Solutions to the exercises in 2.1

SUMMARY OF VECTOR TRANSFORM RULES

Contravariant vectors: $V^{m'} = \dfrac{\partial x^{m'}}{\partial x^m} V^m$

Covariant vectors: $V_{m'} = \dfrac{\partial x^m}{\partial x^{m'}} V_m$

---

1. We have seen many times that when the length scale is scaled down from meters to centimeters, the coordinate transform is

$$\begin{cases} x^{1'} = 100x^1 \\ x^{2'} = 100x^2 \end{cases}$$

Using the formal transformation rules, explicitly show that under this transformation, the components of contravariant vectors scale up, and the components of covariant vectors scale down.

Soln. The necessary derivatives are

$$\frac{\partial x^{1'}}{\partial x^1} = 100\,, \ \frac{\partial x^{1'}}{\partial x^2} = 0\,, \ \frac{\partial x^{2'}}{\partial x^1} = 0\,, \ \frac{\partial x^{2'}}{\partial x^2} = 100$$

For a contravariant vector V^m, we thus have

$$\begin{aligned} V^{1'} &= \frac{\partial x^{1'}}{\partial x^m} V^m \\ &= \frac{\partial x^{1'}}{\partial x^1} V^1 + \frac{\partial x^{1'}}{\partial x^2} V^2 \\ &= 100V^1 + 0V^2 \\ &= 100V^1 \end{aligned}$$

and

$$\begin{aligned} V^{2'} &= \frac{\partial x^{2'}}{\partial x^m} V^m \\ &= \frac{\partial x^{2'}}{\partial x^1} V^1 + \frac{\partial x^{2'}}{\partial x^2} V^2 \\ &= 0V^1 + 100V^2 \\ &= 100V^2 \end{aligned}$$

Hence both componenets of V^m scale up by a factor of 100 when being transformed into $V^{m'}$ in the prime frame.

Covariant vectors are similar: for a covariant vector V_m, we thus have

$$\begin{aligned} V_{1'} &= \frac{\partial x^m}{\partial x^{1'}} V_m \\ &= \frac{\partial x^1}{\partial x^{1'}} V_1 + \frac{\partial x^2}{\partial x^{1'}} V_2 \\ &= \frac{1}{100} V_1 + 0 V_2 \\ &= \frac{1}{100} V_1 \end{aligned}$$

and

$$\begin{aligned} V_{2'} &= \frac{\partial x^m}{\partial x^{2'}} V_m \\ &= \frac{\partial x^1}{\partial x^{2'}} V_1 + \frac{\partial x^2}{\partial x^{2'}} V_2 \\ &= 0 V_1 + \frac{1}{100} V_2 \\ &= \frac{1}{100} V_2 \end{aligned}$$

Note: it is important to realize that m and m' are **completely different, unrelated indices**. m' is an actual index: the contravariant transform equation holds for all $m' = 1', 2', \ldots$. When we wrote down $V^{1'} = \frac{\partial x^{1'}}{\partial x^m} V^m$, we weren't plugging in 1 as m; rather, we were plugging in $1'$ as m'. Essentially we are taking the contravariant equation with $m' = 1'$. $1'$ is the first coordinate of the primed frame. On the other hand, m is simply a dummy summation index. Another way to see that we weren't plugging in 1 as m is to rename the dummy summation index to, say, p. Then the $m' = 1'$ equation is $V^{1'} = \frac{\partial x^{1'}}{\partial x^p} V^p$, and there aren't any m to plug into.

---

2. We have seen how to "derive" the covariant transform from matching indices if you forgot it. Say you have forgotten the contravariant transform. Derive it.

