Tensors

Lecture 7

Physics 411
Classical Mechanics II

September 12th 2007

In Electrodynamics, the implicit law governing the motion of particles is $F^\alpha = m\,\ddot{x}^\alpha$. This is also true, of course, for most of classical physics and the details of the physical principle one is discussing are hidden in F^α, and potentially, its potential. That is what defines the interaction.

In general relativity, the motion of particles will be described by

$$\boxed{\ddot{x}^\mu + \Gamma^\mu{}_{\alpha\beta}\,\dot{x}^\alpha\,\dot{x}^\beta = 0} \tag{7.1}$$

and this will occur in a four-dimensional space-*time* – but that doesn't concern us for now. The point of the above is that it lacks a potential, and can be connected in a natural way to the metric.

7.1 Introduction in Two Dimensions

We will start with some basic examples in two-dimensions for concreteness. Here we will always begin in a Cartesian-parametrized plane, with basis vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$.

7.1.1 Rotation

To begin, consider a simple rotation of the usual coordinate axes through an angle θ (counterclockwise) as shown in Figure 7.1.

From the figure, we can easily relate the coordinates w.r.t. the new axes ($\bar{x}$ and $\bar{y}$) to coordinates w.r.t. the "usual" axes (x and y) – define $\ell =$

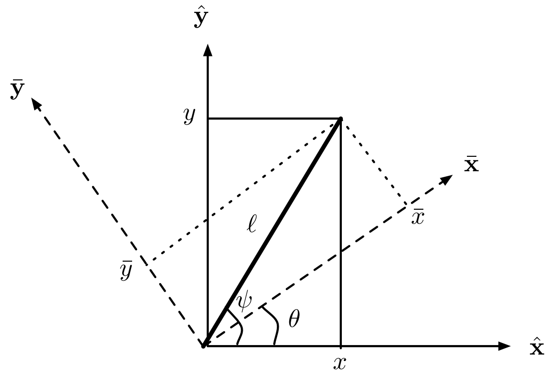

Figure 7.1: Two sets of axes, rotated through an angle θ with respect to each other.

$\sqrt{x^2 + y^2}$, then the invariance of length allows us to write

$$\begin{aligned} \bar{x} &= \ell\,\cos(\psi - \theta) = \ell\,\cos\psi\,\cos\theta + \ell\,\sin\psi\,\sin\theta = x\,\cos\theta + y\,\sin\theta \\ \bar{y} &= \ell\,\sin(\psi - \theta) = \ell\,\sin\psi\,\cos\theta - \ell\,\cos\psi\,\sin\theta = y\,\cos\theta - x\,\sin\theta, \end{aligned} \tag{7.2}$$

which can be written in matrix form as:

$$\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}}_{\equiv \mathbb{R} \doteq R^{\alpha}{}_{\beta}} \begin{pmatrix} x \\ y \end{pmatrix}. \tag{7.3}$$

If we think of infinitesimal displacements centered at the origin, then we would write $d\bar{x}^\alpha = R^\alpha{}_\beta\, dx^\beta$. Notice that in this case, as a matrix, the inverse of $\mathbb{R}$ is

$$\mathbb{R}^{-1} \doteq \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{7.4}$$

and this is also equal to the transpose of $\mathbb{R}$, a defining property of "rotation" or "orthogonal" matrices: $\mathbb{R}^{-1} = \mathbb{R}^T$, in this case coming from the observation that one coordinate system's clockwise rotation is another's counter-clockwise, and symmetry properties of the trigonometric functions.

Scalar

This one is easy – a scalar "does not" transform under coordinate transformations – if we have a function $\phi(x, y)$ in the first coordinate system, then we have:

$$\boxed{\bar{\phi}(\bar{x}, \bar{y}) = \phi(x(\bar{x}, \bar{y}), y(\bar{x}, \bar{y}))} \tag{7.5}$$

so that operationally, all we do is replace the x and y appearing in the definition of ϕ with the relation $x = \bar{x}\,\cos\theta - \bar{y}\,\sin\theta$ and $y = \bar{x}\,\sin\theta + \bar{y}\,\cos\theta$.

