Introduction to tensor calculus

Dr. J. Alexandre
King's College

These lecture notes give an introduction to the formalism used in Special and General Relativity, at an undergraduate level.

Contents

1 Tensor calculus in Special Relativity

1.1 Motivations

An inertial frame is a frame where an observer at rest does not feel any force and where a free motion leads to a constant velocity (direction and intensity). If one wishes to go from an inertial frame to another, i.e. to describe the physics with another point of view, the changes of coordinates in Newtonian mechanics are given by the Galilée transformations:

$$\begin{aligned} t' &= t \\ \vec{r'} &= \vec{r} + \vec{v}t, \end{aligned} \tag{1}$$

where $\vec{v}$ is the velocity of one frame with respect to the other. The transformations (1) imply that the acceleration $\vec{a}$ of a point particle is invariant in this change of frame and thus a force $\vec{f}$ felt by the particle is also invariant if you impose the fundamental law of newtonian mechanics $\vec{f} = m\vec{a}$ to be valid in any inertial frame.

Since the time is an invariant quantity in Galilée transformations, the composition of velocities is obviously given by

$$\frac{d\vec{r'}}{dt'} = \frac{d\vec{r}}{dt} + \vec{v}. \tag{2}$$

It happens that the transformation (2) is not respected by the experiments (historically: Michelson's experiment on diffraction patterns which implies that the speed of light is independent of the inertial frame considered). The issue Einstein found was "simply" to abandon the transformations (1) and to assume that $t' \neq t$.

The independence of the speed of light (that we note c) with respect to the inertial frame leads us to the new postulate: the free motion of a point particle over the distance $|\vec{r}|$ during the time t is such that the quantity $-c^2t^2 + |\vec{r}|^2$ is independent of the inertial observer.

Therefore if the free particle moves over the distance $|\vec{r}|$ measured in a frame $\mathcal{O}$ (and $|\vec{r}'|$ measured in a frame $\mathcal{O}'$) during the time t measured in $\mathcal{O}$ (and t' measured in $\mathcal{O}'$), we will have

$$-c^2t^2 + |\vec{r}|^2 = -c^2t'^2 + |\vec{r'}|^2. \tag{3}$$

Eq.(3) must be considered as the starting point from which everything follows. It has the status of a postulate.

1.2 Proper-time

Before going to the Lorentz transformations, let us define the *proper time* τ as the time measured in a frame by a clock at rest. In another frame, this clock is labelled by $(t, \vec{r})$ and from Eq.(3) for an infinitesimal motion, we can write

$$d\tau^2 = dt^2 - \frac{1}{c^2} d\vec{r}^2, \tag{4}$$

such that

$$\frac{dt}{d\tau} = \frac{1}{\sqrt{1 - \frac{1}{c^2}\left(\frac{d\vec{r}}{dt}\right)^2}}$$

Since the proper-time is measured in the rest frame, $\frac{d\vec{r}}{dt} = \vec{v}$ is the relative velocity between the two frames. We denote then $\gamma = 1/\sqrt{1 - v^2/c^2}$ and have

$$\frac{dt}{d\tau} = \gamma. \tag{5}$$

We will see in the next section the new law of velocities composition from which we can justify that we always have $v \leq c$, such that γ is always defined.

For a general motion and not only uniform, we can consider the tangent inertial frame at any moment and Eq.(5) will be valid all along the trajectory, with γ depending on the time t. Thus the proper time measured on a finite motion between the times t_1 and t_2 will be

$$\Delta\tau = \int_{t_1}^{t_2} \frac{dt}{\gamma(t)} < \Delta t,$$

such that the proper-time is always smaller than the time measured in any other frame.

1.3 Lorentz transformations

We will now look for the new transformations of coordinates, in a change of inertial frame, which are consistent with the assumption (3). For this, we will consider a motion along the x-axis, leaving the coordinates y and z unchanged and we will not suppose that the time is an invariant quantity. Furthermore, we will look for linear transformations since we impose them to be valid uniformly in space and time. We will thus look for transformations of the form

$$\begin{aligned} t' &= Dx + Et \\ x' &= Ax + Bt \\ y' &= y \\ z' &= z. \end{aligned} \tag{6}$$

Plugging the transformations (6) in the assumption (3), we obtain

$$\begin{aligned} c^2D^2 - A^2 &= -1 \\ c^2E^2 - B^2 &= c^2 \\ c^2ED - AB &= 0 \end{aligned}$$

Then if we consider the origin of the spatial coordinates of $\mathcal{O}$, we obtain $x' = Bt$ and $t' = Et$, such that

$$\frac{B}{E} = \frac{x'}{t'} = v,$$

by definition of v. We now have the 4 equations to determine A, B, D, E and find easily the Lorentz transformation

$$\begin{aligned} ct' &= \gamma\left(ct + \frac{v}{c}x\right) \\ x' &= \gamma\left(x + vt\right), \end{aligned} \tag{7}$$

where $\gamma = 1/\sqrt{1 - v^2/c^2}$.

As a first consequence, let us come back to the composition of velocities. From Eqs.(7), we obtain

$$\frac{dx'}{dt'} = \frac{dx + vdt}{dt + \frac{v}{c^2}dx} = \frac{\frac{dx}{dt} + v}{1 + \frac{v}{c^2}\frac{dx}{dt}}.$$

If we suppose that $\frac{dx}{dt} << c$, we are led to $\frac{dx'}{dt'} \simeq \frac{dx}{dt} + v$, which is the non-relativistic limit. Then we see that if $\frac{dx}{dt} = c$, we also have $\frac{dx'}{dt'} = c$, which was expected from the invariance of the speed of light and which indicates that the latter is the maximum speed that one can obtain.

This leads us to the following classification. Two events $(t_1, \vec{r}_1)$ and $(t_2, \vec{r}_2)$ related in a way such that $c^2(t_2 - t_2)^2 > (\vec{r}_2 - \vec{r}_1)^2$ are said to be separated by a time-like interval (they can interact via an information that has a speed smaller than c). If they are such that $c^2(t_2 - t_2)^2 < (\vec{r}_2 - \vec{r}_1)^2$, they are said to be separated by a space-like interval (they cannot interact). Finally, if they are such that $c^2(t_2 - t_2)^2 = (\vec{r}_2 - \vec{r}_1)^2$, they are said to be separated by a light-like interval (they can interact only via a ray of light).

Another consequence of the lorentz transformations is the so-called 'length contraction': if you measure an infinitesimal length dx' in a rest frame with the corresponding length dx measured in another frame at one given time t, you obtain from (7)

$$\frac{dx'}{dx} = \gamma,$$

such that the length measured in the rest frame (the *proper-length*) is always larger than the one measured in any other inertial frame. This implies a 'contraction' of an object when it is measured by a moving observer.