Soln. Say you have forgotten the contravariant transform and are trying to figure it out what $V^{m'}$ should be given a known V^m:

$$V^{m'} = \frac{??}{??} V^m$$

The left hand side has a single upper index m', so the right hand side also needs to have a single upper index m'. The m in V^m right now is an upper index, but there isn't an upper m on the left, so there has to be an lower m to make this m into a dummy index, so that it's not a real index:

$$V^{m'} = \frac{??}{\partial x^m} V^m$$

Since the right hand side also needs to have a single upper index m', it has to go to the numerator:

$$V^{m'} = \frac{\partial x^{m'}}{\partial x^m} V^m$$

---

3. Show that the sum of two contravariant vectors is also contravariant. i.e. given A^m and B^m, both satisfying the contravariant transform, show that $V^m \equiv A^m + B^m$ also satisfies the contravariant transform

Soln. We have

$$A^{m'} = \frac{\partial x^{m'}}{\partial x^m} A^m$$

$$B^{m'} = \frac{\partial x^{m'}}{\partial x^m} B^m$$

Now,

$$\begin{aligned} V^{m'} &= A^{m'} + B^{m'} \\ &= \frac{\partial x^{m'}}{\partial x^m} A^m + \frac{\partial x^{m'}}{\partial x^m} B^m \\ &= \frac{\partial x^{m'}}{\partial x^m} (A^m + B^m) \\ &= \frac{\partial x^{m'}}{\partial x^m} V^m \end{aligned}$$

So V^m is also contravariant.

If the "factoring out" step makes you feel uncomfortable, write out the explicit summation to make sure:

$$\begin{aligned}
\frac{\partial x^{m'}}{\partial x^m}A^m + \frac{\partial x^{m'}}{\partial x^m}B^m &= \left(\frac{\partial x^{m'}}{\partial x^1}A^1 + \frac{\partial x^{m'}}{\partial x^2}A^2 + ...\right) + \left(\frac{\partial x^{m'}}{\partial x^1}B^1 + \frac{\partial x^{m'}}{\partial x^2}B^2\right) \\
&= \frac{\partial x^{m'}}{\partial x^1}(A^1 + B^1) + \frac{\partial x^{m'}}{\partial x^2}(A^2 + B^2) + ... \\
&= \frac{\partial x^{m'}}{\partial x^m}(A^m + B^m)
\end{aligned}$$

---

4. Show that the sum of two covariant vectors is also covariant. i.e. given A_m and B_m, both satisfying the covariant transform, show that $V_m \equiv A_m + B_m$ also satisfies the covariant transform.

Soln.

$$\begin{aligned}
V_{m'} &= A_{m'} + B_{m'} \\
&= \frac{\partial x^m}{\partial x^{m'}}A_m + \frac{\partial x^m}{\partial x^{m'}}B_m \\
&= \frac{\partial x^m}{\partial x^{m'}}(A_m + B_m) \\
&= \frac{\partial x^m}{\partial x^{m'}}V_m
\end{aligned}$$

---

5. Show that if $V_m = 0 \,\forall\, m$ in some frame, then given an arbitrary transform $x^{m'} = x^{m'}(x^1, x^2, ...)$, we have $V_{m'} = 0 \,\forall\, m'$, i.e. if a covariant vector is zero in one frame, then it's zero in all frames.

Soln. Looking at the covariant transform

$$V_{m'} = \frac{\partial x^m}{\partial x^{m'}}V_m = \frac{\partial x^1}{\partial x^{m'}}V_1 + \frac{\partial x^2}{\partial x^{m'}}V_2 + ...$$

we see that if the covariant vector V_m is zero, i.e. $V_1 = V_2 = ... = 0$, then we have $V_{m'} = (...)0 + (...)0 + ... = 0$ for all m'.

---

6. Show that if $V^m = 0 \,\forall\, m$ in some frame , then given an arbitrary transform $x^{m'} = x^{m'}(x^1, x^2, ...)$, we have $V^{m'} = 0 \,\forall\, m'$, i.e. if a contravariant vector is zero in one frame, then it's zero in all frames.

Soln. Looking at the contravariant transform

$$V^{m'} = \frac{\partial x^{m'}}{\partial x^m} V^m = \frac{\partial x^{m'}}{\partial x^1} V^1 + \frac{\partial x^{m'}}{\partial x^2} V^2 + ...$$

we see that if the contravariant vector V^m is zero, i.e. $V^1 = V^2 = ... = 0$, then we have $V^{m'} = (...)0 + (...)0 + ... = 0$ for all m'.

