Vectors, metric and the connection

1 Contravariant and covariant vectors

1.1 Contravariant vectors

Imagine a particle moving along some path in the 2-dimensional flat $x-y$ plane. Let its trajectory be given by $\{x(t), y(t)\}$. We wish to write the velocity vector at some time t_0. The position at this time is $\{x(t_0), y(t_0)\}$. To get the velocity we look at the position at a slightly later time $t_0+\delta t$, when the position will be $\{x_0+\delta x, y_0+\delta y\}$. The components of the velocity vector are defined by

$$\begin{aligned} V^x &= \frac{\delta x}{\delta t} \\ V^y &= \frac{\delta y}{\delta t} \end{aligned} \qquad (1)$$

where it is understood that we should take the limit $\delta t \to 0$.

Now suppose we wish to use polar coordinates $\{r, \theta\}$ instead. How do we write the components of the velocity vector? We will define the required components in a manner similar to what we did for the Cartesian coordinates above

$$\begin{aligned} V^r &= \frac{\delta r}{\delta t} \\ V^\theta &= \frac{\delta \theta}{\delta t} \end{aligned} \qquad (2)$$

Knowing $\{V^x, V^y\}$ we should be able to compute $\{V^r V^\theta\}$. We have

$$\begin{aligned} \delta r &= \frac{\partial r}{\partial x}\delta x + \frac{\partial r}{\partial y}\delta y \\ \delta \theta &= \frac{\partial \theta}{\partial x}\delta x + \frac{\partial \theta}{\partial y}\delta y \end{aligned} \qquad (3)$$

Dividing both sides of the above equation by δt, we find

$$\begin{aligned} V^r &= \frac{\partial r}{\partial x}V^x + \frac{\partial r}{\partial y}V^y \\ V^\theta &= \frac{\partial \theta}{\partial x}V^x + \frac{\partial \theta}{\partial y}V^y \end{aligned} \qquad (4)$$

We have used the example of a velocity vector here, but these transformation rules *define* any vector. Thus if we consider the acceleration vector, then its components in the $\{x, y\}$ coordinate system and in the $\{r, \theta\}$ coordinate system will be related by the same transformation rules. Any set of quantities $\{W^x, W^y\}$ that transform in the manner (4) to give the quantities $\{W^r, W^\theta\}$ will be said to define a *contravariant vector*.

1.2 Covariant vectors

Consider a function

$$f(x, y) = x^2 + 2y^2 \tag{5}$$

We can compute the gradient $\{\nabla f_x, \nabla f_y\}$ of this function

$$\begin{aligned} \nabla f_x &= \frac{\partial f}{\partial x} = 2x \\ \nabla f_y &= \frac{\partial f}{\partial y} = 4y \end{aligned} \tag{6}$$

Let us ask how the gradient looks in polar coordinates. Now we have the two components

$$\begin{aligned} \nabla f_r &= \frac{\partial f}{\partial r} \\ \nabla f_\theta &= \frac{\partial f}{\partial \theta} \end{aligned} \tag{7}$$

There are two ways to compute these components. The first is to write f in polar coordinates, and then carry out the partial derivatives

$$f = r^2 \cos^2\theta + 2r^2 \sin^2\theta = r^2(1 + \sin^2\theta) \tag{8}$$

$$\begin{aligned} \nabla f_r &= \frac{\partial f}{\partial r} = 2r(1 + \sin^2\theta) \\ \nabla f_\theta &= \frac{\partial f}{\partial \theta} = 2r^2 \sin\theta\cos\theta \end{aligned} \tag{9}$$

But we are more interested in seeing for we could have obtained this polar coordinate result directly from the Cartesian coordinate expressions, because we can then use such transformation rules for many other vectors of the same kind. We have

$$\begin{aligned} \nabla f_r &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} = \nabla f_x \frac{\partial x}{\partial r} + \nabla f_y \frac{\partial y}{\partial r} \\ \nabla f_\theta &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta} = \nabla f_x \frac{\partial x}{\partial \theta} + \nabla f_y \frac{\partial y}{\partial \theta} \end{aligned} \tag{10}$$

Substituting the values of $\{\nabla f_x, \nabla f_y\}$ from (6), and putting in the needed partial derivatives, we find that we reproduce the values (9). But our interest is more in the transformation rules that effect the coordinate changes. Comparing (10) with (4) we see that the transformation rules are not the same; the partial derivatives used are 'inverses' of each other. Quantities that transform like the gradient will be called *covariant vectors*.

