Chapter 6

Curvature of a manifold; flat manifolds.

6.1 The commutator of second covariant derivatives.

The covariant derivative was introduced in order that derivatives of tensor densities are still tensor densities. This advantage has, however, a few inconvenient consequences. One of these we have already encountered: the parallel transport defined via covariant differentiation depends on the path. The second inconvenient consequence will appear in this chapter: the second covariant derivatives do not commute, unlike partial derivatives of a function of class C^2.

For a proper scalar (with weight $w = 0$) we have

$$(\nabla_\delta \nabla_\gamma - \nabla_\gamma \nabla_\delta) T = \nabla_\delta (T,_\gamma) - \nabla_\gamma (T,_\delta) = T,_{\gamma\delta} - \Gamma^\rho{}_{\gamma\delta} T,_\rho - T,_{\delta\gamma} + \Gamma^\rho{}_{\delta\gamma} T,_\rho = -2\Omega^\rho{}_{\gamma\delta} T,_\rho \,. \tag{6.1}$$

Hence, the second covariant derivatives of a scalar commute only in torsion-free manifolds.

For a scalar density of weight w we have

$$\begin{aligned}(\nabla_\delta \nabla_\gamma &- \nabla_\gamma \nabla_\delta) \check{T} \\ &= \left(\nabla_\gamma \check{T}\right),_\delta + w\Gamma^\rho{}_{\rho\delta} \nabla_\gamma \check{T} - \Gamma^\rho{}_{\gamma\delta} \nabla_\rho \check{T} - \left(\nabla_\delta \check{T}\right),_\gamma - w\Gamma^\rho{}_{\rho\gamma} \nabla_\delta \check{T} + \Gamma^\rho{}_{\delta\gamma} \nabla_\rho \check{T} \\ &= \left(\check{T},_\gamma + w\Gamma^\rho{}_{\rho\gamma} \check{T}\right),_\delta + w\Gamma^\rho{}_{\rho\delta} \left(\check{T},_\gamma + w\Gamma^\sigma{}_{\sigma\gamma} \check{T}\right) - \Gamma^\rho{}_{\gamma\delta} \nabla_\rho \check{T} \\ &\qquad - \left(\check{T},_\delta + w\Gamma^\rho{}_{\rho\delta} \check{T}\right),_\gamma - w\Gamma^\rho{}_{\rho\gamma} \left(\check{T},_\delta + w\Gamma^\sigma{}_{\sigma\delta} \check{T}\right) + \Gamma^\rho{}_{\delta\gamma} \nabla_\rho \check{T} \\ &= w \left(\Gamma^\rho{}_{\rho\gamma,\delta} - \Gamma^\rho{}_{\rho\delta,\gamma}\right) \check{T} - 2\Omega^\rho{}_{\gamma\delta} \nabla_\rho \check{T}. \end{aligned} \tag{6.2}$$

Let us leave this formula without a comment for a while and let us do the analogous calculation for a covariant vector field:

$$\begin{aligned}(\nabla_\delta \nabla_\gamma &- \nabla_\gamma \nabla_\delta) T_\beta = \nabla_\delta \left(T_{\beta|\gamma}\right) - \nabla_\gamma \left(T_{\beta|\delta}\right) \\ &= \left(T_{\beta|\gamma}\right),_\delta - \Gamma^\sigma{}_{\beta\delta} T_{\sigma|\gamma} - \Gamma^\sigma{}_{\gamma\delta} T_{\beta|\sigma} - \left(T_{\beta|\delta}\right),_\gamma + \Gamma^\sigma{}_{\beta\gamma} T_{\sigma|\delta} + \Gamma^\sigma{}_{\delta\gamma} T_{\beta|\sigma} \\ &= \left(T_{\beta,\gamma} - \Gamma^\rho{}_{\beta\gamma} T_\rho\right),_\delta - \Gamma^\sigma{}_{\beta\delta} \left(T_{\sigma,\gamma} - \Gamma^\rho{}_{\sigma\gamma} T_\rho\right) - \Gamma^\sigma{}_{\gamma\delta} T_{\beta|\sigma} \\ &\quad - \left(T_{\beta,\delta} - \Gamma^\rho{}_{\beta\delta} T_\rho\right),_\gamma + \Gamma^\sigma{}_{\beta\gamma} \left(T_{\sigma,\delta} - \Gamma^\rho{}_{\sigma\delta} T_\rho\right) + \Gamma^\sigma{}_{\delta\gamma} T_{\beta|\sigma} \end{aligned}$$

$$= \left(-\Gamma^{\rho}{}_{\beta\gamma,\delta} + \Gamma^{\rho}{}_{\beta\delta,\gamma} + \Gamma^{\rho}{}_{\sigma\gamma}\Gamma^{\sigma}{}_{\beta\delta} - \Gamma^{\rho}{}_{\sigma\delta}\Gamma^{\sigma}{}_{\beta\gamma}\right) T_{\rho} - 2\Omega^{\sigma}{}_{\gamma\delta} T_{\beta|\sigma}. \tag{6.3}$$

Now let us define

$$B^{\alpha}{}_{\beta\gamma\delta} \stackrel{\text{def}}{=} -\Gamma^{\alpha}{}_{\beta\gamma,\delta} + \Gamma^{\alpha}{}_{\beta\delta,\gamma} + \Gamma^{\alpha}{}_{\sigma\gamma}\Gamma^{\sigma}{}_{\beta\delta} - \Gamma^{\alpha}{}_{\sigma\delta}\Gamma^{\sigma}{}_{\beta\gamma} \equiv -2\Gamma^{\alpha}{}_{\beta[\gamma,\delta]} + 2\Gamma^{\alpha}{}_{\sigma[\gamma}\Gamma^{\sigma}{}_{|\beta|\delta]}. \tag{6.4}$$

The quantity $B^{\alpha}{}_{\beta\gamma\delta}$ is called the **curvature tensor**. In order to note that it is a tensor indeed, let us rewrite eq. (6.3) with (6.4) substituted in it

$$\left(\nabla_{\delta}\nabla_{\gamma} - \nabla_{\gamma}\nabla_{\delta}\right) T_{\beta} = B^{\rho}{}_{\beta\gamma\delta} T_{\rho} - 2\Omega^{\sigma}{}_{\gamma\delta} T_{\beta|\sigma}. \tag{6.5}$$

The left-hand side of (6.5) is a tensor by definition. The quantities $\Omega^{\sigma}{}_{\gamma\delta}$ and $T_{\beta|\sigma}$ are tensors, too, hence the second term on the right is a tensor. It follows that $B^{\rho}{}_{\beta\gamma\delta}T_{\rho}$ is a tensor. Now, since T_{ρ} is a covariant vector, and an arbitrary one in addition, $B^{\rho}{}_{\beta\gamma\delta}$ must be a tensor itself. The tensorial property of $B^{\rho}{}_{\beta\gamma\delta}$ can also be verified using the definition (6.4) and the transformation law of the connection coefficients (4.23).

From (6.4) it follows that

$$B^{\rho}{}_{\rho\gamma\delta} = -\Gamma^{\rho}{}_{\rho\gamma,\delta} + \Gamma^{\rho}{}_{\rho\delta,\gamma}. \tag{6.6}$$

Hence, eq. (6.2) can be rewritten as follows

$$\left(\nabla_{\delta}\nabla_{\gamma} - \nabla_{\gamma}\nabla_{\delta}\right) \check{T} = -wB^{\rho}{}_{\rho\gamma\delta}\check{T} - 2\Omega^{\rho}{}_{\gamma\delta}\check{T}_{|\rho}. \tag{6.7}$$

Finally, for the same commutator acting on a contravariant vector, we obtain

$$\left(\nabla_{\delta}\nabla_{\gamma} - \nabla_{\gamma}\nabla_{\delta}\right) T^{\alpha} = -B^{\alpha}{}_{\rho\gamma\delta} T^{\rho} - 2\Omega^{\sigma}{}_{\gamma\delta} T^{\alpha}{}_{|\sigma}. \tag{6.8}$$

By direct calculation, it can be verified that the commutator of covariant derivatives acts on the tensor product of two tensor densities $T_1 \otimes T_2$ in the following way

$$\left(\nabla_{\delta}\nabla_{\gamma} - \nabla_{\gamma}\nabla_{\delta}\right)\left(T_1 \otimes T_2\right) = \left[\left(\nabla_{\delta}\nabla_{\gamma} - \nabla_{\gamma}\nabla_{\delta}\right) T_1\right] \otimes T_2 + T_1 \otimes \left[\left(\nabla_{\delta}\nabla_{\gamma} - \nabla_{\gamma}\nabla_{\delta}\right) T_2\right]. \tag{6.9}$$

(the weights and indices of T_1 and T_2 are irrelevant here). Hence, the operator $(\nabla_{\delta}\nabla_{\gamma} - \nabla_{\gamma}\nabla_{\delta})$ has the property of an ordinary differentiation. Then, an arbitrary tensor density of type $[w, k, l]$ behaves, under covariant differentiation, like a tensor product of a single scalar density of weight w, k contravariant vectors and l covariant vectors – see (4.34). Hence, we can guess the formula for the commutator of second covariant derivatives acting on an arbitrary tensor density:

$$\begin{aligned}\left(\nabla_{\delta}\nabla_{\gamma} - \nabla_{\gamma}\nabla_{\delta}\right) T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l} = {} & -wB^{\rho}{}_{\rho\gamma\delta}T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l} - \sum_{i=1}^{k} B^{\alpha_i}{}_{\rho_i\gamma\delta} T^{\alpha_1\ldots\rho_i\ldots\alpha_k}_{\beta_1\ldots\beta_l} \\ & + \sum_{j=1}^{l} B^{\rho_j}{}_{\beta_j\gamma\delta} T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\rho_j\ldots\beta_l} - 2\Omega^{\rho}{}_{\gamma\delta} T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l|\rho}.\end{aligned} \tag{6.10}$$

Equation (6.10) is called the **Ricci formula**. Let us note, from (6.4), that

$$B^{\alpha}{}_{\beta\gamma\delta} = B^{\alpha}{}_{\beta[\gamma\delta]}$$

The name "curvature tensor" evokes the right associations. We will see it later (Chap. 7), when we deal with manifolds with a metric – then, curved surfaces in a Euclidean space will become special cases of our considerations.

6.2 The relation between curvature and parallel transport.

By (5.8), a tensor density $T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l}$ is parallely transported along a curve $x^\alpha(\tau)$ when

$$\frac{\mathrm{D}}{\mathrm{d}\tau}T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l} \stackrel{\text{def}}{=} \frac{\mathrm{d}x^\rho}{\mathrm{d}\tau}\nabla_\rho T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l} = 0. \tag{6.11}$$

This is a linear homogeneous set of first-order differential equations with respect to the functions $T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l}(\tau)$. The solutions of such equations are linear functions of the initial conditions. Let the initial conditions for $T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l}(\tau)$ be given at $\tau = 0$. Then the solution of (6.11) can be represented as

$$T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l}(\tau) = P^{\alpha_1\ldots\alpha_k}_{\overline{\alpha_1}\ldots\overline{\alpha_k}}{}^{\overline{\beta_1}\ldots\overline{\beta_l}}_{\beta_1\ldots\beta_l}(\tau)T^{\overline{\alpha_1}\ldots\overline{\alpha_k}}_{\overline{\beta_1}\ldots\overline{\beta_l}}(0), \tag{6.12}$$

where $P^{\alpha_1\ldots\alpha_k}_{\overline{\alpha_1}\ldots\overline{\alpha_k}}{}^{\overline{\beta_1}\ldots\overline{\beta_l}}_{\beta_1\ldots\beta_l}(\tau)$ is the **propagator of parallel transport** – a linear operator that maps the initial value $T^{\overline{\alpha_1}\ldots\overline{\alpha_k}}_{\overline{\beta_1}\ldots\overline{\beta_l}}(0)$ into the running value $T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l}(\tau)$. It has two sets of indices, those with a bar correspond to the initial point $x^\alpha(0)$, those without a bar correspond to the running point $x^\alpha(\tau)$. It depends on the two points and, apart from exceptional cases, on the curve along which the transport occurs. It transforms like a tensor density at the point $x^\alpha(\tau)$ multiplied in the sense of tensor product by a tensor density at the point $x^\alpha(0)$. In consequence of (6.11), the parallel propagator must obey

$$\frac{\mathrm{D}}{\mathrm{d}\tau}P^{\alpha_1\ldots\alpha_k}_{\overline{\alpha_1}\ldots\overline{\alpha_k}}{}^{\overline{\beta_1}\ldots\overline{\beta_l}}_{\beta_1\ldots\beta_l}(\tau) = 0, \tag{6.13}$$

and, in consequence of (6.12), it must obey the initial condition

$$P^{\alpha_1\ldots\alpha_k}_{\overline{\alpha_1}\ldots\overline{\alpha_k}}{}^{\overline{\beta_1}\ldots\overline{\beta_l}}_{\beta_1\ldots\beta_l}(0) = \delta^{\alpha_1}_{\overline{\alpha_1}}\ldots\delta^{\alpha_k}_{\overline{\alpha_k}}\delta^{\overline{\beta_1}}_{\beta_1}\ldots\delta^{\overline{\beta_l}}_{\beta_l}. \tag{6.14}$$

In eq. (6.13), the covariant differentiation applies only to the indices without a bar because the initial point of the transport, to which the barred indices refer, does not depend on τ.

