Tutorial 2

FK8027 - Quantum Field Theory

Wednesday 13^{th} November, 2019

Topics for today

- Vectors and covectors
- The metric tensor and the musical isomorphism

Partially based on these notes from the course homepage, whose reading is recommended.

1 Vectors and covectors

Special relativity is a theory describing physical phenomena by introducing a 4-dimensional (flat) spacetime, having 3 spatial and 1 temporal dimension.[1]

We can consider coordinates in this spacetime and we can denote them with $\left\{x^0 = c \cdot t,\, x^1,\, x^2,\, x^3\right\}$, where c is the speed of light. In natural units, $c = 1$ and $x^0 = t$. We can refer to these coordinates more easily by defining the following object,

$$x^\mu = \left(x^0, x^1, x^2, x^3\right), \quad \mu \in \{0, 1, 2, 3\}. \tag{1}$$

This is just a notation to write the coordinates of a point in Minkowski spacetime in a compact form.

Suppose to consider a general coordinate transformation (GCT) between two coordinate systems x^μ and $\tilde{x}^\mu$. The transformation laws between the two sets of coordinates are,

$$\tilde{x}^\mu = \tilde{x}^\mu\left(x^\mu\right) = \tilde{x}^\mu\left(x^0,\, x^1,\, x^2,\, x^3\right). \tag{2}$$

$$\begin{aligned}
[\text{e.g.,} \quad x &= r\,\sin(\theta)\cos(\phi)\,, \\
y &= r\,\sin(\theta)\sin(\phi)\,, \\
z &= r\,\cos(\theta)\,.]
\end{aligned}$$

By assumption, the functional relationship can be inverted, i.e.,

$$x^\mu = x^\mu\left(\tilde{x}^\mu\right) = x^\mu\left(\tilde{x}^0,\, \tilde{x}^1,\, \tilde{x}^2,\, \tilde{x}^3\right). \tag{3}$$

Using our knowledge from the multivariate calculus, we can write how the differential $\mathrm{d}x^\mu$ and the derivative $\dfrac{\partial}{\partial x^\mu}$ transform under the GCT,

$$\mathrm{d}\tilde{x}^\mu = \sum_{\nu=0}^{3} \left(\frac{\partial \tilde{x}^\mu}{\partial x^\nu}\right) \mathrm{d}x^\nu \tag{4a}$$

[1] The term "flat" signifies that we can cover the entire spacetime with a *unique* coordinate system, thus mapping the entire spacetime to $\mathbb{R}^4$ [1, p.17].

$$\frac{\partial}{\partial \tilde{x}^{\mu}} = \sum_{\nu=0}^{3} \left(\frac{\partial x^{\mu}}{\partial \tilde{x}^{\nu}}\right) \frac{\partial}{\partial x^{\mu}}. \tag{4b}$$

We see that these two quantities transform differently under a GCT. These two types of transformation are the basic types, and we can define two different categories of geometric objects based on them.

Definition 1. A geometrical object whose components transform as those of $\mathrm{d}x^{\mu}$ is called a *vector* (or, old-fashionedly, a "contravariant vector").

Definition 2. A geometrical object whose components transform as those of $\partial/\partial x^{\mu}$ is called a *covector* (or, old-fashionedly, a "covariant vector").

According to these definitions, we can write down how the components of a vector v and a covector ω transform,

$$\tilde{v}^{\mu}(\tilde{x}) = \sum_{\nu=0}^{3} \left(\frac{\partial \tilde{x}^{\mu}}{\partial x^{\nu}}\right) v^{\nu}(x) \tag{5a}$$

$$\tilde{\omega}_{\mu}(\tilde{x}) = \sum_{\nu=0}^{3} \left(\frac{\partial x^{\mu}}{\partial \tilde{x}^{\nu}}\right) \omega_{\nu}(x) \tag{5b}$$

The position of the index, up or down (respectively, a "contravariant index" and a "covariant index"), tells us about the transformation properties of the object. Therefore, vectors have components denoted with upper (contravariant) indices, and covectors have components denoted with lower (covariant) indices. However, a very important remark is that not every object with an upper (lower) index is a vector (covector); e.g., $x^{\mu} = \left(x^0, x^1, x^2, x^3\right)$ does not transform as a vector under a GCT, because the change of coordinates can be arbitrary (nonlinear) [see (2)], hence x^{μ} is not a vector under GCT.

