Coordinates, vectors and tensors

To express properties like location in our world, we need coordinates. How do coordinates work? And what fun things can we do with them? That's what this chapter is about.

1 Coordinates and 1-vectors

1.1 Coordinate systems

Let's consider an n-dimensional space. A **coordinate system** is a function $X(x^1, x^2, \ldots, x^n)$, which assigns to every point in space n numbers $x^1, x^2, \ldots, x^n$. These numbers are called the **coordinates**. A point can have different representations in different coordinate systems.

Suppose we have a coordinate system. We can then draw coordinate lines. **Coordinate lines** are lines for which $n-1$ coordinates are fixed. As the non-fixed coordinate varies, a line is drawn. (Note that these aren't always straight lines.)

1.2 Base vectors

A coordinate system also has **base vectors**. Thse vectors are vectors tangent to the coordinate lines. Together, they form the **(covariant) basis** of the system. The base vectors are defined as

$$\mathbf{e_1} = \frac{\partial \mathbf{X}}{\partial x^1}, \qquad \mathbf{e_2} = \frac{\partial \mathbf{X}}{\partial x^2}, \qquad \ldots, \qquad \mathbf{e_n} = \frac{\partial \mathbf{X}}{\partial x^n}. \tag{1.1}$$

The base vectors can be different at different points in the coordinate system. Also, they do not necessarily have length 1.

A covariant basis also has a corresponding **contravariant basis** (also known as the **dual basis**). Th contravariant base vectors $\mathbf{e^j}$ are defined such that

$$\mathbf{e^j} \cdot \mathbf{e_i} = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases} \tag{1.2}$$

Note that the contravariant basis is denoted by superscripts, while the covariant basis uses subscripts. To find the contravariant basis, you could take the matrix of covariant base vectors $[\mathbf{e_1} \ldots \mathbf{e_n}]$. If we invert it, we get the matrix of contravariant base vectors.

1.3 Normal vectors

Suppose that we have two points A and B. We can indicate their relative position by a vector. We can write down a vector $\mathbf{a}$ in the covariant basis as

$$\mathbf{a} = a^1\mathbf{e_1} + a^2\mathbf{e_2} + \ldots + a^n\mathbf{e_n} = \sum_{i=1}^{n} a^i\mathbf{e_i}. \tag{1.3}$$

The coefficients a^i are called the **contravariant coefficients**, since they have a superscript. We could also express the vector in the contravariant basis. We would then write it as

$$\mathbf{a} = a_1\mathbf{e^1} + a_2\mathbf{e^2} + \ldots + a_n\mathbf{e^n} = \sum_{i=1}^{n} a_i\mathbf{e^i}. \tag{1.4}$$

The coefficient a_i are the **covariant coefficients**, since they have a subscript. Of course, there is a relation between these coefficients. By using the definition of the dual basis, we can find that

$$a_j = \sum_{i=1}^{n} a_i \mathbf{e^i} \cdot \mathbf{e_j} = \mathbf{a} \cdot \mathbf{e_j} = \sum_{i=1}^{n} a^i \mathbf{e_i} \cdot \mathbf{e_j}. \tag{1.5}$$

1.4 The Einstein summation convention

We just saw that, to express a vector, we needed to add up n values. For that, we could use dots ... or the summation sign $\sum$. However, doing this every time could be a bit tiring. Therefore, from now on, we will use the **Einstein summation convention**. When, in a single term, there is both a subscript and an equal superscript, we make a summation. This means that

$$\sum_{i=1}^{n} a^i \mathbf{e_i} \quad \text{means the same as} \quad a^i \mathbf{e_i} \quad \text{and} \quad \sum_{i=1}^{n} a_i \mathbf{e^i} \quad \text{means the same as} \quad a_i \mathbf{e^i} \quad \text{as well.} \tag{1.6}$$

That should save us some ink.

1.5 Change of coordinates

Let's suppose we have a coordinate system $\mathbf{X}(x^1, \ldots, x^n)$. However, we move to a new set of coordinates $\tilde{x}^1, \tilde{x}^2, \ldots, \tilde{x}^n$. The functions $\tilde{x}^i = \tilde{x}^i(x^1, \ldots, x^n)$ are given. In this case, the new base vectors become

$$\mathbf{\tilde{e}_k} = \frac{\partial \mathbf{X}}{\partial \tilde{x}^k} = \frac{\partial \mathbf{X}}{\partial x^i} \frac{\partial x^i}{\partial \tilde{x}^k} = \mathbf{e_i} \frac{\partial x^i}{\partial \tilde{x}^k}, \qquad \text{and similarly,} \qquad \mathbf{e_k} = \mathbf{\tilde{e}_i} \frac{\partial \tilde{x}^i}{\partial x^k}. \tag{1.7}$$

It is important to note that we have used the Einstein summation convention in the above equation. So keep in mind that the above equation actually is a sum.

