GR lecture 2-1

Vectors and covectors as distinct objects; Galilean boosts

I. SOLVING EXERCISES FROM LAST WEEK.

II. VECTORS AS ARROWS, COVECTORS AS STACKS OF PLANES

In the last lecture, we pushed a particular point of view on the issue of upper/lower indices: these are just two mutually conjugate bases in a space where all the usual rules apply. In particular, we've been working in a geometry which has the ordinary scalar product to define lengths and angles, and we just happened to have a non-orthonormal basis. From that point of view, v^i and v_i are just different ways to parameterize a vector $\mathbf{v}$.

There exists a different point of view, which is ultimately more fundamental. There, we treat "vectors" v^i and "covectors" u_i as completely different things. Sure, they can be related by a metric as in $v_i = g_{ij}v^j$, but one might sometimes work in a geometry with no metric, or with more than one metric, or – as in GR – when the metric is another variable field, not given in advance. In fact, even when a metric is given, some quantities are more naturally viewed as vectors, and others as covectors. For example, in a crystal lattice, positions $\mathbf{x}$ are written in the basis $\mathbf{e}_i$, while wavevectors $\mathbf{k}$ are written in the co-basis $\mathbf{e}^i$.

In this new picture, our intuitive notion of a vector as an arrow – i.e. a displacement in space – is captured by the upper-index objects v^i. In particular, coordinates x^i have an upper index. The most basic object with a lower index is then the spatial derivative:

$$\partial_i = \frac{\partial}{\partial x^i} \ . \tag{1}$$

More concretely, for any scalar field $f(x)$, the gradient $\partial_i f$ is a covector – a quantity with a lower index. To verify that this is the case, we simply note that $\partial_i f$ can be contracted with an infinitesimal displacement dx^i, as befits a lower&upper index pair:

$$\partial_i f\, dx^i = \frac{\partial f}{\partial x^1}dx^1 + \frac{\partial f}{\partial x^2}dx^2 + \frac{\partial f}{\partial x^3}dx^3 = df \ . \tag{2}$$

This contraction is simply the change df in the value of f as we move along the interval dx^i. Note that eq. (2) is true for any set of coordinates x^i, regardless of their orthonormality or any other geometric meaning: it is simply the decomposition of a differential in terms of partial derivatives. In other words, it is a "scalar product" that looks the same in every basis – not a metric-dependent contraction of two upper indices, but a fundamental contraction of an upper index with a lower one.

Exercise 1. *Explain now why a "wavevector"* $\mathbf{k}$ *should be thought of as a covector* k_i*. How about momentum* $\mathbf{p}$*?*

What, then, is an appropriate geometric intuition for a covector? It is definitely not an arrow. Absent a metric, absent a notion of lengths and angles, there is no way to produce a scalar by multiplying two arrows. Let us again examine the gradient $\partial_i f$. At first, let's return to familiar metric geometry, where we do know lengths and angles. In calculus, we were taught to think of $\boldsymbol{\partial} f$ as a vector. Recall how this vector is constructed. We draw the surfaces of equal values of f. At each point, the equal-value surfaces have a certain orientation, and a certain density. The magnitude of the gradient vector $\boldsymbol{\partial} f$ is then given by the density of the equal-value surfaces, and its direction is defined as orthogonal to them. Note that this requires the concept of orthogonality, i.e. a metric! However, we can now play the same trick as when we got rid of vector products in Lecture 1-1: we can just skip the last step. The covector $\partial_i f$ should be thought of as the stack of equal-value surfaces of f near the point in which we are interested! More generally, we can think of covectors as stacks of planes: the direction of the covector is captured by the planes' orientation, and the magnitude of the covector is captured by their density. A scalar product $u_i v^i$ now has an intuitive meaning: it counts how many of the planes in the u_i stack are pierced by the v^i arrow. No metric needed! In particular, eq. (2) can now be read as follows: the change df in the value of f is given by the "number" of equal-value surfaces that are traversed as we travel along dx^i.