Finally, a general Lorentz transformation includes also the coordinates y and z, but always in a way such that the proper-time $d\tau$ remains an invariant quantity. We note that the invariance of $d\tau$ is also satisfied by rotations in space, leaving the time unchanged, and not only in a change of inertial frame.

1.4 4-vectors

We now come to the main purpose of this lesson. We wish to take into account these new effects in a formalism showing naturally the conservation of the infinitesimal proper-time (4).

In Galilée transformations, the time is an invariant quantity and we consider the transformations of vectors in space. In Lorentz transformations, the time depends on the inertial observer and thus should be part of a new vector in space-time, having 4 coordinates. We will write such a 4-component vector $x = (ct, \vec{r})$ such that its components are $x^0 = ct$ and $x^k = r^k$ where $k = 1, 2, 3$. Its general components will be noted x^μ where $\mu = 0, 1, 2, 3$. The greek letters will denote space-time indices and the latin letters the space indices.

A 4-vector is defined as a set of 4 quantities which transform under a Lorentz transformation in a change of inertial frame.

By definition, $x = (ct, \vec{r})$ is a 4-vector, called the 4-position (we will note with an arrow the vectors in the 3-dimensional space). Then we can use the proper-time which is an invariant quantity and define the 4-velocity u with components

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{dx^\mu}{dt}\frac{dt}{d\tau} = \gamma\frac{dx^\mu}{dt},$$

such that $u^0 = \gamma c$ and $\vec{u} = \gamma\vec{v}$, where $\vec{v}$ is the velocity in the 3-dimensional space. u is a 4-vector since it is constructed from another 4-vector and an invariant quantity. It must be stressed here that the quantities dx^μ/dt do not form a 4-vector: a Lorentz transformation on these quantities does not lead to the equivalent quantities dx'^μ/dt' in the new frame.

Another example of 4-vector is the 4-momentum $p = mu$, where m is the mass of a particle (a constant under Lorentz transformations). Its components are $p^0 = \gamma mc$ and $\vec{p} = \gamma m\vec{v}$, which is thus the relativistic momentum (in the 3-dimensional space). A Taylor expansion of γ for $v/c << 1$ gives

$$p^0 = \frac{1}{c}\left(mc^2 + \frac{1}{2}mv^2 + ...\right),$$

where we recognize the kinetic energy $mv^2/2$ and the mass energy mc^2 defined as the energy of the particle in the rest frame. We will thus call the energy of the particle the quantity $E = \gamma mc^2$ and the 4-momentum can be written

$$p = \left(\frac{E}{c}, \gamma m\vec{v}\right)$$

Other examples of 4-vectors will be seen after the definition of their scalar product.

1.5 Scalar product in Minkowsky space-time

The idea is to define a scalar product which is left invariant in a change of inertial frame, just as the usual scalar product is left invariant under a rotation in space.

Let x and y be two 4-vectors. We note $x = x^\mu e_\mu$ and $y = y^\nu e_\nu$, where the 4-vectors e_μ form a basis of the 4-dimensional space-time and the summation over repeated indices is understood. We want the scalar product $x.y$ of x and y to satisfy

$$\begin{aligned} x.y &= y.x \\ x.(y+z) &= x.y + x.z \\ (ax).y &= x.(ay) = a(x.y) \end{aligned}$$

where a is any real number. With these properties we have then $x.y = x^\mu y^\nu e_\mu.e_\nu = x^\mu y^\nu \eta_{\mu\nu}$ where

$$\eta_{\mu\nu} = \eta_{\nu\mu} = e_\mu.e_\nu.$$

By definition of the scalar product, we want it to be invariant in a change of frame, such that we have to take

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and $-x^0y^0 + x^ky^k$ is left invariant. $\eta_{\mu\nu}$ is called the metric tensor. We will see the definition of tensors in the next section and see that $\eta_{\mu\nu}$ is indeed a tensor. For the moment, we can consider it as an array 4×4. Note that we could have taken the opposite signs for the components $\eta_{\mu\nu}$. This would not have changed the Physics, of course, and is just a matter of convention.

We define then the new quantities

$$x_\mu = x^\nu \eta_{\mu\nu},$$

which are called the covariant components of x, whereas those with the indice up are called the contravariant components. We have then for any 4-vector x: $x_0 = -x^0$ and $x_k = x^k$ $(k = 1, 2, 3)$. With this definition, we can write the scalar product in a compact way

$$\begin{aligned} x.y &= x^\mu y^\nu \eta_{\mu\nu} \\ &= -x^0 y^0 + x^1 y^1 + x^2 y^2 + x^3 y^3 \\ &= x^0 y_0 + x^1 y_1 + x^2 y_2 + x^3 y_3 \\ &= x^\mu y_\mu = x_\mu y^\mu \end{aligned}$$

We note that in a scalar product we always sum a covariant indice with a contravariant one, and not two of the same kind.

We also define $\eta^{\mu\nu}$ as being the components of the inverse matrix of $\eta_{\mu\nu}$, such that

$$\eta^{\mu\rho}\eta_{\rho\nu} = \delta^\mu_\nu,$$

where δ^μ_ν is the Kronecker symbol (1 if $\mu = \nu$ and 0 otherwise). We have obviously $x^\mu = \eta^{\mu\nu} x_\nu$ and

$$\eta^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Finally, we note that $x^2 = 0$ does not imply $x = 0$, and we can have $x^2 < 0$. The space-time with this scalar product is called the Minkowsky space-time.

Let us check explicitely that this scalar product is invariant under the transformations (7). We have for two 4-vectors $p = (p^0, \vec{p})$ and $q = (q^0, \vec{q})$

$$\begin{aligned} p'^\mu q'_\mu &= -p'^0 q'^0 + p'^1 q'^1 + p'^2 q'^2 + p'^3 q'^3 \\ &= -\gamma^2 \left(p^0 + \frac{v}{c} p^1\right)\left(q^0 + \frac{v}{c} q^1\right) + \gamma^2 \left(p^1 + \frac{v}{c} p^0\right)\left(q^1 + \frac{v}{c} q^0\right) + p^2 q^2 + p^3 q^3 \\ &= -p^0 q^0 + p^1 q^1 + p^2 q^2 + p^3 q^3 \\ &= p^\mu q_\mu \end{aligned}$$

and find the expected result. In general, we define a Lorentz scalar as a quantity invariant under a change of inertial frame.

