol{\varphi}} + \dot{\varphi}\hat{\boldsymbol{\varphi}}, \tag{48}$$

$$\dot{\hat{\boldsymbol{\varphi}}} = -\dot{\theta} \sin\varphi \hat{\boldsymbol{\rho}} - \dot{\theta} \cos\varphi \hat{\boldsymbol{\varphi}}, \tag{49}$$

$$\dot{\hat{\boldsymbol{\theta}}} = -\dot{\varphi} \hat{\boldsymbol{\rho}} - \dot{\theta} \cos\varphi \hat{\boldsymbol{\theta}}. \tag{50}$$

Substituting the derivatives that we computed, we get the following final expression for the velocity vector,

$$\boxed{\boldsymbol{\nu} = \dot{\mathbf{r}} = \dot{\rho}\hat{\boldsymbol{\rho}} + \rho\dot{\theta}\sin\varphi\hat{\boldsymbol{\theta}} + \rho\dot{\varphi}\hat{\boldsymbol{\varphi}}.}$$

We differentiate the velocity vector with respect to t and obtain the acceleration vector $\mathbf{a}$. Lastly, we use (48)-(50) and collect all terms. The final expression for the acceleration $\mathbf{a}$ is,

$$\boxed{\mathbf{a} = \dot{\boldsymbol{\nu}} = \left(\ddot{\rho} - \rho\dot{\theta}^2 \sin^2\varphi - \rho\dot{\varphi}^2\right)\hat{\boldsymbol{\rho}} + \left(2\dot{\rho}\dot{\theta}\sin\varphi + \rho\ddot{\theta}\sin\varphi + 2\rho\dot{\theta}\dot{\varphi}\cos\varphi\right)\hat{\boldsymbol{\theta}} + \left(2\dot{\rho}\dot{\varphi} - \rho\dot{\theta}^2\sin\varphi\cos\varphi + \rho\ddot{\varphi}\right)\hat{\boldsymbol{\varphi}}}$$

In the acceleration equation, the coefficients from left to right represent the radial or centripetal acceleration, the elevation angular acceleration and the circumferential or Coriolis acceleration.

8 Problems

***Problem* 1:** Given $\mathbf{A} = -\hat{\mathbf{x}} + 2\hat{\mathbf{y}} - 2\hat{\mathbf{z}}$, find

(a) its magnitude $A = |A|$,

(b) the expression of the unit vector $\hat{\mathbf{A}}$ in the direction of $\mathbf{A}$, and

(c) the angle that $\mathbf{A}$ makes with the z-axis.

(Solution)

***Problem* 2:** Consider a vector field defined in cylindrical coordinates to be

$$\mathbf{A} = 3\cos\theta\hat{\mathbf{r}} - 2r\hat{\boldsymbol{\theta}} + z\hat{\mathbf{z}}$$

(a) What is the field at $(r, \theta, z) = (4, 60°, 5)$?

(b) Express the position $(r, \theta, z) = (4, 60°, 5)$ in Cartesian coordinates.

(c) Express the vector from part (a) using the Cartesian unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$.

(Solution)

***Problem* 3:** Express the vector $\hat{\mathbf{z}}$ at an arbitrary position (ρ, φ, θ) in spherical coordinates.
(Solution)

***Problem* 4:** Consider the points $P = (r, \theta, z) = (3, 30°, 1)$ and $P' = (r, \theta, z) = (1, 90°, 2)$.

(a) Find the distance from P to P'.

(b) Write the vector $\mathbf{R} = PP'$ at both P and P' in cylindrical coordiates.

(Solution)

***Problem* 5:** Consider a spherical coordinate system with the ordered base vector sequence $\{\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varphi}}\}$ given that the coordinate system is right-handed and the direction of $\hat{\boldsymbol{\rho}}$ and $\hat{\boldsymbol{\theta}}$ follows the same convention as in Section 5.3.1. Sketch the base vectors of the spherical coordinate system at the point $(x, y, z) = (2, 2, 3)$.
(Solution)

9 Solutions

***Solution* 1:** The main concept in this problem is the scalar product. It is defined in Section 2.5 and is revisited in Section 3.

