Lecture 7: Beyond Cartesian Tensors

August 30, 2018

In general the components of a vector maybe measured parallel to the axes or perpendicular to it. This distinction leads to two equivalent representations of a vector. If the components are measured parallel to the axes, the vector is said to be contravariant. Contravariant vectors are denoted by a superscript index as x^i. If the components are perpendicular to the axes, it's called a covariant vector. These are written with a subscript index x_i. The crucial difference between these two types of vectors (two representations of a vector) lies in the transformation rules for these vectors. These can be seen as follows.

1 Transformation rules

Assume that we want to transform a certain vector from an old coordinate system (x_1, x_2, x_3) to a new coordinate system $(\grave{x}_1, \grave{x}_2, \grave{x}_3)$. The new coordinates can be expressed as some functions of the old ones. Hence we have

$$\grave{x}_1 \equiv \grave{x}_1(x_1, x_2, x_3)$$
$$\grave{x}_2 \equiv \grave{x}_2(x_1, x_2, x_3)$$
$$\grave{x}_3 \equiv \grave{x}_3(x_1, x_2, x_3)$$

Then, using chain rule, we can write the components of vector $\underline{d\grave{x}}$ as

$$d\grave{x}^1 = \frac{\partial \grave{x}^1}{\partial x^1} dx^1 + \frac{\partial \grave{x}^1}{\partial x^2} dx^2 + \frac{\partial \grave{x}^1}{\partial x^3} dx^3$$
$$d\grave{x}^2 = \frac{\partial \grave{x}^2}{\partial x^1} dx^1 + \frac{\partial \grave{x}^2}{\partial x^2} dx^2 + \frac{\partial \grave{x}^2}{\partial x^3} dx^3$$
$$d\grave{x}^3 = \frac{\partial \grave{x}^3}{\partial x^1} dx^1 + \frac{\partial \grave{x}^3}{\partial x^2} dx^2 + \frac{\partial \grave{x}^3}{\partial x^3} dx^3$$

or, as per Einstein's summation convention, we can condense these equations as

$$d\grave{x}^i = \frac{\partial \grave{x}^i}{\partial x^j} dx^j$$

Therefore, in general, for any contravariant vector we get

$$\grave{a}^i = \frac{\partial \grave{x}^i}{\partial x^j} a^j$$

which is the transformation rule for contravariant vectors (and hence the superscript indices). More precisely, we say that vectors which transform this way are called contravariant vectors. This can be generalized to contravariant tensors as well

$$\grave{A}^{ij} = \frac{\partial \grave{x}^i}{\partial x^k} \frac{\partial \grave{x}^j}{\partial x^l} A^{kl}$$

The key idea is this: Let there be a vector denoted as $\underline{a}$ in the old coordinate system and $\underline{\grave{a}}$ in the new coordinate system. Then, for contravariant tensors, going from old coordinate representation to the new one ($\underline{a} \to \underline{\grave{a}}$), requires *differentiating the new coordinates with respect to the old ones* ($\grave{x}$ with respect to x).

There is another species of vectors which behave differently from the contravariant vectors. This difference is because the differentials of coordinates ($dx_1 \; dx_2 \; dx_3$) appear in the numerator in vectors like $\underline{dx}$, velocity vector $\underline{u}$ or acceleration vector $\underline{a}$, but they appear in the denominator for the gradient vector $\nabla \phi = (\frac{\partial \phi}{\partial x_1}, \frac{\partial \phi}{\partial x_2}, \frac{\partial \phi}{\partial x_3})$ for some scalar ϕ. How does such a vector transform when we go from an old (unprimed) to a new (primed) coordinate system? Recall that

$$\grave{x}_1 \equiv \grave{x}_1(x_1, x_2, x_3)$$
$$\grave{x}_2 \equiv \grave{x}_2(x_1, x_2, x_3)$$
$$\grave{x}_3 \equiv \grave{x}_3(x_1, x_2, x_3)$$