1.3 Dot products

In ordinary vector calculus you would have learnt that the components of a vector change under a rotation of coordinates, but a scalar is left unchanged. One way to get a scalar is by taking the dot product of two vectors. Now that we have two kinds of vectors, let us see how we should take the dot product.

Start with a coordinate system ξ^i. Take a contravariant vector V^i, and a covariant vector W_j. Then consider

$$V^i W_j \tag{11}$$

where we are using the summation convention. Under a change of coordinates to $\xi'^{i'}$ we will have

$$V'^{i'} = \frac{\partial \xi'^{i'}}{\xi^i} V^i \tag{12}$$

$$W'_{i'} = \frac{\partial \xi^j}{\partial \xi'^{i'}} W_j \tag{13}$$

Thus we have

$$V'^{i'} W'_{i'} = \frac{\partial \xi^j}{\partial \xi'^{i'}} \frac{\partial \xi'^{i'}}{\xi^i} V^i W_j \tag{14}$$

But we have

$$\frac{\partial \xi^j}{\partial \xi'^{i'}} \frac{\partial \xi'^{i'}}{\xi^i} = \delta^j_i \tag{15}$$

and we get

$$V'^{i'} W'_{i'} = \delta^j_i V^i W_j = V^i W_i \tag{16}$$

Thus $V^i W_i$ is a scalar, which does not change its value under a change of coordinates.

2 Connection

Suppose we have two vectors V^i and W^i at the same point. Suppose we want to know if these two vectors are equal. This is easy: we compute the vector $V^i - W^i$ and check if all components of this difference vector vanish.

What if V^i is a vector at one point $\{\xi^1, \dots \xi^d\}$, and W^i is a vector at a slightly different point $\{\xi^1 + \delta\xi^1, \dots \xi^d + \delta\xi^d\}$? Does it make sense to ask if these vectors are the same? If it does, then do we just take the difference of components $V^i - W^i$, or do we have to check something more complicated?

Not only does it make sense to ask this question, we *must* address this question. After all our basic plan is to look for 'geodesics' on curved spacetime, which are curves where we are trying to go 'straight'; i.e., we want to move forward without changing the direction of the velocity vector. We will need some way to know whether the velocity vector at one point of the worldline is the same as the velocity vector an infinitesimal distance further along the worldline.

In this section we will see how to set up the technology for comparing vectors at two infinitesimally separated points. This will lead to a very important concept – the *connection*.

3 Geodesics in the $x - y$ plane

Let us begin with a simple example. Consider the flat $x - y$ plane, and a particle moving (with no force) along a path $\{x(t), y(t)\}$. Since there is no force, the particle will follow a geodesic given by the equations

$$\frac{d^2x}{dt^2} = 0, \quad \frac{d^2y}{dt^2} = 0 \tag{17}$$

which have as their general solution the straight line motion

$$x(t) = V^x t + x_0, \quad y(t) = V^y t + y_0 \tag{18}$$

Now suppose we put polar coordinates $\{r, \theta\}$ on this plane. The motion will still be along straight lines, but we have to see how these lines are described in curvilinear coordinates.

Example: Substituting

$$x = r\cos\theta, \quad y = r\sin\theta \tag{19}$$

in (17), get the geodesic equations for $\{r, \theta\}$.

Solution: We have

$$\begin{aligned} \frac{d^2x}{dt^2} &= \frac{d^2(r\cos\theta)}{dt^2} = \frac{d^2r}{dt^2}\cos\theta - 2\sin\theta\frac{dr}{dt}\frac{d\theta}{dt} - r\sin\theta\frac{d^2\theta}{dt^2} - r\cos\theta(\frac{d\theta}{dt})^2 = 0 \\ \frac{d^2y}{dt^2} &= \frac{d^2(r\sin\theta)}{dt^2} = \frac{d^2r}{dt^2}\sin\theta + 2\cos\theta\frac{dr}{dt}\frac{d\theta}{dt} + r\cos\theta\frac{d^2\theta}{dt^2} - r\sin\theta(\frac{d\theta}{dt})^2 = 0 \end{aligned} \tag{20}$$

These equations give

$$\begin{aligned} \frac{d^2r}{dt^2} - r(\frac{d\theta}{dt})^2 &= 0 \\ \frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt} &= 0 \end{aligned} \tag{21}$$

which are the well known equations for dynamics in polar coordinates incorporating the fictitious 'centrifugal' and 'coreolis' terms.

