3 Metric differential geometry

The main object of study in GR is the ***metric tensor*** $g \in \mathfrak{X}^{(2,0)}(M)$. This is a smooth family

$$g_x : T_xM \times T_xM \to \mathbb{R} \tag{3.1}$$

of metrics as defined in §2.3, where now $V = T_xM$ is indexed by $x \in M$. Thus, repeating the definition, each g_x is *bilinear*, *symmetric* (i.e. $g_x(X_x, Y_x) = g_x(Y_x, X_x)$ for all $X_x, Y_x \in T_xM$), and *nondegenerate* (i.e. $g_x(X_x, Y_x) = 0$ for all $Y_x \in T_xM$ iff $X_x = 0$). *It need not be positive definite.*

The orthonormal basis $(e_a(x)) = (e_1(x), \ldots, e_n(x))$ in which g_x is diagonal (cf. §2.3) may–and typically will–depend on x. But if M is connected, the signature of g_x is independent of x by continuity. Even if M is not connected, we assume it is independent of x. Thus the signature is an intrinsic property not only of each pointwise metric g_x, but even of an entire metric tensor g.

A manifold M with a metric tensor g is called ***semi-Riemannian***, with two special cases:

1. The metric (or manifold) is called ***Riemannian*** if the signature is $(+\cdots+)$. Thus each g_x is positive definite. Given the assumption of symmetry, this *implies* that g_x is nondegenerate, so a metric tensor is Riemannian iff each g_x is symmetric and positive definite.[139]

2. The metric (or manifold) is called ***Lorentzian*** if $\dim(M) = 4$ and $n_- = 1$, i.e. the signature of g is $(-+++)$.[140] Hence with respect to an orthonormal basis (e_a) we have

$$g(e_a, e_b) = \eta_{ab}; \qquad \eta := \text{diag}(-1, 1, 1, 1). \tag{3.2}$$

With some abuse of notation, the symbol η_n, with $\eta \equiv \eta_4$, is also used for the ***Minkowski metric***

$$\eta_n : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}; \qquad \eta_n(X, Y) := \eta_{ab} X^a Y^b = -X^0 Y^0 + \sum_{i=1}^{n-1} X^i Y^i, \tag{3.3}$$

where (X^μ) and (Y^μ) are either meant to be Cartesian coordinates on $\mathbb{R}^4$ seen as a vector space,[141] or, identifying $T_x\mathbb{R}^4 \cong \mathbb{R}^4$, denote components of tangent vectors $X = X^\mu \partial_\mu$ etc. with respect to the basis $(\partial_0, \partial_1, \partial_2, \partial_3)$ defined by the usual coordinates (x) on $\mathbb{R}^4$, seen as our manifold M. Either way, $(\mathbb{R}^4, \eta)$, often written as $(\mathbb{M}, \eta)$, is ***Minkowski space-time***, which is the oldest and simplest example of a Lorentzian manifold. The fact that, in this special case, the metric is defined not only on each tangent space T_xM, as always, but also on M itself, has no analogue for general Lorentzian manifolds. In special relativity, however, lightcones and other causal structures are defined in $M = \mathbb{R}^4 = \mathbb{M}$ itself, which makes it useful to define the metric η on both M and T_xM. Causal theory for general Lorentzian manifolds will be developed in §5.3.

Lorentzian manifolds underlie GR, but we often invoke examples from Riemannian geometry in order to explain some contrast with the Lorentzian case. Furthermore, Riemannian submanifolds of M are often important, e.g. in the Cauchy problem for GR (see chapter 7).

---

[139] The case $(-\cdots-)$ may also be included here, since an overall change of sign in g makes it Riemannian.

[140] This name is sometimes also used in any dimension $d \geq 2$ provided $n_- = \dim(M) - 1$. Furthermore, a similar comment as in the previous footnote applies: we may as well take $n_+ = \dim(M) - 1$. In any $d \geq 2$, a necessary and sufficient condition for a metrizable manifold M to support a Lorentzian metric is that M is either non-compact, or, if it is compact, has zero Euler characteristic. These conditions are equivalent to the existence of a non-vanishing continuous vector field on M (Palomo & Romero 2006, §1.1; Minguzzi, 2019, §1.8). For deeper topological constraints imposed by Lorentzian metrics with additional (causality) properties, see Chernov & Nemirovski (2013). But in GR one often starts with a metric defined by some formula and looks for a manifold supporting it!