You might have noticed that the proof for many covariant and contravariant properties are identical to each other. This is because they are both rank-1 tensors. These properties are, in essence, just tensor properties. The proof for these tensor properties are, as you might have guessed, identical to the ones shown above, just with some more derivatives being multiplied with zero or being factored out.

The three properties we've looked at basically says (1) how to extract the transformation rule given the position of the indices of a tensor, (2) adding two tensors of the same type gives you a third tensor of the same type, and (3) if a tensor is zero (meaning that all the components of this tensor is zero) in one frame, then it's zero in all frames. We will revisit these properties when we talk about tensors in the next note.

7. Show that given a contravariant vector A^m and a covariant vector B_m, under an arbitrary transform $x^{m'} = x^{m'}(x^1, x^2, ...)$, we have $A^m B_m = A^{m'} B_{m'}$, i.e. inner product between a contravariant vector and a covariant vector is the same in all frames.

Hint: the partial derivatives of the base coordinates with each other is zero, and the partial derivative of a base coordinate with itself is one. i.e. in the old language, on a Cartesian plane, we have

$$\frac{\partial x}{\partial x} = \frac{\partial y}{\partial y} = 1, \frac{\partial x}{\partial y} = \frac{\partial y}{\partial x} = 0$$

Soln. Let's use different summation dummies for the two sums to avoid possible confusion.

$$\begin{aligned}
A^{m'} B_{m'} &= \left(\frac{\partial x^{m'}}{\partial x^m} A^m\right)\left(\frac{\partial x^p}{\partial x^{m'}} B_p\right) \\
&= \frac{\partial x^{m'}}{\partial x^m}\frac{\partial x^p}{\partial x^{m'}} A^m B_p \\
&= \frac{\partial x^p}{\partial x^m} A^m B_p \\
&= \delta^p_m A^m B_p \\
&= A^m B_m
\end{aligned}$$

Some explaining might be helpful:
-If the second equal sign makes you uncomfortable, again, convince yourself by writing out the whole sum:

$$\begin{aligned}
\left(\frac{\partial x^{m'}}{\partial x^m} A^m\right)\left(\frac{\partial x^p}{\partial x^{m'}} B_p\right) &= \left(\frac{\partial x^{m'}}{\partial x^1} A^1 + \frac{\partial x^{m'}}{\partial x^2} A^2 + ...\right)\left(\frac{\partial x^1}{\partial x^{m'}} B_1 + \frac{\partial x^2}{\partial x^{m'}} B_2 + ...\right) \\
&= \frac{\partial x^{m'}}{\partial x^1} A^1 \frac{\partial x^1}{\partial x^{m'}} B_1 + \frac{\partial x^{m'}}{\partial x^1} A^1 \frac{\partial x^2}{\partial x^{m'}} B_2 + ... + \frac{\partial x^{m'}}{\partial x^2} A^2 \frac{\partial x^1}{\partial x^{m'}} B_1 + \frac{\partial x^{m'}}{\partial x^2} A^2 \frac{\partial x^2}{\partial x^{m'}} B_2 + ...
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial x^{m'}}{\partial x^m}\frac{\partial x^p}{\partial x^{m'}} A^m B_p &= \frac{\partial x^{m'}}{\partial x^m}\frac{\partial x^1}{\partial x^{m'}} A^m B_1 + \frac{\partial x^{m'}}{\partial x^m}\frac{\partial x^2}{\partial x^{m'}} A^m B_2 + ... \\
&= \left(\frac{\partial x^{m'}}{\partial x^m} A^m\right)\frac{\partial x^1}{\partial x^{m'}} B_1 + \left(\frac{\partial x^{m'}}{\partial x^m} A^m\right)\frac{\partial x^2}{\partial x^{m'}} B_2 + ... \\
&= \left(\frac{\partial x^{m'}}{\partial x^1} A^1 + \frac{\partial x^{m'}}{\partial x^2} A^2 + ...\right)\frac{\partial x^1}{\partial x^{m'}} B_1 + \left(\frac{\partial x^{m'}}{\partial x^1} A^1 + \frac{\partial x^{m'}}{\partial x^2} A^2 + ...\right)\frac{\partial x^2}{\partial x^{m'}} B_2 + ... \\
&= \frac{\partial x^{m'}}{\partial x^1} A^1 \frac{\partial x^1}{\partial x^{m'}} B_1 + \frac{\partial x^{m'}}{\partial x^2} A^2 \frac{\partial x^1}{\partial x^{m'}} B_1 + \frac{\partial x^{m'}}{\partial x^1} A^1 \frac{\partial x^2}{\partial x^{m'}} B_2 + \frac{\partial x^{m'}}{\partial x^2} A^2 \frac{\partial x^2}{\partial x^{m'}} B_2 + ...
\end{aligned}$$