Note that the geodesic equations in the 'curvilinear' polar coordinates are not simply

$$\frac{d^2r}{dt^2} = 0, \ \frac{d^2\theta}{dt^2} = 0 \quad (Wrong!) \tag{22}$$

. In this case we knew the result from the Cartesian description, but how will get the answer in a genuinely curved space where we do not have a Cartesian description?

Looking again at our polar coordinate equations, we realize that the equations are more complicated than (22) because the coordinates themselves change in a nontrivial way as we move along the straight line paths. For instance, the radial direction at any point of a circle depends on what point of the circle we are at. So we need a system to 'remove' such coordinate effects from the description of the vector.

3.1 Transporting a vector

Let us start at the point $\{x = x_0, y = 0\}$, which is $\{r = x_0, \theta = 0\}$ in polar coordinates. Let us take the vector at this point to be $\{V^x, 0\}$ in Cartesian coordinates. We can find its components in polar coordinates by using the transformation law for contravariant vectors

$$\begin{aligned} V'^r &= \frac{\partial r}{\partial x} V^x = \cos\theta V^x = V^x \\ V'^\theta &= \frac{\partial \theta}{\partial x} V^x = -\frac{\sin\theta}{r} V^x = 0 \end{aligned} \tag{23}$$

Now imagine transporting this vector 'without change' a little way around the circle of radius $r = x_0$; this brings the location of the vector to $\{r = x_0, \theta = \delta\theta\}$. Here we need to know what 'without change' means, but since we have the Cartesian system to help us, we know that at least in the Cartesian coordinate system the vector will continue to have the components $\{V^x, 0\}$. What about its components in polar coordinates? We find

$$\begin{aligned} V'^r &= \frac{\partial r}{\partial x} V^x = \cos\theta V^x = V^x \\ V'^\theta &= \frac{\partial \theta}{\partial x} V^x = -\frac{\sin\theta}{r} V^x = -\delta\theta V^x \end{aligned} \tag{24}$$

where in the last steps we have used the approximations valid for small $\delta\theta$. Comparing with (23) we see that even though the vector 'did not change', its components in polar coordinates did, by the amount

$$\delta V^r = 0, \quad \delta V^\theta = -\delta\theta V^x = -\delta\theta V^r \tag{25}$$

There are two things we note about these changes:

(i) The change in the vector components is proportional to the amount of displacement; in this case given by $\delta\theta$.

(ii) The change in the vector components is proportional to components of the vector being transported.

Both these proportionalities make sense: if we move twice as far we will get twice as much change, and if we move a vector that is twice as long then we will get twice the change.

4 The general definition of connection

Let us use the intuition developed from the above example to write down a general expression for transport of vectors. Let the manifold have curvilinear coordinates $\xi^i, i = 1, \ldots d$. (The manifold is arbitrary; it need not be flat.) Start at a point $\{\xi^1, \ldots \xi^d\}$, and a vector at this point with components $\{V^1, \ldots V^d\}$. Imagine transporting this vector 'without change' to a neighboring point $\{\xi^1 + d\xi^1, \ldots \xi^d + \delta\xi^d\}$. Let the components of the vector change to $\{V^1 + \delta V^1, \ldots V^d + \delta V^d\}$. Whatever the definition of 'without change' (we will give the definition later) we can assume quite generally that

$$\begin{aligned} \delta V^i &\propto \delta\xi^j \\ \delta V^i &\propto V^k \end{aligned} \tag{26}$$