We will give a special name to the propagator of parallel transport along a closed curve. Let $0 \le \tau \le 1$ be a parameter on a closed curve and let $x^\alpha(0) = x^\alpha(1)$ be the beginning-end point of the curve. We denote:

$$P^{\alpha_1\ldots\alpha_k}_{\overline{\alpha_1}\ldots\overline{\alpha_k}}{}^{\overline{\beta_1}\ldots\overline{\beta_l}}_{\beta_1\ldots\beta_l}(1) = S^{\alpha_1\ldots\alpha_k}_{\overline{\alpha_1}\ldots\overline{\alpha_k}}{}^{\overline{\beta_1}\ldots\overline{\beta_l}}_{\beta_1\ldots\beta_l}. \tag{6.15}$$

In $S^{\cdots\cdots}_{\cdots\cdots}$ all the indices refer to the same point $x^\alpha(0) = x^\alpha(1)$, hence $S^{\cdots\cdots}_{\cdots\cdots}$ is an ordinary tensor density of type $[w, k+l, k+l]$.

Now let us define elementary propagators: P_w, to transport scalar densities, $P^\alpha_{\overline{\alpha}}$, to transport contravariant vectors, and $P^{\overline{\beta}}_\beta$ to transport covariant vectors. All of them obey (6.13), and the initial conditions

$$P_w(0) = 1, \qquad P^\alpha_{\overline{\alpha}}(0) = \delta^\alpha_{\overline{\alpha}}, \qquad P^{\overline{\beta}}_\beta(0) = \delta^{\overline{\beta}}_\beta. \tag{6.16}$$

The following equation holds

$$P^{\alpha_1\ldots\alpha_k}_{\overline{\alpha_1}\ldots\overline{\alpha_k}}{}^{\overline{\beta_1}\ldots\overline{\beta_l}}_{\beta_1\ldots\beta_l} = P_w P^{\alpha_1}_{\overline{\alpha_1}}\ldots P^{\alpha_k}_{\overline{\alpha_k}} P^{\overline{\beta_1}}_{\beta_1}\ldots P^{\overline{\beta_l}}_{\beta_l}. \tag{6.17}$$

It follows from the fact that its both sides obey the same set of differential equations (6.13) and the same set of initial conditions (6.14), so must be identical.

Similarly, the general propagator of transport along a closed curve, $S^{\cdots\cdots}_{\cdots\cdots}$, is the tensor product of elementary propagators

$$S^{\alpha_1\ldots\alpha_k}_{\overline{\alpha_1}\ldots\overline{\alpha_k}}{}^{\overline{\beta_1}\ldots\overline{\beta_l}}_{\beta_1\ldots\beta_l} = S_w S^{\alpha_1}_{\overline{\alpha_1}}\ldots S^{\alpha_k}_{\overline{\alpha_k}} S^{\overline{\beta_1}}_{\beta_1}\ldots S^{\overline{\beta_l}}_{\beta_l}. \tag{6.18}$$

Hence, in order to investigate the properties of general propagators it is enough to deal with the elementary propagators.

Let us consider the parallel transport along a curve $x^\alpha(\tau)$ from the point $x^\alpha(0)$ to an arbitrary point $x^\alpha(\tau)$ of two vectors, V_α and W^α. From (6.11) it follows that their scalar product, when parallely transported along $x^\alpha(\tau)$,

$$V_\alpha(\tau)W^\alpha(\tau) \equiv P^{\overline{\alpha}}_{\alpha}(\tau)P^{\alpha}_{\overline{\beta}}(\tau)V_{\overline{\alpha}}(0)W^{\overline{\beta}}(0). \tag{6.19}$$

is in fact constant along the curve. Then

$$V_\alpha(\tau)W^\alpha(\tau) = V_{\overline{\alpha}}(0)W^{\overline{\alpha}}(0) = \delta^{\overline{\alpha}}_{\overline{\beta}}V_{\overline{\alpha}}(0)W^{\overline{\beta}}(0). \tag{6.20}$$

Comparing this with the previous equation, we see that

$$\left(P^{\overline{\alpha}}_{\alpha}(\tau)P^{\alpha}_{\overline{\beta}}(\tau) - \delta^{\overline{\alpha}}_{\overline{\beta}}\right)V_{\overline{\alpha}}(0)W^{\overline{\beta}}(0) = 0. \tag{6.21}$$

However, the vectors $V_{\overline{\alpha}}(0)$ and $W^{\overline{\beta}}(0)$ are arbitrary, so

$$P^{\overline{\alpha}}_{\rho}(\tau)P^{\rho}_{\overline{\beta}}(\tau) = \delta^{\overline{\alpha}}_{\overline{\beta}}. \tag{6.22}$$

Now let us contract this equation with $P^{\overline{\beta}}_{\gamma}(\tau)V_{\overline{\alpha}}(0)$. The result is

$$\left(P^{\rho}_{\overline{\beta}}P^{\overline{\beta}}_{\gamma} - \delta^{\rho}_{\gamma}\right)P^{\overline{\alpha}}_{\rho}(\tau)V_{\overline{\alpha}}(0) = \left(P^{\rho}_{\overline{\beta}}P^{\overline{\beta}}_{\gamma} - \delta^{\rho}_{\gamma}\right)V_\rho(\tau) = 0. \tag{6.23}$$

Since $V_\rho(\tau)$ is an arbitrary vector, we have as a consequence

$$P^{\alpha}_{\overline{\beta}}P^{\overline{\beta}}_{\gamma} = \delta^{\alpha}_{\gamma}. \tag{6.24}$$

Equations (6.22) and (6.24) show that the operators $P^{\overline{\alpha}}_{\alpha}$ and $P^{\beta}_{\overline{\beta}}$ are inverse to each other: $P^{\beta}_{\overline{\beta}}$ is not only the propagator of parallel transport of a contravariant vector from $x^{\overline{\alpha}}(0)$ to $x^\alpha(\tau)$, but at the same time the propagator of parallel transport of a covariant vector from $x^\alpha(\tau)$ to $x^\alpha(0)$ *along the same curve.* A similar duality exists for $P^{\overline{\alpha}}_{\alpha}$.

In consequence of (6.22) and (6.24), the following two equations also hold

$$S^{\alpha}_{\overline{\rho}}S^{\overline{\rho}}_{\beta} = \delta^{\alpha}_{\beta}, \qquad S^{\overline{\gamma}}_{\sigma}S^{\sigma}_{\overline{\delta}} = \delta^{\overline{\gamma}}_{\overline{\delta}}. \tag{6.25}$$

Therefore we shall consider only the propagators P_w, S_w, $P_{\beta}^{\overline{\alpha}}$ and $S_{\beta}^{\overline{\alpha}}$, because $P_{\overline{\beta}}^{\alpha}$ and $S_{\overline{\beta}}^{\alpha}$ are algebraically determined by the first four.

Let us consider eq. (6.13) for $P_w(\tau)$. It becomes

$$\frac{\mathrm{d}P_w}{\mathrm{d}\tau} + w\Gamma^{\rho}{}_{\rho\sigma}\frac{\mathrm{d}x^{\sigma}}{\mathrm{d}\tau}P_w = 0. \tag{6.26}$$

This equation, with the initial condition $P_w(0) = 1$, has the following solution:

$$P_w(\tau) = \exp\left(-w\int_{\substack{0\\C}}^{\tau}\Gamma^{\rho}{}_{\rho\sigma}(t)\frac{\mathrm{d}x^{\sigma}}{\mathrm{d}t}(t)\mathrm{d}t\right). \tag{6.27}$$

Hence

$$S_w = \exp\left(-w\oint_C\Gamma^{\rho}{}_{\rho\sigma}(x)\mathrm{d}x^{\sigma}\right). \tag{6.28}$$

Now let us assume that the loop C can be contracted to a point (this assumption will be made in all that follows). Then we can use the Stokes theorem and express the integral along C through the integral over an arbitrary 2-surface leaf S_C spanned on C:

$$\begin{aligned}\oint_C\Gamma^{\rho}{}_{\rho\sigma}(x)\mathrm{d}x^{\sigma} &= \int_{S_C}\mathrm{d}\Gamma^{\rho}{}_{\rho\sigma}\wedge\mathrm{d}x^{\sigma} = \int_{S_C}\Gamma^{\rho}{}_{\rho\delta,\gamma}\mathrm{d}x^{\gamma}\wedge\mathrm{d}x^{\delta}\\ &= -\int_{S_C}\Gamma^{\rho}{}_{\rho[\gamma,\delta]}\mathrm{d}x^{\gamma}\wedge\mathrm{d}x^{\delta} = \frac{1}{2}\int_{S_C}B^{\rho}{}_{\rho\gamma\delta}\mathrm{d}x^{\gamma}\wedge\mathrm{d}x^{\delta}.\end{aligned} \tag{6.29}$$

Hence, in (6.28)

$$S_w = \exp\left(-\frac{1}{2}w\int_{S_C}B^{\rho}{}_{\rho\gamma\delta}\mathrm{d}_2x^{\gamma\delta}\right), \tag{6.30}$$

where $\mathrm{d}_2x^{\gamma\delta}$ denotes the surface element $\mathrm{d}x^{\gamma}\wedge\mathrm{d}x^{\delta}$.

To solve (6.13) for $P_{\beta}^{\overline{\beta}}$ we proceed as follows. We span a 2-surface leaf S_C on the closed loop C, and then we embed the curve C in a one-parameter family of curves $C(\epsilon)$ such that $C(0)$ is the single initial/final point of the loop, $x^{\alpha}(0) = x^{\alpha}(1)$, and $C(1) \equiv C$ (see Fig. 6.1). Points on this surface leaf will be labelled by two parameters, τ and ϵ, in such a way that $x^{\alpha}(0,\epsilon) = x^{\alpha}(1,\epsilon)$ and $x^{\alpha}(\tau_1, 0) = x^{\alpha}(\tau_2, 0)$ for all τ_1 and τ_2.

Then, the surface element under the integral is

$$\mathrm{d}x^{\gamma}\wedge\mathrm{d}x^{\delta} = \left(\frac{\partial x^{\gamma}}{\partial\tau}\frac{\partial x^{\delta}}{\partial\epsilon} - \frac{\partial x^{\gamma}}{\partial\epsilon}\frac{\partial x^{\delta}}{\partial\tau}\right)\mathrm{d}\tau\mathrm{d}\epsilon, \tag{6.31}$$

and so, for an arbitrary function $F(\tau,\epsilon)$

$$\int_{S_C}F(\tau,\epsilon)\mathrm{d}x^{\gamma}\wedge\mathrm{d}x^{\delta} = \int_0^1\mathrm{d}\tau\int_0^1\mathrm{d}\epsilon F(\tau,\epsilon)\left(\frac{\partial x^{\gamma}}{\partial\tau}\frac{\partial x^{\delta}}{\partial\epsilon} - \frac{\partial x^{\gamma}}{\partial\epsilon}\frac{\partial x^{\delta}}{\partial\tau}\right). \tag{6.32}$$

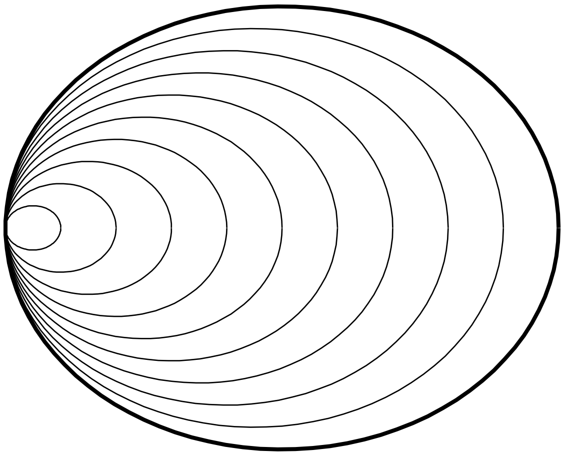

Figure 6.1: Embedding a loop in a one-parameter family of loops. The thicker line is the loop along which we consider the parallel transport; it corresponds to the parameter value $\epsilon = 1$. As ϵ goes down to 0 through all the values in $[0, 1]$, the loop becomes shorter and, in the limit $\epsilon \to 0$ degenerates to the single point where all the loops are tangent. The loops in this figure are ellipses given by the parametric equations $x = \epsilon a(1 - \cos t), y = \epsilon b \sin t$.