Vector (Contravariant)

With the fabulous success and ease of scalar transformations, we are led to a natural definition for an object that transforms as a *vector*. If we write a generic vector in the original coordinate system:

$$\mathbf{f} = f^x(x, y)\,\hat{\mathbf{x}} + f^y(x, y)\,\hat{\mathbf{y}}, \tag{7.6}$$

then we'd like the transformation to be defined as:

$$\boxed{\bar{\mathbf{f}} = \bar{f}^x(\bar{x}, \bar{y})\,\bar{\mathbf{x}} + \bar{f}^y(\bar{x}, \bar{y})\,\bar{\mathbf{y}} = f^x(x, y)\,\hat{\mathbf{x}} + f^y(x, y)\,\hat{\mathbf{y}},} \tag{7.7}$$

i.e. identical to the scalar case, but with the non-trivial involvement of the basis vectors.

From Figure 7.1, we can easily see how to write the basis vector $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ in terms of $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$:

$$\begin{aligned} \bar{\mathbf{x}} &= \cos\theta\,\hat{\mathbf{x}} + \sin\theta\,\hat{\mathbf{y}} \\ \bar{\mathbf{y}} &= -\sin\theta\,\hat{\mathbf{x}} + \cos\theta\,\hat{\mathbf{y}}, \end{aligned} \tag{7.8}$$

and using this, we can write $\bar{\mathbf{f}}$ in terms of the original basis vectors, the elements in front of these will then define f^x and f^y according to our target.

$$\bar{\mathbf{f}} = \left(\bar{f}^x\,\cos\theta - \bar{f}^y\,\sin\theta\right)\,\hat{\mathbf{x}} + \left(\bar{f}^x\,\sin\theta + \bar{f}^y\cos\theta\right)\,\hat{\mathbf{y}}, \tag{7.9}$$

so that we learn (slash demand) that

$$\begin{aligned} f^x &= \bar{f}^x\,\cos\theta - \bar{f}^y\,\sin\theta \\ f^y &= \bar{f}^x\,\sin\theta + \bar{f}^y\,\cos\theta, \end{aligned} \tag{7.10}$$

or, inverting, we have, now in matrix form:

$$\begin{pmatrix} \bar{f}^x \\ \bar{f}^y \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} f^x \\ f^y \end{pmatrix}. \tag{7.11}$$

Notice that the "components" can be written as $\bar{f}^\alpha = R^\alpha{}_\beta \, f^\beta$ from the above, and this leads to the usual statement that "a (contravariant) vector transforms 'like' the coordinates themselves." (more appropriately, like the coordinate differentials, which are actual vectors – recall $d\bar{x}^\alpha = R^\alpha{}_\beta \, dx^\beta$ from above).

The rotation matrix and transformation is just one example, and the form is telling – notice that (somewhat sloppily)

$$d\bar{x}^\alpha = R^\alpha{}_\beta \, dx^\beta \longrightarrow \frac{d\bar{x}^\alpha}{dx^\beta} = R^\alpha{}_\beta \tag{7.12}$$

and so it is tempting to generalize the contravariant vector transformation law to include an arbitrary relation between "new" and "old" coordinates:

$$\boxed{\bar{f}^\alpha = \frac{d\bar{x}^\alpha}{dx^\beta} \, f^\beta,} \tag{7.13}$$

precisely what we defined for a contravariant vector originally.

7.1.2 Non-Orthogonal Axes

We will stick with linear transformations for a moment, but now allow the axes to be skewed – this will allow us to distinguish between the above contravariant transformation law and the covariant (one-form) transformation law.

Consider the two coordinate systems shown in Figure 7.2 – from the figure, we can write the relation for $d\bar{x}^\alpha$ in terms of dx^α:

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} 1 & \cos\theta \\ 0 & \sin\theta \end{pmatrix} \begin{pmatrix} d\bar{x} \\ d\bar{y} \end{pmatrix} \tag{7.14}$$

or, we can write in the more standard form (inverting the above):

$$\begin{pmatrix} d\bar{x} \\ d\bar{y} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & -\cot\theta \\ 0 & \frac{1}{\sin\theta} \end{pmatrix}}_{\equiv \mathbb{T} \doteq T^\alpha{}_\beta} \begin{pmatrix} dx \\ dy \end{pmatrix}, \tag{7.15}$$

so that $d\bar{x}^\alpha = T^\alpha{}_\beta \, dx^\beta$ and the matrix $\mathbb{T}$ is the inverse of the matrix in (7.14).

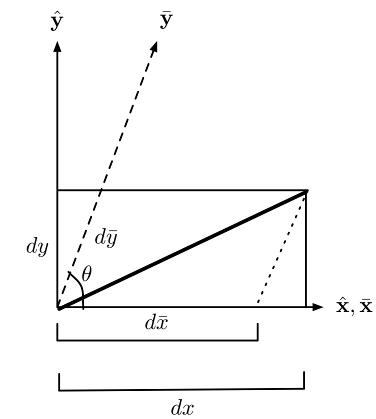

Figure 7.2: Original, orthogonal Cartesian axes, and a skewed set with coordinate differential shown (parallel projection is used since differentials are contravariant).