We can therefore introduce a proportionality constant and write the above two proportionalities as an equality

$$\delta V^i = -\sum_{j=1}^{d}\sum_{k=1}^{d}\Gamma^i_{jk}\delta\xi^j V^k \tag{27}$$

where the negative sign is introduced to follow historical convention. Even though Γ^i_{jk} looks like a tensor (with one contravariant and two covariant indices) we will see shortly that it is not a tensor; it does not change the way a tensor should when we change coordinates. Nevertheless it is convenient to use the notation of upper and lower indices in the sums in (27), and to also use the summation convention that we used for tensors; thus we write

$$\delta V^i = -\Gamma^i_{jk}\delta\xi^j V^k \tag{28}$$

5 Computing the connection in flat space with curvilinear coordinates

Let us start with flat space, where we know what it means for two vectors to be the same at different points. We will follow the method used in the above example of polar coordinates, and derive a general expression for the connection for arbitrary curvilinear coordinates.

Let the flat space have Euclidean coordinates $\{x^1, \ldots, x^d\}$. In these coordinates the metric is

$$g_{ij} = \delta_{ij} \tag{29}$$

and the connection is zero

$$\Gamma^i_{jk} = 0 \tag{30}$$

since the components of a vector do not change when we transport it to a new point. Let the vector be V^i at a point $x = \{x^1, \dots x^d\}$ We will transport the vector to a point $x + dx = \{x^1 + dx^1, \dots x^d + dx^d\}$.

Now consider curviliear coordinates $\{\xi'^1, \dots \xi'^d\}$ on this same space. At a point $\xi' = \{\xi'^1, \dots \xi'^d\}$ we have for V the components

$$V'^{i'} = \frac{\partial \xi'^{i'}}{\partial x^i}(\xi') V^i \tag{31}$$

where we have chosen to use the label ξ to label the point instead of the coordinate x. After transport to the point $\xi' + \xi'$ we have the components

$$V'^{i'} + \delta V'^{i'} = \frac{\partial \xi'^{i'}}{\partial x^i}(\xi' + d\xi') V^i \tag{32}$$

where we have used the fact that the Cartesian components V^i have not changed under the transport. Thus we have

$$\delta V'^{i'} = \frac{\partial}{\partial \xi'^{j'}} \Big(\frac{\partial \xi'^{i'}}{\partial x^i} \Big) \delta \xi'^{j'} V^i \tag{33}$$

We cannot carry out the second order partial derivative directly, since

$$\frac{\partial}{\partial \xi'^{j'}} \Big(\frac{\partial \xi'^{i'}}{\partial x^i} \Big) \neq \frac{\partial^2 \xi'^{i'}}{\partial \xi'^{j'} \partial x^i} \quad !! \tag{34}$$

But we can use a simple trick to convert the first derivative into one with respect to the ξ coordinates, so that at the end we will get both derivatives with respect to the ξ coordinates. We have

$$\frac{\partial \xi'^{i'}}{\partial x^l} \frac{\partial x^l}{\partial \xi'^{k'}} = \delta^{i'}_{k'} \tag{35}$$

Differentiating both sides,

$$\frac{\partial}{\partial \xi'^{j'}} \Big(\frac{\partial \xi'^{i'}}{\partial x^l} \Big) \frac{\partial x^l}{\partial \xi'^{k'}} + \frac{\partial \xi'^{i'}}{\partial x^l} \frac{\partial}{\partial \xi'^{j'}} \Big(\frac{\partial x^l}{\partial \xi'^{k'}} \Big) = 0 \tag{36}$$

Thus

$$\frac{\partial}{\partial \xi'^{j'}} \Big(\frac{\partial \xi'^{i'}}{\partial x^l} \Big) \frac{\partial x^l}{\partial \xi'^{k'}} = - \frac{\partial \xi'^{i'}}{\partial x^l} \frac{\partial^2 x^l}{\partial \xi'^{j'} \xi'^{k'}} \tag{37}$$

where now we have been able to write the doubly differentiated term as a proper second partial derivative. We still have a factor $\frac{\partial x^l}{\partial \xi'^{j'}}$ on the LHS multiplying what we really want, so we ‘strip off’ this factor by multiplying both sides by the ‘inverse derivative’

$$\frac{\partial}{\partial \xi'^{j'}} \Big(\frac{\partial \xi'^{i'}}{\partial x^l} \Big) \frac{\partial x^l}{\partial \xi'^{k'}} \frac{\partial \xi'^{k'}}{\partial x^i} = - \frac{\partial \xi'^{i'}}{\partial x^l} \frac{\partial^2 x^l}{\partial \xi'^{j'} \xi'^{k'}} \frac{\partial \xi'^{k'}}{\partial x^i} \tag{38}$$