(a) $\mathbf{A} \cdot \mathbf{A} = 1 + 4 + 4 = 9 = |\mathbf{A}|^2 \Rightarrow |\mathbf{A}| = 3$;

(b) $\hat{\mathbf{A}} = \frac{\mathbf{A}}{|\mathbf{A}|} = -\frac{1}{3}\hat{\mathbf{x}} + \frac{2}{3}\hat{\mathbf{y}} - \frac{2}{3}\hat{\mathbf{z}}$;

(c) $\mathbf{A} \cdot \hat{\mathbf{z}} = |\mathbf{A}| \; |\hat{\mathbf{z}}| \cos\theta$
$-2 = 3\cos\theta \Rightarrow \theta = 131.8°$.

(Back to problem statement)

***Solution* 2:** The transformation between cylindrical and Cartesian coordinate systems is discussed in Section 5.2.2. Position vectors are discussed in Section 6.

(a) $\mathbf{A}_{cyl} = 3\cos 60°\hat{\mathbf{r}} - 8\hat{\boldsymbol{\theta}} + 5\hat{\mathbf{z}} = \frac{3}{2}\hat{\mathbf{r}} - 8\hat{\boldsymbol{\theta}} + 5\hat{\mathbf{z}}$;

(b) $x = 4\cos 60°$, $y = 4\sin 60° = 2\sqrt{3}$ and $z = 5 \Rightarrow$ Position vector is $(2, 2\sqrt{3}, 5)$;

(c)

$$\begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} = \begin{bmatrix} \cos 60° & -\sin 60° & 0 \\ \sin 60° & \cos 60° & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3/2 \\ -8 \\ 5 \end{bmatrix} = \begin{bmatrix} 7.68 \\ -2.7 \\ 5 \end{bmatrix}$$

(Back to problem statement)

***Solution* 3:** The transformation between spherical and Cartesian coordinate systems is discussed in Section 5.3.2. Position vectors are discussed in Section 6.

$$\begin{bmatrix} A_\rho \\ A_\varphi \\ A_\theta \end{bmatrix} = \begin{bmatrix} \cos\theta\sin\varphi & \sin\theta\sin\varphi & \cos\varphi \\ \cos\theta\cos\varphi & \sin\theta\cos\varphi & -\sin\varphi \\ -\sin\theta & \cos\theta & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\varphi \\ -\sin\varphi \\ 0 \end{bmatrix} = \cos\varphi\hat{\boldsymbol{\rho}} - \sin\varphi\hat{\boldsymbol{\varphi}}.$$

(Back to problem statement)

***Solution* 4:** In computing the distance RR' between any two points in a coordinate system, other than Cartesian, you must first transform the coordinate points into a Cartesian coordinate reference frame, as discussed in Appendix B. The transformation between cylindrical and Cartesian coordinate systems is discussed in Section 5.2.2.

(a) $P(x, y, z) = (3\cos 30°, 3\sin 30°, 1) = (2.6, 1.5, 1)$
$P'(x, y, z) = (\cos 90°, \sin 90°, 2) = (0, 1, 2)$
The vector from P to P' is $\mathbf{R}_{PP'} = -2.6\hat{\mathbf{x}} - 0.5\hat{\mathbf{y}} + \hat{\mathbf{z}}$, and the distance is $|\mathbf{R}_{PP'}| = \sqrt{2.6^2 + 0.5^2 + 1} = 2.83$.

(b) Converting $\mathbf{R}_{PP'}$ from Cartesian coordinate to cylindrical coordinates, we have

$$\begin{bmatrix} R_r \\ R_\theta \\ R_z \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2.6 \\ -0.5 \\ 1 \end{bmatrix}$$

At P, $\theta = 30°$, so $\mathbf{R}_{PP',cyl} = -2.5\hat{\mathbf{r}} + 0.87\hat{\boldsymbol{\theta}} + \hat{\mathbf{z}}$.
At P', $\theta = 90°$, so $\mathbf{R}_{PP',cyl} = -0.5\hat{\mathbf{r}} + 2.6\hat{\boldsymbol{\theta}} + \hat{\mathbf{z}}$.