This index notation allows us to handle more complex geometric objects in a vary handy way.

Definition 3. A *tensor T of rank $m+n$*, with components having m contravariant indices and n covariant indices and denoted by $T^{\mu_1 \cdots \mu_n}{}_{\nu_1 \ldots \nu_n}$, is a geometrical object whose components transform as follows under a GCT,

$$\tilde{T}^{\tilde{\mu}_1 \ldots \tilde{\mu}_n}{}_{\tilde{\nu}_1 \ldots \tilde{\nu}_n} = \left(\frac{\partial \tilde{x}^{\tilde{\mu}_1}}{\partial x^{\mu_1}} \cdots \frac{\partial \tilde{x}^{\tilde{\mu}_m}}{\partial x^{\mu_m}}\right) T^{\mu_1 \ldots \mu_n}{}_{\nu_1 \ldots \nu_n} \left(\frac{\partial x^{\nu_1}}{\partial \tilde{x}^{\tilde{\nu}_1}} \cdots \frac{\partial x^{\nu_n}}{\partial \tilde{x}^{\tilde{\nu}_n}}\right), \tag{6}$$

where we are using the *Einstein's summation convention*, i.e. *a summation is implied whenever an index is repeated in upper and lower position in an expression.*

Note that upper indices appearing in denominators count as lower indices, and viceversa (see Exercise 1)—you can actually see this explicitly in (6). We remind that each index run from 0 to 3, so we have, e.g.,

$$\frac{\partial \tilde{x}^{\mu}}{\partial x^{\mu}} = \frac{\partial \tilde{x}^{0}}{\partial x^{0}} + \frac{\partial \tilde{x}^{1}}{\partial x^{1}} + \frac{\partial \tilde{x}^{2}}{\partial x^{2}} + \frac{\partial \tilde{x}^{3}}{\partial x^{3}}. \tag{7}$$

Since $\partial \tilde{x}^\mu / \partial x^\nu$ and $\partial x^\mu / \partial \tilde{x}^\nu$ are functions, respectively, of x^μ and $\tilde{x}^\mu$, they are numbers at each point of the spacetime. Hence, they can be regarded as element of matrices (that depend on the point of the spacetime). We can define,

$$M^\mu{}_\nu\left(x^\rho\right) := \frac{\partial \tilde{x}^\mu}{\partial x^\nu}\left(x^\rho\right), \quad N^\mu{}_\nu\left(\tilde{x}^\rho\right) := \frac{\partial x^\mu}{\partial \tilde{x}^\nu}\left(\tilde{x}^\rho\right), \tag{8}$$

where the index labeling the rows of the matrices is the first one (on the left), the one labeling the columns of the matrices is the second one (on the right). The height of the index only defines the transformation properties.

Then, in matrix notation, the transformation laws of vectors and covectors can be written

$$\tilde{v}^\mu = M^\mu{}_\nu v^\nu, \qquad \text{or} \quad \tilde{v} = M \cdot v, \tag{9a}$$

$$\tilde{\omega}_\mu = N^\nu{}_\mu \omega_\nu = (N^\intercal)_\mu{}^\nu \omega_\nu, \quad \text{or} \quad \tilde{\omega} = \omega \cdot N. \tag{9b}$$

Using the chain rule, we can write

$$\sum_{\nu=0}^{3} \frac{\partial \tilde{x}^\mu}{\partial x^\nu} \frac{\partial x^\nu}{\partial \tilde{x}^\sigma} = \frac{\partial \tilde{x}^\mu}{\partial \tilde{x}^\sigma} = \delta^\mu{}_\sigma \tag{10}$$

with $\delta^\mu{}_\sigma$ Kronecker delta. It follows that

$$\delta^\mu{}_\sigma = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} \frac{\partial x^\nu}{\partial \tilde{x}^\sigma} = M^\mu{}_\nu N^\nu{}_\sigma \Longrightarrow N^\nu{}_\sigma = \left(M^{-1}\right)^\nu{}_\sigma, \text{ or } N = M^{-1}, \tag{11}$$

as one might expect.