Then, using chain rule,

$$\frac{\partial \phi}{\partial \grave{x}^1} = \frac{\partial \phi}{\partial x^1} \frac{\partial x^1}{\partial \grave{x}^1} + \frac{\partial \phi}{\partial x^2} \frac{\partial x^2}{\partial \grave{x}^1} + \frac{\partial \phi}{\partial x^3} \frac{\partial x^3}{\partial \grave{x}^1}$$
$$\frac{\partial \phi}{\partial \grave{x}^2} = \frac{\partial \phi}{\partial x^1} \frac{\partial x^1}{\partial \grave{x}^2} + \frac{\partial \phi}{\partial x^2} \frac{\partial x^2}{\partial \grave{x}^2} + \frac{\partial \phi}{\partial x^3} \frac{\partial x^3}{\partial \grave{x}^2}$$
$$\frac{\partial \phi}{\partial \grave{x}^3} = \frac{\partial \phi}{\partial x^1} \frac{\partial x^1}{\partial \grave{x}^3} + \frac{\partial \phi}{\partial x^2} \frac{\partial x^2}{\partial \grave{x}^3} + \frac{\partial \phi}{\partial x^3} \frac{\partial x^3}{\partial \grave{x}^3}$$

or, in index notation,
$$\frac{\partial \phi}{\partial \grave{x}^i} = \frac{\partial x^j}{\partial \grave{x}^i} \frac{\partial \phi}{\partial x^j}$$

Then, instead of $\frac{\partial \phi}{\partial x^j}$, if there was a general covariant vector a, we would get
$$\grave{a}_j = \frac{\partial x^i}{\partial \grave{x}^j} a_i$$

This can be generalized for a second order tensor
$$\grave{a}_{ij} = \frac{\partial x^k}{\partial \grave{x}^i} \frac{\partial x^l}{\partial \grave{x}^j} a_{kl}$$

In essence, for a covariant vector $\underline{a}$, going from the old to the new coordinate system requires *differentiation of the old coordinates with respect to the new ones* and therein lies the key difference. Finally, we can also note the transformation rule for mixed tensors
$$\grave{a}_j^i = \frac{\partial x^k}{\partial \grave{x}^i} \frac{\partial \grave{x}^j}{\partial x^l} a_l^k$$

2 Index contraction and the metric tensor

Knowing how the vectors transform under coordinate transformations, we can show an important result: *Contraction of indices always occurs between a contravariant and a covariant index*. Two contravariant or two covariant indices cannot contract amongst each other.

To show this, let there be a scalar $S = a^i b_i$. Being a scalar, it will be invariant to coordinate transformations. Hence we must also have $S = \grave{a}^i \grave{b}_i$. Let's assume that $\grave{a}^i \grave{b}_i = \grave{S}$ for some $\grave{S}$. Then,

$$\grave{S} = \grave{a}^i \grave{b}_i = \frac{\partial \grave{x}^i}{\partial x^k} a^k \frac{\partial x^l}{\partial \grave{x}^i} b_l = \frac{\partial x^l}{\partial x^k} a^k b_l = \delta_k^l a^k b_l = a^k b_k = S$$

Note that the key step in this process is

$$\frac{\partial \grave{x}^i}{\partial x^k} a^k \frac{\partial x^l}{\partial \grave{x}^i} b_l = \frac{\partial x^l}{\partial x^k} a^k b_l$$

This is only possible because one of the contracting vectors is contravariant and the other is covariant so that one of the primed coordinate can cancel the other. Hence, contractions can only occur between two different types of vectors. In general, this idea can be used to lower the indices of higher order

tensors as we will soon see. Another thing to note is that δ^i_j is, in fact, a mixed second order tensor- it has one contravariant and one covariant index.

This idea allows us to define the metric tensor. To begin with, consider the expression for a short length ds in Cartesian coordinate system

$$ds^2 = {dx_1}^2 + {dx_2}^2 + {dx_3}^2$$

For cylindrical coordinates $(x_1 = r, x_2 = \phi, x_3 = z)$, the expression for ds^2 is given by

$$ds^2 = dr^2 + r^2 d\phi^2 + dz^2$$

Here r is the *metric factor*. It converts the coordinate differential $d\phi$ to a small distance $r d\phi$ in cylindrical coordinates. More generally, in an arbitrary coordinate system, we can write

$$ds^2 = h_1^2 {dx_1}^2 + h_2^2 {dx_2}^2 + h_3^2 {dx_3}^2$$

where h_1, h_2, h_3 are the metric factors corresponding to each of the coordinates. Yet, this is not the most general form for ds^2. This is the most general form for the metric in an orthogonal coordinate system. If the coordinate system is non-orthogonal, the form for the metric is given by a more general expression