Along each loop $C(\epsilon)$ the propagator $P^{\overline{\beta}}_{\beta}(\tau, \epsilon)$ is defined by the equation

$$\frac{\mathrm{D}}{\partial \tau} P^{\overline{\beta}}_{\beta}(\tau, \epsilon) \underset{\epsilon}{\equiv} 0, \tag{6.33}$$

with the initial condition

$$P^{\overline{\beta}}_{\beta}(0, \epsilon) \underset{\epsilon}{\equiv} \delta^{\overline{\beta}}_{\beta}. \tag{6.34}$$

We also have

$$S^{\overline{\beta}}_{\beta}(\epsilon) = P^{\overline{\beta}}_{\beta}(1, \epsilon) \tag{6.35}$$

($S^{\overline{\beta}}_{\beta}$ in general depends on the curve, so it is a function of ϵ).

On the leaf S_C we can consider also the covariant derivatives by the parameter ϵ (i.e. along the curves given by $\tau =$ constant). We have

$$\frac{\mathrm{D}}{\partial \epsilon} S^{\overline{\beta}}_{\beta}(\epsilon) = \frac{\mathrm{d}}{\mathrm{d}\epsilon} S^{\overline{\beta}}_{\beta}(\epsilon) \tag{6.36}$$

because under a transformation of the parameter ϵ the propagator $S^{\overline{\beta}}_{\beta}(\epsilon)$ transforms like a scalar (by substitution only).

Now we use (6.10) to calculate the commutator of the second covariant derivatives by τ and ϵ acting on the propagator $P^{\overline{\beta}}_{\beta}(\tau, \epsilon)$:

$$\left(\frac{\mathrm{D}}{\partial \tau} \frac{\mathrm{D}}{\partial \epsilon} - \frac{\mathrm{D}}{\partial \epsilon} \frac{\mathrm{D}}{\partial \tau} \right) P^{\overline{\beta}}_{\beta}(\tau, \epsilon) = - \frac{\partial x^{\gamma}}{\partial \tau} \frac{\partial x^{\delta}}{\partial \epsilon} B^{\rho}{}_{\beta\gamma\delta} P^{\overline{\beta}}_{\rho}(\tau, \epsilon) \tag{6.37}$$

(the left-hand side of (6.37) produces first covariant derivatives of $(\partial x^\gamma/\partial\tau)$ and $\left(\partial x^\delta/\partial\epsilon\right)$, and their contribution cancels the torsion term of (6.10)). But the second term on the left-hand side is zero in virtue of (6.33). Knowing this, we contract (6.37) with $P^{\beta}_{\overline{\alpha}}$. Using (6.33) again we obtain

$$\frac{\mathrm{D}}{\partial\tau}\left(P^{\beta}_{\overline{\alpha}}(\tau,\epsilon)\frac{\mathrm{D}}{\partial\epsilon}P^{\overline{\beta}}_{\beta}(\tau,\epsilon)\right) = -\frac{\partial x^\gamma}{\partial\tau}\frac{\partial x^\delta}{\partial\epsilon}B^{\rho}{}_{\beta\gamma\delta}P^{\overline{\beta}}_{\rho}(\tau,\epsilon)P^{\beta}_{\overline{\alpha}}(\tau,\epsilon). \tag{6.38}$$

The expression in large parentheses on the left-hand side is a scalar with respect to the non-barred indices, so it is a scalar for the covariant differentiation with respect to τ. Hence, the covariant derivative on the left of (6.38) reduces to the partial derivative by τ. Then, we integrate the resulting equation by τ from 0 to 1, we make use of (6.34), of the covariant constancy of $\delta^{\overline{\beta}}_{\beta}$ (axiom 4 of the covariant derivative) and of (6.35). The result is

$$S^{\beta}_{\overline{\alpha}}(\epsilon)\frac{\mathrm{D}}{\partial\epsilon}S^{\overline{\beta}}_{\beta}(\epsilon) = -\oint_0^1 \mathrm{d}\tau\frac{\partial x^\gamma}{\partial\tau}\frac{\partial x^\delta}{\partial\epsilon}B^{\rho}{}_{\beta\gamma\delta}P^{\overline{\beta}}_{\rho}(\tau,\epsilon)P^{\beta}_{\overline{\alpha}}(\tau,\epsilon). \tag{6.39}$$

Now we contract both sides of the above equation with $S^{\overline{\alpha}}_{\alpha}$, use (6.25) and (6.36), integrate the result by ϵ from 0 to 1 and observe that $S^{\overline{\beta}}_{\alpha} = \delta^{\overline{\beta}}{}_{\alpha}$ at $\epsilon = 0$. The result is

$$\begin{aligned} S^{\overline{\beta}}_{\alpha} - \delta^{\overline{\beta}}_{\alpha} &= -\int_0^1 \mathrm{d}\epsilon\oint_0^1 \mathrm{d}\tau\frac{\partial x^\gamma}{\partial\tau}\frac{\partial x^\delta}{\partial\epsilon}B^{\rho}{}_{\beta\gamma\delta}P^{\overline{\beta}}_{\rho}(\tau,\epsilon)P^{\beta}_{\overline{\alpha}}(\tau,\epsilon)S^{\overline{\alpha}}_{\alpha}(\epsilon) \\ &= -\frac{1}{2}\int_0^1 \mathrm{d}\epsilon\oint_0^1 \mathrm{d}\tau B^{\rho}{}_{\beta\gamma\delta}P^{\overline{\beta}}_{\rho}(\tau,\epsilon)P^{\beta}_{\overline{\alpha}}(\tau,\epsilon)S^{\overline{\alpha}}_{\alpha}(\epsilon)\left(\frac{\partial x^\gamma}{\partial\tau}\frac{\partial x^\delta}{\partial\epsilon} - \frac{\partial x^\gamma}{\partial\epsilon}\frac{\partial x^\delta}{\partial\tau}\right) \\ &= -\frac{1}{2}\int_{S_C} B^{\rho}{}_{\beta\gamma\delta}P^{\overline{\beta}}_{\rho}(\tau,\epsilon)P^{\beta}_{\overline{\alpha}}(\tau,\epsilon)S^{\overline{\alpha}}_{\alpha}(\epsilon)\mathrm{d}_2x^{\gamma\delta}. \end{aligned} \tag{6.40}$$

Now it can be seen that if $B^{\alpha}{}_{\beta\gamma\delta} = 0$, then $S^{\overline{\beta}}_{\alpha} = \delta^{\overline{\beta}}_{\alpha}$, which means that the parallel transport of an arbitrary covariant vector along a loop that can be contracted to a point reproduces the initial vector. Hence, in consequence of (6.25), the same is true for an arbitrary contravariant vector, and, in consequence of (6.30), for any scalar density. Then, in consequence of (6.18), when $B^{\alpha}{}_{\beta\gamma\delta} = 0$, *any* tensor density will return to its initial value when transported around any loop that can be contracted to a point.

The converse theorem also holds: if the parallel transport reproduces the initial tensor density *for every closed loop that can be contracted to a point*, then $B^{\alpha}{}_{\beta\gamma\delta} = 0$. This is because the integral in (6.40) is then zero for any arbitrary surface leaf S_C. We have thus

Theorem 6.1 *Parallel transport along any closed loop that can be contracted to a point reproduces the initial value of every tensor density if and only if the curvature tensor is zero.*

This is at the same time a necessary and sufficient condition for the result of parallel transport to be independent of the curve along which the transport is done.

The manifolds for which $B^{\alpha}{}_{\beta\gamma\delta} = 0$ are called **flat**.

6.3 Covariantly constant fields of vector bases.

Suppose a set of n vector fields exists on a manifold M_n such that at every point of M_n the vectors defined by these fields are linearly independent, and moreover all the fields are covariantly constant. Then ${e_a}^\alpha{}_{|\beta} \underset{a}{\equiv} 0, a = 1, \ldots, n$, and so

$$0 = {e_a}^\alpha{}_{|\gamma\delta} - {e_a}^\alpha{}_{|\delta\gamma} = -B^\alpha{}_{\rho\gamma\delta}{e_a}^\rho.$$

Contracting both sides of this with $e^a{}_\beta$ we obtain $B^\alpha{}_{\beta\gamma\delta} = 0$ as the necessary condition for the existence of such a field of bases.

We will show that this is also a sufficient condition. Assume that $B^\alpha{}_{\beta\gamma\delta} = 0$. Then the result of parallel transport does not depend on the path. Hence, with a fixed initial point and fixed initial data at that point, this operation will give a unique result at any other point of the manifold. So, choose an arbitrary point $x_0 \in M_n$, and in the tangent space to M_n at x_0 choose an arbitrary basis of vectors ${e_a}^\alpha(x_0)$. Then define bases in spaces tangent to M_n at other points as sets of vectors obtained from ${e_a}^\alpha(x_0)$ by parallel transport. This set of vector fields will be covariantly constant along every curve, and so covariantly constant on M_n.

The field of bases $e^a{}_\alpha$, dual to the covariantly constant field of bases ${e_a}^\alpha$, is also covariantly constant. Choosing such bases to calculate the connection coefficients we obtain

$$\Gamma^\alpha{}_{\beta\gamma} = {e_s}^\alpha e^s{}_{\beta,\gamma} = -e^s{}_\beta {e_s}^\alpha{}_{,\gamma}. \tag{6.41}$$

6.4 A torsion-free flat manifold.

If a manifold is not only flat, but also torsion-free, then for a covariantly constant field of bases we get from (6.41):

$$0 = {e_s}^\alpha \left(e^s{}_{\beta,\gamma} - e^s{}_{\gamma,\beta}\right). \tag{6.42}$$

Contracting this with $e^a{}_\alpha$ we obtain $e^a{}_{\beta,\gamma} - e^a{}_{\gamma,\beta} \underset{a}{\equiv} 0$. This equation implies that, in a neighbourhood of every point on the manifold, a set of scalar functions $\phi^a(x)$ exists such that $e^a{}_\beta = \phi^a{}_{,\beta}$, $a = 1, \ldots, n$. Since we assumed that the vectors $e^a{}_\alpha$ are linearly independent at every point, i.e. $\det ||e^a{}_\alpha|| \neq 0$, this now implies $\partial(\phi^1, \ldots, \phi^n)/\partial(x^1, \ldots, x^n) \neq 0$. This means that the connection between the functions $\{\phi^a\}_{a=1,\ldots,n}$ and the coordinates $\{x^\alpha\}$ is unique and thus reversible, and so the functions $\{\phi^a\}$ can be chosen as the new coordinates. Then let $x^{\alpha'}(x) = \phi^{\alpha'}(x)$. In these coordinates, $e^a{}_{\alpha'} = \phi^a{}_{,\alpha'} = \partial\phi^a/\partial\phi^{\alpha'} = \delta^a{}_{\alpha'}$ and ${e_a}^{\alpha'} = \partial\phi^{\alpha'}/\partial\phi^a = \delta^{\alpha'}{}_a$. Hence, in (6.41) we have $\Gamma^{\alpha'}{}_{\beta'\gamma'} = 0$.

In a flat torsion-free manifold one can thus choose such coordinates in which the connection coefficients are zero. In such coordinates (which we shall call Cartesian), the covariant derivatives reduce to ordinary partial derivatives. A special case (but not the only one) of a flat torsion-free manifold is a Euclidean space (of any dimension).

6.5 Parallel transport in a flat manifold.

Since parallel transport in a flat manifold does not depend on the path, the propagators of parallel transport are functions that depend only on the initial point $\overline{x}$ and on the final point x, but not on the curve that is chosen to join $\overline{x}$ to x. For covariantly constant fields of bases $\{e_A{}^\alpha\}$ and $\{e^A{}_\alpha\}$ we then have

$$P_w = [e(x)/e(\overline{x})]^w, \qquad P^\alpha{}_{\overline{\alpha}} = e_S{}^\alpha(x) e^S{}_{\overline{\alpha}}(\overline{x}), \qquad P^{\overline{\beta}}{}_\beta = e^S{}_\beta(x) e_S{}^{\overline{\beta}}(\overline{x}). \tag{6.43}$$

These equations follow from the fact that their right-hand sides obey the differential equations for propagators and the corresponding initial conditions.

If the manifold is torsion-free in addition, then the fields $e^A{}_\alpha$ are gradients of scalar functions. Choosing these functions as new coordinates we get

$$P_w(x, \overline{x}) = 1, \qquad P^\alpha{}_{\overline{\alpha}} = \delta^\alpha{}_{\overline{\alpha}}, \qquad P^{\overline{\beta}}{}_\beta = \delta_\beta{}^{\overline{\beta}}. \tag{6.44}$$

Thus, in a flat torsion-free manifold, in Cartesian coordinates, a vector that is transported parallely has the same components at every point, so the point to which the vector is attached becomes irrelevant.