Let us first check the contravariant form $d\mathbf{x} = dx\,\hat{\mathbf{x}} + dy\,\hat{\mathbf{y}}$ transforms appropriately. We need the basis vectors in the skewed coordinate system – from the figure, these are $\bar{\mathbf{x}} = \hat{\mathbf{x}}$ and $\bar{\mathbf{y}} = \cos\theta\,\hat{\mathbf{x}} + \sin\theta\,\hat{\mathbf{y}}$, then

$$\begin{aligned} d\bar{\mathbf{x}} &= d\bar{x}\,\bar{\mathbf{x}} + d\bar{y}\,\bar{\mathbf{y}} \\ &= (dx - \cot\theta\,dy)\,\hat{\mathbf{x}} + \frac{dy}{\sin\theta}(\cos\theta\,\hat{\mathbf{x}} + \sin\theta\,\hat{\mathbf{y}}) \\ &= dx\,\hat{\mathbf{x}} + dy\,\hat{\mathbf{y}} \end{aligned} \tag{7.16}$$

as expected.

Recall that we developed the metric for this type of situation previously – if we take the scalar length $dx^2 + dy^2$, and write it in terms of the $d\bar{x}$ and $d\bar{y}$ infinitesimals, then

$$dx^2 + dy^2 = d\bar{x}^2 + d\bar{y}^2\, 2\, d\bar{x}\, d\bar{y}\, \cos\theta \doteq \begin{pmatrix} d\bar{x} & d\bar{y} \end{pmatrix} \underbrace{\begin{pmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{pmatrix}}_{\equiv \bar{g}_{\mu\nu}} \begin{pmatrix} d\bar{x} \\ d\bar{y} \end{pmatrix}, \tag{7.17}$$

defining the metric. We have the relation $dx^\alpha g_{\alpha\beta}\, dx^\beta = d\bar{x}^\alpha\, \bar{g}_{\alpha\beta}\, d\bar{x}^\beta$, which tells us how the metric itself transforms. From the definition of the transformation, $d\bar{x}^\alpha = T^\alpha_{\ \beta}\, dx^\beta$, we have $\frac{d\bar{x}^\alpha}{dx^\beta} = T^\alpha_{\ \beta}$ as usual, but we can also write the inverse transformation, making use of the inverse of the matrix $\mathbb{T}$

in (7.14). Letting $\tilde{T}^\alpha_{\ \beta} \doteq \mathbb{T}^{-1}$,

$$dx^\alpha = \tilde{T}^\alpha_{\ \beta}\, d\bar{x}^\beta \longrightarrow \frac{dx^\alpha}{d\bar{x}^\beta} = \tilde{T}^\alpha_{\ \beta}. \tag{7.18}$$

Now going back to the invariant,

$$d\bar{x}^\alpha\, \bar{g}_{\alpha\beta}\, d\bar{x}^\beta = (T^\alpha_{\ \gamma}\, dx^\gamma)\, \bar{g}_{\alpha\beta}\, (T^\beta_{\ \rho}\, dx^\rho). \tag{7.19}$$

For this to expression to equal $dx^\alpha g_{\alpha\beta}\, dx^\beta$, we must have

$$\bar{g}_{\alpha\beta} = \tilde{T}^\sigma_{\ \alpha}\, g_{\sigma\delta}\, \tilde{T}^\delta_{\ \beta}, \tag{7.20}$$

which we can put in as a check:

$$\begin{aligned} d\bar{x}^\alpha\, \bar{g}_{\alpha\beta}\, d\bar{x}^\beta &= (T^\alpha_{\ \gamma}\, dx^\gamma)\, \bar{g}_{\alpha\beta}\, (T^\beta_{\ \rho}\, dx^\rho) \\ &= (T^\alpha_{\ \gamma}\, dx^\gamma) \left(\tilde{T}^\sigma_{\ \alpha}\, g_{\sigma\delta}\, \tilde{T}^\delta_{\ \beta}\right) (T^\beta_{\ \rho}\, dx^\rho) \\ &= \delta^\sigma_{\ \gamma}\, \delta^\delta_{\ \rho}\, g_{\sigma\delta}\, dx^\gamma\, dx^\rho \\ &= dx^\sigma\, g_{\sigma\delta}\, dx^\delta. \end{aligned} \tag{7.21}$$

The metric transformation law (7.20) is different from that of a vector – first of all, there are two indices, but more important, rather than transforming with $T^\alpha_{\ \beta}$, the covariant indices transform with the inverse of the transformation. This type of transformation is not obvious under rotations since rotations leave the metric itself invariant ($\mathbb{R}^T\, \boldsymbol{g}\, \mathbb{R} = \boldsymbol{g}$), and we don't notice the transformation of the metric.