-The third equal sign is chain rule: $\dfrac{\partial x^{m'}}{\partial x^m}\dfrac{\partial x^p}{\partial x^{m'}} = \dfrac{\partial x^p}{\partial x^m}$

-The forth equal sign uses the hint: $\dfrac{\partial x^p}{\partial x^m}$ are the partial derivatives of the base coordinates with themselves. If $p = m$, then we have the derivative of a coordinate with itself, which is one; if $p \neq m$, then we have the derivative of a coordinate with another coordinate, which is zero. This reduces this term to the Kronecker delta [1]. When writing the Kronecker delta we keep and upstairs-ness of p and downstairs-ness of m so that the two implied summations in

$$\frac{\partial x^p}{\partial x^m} A^m B_p = \delta^p_m A^m B_p$$

is not broken. In the future we will see similar arrangements like δ_{mn} or δ^{mn}; they all refer to the Kronecker delta, i.e. 1 if the indices are the same, zero otherwise.

-The fifth equal sign is summation with Kronecker delta: in $\delta^p_m A^m B_p$, when p is summed over, all terms with $p \neq m$ are zero due to the δ^p_m, so the only term that survives after the p summation is the term with $p = m$, i.e. the term $\delta^m_m A^m B_m$. Since $\delta^m_m = 1$, so this is just $A^m B_m$.

This is an important result. It can be phrased in slightly different ways:
- The inner product of a covariant vector and a contravariant vector is the same in all frames;
- The inner product of a covariant vector and a contravariant vector is frame-invariant;
- The inner product of a covariant vector and a contravariant vector is a scalar [2].

These statements all refer to the same idea.

[1] The Kronecker delta is $\delta_{mn} = 1$ if $m = n$ and 0 if $m \neq n$

[2] Recall that scalars are quantities that do not transform when we change coordinates

8. We know the transformation between polar coordinates and Cartesian coordinates is

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

or, equivalently,

$$\begin{cases} r = \sqrt{x^2+y^2} \\ \\ \theta = \tan^{-1}(\frac{y}{x}) \end{cases}$$

(a) Consider a contravariant vector in polar coordinates $A^m = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. This means its r-component is 0 and the θ-component is 1. Find its components $A^{m'}$ in Cartesian coordinates. Does your result make sense?

(b) Consider a covariant vector in polar coordinates $B_m = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. This means its r-component is 0 and the θ-component is 1. Find its components $B_{m'}$ in Cartesian coordinates. Does your result make sense?

(c) Explicitly verify that the inner product is frame-invariant, i.e. $A^m B_m = A^{m'} B_{m'}$. Physically, what does this result mean?

Soln.
The necessary derivatives are [3]

$$\frac{\partial x}{\partial r} = \cos\theta\,, \ \frac{\partial x}{\partial \theta} = -r\sin\theta\,, \ \frac{\partial y}{\partial r} = \sin\theta\,, \ \frac{\partial y}{\partial \theta} = r\cos\theta$$

$$\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2+y^2}}\,, \ \frac{\partial r}{\partial y} = \frac{y}{\sqrt{x^2+y^2}}$$

$$\frac{\partial \theta}{\partial x} = \frac{-y/x^2}{1+y^2/x^2} = -\frac{y}{x^2+y^2}\,, \ \frac{\partial \theta}{\partial y} = \frac{1/x}{1+y^2/x^2} = \frac{x}{x^2+y^2}$$

[3] Reminder: derivative of $\tan^{-1}(x)$ is $\frac{1}{1+x^2}$.