[141] Seen as Minkowski space-time, it is conventional to relabel the usual coordinates of $\mathbb{R}^4$ as (x^0, x^1, x^2, x^3), where $x^0 = t$ denotes time. In diagrams, the time axis is always drawn vertically. We also introduce a convention often used in the (physics) literature: Greek indices μ, ν etc. run from 0 to 3, whereas Latin indices i, j etc. run from 1 to 3. Both Greek and Latin indices midway in the alphabet usually refer to the canonical coordinate basis $\partial_\mu = \partial/\partial x^\mu$ or $\partial_i = \partial/\partial x^i$, whereas indices a, b etc. typically refer to other bases (e_a), often orthonormal ones.

3.1 Lowering and raising indices

Let (M,g) be a (semi) Riemannian manifold. Since each g_x is a metric, the distinction between vectors and covectors is blurred, because as in §2.3 we have "musical" isomorphisms

$$\flat_x : T_xM \to T_x^*M; \qquad \flat_x(X) \equiv X^\flat; \qquad X^\flat(Y) := g_x(X,Y); \tag{3.4}$$
$$\sharp_x : T_x^*M \to T_xM; \qquad \sharp_x(\theta) \equiv \theta_\sharp; \qquad g_x(\theta_\sharp, X) := \theta(X), \tag{3.5}$$

which are each other's inverse. These pointwise isomorphisms induce mutually inverse maps

$$\flat : \mathfrak{X}(M) \to \Omega(M); \qquad \sharp : \Omega(M) \to \mathfrak{X}(M), \tag{3.6}$$

by pointwise application. This leads to the ***lowering and raising of indices***, which is crucial to almost any computation in GR. At any point x (which we omit) we define (g^{ab}) as the inverse (matrix) to (g_{ab}), where $g_{ab} = g(e_a,e_b)$ in some basis e_a (so that $g^{ab}g_{bc} = \delta^a_c$). Then

$$X_a^\flat = g_{ab}X^b; \qquad \theta_\sharp^a = g^{ab}\theta_b, \tag{3.7}$$

which notation may then be extended to any tensor, *where the "sharp" and "flat" signs are usually omitted*. For example, (3.7) is simply written as $X_a = g_{ab}X^b$ and $\theta^a = g^{ab}\theta_b$.

The above definition of (g^{ab}) is consistent with the following one. Extending $\sharp_x$ to a map

$$\sharp_x \otimes \sharp_x : T_x^*M \otimes T_x^*M \to T_xM \otimes T_xM \tag{3.8}$$

in the obvious way, i.e., by linear extension of $\theta \otimes \eta \mapsto \theta_\sharp \otimes \eta_\sharp$, we obtain

$$\sharp_x \otimes \sharp_x(g_x) \in T_x^{(0,2)}M = \mathrm{Hom}(T_x^*M \times T_x^*M, \mathbb{R}). \tag{3.9}$$

If (ω^a) is the dual basis to (e_a), then $g_x^{ab} = \sharp_x \otimes \sharp_x(g_x)(\omega^a(x), \omega^b(x))$, as the reader will verify.

More generally, lowering and raising of specified indices are maps defined, respectively, by

$$\flat : \mathfrak{X}^{(k,l)}(M) \to \mathfrak{X}^{(k+1,l-1)}(M); \qquad \sharp : \mathfrak{X}^{(k,l)}(M) \to \mathfrak{X}^{(k-1,l+1)}(M), \tag{3.10}$$

provided $l > 0$ in the first and $k > 0$ in the second case. Taking the first index for example gives

$$T^\flat(X_1,\dots,X_{k+1};\theta^1,\dots,\theta^{l-1}) = T(X_2,\dots,X_{k+1};X_1^\flat,\theta^1,\dots,\theta^{l-1}); \tag{3.11}$$
$$T_\sharp(X_1,\dots,X_{k-1};\theta^1,\dots,\theta^{l+1}) = T(\theta^1_\sharp,X_1,\dots,X_{k-1};\theta^2,\dots,\theta^{l+1}). \tag{3.12}$$