We can also express any vector $\mathbf{a}$ in our new coordinates. We simply need to find the new coefficients $\tilde{a}^i$. In this case, we have

$$\mathbf{a} = a^k \mathbf{e_k} = a^k \left(\mathbf{\tilde{e}_i} \frac{\partial \tilde{x}^i}{\partial x^k} \right) = \left(a^k \frac{\partial \tilde{x}^i}{\partial x^k} \right) \mathbf{\tilde{e}_i} = \tilde{a}^i \mathbf{\tilde{e}_i}, \qquad \text{which implies that} \qquad \tilde{a}^i = a^k \frac{\partial \tilde{x}^i}{\partial x^k}. \tag{1.8}$$

And the transformation is complete.

2 Multi-vectors

2.1 2-vectors

We now know how to describe points (with three coordinates) and lines (with a vector). But how would we describe a surface? For this, we use **2-vectors**. We do this using the **wedge operator** $\wedge$. Let's suppose we have two vectors $\mathbf{a}$ and $\mathbf{b}$. Together, they can form the 2-vector $(\mathbf{a} \wedge \mathbf{b})$.

The two-vector is subject to several rules. The most important rules are

$$c_1(\mathbf{a} \wedge \mathbf{b}) + c_2(\mathbf{a} \wedge \mathbf{b}) = (c_1 + c_2)(\mathbf{a} \wedge \mathbf{b}), \tag{2.1}$$
$$(\mathbf{a} \wedge \mathbf{d}) + (\mathbf{b} \wedge \mathbf{d}) = ((\mathbf{a} + \mathbf{b}) \wedge \mathbf{d}), \tag{2.2}$$
$$(c\mathbf{a} \wedge \mathbf{b}) = (\mathbf{a} \wedge c\mathbf{b}) = c(\mathbf{a} \wedge \mathbf{b}), \tag{2.3}$$
$$(\mathbf{a} \wedge \mathbf{b}) = -(\mathbf{b} \wedge \mathbf{a}), \tag{2.4}$$
$$(\mathbf{a} \wedge \mathbf{a}) = 0. \tag{2.5}$$

One way to think of the 2-vector $(\mathbf{a} \wedge \mathbf{b})$ is as the surface spanned by the two vectors $\mathbf{a}$ and $\mathbf{b}$. It then also makes sense why $(\mathbf{a} \wedge \mathbf{a}) = 0$. A single vector can't span a surface by itself.

Let's suppose that $\mathbf{u} = a\mathbf{e_1} + b\mathbf{e_2}$ and $\mathbf{v} = c\mathbf{e_1} + d\mathbf{e_2}$. We can then simplify $(\mathbf{u} \wedge \mathbf{v})$ to

$$((a\mathbf{e_1}+b\mathbf{e_2})\wedge(c\mathbf{e_1}+d\mathbf{e_2})) = ac(\mathbf{e_1}\wedge\mathbf{e_1})+ad(\mathbf{e_1}\wedge\mathbf{e_2})+bc(\mathbf{e_2}\wedge\mathbf{e_1})+bd(\mathbf{e_2}\wedge\mathbf{e_2}) = (ad-bc)(\mathbf{e_1}\wedge\mathbf{e_2}). \tag{2.6}$$

Another way to represent a surface, is by using the normal vector. Let's examine the surface $(\mathbf{e_1}, \mathbf{e_2})$. The normal vector of this surface is $\mathbf{e_1} \times \mathbf{e_2} = \mathbf{e_3}$. (Similarly, $\mathbf{e_2} \times \mathbf{e_3} = \mathbf{e_1}$ and $\mathbf{e_3} \times \mathbf{e_1} = \mathbf{e_2}$. So, instead of taking the wedge operator, we could use the cross product to respresent surfaces. In this case, we would also find that

$$\mathbf{u} \times \mathbf{v} = (a\mathbf{e_1} + b\mathbf{e_2}) \times (c\mathbf{e_1} + d\mathbf{e_2}) = (ad - bc)\mathbf{e_3}. \tag{2.7}$$

We see that this matches with what we found earlier.

2.2 3-vectors

Just like a 2-vector represents a surface, so does a 3-vector represent a volume. We denote such a 3-vector by $(\mathbf{u} \wedge \mathbf{v} \wedge \mathbf{w})$. There are rules for 3-vectors as well. The most important ones are

$$a(\mathbf{u} \wedge \mathbf{v} \wedge \mathbf{w}) + b(\mathbf{u}' \wedge \mathbf{v} \wedge \mathbf{w}) = ((a\mathbf{u} + b\mathbf{u}') \wedge \mathbf{v} \wedge \mathbf{w}), \tag{2.8}$$
$$(\mathbf{u} \wedge \mathbf{u} \wedge \mathbf{v}) = (\mathbf{u} \wedge \mathbf{v} \wedge \mathbf{u}) = (\mathbf{v} \wedge \mathbf{u} \wedge \mathbf{u}) = 0, \tag{2.9}$$
$$(\mathbf{u} \wedge \mathbf{v} \wedge \mathbf{w}) = (\mathbf{w} \wedge \mathbf{u} \wedge \mathbf{v}) = (\mathbf{v} \wedge \mathbf{w} \wedge \mathbf{u}) = -(\mathbf{w} \wedge \mathbf{v} \wedge \mathbf{u}). \tag{2.10}$$