We can define an operation between a vector and a covector, called "contraction", as

$$\text{contraction}\left(v, \omega\right) := v^\mu \omega_\mu = v^0 \omega_0 + v^1 \omega_1 + v^2 \omega_2 + v^3 \omega_3. \tag{12}$$

The result of the contraction does not change under a GCT,

$$\tilde{v}^\mu \tilde{\omega}_\mu = \frac{\partial \tilde{x}^\mu}{\partial x^\nu} v^\nu \frac{\partial x^\lambda}{\partial \tilde{x}^\mu} \omega_\lambda = v^\nu \frac{\partial \tilde{x}^\mu}{\partial x^\nu} \frac{\partial x^\lambda}{\partial \tilde{x}^\mu} \omega_\lambda = v^\nu \delta_\nu{}^\lambda \omega_\lambda = v^\nu \omega_\nu. \tag{13}$$

The contracted indices are called "dummy" indices. Not contracted indices are called "free" indices. Only free indices determine the transformation properties of a geometric object, hence only free indices establish the rank of a tensor. The contraction can be defined between any rank-m tensor and rank-n tensor, e.g.,

$$\begin{aligned} T^{\mu\nu\rho\sigma} S_{\alpha\mu\beta\sigma} = {} & T^{0\nu\rho 0} S_{\alpha 0 \beta 0} + T^{1\nu\rho 0} S_{\alpha 1 \beta 0} + T^{2\nu\rho 0} S_{\alpha 2 \beta 0} + T^{3\nu\rho 0} S_{\alpha 3 \beta 0} \\ & + T^{0\nu\rho 1} S_{\alpha 0 \beta 1} + T^{1\nu\rho 1} S_{\alpha 1 \beta 1} + T^{2\nu\rho 1} S_{\alpha 2 \beta 1} + T^{3\nu\rho 1} S_{\alpha 3 \beta 1} \\ & + T^{0\nu\rho 2} S_{\alpha 0 \beta 2} + T^{1\nu\rho 2} S_{\alpha 1 \beta 2} + T^{2\nu\rho 2} S_{\alpha 2 \beta 2} + T^{3\nu\rho 2} S_{\alpha 3 \beta 2} \end{aligned}$$

$$+T^{0\nu\rho 3}S_{\alpha 0\beta 3}+T^{1\nu\rho 3}S_{\alpha 1\beta 3}+T^{2\nu\rho 3}S_{\alpha 2\beta 3}+T^{3\nu\rho 3}S_{\alpha 3\beta 3} \quad (14)$$

Notice that the dummy indices can be renamed at any time, but they cannot be renamed with a letter which already indicates another contraction, or a free index. For example, $T^{\mu\nu\rho\sigma}S_{\alpha\mu\beta\sigma}$ is a well defined expression. We can write it also as $T^{\gamma\nu\rho\delta}S_{\alpha\gamma\beta\delta}$ and it is still well-defined. However, $T^{\mu\nu\rho\mu}S_{\alpha\mu\beta\mu}$ is not well-defined, because we do not know, e.g., if the first μ index of T is contracted with the first or second index μ of S. Analogously, the expression $T^{\rho\nu\rho\sigma}S_{\alpha\rho\beta\sigma}$ is not well-defined.

We defined the contraction as an operation involving the components of a vector and a covector. The abstract meaning of this operation is that a covector is a function which takes a vector as its argument and returns a real number, equal to the output of the contraction (12),

$$\omega_\mu \text{ takes } v^\mu \text{ and gives } \omega_\mu v^\mu \iff \omega(v)=\omega_\mu v^\mu \quad (15)$$

This definition of a covector is independent of the coordinates and is the one used in linear algebra and differential geometry. More in detail, suppose we have a vector space V; we define its dual space as the space of all linear functions $\omega : V \to \mathbb{R}$. The elements of this space are the covectors. Analogously, vectors are linear functions from the dual space V^* of covectors to $\mathbb{R}$, so they take a covector as argument and return a real number (again equal to the result of (12)).

$$v^\mu \text{ takes } \omega_\mu \text{ and gives } v^\mu\omega_\mu \iff v(\omega)=v^\mu\omega_\mu \quad (16)$$

This abstract definition can be extended to tensors of any rank. For example, the tensor P with components $P^{\mu\nu}{}_{\rho\sigma\alpha}$ is a linear function which takes as arguments 2 covectors—because it has 2 contravariant indices that can be contracted with 2 covariant indices—and 3 vectors, and returns the real number given by the contraction $\omega_\mu\gamma_\nu\, P^{\mu\nu}{}_{\rho\sigma\alpha}\, v^\rho u^\sigma w^\alpha$. The index notation is powerful because it allows us to perform all these abstract operations in an easy way by contracting indices, using only the components of the tensors.

