

Chapter 5

Riemannian Geometry



In the previous chapters, we studied non-gravitational phenomena in inertial reference frames, and often we limited our discussion to Cartesian coordinate systems. Now we want to include gravity, non-inertial reference frames, and general coordinate systems. The aim of this chapter is to introduce some mathematical tools necessary to achieve this goal. We will follow quite a heuristic approach. The term *Riemannian geometry* is used when we deal with a differentiable manifold equipped with a metric tensor (see Appendix C for the definition of the concept of differentiable manifold).

5.1 Motivations

As we will see better in the next chapter, gravity has quite a special property: for the same initial conditions, any test-particle¹ in an external gravitational field follows the same trajectory, regardless of its internal structure and composition. To be more explicit, we can consider the Newtonian case. Newton's Second Law reads $m_i \ddot{\mathbf{x}} = \mathbf{F}$, where m_i is the *inertial mass* of the particle. If \mathbf{F} is the gravitational force on our particle generated by a point-like body with mass M , we have

$$\mathbf{F} = G_N \frac{M m_g}{r^2} \hat{\mathbf{r}}, \quad (5.1)$$

where m_g is the *gravitational mass* of the particle (and we are assuming that $m_g \ll M$). In principle, m_i and m_g may be different, because the former has nothing to do with the gravitational force (it is well defined even in the absence of gravity!) and

¹A test-particle must have a sufficiently small mass, size, etc. such that its mass does not significantly alter the background gravitational field, tidal forces can be ignored, etc.

the second is the “gravitational charge” of the particle. For instance, in the case of an electrostatic field, the force is given by the Coulomb force, which is proportional to the product of the electric charges of the two objects. The electric charge is completely independent of the inertial mass of a body. On the contrary, the ratio between the inertial and the gravitational masses, m_i/m_g , is a constant independent of the particle. This is an experimental result! We can thus choose units in which $m_i = m_g = m$, where m is just the mass of the particle. At this point, Newton’s Second Law reads

$$\ddot{\mathbf{x}} = G_N \frac{M}{r^2} \hat{\mathbf{r}}, \quad (5.2)$$

and the solution is independent of m and the internal structure and composition of the particle: any test-particle follows the same trajectory for the same initial conditions.

The trajectory of a particle can be obtained by minimizing the path length between two events of the spacetime. We can thus think of writing an effective metric such that the equations of motion of the particle take into account the effect of the gravitational field. The example below can better illustrate this point.

In Newtonian mechanics, the Lagrangian of a particle in a gravitational field is $L = T - V$, where T is the particle kinetic energy, $V = m\Phi$ is the gravitational potential energy, and Φ is the gravitational potential; see Sect. 1.8. As seen in Chap. 3, in special relativity, for small velocities T is replaced by Eq. (3.11). The Lagrangian of a non-relativistic particle in a Newtonian gravitational field is thus

$$L = -mc^2 + \frac{1}{2}m\mathbf{v}^2 - m\Phi. \quad (5.3)$$

Since

$$mc^2 \gg \frac{1}{2}m\mathbf{v}^2, \quad -m\Phi, \quad (5.4)$$

we can rewrite Eq. (5.3) as

$$L = -mc\sqrt{c^2 - \mathbf{v}^2 + 2\Phi}, \quad (5.5)$$

and the corresponding action as

$$\begin{aligned} S &= -mc \int \sqrt{\left(1 + \frac{2\Phi}{c^2}\right) c^2 - \dot{x}^2 - \dot{y}^2 - \dot{z}^2} dt \\ &= -mc \int \sqrt{-g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} dt, \end{aligned} \quad (5.6)$$

where we have introduced the metric tensor $g_{\mu\nu}$ defined as

$$\|g_{\mu\nu}\| = \begin{pmatrix} -\left(1 + \frac{2\phi}{c^2}\right) & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.7)$$

If we apply the Least Action Principle to the action in Eq. (5.6), we obtain the geodesic equations for the metric $g_{\mu\nu}$. They are equivalent to the Euler–Lagrange equations for the Lagrangian in (5.3) by construction. So we can describe the gravitational field as a geometrical property of the spacetime.

With this simple example, we see how we can “absorb” the gravitational field into the metric tensor $g_{\mu\nu}$. The particle trajectories provided by the geodesic equations for the metric $g_{\mu\nu}$ are not straight lines, because $g_{\mu\nu}$ takes gravity into account. Note that $g_{\mu\nu}$ cannot be reduced to the Minkowski metric in the whole spacetime with a coordinate transformation and we say that the spacetime is *curved*. On the contrary, if we can recover the Minkowski metric $\eta_{\mu\nu}$ in the whole spacetime with a coordinate transformation, the spacetime is *flat*. In this second case, the reference frame in which the metric is not $\eta_{\mu\nu}$ either employs non-Cartesian coordinates or is a non-inertial reference frame (or both).

5.2 Covariant Derivative

The partial derivative of a scalar is a dual vector and it is easy to see that it transforms as a dual vector under a coordinate transformation

$$\frac{\partial\phi}{\partial x^\mu} \rightarrow \frac{\partial\phi}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial\phi}{\partial x^\nu}. \quad (5.8)$$

The partial derivative of the components of a vector field is not a tensor field. Let V^μ be a vector and $x^\mu \rightarrow x'^\mu$ a coordinate transformation. We have

$$\begin{aligned} \frac{\partial V^\mu}{\partial x^\nu} &\rightarrow \frac{\partial V'^\mu}{\partial x'^\nu} = \frac{\partial x^\rho}{\partial x'^\nu} \frac{\partial}{\partial x^\rho} \left(\frac{\partial x'^\mu}{\partial x^\sigma} V^\sigma \right) \\ &= \frac{\partial x^\rho}{\partial x'^\nu} \frac{\partial x'^\mu}{\partial x^\sigma} \frac{\partial V^\sigma}{\partial x^\rho} + \frac{\partial x^\rho}{\partial x'^\nu} \frac{\partial^2 x'^\mu}{\partial x^\rho \partial x^\sigma} V^\sigma. \end{aligned} \quad (5.9)$$

If the relation between the two coordinate systems is not linear, we have also the second term on the right hand side and we see that $\partial V^\mu / \partial x^\nu$ cannot be a vector. The reason is that dV^μ is the difference between two vectors at different points. With the terminology of Appendix C, vectors at different points belong to different tangent spaces. dV^μ transforms as

$$\begin{aligned}
 dV^\mu &\rightarrow dV'^\mu = V'^\mu(x + dx) - V'^\mu(x) \\
 &= \left(\frac{\partial x'^\mu}{\partial x^\alpha} \right)_{x+dx} V^\alpha(x + dx) - \left(\frac{\partial x'^\mu}{\partial x^\alpha} \right)_x V^\alpha(x). \quad (5.10)
 \end{aligned}$$

If $\partial x'^\mu / \partial x^\alpha$ in front of $V^\alpha(x + dx)$ were the same as that in front of $V^\alpha(x)$, then we would have

$$dV'^\mu = \frac{\partial x'^\mu}{\partial x^\alpha} [V^\alpha(x + dx) - V^\alpha(x)] = \frac{\partial x'^\mu}{\partial x^\alpha} dV^\alpha. \quad (5.11)$$

However, in general this is not the case: $\partial x'^\mu / \partial x^\alpha$ in front of $V^\alpha(x + dx)$ is evaluated at $x + dx$, that in front of $V^\alpha(x)$ is evaluated at x . In this section we want to introduce the concept of covariant derivative, which is the natural generalization of partial derivative in the case of arbitrary coordinates.

5.2.1 Definition

We know that dx^μ is a 4-vector and that the 4-velocity of a particle, $u^\mu = dx^\mu / d\tau$ is a 4-vector too, since $d\tau$ is a scalar. However, we know from Eq. (5.10) that du^μ is not a 4-vector.