(Back to problem statement)

***Solution* 5:** Right-handed system in spherical are discussed in Section 5.3.1.

The base vector $\hat{\boldsymbol{\varphi}}$ points upward with respect to the surface of the cone $\varphi = \varphi_0$. φ is measured with respect to the z axis in the counterclockwise sense.

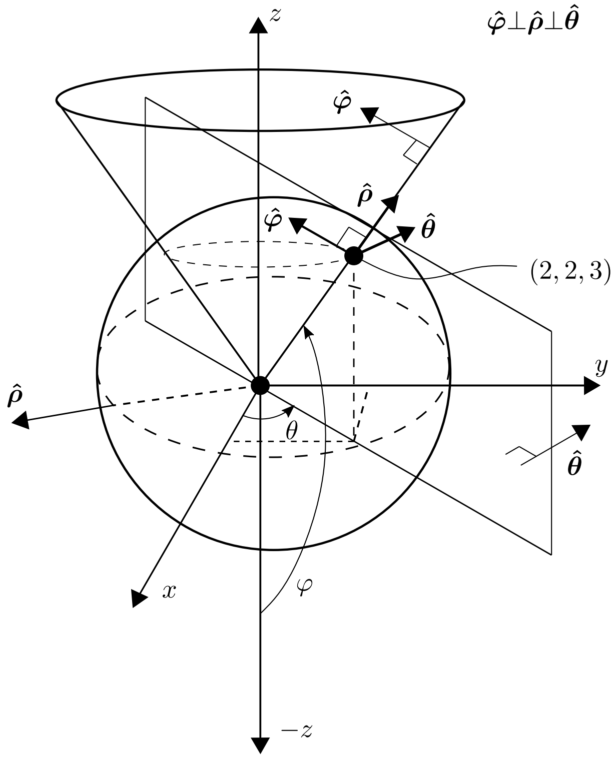


Figure 33: Right-handed coordinate system for the ordered sequence $\{\hat{\boldsymbol{\rho}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varphi}}\}$

(Back to problem statement)

Appendix A Motivation for scalar triple product

In the main section we showed that the scalar triple product has a geometrical significance. Here we show that it has another use which relates to the creation of a dual vector space with specific properties in relation to the original vector space. Recall that the choice of basis vectors is arbitrary.

Consider now an alternative basis vector set that is related to the initial basis vectors in a specific manner. This alternate base vector space is referred to as a reciprocal or dual vector space as will become evident later.

The scalar triple product appears naturally when we design a physical coordinate system in 3D. It is notably used when we ensure certain properties between a primal (normal) and a dual (reciprocal) vector space in a coordinate system. We want a transformation matrix that allows us to convert from the primal space to the dual space with some physical constraints applied.

Recall from Section 4.2.2 that we can re-express vectors using a matrix made of projection. Let **a**, **b**, **c** be the primal vectors and **v**, **u** and **w** be the dual vectors. These vectors are represented in Figure A.1. The dual of a given vector possesses an important property. A base vector in the dual space is perpendicular to the plane generated by two basis vectors in the primal space. There are three combinations of pairs of base vectors in the primal space so there will be three base vectors in the dual space. The cross product of any two primal base vectors will be parallel to one of the vectors in the dual vector space. The transformation matrix, $\mathbf{T}$, has the form,

$$\mathbf{T} = \begin{bmatrix} \mathbf{u}\cdot\mathbf{a} & \mathbf{v}\cdot\mathbf{a} & \mathbf{w}\cdot\mathbf{a} \\ \mathbf{u}\cdot\mathbf{b} & \mathbf{v}\cdot\mathbf{b} & \mathbf{w}\cdot\mathbf{b} \\ \mathbf{u}\cdot\mathbf{c} & \mathbf{v}\cdot\mathbf{c} & \mathbf{w}\cdot\mathbf{c} \end{bmatrix}$$

which is the projection of one set of vectors onto the dual vector space. The dot product is commutative, hence we can re-write

$$\mathbf{T} = \begin{bmatrix} \mathbf{a}\cdot\mathbf{u} & \mathbf{a}\cdot\mathbf{v} & \mathbf{a}\cdot\mathbf{w} \\ \mathbf{b}\cdot\mathbf{u} & \mathbf{b}\cdot\mathbf{v} & \mathbf{b}\cdot\mathbf{w} \\ \mathbf{c}\cdot\mathbf{u} & \mathbf{c}\cdot\mathbf{v} & \mathbf{c}\cdot\mathbf{w} \end{bmatrix}$$