(a) We have $A^m = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in polar coordinates, which means we are using unprimed for polar and thus primed for Cartesian. Therefore, using $A^{m'} = \dfrac{\partial x^{m'}}{\partial x^m} A^m$, we have

$$\begin{aligned}
A^x &= \frac{\partial x}{\partial x^m} A^m \\
&= \frac{\partial x}{\partial r} A^r + \frac{\partial x}{\partial \theta} A^\theta \\
&= \frac{\partial x}{\partial \theta} \quad (\text{since } A^r = 0 \text{ and } A^\theta = 1) \\
&= -r \sin\theta \\
&= -y
\end{aligned}$$

and

$$\begin{aligned}
A^y &= \frac{\partial y}{\partial x^m} A^m \\
&= \frac{\partial y}{\partial r} A^r + \frac{\partial x}{\partial \theta} A^\theta \\
&= \frac{\partial y}{\partial \theta} \quad (\text{since } A^r = 0 \text{ and } A^\theta = 1) \\
&= r \cos\theta \\
&= x
\end{aligned}$$

Therefore, $A^{m'} = \begin{pmatrix} A^x \\ A^y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}$

Physical meaning: the vector $A^m = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in polar coordinates is contravariant, so we should think of it as a displacement vector. Thus it means a displacement between two points, with the two points having the same r coordinate and their θ coordinate differ by 1.

Thus, the displacement vector doesn't move in the r direction and takes one unit step in the θ direction [4]. In addition, notice that as we go further away from the origin (i.e. as r gets bigger), a unit θ step would correspond to a larger distance. It looks like the following:

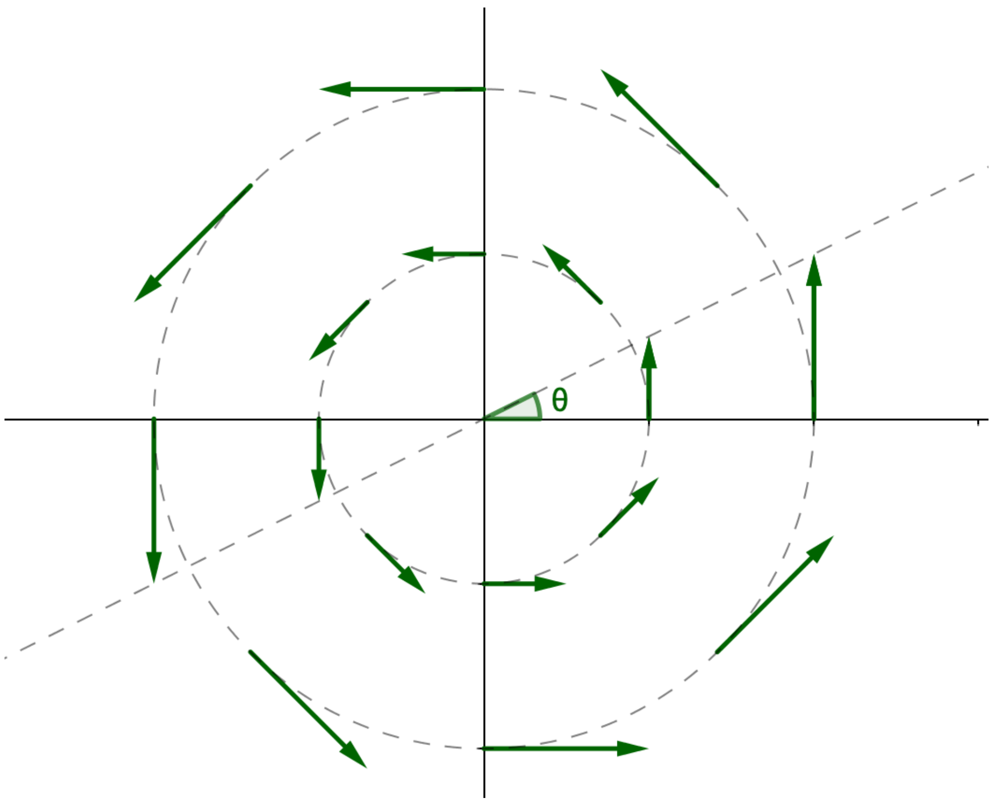

Figure 1: The displacement vector $A^m = (0, 1)$ in polar coordinates, not drawn to scale. Notice that even if the difference in θ is the same, for larger r the actual vector is longer.