$$ds^2 = g_{ij} dx^i dx^j$$

For the Cartesian coordinates, $g_{ij} = \delta_{ij}$. For cylindrical coordinates, g_{ij} is such that $\{g_{11} = 1(= h_1^2), g_{22} = r(= h_2^2), g_{33} = 1(= h_3^2)\}$ and the rest of the terms are 0. The fact that g_{ij} is diagonal is an artifact of the orthogonality of the coordinate system. If the coordinate system is non-orthogonal, then the off-diagonal terms can also be non-zero. How does g_{ij} transform upon change of coordinates? Let there be a metric ds^2 defined in the $\grave{x}$ coordinate system as

$$ds^2 = \grave{g}_{ij} d\grave{x}^i d\grave{x}^j$$

Now, in order to migrate to the old coordinates, we impose the transformation rules for contravariant vectors (ds^2, being a scalar, remains invariant to coordinate transformations)

$$ds^2 = \grave{g}_{ij} \frac{\partial \grave{x}^i}{\partial x^k} dx^k \frac{\partial \grave{x}^j}{\partial x^l} dx^l$$

$$\Rightarrow ds^2 = \grave{g}_{ij} \frac{\partial \grave{x}^i}{\partial x^k} \frac{\partial \grave{x}^j}{\partial x^l} dx^k dx^l$$

Comparing this equation with the definition of metric in the old coordinate system $ds^2 = g_{kl}dx^k dx^l$ we get the transformation rule for g as

$$g_{kl} = \frac{\partial \grave{x}^i}{\partial x^k}\frac{\partial \grave{x}^j}{\partial x^l}\grave{g}_{ij}$$

Similarly we can obtain the inverse transformation as

$$\grave{g}_{ij} = \frac{\partial x^k}{\partial \grave{x}^i}\frac{\partial x^l}{\partial \grave{x}^j}g_{kl}$$

Therefore, allowing us to find the metric tensor for the new space from the knowledge of metric tensor for the old space. Specifically, if the old coordinate system is Cartesian, then $g_{kl} = \delta_{kl}$. Let's denote the Cartesian coordinates with y_i's and the new coordinate system with x_i's. Then, the transformation becomes

$$\begin{aligned}
g_{ij} &= \frac{\partial y^k}{\partial x^i}\frac{\partial y^l}{\partial x^j}\delta_{kl} \\
&= \frac{\partial y^k}{\partial x^i}\frac{\partial y^k}{\partial x^j}
\end{aligned}$$

This will allow us to obtain the metric tensor when we go from Cartesian to a general coordinate system.

Let us also define an inverse contravariant metric tensor g^{ij} such that

$$g_{ik}g^{kj} = \delta_j^i$$

This will allow us to perform operations like lowering and raising of indices i.e. to transform a covariant vector into its contravariant representation and vice versa.

$$\begin{aligned}
a_i &= g_{ij}a^j \\
a^i &= g^{ij}a_j
\end{aligned}$$

With this formalism, we can consider a simple application.

3   The Convection-Diffusion equation

Let's say we have simple shear flow in a channel. Let's define two coordinate systems- the Cartesian coordinates which are fixed in space (y_i's) and the

convected coordinates which deform with the flow (x_i's). Then, we have

$$u_1 = \dot{\gamma} y_2$$
$$u_2 = 0$$
$$\Rightarrow y_1 = \dot{\gamma} y_2 t + c_1$$
$$y_2 = c_2$$
$$\Rightarrow c_1 = y_1 - \dot{\gamma} t y_2$$
$$c_2 = y_2$$

c_1 and c_2 are in fact the convected coordinates. At $t = 0$, $c_1 = y_1, c_2 = y_2$. For later times, c_1 remains y_1 while c_2 convects with the flow. Therefore, our convected coordinates become

$$x_1 = y_1 - \dot{\gamma} t y_2$$
$$x_2 = y_2$$

Additionally, we can also find the new basis vectors $(\underline{g}_1, \underline{g}_2)$ using transformation law for vectors

$$\grave{a}_j = \frac{\partial x^i}{\partial \grave{x}^j} a_i$$
$$\Rightarrow g_i = \frac{\partial y^j}{\partial x^i} \underline{1}_j$$
$$\Rightarrow g_1 = \underline{1}_1$$
$$g_2 = \dot{\gamma} t \underline{1}_1 + \underline{1}_2$$

Using chain rule, we can express the derivatives in the new coordinates

$$\frac{\partial}{\partial y_1} = \frac{\partial}{\partial x^1}$$
$$\frac{\partial}{\partial y_2} = -\dot{\gamma} t \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}$$