We want to ensure the following property between the primal and dual vectors in our system in order to make the matrix $\mathbf{T}$ diagonal and a multiple of the identity matrix. This type of construct is easy to work with and appears in practice.

$$\mathbf{a}\cdot\mathbf{u} = A, \quad \mathbf{b}\cdot\mathbf{v} = A, \quad \mathbf{c}\cdot\mathbf{w} = A, \tag{A.1}$$

where A is a scalar number and

$$\mathbf{a}\cdot\mathbf{v} = \mathbf{a}\cdot\mathbf{w} = \mathbf{b}\cdot\mathbf{u} = \mathbf{b}\cdot\mathbf{w} = \mathbf{c}\cdot\mathbf{u} = \mathbf{c}\cdot\mathbf{v} = 0. \tag{A.2}$$

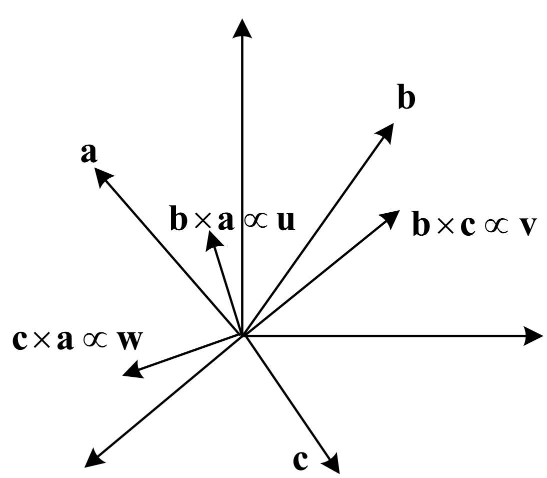


Figure A.1: Primal and dual vectors

The only dual vectors that will respect these constraints are the vectors defined as,

$$\mathbf{u} = A\frac{\mathbf{b}\times\mathbf{c}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}}, \quad \mathbf{v} = A\frac{\mathbf{c}\times\mathbf{a}}{\mathbf{b}\cdot\mathbf{c}\times\mathbf{a}}, \quad \mathbf{w} = A\frac{\mathbf{a}\times\mathbf{b}}{\mathbf{c}\cdot\mathbf{a}\times\mathbf{b}}, \tag{A.3}$$

where the denominator is by definition the scalar triple product. Notice that (A.3) gives dual vectors parallel to the cross product of the two primal vectors which respect to both (A.1) and (A.2). We can check this by substituting the dual vectors given in A.3 in both (A.1) and (A.2).

Appendix B Addition of two position vectors in polar coordinates

Position vectors introduced in Section 6 and described in coordinate systems other than Cartesian coordinate system cannot be directly added component-wise like general vectors. The simplest approach to add two positions vectors is always to convert them to Cartesian then add them and finally convert the resultant vector to the desired coordinate system. This can be done using the transformation matrices introduced in Section 4.2.2 in the 2D polar coordinates case and in Section 5.2.2 and in Section 5.3.2 in the 3D cylindrical and spherical coordinates case respectively.

In this section, we show that the sum of two position vectors in a non Cartesian coordinate system is not given by the component-wise addition of its parameter. We will use a polar coordinate system for demonstration purposes. It will also become clear from this example that to sum vectors in a non Cartesian system becomes a futile exercise for coordinate systems with higher dimensionality. Consequently, the recommended approach for adding vectors in non Cartesian coordinate systems is the easiest way forward.