To give another illustration of this invariance, let us compute the square of the 4-momentum:

$$p^2 = p^\mu p_\mu = -(\gamma m c)^2 + (\gamma m \vec{v})^2 = -m^2 c^2,$$

which is indeed an invariant quantity. This last relation can also be written

$$E^2 = c^2|\vec{p}|^2 + m^2c^4.$$

Let us now see an example taken from classical electrodynamics. Maxwell's equations are of course relativistic since they describe the propagation of light and are at the origin of the theory of relativity! Thus the scalar quantities appearing in the solutions of Maxwell's equations must be Lorentz scalars. We have for example the phase of a monochromatic plane wave which can be written

$$\theta = -\omega t + \vec{k}\vec{r} = k^\mu x_\mu,$$

where $k^\mu = (\omega/c, \vec{k})$. Since x_μ is a 4-vector and its scalar product with k^μ gives a Lorentz scalar, we conclude that k^μ is a 4-vector. We can derive from this example the relativistic Doppler effect. Suppose that a source of light of proper frequency ω_0 (i.e. in a frame where the source is at rest) and proper-time τ moves with a constant velocity $\vec{v}$ with respect to an inertial observer who measures the frequency ω, the time t and the position $\vec{r}$ of the source. If $\vec{n}$ is the unit vector pointing from the observer towards the source at the time of emission, we know from the invariance of the phase θ that during an infinitesimal motion

$$\omega_0 d\tau = \omega dt - \vec{k}.d\vec{r} = \omega dt + \frac{\omega}{c}\vec{n}.d\vec{r},$$

such that,

$$\omega_0 = \omega\left(\frac{dt}{d\tau} + \frac{\vec{n}.d\vec{r}}{cdt}\frac{dt}{d\tau}\right) = \gamma\omega\left(1 + \frac{\vec{n}.\vec{v}}{c}\right)$$

We note that, unlike the non-relativistic case, we have a Doppler effect when $\vec{n}.\vec{v} = 0$.

Let us now consider the situation of the derivatives with respect to space-time indices. We have

$$\begin{aligned}
\frac{\partial(x^2)}{\partial x^\mu} &= \frac{\partial}{\partial x^\mu}(x^\rho x^\sigma \eta_{\rho\sigma}) \\
&= \eta_{\rho\sigma}(\delta^\rho_\mu x^\sigma + \delta^\sigma_\mu x^\rho) \\
&= \eta_{\sigma\mu}x^\sigma + \eta_{\mu\rho}x^\rho \\
&= 2x_\mu
\end{aligned}$$

such that the derivative with respect to a contravariant component is a covariant component. We can see in the same way that the derivative with respect to a covariant component is a contravariant component. Thus we note

$$\frac{\partial}{\partial x^\mu} = \partial_\mu \quad \text{and} \quad \frac{\partial}{\partial x_\mu} = \partial^\mu, \tag{8}$$

and the derivatives (8) are 4-vectors. If we come back to classical electrodynamics, the conservation of the electric charge can be written

$$0 = \frac{\partial \rho}{\partial t} + \vec{\nabla}.\vec{\mathrm{j}} = \frac{\partial (c\rho)}{\partial x^0} + \frac{\partial j^k}{\partial x^k} = \partial_\mu j^\mu \tag{9}$$

where $j^\mu = (c\rho, \vec{\mathrm{j}})$ is a 4-vector (the 4-current) since its scalar product with ∂_μ gives zero in any inertial frame (the 'continuity equation' (9) is of course Lorentz invariant since it is derived from Maxwell's equations).

1.6 Tensors in Minkowsky space-time

Now that we have defined 4-vectors, we are naturally led to a generalization of this classification of physical quantities with respect to their transformation under a change of inertial frame, what will be done with tensors. Let us first define for this the tensorial product $\mathcal{M} = M \otimes M$, where M is the Minkowsky space-time. As M, $\mathcal{M}$ is a vector space. Its dimension is 4×4 and its elements have the following properties

$$\begin{aligned} x \otimes (y + z) &= x \otimes y + x \otimes z \\ (ax) \otimes y &= x \otimes (ay) = a(x \otimes y), \end{aligned}$$

where x, y and z are 4-vectors and a is a scalar. With these properties, we can write in terms of components

$$x \otimes y = x^\mu y^\nu e_\mu \otimes e_\nu \tag{10}$$

where $e_\mu \otimes e_\nu$ are 4×4 basis tensors of $\mathcal{M}$. The 16 components of $x \otimes y$ are thus given by $x^\mu y^\nu$. A general tensor of rank two is a linear combination of terms like (10) and can be written

$$T_{(2)} = T^{\mu\nu} e_\mu \otimes e_\nu, \tag{11}$$

where $T^{\mu\nu}$ are its contravariant components. By making r tensorial products, one can define a tensor of rank r which has the general form

$$T_{(r)} = T^{\mu_1 \ldots \mu_r} e_{\mu_1} \otimes \ldots \otimes e_{\mu_r}.$$

We define the covariant components of a tensor as we did with the 4-vectors:

$$T_\mu^{\ \ \nu} = T^{\rho\nu} \eta_{\mu\rho} \quad \text{and} \quad T_{\mu\nu} = T^{\rho\sigma} \eta_{\mu\rho} \eta_{\nu\sigma},$$

and similarily for higher order tensors. Finally we note that a 4-vector is a tensor of rank 1.

Let us now come to a fundamental property of tensors. This property deals with the transformation law of a tensor under a change of inertial frame and can actually be seen as a definition of tensors in Minkowsky space-time. In a general Lorentz transformation, the components of a 4-vector transform like

$$x^\mu \longrightarrow x'^\mu = \Lambda^\mu_\nu x^\nu.$$

The specific example that was derived in section 3 is a particular case which took into account the transformations of x^0 and x^1 only. In this situation, the matrix Λ^μ_ν was

$$\Lambda^\mu_\nu = \left(\begin{array}{cccc} \gamma & \frac{v}{c}\gamma & 0 & 0 \\ \frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right),$$

but we consider now a general Lorentz transformation, affecting all the 4 components x^μ. The coefficients Λ^μ_ν being constant, we have $\partial x'^\mu / \partial x^\nu = \Lambda^\mu_\nu$ and thus

$$x'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} x^\nu. \tag{12}$$

A general 4-vector can then be written in the two different frames

$$x = x^\nu e_\nu \quad \text{and} \quad x = x'^\mu e'_\mu = \frac{\partial x'^\mu}{\partial x^\nu} x^\nu e'_\mu,$$

such that the transformation of the basis 4-vectors is given by

$$e_\nu = \frac{\partial x'^\mu}{\partial x^\nu} e'_\mu.$$

Then we conclude from the definition (11) that the contravariant components of a tensor of rank 2 transform as

$$T^{\mu\nu} \longrightarrow T'^{\mu\nu} = T^{\rho\sigma} \frac{\partial x'^\mu}{\partial x^\rho} \frac{\partial x'^\nu}{\partial x^\sigma}. \tag{13}$$

Finally, the contravariant components of a tensor of rank r obviously transform as

$$T^{\mu_1 \cdots \mu_r} \longrightarrow T'^{\mu_1 \cdots \mu_r} = T^{\rho_1 \cdots \rho_r} \frac{\partial x'^{\mu_1}}{\partial x^{\rho_1}} \cdots \frac{\partial x'^{\mu_r}}{\partial x^{\rho_r}},$$

and we recover the fact that a tensor of rank 1 (4-vector) transforms as Eq.(12).