In Sect. 1.7, we introduced the geodesic equations. Since $u^\mu = dx^\mu / d\tau$, we can rewrite the geodesic equations as

$$\frac{du^\mu}{d\tau} + \Gamma_{\nu\rho}^\mu u^\rho \frac{dx^\nu}{d\tau} = 0, \quad (5.12)$$

and also as

$$\frac{Du^\mu}{d\tau} = 0, \quad (5.13)$$

where we have defined Du^μ as

$$Du^\mu = du^\mu + \Gamma_{\nu\rho}^\mu u^\rho dx^\nu = \left(\frac{\partial u^\mu}{\partial x^\nu} + \Gamma_{\nu\rho}^\mu u^\rho \right) dx^\nu. \quad (5.14)$$

We will now show that Du^μ is the natural generalization of du^μ for general coordinate systems and that the partial derivative ∂_μ generalizes to the *covariant derivative* ∇_μ . In the case of a 4-vector like u^μ , we have $Du^\mu = (\nabla_\nu u^\mu) dx^\nu$, where ∇_ν is defined as

$$\nabla_\nu u^\mu = \frac{\partial u^\mu}{\partial x^\nu} + \Gamma_{\nu\rho}^\mu u^\rho. \quad (5.15)$$

First, we check that the components of $\nabla_{\nu}u^{\mu}$ transform as a tensor. The first term on the right hand side in Eq. (5.15) transforms as

$$\begin{aligned}\frac{\partial u^{\mu}}{\partial x^{\nu}} &\rightarrow \frac{\partial u'^{\mu}}{\partial x'^{\nu}} = \frac{\partial x^{\alpha}}{\partial x'^{\nu}} \frac{\partial}{\partial x^{\alpha}} \left(\frac{\partial x'^{\mu}}{\partial x^{\beta}} u^{\beta} \right) \\ &= \frac{\partial x^{\alpha}}{\partial x'^{\nu}} \frac{\partial x'^{\mu}}{\partial x^{\beta}} \frac{\partial u^{\beta}}{\partial x^{\alpha}} + \frac{\partial x^{\alpha}}{\partial x'^{\nu}} \frac{\partial^2 x'^{\mu}}{\partial x^{\alpha} \partial x^{\beta}} u^{\beta}.\end{aligned}\quad (5.16)$$

The Christoffel symbols transform as

$$\begin{aligned}\Gamma_{\nu\rho}^{\mu} &\rightarrow \Gamma'_{\nu\rho}{}^{\mu} = \frac{1}{2} g'^{\mu\sigma} \left(\frac{\partial g'_{\sigma\rho}}{\partial x'^{\nu}} + \frac{\partial g'_{\nu\sigma}}{\partial x'^{\rho}} - \frac{\partial g'_{\nu\rho}}{\partial x'^{\sigma}} \right) \\ &= \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial x^{\gamma}}{\partial x'^{\nu}} \frac{\partial}{\partial x^{\gamma}} \left(\frac{\partial x^{\delta}}{\partial x'^{\sigma}} \frac{\partial x^{\epsilon}}{\partial x'^{\rho}} g_{\delta\epsilon} \right) \\ &\quad + \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial x^{\gamma}}{\partial x'^{\rho}} \frac{\partial}{\partial x^{\gamma}} \left(\frac{\partial x^{\delta}}{\partial x'^{\nu}} \frac{\partial x^{\epsilon}}{\partial x'^{\sigma}} g_{\delta\epsilon} \right) \\ &\quad - \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial x^{\gamma}}{\partial x'^{\sigma}} \frac{\partial}{\partial x^{\gamma}} \left(\frac{\partial x^{\delta}}{\partial x'^{\nu}} \frac{\partial x^{\epsilon}}{\partial x'^{\rho}} g_{\delta\epsilon} \right).\end{aligned}\quad (5.17)$$

Since the calculations become long, we consider the three terms on the right hand side in Eq. (5.17) separately. The first term is

$$\begin{aligned}&\frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial x^{\gamma}}{\partial x'^{\nu}} \frac{\partial}{\partial x^{\gamma}} \left(\frac{\partial x^{\delta}}{\partial x'^{\sigma}} \frac{\partial x^{\epsilon}}{\partial x'^{\rho}} g_{\delta\epsilon} \right) \\ &= \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial x^{\gamma}}{\partial x'^{\nu}} \frac{\partial^2 x^{\delta}}{\partial x'^{\tau} \partial x'^{\sigma}} \frac{\partial x^{\tau}}{\partial x^{\gamma}} \frac{\partial x^{\epsilon}}{\partial x'^{\rho}} g_{\delta\epsilon} \\ &\quad + \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial x^{\gamma}}{\partial x'^{\nu}} \frac{\partial x^{\delta}}{\partial x'^{\sigma}} \frac{\partial^2 x^{\epsilon}}{\partial x'^{\tau} \partial x'^{\rho}} \frac{\partial x^{\tau}}{\partial x^{\gamma}} g_{\delta\epsilon} \\ &\quad + \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial x^{\gamma}}{\partial x'^{\nu}} \frac{\partial x^{\delta}}{\partial x'^{\sigma}} \frac{\partial x^{\epsilon}}{\partial x'^{\rho}} \frac{\partial g_{\delta\epsilon}}{\partial x^{\gamma}} \\ &= \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial^2 x^{\delta}}{\partial x'^{\nu} \partial x'^{\sigma}} \frac{\partial x^{\epsilon}}{\partial x'^{\rho}} g_{\delta\epsilon} + \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} g^{\alpha\beta} \frac{\partial^2 x^{\epsilon}}{\partial x'^{\nu} \partial x'^{\rho}} g_{\beta\epsilon} \\ &\quad + \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} g^{\alpha\beta} \frac{\partial x^{\gamma}}{\partial x'^{\nu}} \frac{\partial x^{\epsilon}}{\partial x'^{\rho}} \frac{\partial g_{\beta\epsilon}}{\partial x^{\gamma}}.\end{aligned}\quad (5.18)$$

For the second term we have

$$\begin{aligned}
& \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial x^{\gamma}}{\partial x'^{\rho}} \frac{\partial}{\partial x^{\gamma}} \left(\frac{\partial x^{\delta}}{\partial x'^{\nu}} \frac{\partial x^{\varepsilon}}{\partial x'^{\sigma}} g_{\delta\varepsilon} \right) \\
&= \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} g^{\alpha\beta} \frac{\partial^2 x^{\delta}}{\partial x'^{\nu} \partial x'^{\rho}} g_{\delta\beta} + \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial x^{\delta}}{\partial x'^{\nu}} \frac{\partial^2 x^{\varepsilon}}{\partial x'^{\rho} \partial x'^{\sigma}} g_{\delta\varepsilon} \\
&\quad + \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} g^{\alpha\beta} \frac{\partial x^{\gamma}}{\partial x'^{\rho}} \frac{\partial x^{\delta}}{\partial x'^{\nu}} \frac{\partial g_{\delta\beta}}{\partial x^{\gamma}}.
\end{aligned} \tag{5.19}$$

Lastly, the third term becomes

$$\begin{aligned}
& -\frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial x^{\gamma}}{\partial x'^{\sigma}} \frac{\partial}{\partial x^{\gamma}} \left(\frac{\partial x^{\delta}}{\partial x'^{\nu}} \frac{\partial x^{\varepsilon}}{\partial x'^{\rho}} g_{\delta\varepsilon} \right) \\
&= -\frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial^2 x^{\delta}}{\partial x'^{\nu} \partial x'^{\sigma}} \frac{\partial x^{\varepsilon}}{\partial x'^{\rho}} g_{\delta\varepsilon} - \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\sigma}}{\partial x^{\beta}} g^{\alpha\beta} \frac{\partial x^{\delta}}{\partial x'^{\nu}} \frac{\partial^2 x^{\varepsilon}}{\partial x'^{\rho} \partial x'^{\sigma}} g_{\delta\varepsilon} \\
&\quad - \frac{1}{2} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} g^{\alpha\beta} \frac{\partial x^{\delta}}{\partial x'^{\nu}} \frac{\partial x^{\varepsilon}}{\partial x'^{\rho}} \frac{\partial g_{\delta\varepsilon}}{\partial x^{\beta}}.
\end{aligned} \tag{5.20}$$

If we combine the results in Eqs. (5.18)–(5.20), we find the transformation rule for the Christoffel symbols

$$\begin{aligned}
\Gamma_{\nu\rho}^{\mu} \rightarrow \Gamma'_{\nu\rho}{}^{\mu} &= \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x^{\delta}}{\partial x'^{\nu}} \frac{\partial x^{\varepsilon}}{\partial x'^{\rho}} \frac{1}{2} g^{\alpha\beta} \left(\frac{\partial g_{\beta\varepsilon}}{\partial x^{\delta}} + \frac{\partial g_{\delta\beta}}{\partial x^{\varepsilon}} - \frac{\partial g_{\delta\varepsilon}}{\partial x^{\beta}} \right) \\
&\quad + \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial^2 x^{\delta}}{\partial x'^{\nu} \partial x'^{\rho}} g^{\alpha\beta} g_{\delta\beta} \\
&= \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x^{\delta}}{\partial x'^{\nu}} \frac{\partial x^{\varepsilon}}{\partial x'^{\rho}} \Gamma_{\delta\varepsilon}^{\alpha} + \frac{\partial x^{\alpha}}{\partial x'^{\nu} \partial x'^{\rho}} \frac{\partial x'^{\mu}}{\partial x^{\alpha}}.
\end{aligned} \tag{5.21}$$

We see here that the Christoffel symbols do not transform as the components of a tensor, so they cannot be the components of a tensor.