Consider $\mathbf{r}_1$ and $\mathbf{r}_2$ as the vectors between the origin and the points (r_1, θ_1) and (r_2, θ_2) respectively. $\mathbf{r}_1$ is hence a vector with magnitude r_1 making an angle θ_1 with the x-axis. Similarly, $\mathbf{r}_2$ is a vector with magnitude r_2 making an angle θ_2 with the x-axis. For simplicity, we consider $0 \leq \theta_1 \leq \theta_2 \leq \pi/2$. We want to show that the sum of vectors in a non Cartesian coordinate system is not simply done component-wise, hence this assumption is not to restrictive because only a counterexample should suffice to demonstrate our point. Our goal is to compute $\mathbf{r} = \mathbf{r}_1 + \mathbf{r}_2$ which in polar coordinates can be expressed by its magnitude r and angle θ and hence by the pair (r, θ) with $\theta_1 \leq \theta \leq \pi$. We show that $\langle r, \theta\rangle \neq \langle r_1 + r_2, \theta_1 + \theta_2\rangle$ like a general vector in a Cartesian coordinate system; for position vectors, the component-wise addition is only valid when vectors are expressed in Cartesian coordinates. In a Cartesian coordinate system, we have $\langle x_1, y_1, z_1\rangle + \langle x_2, y_2, z_2\rangle = \langle x_1 + x_2, y_1 + y_2, z_1 + z_2\rangle$.

Let us now outline the procedure for computing $\langle r, \theta\rangle$ directly in polar coordinates. We can first directly find the magnitude of the resulting vector $\mathbf{r}$. Recall that $\|\mathbf{r}\|^2 = \mathbf{r}\cdot\mathbf{r}$. Then

$$\begin{aligned} r^2 &= (\mathbf{r}_1 + \mathbf{r}_2)\cdot(\mathbf{r}_1 + \mathbf{r}_2) \\ &= \mathbf{r}_1\cdot\mathbf{r}_1 + \mathbf{r}_2\cdot\mathbf{r}_2 + 2\mathbf{r}_1\cdot\mathbf{r}_2, \end{aligned}$$

where we use the commutativity property ($\mathbf{r}_1\cdot\mathbf{r}_2 = \mathbf{r}_2\cdot\mathbf{r}_1$) of the dot product in the last equation. Using the general definition of the scalar product leads to,

$$\begin{aligned} r^2 &= r_1^2 + r_2^2 + 2r_1 r_2 \cos(\theta_1 - \theta_2) \\ \Leftrightarrow r &= \sqrt{r_1^2 + r_2^2 + 2r_1 r_2 \cos(\theta_2 - \theta_1)}, \end{aligned} \tag{B.1}$$

the magnitude of the resulting vector. We note that the r-component of the position vector in polar coordinates is *not* simply $r_1 + r_2$, the component-wise addition of the r-components. Then, let's continue with the angle θ of the resulting vector.

Our strategy is to use the dot product between $\mathbf{r}_1$ and $\mathbf{r}$ to determine θ. We have,

$$\begin{aligned}
\mathbf{r}_1 \cdot \mathbf{r} &= \mathbf{r}_1 \cdot (\mathbf{r}_1 + \mathbf{r}_2) \\
\Leftrightarrow r r_1 \cos(\theta - \theta_1) &= r_1^2 + \mathbf{r}_1 \cdot \mathbf{r}_2, \\
r r_1 \cos(\theta - \theta_1) &= r_1^2 + r_1 r_2 \cos(\theta_2 - \theta_1), \\
\cos(\theta - \theta_1) &= \frac{r_1 + r_2 \cos(\theta_2 - \theta_1)}{r}, \\
&= \frac{r_1 + r_2 \cos(\theta_2 - \theta_1)}{\sqrt{r_1^2 + r_2^2 + 2 r_1 r_2 \cos(\theta_2 - \theta_1)}}.
\end{aligned}$$

This leads to,

$$\theta = \theta_1 + \arccos\left(\frac{r_1 + r_2 \cos(\theta_2 - \theta_1)}{\sqrt{r_1^2 + r_2^2 + 2 r_1 r_2 \cos(\theta_2 - \theta_1)}}\right) \tag{B.2}$$

Once again, this differs significantly from simply adding both θ_1 and θ_2.