We can check here that the metric tensor $\eta_{\mu\nu}$ satisfies the transformation law (13): if x is a 4-vector, its square does not depend on the inertial frame and thus

$$x'^2 = x'^\mu x'^\nu \eta_{\mu\nu} = x^\rho x^\sigma \frac{\partial x'^\mu}{\partial x^\rho} \frac{\partial x'^\nu}{\partial x^\sigma} \eta_{\mu\nu}.$$

But we know that

$$x'^2 = x^2 = x^\rho x^\sigma \eta_{\rho\sigma},$$

such that we must have

$$\eta_{\mu\nu} \frac{\partial x'^\mu}{\partial x^\rho} \frac{\partial x'^\nu}{\partial x^\sigma} = \eta_{\rho\sigma} = \eta'_{\rho\sigma},$$

since the metric tensor is the same in all the inertial frames, what gives us the expected result. We note that $\eta^\mu{}_\nu = \delta^\mu_\nu$.

We can also check that the 4-vector ∂^μ satisfies the expected transformation law (12). The chain rule tells us that

$$\partial'^\mu = \frac{\partial}{\partial x'_\mu} = \frac{\partial x_\rho}{\partial x'_\mu} \frac{\partial}{\partial x_\rho} = \frac{\partial x_\rho}{\partial x'_\mu} \partial^\rho$$

To see that this is consistent with the expected law of transformation of a 4-vector, we can take a derivative of the identity $x'^2 = x^2$ with respect to x'_μ to see that

$$x'^\mu = \frac{\partial x_\rho}{\partial x'_\mu} x^\rho,$$

so as, together with Eq.(12), we find

$$\frac{\partial x_\rho}{\partial x'_\mu} = \frac{\partial x'^\mu}{\partial x^\rho}$$

and obtain that ∂_μ transforms as expected:

$$\partial'^\mu = \frac{\partial x'^\mu}{\partial x^\rho} \partial^\rho.$$

We defined the scalar product for 4-vectors and we can now extend this notion to tensors. We call *contraction* of two indices the following operation. Let $T_{(2)}$ and $K_{(2)}$ be two tensors of rank 2. Starting from their contravariant components, we can construct

$$R^{\mu\nu} = T^{\mu\rho}K^{\sigma\nu}\eta_{\rho\sigma} = T^{\mu\rho}K_{\rho}{}^{\nu} = T^{\mu}{}_{\rho}K^{\rho\nu},$$

which are the contravariant components of the new tensor R of rank 2. The contraction is Lorentz invariant. This operation can be extended to any tensor of any rank and can also be be done within the set of indices of a given tensor: for example, starting from a tensor $T_{(4)}$ of rank 4, we can construct

$$P^{\mu\nu} = T^{\mu\nu\rho\sigma}\eta_{\rho\sigma} = T^{\mu\nu\rho}{}_{\rho} = T^{\mu\nu}{}_{\rho}{}^{\rho},$$

which are the components of a tensor $P_{(2)}$ of rank 2. The trace of a tensor of rank 2 is defined as

$$\text{tr } T_{(2)} = T^{\mu\nu}\eta_{\mu\nu} = T^{\mu}{}_{\mu} = T_{\mu}{}^{\mu},$$

and is a Lorentz scalar. For example, the trace of the metric tensor is

$$\text{tr } \eta = \eta^{\mu\nu}\eta_{\mu\nu} = \delta^{\mu}_{\mu} = 4.$$

1.7 Maxwell's equations in tensor notations

We come back to the origin of special relativity and claim that Maxwell's equation should be written in a compact way using tensors, since they are invariant under a change of inertial frame. In what follows, we will take the constants $\mu_0 = \varepsilon_0 = c = 1$.

Let us look at the electric field

$$\vec{E} = -\frac{\partial \vec{A}}{\partial t} - \vec{\nabla} V,$$

whose components can be written

$$\begin{aligned} E^k &= -\frac{\partial A^k}{\partial x^0} - \frac{\partial V}{\partial x^k} \\ &= -\partial_0 A^k - \partial_k V \\ &= \partial^0 A^k - \partial^k A^0. \end{aligned}$$

where we defined $A^0 = V$. The magnetic field

$$\vec{B} = \vec{\nabla} \times \vec{A}$$

has the components

$$\begin{aligned} B^k &= \varepsilon^{kij}\partial_i A_j \\ &= \partial^i A^j - \partial^j A^i, \end{aligned} \tag{14}$$

where ε^{kij} is the totally antisymmetric tensor ($\varepsilon^{123} = 1$ and $\varepsilon^{132} = -1$ as well as the circular permutations, and the other components vanish).

Thus we are led to define the quantities

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu,$$

and have $E^k = F^{0k}$ and $B^k = F^{ij}$ (where i, j and k are such that $\varepsilon^{kij} = 1$).

Let us now consider the Maxwell-Gauss equation $\vec{\nabla}.\vec{E} = 4\pi\rho$. The latter can be written

$$\frac{\partial E^k}{\partial x^k} = \partial_k F^{0k} = 4\pi j^0, \tag{15}$$

where we have defined the 4-vector j^μ in section 5. Since $F^{00} = 0$, Eq.(15) also reads

$$\partial_\nu F^{0\nu} = 4\pi j^0. \tag{16}$$

The Maxwell-Ampère equation $\vec{\nabla} \times \vec{B} = 4\pi\vec{j} + \partial\vec{E}/\partial t$ can be written

$$\varepsilon^{kij}\partial_i B_j = \varepsilon^{kij}\varepsilon_{jpq}\partial_i\partial^p A^q = 4\pi j^k + \partial_0 E^k. \tag{17}$$

It is easy to see that $\varepsilon^{kij}\varepsilon_{jpq} = \delta^k_p\delta^i_q - \delta^k_q\delta^i_p$, such that Eq.(17) can be written

$$\partial_i\left(\partial^k A^i - \partial^i A^k\right) = \partial_i F^{ki} = 4\pi j^k + \partial_0 F^{0k},$$

or

$$\partial_\nu F^{k\nu} = 4\pi j^k. \tag{18}$$

Finally, Eqs.(16) and (18) can be put together to read

$$\partial_\nu F^{\mu\nu} = 4\pi j^\mu.$$

which is the final form of these two Maxwell's equations. The other two equations are a consequence of the definition of $F^{\mu\nu}$. We now see that $F^{\mu\nu}$ must be the components of a tensor, since its contraction with ∂_ν gives the 4-vector j^μ. We call it the field strength tensor. We also find that A^μ are the components of a 4-vector (the 4-potential), since $F^{\mu\nu}$ is composed out of ∂^μ and A^ν.