Let us now combine the results in Eqs. (5.16) and (5.21). We find

$$\begin{aligned}
\frac{\partial u^{\mu}}{\partial x^{\nu}} + \Gamma_{\nu\rho}^{\mu} u^{\rho} \rightarrow \frac{\partial u'^{\mu}}{\partial x'^{\nu}} + \Gamma'_{\nu\rho}{}^{\mu} u'^{\rho} &= \frac{\partial x^{\alpha}}{\partial x'^{\nu}} \frac{\partial x'^{\mu}}{\partial x^{\beta}} \frac{\partial u^{\beta}}{\partial x^{\alpha}} + \frac{\partial x^{\alpha}}{\partial x'^{\nu}} \frac{\partial^2 x'^{\mu}}{\partial x^{\alpha} \partial x^{\beta}} u^{\beta} \\
&\quad + \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x^{\beta}}{\partial x'^{\nu}} \frac{\partial x^{\gamma}}{\partial x'^{\rho}} \Gamma_{\beta\gamma}^{\alpha} \frac{\partial x'^{\rho}}{\partial x^{\delta}} u^{\delta} \\
&\quad + \frac{\partial^2 x^{\alpha}}{\partial x'^{\nu} \partial x'^{\rho}} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\rho}}{\partial x^{\beta}} u^{\beta} \\
&= \frac{\partial x^{\alpha}}{\partial x'^{\nu}} \frac{\partial x'^{\mu}}{\partial x^{\beta}} \frac{\partial u^{\beta}}{\partial x^{\alpha}} + \frac{\partial x^{\alpha}}{\partial x'^{\nu}} \frac{\partial^2 x'^{\mu}}{\partial x^{\alpha} \partial x^{\beta}} u^{\beta} \\
&\quad + \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x^{\beta}}{\partial x'^{\nu}} \Gamma_{\beta\gamma}^{\alpha} u^{\gamma} \\
&\quad + \frac{\partial^2 x^{\alpha}}{\partial x'^{\nu} \partial x'^{\rho}} \frac{\partial x'^{\mu}}{\partial x^{\alpha}} \frac{\partial x'^{\rho}}{\partial x^{\beta}} u^{\beta}
\end{aligned} \tag{5.22}$$

Note that

$$\begin{aligned}\frac{\partial x^\alpha}{\partial x'^\nu} \frac{\partial x'^\mu}{\partial x^\alpha} &= \delta_\nu^\mu, \\ \frac{\partial}{\partial x^\beta} \left(\frac{\partial x^\alpha}{\partial x'^\nu} \frac{\partial x'^\mu}{\partial x^\alpha} \right) &= \frac{\partial}{\partial x^\beta} \delta_\nu^\mu, \\ \frac{\partial^2 x^\alpha}{\partial x'^\nu \partial x'^\rho} \frac{\partial x'^\rho}{\partial x^\beta} \frac{\partial x'^\mu}{\partial x^\alpha} + \frac{\partial x^\alpha}{\partial x'^\nu} \frac{\partial^2 x'^\mu}{\partial x^\beta \partial x^\alpha} &= 0,\end{aligned}\tag{5.23}$$

and therefore

$$\frac{\partial^2 x^\alpha}{\partial x'^\nu \partial x'^\rho} \frac{\partial x'^\rho}{\partial x^\beta} \frac{\partial x'^\mu}{\partial x^\alpha} = - \frac{\partial x^\alpha}{\partial x'^\nu} \frac{\partial^2 x'^\mu}{\partial x^\beta \partial x^\alpha}.\tag{5.24}$$

We use Eq. (5.24) in the last term on the right hand side in Eq. (5.22) and we see that the second and the last terms cancel each other. Equation (5.22) can thus be written as

$$\frac{\partial u'^\mu}{\partial x'^\nu} + \Gamma_{\nu\rho}^{\mu} u'^\rho = \frac{\partial x^\alpha}{\partial x'^\nu} \frac{\partial x'^\mu}{\partial x^\beta} \left(\frac{\partial u^\beta}{\partial x^\alpha} + \Gamma_{\alpha\gamma}^{\beta} u^\gamma \right),\tag{5.25}$$

and we see that $\nabla_\nu u'^\mu$ s transform as the components of a tensor of type (1, 1).

5.2.2 Parallel Transport

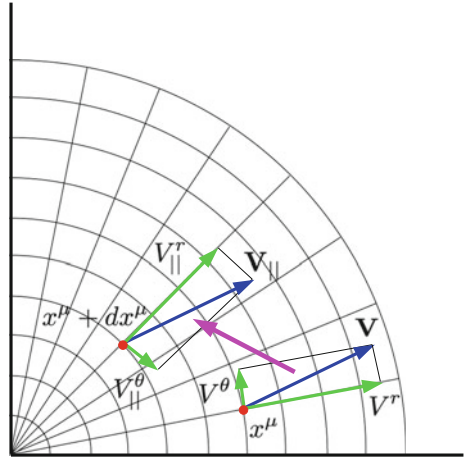
As it was pointed out at the beginning of this section, the partial derivative of a vector computes the difference of two vectors belonging to different points and for this reason the new object is not a tensor. The sum or the difference of two vectors is another vector if the two vectors belong to the same vector space, but this is not the case here. Intuitively, we should “transport” one of the two vectors to the point of the other vector and compute the difference there. This is what the covariant derivative indeed does and involves the concept of *parallel transport*.

Let us consider the example illustrated in Fig. 5.1. We have a 2-dimensional Euclidean space and we consider both Cartesian coordinates (x, y) and polar coordinates (r, θ) . The vector \mathbf{V} is at point $A = x_A = \{x^\mu\}$. Its components are (V^x, V^y) in Cartesian coordinates and (V^r, V^θ) in polar coordinates. If we think of “rigidly” transporting the vector \mathbf{V} from point A to point $B = x_B = \{x^\mu + dx^\mu\}$, as shown in Fig. 5.1, the Cartesian coordinates do not change

$$(V^x, V^y) \rightarrow (V^x, V^y).\tag{5.26}$$

However, the polar coordinates change. This operation is called parallel transport and in what follows we want to show that the components of the parallel transported

Fig. 5.1 The vector \mathbf{V} at x^μ is parallel transported to the point $x^\mu + dx^\mu$. While the vector \mathbf{V} does not change under parallel transport, its components change in general



vector are given by

$$V_{A \rightarrow B}^\mu = V_A^\mu - \Gamma_{\nu\rho}^\mu(x_A) V_A^\nu dx^\rho . \tag{5.27}$$

First, the relations between Cartesian and polar coordinates are

$$x = r \cos \theta , \quad y = r \sin \theta , \tag{5.28}$$

$$r = \sqrt{x^2 + y^2} , \quad \theta = \arctan \frac{y}{x} . \tag{5.29}$$

If the vector \mathbf{V} has Cartesian coordinates (V^x, V^y) , its polar coordinates (V^r, V^θ) are

$$\begin{aligned} V^r &= \frac{\partial r}{\partial x} V^x + \frac{\partial r}{\partial y} V^y = \cos \theta V^x + \sin \theta V^y , \\ V^\theta &= \frac{\partial \theta}{\partial x} V^x + \frac{\partial \theta}{\partial y} V^y = -\frac{\sin \theta}{r} V^x + \frac{\cos \theta}{r} V^y . \end{aligned} \tag{5.30}$$

If we parallel transport the vector \mathbf{V} from the point (r, θ) to the point $(r + dr, \theta + d\theta)$, we have the vector \mathbf{V}_\parallel in Fig. 5.1. The radial coordinate V_\parallel^r is

$$\begin{aligned} V_\parallel^r &= \cos(\theta + d\theta) V^x + \sin(\theta + d\theta) V^y \\ &= (\cos \theta - \sin \theta d\theta) V^x + (\sin \theta + \cos \theta d\theta) V^y . \end{aligned} \tag{5.31}$$

where we have neglected $O(d\theta^2)$ terms

$$\begin{aligned}
\cos(\theta + d\theta) &= \cos\theta \cos d\theta - \sin\theta \sin d\theta \\
&= \cos\theta - \sin\theta d\theta + O(d\theta^2), \\
\sin(\theta + d\theta) &= \sin\theta \cos d\theta + \cos\theta \sin d\theta \\
&= \sin\theta + \cos\theta d\theta + O(d\theta^2).
\end{aligned} \tag{5.32}$$

For the polar coordinate V_{\parallel}^{θ} , we have

$$V_{\parallel}^{\theta} = -\frac{\sin(\theta + d\theta)}{(r + dr)} V^x + \frac{\cos(\theta + d\theta)}{(r + dr)} V^y. \tag{5.33}$$

Since

$$\begin{aligned}
\frac{\sin(\theta + d\theta)}{(r + dr)} &= \frac{\sin\theta}{r} + \frac{\cos\theta}{r} d\theta - \frac{\sin\theta}{r^2} dr + O(dr^2, drd\theta, d\theta^2), \\
\frac{\cos(\theta + d\theta)}{(r + dr)} &= \frac{\cos\theta}{r} - \frac{\sin\theta}{r} d\theta - \frac{\cos\theta}{r^2} dr + O(dr^2, drd\theta, d\theta^2),
\end{aligned} \tag{5.34}$$