One can also think of covectors more abstractly: a covector u_i can be defined by its scalar product $u_i v^i$ with all possible vectors (of course, a spanning set of v^i's is enough for this). In other words, covectors are simply linear functions from vectors into scalars. Note that this definition is completely symmetric: we could also define a vector as a linear function from covectors into scalars. The asymmetry between vectors and covectors must be introduced separately, e.g. when we define that vectors, not covectors, describe displacements ("arrows") in space.

III. COMPONENTS VS. BASES

The picture we laid out above is quite consistent and self-sufficient: dx^i is a vector, and $\partial_i = \partial/\partial x^i$ is a covector. However, in some texts, one encounters a language which, at first sight, seems opposite. It is nice to have the mental agility to switch back and forth between different points of view on the same concepts. The seemingly opposite language has to do with the fact that, as we've seen, components in one basis have the same index structure (upper/lower) as basis elements of the opposite basis. Thus, dx^i can be thought of as the components of a vector, or as a basis of covectors. Similarly, ∂_i can be thought of as the components of a covector, or as a basis of vectors. This makes more sense than it may seem at first: if ∂_i is a basis of vectors, then $v^i\partial_i$ is a vector with components v^i: the components have an upper index, as they should. Geometrically, $v^i\partial_i$ is a directional derivative, i.e. a derivative in the particular direction of v^i. It makes sense to identify it with the vector v^i itself. For example, $\partial/\partial x^1$ can be intuitively identified with an arrow pointing in the x^1 direction, i.e. in the direction of changing x^1 with (x^2, x^3) held fixed. Similarly, there are two ways to make sense of the fact that dx^i is a basis of vectors. First, we can just say that $u_i dx^i$ is a linear map that makes a scalar out of any displacement dx^i, and thus we can identify it with a covector u_i. Second, we can geometrically consider e.g. dx^1 as the gradient of the coordinate x^1. In other words, a covector whose direction is set by the surfaces of constant x^1, and whose magnitude is set by how fast the coordinate x^1 changes as we move in space (with no regard to what the other coordinates are doing, which is very different from the case of the partial derivative [=vector] $\partial/\partial x^1$).

IV. GALILEAN SYMMETRY AND ITS DOWNFALL

Let's get back to physics. In the above, we've been building mathematical tools to work with flat space, with an eye towards curved space. GR, however, is not just about curved space: it is about curved spacetime. So, we should first make sure that we understand flat spacetime, also known as Minkowski space. That is the subject of Special Relativity.

Let us start from historical basics. The laws of Nature are invariant under a number of geometric transformations: translations in time (there is no preferred $t = 0$), translations in space (there is no preferred $\mathbf{x} = 0$), rotations in space (there is no preferred orientation for the (x, y, z) axes). The great Galilei discovered another such symmetry – the symmetry of passing into a reference frame that moves with velocity $\mathbf{v}$ with respect to our old one. In modern terminology, we call such transformations boosts:

$$t \ \rightarrow \ t \ ; \quad \mathbf{x} \ \rightarrow \ \mathbf{x} - \mathbf{v}t \ . \tag{3}$$

As long as people studied elementary mechanics, this symmetry was near and dear to their hearts. However, gradually, they turned their attention to other phenomena, such as heat, sound, fluids, solids, various waves... Many of these are associated with a medium, and thus with a preferred reference frame: the frame in which the medium is at rest. After a few decades of thinking about this stuff, people kind of forgot that velocity is relative, that boosts are supposed to be a symmetry at all. In particular, they weren't duly surprised when they discovered the magnetic force law $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$, or, more concretely, the magnetic force between two parallel currents:

$$\mathbf{F}_{\text{mag}} = \frac{\mu_0 I_1 I_2 \mathbf{r}}{2r^2} = \frac{\mu_0 \lambda_1 v_1 \lambda_2 v_2 \mathbf{r}}{2r^2} \ , \tag{4}$$

where $\lambda_{1,2}$ are the densities of moving charges per unit length, and $v_{1,2}$ are their velocities. The magnetic force clearly breaks the Galilean boost symmetry (3), and so do the Maxwell equations. In the historical context, people didn't worry about this too much: electromagnetism was probably just another of those things like sound or hydrodynamics – it probably happened in some medium, with a preferred reference frame.

– TO BE CONTINUED –