When dealing with the components of tensors having a rank ≥ 2, we can define some maps, other than the tensors themselves, but having the same components. For example, suppose that we contract $P^{\mu\nu}{}_{\rho\sigma\alpha}$ with 2 covectors ω_μ, σ_μ and 2 vectors u^μ, v^μ,

$$P^{\mu\nu}{}_{\rho\sigma\alpha}\,\omega_\mu\sigma_\nu u^\rho v^\sigma. \quad (17)$$

The result is an object with a covariant free index α, i.e., a covector. Then, we can think of another map $\tilde{P}$, with the *same* components of the tensor P, which takes as arguments 2 covectors and 2 vectors, and returns a covector. This concept is important to understand what the *metric tensor* does.

2 The metric tensor and the musical isomorphism

From the geometry and linear algebra course, we know that in a vector space V we can define a "scalar product" $\cdot$ (or "inner product") between two vectors u and v as $u \cdot v = x$, with $x \in \mathbb{R}$. Let's consider the components $g_{\mu\nu}$ of a nondegenerate tensor such that their contraction with the components of any two vectors u^μ and v^μ gives their scalar product.[2] In formulas,

$$g_{\mu\nu} u^\mu v^\nu := u \cdot v = x \in \mathbb{R}, \quad \forall u, v \in V. \tag{18}$$

The tensor g whose components satisfy (18) is called the *metric tensor*. Usually, one first defines a certain metric tensor, which implies the existence of a scalar product. Notice that, since the scalar product is symmetric, i.e., $u \cdot v = v \cdot u$, the metric tensor itself is symmetric, and this implies that its components satisfy $g_{\mu\nu} = g_{\nu\mu}$,

$$u \cdot v = v \cdot u \Longrightarrow g_{\mu\nu} u^\mu v^\nu = g_{\mu\nu} v^\mu u^\nu \overset{\substack{\text{rename}\\\text{dummy}\\\text{indices}}}{=} g_{\nu\mu} u^\mu v^\nu \Longrightarrow g_{\mu\nu} = g_{\nu\mu}. \tag{19}$$

As we pointed out in the previous section, we can define a map $\tilde{g}$ with the same components of the metric tensor $g_{\mu\nu}$, which sends a vector to a covector. It is possible to prove [2, p. 282] that this map is bijective and that it is actually an isomorphism between the vector space of vectors and the dual vector space of covectors.[3] Therefore, *if we have a metric*, vectors and covectors can be regarded as equivalent.

Consider the vector v with components v^μ. We identify the covector having components $g_{\mu\nu} v^\nu$ with v; then, we denote the components of the covector with

$$v_\mu := g_{\mu\nu} v^\nu. \tag{20}$$

We use the same symbol v for the vector and the covector, by virtue of the isomorphism induced by the metric g, which tells us that the two objects are equivalent.

It is very common, in physics (and in old-fashioned mathematics), to talk about this topic in terms of "rasing and lowering" indices. Look at (20): the map $\tilde{g}$ is just "lowering" the index of v, so it *appears* that we are considering different "versions" of the same geometric object. However, you should remember that we are really referring to different objects when we write v^μ or v_μ; it is not the same object with different indices. In the first case we refer to the components of a vector, in the second case we refer to the components of its dual covector, *which we call with the same name.*

[2]Nondegenerate means that the matrix of components $g_{\mu\nu}$ has a nonzero determinant, i.e., it is invertible.

[3]Two isomorphic structures, roughly speaking, can be considered as equivalent or, if you prefer, "indistinguishable".