1.8 Exercises

Exercise 1 Let p^μ be the components of a 4-vector and define the tensors L and T with components

$$L^{\mu\nu} = \frac{p^\mu p^\nu}{p^2} \quad \text{and} \quad T^{\mu\nu} = \eta^{\mu\nu} - L^{\mu\nu}.$$

Compute the components $(Tp)^\mu$ and $(Lp)^\mu$ of the 4-vectors Tp and Lp and the components $(T^2)^{\mu\nu}$, $(L^2)^{\mu\nu}$ and $(TL)^{\mu\nu}$ of the tensors T^2, L^2 and TL. Conclude on the nature of T and L. Show that the inverse of $T + \alpha L$ is $T + \alpha^{-1}L$, where α is any non-zero real number.

Exercise 2 Let $A^\mu(x)$ be the components of the 4-potential depending on the 4-position x^μ and define the field strength tensor $F(x)$ with components $F^{\mu\nu}(x) = \partial^\mu A^\nu(x) - \partial^\nu A^\mu(x)$. Show with a Fourier transform that (the integrals are taken over the space-time)

$$\int d^4x F^{\mu\nu}(x)F_{\mu\nu}(x) = 2\int \frac{d^4p}{(2\pi)^4} p^2 T^{\mu\nu}(p) A_\mu(p) A_\nu(-p),$$

where $p^2T^{\mu\nu}(p) = p^2\eta^{\mu\nu} - p^\mu p^\nu$ and the Fourier transform of a function ϕ is

$$\phi(x) = \int \frac{d^4p}{(2\pi)^4} \phi(p) e^{ip^\mu x_\mu},$$

such that the Dirac distribution in momentum space is defined by

$$(2\pi)^4\delta^{(4)}(p) = (2\pi)^4\delta(p^0)\delta(p^1)\delta(p^2)\delta(p^3) = \int d^4x e^{ip^\mu x_\mu}.$$

Exercise 3 Let f be any function such that the integral

$$I^{\mu\nu} = \int d^4x \; x^\mu x^\nu f(x^2)$$

is defined (the domain of integration is the whole space-time). Show that

$$I^{\mu\nu} = \frac{1}{4}\eta^{\mu\nu} \int d^4x \; x^2 f(x^2).$$

2 Tensor calculus in General Relativity

2.1 General setting

For reasons explained in the lesson on General Relativity, we need to describe gravitation in a curved space-time. In Special Relativity, we imposed the transformation laws to go from an inertial frame to another to be uniform in space and time, leading to linear transformations of coordinates and thus of tensors. In a curved space-time, the tensors (which are still defined as elements of a vector space) will be defined at a given point in the *tangent* space-time, which is indeed a vector space.

Let us put this in a more quantitative way. We define first a manifold $\mathcal{S}$ of dimension n as a topological space which can be decomposed as a union of open sets that can each be mapped onto an open set of the vector space $\mathcal{R}^n$, in a differentiable way. Therefore a manifold is a 'smoothly curved surface' that *locally* looks like a 'flat surface'.

A sphere is an example of manifold of dimension 2, but a cube is not: any open set including an edge or a corner cannot be mapped onto an open set of $\mathcal{R}^2$ in a differentiable way.

Suppose now that the space-time is a manifold $\mathcal{S}$ of dimension 4, labeled by the coordinates x^μ. Let M be a point of $\mathcal{S}$ and consider the tangent vector space S of $\mathcal{S}$ at the point M, spanned by the 4-vectors

$$e_\mu = \frac{\partial M}{\partial x^\mu}.$$

The Strong Equivalence Principle ensures that this flat tangent space-time exists. Once we have S, we can define the 4-vectors p and q and their scalar product $p.q$ with the properties that we had in Special Relativity and:

$$p.q = p^\mu q^\nu e_\mu.e_\nu = p^\mu q^\nu g_{\mu\nu} = p^\mu q^\nu g_{\nu\mu},$$

where $g_{\nu\mu}$ are the components of the metric tensor, depending on the coordinates of M. We impose this scalar product to be independent of the parametrization of space-time, i.e. of the frame where it is computed. It was already the case in Special Relativity for uniform (linear) transformations and is now valid for any differentiable transformation of coordinates.

We define the covariant components as $x_\mu = x^\nu g_{\mu\nu}$ and we define also $g^{\mu\nu}$ as the components of the inverse metric tensor, such that

$$g^{\mu\rho} g_{\rho\nu} = \delta^\mu_\nu.$$

We stress again that unlike the situation of Special Relativity, the metric tensor depends here on the coordinates of the point M that we consider in the curved space-time.

Let us look now at the transformation of a tensor under a change of frame. If we take other coordinates on the curved space-time, we can define another basis of 4-vectors in the tangent space as

$$e'_\mu = \frac{\partial M}{\partial x'^\mu} = \frac{\partial M}{\partial x^\rho}\frac{\partial x^\rho}{\partial x'^\mu} = \frac{\partial x^\rho}{\partial x'^\mu} e_\rho, \tag{19}$$

such that, with similar arguments that were made in Special Relativity, the contravariant components of a tensor of rank r transform as

$$T'^{\mu_1 \dots \mu_r} = T^{\rho_1 \dots \rho_r} \frac{\partial x'^{\mu_1}}{\partial x^{\rho_1}} \dots \frac{\partial x'^{\mu_r}}{\partial x^{\rho_r}}, \tag{20}$$

with the coefficients $\partial x'^{\mu_i}/\partial x^{\rho_i}$ being functions of coordinates, and not constants as they were in flat space-time.

2.2 Curvilinear coordinates

Before going to a more specific study of curved space-times, let us derive some important properties of curvilinear coordinates, valid in curved or flat spaces.

Lengths

Once we have defined the scalar product, the infinitesimal distance ds which separates two points x and $x + dx$ is by definition given by

$$ds^2 = dx^\mu dx_\mu = g_{\mu\nu} dx^\mu dx^\nu,$$

such that the length l of a curve parametrized by $x^\mu(s)$ between s_1 and s_2 is

$$l = \int_{s_1}^{s_2} ds \sqrt{g_{\mu\nu}(x(s)) \frac{dx^\mu(s)}{ds} \frac{dx^\nu(s)}{ds}}$$

This length is of course independent of the choice of coordinates.