6.6 Geodesic deviation.

As we will learn later, the inertial forces acting on a body moving along a geodesic cancel the gravitational forces. On the other hand, the geodesics are the only curves privileged by the geometry of a manifold with connection, hence only they can be used as a privileged reference system. Since a gravitational field cannot be detected on a geodesic, one can only observe neighbouring geodesics from the given one and thereby make conclusions about the gravitational field. The vector field of geodesic deviation is a measure of the position of a particle on a neighbouring geodesic with respect to a particle on a given geodesic.

Let $U \subset M_n$ be an open subset of a manifold M_n with connection. Choose in U a one-parameter family of geodesics G labelled by the parameter ϵ. Let $\epsilon = 0$ correspond to that geodesic from which we will be observing the neighbourhood. Every point on any $G(\epsilon)$ may be identified by the value of ϵ (which identifies the geodesic) and of the affine parameter s on $G(\epsilon)$, $x^\alpha|_U = x^\alpha(s, \epsilon)$. The **geodesic deviation** is the vector field

$$\delta x^\alpha(s) = \left.\frac{\partial x^\alpha}{\partial \epsilon}\right|_{\epsilon=0} \tag{6.45}$$

defined along $G(0)$. We shall find how $\delta x^\alpha(s)$ changes as we move down this geodesic. Since the set $\{x^\alpha(s, \epsilon)\}$ with fixed ϵ and changing s is a geodesic, the following is true

$$\left.\frac{D}{\mathrm{d}s}\frac{\partial x^\alpha}{\partial s}\right|_{\epsilon=\mathrm{const}} = 0. \tag{6.46}$$

Then, from (6.10)

$$\left(\frac{D}{\mathrm{d}s}\frac{D}{\mathrm{d}\epsilon} - \frac{D}{\mathrm{d}\epsilon}\frac{D}{\mathrm{d}s}\right)\frac{\partial x^\alpha}{\partial s} = -B^\alpha{}_{\rho\mu\nu}\frac{\partial x^\rho}{\partial s}\frac{\partial x^\mu}{\partial s}\frac{\partial x^\nu}{\partial \epsilon}. \tag{6.47}$$

But, in virtue of (6.46), the second term on the left is zero. Now we note that

$$\begin{aligned}\frac{D}{\mathrm{d}\epsilon}\frac{\partial x^\alpha}{\partial s} &= \frac{\partial^2 x^\alpha}{\partial\epsilon\partial s} + \Gamma^\alpha{}_{\rho\sigma}\frac{\partial x^\rho}{\partial s}\frac{\partial x^\sigma}{\partial\epsilon} = \frac{\partial^2 x^\alpha}{\partial s\partial\epsilon} + \Gamma^\alpha{}_{\sigma\rho}\frac{\partial x^\sigma}{\partial\epsilon}\frac{\partial x^\rho}{\partial s} + 2\Omega^\alpha{}_{\rho\sigma}\frac{\partial x^\rho}{\partial s}\frac{\partial x^\sigma}{\partial\epsilon} \\ &= \frac{D}{\mathrm{d}s}\frac{\partial x^\alpha}{\partial\epsilon} + 2\Omega^\alpha{}_{\rho\sigma}\frac{\partial x^\rho}{\partial s}\frac{\partial x^\sigma}{\partial\epsilon}. \end{aligned} \tag{6.48}$$

Making use of (6.48) in (6.47) we get

$$\frac{D^2}{\mathrm{d}s^2}\frac{\partial x^\alpha}{\partial\epsilon} + 2\frac{\partial x^\rho}{\partial s}\frac{D}{\mathrm{d}s}\left(\Omega^\alpha{}_{\rho\sigma}\frac{\partial x^\sigma}{\partial\epsilon}\right) = -B^\alpha{}_{\rho\mu\nu}\frac{\partial x^\rho}{\partial s}\frac{\partial x^\mu}{\partial s}\frac{\partial x^\nu}{\partial\epsilon}. \tag{6.49}$$

Denoting $k^\alpha = \frac{\partial x^\alpha}{\partial s}$ (the vector field tangent to the geodesic) and δx^α as in (6.45), then substituting $\epsilon = 0$ in the above, we finally get

$$\frac{D^2}{\mathrm{d}s^2}\delta x^\alpha + 2k^\rho\frac{D}{\mathrm{d}s}\left(\Omega^\alpha{}_{\rho\sigma}\delta x^\sigma\right) + B^\alpha{}_{\rho\mu\nu}k^\rho k^\mu\delta x^\nu = 0. \tag{6.50}$$

This is the **geodesic deviation equation**. It allows one (in principle – in practice this is extremely difficult) to determine the curvature using measurements of relative displacements of bodies moving along neighbouring geodesics.

In a flat torsion-free manifold and in the Cartesian coordinates, eq. (6.50) simplifies to $\mathrm{d}^2\left(\delta x^\alpha\right)/\mathrm{d}s^2 = 0$, so its solution is $\delta x^\alpha = A^\alpha s + B^\alpha$, where A^α and B^α are constant vectors. From (6.45) then, the position of the point under observation, $x^\alpha(s,\epsilon_0)$, relative to the point $x^\alpha(s,0)$ of the geodesic $\epsilon = 0$ is given by

$$x^\alpha(s,\epsilon_0) = x^\alpha(s,0) + \epsilon_0\delta x^\alpha. \tag{6.51}$$

Hence, in a flat torsion-free manifold two neighbouring geodesics either diverge or converge with a constant velocity, or stay parallel (when $A^\alpha = 0$). In curved manifolds and in manifolds with torsion, solutions of the geodesic deviation equation are more complicated. (Actually, it is rarely possible to find the solution explicitly. Usually, solutions are found numerically or by perturbations.)

The form of (6.50) does not change when ϵ is transformed by $\epsilon = \epsilon(\epsilon')$. The derivatives $\mathrm{d}\epsilon'/\mathrm{d}\epsilon$ are constant along geodesics, hence they will cancel out in (6.50). Changing the parametrisation only means choosing a different basis in the space of solutions of eq. (6.50).

6.7 Algebraic and differential identities obeyed by the curvature tensor.

We have already noted that $B^\alpha{}_{\beta\gamma\delta} = B^\alpha{}_{\beta[\gamma\delta]}$ from the definition (6.4). The curvature tensor obeys two other important sets of identities. Let us calculate

$$\begin{aligned} B^\alpha{}_{[\beta\gamma\delta]} &= \frac{1}{3!}\delta^{\rho\sigma\mu}_{\beta\gamma\delta}B^\alpha{}_{\rho\sigma\mu} = \frac{1}{3!}\delta^{\rho\sigma\mu}_{\beta\gamma\delta}\left(-2\Gamma^\alpha{}_{\rho[\sigma,\mu]} + 2\Gamma^\alpha{}_{\nu[\sigma}\Gamma^\nu{}_{|\rho|\mu]}\right) \\ &= \frac{1}{3}\delta^{\rho\sigma\mu}_{\beta\gamma\delta}\left(-\Gamma^\alpha{}_{\rho\sigma,\mu} + \Gamma^\alpha{}_{\nu\sigma}\Gamma^\nu{}_{\rho\mu}\right) = \frac{1}{3}\delta^{\rho\sigma\mu}_{\beta\gamma\delta}\left(-\Omega^\alpha{}_{\rho\sigma,\mu} + \Gamma^\alpha{}_{\nu\sigma}\Omega^\nu{}_{\rho\mu}\right) \end{aligned}$$

$$= \frac{1}{3}\delta^{\rho\sigma\mu}_{\beta\gamma\delta}\left(-\Omega^{\alpha}{}_{\rho\sigma|\mu} + \Gamma^{\alpha}{}_{\nu\mu}\Omega^{\nu}{}_{\rho\sigma} - \Gamma^{\nu}{}_{\rho\mu}\Omega^{\alpha}{}_{\nu\sigma} - \Gamma^{\nu}{}_{\sigma\mu}\Omega^{\alpha}{}_{\rho\nu} + \Gamma^{\alpha}{}_{\nu\sigma}\Omega^{\nu}{}_{\rho\mu}\right). \quad (6.52)$$

In the transformations above we used the fact that in the contraction over three indices with the antisymmetric $\delta^{\rho\sigma\mu}_{\beta\gamma\delta}$ the antisymmetrisation with respect to $\sigma\mu$ occurs automatically, and we expressed the partial derivative of $\Omega^{\alpha}{}_{\rho\sigma}$ through the corresponding covariant derivative. Now let us note that by interchanging the names of the indices σ and μ and interchanging their positions in the $\delta^{\cdots}_{\cdots}$ contracted with the last term above, the last term turns out to be equal to the second with an opposite sign. Thus we have

$$\begin{aligned} B^{\alpha}{}_{[\beta\gamma\delta]} &= \frac{1}{3}\delta^{\rho\sigma\mu}_{\beta\gamma\delta}\left(-\Omega^{\alpha}{}_{\rho\sigma|\mu} - \Omega^{\nu}{}_{\rho\mu}\Omega^{\alpha}{}_{\nu\sigma} - \Omega^{\nu}{}_{\sigma\mu}\Omega^{\alpha}{}_{\rho\nu}\right) \\ &= \frac{1}{3}\delta^{\rho\sigma\mu}_{\beta\gamma\delta}\left(-\Omega^{\alpha}{}_{\rho\sigma|\mu} - \Omega^{\alpha}{}_{\nu\sigma}\Omega^{\nu}{}_{\rho\mu} + \Omega^{\alpha}{}_{\sigma\nu}\Omega^{\nu}{}_{\rho\mu}\right) \\ &= \frac{1}{3}\delta^{\rho\sigma\mu}_{\beta\gamma\delta}\left(-\Omega^{\alpha}{}_{\rho\sigma|\mu} - 2\Omega^{\alpha}{}_{\nu\sigma}\Omega^{\nu}{}_{\rho\mu}\right) = -2\Omega^{\alpha}{}_{[\beta\gamma|\delta]} - 4\Omega^{\alpha}{}_{\nu[\gamma}\Omega^{\nu}{}_{\beta\delta]}. \quad (6.53) \end{aligned}$$

The curvature tensor obeys also the following set of differential identities:

$$B^{\alpha}{}_{\beta[\gamma\delta|\epsilon]} = -2\Omega^{\nu}{}_{[\gamma\delta}B^{\alpha}{}_{|\beta|\epsilon]\nu}, \quad (6.54)$$

called the **Bianchi identities**. Their derivation is left as a Exercise 3 for the readers.

Recall that the curvature tensor arose by calculating the commutator of covariant derivatives. Commutators of linear operators obey the well-known Jacobi identity. The Bianchi identities arise from the commutators of covariant derivatives in the same way as the Jacobi identity arises from commutators of linear operators.

The Bianchi identities are important for the physical interpretation of the relativity theory. They assure that equations of motion of material media follow from the field equations and need not be postulated separately.

6.8 Exercises

1. Observe that coordinates can always be adapted to *one* vector field k^{α} so that, in the new coordinates, the field has only one nonzero component, and this one component is equal to 1. For example, let that preferred component be $k'^1 = 1$; to achieve this one has to solve the set of equations $k^{\alpha}x^{\alpha'}{}_{,\alpha} = \delta^{\alpha'}_1$ where $x^{\alpha'}(x)$ are the unknown functions. Calling $x'^1 = \tau$ and taking the inverse transformation we get $k^{\alpha} = (\partial x^{\alpha}) / (\partial \tau)$.

2. Using the result of the previous exercise show that coordinates can be simultaneously adapted to two vector fields k^{α} and l^{α} so that $k'^{\alpha} = \delta^{\alpha}_1$ and $l'^{\alpha} = \delta^{\alpha}_2$ if and only if the two vector fields commute, $[k,l]^{\alpha} \overset{\text{def}}{=} k^{\rho}l^{\alpha}{}_{,\rho} - l^{\rho}k^{\alpha}{}_{,\rho} = 0$. It follows that if i vector fields all commute, then coordinates can be adapted simultaneously to all of them.

Hint: That this is a necessary condition (i.e. adapted coordinates exist $\Longrightarrow [k,l] = 0$) is easy to prove. To show that it is also sufficient, adapt the coordinates to k^{α}, then find the transformations of coordinates that preserve the property $k'^{\alpha} = \delta_1^{\alpha}$, and, using only these transformations, try to adapt the coordinates to l^{α}. You will find that if the commutator is nonzero, then this would require solving a set of n equations for n functions of $(n-1)$ variables, but the coefficients of the set depend on all the n variables.

3. Derive the Bianchi identities (6.54).

Hint: use Kronecker 3-deltas to calculate antisymmetrisations, in the same way in which they were used in deriving (6.53).

Chapter 7

Riemannian geometry.