Covariant "vector" Transformation

If we generalize the above covariant transformation to non-linear coordinate relations, we define covariant vector transformation as:

$$\bar{f}_\alpha = \frac{dx^\beta}{d\bar{x}^\alpha}\, f_\beta. \tag{7.22}$$

For a second rank covariant tensor, like the metric $g_{\mu\nu}$, we just introduce a copy of the transformation for each index (the same is true for the contravariant transformation law):

$$\bar{g}_{\mu\nu} = \frac{dx^\alpha}{d\bar{x}^\mu}\, \frac{dx^\beta}{d\bar{x}^\nu}\, g_{\alpha\beta} \tag{7.23}$$

Now the question becomes: For contravariant vectors, the model was dx^α, what is the model covariant vector? The answer is, "the gradient". Consider a scalar $\phi(x, y)$ – we can write its partial derivaties w.r.t. the coordinates as:

$$\frac{\partial \phi(x,y)}{\partial x^\alpha} \equiv \phi_{,\alpha}, \tag{7.24}$$

and ask how this transforms under any coordinate transformation – the answer is provided by the chain rule. We have the usual scalar transformation $\bar\phi(\bar x, \bar y) = \phi(x(\bar x, \bar y), y(\bar x, \bar y))$, so

$$\bar\phi_{,\alpha} = \frac{\partial \bar\phi}{\partial \bar x^\alpha} = \frac{\partial \phi}{\partial x^\beta}\,\frac{\partial x^\beta}{\partial \bar x^\alpha} = \frac{\partial x^\beta}{\partial \bar x^\alpha}\,\phi_{,\beta}, \tag{7.25}$$

precisely the rule for covariant vector transformation.

How does this compare with our usual $\nabla\phi = \frac{\partial\phi}{\partial x}\,\hat{\mathbf{x}} + \frac{\partial\phi}{\partial y}\,\hat{\mathbf{y}}$? Almost automatically, we realize that this is a contravariant vector – it is written in terms of the basis vector components just as f^α was, so we expect to find that $\nabla\phi$ is shorthand for $\phi_{,}{}^{\alpha}$ rather than $\phi_{,\alpha}$. There are deep distinctions between covariant and contravariant tensors, and we have not addressed them in full abstraction here – for our purposes, there is a one-to-one map between the two provided by the metric, and that makes them effectively equivalent. It is important, in some settings, to keep track of the fundamental nature of an object, covariant or contravariant (as is the case with, for example, momentum), but for now, it suffices to define the relation between contravariant and covariant objects via the metric: For a contravariant vector f^α,

$$f_\alpha \equiv g_{\alpha\beta}\, f^\beta. \tag{7.26}$$

We further define $g^{\alpha\beta}$ to be the contravariant form of the metric, its numerical matrix inverse: $g^{\alpha\beta}\, g_{\beta\gamma} = \delta^\alpha_\gamma$.

Now we can see that:

$$\bar\phi_{,}{}^{\alpha} = \bar g^{\alpha\beta}\,\bar\phi_{,\beta} \tag{7.27}$$

and for our skew-axis transformation, we have:

$$\bar g_{\alpha\beta} \doteq \begin{pmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{pmatrix} \bar g^{\alpha\beta} \doteq \begin{pmatrix} \frac{1}{\sin^2\theta} & -\frac{\cos\theta}{\sin^2\theta} \\ -\frac{\cos\theta}{\sin^2\theta} & \frac{1}{\sin^2\theta} \end{pmatrix} \tag{7.28}$$

so that

$$\bar g^{\alpha\beta}\,\bar\phi_{,\beta} \doteq \frac{1}{\sin\theta}\left(\frac{1}{\sin\theta}\,\frac{\partial\bar\phi}{\partial\bar x} - \frac{\cos\theta}{\sin\theta}\,\frac{\partial\bar\phi}{\partial\bar y}\right)\bar{\mathbf{x}} + \frac{1}{\sin\theta}\left(-\frac{\cos\theta}{\sin\theta}\,\frac{\partial\bar\phi}{\partial\bar x} + \frac{1}{\sin\theta}\,\frac{\partial\bar\phi}{\partial\bar y}\right)\bar{\mathbf{y}}, \tag{7.29}$$