Now the LHS becomes

$$\frac{\partial}{\partial\xi'^{j'}}\Big(\frac{\partial\xi'^{i'}}{\partial x^l}\Big)\frac{\partial x^l}{\partial\xi'^{k'}}\frac{\partial\xi'^{k'}}{\partial x^i}=\frac{\partial}{\partial\xi'^{j'}}\Big(\frac{\partial\xi'^{i'}}{\partial x^l}\Big)\delta^l_i=\frac{\partial}{\partial\xi'^{j'}}\Big(\frac{\partial\xi'^{i'}}{\partial x^i}\Big) \tag{39}$$

so that we finally get

$$\frac{\partial}{\partial\xi'^{j'}}\Big(\frac{\partial\xi'^{i'}}{\partial x^i}\Big)=-\frac{\partial\xi'^{i'}}{\partial x^l}\frac{\partial^2 x^l}{\partial\xi'^{j'}\xi'^{k'}}\frac{\partial\xi'^{k'}}{\partial x^i} \tag{40}$$

Using this in (33), we get

$$\delta V'^{i'}=-\frac{\partial\xi'^{i'}}{\partial x^l}\frac{\partial^2 x^l}{\partial\xi'^{j'}\xi'^{k'}}\frac{\partial\xi'^{k'}}{\partial x^i}\delta\xi'^{j'}V^i=-\frac{\partial\xi'^{i'}}{\partial x^l}\frac{\partial^2 x^l}{\partial\xi'^{j'}\xi'^{k'}}\frac{\partial\xi'^{k'}}{\partial x^i}\delta\xi'^{j'}\frac{\partial x^i}{\partial\xi'^{l'}}V'^{l'}=-\frac{\partial\xi'^{i'}}{\partial x^l}\frac{\partial^2 x^l}{\partial\xi'^{j'}\xi'^{k'}}\delta\xi'^{j'}V'^{k'} \tag{41}$$

where in the last step we have used $\frac{\partial\xi'^{k'}}{\partial x^i}\frac{\partial x^i}{\partial\xi'^{l'}}=\delta^{k'}_{l'}$. Comparing with our definition (28) we get

$$\Gamma'^{i'}_{j'k'}=\frac{\partial\xi'^{i'}}{\partial x^l}\frac{\partial^2 x^l}{\partial\xi'^{j'}\xi'^{k'}} \tag{42}$$

for the curvilinear coordinates ξ'.

We note an important fact from this expression:

$$\Gamma'^{i'}_{j'k'}=\Gamma'^{i'}_{k'j'} \tag{43}$$

Thus the connection is symmetric in its two lower indices, at least in this case where we have placed curvilinear coordinates in flat space. When we move to genuinely curved space we will keep this symmetry, though it should be noted that sometimes we define other kinds of connections that do *not* have this symmetry; the antisymmetric part of the connection in those cases is called the *torsion*.

6 Getting a better form for the connection

The above expression for $\Gamma'^{i'}_{j'k'}$ is a little unwieldy, so let us see if we can replace the partial derivatives here by something else. The metric in Cartesian coordinates was simple (29), but the metric in the coordinates ξ is

$$g'_{i'k'}=\frac{\partial x^i}{\partial\xi'^{i'}}\frac{\partial x^k}{\partial\xi'^{k'}}\delta_{ik} \tag{44}$$

Thus we have

$$\frac{\partial}{\partial\xi'^{j'}}g'_{i'k'}\equiv g'_{i'k',j'}=\frac{\partial^2 x^i}{\partial\xi'^{j'}\partial\xi'^{i'}}\frac{\partial x^k}{\partial\xi'^{k'}}\delta_{ik}+\frac{\partial x^i}{\partial\xi'^{i'}}\frac{\partial^2 x^k}{\partial\xi'^{j'}\xi'^{k'}}\delta_{ik} \tag{45}$$