Now, the convection-diffusion equation equation, in Cartesian coordinates, is

$$\frac{\partial T}{\partial t} + u_\alpha \frac{\partial T}{\partial y_\alpha} = \kappa \frac{\partial^2 T}{\partial y_\alpha y_\alpha}$$

Using the expressions for shear-flow velocity profile

$$\frac{\partial T}{\partial t} + \dot{\gamma} y_2 \frac{\partial T}{\partial y_1} = \kappa \left(\frac{\partial^2 T}{\partial y_1^2} + \frac{\partial^2 T}{\partial y_2^2} \right)$$

6

Expressing the derivatives in the new coordinates, we finally get the convection-diffusion equation in convected coordinates

$$\frac{\partial T}{\partial t} = \kappa \left((1 + (\dot{\gamma}t)^2) \frac{\partial^2 T}{\partial x_1^2} + \frac{\partial^2 T}{\partial x_2^2} - 2\dot{\gamma}t \frac{\partial^2 T}{\partial x^1 \partial x^2} \right)$$

What does this equation tell us? Notice that there is no convection term anymore. We only have the first time derivative proportional to the second order spatial derivatives which is a telltale sign of pure diffusion. Therefore, in the convected coordinates, the convection-diffusion equation appears as just the diffusion equation. But there is one important difference. The last term includes partial derivatives with respect to both x_1 and x_2. Presence of cross-derivatives is an artifact of non-orthogonality in the coordinate system. As time increases, the coefficient of this term increases in magnitude and hence the non-orthogonality in the coordinate system.

It also makes intuitive sense. Since, the coordinates move with the flow, an observer moving with the coordinates will not perceive any flow.

3.1 Method II

In the previous section we arrived at the result using just the coordinate transformation. Now we arrive at the same result using the isotropy of metric tensor. The heat equation is given by $\frac{\partial T}{\partial t} = \kappa \frac{\partial q^i}{\partial x^i}$. Using Fourier's law ($q^i = -\kappa g^{ij} \frac{\partial T}{\partial x^j}$), it becomes

$$\frac{\partial T}{\partial t} = \kappa \frac{\partial}{\partial x^i} \left(g^{ij} \frac{\partial T}{\partial x^j} \right)$$

Note that heat flux is a contravariant vector. This is because its contraction with a covariant vector, $\nabla \cdot q$ is a scalar. Since gradient vector $\frac{\partial T}{\partial x^j}$ is obviously a covariant vector, this implies that we need an isotropic second order contravariant tensor to carry out the transformation. Hence, g^{ij} is a contravariant tensor. Now all we need to do is find g^{ij}.

Firstly, we find the elements of the metric tensor g_{ij}

$$g_{ij} = \sum_{\alpha=1}^{2} \frac{\partial y^k}{\partial x_i} \frac{\partial y^k}{\partial x^j}$$

$$\Rightarrow g_{11} = 1$$
$$g_{22} = 1 + (\dot{\gamma}t)^2$$
$$g_{12} = g_{21} = \dot{\gamma}t$$

Using $g_{ik}g^{kj} = \delta^i_j$, we can find the contravariant metric tensor g^{ij}. Its elements turn out to be

$$g^{11} = 1 + (\dot{\gamma}t)^2$$
$$g^{22} = 1$$
$$g^{12} = g^{21} = -\dot{\gamma}t$$

Substituting the values in the heat equation, we get exactly the same equation as before

$$\frac{\partial T}{\partial t} = \kappa\left((1 + (\dot{\gamma}t)^2)\frac{\partial^2 T}{\partial x_1^2} + \frac{\partial^2 T}{\partial x_2^2} - 2\dot{\gamma}t\frac{\partial^2 T}{\partial x^1 \partial x^2}\right)$$

However, there is a glitch in this process. In order to carry out the outer differentiation in the term $\kappa\frac{\partial}{\partial x^i}(g^{ij}\frac{\partial T}{\partial x^j})$, we need the covariant derivative. A covariant derivative is a form of differentiation that, besides considering the change in components of the vector, also accounts for change in the direction of unit vectors as we traverse a small distance in the space. It would not matter in the above example because the metric tensor g^{ij} does not vary with space. Thus, at any given time, the direction of unit vectors is same at all points in the space. But for a general metric, we must use the covariant derivative to carry out the differentiation in the Fourier's law.

The above said differentiation requires us to differentiate a contravariant tensor and a covariant vector. Once we incorporate the laws for covariant derivative of the contravariant tensor and covariant derivative of the covariant vector, we can write down the heat-equation in a general, non-orthogonal coordinate system.