Equation (5.33) becomes

$$\begin{aligned}
V_{\parallel}^{\theta} &= \left(-\frac{\sin\theta}{r} - \frac{\cos\theta}{r} d\theta + \frac{\sin\theta}{r^2} dr \right) V^x \\
&\quad + \left(\frac{\cos\theta}{r} - \frac{\sin\theta}{r} d\theta - \frac{\cos\theta}{r^2} dr \right) V^y.
\end{aligned} \tag{5.35}$$

In polar coordinates, the line element reads

$$dl^2 = dr^2 + r^2 d\theta^2. \tag{5.36}$$

As we have seen in Sect. 1.7, the Christoffel symbols can be more quickly calculated from the comparison of the Euler–Lagrange equations for a free particle with the geodesic equations. The Lagrangian to employ is

$$L = \frac{1}{2} (\dot{r}^2 + r^2 \dot{\theta}^2). \tag{5.37}$$

The Euler–Lagrange equation for the Lagrangian coordinate r is

$$\begin{aligned}
\frac{d\dot{r}}{dt} - r\dot{\theta}^2 &= 0, \\
\ddot{r} - r\dot{\theta}^2 &= 0.
\end{aligned} \tag{5.38}$$

For the Lagrangian coordinate θ , we have

$$\begin{aligned}\frac{d}{dt}(r^2\dot{\theta}) &= 0, \\ \ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} &= 0.\end{aligned}\tag{5.39}$$

If we compare Eqs. (5.38) and (5.39) with the geodesic equations, we see that the non-vanishing Christoffel symbols are

$$\Gamma_{\theta\theta}^r = -r, \quad \Gamma_{r\theta}^\theta = \Gamma_{\theta r}^\theta = \frac{1}{r}.\tag{5.40}$$

Now we want to see that, if we parallel transport the vector \mathbf{V} from point A to point B , the coordinates of the vector at B are given by

$$V_{A \rightarrow B}^\mu = V_A^\mu - \Gamma_{\nu\rho}^\mu(x_A)V_A^\nu dx^\rho.\tag{5.41}$$

For the radial coordinate we have

$$\begin{aligned}V_{A \rightarrow B}^r &= V_A^r + rV_A^\theta d\theta \\ &= \cos\theta V^x + \sin\theta V^y + r\left(-\frac{\sin\theta}{r}V^x + \frac{\cos\theta}{r}V^y\right)d\theta \\ &= (\cos\theta - \sin\theta d\theta)V^x + (\sin\theta + \cos\theta d\theta)V^y.\end{aligned}\tag{5.42}$$

For the polar coordinate we have

$$\begin{aligned}V_{A \rightarrow B}^\theta &= V_A^\theta - \frac{1}{r}V_A^r d\theta - \frac{1}{r}V_A^\theta dr \\ &= -\frac{\sin\theta}{r}V^x + \frac{\cos\theta}{r}V^y - \frac{1}{r}(\cos\theta V^x + \sin\theta V^y)d\theta \\ &\quad - \frac{1}{r}\left(-\frac{\sin\theta}{r}V^x + \frac{\cos\theta}{r}V^y\right)dr \\ &= \left(-\frac{\sin\theta}{r} - \frac{\cos\theta}{r}d\theta + \frac{\sin\theta}{r^2}dr\right)V^x \\ &\quad + \left(\frac{\cos\theta}{r} - \frac{\sin\theta}{r}d\theta - \frac{\cos\theta}{r^2}dr\right)V^y.\end{aligned}\tag{5.43}$$

We thus see that we recover the results in Eqs. (5.31) and (5.35).

When we compute the covariant derivative of a vector V^μ we are calculating

$$\begin{aligned}\nabla_\nu V^\mu &= \frac{\partial V^\mu}{\partial x^\nu} + \Gamma_{\nu\rho}^\mu V^\rho \\ &= \lim_{dx^\nu \rightarrow 0} \frac{V^\mu(x + dx) - V^\mu(x) + \Gamma_{\nu\rho}^\mu V^\rho dx^\nu}{dx^\nu}\end{aligned}$$

$$= \lim_{dx^v \rightarrow 0} \frac{V_{x+dx \rightarrow x}^\mu(x) - V^\mu(x)}{dx^v}, \quad (5.44)$$

that is, we calculate the difference between the vector $V^\mu(x + dx)$ parallel transported to x and the vector $V^\mu(x)$. Note that now the sign in front of the Christoffel symbols is plus because we are transporting a vector from $x + dx$ to x , while in Eq. (5.41) we have the opposite case.

5.2.3 Properties of the Covariant Derivative

From the discussion above, it is clear that the covariant derivative of a scalar reduces to the ordinary partial derivative

$$\nabla_\mu \phi = \frac{\partial \phi}{\partial x^\mu}. \quad (5.45)$$

Indeed a scalar is just a number and the parallel transport is trivial: $\phi_{A \rightarrow B} = \phi_A$.

The covariant derivative of a dual vector is given by

$$\nabla_\mu V_\nu = \frac{\partial V_\nu}{\partial x^\mu} - \Gamma_{\mu\nu}^\rho V_\rho. \quad (5.46)$$

Indeed, if we consider any vector V^μ and dual vector W_μ , $V^\mu W_\mu$ is a scalar. Enforcing Leibniz's rule for the covariant derivative, we have

$$\begin{aligned} \nabla_\mu (V^\nu W_\nu) &= (\nabla_\mu V^\nu) W_\nu + V^\nu (\nabla_\mu W_\nu) \\ &= \frac{\partial V^\nu}{\partial x^\mu} W_\nu + \Gamma_{\mu\rho}^\nu V^\rho W_\nu + V^\nu (\nabla_\mu W_\nu). \end{aligned} \quad (5.47)$$

Since ∇_μ becomes ∂_μ for a scalar function

$$\nabla_\mu (V^\nu W_\nu) = \frac{\partial V^\nu}{\partial x^\mu} W_\nu + V^\nu \frac{\partial W_\nu}{\partial x^\mu}, \quad (5.48)$$

and therefore, equating Eqs. (5.47) and (5.48), we must have

$$\nabla_\mu W_\nu = \frac{\partial W_\nu}{\partial x^\mu} - \Gamma_{\mu\nu}^\rho W_\rho. \quad (5.49)$$

The generalization to tensors of any type is straightforward and we have

$$\nabla_\lambda T_{v_1 v_2 \dots v_s}^{\mu_1 \mu_2 \dots \mu_r} = \frac{\partial}{\partial x^\lambda} T_{v_1 v_2 \dots v_s}^{\mu_1 \mu_2 \dots \mu_r}$$

$$\begin{aligned}
& \underbrace{+\Gamma_{\lambda\sigma}^{\mu_1} T_{v_1 v_2 \dots v_s}^{\sigma \mu_2 \dots \mu_r} + \Gamma_{\lambda\sigma}^{\mu_2} T_{v_1 v_2 \dots v_s}^{\mu_1 \sigma \dots \mu_r} + \dots + \Gamma_{\lambda\sigma}^{\mu_r} T_{v_1 v_2 \dots v_s}^{\mu_1 \mu_2 \dots \sigma}}_{r \text{ terms}} \\
& \underbrace{-\Gamma_{\lambda v_1}^{\sigma} T_{\sigma v_2 \dots v_s}^{\mu_1 \mu_2 \dots \mu_r} - \Gamma_{\lambda v_2}^{\sigma} T_{v_1 \sigma \dots v_s}^{\mu_1 \mu_2 \dots \mu_r} - \dots - \Gamma_{\lambda v_s}^{\sigma} T_{v_1 v_2 \dots \sigma}^{\mu_1 \mu_2 \dots \mu_r}}_{s \text{ terms}}
\end{aligned} \tag{5.50}$$

It is worth noting that the covariant derivative of the metric tensor vanishes. If we calculate $\nabla_{\mu} g_{\nu\rho}$, we find

$$\begin{aligned}
\nabla_{\mu} g_{\nu\rho} &= \frac{\partial g_{\nu\rho}}{\partial x^{\mu}} - \Gamma_{\mu\nu}^{\kappa} g_{\kappa\rho} - \Gamma_{\mu\rho}^{\kappa} g_{\nu\kappa} \\
&= \frac{\partial g_{\nu\rho}}{\partial x^{\mu}} - \frac{1}{2} g^{\kappa\lambda} \left(\frac{\partial g_{\lambda\nu}}{\partial x^{\mu}} + \frac{\partial g_{\mu\lambda}}{\partial x^{\nu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\lambda}} \right) g_{\kappa\rho} \\
&\quad - \frac{1}{2} g^{\kappa\lambda} \left(\frac{\partial g_{\lambda\rho}}{\partial x^{\mu}} + \frac{\partial g_{\mu\lambda}}{\partial x^{\rho}} - \frac{\partial g_{\mu\rho}}{\partial x^{\lambda}} \right) g_{\nu\kappa}
\end{aligned} \tag{5.51}$$