Excellent sources for the subject of this chapter are still [7] and [8], in spite of their age. Ref. [7] is a presentation of differential geometry of curves and two-dimensional surfaces in a Euclidean space, Ref. [8] is a textbook on Riemannian geometry in n dimensions.

7.1 The metric tensor.

Up to here, we have dealt with such manifolds on which the only additional object were the affine connection coefficients. That structure allowed us to define the parallel transport and thereby to compare directions of vectors attached to different points. However, we have had no means to calculate distances between points or angles between vectors.

Now we will add a new object that will allow for just that: a symmetric second-rank covariant tensor, $g_{\alpha\beta} = g_{(\alpha\beta)}$, called the **metric tensor**. Using it, we can define the **metric form**, also called **metric**,

$$\mathrm{d}s^2 = g_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta, \tag{7.1}$$

which is a convenient shorthand notation for an expression (formally a scalar) that makes sense only under an integral, e.g. along a curve $x^\alpha(\lambda)$:

$$l_{\lambda_0\lambda_1} = \int_{\lambda_0}^{\lambda_1} \mathrm{d}s = \int_{\lambda_0}^{\lambda_1} \left| g_{\alpha\beta}(\lambda) \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} \right|^{1/2} \mathrm{d}\lambda. \tag{7.2}$$

(We take the absolute value above because we do not assume that the form is positive-definite. We will explain later the meaning of the different possible signs of this form.) The $l_{\lambda_0\lambda_1}$ is the length of the arc of the curve $x^\alpha(\lambda)$ between the points $x^\alpha(\lambda_0)$ and $x^\alpha(\lambda_1)$.

An example of a manifold with a metric tensor is a Euclidean space (of arbitrary dimension). Its metric tensor in rectangular Cartesian coordinates is the unit matrix, and its metric form is

$$\mathrm{d}s^2 = \left(\mathrm{d}x^1\right)^2 + \left(\mathrm{d}x^2\right)^2 + \cdots + \left(\mathrm{d}x^n\right)^2. \tag{7.3}$$

[7] L. P. Eisenhart, *An introduction to differential geometry with use of tensor calculus.* Princeton University Press, Princeton 1940.

[8] L. P. Eisenhart, *Riemannian geometry.* Princeton University Press, Princeton 1964.

7.2 Riemann spaces.

It is logical to require that any vector k^α preserves its length when parallely transported along an arbitrary curve in the affine parametrisation. (This cannot be required with just any parametrisation because even the tangent vector to a geodesic would not obey this in general.) What condition does such a requirement impose on the metric tensor?

Let s be an affine parameter. The length of a parallely transported vector $k^\alpha(s)$ will not change when

$$0 = \frac{\mathrm{d}}{\mathrm{d}s}\left(g_{\alpha\beta}k^\alpha k^\beta\right) = \frac{D}{\mathrm{d}s}\left(g_{\alpha\beta}k^\alpha k^\beta\right) = \frac{Dg_{\alpha\beta}}{\mathrm{d}s}k^\alpha k^\beta \tag{7.4}$$

(since $Dk^\alpha/\mathrm{d}s = 0$ under parallel transport). If (7.4) holds for every vector field k^α, then

$$0 = \frac{Dg_{\alpha\beta}}{\mathrm{d}s} = \frac{\mathrm{d}x^\rho}{\mathrm{d}s}g_{\alpha\beta|\rho}. \tag{7.5}$$

Now, this equation should hold for every curve, which implies in the end

$$g_{\alpha\beta|\rho} = 0, \tag{7.6}$$

or, explicitly

$$g_{\alpha\beta,\rho} - \Gamma^\sigma{}_{\alpha\rho}g_{\sigma\beta} - \Gamma^\sigma{}_{\beta\rho}g_{\alpha\sigma} = 0. \tag{7.7}$$

Those manifolds, on which a metric tensor obeying (7.6) is defined, and on which in addition

$$\Omega^\alpha{}_{\beta\gamma} = 0 \Longrightarrow \Gamma^\alpha{}_{\beta\gamma} = \Gamma^\alpha{}_{(\beta\gamma)} \tag{7.8}$$

are called **Riemann spaces**, they will be denoted by V_n. For brevity, we shall use the same name also for those manifolds on which the metric tensor is not positive-definite, although this does not agree with the general habit of mathematicians. In mathematics, those manifolds for which $g_{\alpha\beta}$ is not positive-definie are called *pseudoriemannian* in general, and sometimes have special names, e.g. the metric used in the relativity theory (both special and general) is called **Lorentzian**.

7.3 Signature of a metric, degenerate metrics.

The expressions $g_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta$ and $g_{\alpha\beta}k^\alpha k^\beta$, considered in the previous section, are quadratic forms. A coordinate transformation $\{x\} \to \{x'\}$ causes that the form $g_{\alpha\beta}k^\alpha k^\beta$ changes into $g_{\alpha'\beta'}k^{\alpha'}k^{\beta'}$, where $k^{\alpha'} = x^{\alpha'}{}_{,\alpha}\,k^\alpha$ and $g_{\alpha'\beta'} = x^\alpha{}_{,\alpha'}\,x^\beta{}_{,\beta'}\,g_{\alpha\beta}$. Let X denote the matrix $\left|\left|x^{\alpha'}{}_{,\alpha}\right|\right|$, let g denote the matrix $||g_{\alpha\beta}||$, and let k denote the vector k^α. In the terminology of quadratic forms, the transition $k \to k' = Xk$, $g \to g' = \left(X^{-1}\right)^T gX^{-1}$ is called the transformation of a basis. For matrices of quadratic forms, the Sylvester *theorem on the inertia of forms* holds. It says that whatever method is used to diagonalise a real symmetric matrix, the numbers of diagonal elements that are positive, zero and negative are always the same. This set of positive, negative and zero values can be denoted by the symbol

$$(\underbrace{+,\ldots,+}_{n_1\ \text{times}},\underbrace{-,\ldots,-}_{n_2\ \text{times}},\underbrace{0,\ldots,0}_{n_3\ \text{times}}), \tag{7.9}$$

where n_1 is the number of positive elements, n_2 is the number of negative elements, and n_3 is the number of zero elements. Of course, $n_1 + n_2 + n_3 = n = \dim V_n$. The symbol (7.9) is a basis-independent characteristic property of every matrix, and so in the case of the metric tensor – a coordinate-independent characteristic of a point of a Riemann space. This symbol is called the **signature** of the metric tensor. When $n_3 > 0$, the metric is called degenerate. When $n_2 = n_3 = 0$, the metric is positive-definite.

In the relativity theory, 4-dimensional Riemann spaces are used, with the signature $(+---)$ or $(-+++)$ or $(+++-)$, depending on the convention (in this text, the signature will always be $(+---)$). Since such a metric is not positive-definite, the Riemann spaces of relativity are not metric spaces. If the distance $d(x,y)$ between the points $x \in V_n$ and $y \in V_n$ is defined by (7.2) along any curve (e.g. a geodesic), then $d(x,y) = 0$ does not imply $x = y$. The integral can vanish also for such (x,y) pairs that do not coincide. The geometrical and physical meaning of $l_{\lambda_0\lambda_1} = 0$ in (7.2) will be explained further.

Coordinate transformations preserve the signature of the metric tensor at any single point, but no theorem guarantees that the signature must be the same at all points of the manifold. However, a region with a signature different from $(+---)$ (or any of its equivalent counterparts) would have no physical interpretation. Usually, the subset on which the signature would change to an unphysical one is, in one sense or another, an "edge" of the manifold or of the allowed coordinate patch. However, transitions of the type $(+---) \to (-+--)$ do occur, in which two coordinates interchange their roles and another becomes the privileged one. This happens, for example, at the horizon of a black hole (see Sec. 11.8).

For a degenerate metric, the determinant of $g_{\alpha\beta}$ is zero and the inverse matrix to $g_{\alpha\beta}$ does not exist. Then, the mapping $k^\alpha \to g_{\alpha\beta}k^\beta$ is analogous to projecting a vector on a subspace of lower dimension, so no inverse mapping exists. For nondegenerate metrics, the matrix inverse to $g_{\alpha\beta}$ does exist, and is denoted $g^{\alpha\beta}$; it obeys the equation $g^{\alpha\rho}g_{\rho\beta} = \delta^\alpha{}_\beta$. Thus, in a Riemann space with a nondegenerate metric, for every contravariant vector k^α there exists the corresponding covariant vector $g_{\alpha\beta}k^\beta$, denoted k_α, and for every covariant vector m_α there exists the corresponding contravariant vector $g^{\alpha\beta}m_\beta$, denoted m^α. The mapping $k^\alpha \to k_\alpha = g_{\alpha\beta}k^\beta$ is called **lowering the index**, the inverse mapping $m_\alpha \to m^\alpha = g^{\alpha\beta}m_\beta$ is called **raising the index**. In such a Riemann space, we can thus consider contravariant and covariant *components* of the same vector, while in a manifold with no metric there is no relation between covariant and contravariant vectors.

7.4 Christoffel symbols.

Now we shall solve eq. (7.7) for $\Gamma^\alpha{}_{\beta\gamma}$ assuming that the metric tensor is nondegenerate. Let us rewrite (7.7) 3 times, each time with a different permutation of the indices:

$$g_{\alpha\beta,\rho} - \Gamma^\sigma{}_{\alpha\rho}g_{\sigma\beta} - \Gamma^\sigma{}_{\beta\rho}g_{\alpha\sigma} = 0, \tag{7.10}$$
$$g_{\beta\rho,\alpha} - \Gamma^\sigma{}_{\beta\alpha}g_{\sigma\rho} - \Gamma^\sigma{}_{\rho\alpha}g_{\beta\sigma} = 0, \tag{7.11}$$
$$g_{\rho\alpha,\beta} - \Gamma^\sigma{}_{\rho\beta}g_{\sigma\alpha} - \Gamma^\sigma{}_{\alpha\beta}g_{\rho\sigma} = 0. \tag{7.12}$$

Let us calculate the combination (7.11) + (7.12) − (7.10). Using (7.8) and contracting the result with $g^{\gamma\rho}$ we obtain

$$\Gamma^{\gamma}{}_{\alpha\beta} = \frac{1}{2} g^{\gamma\rho} \left(g_{\alpha\rho,\beta} + g_{\beta\rho,\alpha} - g_{\alpha\beta,\rho} \right). \tag{7.13}$$

The affine connection coefficients built of the metric tensor in this way are called **Christoffel symbols** and are denoted $\left\{ {\gamma \atop \alpha\beta} \right\}$.

7.5 Curvature of a Riemann space.

The curvature tensor of a Riemann space, built of the Christoffel symbols standing in place of the connection coefficients, is called the **Riemann tensor** and denoted $R^{\alpha}{}_{\beta\gamma\delta}$. Since the Riemann spaces are torsion-free, the identities (6.53) and (6.54) reduce here to

$$R^{\alpha}{}_{[\beta\gamma\delta]} = 0, \tag{7.14}$$

$$R^{\alpha}{}_{\beta[\gamma\delta;\epsilon]} = 0 \tag{7.15}$$

(the covariant derivative in a Riemann space will be denoted by a semi-colon).

The Riemann tensor obeys one more set of identities. Using (6.10) and (7.5) we obtain

$$0 = g_{\alpha\beta;\gamma\delta} - g_{\alpha\beta;\delta\gamma} = R^{\rho}{}_{\alpha\gamma\delta} g_{\rho\beta} + R^{\rho}{}_{\beta\gamma\delta} g_{\alpha\rho}. \tag{7.16}$$

This can be written as

$$R_{\beta\alpha\gamma\delta} + R_{\alpha\beta\gamma\delta} = 0 \Longleftrightarrow R_{\alpha\beta\gamma\delta} = R_{[\alpha\beta]\gamma\delta}. \tag{7.17}$$

From the identities written above, one more identity can be derived. Equation (7.14), in consequence of $R_{\alpha\beta\gamma\delta} = R_{\alpha\beta[\gamma\delta]}$, can be rewritten in the form

$$R_{\alpha\beta\gamma\delta} + R_{\alpha\gamma\delta\beta} + R_{\alpha\delta\beta\gamma} = 0. \tag{7.18}$$

Let us rewrite (7.18) three more times, taking cyclic permutations of all four indices

$$\begin{aligned} R_{\beta\gamma\delta\alpha} + R_{\beta\delta\alpha\gamma} + R_{\beta\alpha\gamma\delta} &= 0, && (7.19)\\ R_{\gamma\delta\alpha\beta} + R_{\gamma\alpha\beta\delta} + R_{\gamma\beta\delta\alpha} &= 0, && (7.20)\\ R_{\delta\alpha\beta\gamma} + R_{\delta\beta\gamma\alpha} + R_{\delta\gamma\alpha\beta} &= 0. && (7.21) \end{aligned}$$

Now let us calculate the combination (7.18) + (7.21) − (7.19) − (7.20). Taking into account (7.17) and the antisymmetry in the last pair of indices, we get $2\left(R_{\alpha\beta\gamma\delta} - R_{\gamma\delta\alpha\beta}\right) = 0$, i.e.

$$R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta}, \tag{7.22}$$

so the Riemann tensor is symmetric with respect to the interchange of the first pair of indices with the second.