Thus differentiating the metric produces the same kind of partial derivatives that appear in the expression for the connection. Thus we can try to write the connection in terms of derivatives of the metric. There are two terms on the RHS of (45). Since the second is symmetric in j', k' We can remove the second by antisymmetrizing in these variables

$$g'_{i'k',j'} - g'_{i'j',k'} = \frac{\partial^2 x^i}{\partial \xi'^{j'} \partial \xi'^{i'}} \frac{\partial x^k}{\partial \xi'^{k'}} \delta_{ik} - \frac{\partial^2 x^i}{\partial \xi'^{k'} \partial \xi'^{i'}} \frac{\partial x^k}{\partial \xi'^{j'}} \delta_{ik} \tag{46}$$

We still have two terms on the RHS, but note that

$$g'_{j'k',i'} = \frac{\partial^2 x^i}{\partial \xi'^{i'} \partial \xi'^{j'}} \frac{\partial x^k}{\partial \xi'^{k'}} \delta_{ik} + \frac{\partial x^i}{\partial \xi'^{j'}} \frac{\partial^2 x^k}{\partial \xi'^{i'} \xi'^{k'}} \delta_{ik} = \frac{\partial^2 x^i}{\partial \xi'^{i'} \partial \xi'^{j'}} \frac{\partial x^k}{\partial \xi'^{k'}} \delta_{ik} + \frac{\partial x^k}{\partial \xi'^{j'}} \frac{\partial^2 x^i}{\partial \xi'^{i'} \xi'^{k'}} \delta_{ik} \tag{47}$$

where in the second step we have interchanged the dummy indices i, k on the second term. Then we find

$$g'_{i'k',j'} - g'_{i'j',k'} + g_{j'k',i'} = 2 \frac{\partial^2 x^i}{\partial \xi'^{j'} \partial \xi'^{i'}} \frac{\partial x^k}{\partial \xi'^{k'}} \delta_{ik} \tag{48}$$

and on the RHS we see the kind of second order partial derivative that we need in the expression for the connection. The first order derivative term is however not quite the one we need. We 'strip it off' in the usual way

$$[g'_{i'k',j'} - g'_{i'j',k'} + g_{j'k',i'}] \frac{\partial \xi'^{k'}}{\partial x^l} = 2 \frac{\partial^2 x^i}{\partial \xi'^{j'} \partial \xi'^{i'}} \frac{\partial x^k}{\partial \xi'^{k'}} \frac{\partial \xi'^{k'}}{\partial x^l} \delta_{ik} = 2 \frac{\partial^2 x^i}{\partial \xi'^{j'} \partial \xi'^{i'}} \delta_{ik} \delta^k_l = 2 \frac{\partial^2 x^i}{\partial \xi'^{j'} \partial \xi'^{i'}} \delta_{il} \tag{49}$$

Multiplying both sides by $\frac{\partial \xi'^{l'}}{\partial x^m} \delta^{ml}$ we get

$$[g'_{i'k',j'} - g'_{i'j',k'} + g_{j'k',i'}] \frac{\partial \xi'^{l'}}{\partial x^m} \frac{\partial \xi'^{k'}}{\partial x^l} \delta^{ml} = 2 \frac{\partial^2 x^i}{\partial \xi'^{j'} \partial \xi'^{i'}} \frac{\partial \xi'^{l'}}{\partial x^m} \delta_{il} \delta^{ml} \tag{50}$$

Note that

$$\frac{\partial \xi'^{l'}}{\partial x^m} \frac{\partial \xi'^{k'}}{\partial x^l} \delta^{ml} = g'^{l'k'} \tag{51}$$

Thus we finally get

$$\frac{1}{2} g'^{l'k'} [g'_{i'k',j'} - g'_{i'j',k'} + g_{j'k',i'}] = \frac{\partial^2 x^m}{\partial \xi'^{j'} \partial \xi'^{i'}} \frac{\partial \xi'^{l'}}{\partial x^m} \tag{52}$$

Dropping the primes (and interchanging $i \leftrightarrow l$), get our expression for the connection

$$\Gamma^i_{jk} = \frac{1}{2} g^{il} [g_{lj,k} + g_{lk,j} - g_{jk,l}] \tag{53}$$