Since $g^{\kappa\lambda} g_{\kappa\rho} = \delta_{\rho}^{\lambda}$ and $g^{\kappa\lambda} g_{\nu\kappa} = \delta_{\nu}^{\lambda}$, we obtain

$$\begin{aligned}
\nabla_{\mu} g_{\nu\rho} &= \frac{\partial g_{\nu\rho}}{\partial x^{\mu}} - \frac{1}{2} \frac{\partial g_{\rho\nu}}{\partial x^{\mu}} - \frac{1}{2} \frac{\partial g_{\mu\rho}}{\partial x^{\nu}} + \frac{1}{2} \frac{\partial g_{\mu\nu}}{\partial x^{\rho}} \\
&\quad - \frac{1}{2} \frac{\partial g_{\nu\rho}}{\partial x^{\mu}} - \frac{1}{2} \frac{\partial g_{\mu\nu}}{\partial x^{\rho}} + \frac{1}{2} \frac{\partial g_{\mu\rho}}{\partial x^{\nu}} = 0.
\end{aligned} \tag{5.52}$$

5.3 Useful Expressions

In this section we will derive a number of useful expressions and identities involving the metric tensor, the Christoffel symbols, and the covariant derivative.

We indicate with g the determinant of the metric tensor $g_{\mu\nu}$ and with $\tilde{g}_{\mu\nu}$ the cofactor (μ, ν) . The determinant g is defined as

$$g = \sum_{\nu} g_{\mu\nu} \tilde{g}_{\mu\nu} \quad (\text{no summation over } \mu). \tag{5.53}$$

The cofactor $\tilde{g}_{\mu\nu}$ can be written in terms of the determinant g and of the inverse of the metric tensor as

$$\tilde{g}_{\mu\nu} = g g^{\mu\nu}. \tag{5.54}$$

Indeed, if we have an invertible square matrix A , its inverse is given by

$$A^{-1} = \frac{1}{\det(A)} C^T, \tag{5.55}$$

where C is the cofactor matrix and C^T is the transpose of C (the proof of this formula can be found in a textbook on linear algebra). If we apply Eq. (5.55) to the metric tensor, we recover Eq. (5.54). If we plug Eq. (5.54) into (5.53), we find $g = g$, which confirms that Eq. (5.54) is correct.

With the help of Eq. (5.54) we can write

$$\frac{\partial g}{\partial g_{\mu\nu}} = \tilde{g}^{\mu\nu} = g g^{\mu\nu}. \quad (5.56)$$

as well as

$$\frac{\partial g}{\partial x^\sigma} = \frac{\partial g}{\partial g_{\mu\nu}} \frac{\partial g_{\mu\nu}}{\partial x^\sigma} = g g^{\mu\nu} \frac{\partial g_{\mu\nu}}{\partial x^\sigma}. \quad (5.57)$$

This last expression will be used later.

Let us write the formula for Christoffel symbols

$$\Gamma_{\nu\sigma}^\kappa = \frac{1}{2} g^{\kappa\lambda} \left(\frac{\partial g_{\lambda\sigma}}{\partial x^\nu} + \frac{\partial g_{\nu\lambda}}{\partial x^\sigma} - \frac{\partial g_{\nu\sigma}}{\partial x^\lambda} \right). \quad (5.58)$$

We multiply both sides in Eq. (5.58) by $g_{\kappa\mu}$ and we get

$$g_{\kappa\mu} \Gamma_{\nu\sigma}^\kappa = \frac{1}{2} \left(\frac{\partial g_{\mu\sigma}}{\partial x^\nu} + \frac{\partial g_{\nu\mu}}{\partial x^\sigma} - \frac{\partial g_{\nu\sigma}}{\partial x^\mu} \right). \quad (5.59)$$

We rewrite Eq. (5.59) exchanging the indices μ and ν

$$g_{\kappa\nu} \Gamma_{\mu\sigma}^\kappa = \frac{1}{2} \left(\frac{\partial g_{\nu\sigma}}{\partial x^\mu} + \frac{\partial g_{\mu\nu}}{\partial x^\sigma} - \frac{\partial g_{\mu\sigma}}{\partial x^\nu} \right). \quad (5.60)$$

We sum Eq. (5.59) with (5.60) and we obtain

$$\frac{\partial g_{\mu\nu}}{\partial x^\sigma} = g_{\kappa\mu} \Gamma_{\nu\sigma}^\kappa + g_{\kappa\nu} \Gamma_{\mu\sigma}^\kappa. \quad (5.61)$$

Now we can multiply both sides in Eq. (5.61) by $g g^{\mu\nu}$ and employ Eq. (5.57)

$$\frac{\partial g}{\partial x^\sigma} = g g^{\mu\nu} (g_{\kappa\mu} \Gamma_{\nu\sigma}^\kappa + g_{\kappa\nu} \Gamma_{\mu\sigma}^\kappa) = g (\Gamma_{\sigma\nu}^\nu + \Gamma_{\sigma\mu}^\mu) = 2g \Gamma_{\sigma\mu}^\mu. \quad (5.62)$$

We rewrite Eq. (5.62) as

$$\frac{\partial}{\partial x^\sigma} \ln \sqrt{-g} = \frac{1}{2g} \frac{\partial g}{\partial x^\sigma} = \Gamma_{\sigma\mu}^\mu. \quad (5.63)$$

With the help of Eq. (5.63), the covariant divergence of a generic vector A^μ can be written as

$$\begin{aligned}\nabla_\mu A^\mu &= \frac{\partial A^\mu}{\partial x^\mu} + \Gamma_{\sigma\mu}^\mu A^\sigma = \frac{\partial A^\mu}{\partial x^\mu} + A^\sigma \frac{\partial}{\partial x^\sigma} \ln \sqrt{-g} \\ &= \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\mu} (A^\mu \sqrt{-g}) .\end{aligned}\quad (5.64)$$

In the case of a generic tensor $A^{\mu\nu}$ of type (2, 0), we can write

$$\begin{aligned}\nabla_\mu A^{\mu\nu} &= \frac{\partial A^{\mu\nu}}{\partial x^\mu} + \Gamma_{\sigma\mu}^\mu A^{\sigma\nu} + \Gamma_{\sigma\mu}^\nu A^{\mu\sigma} \\ &= \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\mu} (A^{\mu\nu} \sqrt{-g}) + \Gamma_{\sigma\mu}^\nu A^{\mu\sigma} .\end{aligned}\quad (5.65)$$

Note that, if $A^{\mu\nu}$ is antisymmetric, namely $A^{\mu\nu} = -A^{\nu\mu}$, $\Gamma_{\sigma\mu}^\nu A^{\mu\sigma} = 0$, and Eq. (5.65) simplifies to

$$\nabla_\mu A^{\mu\nu} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\mu} (A^{\mu\nu} \sqrt{-g}) .\quad (5.66)$$

5.4 Riemann Tensor

5.4.1 Definition

We know that if the first partial derivatives are differentiable then the partial derivatives commute, i.e. $\partial_\mu \partial_\nu = \partial_\nu \partial_\mu$; see e.g. [1]. In general, covariant derivatives do not. We can introduce the *Riemann tensor* $R^\lambda_{\rho\nu\mu}$ as the tensor of type (1, 3) defined as

$$\nabla_\mu \nabla_\nu A_\rho - \nabla_\nu \nabla_\mu A_\rho = R^\lambda_{\rho\nu\mu} A_\lambda ,\quad (5.67)$$

where A_μ is a generic dual vector. $R^\lambda_{\rho\nu\mu}$ is a tensor because $\nabla_\mu \nabla_\nu A_\rho$ and $\nabla_\nu \nabla_\mu A_\rho$ are tensors.