Volumes

Let us first consider a vector space of dimension n and a set of n independent vectors $(\vec{e}_1, ..., \vec{e}_n)$ which form a parallelepipedic area $\mathcal{P}$ and form also a basis of the vector space, with coordinates $(x^1, ..., x^n)$. If $(\vec{\imath}_1, ..., \vec{\imath}_n)$ is an orthonormal basis with coordinates $(\xi^1, ..., \xi^n)$, the volume of $\mathcal{P}$ is

$$V = \int_{\mathcal{P}} d\xi^1 \dots d\xi^n.$$

We now make a change of variables from $(\xi^1, ..., \xi^n)$ to $(x^1, ..., x^n)$ and obtain

$$V = \int_0^1 \left| \det \left(\frac{\partial \xi^k}{\partial x^l} \right) \right| dx^1 ... dx^n,$$

where the determinant contains the coordinates $e^k_l = \vec{e}_l . \vec{\imath}_k$ which are constants and therefore

$$V = \left| \det \left(\frac{\partial \xi^k}{\partial x^l} \right) \right| = \begin{vmatrix} e^1_1 & ... & e^n_1 \\ e^1_2 & ... & e^n_2 \\ ... & ... & ... \\ e^1_n & ... & e^n_n \end{vmatrix}.$$

It is more interesting to give for V a component-independent expression:

$$V^2 = |\det A|^2 = |\det A||\det A^\dagger| = |\det AA^\dagger| = \begin{vmatrix} \vec{e}_1.\vec{e}_1 & ... & \vec{e}_1.\vec{e}_n \\ \vec{e}_2.\vec{e}_1 & ... & \vec{e}_2.\vec{e}_n \\ ... & ... & ... \\ \vec{e}_n.\vec{e}_1 & ... & \vec{e}_n.\vec{e}_n \end{vmatrix}. \quad (21)$$

We now come to curvilinear coordinates where an infinitesimal volume dV can be considered as a parallelepipedic area spanned by the vectors $(dx^1\vec{e}_1, ..., dx^n\vec{e}_n)$. We have then with the multilinearity property of the determinants

$$(dV)^2 = (dx^1 ... dx^n)^2 |\det(e_\mu . e_\nu)|,$$

such that

$$dV = \sqrt{|g|} dx^1 ... dx^n, \quad (22)$$

where g is the determinant of the metric tensor. Eq.(22) is the final expression of the infinitesimal element of volume integration in curvilinear coordinates.

Example: spherical coordinates

Let us take the example of spherical coordinates in 3-dimensional (flat) space. The curvilinear coordinates (r, θ, ϕ) are defined from the euclidean coordinates (x, y, z) by

$$\begin{aligned} x &= r \sin\theta \cos\phi \\ y &= r \sin\theta \sin\phi \\ z &= r \cos\theta. \end{aligned}$$

The basis vectors $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ corresponding to the system of curvilinear coordinates are defined as in Eq.(19) from the orthonormal basis vectors $(\vec{\imath}, \vec{\jmath}, \vec{k})$ by

$$
\begin{aligned}
\vec{e}_1 &= \frac{\partial x}{\partial r}\vec{\imath} + \frac{\partial y}{\partial r}\vec{\jmath} + \frac{\partial z}{\partial r}\vec{k} = \vec{e}_r \\
\vec{e}_2 &= \frac{\partial x}{\partial \phi}\vec{\imath} + \frac{\partial y}{\partial \phi}\vec{\jmath} + \frac{\partial z}{\partial \phi}\vec{k} = r\sin\theta\vec{e}_\phi \\
\vec{e}_3 &= \frac{\partial x}{\partial \theta}\vec{\imath} + \frac{\partial y}{\partial \theta}\vec{\jmath} + \frac{\partial z}{\partial \theta}\vec{k} = r\vec{e}_\theta
\end{aligned}
$$

where $(\vec{e}_r, \vec{e}_\phi, \vec{e}_\theta)$ form an orthonormal basis. The metric tensor is thus

$$
g_{kl} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2\sin^2\theta & 0 \\ 0 & 0 & r^2 \end{pmatrix},
$$

and its inverse

$$
g^{kl} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2\sin^2} & 0 \\ 0 & 0 & \frac{1}{r^2} \end{pmatrix}. \tag{23}
$$

From this result, we find that the determinant of the metric tensor is $g = r^4\sin^2\theta$ and thus the infinitesimal volume is $dV = r^2\sin\theta dr d\theta d\phi$, as well known.

We can also compute the gradient of a scalar function f in spherical coordinates. The following definition of the gradient is independent of the basis:

$$
\vec{\nabla} f = g^{kl}\partial_k f \vec{e}_l,
$$

and we find from the inverse metric (23) that in spherical coordinates it reads

$$
\begin{aligned}
\vec{\nabla} f &= g^{11}\frac{\partial f}{\partial r}\vec{e}_1 + g^{22}\frac{\partial f}{\partial \phi}\vec{e}_2 + g^{33}\frac{\partial f}{\partial \theta}\vec{e}_3 \\
&= \frac{\partial f}{\partial r}\vec{e}_r + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}\vec{e}_\phi + \frac{1}{r}\frac{\partial f}{\partial \theta}\vec{e}_\theta,
\end{aligned}
$$

as also well-known.

2.3 Covariant derivatives

Since the basis 4-vectors e_μ depend on space-time coordinates, a 4-vector $v = v^\nu e_\nu$ will depend on space-time coordinates via its components v^ν *and* via e_ν. We define then the covariant derivative ∇_μ of v^ν as

$$
dv = \nabla_\mu v^\nu dx^\mu e_\nu. \tag{24}
$$

To compute this covariant derivative, we first define the *Christoffel symbol* $\Gamma^\rho{}_{\nu\mu}$ by

$$de_\mu = \partial_\nu e_\mu dx^\nu = \Gamma^\rho{}_{\nu\mu}\ e_\rho dx^\nu \tag{25}$$

and write then

$$\begin{aligned} dv &= d(v^\nu e_\nu) = dv^\nu e_\nu + v^\nu de_\nu \\ &= \frac{\partial v^\nu}{\partial x^\mu} dx^\mu e_\nu + \Gamma^\rho{}_{\mu\nu}\ v^\nu dx^\mu e_\rho, \end{aligned}$$

to conclude that

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu{}_{\mu\rho}\ v^\rho. \tag{26}$$

The Christoffel symbols contain the information about the space-time geometry and its properties will be seen in the next section. We can say here that in a flat space-time we can always find a coordinate system where all the $\Gamma^\nu{}_{\mu\rho}$ vanish (this can be seen as a definition of flat space-time).