Now let us redo the same example using the recommended approach that is to convert the position vector into Cartesian first. At this point we obtain a vector of the form: $\langle r_{1,x}, r_{1,y}\rangle$ and $\langle r_{2,x}, r_{2,y}\rangle$ where $r_{i,x}$ and $r_{i,y}$ for $i = 1, 2$ are the x and y component of the vector $\mathbf{r}_i$. Then we add the components of the vector component wise and obtain the resulting vector $\langle r_{1,x} + r_{2,x}, r_{1,y} + r_{2,y}\rangle$. Finally, we convert the resulting vector back to polar coordinates, $\langle r, \theta, \rangle$.

We start by converting $\mathbf{r}_1$ and $\mathbf{r}_2$ to Cartesian. Recall from Section 6 that a position vector in polar coordinates (cylindrical with $z = 0$) is given by: $\mathbf{r}_1 = r_1 \hat{\mathbf{r}}$ and $\mathbf{r}_2 = r_2 \hat{\mathbf{r}}$ where the base vector $\hat{\mathbf{r}}$ is a function of θ. We use the transformation matrix given in (23) and we obtain,

$$\begin{aligned}
\mathbf{r}_{1,\text{cart}} &= \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} r_1 \\ 0 \end{bmatrix} \\
&= r_1 \cos\theta_1 \hat{\mathbf{x}} + r_1 \cos\theta_1 \hat{\mathbf{y}}
\end{aligned}$$

And similarly,

$$\mathbf{r}_{2,\text{cart}} = r_2 \cos\theta_2 \hat{\mathbf{x}} + r_2 \cos\theta_2 \hat{\mathbf{y}}$$

As mentioned previously, position vectors can be added component-wise in Cartesian since the base vectors are not a function of the position. This leads to,

$$\mathbf{r}_{\text{cart}} = (r_1 \cos\theta_1 + r_2 \cos\theta_2)\,\hat{\mathbf{x}} + (r_1 \sin\theta_1 + r_2 \sin\theta_2)\,\hat{\mathbf{y}}$$

where the coefficients of $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ represent the x and y position of the new vector. We can retrieve the pair (r, θ), by computing the norm of $\mathbf{r}$ and the angle between the x and y components or between the $\mathbf{r}$ and its x component.

$$\begin{aligned}
r &= \|\mathbf{r}_{\text{cart}}\| \\
&= \sqrt{(r_1 \cos\theta_1 + r_2 \cos\theta_2)^2 + (r_1 \sin\theta_1 + r_2 \sin\theta_2)^2}, \\
&= \sqrt{r_1^2 + r^2 + 2 r_1 r_2 \cos(\theta_1)\cos(\theta_2) + 2 r_1 r_2 \sin(\theta_1)\sin(\theta_2)}, \\
&= \sqrt{r_1^2 + r^2 + 2 r_1 r_2 \frac{1}{2}(\cos(\theta_2 - \theta_1) + \cos(\theta_2 + \theta_1)) + 2 r_1 r_2 \frac{1}{2}(\cos(\theta_2 - \theta_1) + \cos(\theta_2 - \theta_1))}, \\
&= \sqrt{r_1^2 + r^2 + 2 r_1 r_2 \cos(\theta_2 - \theta_1)}
\end{aligned} \tag{B.3}$$

Hence, r, as expected, is the same as (B.3) and in (B.1). Then, for the angle, we have,

$$\theta = \arccos\left(\frac{r_1 \cos\theta_1 + r_2 \cos\theta_2}{\sqrt{r_1^2 + r^2 + 2 r_1 r_2 \cos(\theta_2 - \theta_1)}}\right), \tag{B.4}$$

where it can be shown with a bit of work that (B.4) is equivalent to (B.2). The position vector $\mathbf{r} = \langle r, \theta\rangle$ can therefore be computed directly by converting each vector in polar coordinate into Cartesian coordinates, summing the vectors and then converting the resulting vector back into polar coordinates.

References

[1] D. K. Cheng *et al.*, *Field and Wave Electromagnetics.* Pearson Education India, 1989.

[2] D. J. Griffiths, "Introduction to Electrodynamics," 2005.

[3] W. L. Briggs, L. Cochran, and B. Gillett, *Calculus for Scientists and Engineers: Early Transcendentals.* Pearson Higher Ed, 2012.