In order to find the explicit expression of the Riemann tensor, first we calculate $\nabla_\mu \nabla_\nu A_\rho$

$$\begin{aligned}\nabla_\mu \nabla_\nu A_\rho &= \nabla_\mu \left(\frac{\partial A_\rho}{\partial x^\nu} - \Gamma_{\nu\rho}^\lambda A_\lambda \right) \\ &= \frac{\partial}{\partial x^\mu} \left(\frac{\partial A_\rho}{\partial x^\nu} - \Gamma_{\nu\rho}^\lambda A_\lambda \right) - \Gamma_{\mu\rho}^\kappa \left(\frac{\partial A_\kappa}{\partial x^\nu} - \Gamma_{\kappa\nu}^\lambda A_\lambda \right) - \Gamma_{\mu\nu}^\kappa \left(\frac{\partial A_\rho}{\partial x^\kappa} - \Gamma_{\kappa\rho}^\lambda A_\lambda \right) \\ &= \frac{\partial^2 A_\rho}{\partial x^\mu \partial x^\nu} - \frac{\partial \Gamma_{\nu\rho}^\lambda}{\partial x^\mu} A_\lambda - \Gamma_{\nu\rho}^\lambda \frac{\partial A_\lambda}{\partial x^\mu} - \Gamma_{\mu\rho}^\kappa \frac{\partial A_\kappa}{\partial x^\nu} + \Gamma_{\mu\rho}^\kappa \Gamma_{\kappa\nu}^\lambda A_\lambda\end{aligned}$$

$$- \Gamma_{\mu\nu}^{\kappa} \frac{\partial A_{\rho}}{\partial x^{\kappa}} + \Gamma_{\mu\nu}^{\kappa} \Gamma_{\kappa\rho}^{\lambda} A_{\lambda}. \quad (5.68)$$

The expression for $\nabla_{\nu} \nabla_{\mu} A_{\rho}$ is

$$\begin{aligned} \nabla_{\nu} \nabla_{\mu} A_{\rho} &= \frac{\partial^2 A_{\rho}}{\partial x^{\nu} \partial x^{\mu}} - \frac{\partial \Gamma_{\mu\rho}^{\lambda}}{\partial x^{\nu}} A_{\lambda} - \Gamma_{\mu\rho}^{\lambda} \frac{\partial A_{\lambda}}{\partial x^{\nu}} - \Gamma_{\nu\rho}^{\kappa} \frac{\partial A_{\kappa}}{\partial x^{\mu}} + \Gamma_{\nu\rho}^{\kappa} \Gamma_{\kappa\mu}^{\lambda} A_{\lambda} \\ &\quad - \Gamma_{\nu\mu}^{\kappa} \frac{\partial A_{\rho}}{\partial x^{\kappa}} + \Gamma_{\nu\mu}^{\kappa} \Gamma_{\kappa\rho}^{\lambda} A_{\lambda}. \end{aligned} \quad (5.69)$$

We combine Eqs. (5.68) and (5.69) and we find

$$\nabla_{\mu} \nabla_{\nu} A_{\rho} - \nabla_{\nu} \nabla_{\mu} A_{\rho} = \left(\frac{\partial \Gamma_{\mu\rho}^{\lambda}}{\partial x^{\nu}} - \frac{\partial \Gamma_{\nu\rho}^{\lambda}}{\partial x^{\mu}} + \Gamma_{\mu\rho}^{\kappa} \Gamma_{\kappa\nu}^{\lambda} - \Gamma_{\nu\rho}^{\kappa} \Gamma_{\kappa\mu}^{\lambda} \right) A_{\lambda}. \quad (5.70)$$

We can now write the Riemann tensor $R^{\lambda}_{\rho\nu\mu}$ in terms of the Christoffel symbols as follows

$$R^{\mu}_{\nu\rho\sigma} = \frac{\partial \Gamma_{\nu\sigma}^{\mu}}{\partial x^{\rho}} - \frac{\partial \Gamma_{\nu\rho}^{\mu}}{\partial x^{\sigma}} + \Gamma_{\nu\sigma}^{\lambda} \Gamma_{\rho\lambda}^{\mu} - \Gamma_{\nu\rho}^{\lambda} \Gamma_{\sigma\lambda}^{\mu}. \quad (5.71)$$

It is also useful to have the explicit expression of $R_{\mu\nu\rho\sigma}$. From Eq. (5.71), we lower the upper index with the metric tensor

$$\begin{aligned} R_{\mu\nu\rho\sigma} &= g_{\mu\lambda} R^{\lambda}_{\nu\rho\sigma} \\ &= g_{\mu\lambda} \left(\frac{\partial \Gamma_{\nu\sigma}^{\lambda}}{\partial x^{\rho}} - \frac{\partial \Gamma_{\nu\rho}^{\lambda}}{\partial x^{\sigma}} + \Gamma_{\nu\sigma}^{\kappa} \Gamma_{\rho\kappa}^{\lambda} - \Gamma_{\nu\rho}^{\kappa} \Gamma_{\sigma\kappa}^{\lambda} \right), \end{aligned} \quad (5.72)$$

The first term on the right hand side in Eq. (5.72) can be written as

$$\begin{aligned} g_{\mu\lambda} \frac{\partial \Gamma_{\nu\sigma}^{\lambda}}{\partial x^{\rho}} &= g_{\mu\lambda} \frac{\partial}{\partial x^{\rho}} \left[\frac{1}{2} g^{\lambda\kappa} \left(\frac{\partial g_{\kappa\sigma}}{\partial x^{\nu}} + \frac{\partial g_{\nu\kappa}}{\partial x^{\sigma}} - \frac{\partial g_{\nu\sigma}}{\partial x^{\kappa}} \right) \right] \\ &= \frac{1}{2} g_{\mu\lambda} \frac{\partial g^{\lambda\kappa}}{\partial x^{\rho}} \left(\frac{\partial g_{\kappa\sigma}}{\partial x^{\nu}} + \frac{\partial g_{\nu\kappa}}{\partial x^{\sigma}} - \frac{\partial g_{\nu\sigma}}{\partial x^{\kappa}} \right) \\ &\quad + \frac{1}{2} g_{\mu}^{\kappa} \left(\frac{\partial^2 g_{\kappa\sigma}}{\partial x^{\rho} \partial x^{\nu}} + \frac{\partial^2 g_{\nu\kappa}}{\partial x^{\rho} \partial x^{\sigma}} - \frac{\partial^2 g_{\nu\sigma}}{\partial x^{\rho} \partial x^{\kappa}} \right) \\ &= -\frac{1}{2} \frac{\partial g_{\mu\lambda}}{\partial x^{\rho}} g^{\lambda\kappa} \left(\frac{\partial g_{\kappa\sigma}}{\partial x^{\nu}} + \frac{\partial g_{\nu\kappa}}{\partial x^{\sigma}} - \frac{\partial g_{\nu\sigma}}{\partial x^{\kappa}} \right) \\ &\quad + \frac{1}{2} \left(\frac{\partial^2 g_{\mu\sigma}}{\partial x^{\rho} \partial x^{\nu}} + \frac{\partial^2 g_{\nu\mu}}{\partial x^{\rho} \partial x^{\sigma}} - \frac{\partial^2 g_{\nu\sigma}}{\partial x^{\rho} \partial x^{\mu}} \right) \\ &= -\Gamma_{\nu\sigma}^{\lambda} \frac{\partial g_{\mu\lambda}}{\partial x^{\rho}} + \frac{1}{2} \left(\frac{\partial^2 g_{\mu\sigma}}{\partial x^{\rho} \partial x^{\nu}} + \frac{\partial^2 g_{\nu\mu}}{\partial x^{\rho} \partial x^{\sigma}} - \frac{\partial^2 g_{\nu\sigma}}{\partial x^{\rho} \partial x^{\mu}} \right). \end{aligned} \quad (5.73)$$

We use Eq. (5.61) to rewrite $\partial g_{\mu\lambda}/\partial x^\rho$ and Eq. (5.73) becomes

$$g^{\mu\lambda} \frac{\partial \Gamma_{\nu\sigma}^\lambda}{\partial x^\rho} = \frac{1}{2} \left(\frac{\partial^2 g_{\mu\sigma}}{\partial x^\rho \partial x^\nu} + \frac{\partial^2 g_{\nu\mu}}{\partial x^\rho \partial x^\sigma} - \frac{\partial^2 g_{\nu\sigma}}{\partial x^\mu \partial x^\rho} \right) - g_{\mu\kappa} \Gamma_{\nu\sigma}^\lambda \Gamma_{\lambda\rho}^\kappa - g_{\kappa\lambda} \Gamma_{\nu\sigma}^\lambda \Gamma_{\mu\rho}^\kappa. \quad (5.74)$$

In the same way, we can rewrite the second term on the right hand side of Eq. (5.72)

$$g^{\mu\lambda} \frac{\partial \Gamma_{\nu\rho}^\lambda}{\partial x^\sigma} = \frac{1}{2} \left(\frac{\partial^2 g_{\mu\rho}}{\partial x^\sigma \partial x^\nu} + \frac{\partial^2 g_{\nu\mu}}{\partial x^\sigma \partial x^\rho} - \frac{\partial^2 g_{\nu\rho}}{\partial x^\mu \partial x^\sigma} \right) - g_{\mu\kappa} \Gamma_{\nu\rho}^\lambda \Gamma_{\lambda\sigma}^\kappa - g_{\kappa\lambda} \Gamma_{\nu\rho}^\lambda \Gamma_{\mu\sigma}^\kappa. \quad (5.75)$$