and

$$\begin{aligned} \frac{\partial \bar{\phi}}{\partial \bar{x}} &= \frac{\partial \phi}{\partial x}\frac{\partial x}{\partial \bar{x}} + \frac{\partial \phi}{\partial y}\frac{\partial y}{\partial \bar{x}} = \frac{\partial \phi}{\partial x} \\ \frac{\partial \bar{\phi}}{\partial \bar{y}} &= \frac{\partial \phi}{\partial x}\frac{\partial x}{\partial \bar{y}} + \frac{\partial \phi}{\partial y}\frac{\partial y}{\partial \bar{y}} = \frac{\partial \phi}{\partial x}\cos\theta + \frac{\partial \phi}{\partial y}\sin\theta, \end{aligned} \tag{7.30}$$

which we can input into (7.29):

$$\begin{aligned} \bar{g}^{\alpha\beta}\bar{\phi}_{,\beta} &\doteq \frac{1}{\sin\theta}\left(\frac{1}{\sin\theta}\frac{\partial \phi}{\partial x} - \frac{\partial \phi}{\partial x}\frac{\cos^2\theta}{\sin\theta} - \frac{\partial \phi}{\partial y}\cos\theta\right)\bar{\mathbf{x}} + \frac{1}{\sin\theta}\left(-\frac{\cos\theta}{\sin\theta}\frac{\partial \phi}{\partial x} + \frac{\cos\theta}{\sin\theta}\frac{\partial \phi}{\partial x} + \frac{\partial \phi}{\partial y}\right)\bar{\mathbf{y}} \\ &= \left(\frac{\partial \phi}{\partial x} - \cot\theta\frac{\partial \phi}{\partial y}\right)\bar{\mathbf{x}} + \frac{1}{\sin\theta}\frac{\partial \phi}{\partial y}\bar{\mathbf{y}}, \end{aligned} \tag{7.31}$$

or, finally, replacing the unit vectors with the original set

$$\begin{aligned} \bar{g}^{\alpha\beta}\,\bar{\phi}_{,\beta} &\doteq \left(\frac{\partial \phi}{\partial x} - \cot\theta\frac{\partial \phi}{\partial y}\right)\bar{\mathbf{x}} + \frac{1}{\sin\theta}\frac{\partial \phi}{\partial y}\bar{\mathbf{y}} \\ &= \frac{\partial \phi}{\partial x}\hat{\mathbf{x}} + \frac{\partial \phi}{\partial y}\hat{\mathbf{y}} \end{aligned} \tag{7.32}$$

and we see that this is the correct expression for a contravariant vector: $\bar{\nabla}\bar{\phi} = \nabla\phi$.

7.2 Transformation and Basis

We have discussed the transformation laws for our various tensors, but let's look in more detail: consider a vector written in the usual Cartesian coordinates

$$x^\alpha \doteq \mathbf{x} \doteq x\,\hat{x} + y\,\hat{y} + z\,\hat{z}. \tag{7.33}$$

Our aim is to transform this to a new set of coordinates (r, θ, ϕ) related to these in some manner. The transformation laws for tensors involve objects like $\frac{\partial x'^\alpha}{\partial x^\beta}$, so we must be able to form $x^\alpha(x')$ and vice-versa (a statement about the non-singular nature of the transformation)

$$\begin{aligned} x^1(r,\theta,\phi) &= r\,\sin\theta\,\cos\phi \\ x^2 &= r\,\sin\theta\,\sin\phi \\ x^3 &= r\,\cos\theta \end{aligned} \tag{7.34}$$

Just by the chain rule, then, we have

$$dx^\alpha = \frac{\partial x^\alpha}{\partial x'^\beta}\,dx'^\beta. \tag{7.35}$$

In matrix form

$$\begin{aligned} dx^\alpha &= \frac{\partial x^\alpha}{\partial x'^\beta}\, dx'^\beta \\ &\doteq \begin{pmatrix} \sin\theta\,\cos\phi & r\,\cos\theta\,\cos\phi & -r\,\sin\theta\,\sin\phi \\ \sin\theta\,\sin\phi & r\,\cos\theta\,\sin\phi & r\,\sin\theta\,\cos\phi \\ \cos\theta & -r\,\sin\theta & 0 \end{pmatrix} \begin{pmatrix} dr \\ d\theta \\ d\phi \end{pmatrix} \\ &= dr\,\mathbf{e_r} + d\theta\,\mathbf{e}_\theta + d\phi\,\mathbf{e}_\phi. \end{aligned} \tag{7.36}$$