The tensor $R_{\alpha\beta\gamma\delta}$ has n^4 components, i.e. 256 in a 4-dimensional space. However, in consequence of its symmetries, only some of the components are algebraically independent.

Antisymmetry in (α,β) provides $n(n+1)/2$ equations $R_{(\alpha\beta)\gamma\delta}=0$ for each set (γ,δ), which leaves us with at most $n^4-n^2\times n(n+1)/2=n^2\times n(n-1)/2$ independent components. Because of antisymmetry in (γ,δ), we get $[n(n+1)/2]\times[n(n-1)/2]$ additional equations $R_{\alpha\beta(\gamma\delta)}=0$, which leaves us with at most $[n(n-1)/2]^2$ independent components. From this, we have to subtract the $n\begin{pmatrix} n \\ 3 \end{pmatrix}$ equations (7.14). This gives finally

$$[n(n-1)/2]^2 - n\begin{pmatrix} n \\ 3 \end{pmatrix} = \frac{1}{12}n^2\left(n^2-1\right) \tag{7.23}$$

independent components, i.e. only 20 when $n=4$.

7.6 Flat Riemann spaces.

If $R^{\alpha}{}_{\beta\gamma\delta}=0$, then the results of Section 6.4 apply – one can choose coordinates so that $\left\{ {\alpha \atop \beta\gamma} \right\} \underset{*}{=} 0$. This equation is preserved by all linear coordinate transformations. In such coordinates, as seen from (7.7), $g_{\alpha\beta}$ has constant components. Then, further linear coordinate transformations can be used so that $g_{\alpha\beta}$ becomes diagonal and all its diagonal elements are either $+1$, -1 or 0, thus

$$\mathrm{d}s^2 \underset{*}{=} \sum_{i=1}^{n} \epsilon_i \left(dx^i\right)^2,$$

where $\epsilon_i=+1,0$ or -1.

7.7 Subspaces of a Riemann space.

Let a subspace S_m of a Riemann space V_n be given by the equations $x^\alpha = f^\alpha\left(\tau^1,\tau^2,\ldots,\tau^m\right)$, $\alpha=1,\ldots,n$. For the pair of points $\{\tau_0^a\}\in S_m$ and $\{\tau_0^a+\mathrm{d}\tau^a\}\in S_m$ we have

$$g_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta = g_{\alpha\beta}\frac{\partial x^\alpha}{\partial\tau^a}\frac{\partial x^\beta}{\partial\tau^b}\mathrm{d}\tau^a\mathrm{d}\tau^b \tag{7.24}$$

Hence, the metric tensor $g_{\alpha\beta}$ of V_n **induces** the metric tensor h_{ab} in S_m by the formula

$$h_{ab} = g_{\alpha\beta}\frac{\partial x^\alpha}{\partial\tau^a}\frac{\partial x^\beta}{\partial\tau^b}. \tag{7.25}$$

In a 3-dimensional Euclidean space, in rectangular Cartesian coordinates, the length of a curve is

$$l_{\lambda_0,\lambda_1} = \int_{\lambda_0}^{\lambda_1}\left[\left(\frac{\mathrm{d}x}{\mathrm{d}\lambda}\right)^2+\left(\frac{\mathrm{d}y}{\mathrm{d}\lambda}\right)^2+\left(\frac{\mathrm{d}z}{\mathrm{d}\lambda}\right)^2\right]^{1/2}\mathrm{d}\lambda. \tag{7.26}$$

Thus, the metric form here is $\mathrm{d}s^2=\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2$, and the metric tensor is the unit matrix. Knowing this, we can use (7.25) to find the metric tensor of an arbitrary 2-dimensional surface in the Euclidean space. For example, for a sphere of radius a, with the centre at $x=y=z=0$, the parametric equations in spherical coordinates are:

$$x = a \sin\vartheta \cos\varphi, \qquad y = a \sin\vartheta \sin\varphi, \qquad z = a \cos\vartheta. \tag{7.27}$$

Consequently,

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = a^2 \left(\mathrm{d}\vartheta^2 + \sin^2\vartheta \mathrm{d}\varphi^2\right), \tag{7.28}$$

and the associated metric tensor is the matrix

$$||h_{ab}|| = \begin{pmatrix} a^2 & 0 \\ 0 & a^2 \sin^2\vartheta \end{pmatrix}. \tag{7.29}$$

7.8 The Riemann curvature vs. the normal curvature of a surface.

In an n-dimensional Riemann space, the curvature tensor has $n^2(n^2-1)/12$ algebraically independent components, which makes just 1 when $n=2$. Hence, the curvature of a 2-dimensional surface is determined by one function, for example $g^{\alpha\beta} R^{\rho}{}_{\alpha\rho\beta}$. This quantity is the Gauss curvature, whose description is given below. For the full reasoning see Ref. [7].

I assume the reader knows the definition of the curvature of a curve. Now consider a 2-dimensional surface S and a point $p \in S$. Draw the straight line L through p perpendicular to S, and then consider any plane F containing L. The plane intersects S along a curve C called **normal section** of S. Imagine F being rotated around L. It can happen that the curvatures of all the normal sections will be equal (like at every point of a sphere, or when L is a symmetry axis of S). But in general the curvature changes when F is rotated, and in the collection of all curvatures there is a maximal and a minimal value. (If the point p is nonsingular, then all the curvatures will be finite.) The Gauss curvature of S at p is the product of the greatest curvature of C by the smallest one. When all curvatures are equal, the curvature of S at p is the square of the curvature of each normal section. Knowing this, one can verify that the Gauss curvature of S at p is $g^{\alpha\beta} R^{\rho}{}_{\alpha\rho\beta}$. This shows that the name "curvature tensor" evokes the right association.

Examples:

(1) The curvature of a cylinder is equal to zero: at its any point one of the normal sections is a straight line whose curvature is zero, while all the other normal sections have positive curvatures. Hence, the smallest curvature of a normal section is zero.

(2) On a one-sheeted hyperboloid two straight lines are among the normal sections at each point. However, the other normal sections have positive and negative curvatures, so the zero curvature is neither maximal nor minimal. Consequently, the curvature of a one-sheeted hyperboloid is negative (see Exercise 1).

7.9 The geodesic line as the line of extremal distance.

If the length of a curve arc is defined on a manifold, then in the collection of all arcs joining two given points we can look for the arc of greatest or smallest length. Thus, we look for a curve on which, with fixed λ_0 and λ_1, the quantity

$$\int_{\lambda_0}^{\lambda_1} \left| g_{\rho\sigma}(x) \frac{\mathrm{d}x^\rho}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\sigma}{\mathrm{d}\lambda} \right|^{1/2} \mathrm{d}\lambda \tag{7.30}$$

takes the extremal value. This curve must obey the Euler – Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \left(\frac{\partial f}{\partial \dot{x}^\gamma} \right) - \frac{\partial f}{\partial x^\gamma} = 0, \tag{7.31}$$

where f is the integrand in (7.30), and $\dot{x}^\gamma \stackrel{\text{def}}{=} \mathrm{d}x^\gamma/\mathrm{d}\lambda$. For convenience, let us write

$$|g_{\rho\sigma}\dot{x}^\rho \dot{x}^\sigma| = \varepsilon g_{\rho\sigma}\dot{x}^\rho \dot{x}^\sigma, \tag{7.32}$$

where $\varepsilon = +1$ or $\varepsilon = -1$, as appropriate. Then, from (7.31)

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \left(\frac{g_{\gamma\sigma}\dot{x}^\sigma}{\sqrt{\varepsilon g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}} \right) - \frac{1}{2} \frac{g_{\rho\sigma,\gamma}\dot{x}^\rho \dot{x}^\sigma}{\sqrt{\varepsilon g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}} = 0. \tag{7.33}$$

The case $g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = 0$ requires separate treatment. We shall do this further on; for now we assume that $g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu \neq 0$. Let us introduce the new parameter $s(\lambda)$ defined by

$$\frac{\mathrm{d}s}{\mathrm{d}\lambda} = \sqrt{\varepsilon g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}. \tag{7.34}$$

Then we have in (7.33):

$$\frac{\mathrm{d}s}{\mathrm{d}\lambda} \frac{\mathrm{d}}{\mathrm{d}s} \left(g_{\gamma\sigma} \frac{\mathrm{d}x^\sigma}{\mathrm{d}s} \right) - \frac{1}{2} g_{\rho\sigma,\gamma} \frac{\mathrm{d}s}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\rho}{\mathrm{d}s} \frac{\mathrm{d}x^\sigma}{\mathrm{d}s} = 0. \tag{7.35}$$

Hence

$$g_{\gamma\sigma,\rho} \frac{\mathrm{d}x^\rho}{\mathrm{d}s} \frac{\mathrm{d}x^\sigma}{\mathrm{d}s} + g_{\gamma\sigma} \frac{\mathrm{d}^2 x^\sigma}{\mathrm{d}s^2} - \frac{1}{2} g_{\rho\sigma,\gamma} \frac{\mathrm{d}x^\rho}{\mathrm{d}s} \frac{\mathrm{d}x^\sigma}{\mathrm{d}s} = 0. \tag{7.36}$$

After contracting this with $g^{\alpha\gamma}$ we get

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}s^2} + \left\{ {\alpha \atop \rho\sigma} \right\} \frac{\mathrm{d}x^\rho}{\mathrm{d}s} \frac{\mathrm{d}x^\sigma}{\mathrm{d}s} = 0. \tag{7.37}$$

This is the geodesic equation (5.13). Thus, in Riemann spaces, a geodesic has another characteristic property: it extremises the arc length.

Note that the geodesic extremises not only (7.30), but also the functional $\int_{\lambda_0}^{\lambda_1} g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu \mathrm{d}\lambda$. For this functional, the Euler – Lagrange equations again lead to (7.37), and the assumption $g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu \neq 0$ is not necessary.

7.10 Timelike, null and spacelike intervals in a 4-dimensional spacetime.

Physically important are 4-dimensional Riemann spaces with the signature $(+ - - -)$, called **spacetimes**. Consider the following equation in a spacetime:

$$\mathrm{d}s^2 = g_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta = 0. \tag{7.38}$$

Let us choose a point $C \in V_4$. Apart from special cases which we will not consider, we can choose coordinates so that at C

$$g_{\alpha\beta}(C) = \begin{bmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix},$$

i.e. the metric becomes the Minkowski metric of special relativity. (Note: this happens only at C, but the metric becomes approximately Minkowskian in a small neighbourhood of C.) In these coordinates, called **locally Cartesian**, eq. (7.38) taken at C becomes

$$\mathrm{d}s^2 = \left(\mathrm{d}x^0\right)^2 - \left(\mathrm{d}x^1\right)^2 - \left(\mathrm{d}x^2\right)^2 - \left(\mathrm{d}x^3\right)^2 = 0. \tag{7.39}$$

In a 3-dimensional Euclidean space, the equation $(z - z_0)^2 - (x - x_0)^2 - (y - y_0)^2 = 0$ describes a cone with the vertex at (x_0, y_0, z_0) and the axis parallel to the z-axis. By analogy, the hypersurface in spacetime determined by (7.38) is called a **light cone** (see Fig. 7.1). The x^0 is called the **time coordinate**, (x^1, x^2, x^3) are **space coordinates**. The light cone divides all of spacetime, relative to C, into three disjoint regions. Each point lying on the light cone can be connected to C by a geodesic arc of zero length. All curves of zero length, geodesic or not, are called **null curves**, and the points on the cone are said to be in a **null relation** to C. The tangent vector to a null curve at any point has zero length, such vectors are called **null vectors**. Each point in the regions F and P inside the cone can be connected to C by a curve $x^\alpha(\lambda)$ on which $g_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} > 0$ everywhere. These points are said to be in a **timelike relation** with C, and the vectors of positive length are called **timelike vectors**. Finally, each point in the region E outside the light cone can be connected to C by a curve on which $g_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} < 0$ everywhere. These points are said to be in a **spacelike relation** to C, and vectors with negative length are called **spacelike vectors**. Justifications of these names come from special relativity. The region F, in which the x^0 coordinates of all the points are greater than the x^0-coordinate of C is called the **future** of C. The region P is called the **past** of C. Finally, the region E does not have a name and is colloquially called "elsewhere" with respect to C.