From this picture it is clear that in Cartesian coordinates, this displacement would be $(-y, x)$. Notice that the magnitude (squared) of the Cartesian vector $(-y, x)$, which is $x^2 + y^2$, increases as we go further away from the origin.

[4] For simplicity of illustration we take the vector to lie along the tangent line of the circle, so it vector itself is not "curved". For infinitesimal vectors, it doesn't matter whether the vector is strictly lying on the circle as a curved vector, or a straight one lying tangent to the circle.

(b) We have $B_m = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in polar coordinates, which means we are using unprimed for polar and thus primed for Cartesian. Therefore, using $B_{m'} = \dfrac{\partial x^m}{\partial x^{m'}} B_m$, we have

$$\begin{aligned}
B_x &= \frac{\partial x^m}{\partial x} B_m \\
&= \frac{\partial r}{\partial x} B_r + \frac{\partial \theta}{\partial x} B_\theta \\
&= \frac{\partial \theta}{\partial x} \quad (\text{since } B_r = 0 \text{ and } B_\theta = 1) \\
&= -\frac{y}{x^2 + y^2}
\end{aligned}$$

and

$$\begin{aligned}
B_y &= \frac{\partial x^m}{\partial y} B_m \\
&= \frac{\partial r}{\partial y} B_r + \frac{\partial \theta}{\partial y} B_\theta \\
&= \frac{\partial \theta}{\partial y} \quad (\text{since } B_r = 0 \text{ and } B_\theta = 1) \\
&= \frac{x}{x^2 + y^2}
\end{aligned}$$

Therefore, $B_{m'} = \begin{pmatrix} B_x \\ B_y \end{pmatrix} = \begin{pmatrix} -\dfrac{y}{x^2 + y^2} \\ \\ \dfrac{x}{x^2 + y^2} \end{pmatrix}$

Physical meaning: the vector $B_m = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in polar coordinates is covariant, so we should think of it as a gradient vector, i.e. a force. Thus it means a force that doesn't do any work in the r direction, and does 1 unit of work per 1 unit displacement in the θ direction [5].

We know that as r increases, 1 unit displacement in the θ direction gets longer. Thus, as r increases, that 1 unit of work done over the longer unit θ distance means a smaller actual force.

The Cartesian vector for the force is $(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2})$, which does get smaller as we go further away from the origin.

(c) We can explicitly verify the invariance of inner product.

$$\begin{aligned}
A^{m'}B_{m'} &= A^x B_x + A^y B_y \\
&= (-y)(-\frac{y}{x^2+y^2}) + (x)(\frac{x}{x^2+y^2}) \\
&= \frac{y^2}{x^2+y^2} + \frac{x^2}{x^2+y^2} \\
&= 1
\end{aligned}$$

and

$$\begin{aligned}
A^m B_m &= A^r B_r + A^\theta B_\theta \\
&= (0)(1) + (0)(1) \\
&= 1
\end{aligned}$$

So inner product is indeed invariant.

The inner product between a displacement (contravariant, i.e. the A^m) and a force (covariant, i.e. the B_m) is the work done when the force moves something over that displacement. Since it's the same force and the same displacement, the work done is of course the same [6]. Again, the vectors in polar coordinates and Cartesian coordinates are the *same* vectors; they *look* different (as in, they have different explicit numerical components) only because we are describing them with different coordinate systems.

[5] Recall that covariant vectors have "something per distance" as their units. For forces, that's "joules per meter". Here instead of 1 meter, since it's 1 in the θ component, the unit distance is the distance represented by a displacement of 1 in the θ direction.

[6] This is **not** the conservation of energy; it's simply looking at the same thing from different frames.