This last line is important – keep in mind that matrix-vector multiplication can be viewed as $A_{ij}\,x_j = [A_k x_k]$ where the right-hand side means "multiply the k^{th} column of $\mathbb{A}$ by x_k. So the above tells us what we mean by the *vectors* $\mathbf{e_r}$, $\mathbf{e}_\theta$, $\mathbf{e}_\phi$. These are "basis vectors" as usual. They look a little funny sitting here, we are used to a slightly different type of basis vector. If you consult Griffiths, you will find that his spherical basis vectors are written, in terms of ours, as

$$\boxed{\hat{\mathbf{e}}_\mathbf{r} = \mathbf{e_r} \qquad \hat{\mathbf{e}}_\theta = \frac{1}{r}\,\mathbf{e}_\theta \qquad \hat{\mathbf{e}}_\phi = \frac{1}{r\sin\theta}\,\mathbf{e}_\phi} \tag{7.37}$$

i.e. they are *normalized*.

So far, so good – if we want to represent $d\mathbf{x} = dx^1\,\hat{x} + dx^2\,\hat{y} + dx^3\,\hat{z}$ in terms of (r, θ, ϕ) and $(\hat{\mathbf{e}}_\mathbf{r}, \hat{\mathbf{e}}_\theta, \hat{\mathbf{e}}_\phi)$, we know how to do it. Now the question is, if we had a vector f^α on the Cartesian side, so that $\mathbf{f} = f^1\,\hat{\mathbf{x}} + f^2\,\hat{\mathbf{y}} + f^3\,\hat{\mathbf{z}}$, what is this vector written in terms of the spherical coordinate system and basis? Well, we know how to write the Cartesian basis in terms of the spherical one, and we get:

$$\begin{aligned} \mathbf{f}' =& \left(\left(f^1\,\cos\phi + f^2\,\sin\phi\right)\,\sin\theta + f^3\,\cos\theta\right)\,\hat{\mathbf{e}}_\mathbf{r} \\ &+ \left(f^1\,\cos\phi\,\cos\theta + f^2\,\cos\theta\,\sin\phi - f^3\,\sin\theta\right)\,\hat{\mathbf{e}}_\theta \\ &+ \left(f^2\,\cos\phi - f^1\,\sin\phi\right)\,\hat{\mathbf{e}}_\phi \end{aligned} \tag{7.38}$$

and we can further write this in terms of the more natural basis $(\mathbf{e_r}, \mathbf{e}_\theta, \mathbf{e}_\phi)$:

$$\begin{aligned} \mathbf{f}' =& \left(f^3\,\cos\theta + \left(f^1\,\cos\phi + f^2\,\sin\phi\right)\,\sin\theta\right)\,\mathbf{e_r} \\ &+ \frac{1}{r}\left(f^1\,\cos\phi\,\cos\theta + f^2\,\cos\theta\,\sin\phi - f^3\,\sin\theta\right)\,\mathbf{e}_\theta \\ &+ \frac{1}{r\sin\theta}\left(f^2\,\cos\phi - f^1\,\sin\phi\right)\,\mathbf{e}_\phi, \end{aligned} \tag{7.39}$$

which leads us to *the point* of all this – we can call the elements of the above:

$$
\begin{aligned}
f'^\alpha \doteq & \begin{pmatrix} f^3\,\cos\theta + \left(f^1\,\cos\phi + f^2\,\sin\phi\right)\,\sin\theta \\ \frac{1}{r}\left(f^1\,\cos\phi\,\cos\theta + f^2\,\cos\theta\,\sin\phi - f^3\,\sin\theta\right) \\ \frac{1}{r\,\sin\theta}\left(f^2\,\cos\phi - f^1\,\sin\phi\right) \end{pmatrix} \\
\doteq & \begin{pmatrix} \sin\theta\,\cos\phi & \sin\theta\,\sin\phi & \cos\theta \\ \frac{1}{r}\,\cos\theta\,\cos\phi & \frac{1}{r}\,\cos\theta\,\sin\phi & -\frac{1}{r}\,\sin\theta \\ -\frac{1}{r\,\sin\theta}\,\sin\phi & \frac{1}{r\,\sin\theta}\,\cos\phi & 0 \end{pmatrix} \begin{pmatrix} f^1 \\ f^2 \\ f^3 \end{pmatrix}.
\end{aligned}
\tag{7.40}
$$