Not every curve lying on a light cone has zero length. Only the generators of the cone, which are **null geodesics**, have this property. Other curves on the cone, for example spirals winding on its surface toward C, are spacelike curves whose tangent vectors ℓ^α have $g_{\alpha\beta}\ell^\alpha\ell^\beta < 0$. Curves that are null but not geodesic are at each point q tangent to the light cone of q, but veer from one cone to another. Fig. 7.2 shows a broken null line whose straight segments are null and geodesic, but at the corners it goes from one cone to another. A general nongeodesic null line can be imagined as a limit of a sequence of such broken null lines as the extent of each geodesic segment (as measured, for example, by the range of the affine parameter) goes to zero. Such null lines through C enter the region F, or reach C from within P.

Similarly, not every curve arc in the regions F and P will be timelike. These regions contain also null nongeodesic curves, such as the one described above, and spacelike curves. The characteristic property of the region $F \cup P$ is that for each of its points a timelike curve joining it to C *exists*. Such curves do not exist on the light cone and in E. Curves

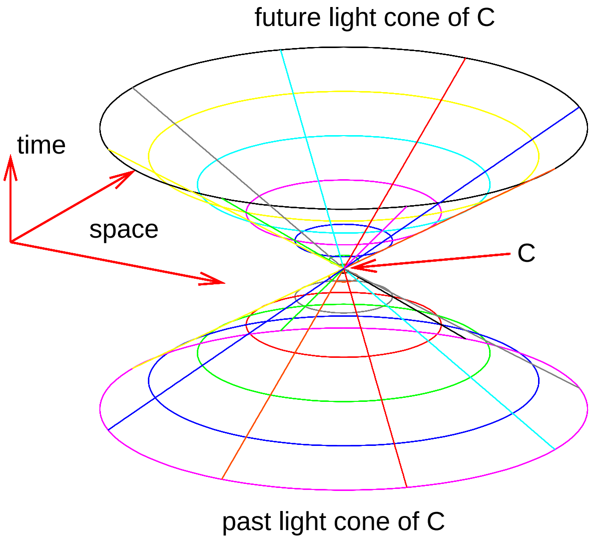

Figure 7.1: Equation (7.38) determines a light cone that divides the spacetime relative to C into the future of C, the past of C and the "elsewhere". A light cone looks so simple only in the flat spacetime in which coordinates with the property (7.39) can be introduced globally.

on the light cone that reach C are either null or spacelike, and every curve arc in E that reaches C will have at least a segment that is spacelike.

It should now be clear that a light cone exists at every point of a spacetime, and the analysis above applies to each point. The equation (7.38) is covariant, and so the light cone is a geometric object, independent of the choice of coordinates.

The equation of a light cone can be transformed to the form (7.39) only at each point separately. Light cones in curved spacetimes are not as simple as in the Minkowski space – in general they are not axially symmetric, and their spacelike diameters do not uniformly increase with x^0. They can have self-intersections (caustics) and the structure of a multi-sheeted surface. Figures 7.3 and 7.4 show examples of such deformed light cones. The left panel of Fig. 7.3 shows the shape of the past light cone of an observer in a spacetime that is now believed to be the large-scale model of our real Universe.

7.11 Embedding a Riemann space V_n in a Riemann space V_{n+1}

This section presents a simplified version of the embedding conditions, adapted to the problem that is often considered in relativity, namely, matching regions of two spacetimes across a 3-dimensional hypersurface. In that situation, $n = 3$. However, the generalisation to arbitrary n is nearly trivial, so we consider it here.

Let V_n be a subspace of the Riemann space V_{n+1}, and let the metric tensor of V_{n+1} be $g_{\alpha\beta}$. Let V_n be defined in V_{n+1} by the parametric equations $x^\alpha = f^\alpha(\tau^1, \ldots, \tau^n)$. Then, as we observed in Section 7.7, V_n is itself a Riemann space with the metric tensor (7.25).

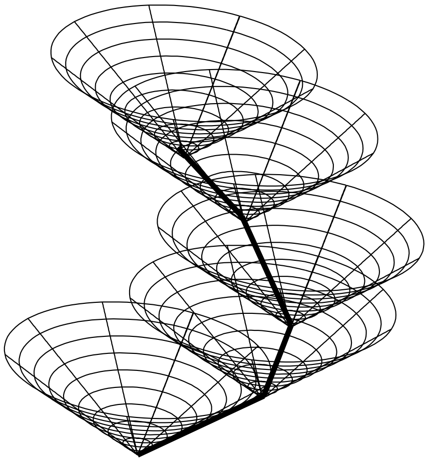

Figure 7.2: Future light cones along a non-geodesic null line with geodesic segments (the thicker line). The line is tangent to each light cone at its vertex, but passes from one cone to another as it proceeds.

Now we ask the reciprocal question: When can a given Riemann space V_n be a subspace of a Riemann space V_{n+1}? We do not consider *null subspaces* (i.e. V_n-s in which some tangent vectors have zero length) because in them the determinant of the metric is zero, and they require a separate treatment. Also, we will investigate only *local embeddings*, in which an open subset of V_n can be embedded in V_{n+1}. Global embeddings pose additional problems that we will not discuss. If V_n is a subspace of V_{n+1}, then a set of functions

$$x^\alpha = f^\alpha\left(\tau^1, \ldots, \tau^n\right), \qquad \alpha = 1, \ldots, n+1, \tag{7.40}$$

should exist such that

$$h_{ab} = g_{\alpha\beta} x^\alpha{}_{,a}\, x^\beta{}_{,b}\,, \tag{7.41}$$

where τ^a are coordinates in V_n, $x^\alpha{}_{,a} \overset{\text{def}}{=} \partial x^\alpha / \partial \tau^a$ and h_{ab} is the metric tensor of V_n. Equations (7.40) are parametric equations of our V_n as a subspace of V_{n+1}. Note that x^α and $g_{\alpha\beta}$ are scalars with respect to coordinate transformations in V_n.

The question of whether a given V_n can be embedded in a V_{n+1} can be answered with the help of the reasoning presented below (see Ref. [8] for a more detailed exposition). Let X^α be a vector field on V_{n+1} that is orthogonal to V_n and has unit length:

$$g_{\alpha\beta} X^\alpha X^\beta = \varepsilon = \pm 1. \tag{7.42}$$

Since $x^\alpha{}_{,a}$ are tangent to V_n and X^α is orthogonal to V_n, we have

$$g_{\alpha\beta} x^\alpha{}_{,a}\, X^\beta = 0 \tag{7.43}$$

for all a. We differentiate (7.41) covariantly by τ^c. (From now on the semi-colon will denote covariant derivatives in V_n. We have to watch the difference between these and the

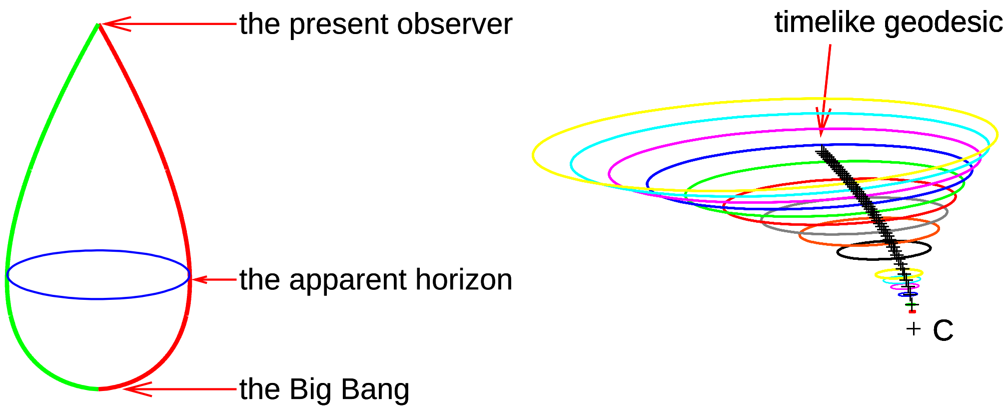

Figure 7.3: **Left panel:** The past light cone of the present observer in the now-standard model of the Universe. **Right panel:** The future light cone of event C in a spacetime with a gravitational field (a fictitious picture).

covariant derivatives in V_{n+1} that will be denoted by the vertical stroke $|$.) Since h_{ab} is covariantly constant with respect to τ^a, while x^α and $g_{\alpha\beta}$ are scalars in V_n, we obtain

$$\frac{\partial g_{\alpha\beta}}{\partial x^\gamma} x^\alpha{}_{,a}\, x^\beta{}_{,b}\, x^\gamma{}_{,c} + g_{\alpha\beta}\left(x^\alpha{}_{;ac}\, x^\beta{}_{,b} + x^\alpha{}_{,a}\, x^\beta{}_{;bc}\right) = 0. \tag{7.44}$$

Rewriting this equation with indices and signs permuted as $-\{abc\} + \{acb\} + \{bca\}$ and adding all three equations we obtain[1]

$$g_{\alpha\beta} x^\beta{}_{,c} \left(x^\alpha{}_{;ab} + \left\{ {\alpha \atop \mu\nu} \right\}_g x^\mu{}_{,a}\, x^\nu{}_{,b}\right) = 0. \tag{7.45}$$

For a fixed c, $x^\beta{}_{,c}$ is the collection of components of a vector in V_{n+1} tangent to V_n, and, since c runs through all n values, the collection $\left\{x^\beta{}_{,c}\right\}_{c=1,\ldots,n}$ is a basis of the tangent space to V_n at a fixed point $\{\tau^a\} \in V_n$. Equation (7.45) means that the object in parentheses is orthogonal to all the n tangent vectors to V_n, so must be collinear with the vector X^α:

$$x^\alpha{}_{;ab} + \left\{ {\alpha \atop \mu\nu} \right\}_g x^\mu{}_{,a}\, x^\nu{}_{,b} = \varepsilon \Omega_{ab} X^\alpha, \tag{7.46}$$

where Ω_{ab} are coefficients to be determined below and ε is the same as in (7.42). The collection of all Ω_{ab} is a tensor in V_n, symmetric in (ab), and a collection of scalars in V_{n+1}. Using (7.41) – (7.43), we find:

$$\Omega_{ab} = g_{\alpha\beta} X^\beta \left(x^\alpha{}_{;ab} + \left\{ {\alpha \atop \mu\nu} \right\}_g x^\mu{}_{,a}\, x^\nu{}_{,b}\right). \tag{7.47}$$

The Ω_{ab} is called the **second fundamental form** of the subspace V_n. Note that the term $\left\{ {s \atop ab} \right\}_h x^\alpha{}_{,s}$ arising from $x^\alpha{}_{;ab}$ in (7.47) vanishes when contracted with $g_{\alpha\beta} X^\beta$, and

$$x^\alpha{}_{,ab} + \left\{ {\alpha \atop \mu\nu} \right\}_g x^\mu{}_{,a}\, x^\nu{}_{,b} \equiv \left(x^\alpha{}_{,a}\right)_{|\nu} x^\nu{}_{,b}$$

[1] $\left\{ {\alpha \atop \mu\nu} \right\}_g$ denotes the Christoffel symbols calculated for the metric $g_{\alpha\beta}$ of the space V_{n+1}. In verifying (7.45) one has to interchange the names of the summation indices in a few places.

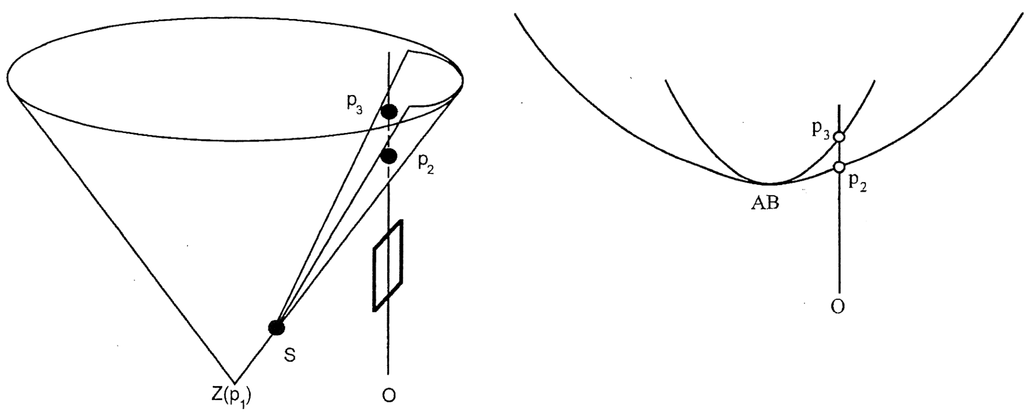

Figure 7.4: A schematic picture of a light cone with self-intersections (this is a typical situation in a neighbourhood of a gravitational lens). **Left panel:** The future light cone of the event Z(p_1). The lens is at S, the vertical line O intersects the cone in two points, p_2 and p_3. **Right panel:** The cross-section of the light cone with the plane marked in the left panel.

is the directional covariant derivative of $x^\alpha{}_{,a}$ along $x^\nu{}_{,b}$. Consequently, this quantity measures the rate of change of the a-th tangent vector to V_n, $x^\alpha{}_{,a}$, as we move along the b-th tangent vector field $x^\nu{}_{,b}$. Then, Ω_{ab} is the projection of this rate of change on the normal vector field to V_n. For this reason, Ω_{ab} is sometimes called the **extrinsic curvature** of the subspace V_n embedded in V_{n+1}. It allows us to 'view' the geometry of V_n from an enveloping space and to see a difference between Riemann spaces that have the same intrinsic Riemann geometry. For example, a plane and a cylinder in a Euclidean 3-space that have identical intrinsic geometries have different second fundamental forms.