We can then derive the covariant derivatives of the contravariant components of a tensor. Let us consider a tensor of rank 2 for the sake of clarity. We have

$$\begin{aligned} dT_{(2)} &= dT^{\mu\nu} e_\mu \otimes e_\nu + T^{\mu\nu}\left[de_\mu \otimes e_\nu + e_\mu \otimes de_\nu\right] \\ &= \partial_\rho T^{\mu\nu} dx^\rho e_\mu \otimes e_\nu + T^{\mu\nu}\Gamma^\sigma{}_{\rho\mu} e_\sigma \otimes e_\nu dx^\rho + T^{\mu\nu}\Gamma^\sigma{}_{\rho\nu} e_\mu \otimes e_\sigma dx^\rho \\ &= \left[\partial_\rho T^{\mu\nu} + T^{\sigma\nu}\Gamma^\mu{}_{\rho\sigma} + T^{\mu\sigma}\Gamma^\nu{}_{\rho\sigma}\right] e_\mu \otimes e_\nu dx^\rho, \end{aligned}$$

such that finally

$$\nabla_\rho T^{\mu\nu} = \partial_\rho T^{\mu\nu} + \Gamma^\mu{}_{\rho\sigma}\ T^{\sigma\nu} + \Gamma^\nu{}_{\rho\sigma}\ T^{\mu\sigma}. \tag{27}$$

For a tensor of rank r, we will have a contribution $\Gamma^{\mu_i}_{\rho\sigma}$ for every indice μ_i with $i = 1, .., r$.

We note that the covariant derivative of a scalar is its usual derivative, since by definition a scalar does not depend on the frame (the basis 4-vectors e_μ do not appear in its definition).

We can also define the covariant derivatives of the covariant components of a 4-vector. Let us consider for this a constant 4-vector w, i.e. with vanishing covariant derivative: $\nabla_\mu w^\nu = 0$, such that $\partial_\mu w^\nu = -\Gamma^\nu{}_{\mu\rho}\ w^\rho$. The scalar product with a 4-vector v is then

$$d(v.w) = \partial_\mu(v.w)dx^\mu = w^\nu\left(\partial_\mu v_\nu - \Gamma^\rho{}_{\mu\nu}\ v_\rho\right)dx^\mu.$$

But we also have $d(v.w) = w.dv = w^\nu \nabla_\mu v_\nu dx^\mu$, so that

$$\nabla_\mu v_\nu = \partial_\mu v_\nu - \Gamma^\rho{}_{\mu\nu}\ v_\rho$$

We can show in a similarily way that for a tensor of rank 2

$$\nabla_\rho T_{\mu\nu} = \partial_\rho T_{\mu\nu} - \Gamma^\sigma{}_{\rho\mu}\, T_{\sigma\nu} - \Gamma^\sigma{}_{\rho\nu}\, T_{\mu\sigma}. \tag{28}$$

To conclude this section, let us look at the covariant derivatives of the metric tensor. From Eq.(28) and the definition of $g_{\mu\nu}$ we have

$$\begin{aligned}
\nabla_\rho g_{\mu\nu} &= \partial_\rho g_{\mu\nu} - \Gamma^\sigma{}_{\rho\mu}\, g_{\sigma\nu} - \Gamma^\sigma{}_{\rho\nu}\, g_{\mu\sigma} \\
&= \partial_\rho (e_\mu . e_\nu) - \Gamma^\sigma{}_{\rho\mu}\, g_{\sigma\nu} - \Gamma^\sigma{}_{\rho\nu}\, g_{\mu\sigma} \\
&= \Gamma^\sigma{}_{\rho\mu}\, e_\sigma . e_\nu + \Gamma^\sigma{}_{\rho\nu}\, e_\mu . e_\sigma - \Gamma^\sigma{}_{\rho\mu}\, g_{\sigma\nu} - \Gamma^\sigma{}_{\rho\nu}\, g_{\mu\sigma},
\end{aligned}$$

such that finally

$$\nabla_\rho g_{\mu\nu} = 0.$$

This important property is the analogue of $\partial_\rho \eta_{\mu\nu} = 0$ in flat space-time. We can show now the similar property for the contravariant components $g^{\mu\nu}$. For a general matrix A depending on a parameter t, we have by definition $AA^{-1} = \mathbf{1}$, where $\mathbf{1}$ is the unity matrix. Taking a derivative with respect to t gives $(dA/dt)A^{-1} + A(dA^{-1}/dt) = 0$ such that

$$\frac{dA^{-1}}{dt} = -A^{-1}\frac{dA}{dt}A^{-1}.$$

Applying this result to $g^{\mu\nu}$ we obtain

$$\partial_\rho g^{\mu\nu} = -g^{\mu\sigma}(\partial_\rho g_{\sigma\lambda})g^{\lambda\nu}.$$

Thus we can write with the help of Eq.(27)

$$\begin{aligned}
\nabla_\rho g^{\mu\nu} &= -g^{\mu\sigma}(\partial_\rho g_{\sigma\lambda})g^{\lambda\nu} + g^{\sigma\nu}\Gamma^\mu{}_{\rho\sigma} + g^{\mu\sigma}\Gamma^\nu{}_{\rho\sigma} \\
&= -g^{\mu\sigma}\left[\partial_\rho g_{\sigma\lambda} - g_{\sigma\delta}\Gamma^\delta{}_{\rho\lambda} - g_{\delta\lambda}\Gamma^\delta{}_{\rho\sigma}\right] g^{\lambda\nu} \\
&= -g^{\mu\sigma}(\nabla_\rho g_{\sigma\lambda})g^{\lambda\nu} \\
&= 0
\end{aligned}$$

2.4 Properties of the Cristoffel symbol

An important property of the Cristoffel symbol deals with its symmetry under the exchange of its two lower indices. To see this, we use the commutativity of the derivatives and write

$$\partial_\mu \partial_\nu M = \partial_\nu \partial_\mu M,$$

so that $\partial_\mu e_\nu = \partial_\nu e_\mu$. Then from the definition (25) we conclude that

$$\Gamma^\rho{}_{\mu\nu} = \Gamma^\rho{}_{\nu\mu}. \tag{29}$$

Let us now give the expression of the Cristoffel symbols in terms of the metric tensor. We have

$$dg_{\mu\nu} = de_\mu.e_\nu + e_\mu.de_\nu = \Gamma^\rho{}_{\sigma\mu}\ dx^\sigma e_\rho.e_\nu + \Gamma^\rho{}_{\sigma\nu}\ dx^\sigma e_\mu.e_\rho,$$

which implies

$$\partial_\sigma g_{\mu\nu} = \Gamma^\rho{}_{\sigma\mu}\ g_{\rho\nu} + \Gamma^\rho{}_{\sigma\nu}\ g_{\mu\rho}.$$

We define then $\Gamma_{\sigma\mu\nu} = \Gamma^\rho{}_{\mu\nu}\ g_{\rho\sigma}$ and can write

$$\partial_\sigma g_{\mu\nu} = \Gamma_{\mu\sigma\nu} + \Gamma_{\nu\sigma\mu}. \tag{30}$$