With Eqs. (5.74) and (5.75) we can rewrite $R_{\mu\nu\rho\sigma}$ as

$$R_{\mu\nu\rho\sigma} = \frac{1}{2} \left(\frac{\partial^2 g_{\mu\sigma}}{\partial x^\nu \partial x^\rho} + \frac{\partial^2 g_{\nu\rho}}{\partial x^\mu \partial x^\sigma} - \frac{\partial^2 g_{\mu\rho}}{\partial x^\nu \partial x^\sigma} - \frac{\partial^2 g_{\nu\sigma}}{\partial x^\mu \partial x^\rho} \right) + g_{\kappa\lambda} (\Gamma_{\nu\rho}^\lambda \Gamma_{\mu\sigma}^\kappa - \Gamma_{\nu\sigma}^\lambda \Gamma_{\mu\rho}^\kappa). \quad (5.76)$$

The Riemann tensor $R_{\mu\nu\rho\sigma}$ is antisymmetric in the first and second indices as well as in the third and fourth indices, while it is symmetric if we exchange the first and second indices respectively with the third and fourth indices:

$$R_{\mu\nu\rho\sigma} = -R_{\nu\mu\rho\sigma} = -R_{\mu\nu\sigma\rho} = R_{\rho\sigma\mu\nu}. \quad (5.77)$$

Note that, if $R_{\nu\rho\sigma}^\mu = 0$ in a certain coordinate system, it vanishes in any coordinate system. This follows from the transformation rule of tensors

$$R_{\nu\rho\sigma}^\mu \rightarrow R'_{\nu\rho\sigma}{}^\mu = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x^\beta}{\partial x'^\nu} \frac{\partial x^\gamma}{\partial x'^\rho} \frac{\partial x^\delta}{\partial x'^\sigma} R_{\beta\gamma\delta}^\alpha. \quad (5.78)$$

In particular, since in flat spacetime in Cartesian coordinates the Riemann tensor vanishes, it vanishes in any coordinate system even if the Christoffel symbols may not vanish. So in flat spacetime all the components of the Riemann tensor are identically zero. Note that such a statement is not true for the Christoffel symbols, because they transform with the rule in Eq. (5.21), where the last term may be non-zero under a certain coordinate transformation.

5.4.2 Geometrical Interpretation

Here we want to show that the result of parallel transport of a vector depends on the path. With reference to Fig. 5.2, we have the vector \mathbf{V} at point $A = x_A = \{x^\mu\}$. The

vector has components

$$V_A^\mu = V^\mu . \quad (5.79)$$

Let us now parallel transport the vector to point $B = x_B = \{x^\mu + p^\mu\}$, where p^μ is an infinitesimal displacement. After parallel transport, the components of the vector are

$$V_{A \rightarrow B}^\mu = V^\mu - \Gamma_{\nu\rho}^\mu(x_A) V^\nu p^\rho . \quad (5.80)$$

Lastly, we parallel transport the vector to point $D = x_D = \{x^\mu + p^\mu + q^\mu\}$, where q^μ is an infinitesimal displacement too. At point D the components of the vector are

$$\begin{aligned} V_{A \rightarrow B \rightarrow D}^\mu &= V_{A \rightarrow B}^\mu - \Gamma_{\nu\rho}^\mu(x_B) V_{A \rightarrow B}^\nu q^\rho \\ &= V^\mu - \Gamma_{\nu\rho}^\mu(x_A) V^\nu p^\rho - \left[\Gamma_{\nu\rho}^\mu(x_A) + \frac{\partial \Gamma_{\nu\rho}^\mu}{\partial x^\sigma}(x_A) p^\sigma \right] [V^\nu - \Gamma_{\tau\nu}^\nu(x_A) V^\tau p^\nu] q^\rho \\ &= V^\mu - \Gamma_{\nu\rho}^\mu(x_A) V^\nu p^\rho - \Gamma_{\nu\rho}^\mu(x_A) V^\nu q^\rho - \frac{\partial \Gamma_{\nu\rho}^\mu}{\partial x^\sigma}(x_A) p^\sigma V^\nu q^\rho \\ &\quad + \Gamma_{\nu\rho}^\mu(x_A) \Gamma_{\tau\nu}^\nu(x_A) V^\tau p^\nu q^\rho , \end{aligned} \quad (5.81)$$

where we have neglected terms of order higher than second in the infinitesimal displacements p^μ and q^μ .

Let us now do the same changing path. We start from the vector \mathbf{V} at point A and we parallel transport it to point $C = x_C = \{x^\mu + q^\mu\}$, as shown in Fig. 5.2. The result is the vector $\mathbf{V}_{A \rightarrow C}$. We continue and we parallel transport the vector to point D , where the vector components are

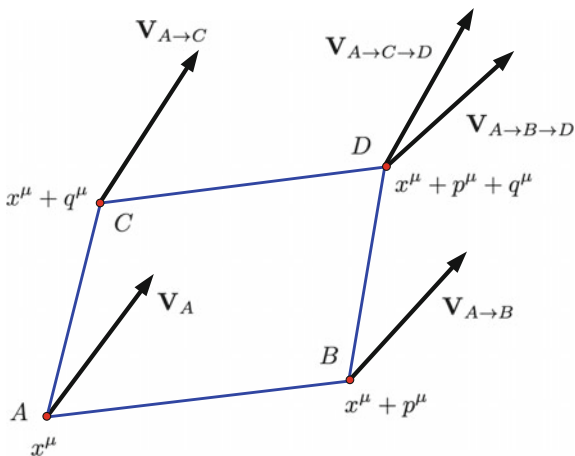
$$\begin{aligned} V_{A \rightarrow C \rightarrow D}^\mu &= V^\mu - \Gamma_{\nu\rho}^\mu(x_A) V^\nu q^\rho - \Gamma_{\nu\rho}^\mu(x_A) V^\nu p^\rho - \frac{\partial \Gamma_{\nu\rho}^\mu}{\partial x^\sigma}(x_A) q^\sigma V^\nu p^\rho \\ &\quad + \Gamma_{\nu\rho}^\mu(x_A) \Gamma_{\tau\nu}^\nu(x_A) V^\tau q^\nu p^\rho . \end{aligned} \quad (5.82)$$

If we compare Eqs. (5.81) and (5.82), we find that

$$\begin{aligned} V_{A \rightarrow B \rightarrow D}^\mu - V_{A \rightarrow C \rightarrow D}^\mu &= \left(\frac{\partial \Gamma_{\tau\rho}^\mu}{\partial x^\nu} - \frac{\partial \Gamma_{\tau\nu}^\mu}{\partial x^\rho} + \Gamma_{\nu\nu}^\mu \Gamma_{\tau\rho}^\nu - \Gamma_{\nu\rho}^\mu \Gamma_{\tau\nu}^\nu \right)_{x=x_A} V^\tau q^\nu p^\rho \\ &= R_{\tau\nu\rho}^\mu V^\tau q^\nu p^\rho . \end{aligned} \quad (5.83)$$

The difference in the parallel transport between the two paths is regulated by the Riemann tensor. In flat spacetime, the Riemann tensor vanishes, and, indeed, if we parallel transport a vector from one point to another the result is independent of the choice of the path.

Fig. 5.2 If the vector \mathbf{V}_A at $A = \{x^\mu\}$ is parallel transported to point $B = \{x^\mu + p^\mu\}$ and then to point $D = \{x^\mu + p^\mu + q^\mu\}$, we obtain the vector $\mathbf{V}_{A \rightarrow B \rightarrow D}$. If \mathbf{V}_A is parallel transported to point $C = \{x^\mu + q^\mu\}$ and then to point $D = \{x^\mu + p^\mu + q^\mu\}$, we obtain the vector $\mathbf{V}_{A \rightarrow C \rightarrow D}$. In general, $\mathbf{V}_{A \rightarrow B \rightarrow D}$ and $\mathbf{V}_{A \rightarrow C \rightarrow D}$ are not the same vector



5.4.3 Ricci Tensor and Scalar Curvature

From the Riemann tensor, we can define the Ricci tensor and the scalar curvature after contracting its indices. The *Ricci tensor* is a tensor of second order defined as

$$R_{\mu\nu} = R^\lambda_{\mu\lambda\nu} = \frac{\partial \Gamma^\lambda_{\mu\nu}}{\partial x^\lambda} - \frac{\partial \Gamma^\lambda_{\mu\lambda}}{\partial x^\nu} + \Gamma^\kappa_{\mu\nu} \Gamma^\lambda_{\kappa\lambda} - \Gamma^\kappa_{\mu\lambda} \Gamma^\lambda_{\nu\kappa}. \tag{5.84}$$

The Ricci tensor is symmetric

$$R_{\mu\nu} = R_{\nu\mu}. \tag{5.85}$$

Contracting the indices of the Ricci tensor we obtain the *scalar curvature*

$$R = R^\mu_{\mu} = g^{\mu\nu} R_{\mu\nu}. \tag{5.86}$$

With the Ricci tensor and the scalar curvature we can define the *Einstein tensor* as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R. \tag{5.87}$$

Since both $R_{\mu\nu}$ and $g_{\mu\nu}$ are symmetric tensors of second order, the Einstein tensor is a symmetric tensor of second order as well.