Now let's suppose $\mathbf{u} = a\mathbf{e_1} + b\mathbf{e_2} + c\mathbf{e_3}$, $\mathbf{v} = k\mathbf{e_1} + l\mathbf{e_2} + m\mathbf{e_3}$ and $\mathbf{v} = p\mathbf{e_1} + q\mathbf{e_2} + e\mathbf{e_3}$. We can then simplify their wedge product to

$$(\mathbf{u} \wedge \mathbf{v} \wedge \mathbf{w}) = \det \begin{vmatrix} a & b & c \\ k & l & m \\ p & q & r \end{vmatrix} (\mathbf{e_1} \wedge \mathbf{e_2} \wedge \mathbf{e_3}). \tag{2.11}$$

3 Tensors

3.1 Tensor definitions

Tensors can be used to transform one vector to another. For example, we can say that the tensor **A** transforms vector $\mathbf{a}$ to vector $\mathbf{b}$. We write this as $\mathbf{b} = A\mathbf{a}$. We assume the tensor transforms vectors linearly. So,

$$A(c\mathbf{a}) = cA\mathbf{a} \qquad \text{and} \qquad A(\mathbf{a} + \mathbf{b}) = A\mathbf{a} + A\mathbf{b}. \tag{3.1}$$

Let's suppose that we have $\mathbf{b} = A\mathbf{a}$. The **inverse** of a tensor A (denoted by A^{-1}) is defined such that $\mathbf{a} = A^{-1}\mathbf{b}$, for every $\mathbf{a}$ and $\mathbf{b}$. The **transpose** of a vector A (denoted by A^T) is the vector which satisfies $\mathbf{b} \cdot A\mathbf{a} = \mathbf{a} \cdot A^T\mathbf{b}$, for every $\mathbf{a}$ and $\mathbf{b}$. If a tensor satisfies $A^T = A$, then it is called **symmetric**. If $A^T = -A$, then it is **anti-symmetric** (also known as **skew-symmetric**). If we have $A^T = A^{-1}$, then A is an **orthogonal** tensor. Finally, we define the identity tensor I as the tensor satisfying $\mathbf{a} = I\mathbf{a}$, for every $\mathbf{a}$.

3.2 Adding coordinate systems

The rules of the previous paragraph don't require any coordinate system. If we, however, do add a coordinate system, then we can represent a tensor A as a matrix. You should be careful with this though, as the matrix differs per coordinate system.

Let's suppose we know how a tensor A transforms vectors. How can we find the appropriate matrix? To find that out, we examine $\mathbf{b} = A\mathbf{a}$. Rewriting this, using the Einstein summation convention, gives

$$b^j \mathbf{e_j} = A\left(a^i \mathbf{e_i}\right) = a^i \left(A\mathbf{e_i}\right). \tag{3.2}$$

Left-multiplying by the dual basis vector $\mathbf{e^k}$ gives

$$b^k = b^j \mathbf{e^k} \cdot \mathbf{e_j} = a^i \left(\mathbf{e^k} \cdot A\mathbf{e_i}\right). \tag{3.3}$$

We thus find that A_i^k (being the component of A in the k-th row and the i-th column) is

$$A_i^k = \mathbf{e^k} \cdot A\mathbf{e_i}. \tag{3.4}$$

Note that we can now also write $\mathbf{b^k} = A_i^k a^i$.

3.3 Change of variables

Let's suppose we know all the coefficients A_i^k. But now we move to a new coordinate system, having coordinates $\tilde{x}^1, \tilde{x}^2, \ldots, \tilde{x}^n$. Again, the functions $\tilde{x}^i = \tilde{x}^i(x^1, \ldots, x^n)$ are given. How can we find the new components of the transformation matrix $\tilde{A}_i^k$?

To do this, we write $\tilde{A}_i^k$ as $\tilde{\mathbf{e}}^{\mathbf{k}} A \tilde{\mathbf{e}}_{\mathbf{i}}$. We can then apply the change of base vector equation (1.7) for base vectors. If we also work things out, we will find that

$$\tilde{A}_i^k = \tilde{\mathbf{e}}^{\mathbf{k}} A \tilde{\mathbf{e}}_{\mathbf{i}} = \left(\mathbf{e^l} \frac{\partial x^l}{\partial \tilde{x}^k}\right) A \left(\mathbf{e_j} \frac{\partial x_j}{\partial \tilde{x}_i}\right) = \frac{\partial x^l}{\partial \tilde{x}^k} \left(\mathbf{e^l} A \mathbf{e_j}\right) \frac{\partial x_j}{\partial \tilde{x}_i} = \frac{\partial x^l}{\partial \tilde{x}^k} \frac{\partial x_j}{\partial \tilde{x}_i} A_j^l. \tag{3.5}$$

The final relation above may look simple. But do remember that you need to sum up 9 individual parts to find the single component $\tilde{A}_i^k$, due to the summation convention. We could, of course, also reverse the above relation. We then would have

$$A_j^l = \frac{\partial \tilde{x}^k}{\partial x^l} \frac{\partial \tilde{x}_i}{\partial x_j} \tilde{A}_i^k. \tag{3.6}$$