If we make a circular permutation of the indices, we obtain the similar equations

$$\begin{aligned} \partial_\nu g_{\sigma\mu} &= \Gamma_{\sigma\nu\mu} + \Gamma_{\mu\nu\sigma} \\ \partial_\mu g_{\nu\sigma} &= \Gamma_{\nu\mu\sigma} + \Gamma_{\sigma\mu\nu}. \end{aligned} \tag{31}$$

Using then the symmetry property (29) we obtain from Eqs.(30) and (31)

$$\Gamma_{\sigma\mu\nu} = \frac{1}{2}\left(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}\right) \tag{32}$$

Finally it is important to stress that $\Gamma_{\mu\sigma\nu}$ are *not* the components of a tensor, what can be checked by making a change of basis on the metric tensor. This comes from the fact that the coefficients $\partial x'^\mu/\partial x^\rho$ appearing in the transformation law (20) are not constants and thus will give an additional contribution in the derivatives of (32).

2.5 Parallel transport

We wish to generalize the notion of prallelism to curved space-time. Let $\mathcal{C}$ be a curve parametrized by s. The tangent 4-vector t of $\mathcal{C}$ is defined at each point by

$$t^\mu = \frac{dx^\mu(s)}{ds}. \tag{33}$$

We say that a given 4-vector v depending on space-time coordinates is parallely transported along $\mathcal{C}$ if each of its component v^ν is a constant along $\mathcal{C}$, i.e. if its (covariant) gradient is always perpendicular to the tangent t: $t^\mu \nabla_\mu v^\nu = 0$ (remember that the gradient of a function ϕ is perpendicular to the surfaces ϕ =constant). With the definition of t and the expression of the covariant derivative, the condition of parallel transport also reads

$$\frac{dv^\nu}{ds} + \Gamma^\nu{}_{\mu\rho}\; t^\mu v^\rho = 0,$$

since we have $t^\mu \partial_\mu = d/ds$.

We can define a geodesic as a curve whose tangent is parallely transported, such that for any ν we have $t^\mu \nabla_\mu t^\nu = 0$. With the definition (33) of the tangent, the equation of a geodesic is then given by the equation

$$\frac{d^2 x^\nu}{ds^2} + \Gamma^\nu{}_{\mu\rho}\; \frac{dx^\mu}{ds}\frac{dx^\rho}{ds} = 0. \tag{34}$$

We recognize in Eq.(34) the equation of the shortest paths in a curved space, as can be found in the course of General Relativity. Thus the tangent of the trajectory of a ray of light (which is a geodesic) is parallely transported.

2.6 Curvature tensor

To conclude with this introduction to tensor calculus in curved space-time, we introduce the curvature tensor which plays a fundamental role in General Relativity since it appears in the Einstein equation which describes the interplay between space-time geometry and matter distribution.

In a flat space, if a vector is parallely transported along a closed curve, it will come back to its initial configuration at the end of the transport. In a curved space it will not, and the 'curvature tensor' describes this change of configuration.

Consider a 4-vector parallely transported along a closed curve $\mathcal{C}$ of tangent $t^\nu = dx^\nu/ds$, such that

$$dx^\nu \nabla_\nu A^\mu = dx^\nu \partial_\nu A^\mu + dx^\nu \Gamma^\mu{}_{\nu\rho}\; A^\rho = 0 \tag{35}$$

Its variation along $\mathcal{C}$ (circulation of its derivative) is then

$$\Delta A^\mu = \oint_{\mathcal{C}} dx^\nu \partial_\nu A^\mu = -\oint_{\mathcal{C}} dx^\nu \Gamma^\mu{}_{\nu\rho}\; A^\rho \tag{36}$$

We use then the generalization of Stokes' theorem to 4 dimensions and write the circulation (36) as a flux across a surface $\mathcal{S}$ limited by $\mathcal{C}$:

$$\Delta A^\mu = -\int_{\mathcal{S}} dS^{\sigma\lambda} \left[\partial_\sigma \left(\Gamma^\mu{}_{\lambda\rho}\; A^\rho\right) - \partial_\lambda \left(\Gamma^\mu{}_{\sigma\rho}\; A^\rho\right)\right],$$

where $dS^{\sigma\lambda} = dx^\sigma dx'^\lambda - dx^\lambda dx'^\sigma$ is the infinitesimal surface spanned by two infinitesimal 4-vectors dx and dx' in the plane (σ, λ) (the justification if the same as the one which was given in section 2.2: $dS^{\sigma\lambda}$ is a two-dimensionnal 'volume'). Using again Eq.(35), we find

$$\Delta A^\mu = \int_{\mathcal{S}} dS^{\sigma\lambda}\; R^\mu{}_{\rho\sigma\lambda}\; A^\rho,$$

where we define the curvature tensor $R_{(4)}$ of rank 4 with components

$$R^{\mu}{}_{\rho\sigma\lambda} = \partial_\sigma \Gamma^{\mu}{}_{\rho\lambda} - \partial_\lambda \Gamma^{\mu}{}_{\sigma\rho} + \Gamma^{\mu}{}_{\lambda\eta}\Gamma^{\eta}{}_{\sigma\rho} - \Gamma^{\mu}{}_{\sigma\eta}\Gamma^{\eta}{}_{\lambda\rho} \tag{37}$$

Note the following important properties of R. First, R *is* a tensor, whereas the Cristoffel symbols are not. This is because the unwanted terms appearing in $\Gamma^{\mu}{}_{\rho\lambda}$ in a change of frame cancel in the expression (37). Then we have obviously the antisymmetric properties

$$R^{\mu}{}_{\rho\sigma\lambda} = -R^{\mu}{}_{\rho\lambda\sigma} \quad \text{and} \quad R^{\mu}{}_{\rho\sigma\lambda} = -R^{\rho}{}_{\mu\sigma\lambda}$$

We define then the Ricci tensor $R_{(2)}$ which appears in the Einstein equation: this tensor of rank 2 has components

$$R_{\mu\nu} = R^{\rho}{}_{\mu\sigma\nu}\; g_{\rho}{}^{\sigma} = R^{\rho}{}_{\mu\sigma\nu}\; \delta^{\sigma}_{\rho} = R^{\rho}{}_{\mu\rho\nu}$$

Finally, the scalar curvature is defined as the trace of the Ricci tensor:

$$R = R_{\mu\nu}\; g^{\mu\nu} = R_{\mu}{}^{\mu} = R^{\mu}{}_{\mu},$$

and its value is of course independent of the frame in which it is computed (a scalar is by definition independent of the coordinate system and can be seen as a tensor of rank 0).