5.4.4 Bianchi Identities

The Bianchi Identities are two important identities involving the Riemann tensor. The *First Bianchi Identity* reads

$$R^\mu{}_{\nu\rho\sigma} + R^\mu{}_{\rho\sigma\nu} + R^\mu{}_{\sigma\nu\rho} = 0, \quad (5.88)$$

and it can be easily verified by using the explicit expression of the Riemann tensor in Eq. (5.71). Indeed we have

$$\begin{aligned} R^\mu{}_{\nu\rho\sigma} + R^\mu{}_{\rho\sigma\nu} + R^\mu{}_{\sigma\nu\rho} &= \frac{\partial\Gamma^\mu_{\nu\sigma}}{\partial x^\rho} - \frac{\partial\Gamma^\mu_{\nu\rho}}{\partial x^\sigma} + \Gamma^\lambda_{\nu\sigma}\Gamma^\mu_{\rho\lambda} - \Gamma^\lambda_{\nu\rho}\Gamma^\mu_{\sigma\lambda} \\ &\quad + \frac{\partial\Gamma^\mu_{\rho\nu}}{\partial x^\sigma} - \frac{\partial\Gamma^\mu_{\rho\sigma}}{\partial x^\nu} + \Gamma^\lambda_{\rho\nu}\Gamma^\mu_{\sigma\lambda} - \Gamma^\lambda_{\rho\sigma}\Gamma^\mu_{\nu\lambda} \\ &\quad + \frac{\partial\Gamma^\mu_{\sigma\rho}}{\partial x^\nu} - \frac{\partial\Gamma^\mu_{\sigma\nu}}{\partial x^\rho} + \Gamma^\lambda_{\sigma\rho}\Gamma^\mu_{\nu\lambda} - \Gamma^\lambda_{\sigma\nu}\Gamma^\mu_{\rho\lambda} \\ &= 0. \end{aligned} \quad (5.89)$$

The *Second Bianchi Identity* reads

$$\nabla_\mu R^\kappa{}_{\lambda\nu\rho} + \nabla_\nu R^\kappa{}_{\lambda\rho\mu} + \nabla_\rho R^\kappa{}_{\lambda\mu\nu} = 0. \quad (5.90)$$

The first term on the left hand side in Eq. (5.90) can be written as

$$\nabla_\mu R^\kappa{}_{\lambda\nu\rho} = \nabla_\mu \left(\frac{\partial\Gamma^\kappa_{\lambda\rho}}{\partial x^\nu} - \frac{\partial\Gamma^\kappa_{\lambda\nu}}{\partial x^\rho} + \Gamma^\sigma_{\lambda\rho}\Gamma^\kappa_{\nu\sigma} - \Gamma^\sigma_{\lambda\nu}\Gamma^\kappa_{\rho\sigma} \right). \quad (5.91)$$

If we choose a coordinate system in which the Christoffel symbols vanish at a certain point (this is always possible, see Sect. 6.4.2), at that point Eq. (5.91) becomes

$$\nabla_\mu R^\kappa{}_{\lambda\nu\rho} = \frac{\partial^2\Gamma^\kappa_{\lambda\rho}}{\partial x^\mu\partial x^\nu} - \frac{\partial^2\Gamma^\kappa_{\lambda\nu}}{\partial x^\mu\partial x^\rho}, \quad (5.92)$$

because in the case of vanishing Christoffel symbols the covariant derivative reduces to the partial one. The second and third terms on the left hand side in Eq. (5.90) read, respectively,

$$\nabla_\nu R^\kappa{}_{\lambda\rho\mu} = \frac{\partial^2\Gamma^\kappa_{\lambda\mu}}{\partial x^\nu\partial x^\rho} - \frac{\partial\Gamma^\kappa_{\lambda\rho}}{\partial x^\nu\partial x^\mu}, \quad (5.93)$$

$$\nabla_\rho R^\kappa{}_{\lambda\mu\nu} = \frac{\partial\Gamma^\kappa_{\lambda\nu}}{\partial x^\rho\partial x^\mu} - \frac{\partial\Gamma^\kappa_{\lambda\mu}}{\partial x^\rho\partial x^\nu}. \quad (5.94)$$

The Second Bianchi Identity can thus be written as

$$\begin{aligned}
\nabla_{\mu} R^{\kappa}_{\lambda\nu\rho} + \nabla_{\nu} R^{\kappa}_{\lambda\rho\mu} + \nabla_{\rho} R^{\kappa}_{\lambda\mu\nu} &= \frac{\partial^2 \Gamma^{\kappa}_{\lambda\rho}}{\partial x^{\mu} \partial x^{\nu}} - \frac{\partial^2 \Gamma^{\kappa}_{\lambda\nu}}{\partial x^{\mu} \partial x^{\rho}} \\
&+ \frac{\partial^2 \Gamma^{\kappa}_{\lambda\mu}}{\partial x^{\nu} \partial x^{\rho}} - \frac{\partial \Gamma^{\kappa}_{\lambda\rho}}{\partial x^{\nu} \partial x^{\mu}} \\
&+ \frac{\partial \Gamma^{\kappa}_{\lambda\nu}}{\partial x^{\rho} \partial x^{\mu}} - \frac{\partial \Gamma^{\kappa}_{\lambda\mu}}{\partial x^{\rho} \partial x^{\nu}} \\
&= 0.
\end{aligned} \tag{5.95}$$

Since the left hand side is a tensor, if all its components vanish in a certain coordinate system they vanish in any coordinate system, and this concludes the proof of the identity.

From the Second Bianchi Identity we find that the covariant divergence of the Einstein tensor vanishes. As we will see in Sect. 7.1, this is of fundamental importance in Einstein's gravity. If we multiply the Second Bianchi Identity in Eq. (5.90) by g^{ν}_{κ} and we sum over the indices κ and ν , we find

$$\begin{aligned}
g^{\nu}_{\kappa} (\nabla_{\mu} R^{\kappa}_{\lambda\nu\rho} + \nabla_{\nu} R^{\kappa}_{\lambda\rho\mu} + \nabla_{\rho} R^{\kappa}_{\lambda\mu\nu}) &= 0, \\
\nabla_{\mu} (g^{\nu}_{\kappa} R^{\kappa}_{\lambda\nu\rho}) + g^{\nu}_{\kappa} \nabla_{\nu} R^{\kappa}_{\lambda\rho\mu} - \nabla_{\rho} (g^{\nu}_{\kappa} R^{\kappa}_{\lambda\nu\mu}) &= 0, \\
\nabla_{\mu} R_{\lambda\rho} + \nabla_{\kappa} R^{\kappa}_{\lambda\rho\mu} - \nabla_{\rho} R_{\lambda\mu} &= 0.
\end{aligned} \tag{5.96}$$

We multiply by $g^{\lambda\rho}$ and we sum over the indices λ and ρ

$$\begin{aligned}
g^{\lambda\rho} (\nabla_{\mu} R_{\lambda\rho} + \nabla_{\kappa} R^{\kappa}_{\lambda\rho\mu} - \nabla_{\rho} R_{\lambda\mu}) &= 0, \\
\nabla_{\mu} (g^{\lambda\rho} R_{\lambda\rho}) - \nabla_{\kappa} (g^{\lambda\rho} g^{\kappa\sigma} R_{\lambda\sigma\rho\mu}) - \nabla_{\rho} (g^{\lambda\rho} R_{\lambda\mu}) &= 0, \\
\nabla_{\mu} R - \nabla_{\kappa} R^{\kappa}_{\mu} - \nabla_{\rho} R^{\rho}_{\mu} &= 0, \\
\nabla_{\kappa} (g^{\kappa}_{\mu} R - 2R^{\kappa}_{\mu}) &= 0.
\end{aligned} \tag{5.97}$$

This is equivalent to

$$\nabla_{\mu} G^{\mu\nu} = 0. \tag{5.98}$$

Problems

5.1 Write the components of the following tensors:

$$\nabla_{\mu} A_{\alpha\beta}, \quad \nabla_{\mu} A^{\alpha\beta}, \quad \nabla_{\mu} A^{\alpha}_{\beta}, \quad \nabla_{\mu} A^{\beta}_{\alpha}, \tag{5.99}$$

5.2 Write the non-vanishing components of the Riemann tensor, the Ricci tensor, and the scalar curvature for the Minkowski spacetime in spherical coordinates.

5.3 Check that the Ricci tensor is symmetric.

Reference

1. M.H. Protter, C.B. Morrey, *A First Course in Real Analysis* (Springer, New York, 1991)