Comparing this to (7.36) (invert the matrix you find there to verify that $\frac{\partial x^\alpha}{\partial x'^\beta}\,\frac{\partial x'^\beta}{\partial x^\gamma} = \delta^\alpha_\gamma$) we see that

$$
f'^\alpha = \frac{\partial x'^\alpha}{\partial x^\beta}\,f^\beta. \tag{7.41}
$$

7.3 Derivatives

So, we see that we might write: $\mathbf{f} = f^\beta\,\mathbf{e}_\beta$ as a representation of the vector in a basis (or “frame”). The “gradient” of $\mathbf{f}$ is

$$
\frac{\partial \mathbf{f}}{\partial x^\alpha} = \frac{\partial f^\beta}{\partial x^\alpha}\,\mathbf{e}_\beta + f^\beta\,\frac{\partial \mathbf{e}_\beta}{\partial x^\alpha} \tag{7.42}
$$

which is the usual sort of statement: we must take the derivatives of the function *and* the basis if it depends on coordinates.

Let us be clear, and I’ll do this on the tensor side entirely, since I don’t like mixing it up. Think of a contravariant vector transformation:

$$
f'^\alpha = \frac{\partial x'^\alpha}{\partial x^\beta}\,f^\beta, \tag{7.43}
$$

and take $\frac{\partial}{\partial x'^\gamma}$ of this, then by the chain rule:

$$
\frac{\partial f'^\alpha}{\partial x'^\gamma} = \frac{\partial x'^\alpha}{\partial x^\beta}\left(\frac{\partial x^\sigma}{\partial x'^\gamma}\,\frac{\partial f^\beta}{\partial x^\sigma}\right) + \frac{\partial^2 x'^\alpha}{\partial x^\beta \partial x^\rho}\,\frac{\partial x^\rho}{\partial x'^\gamma}\,f^\beta. \tag{7.44}
$$

Now consider the transformation rule for a second rank, mixed tensor:

$$
\boxed{f'^\alpha{}_\beta = \frac{\partial x'^\alpha}{\partial x^\gamma}\,\frac{\partial x^\sigma}{\partial x'^\beta}\,f^\gamma{}_\sigma} \tag{7.45}
$$

so comparing to (7.44), we see that the second term is the problem one!

That is to say, the object $f^\alpha{}_{,\gamma}$ is *not* a tensor. This must be fixed immediately.

7.4 What is a Tensorial Derivative?

The usual definition of a derivative is in terms of a limit like

$$f'(x) = \lim_{\delta x \to 0} \frac{f(x + \delta x) - f(x)}{\delta x}. \tag{7.46}$$

Suppose we expand a tensor $f^\alpha(x+\delta x)$ about the point P (with coordinate x), this is just Taylor's theorem

$$f^\alpha(x + \delta x) = f^\alpha(x) + \delta x^\gamma \, f^\alpha_{\ ,\gamma} + O(\delta x^2) \tag{7.47}$$

and we would say that the (non-tensorial) change of the vector is given by $\Delta f^\alpha \equiv f^\alpha(x + \delta x) - f(x) = \delta x^\gamma \, f^\alpha_{\ ,\gamma}$.

Instead of using this Δf^α, which, since it involves the "usual" derivative cannot be a tensor, we introduce a fudge factor. So we will compare $f^\alpha(x + \delta x)$ with $f^\alpha(x) + \delta f^\alpha(x)$ – that is, we will define the "covariant" derivative in terms of

$$f^\alpha_{\ ;\gamma}(x) \equiv \lim_{\delta x \to 0} \frac{f^\alpha(x + \delta x) - (f^\alpha(x) + \delta f^\alpha(x))}{\delta x^\gamma}. \tag{7.48}$$

What form should this "extra bit" $\delta f^\alpha(x)$ take? Well, I suggest that in the limiting case that $f^\alpha(x) = 0$, it should be zero, and it should also tend to zero with $\delta x \to 0$, i.e. if we don't move from the point P, we should recover the usual derivative. This suggests an object of the form $\sim \delta x^\rho \, f^\tau$, and we still need a contravariant index to form δf^α – so we propose:

$$\delta f^\alpha \equiv -C^\alpha_{\ \rho\tau} \, \delta x^\rho \, f^\tau \tag{7.49}$$