In order to know whether a given V_n can be embedded in a given V_{n+1}, we have to find out whether the functions x^α obeying (7.41) exist. They must obey (7.46), which determine the second covariant derivatives of x^α in V_n. These equations will be solvable if the integrability condition – the Ricci formula (6.9) – is fulfilled:

$$x^\alpha{}_{;abc} - x^\alpha{}_{;acb} = R^s{}_{abc}(h) x^\alpha{}_{,s}\,, \tag{7.48}$$

where $R^s{}_{abc}(h)$ is the Riemann tensor of V_n. In calculating $x^\alpha{}_{;abc}$ from (7.46) we will encounter the first derivatives of X^α, so we need to know more about them.

In the calculation below we will use the property $g_{\alpha\beta;\gamma} = 0$, which can be written as

$$g_{\alpha\beta,\gamma} = \left\{ {\rho \atop \alpha\gamma} \right\}_g g_{\rho\beta} + \left\{ {\rho \atop \beta\gamma} \right\}_g g_{\alpha\rho}. \tag{7.49}$$

We differentiate (7.43) covariantly by τ^b, eliminate $g_{\alpha\beta} x^\alpha{}_{;ab} X^\beta$ using (7.47) and use (7.49). As a result, we obtain another expression for Ω_{ab}, equivalent to (7.47):

$$\Omega_{ab} = -g_{\alpha\beta} x^\alpha{}_{,a} X^\beta{}_{,b} - g_{\alpha\rho} \left\{ {\rho \atop \beta\gamma} \right\}_g x^\alpha{}_{,a} x^\gamma{}_{,b} X^\beta. \tag{7.50}$$

This can be equivalently written as

$$-\Omega_{ab} = g_{\alpha\beta} x^\alpha{}_{,a} \; X^\beta{}_{|\gamma} \, x^\gamma{}_{,b} \,, \tag{7.51}$$

which is a covariant derivative of X^β in V_{n+1} projected onto vectors tangent to V_n. This provides another interpretation of Ω_{ab}: $X^\beta{}_{|\gamma} x^\gamma{}_{,b}$ is the rate of change of the normal vector to V_n as we move along the b-th tangent field $x^\gamma{}_{,b}$, and Ω_{ab} is the projection of that rate of change on the a-th tangent vector field $x^\alpha{}_{,a}$. Equations (7.47) and (7.51) clearly show that Ω_{ab} are scalars in V_{n+1}.

The derivatives $X^\beta{}_{,b}$ are not tensors in V_{n+1}. Still, they are objects with one contravariant index in V_{n+1}, and so in every fixed coordinate system they can be decomposed in the vector basis $\{x^\beta{}_{,b} \,, X^\beta\}$; only the coefficients of the decomposition will not be scalars:

$$X^\beta{}_{,b} = A_b{}^s x^\beta{}_{,s} + B_b X^\beta. \tag{7.52}$$

The $A_b{}^s$ are determined when (7.52) is substituted in (7.50). Contracting the result of that substitution with h^{ac} and making use of (7.43) we obtain

$$A_b{}^c = -h^{ac}\Omega_{ab} - h^{ac} g_{\alpha\rho} \left\{ \begin{matrix} \rho \\ \beta\gamma \end{matrix} \right\}_g x^\alpha{}_{,a} \, x^\gamma{}_{,b} \, X^\beta. \tag{7.53}$$

To determine B_b we contract (7.52) with $\varepsilon g_{\alpha\beta} X^\alpha$, and then observe that, from (7.42)

$$0 = \left(g_{\alpha\beta} X^\alpha X^\beta \right)_{,b} \equiv g_{\alpha\beta,b} X^\alpha X^\beta + 2 g_{\alpha\beta} X^\alpha X^\beta{}_{,b} \,,$$

so $g_{\alpha\beta} X^\alpha X^\beta{}_{,b} = -\frac{1}{2} g_{\alpha\beta,b} X^\alpha X^\beta$. Using (7.49) we then obtain

$$B_b = -\varepsilon g_{\beta\rho} \left\{ \begin{matrix} \rho \\ \mu\nu \end{matrix} \right\}_g x^\mu{}_{,b} \, X^\nu X^\beta. \tag{7.54}$$

Note now that the set $\{x^\alpha{}_{,a} \,, X^\alpha\}$, $a = 1, \ldots, n$, is a field of vector bases on V_{n+1}, of exactly the kind we used in Section 4.3. In agreement with our considerations there, the metric tensor $g_{\alpha\beta}$ in V_{n+1} can be represented through its scalar components in these bases:

$$\begin{aligned} \widehat{g}_{ab} &= x^\alpha{}_{,a} \, x^\beta{}_{,b} \, g_{\alpha\beta} = h_{ab}, \qquad \widehat{g}_{(n+1)b} = X^\alpha x^\beta{}_{,b} \, g_{\alpha\beta}, \\ \widehat{g}_{(n+1)(n+1)} &= X^\alpha X^\beta g_{\alpha\beta}. \end{aligned} \tag{7.55}$$

Since X^α is orthogonal to all $x^\alpha{}_{,a}$ by (7.43) and has unit length by (7.42), we have:

$$\widehat{g}_{(n+1)b} = 0, \qquad \widehat{g}_{(n+1)(n+1)} = \varepsilon. \tag{7.56}$$

Therefore, the inverse metric $\widehat{g}^{\alpha\beta}$ has the same block-diagonal form

$$\widehat{g}^{ab} = h^{ab}, \qquad \widehat{g}^{(n+1)b} = 0, \qquad \widehat{g}^{(n+1)(n+1)} = \varepsilon. \tag{7.57}$$

The same coefficients $\{x^\alpha{}_{,a} \,, X^\alpha\}$ can then be used to represent the inverse metric $g^{\alpha\beta}$ via the $\widehat{g}^{ab}$, $\widehat{g}^{(n+1)b}$ and $\widehat{g}^{(n+1)(n+1)}$. Adapting Eq. (4.16) to our present notation, we have

$$g^{\alpha\beta} = x^\alpha{}_{,a} \, x^\beta{}_{,b} \, h^{ab} + \varepsilon X^\alpha X^\beta. \tag{7.58}$$

Using this to eliminate $x^\alpha{}_{,a}\, x^\beta{}_{,c}\, h^{ac}$, we find from (7.53):

$$A_b{}^s x^\beta{}_{,s} = -h^{as}\Omega_{ab}x^\beta{}_{,s} - \left\{ \begin{matrix} \beta \\ \sigma\gamma \end{matrix} \right\}_g x^\gamma{}_{,b}\, X^\sigma + \varepsilon g_{\alpha\rho} \left\{ \begin{matrix} \rho \\ \sigma\gamma \end{matrix} \right\}_g x^\gamma{}_{,b}\, X^\sigma X^\alpha X^\beta. \tag{7.59}$$

Using this and (7.54) in (7.52) we obtain

$$X^\beta{}_{,b} = -h^{sc}\Omega_{sb}x^\beta{}_{,c} - \left\{ \begin{matrix} \beta \\ \sigma\gamma \end{matrix} \right\}_g x^\gamma{}_{,b}\, X^\sigma. \tag{7.60}$$

Now we can use (7.48). Substituting for $x^\alpha{}_{;ab}$ from (7.46), then using (7.46) and (7.60) to eliminate the second derivatives of x^α and the derivatives of X^α, we obtain:

$$\begin{aligned} & R^\alpha{}_{\mu\nu\rho}(g)x^\mu{}_{,a}\, x^\nu{}_{,b}\, x^\rho{}_{,c} + \varepsilon X^\alpha \left(\Omega_{ab;c} - \Omega_{ac;b}\right) \\ + \quad & \varepsilon x^\alpha{}_{,r}\, h^{rm} \left(\Omega_{ac}\Omega_{mb} - \Omega_{ab}\Omega_{mc}\right) - R^r{}_{abc}(h) x^\alpha{}_{,r} = 0, \end{aligned} \tag{7.61}$$

where $R^\alpha{}_{\mu\nu\rho}(g)$ is the Riemann tensor of V_{n+1}. Since $\{x^\alpha{}_{,a}\, , X^\alpha\}$ are a basis of the tangent space to V_{n+1}, Eqs. (7.61) are equivalent to the collection of projections of (7.61) on $\{x^\alpha{}_{,a}\}$ and X^α. Contracting (7.61) with $g_{\alpha\sigma}x^\sigma{}_{,d}$ and with $g_{\alpha\sigma}X^\sigma$ and using (7.41) – (7.43) we get, respectively

$$R_{dabc}(h) = R_{\sigma\mu\nu\rho}(g)x^\sigma{}_{,d}\, x^\mu{}_{,a}\, x^\nu{}_{,b}\, x^\rho{}_{,c} + \varepsilon \left(\Omega_{ac}\Omega_{db} - \Omega_{ab}\Omega_{dc}\right), \tag{7.62}$$

$$\Omega_{ab;c} - \Omega_{ac;b} = -R_{\sigma\mu\nu\rho}(g)X^\sigma x^\mu{}_{,a}\, x^\nu{}_{,b}\, x^\rho{}_{,c}\,. \tag{7.63}$$

Equations (7.62) – (7.63) are called the **Gauss – Codazzi equations**. This is the full set of necessary and sufficient conditions for V_n to be embeddable in V_{n+1}. In relativity, Eqs. (7.62) and (7.63) appear most often with $n = 3$, i.e. for hypersurfaces in spacetime.

The expression (7.51) may be simplified further when the coordinates in V_{n+1} are adapted to V_n as follows. Through every point of V_n we run a curve $C(p)$ in V_{n+1} orthogonal to V_n and choose the arc length s on the curves as the x^{n+1} coordinate in V_{n+1} with $x^{n+1} = A =$ constant on V_n. The equations $A \neq x^{n+1} =$ constant then define other hypersurfaces in V_{n+1}. The intrinsic coordinates of V_n, $x^a = \tau^a$, $a = 1, \ldots, n$ are chosen as the remaining $\{x^1, \ldots, x^n\}$ coordinates in V_{n+1}. In such coordinates $g_{(n+1)a}(V_n) = 0$ and $g_{ab}(V_n) = h_{ab}$. In these coordinates, (7.51) may be written as

$$\Omega_{ab} = -X_{a;b}. \tag{7.64}$$

In some textbooks, Eq. (7.64) is used as the definition of Ω_{ab}. It has the appearance of being covariant, but it is not – it holds only in the adapted coordinates. Moreover, the notation of (7.64) is misleading: the semicolon denotes not the covariant derivative in V_n, but the covariant derivative in V_{n+1} projected on V_n by $x^\mu{}_{,a}$.

In the adapted coordinates, Eq. (7.50) reduces to another useful, although non-covariant, form. In these coordinates, X^β has only the $(n+1)$-st component everywhere in V_{n+1}, thus $X^a = 0$. Since $x^a = \tau^a$, the first term in (7.50) becomes $-g_{a(n+1)}X^{n+1}{}_{,b}$ and is zero because $g_{a(n+1)} = 0$. The second term becomes $\left(-h_{ar}\left\{ \begin{matrix} r \\ (n+1)b \end{matrix} \right\}_g X^{n+1}\right)$. It is easy to verify that $\left\{ \begin{matrix} r \\ (n+1)b \end{matrix} \right\}_g = \frac{1}{2}h^{rs}h_{bs,(n+1)}$, thus

$$\Omega_{ab} = -\frac{1}{2}h_{ab,(n+1)}X^{n+1}. \tag{7.65}$$

Thus, in the adapted coordinates, Ω_{ab} is proportional to the directional derivative of the metric in V_n along the normal vector.

7.12 Exercises

1. Find the metric tensors for the following surfaces: (*i*) A cone; (*ii*) A paraboloid of revolution; (*iii*) One- and two-sheeted hyperboloids of revolution; (*iv*) A torus; (*v*) A cylinder. Calculate the curvature of the cone, the cylinder and the torus.

2. Show that the geodesic lines on a 2-dimensional sphere are great circles.

Hint: First derive the equation of a great circle in spherical coordinates, then solve the geodesic equation in the same coordinates.

3. Solve the geodesic equation on the surface of a cylinder. What curves are the geodesics?

4. Find the second fundamental form for a plane and for a cylinder embedded in a Euclidean E^3.