We assumed that the metric g is invertible. We denote the inverse with g^{-1} and its components with $g^{\mu\nu}$. It holds, by definition,

$$g_{\mu\rho}g^{\rho\nu} = g_{\mu}{}^{\nu} = \delta_{\mu}{}^{\nu}. \tag{21}$$

In the same way $\tilde{g}$ lowers indices, $\tilde{g}^{-1}$ raises them,

$$v^{\mu} := g^{\mu\nu}v_{\nu}, \tag{22}$$

compatibly with (21).

The raising and lowering of the indices of a tensor resembles the raising and lowering of a musical tone (respectively, sharp and flat tone). Because of this, the isomorphism $\tilde{g}$ between vectors and covectors, induced by a metric g, is called the "musical isomorphism". If you want to deepen these topics, a good reference is [2, Chapter 6, p. 124; Chapter 11, p. 282].

In classical physics, the metric tensor is the euclidean metric, i.e., the identity,

$$\delta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{23}$$

therefore there is no difference at all between the component of a vector and those of its dual covector (see next section). In special relativity and in QFT, the metric tensor is the Minkowski one, denoted by $\eta_{\mu\nu}$

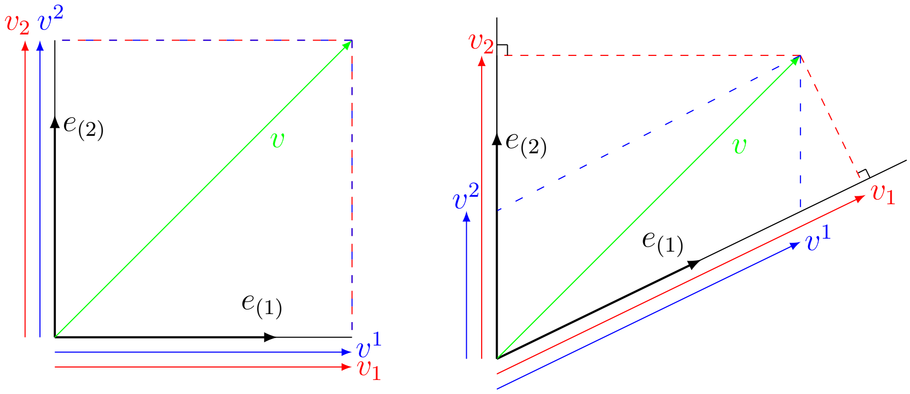
$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{24}$$

Hence, the time components of a vector and its dual are the same, but the spatial ones differ in sign.

Let's now define the trace operation in index notation. We know that, if we have a matrix A, its trace is the sum of its diagonal components. In index notation, a sum in indicated by a repeated index in upper and lower position (Einstein's summation convention), therefore the trace of a rank-2 tensor A is given by

$$\mathrm{tr}\,(A) := A^{\mu}{}_{\mu} = A^{0}{}_{0} + A^{1}{}_{1} + A^{2}{}_{2} + A^{3}{}_{3}, \tag{25}$$

which is a contraction of its own indices. Note that, if you have a tensor with both covariant (contravariant) indices and you want to take its trace, you must always raise (lower) one of its indices before contracting them. As



(a) Orthogonal frame in 2-dimensional Euclidean space.

(b) Non-orthogonal frame in 2-dimensional Euclidean space (the same happens with an orthogonal frame in a curved spacetime).

Figure 1: Components of the vector v (blue) and of its dual covector (red) with respect to the frame $\{e_{(1)}, e_{(2)}\}$.

an example, we compute the trace of the metric tensor, which is not $g_{\mu\mu}$ but rather,

$$\mathrm{tr}(g) := g^{\mu\rho} g_{\rho\mu} = g^{\mu}{}_{\mu} = \delta^{\mu}{}_{\mu} = 4. \tag{26}$$

We have first raised the first index of the metric with the inverse metric and obtained the identity (whose components are given by the Kronecker delta), and only afterwards we have contracted the upper and lower indices of the delta. The trace of any metric is always equal to the dimension of the spacetime, i.e., 4 in our case.

