Office hours 1

Joe Keir

Joseph.Keir@maths.ox.ac.uk

The relationship between component notation and vectors/matrices

What is the relationship between notation like m_{ab}, and the matrix $\mathrm{diag}(-1, 1, 1, 1)$? I haven't been very explicit with this – it's up to you the conventions you use – but let's say that the first index corresponds to the row and the second to the column.

Then for a matrix M times a vector X (on the right), we could write something like

$$M_{ab}X^b.$$

If we wanted to multiply by another vector Y *on the left*, then we would write

$$\boldsymbol{Y}^T M \boldsymbol{X} = Y^a M_{ab} X^b = M_{ab} Y^a X^b.$$

What are these objects X^a? Choose a point $p \in \mathbb{M}^4$, then the tangent space at p is the set of vectors "starting at" p, i.e. objects like $(q - p)$ with $q \in \mathbb{M}^4$. This is a 4-dimensional vector space.

Choosing "inertial coordinates centred at p" means choosing a special kind of basis for this vector space, e_0, e_1, e_2, e_3. Then the components of the vector $(q - p)$ are given by

$$q - p = (q - p)^a e_a.$$

The "coordinates of the point q" in these inertial coordinates are just $(q - p)^a$ (so that the coordinates of p are $(0, 0, 0, 0)$).

What makes the choice of basis "special"?

The basis is "special" if the components of the Minkowski metric in this basis are $\mathrm{diag}(-1, 1, 1, 1)$. What do we mean by the components of the metric?

The metric is a bilinear form, i.e. something which takes in two vectors (say at the point p) X and Y, and spits out a real number, $m(X, Y)$. It is linear in both arguments, so

$$\begin{aligned} m(a_1 X_1 + a_2 X_2, b_1 Y_1 + b_2 Y_2) = a_1 b_1 m(X_1, Y_1) + a_1 b_2 m(X_1, Y_2) \\ + a_2 b_1 m(X_2, Y_1) + a_2 b_2 m(X_2, Y_2). \end{aligned}$$

The *components* of the bilinear form, m_{ab}, are just the numbers we get by acting on the basis vectors:

$$m_{ab} := m(e_a, e_b).$$

So inertial coordinates $\leftrightarrow$ choosing a (pseudo-)orthonormal basis with respect to the inner product $\langle X, Y \rangle = m(X, Y)$.

Now it's easy to see that

$$
\begin{aligned}
m(X, Y) &= m(X^a e_a, Y^b e_b) && \text{(expanding the vectors in a basis)} \\
&= X^a Y^b m(e_a, e_b) && \text{(by linearity)} \\
&= X^a Y^b m_{ab} && \text{(definition of } m_{ab}\text{).}
\end{aligned}
$$

A Lorentz transformation (centred at p) is a change of basis vectors, from e_a to $e'_{a'}$, "preserving the form of the metric". Since the basis vectors $e'_{a'}$ form a basis, we can write

$$e_a = \Lambda_a{}^{a'} e'_{a'},$$

and then the vector X can be expanded in terms of this new basis:

$$X = X^a e_a = X^a \Lambda_a{}^{a'} e'_{a'} = (X^a \Lambda_a{}^{a'}) e'_{a'} = (X'^{a'}) e'_{a'},$$

so the new components of X are

$$X'^{a'} = \Lambda_a{}^{a'} X^a.$$

What about the new components of the metric, $m'_{a'b'}$? These are

$$
\begin{aligned}
m'_{a'b'} &= m(e'_{a'}, e'_{b'}) \\
&= m\left((\Lambda^{-1})_{a'}{}^{a} e_a, (\Lambda^{-1})_{b'}{}^{b} e_b\right) \\
&= (\Lambda^{-1})_{a'}{}^{a} (\Lambda^{-1})_{b'}{}^{b} m(e_a, e_b) \\
&= (\Lambda^{-1})_{a'}{}^{a} (\Lambda^{-1})_{b'}{}^{b} m_{ab}.
\end{aligned}
$$

This means multiplying on both the left and the right by Λ^{-1}: in matrix notation,

$$m' = (\Lambda^{-1})^T m \Lambda^{-1}.$$

An alternative way to derive this formula: the quantity $m(X, Y)$ doesn't "know" anything about which basis we are working in. Hence we must have

$$m(X, Y) = m_{ab} X^a Y^b$$

$$= m'_{a'b'} (X')^{a'} (Y')^{b'}$$

$$= (m'_{a'b'} \Lambda_a{}^{a'} \Lambda_b{}^{b'}) X^a Y^b,$$

and so

$$m_{ab} = (m'_{a'b'} \Lambda_a{}^{a'} \Lambda_b{}^{b'})$$

$$\Rightarrow m'_{a'b'} = m_{ab} (\Lambda^{-1})_{a'}{}^a (\Lambda^{-1})_{b'}{}^b.$$

What is ∂^a?

Remember, the covector $\mathrm{d}\phi$ has components (in an inertial coordinate system) $\partial_a\phi$, which are just the partial derivatives of ϕ with respect to the inertial coordinates. We showed that this definition doesn't depend on the choice of inertial coordinate system.

The same thing works for higher rank tensors – e.g. given a $(0, 2)$ tensor T, the tensor ∂T is the tensor with components $\partial_a T_{bc}$, which are the partial derivatives of the components of T with respect to some inertial coordinate system. This also doesn't depend on which inertial coordinates we choose.

What about the notation $\partial^a \phi$? Remember, we lower and raise indices using the metric and its inverse. So $\partial^a \phi$ are the components of the vector $(\mathrm{d}\phi)^\sharp$: in index notation

$$\partial^a \phi = (m^{-1})^{ab} \partial_b \phi.$$

In fact, the components of $\partial^a \phi$ are the same as the components of $\partial_a \phi$, except that the zero-th components has an extra minus sign – if, in some chosen inertial coordinates, $\partial_a \phi = (A, B, C, D)$, then $\partial^a \phi = (-A, B, C, D)$.

Vectors vs components

Many textbooks don't distinguish between a vector V and its components V^a. People will often think of a vector as a list of components V^a, which transforms in a certain way under Lorentz transformations.

For us, a vector ("at p") is an ordered pair of points in $\mathbb{M}^4$, which we write as $V = q - p$. The components of the vector V^a are the numbers $V^1, \ldots, V^3$ such that

$$V = V^a e_a,$$

where e_a are a set of (pseudo)-orthonormal basis vectors, $m(e_a, e_b) = m_{ab}$. The inertial coordinates x^a associated with the basis vectors e_a are coordinates on $\mathbb{M}^4$, chosen such that

$$q - p = x^a(q) e_a.$$

In other words, the coordinates of the point q are the same as the components of the vector $q - p$.

Using this, we can *derive* the transformation law for the components of a vector – we don't just define a vector as a list of numbers that transforms as $X'^{(}a') = \Lambda_a{}^{a'} X^a$.

We will follow the same pattern in GR – rather than defining vectors, tensors etc. as lists or arrays which transform in certain ways under coordinate transformations, we will define certain vector spaces and then *derive* the transformation laws.

Advantages of this approach:

- It is clear that at least some examples of these objects (vectors, tensors etc.) exist as mathematical objects.
- It helps to avoid confusion: e.g. when we change coordinates, the *components* of a vector change, but the vector itself is invariant.
- The relationship between vectors, tangents to curves, partial derivatives etc. is made clear.
- It leads to an easy way to *compute* the transformation laws for various objects, including non-tensorial objects like the Christoffel symbols. In the alternative approach, we often demand that a certain object transforms "in the way it should", and then use this to derive the appropriate transformation laws for various objects.