(the negative sign is convention) where $C^\alpha_{\ \rho\tau}$ is just some triply-indexed object that we don't know anything about yet. Our new derivative, the "covariant derivative" will be written

$$\boxed{\begin{aligned} f^\alpha_{\ ;\gamma}(x) &\equiv \lim_{\delta x \to 0} \frac{f^\alpha(x + \delta x) - (f^\alpha(x) + \delta f^\alpha(x))}{\delta x^\gamma} \\ &= \lim_{\delta x \to 0} \frac{f^\alpha(x + \delta x) - f^\alpha(x)}{\delta x^\gamma} + C^\alpha_{\ \gamma\tau} \, f^\tau \\ &= f^\alpha_{\ ,\gamma}(x) + C^\alpha_{\ \gamma\tau}(x) \, f^\tau(x) \end{aligned}} \tag{7.50}$$

But I haven't really done anything, just suggested that we move a vector from point to point in some $C^\alpha_{\ \beta\gamma}$-dependent way rather than the usual transport operator (the derivative). The whole motivation for doing this was to ensure that our notion of derivative was tensorial. This imposes a constraint on $C^\alpha_{\ \beta\gamma}$ (which is arbitrary right now) – let's work it out. Using (7.44)

$$\begin{aligned} f'^\alpha_{\ ;\gamma} &= \frac{\partial x'^\alpha}{\partial x^\beta}\frac{\partial x^\sigma}{\partial x'^\gamma}\frac{\partial f^\beta}{\partial x^\sigma} + \frac{\partial^2 x'^\alpha}{\partial x^\beta\,\partial x^\rho}\frac{\partial x^\rho}{\partial x'^\gamma} f^\beta + C'^\alpha_{\ \beta\gamma}\frac{\partial x'^\beta}{\partial x^\sigma} f^\sigma \\ &= \frac{\partial x'^\alpha}{\partial x^\beta}\frac{\partial x^\sigma}{\partial x'^\gamma}\left(f^\beta_{\ ;\sigma} - f^\rho\, C^\beta_{\ \sigma\rho}\right) + \frac{\partial^2 x'^\alpha}{\partial x^\beta\,\partial x^\rho}\frac{\partial x^\rho}{\partial x'^\gamma} f^\beta + C'^\alpha_{\ \beta\gamma}\frac{\partial x'^\beta}{\partial x^\sigma} f^\sigma, \end{aligned} \tag{7.51}$$

the term in parenthesis comes from noting that $f^\alpha_{\ ,\gamma} = f^\alpha_{\ ;\gamma} - C^\alpha_{\ \beta\gamma} f^\beta$. Collecting everything that is "wrong" in the above, we can define the transformation rule for $C^\alpha_{\ \beta\gamma}$

$$f'^\alpha_{\ ;\gamma} = \frac{\partial x'^\alpha}{\partial x^\beta}\frac{\partial x^\sigma}{\partial x'^\gamma} f^\beta_{\ ;\sigma} + \left(-\frac{\partial x'^\alpha}{\partial x^\beta}\frac{\partial x^\tau}{\partial x'^\gamma} C^\beta_{\ \tau\rho} + \frac{\partial^2 x'^\alpha}{\partial x^\rho\,\partial x^\tau}\frac{\partial x^\tau}{\partial x'^\gamma} + C'^\alpha_{\ \beta\gamma}\frac{\partial x'^\beta}{\partial x^\rho}\right) f^\rho. \tag{7.52}$$

In order to kill the offending term in parentheses, we define

$$\boxed{C'^\alpha_{\ \tau\gamma} = \frac{\partial x^\lambda}{\partial x'^\tau}\frac{\partial x'^\alpha}{\partial x^\rho}\frac{\partial x^\sigma}{\partial x'^\gamma} C^\rho_{\ \sigma\lambda} - \frac{\partial^2 x'^\alpha}{\partial x^\lambda\,\partial x^\rho}\frac{\partial x^\rho}{\partial x'^\gamma}\frac{\partial x^\lambda}{\partial x'^\tau}} \tag{7.53}$$

and from this we conclude that $C^\alpha_{\ \beta\gamma}$ is itself not a tensor. Anything that transforms according to (7.53) is called a "connection". What we've done in (7.50) is add two things, neither of which is a tensor, in a sum that produces a tensor. Effectively, the non-tensor parts of the two terms kill each other (by construction, of course, that's what gave us (7.53)), we just needed a little extra freedom in our definition of derivative. For better or worse, we have it! This $f^\alpha_{\ ;\beta}$ is called the covariant derivative, and plays the role, in tensor equations, of the usual $f^\alpha_{\ ,\beta}$.