3 A graphical way to visualize the components of a vector and those of its dual

Let's try to have a geometric intuition of what we have talked about. Suppose we have a vector v. Its components with respect to a basis $\{e_{(1)}, e_{(2)}\}$ are obtained from the vectorial sum law (or parallelogram law), i.e.,

$$v = v^1 e_{(1)} + v^2 e_{(2)}. \tag{27}$$

The components of its dual covector are $v_\mu = g_{\mu\nu} v^\nu$, which graphically means that they are determined not by using the parallelogram law, but insted by orthogonal projection to the basis vectors. In Euclidean spacetime with an orthogonal frame, the two methods lead to the same result, as we can see in Figure 1a. However, in the other cases the two sets of components differ; in Figure 1b we see what happens if the reference frame is not orthogonal with respect to the Euclidean metric.

References

[1] M. Tsamparlis, *Special relativity: an introduction with 200 problems and solutions* (Springer Berlin Heidelberg, 2010) (cit. on p. 1).

[2] J. Lee, *Introduction to smooth manifolds*, Graduate Texts in Mathematics (Springer, 2003) (cit. on pp. 5–6).

Exercises

Exercise 1. Prove that $\partial_\mu \equiv \partial/\partial x^\mu$ transforms as a covector.

Solution. The chain rule of derivatives tells us that

$$\partial'_\mu = \frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu}\frac{\partial}{\partial x^\nu} = \frac{\partial x^\nu}{\partial x'^\mu}\partial_\nu. \tag{28}$$

This justifies the statement that contravariant indices at the denominator count as covariant ones. The proof that $\partial/\partial x^\mu$ transforms as a vector is left to the reader.

Exercise 2. Prove that $\mathrm{d}x^\mu$ transforms as a vector.

Solution. The differential of a function $f : \mathbb{R}^4 \to \mathbb{R}$ is defined as,

$$\mathrm{d}f \coloneqq \frac{\partial f}{\partial x^\mu}\mathrm{d}x^\mu. \tag{29}$$

Hence, the differential of the new coordinate functions $\tilde{x}^\mu(x^\nu)$ is given by

$$\mathrm{d}\tilde{x}^\mu \coloneqq \frac{\partial \tilde{x}^\mu}{\partial x^\nu}\mathrm{d}x^\nu, \tag{30}$$

which is the transformation law of $\mathrm{d}x^\mu$ under a GCT.

Exercise 3. Show that the trace of a mixed tensor $T^\mu{}_\nu$ is a scalar, i.e., it holds $\mathrm{tr}\,(T') = \mathrm{tr}\,(T)$ (the trace does not change under a change of coordinates).

Solution.

$$T'^\mu{}_\mu = \frac{\partial x^\mu}{\partial x'^\nu}T^\nu{}_\rho\frac{\partial x'^\mu}{\partial x^\nu} = T^\nu{}_\rho\frac{\partial x'^\rho}{\partial x'^\nu} = T^\nu{}_\rho\delta^\rho{}_\nu = T^\nu{}_\nu.$$

Exercise 4. If A^μ and B^ν are vectors, prove that $T^{\mu\nu} \equiv A^\mu B^\nu$ is a tensor.

Solution.

$$T'^{\mu\nu} = A'^\mu B'^\nu = \left(\frac{\partial x'^\mu}{\partial x^\rho}A^\rho\right)\left(\frac{\partial x'^\nu}{\partial x^\sigma}B^\sigma\right) = \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma}A^\rho B^\sigma = \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma}T^{\rho\sigma}.$$

Exercise 5. If B^μ is a vector and $A_{\mu\nu}$ is a tensor, prove that $X_\mu = B^\nu A_{\mu\nu}$ is a covector.

Solution.

$$\begin{aligned} X'_\mu = B'^\mu A'_{\mu\nu} &= \frac{\partial x'^\nu}{\partial x^\sigma}B^\sigma\frac{\partial x^\rho}{\partial x'^\mu}A_{\rho\gamma}\frac{\partial x^\gamma}{\partial x'^\nu} \\ &= \frac{\partial x^\rho}{\partial x'^\mu}\delta_\sigma{}^\gamma\left(B^\sigma A_{\rho\gamma}\right) = \frac{\partial x^\rho}{\partial x'^\mu}\left(B^\sigma A_{\rho\sigma}\right) = \frac{\partial x^\rho}{\partial x'^\mu}X_\rho. \end{aligned}$$