Curvature will described by the Riemann tensor $R \in \mathfrak{X}^{(3,1)}(M)$, of which the only upper index is usually written first. This index may then be lowered, so that $R^\flat \in \mathfrak{X}^{(4,0)}(M)$ has components

$$R^\flat_{abcd} \equiv R_{abcd} = g_{ae}R^e_{bcd}. \tag{3.13}$$

The contraction process explained at the end of the previous chapter, which in principle has nothing to do with the metric, may now elegantly be rewritten in terms of the metric by, e.g.,

$$R_{ab} = R^c_{acb} = g^{cd}R^\flat_{dacb} \equiv g^{cd}R_{dacb}. \tag{3.14}$$

Metric contraction may be done even in case where the original version does not apply, as in

$$R = R^a_{\sharp a} = g^{ab}R_{ab}. \tag{3.15}$$

If $R \in \mathfrak{X}^{(3,1)}(M)$ is the Riemann tensor, so that its first contraction $R \in \mathfrak{X}^{(2,0)}(M)$ is the Ricci tensor, this second contraction yields the ***Ricci scalar***, which again plays a central role in GR.[142]

[142]Our use of the same letter R for the Riemann tensor, the Ricci tensor, and the Ricci scalar will never lead to confusion, as all relevant instances contain indices distinguishing them. For experts: we do *not* use Penrose's abstract index notation, which may clarify things but ever so often leads to typographically awkward expressions.

3.2 Geodesics

Intuitively, geodesics are paths of shortest lengths between two given points.[143] This idea only makes direct sense in the Riemannian case (as opposed to the semi-Riemannian case), with which we therefore start. We will then find a redefinition of a geodesic that does make sense also on semi-Riemannian manifolds. *Throughout this section* (M,g) is a *Riemannian* manifold. It will now be convenient to use *closed* intervals $I=[a,b]$ as the domains of curves $\gamma: I \to M$.

1. The ***length*** of a curve $\gamma: [a,b] \to M$ is defined as

$$L(\gamma) := \int_a^b dt \sqrt{g_{\gamma(t)}(\dot\gamma(t),\dot\gamma(t))} \equiv \int_a^b dt\, \|\dot\gamma(t)\|, \tag{3.16}$$

   where $\dot\gamma(t) \in T_{\gamma(t)}M$ is the tangent vector to the curve, cf. (2.25). So in coordinates one has $\gamma(t) = (\gamma^1(t),\dots,\gamma^n(t))$, where $\gamma^i : [a,b] \to \mathbb{R}$, and hence

$$g_{\gamma(t)}(\dot\gamma(t),\dot\gamma(t)) = g_{ij}(\gamma(t))\frac{d\gamma^i(t)}{dt}\frac{d\gamma^j(t)}{dt} \equiv g_{ij}(\gamma(t))\dot\gamma^i(t)\dot\gamma^j(t). \tag{3.17}$$

   Using a change of variables in the integral (3.16), it is easy to show that the length of γ is independent of its parametrization, so that it only depends on the image $\gamma([a,b])$ in M.

2. If M is connected, any two points can be connected by a smooth curve, and hence we can define the ***distance*** $d(x,y)$ between $x,y \in M$ as the infimum of $L(\gamma)$ over all smooth curves $\gamma: [0,1] \to M$ with $\gamma(0) = x$ and $\gamma(1) = y$ (one may equivalently use piecewise smooth curves here, since these can be smoothened, cf. Lemma 5.8 below). This is a metric on M, whose metric topology coincides with the original topology of M.[144]

3. A ***geodesic*** is a curve of extremal length (with a specific parametrization, see below).

We will not precisely explain what this problem in the calculus of variations means, since our goal is merely to motivate Definition 3.1 below, which also applies to the semi-Riemannian case. Therefore, we just outline how this extremal problem is solved. In general, a functional

$$S(\gamma) = \int_a^b dt\, \mathscr{L}(\gamma(t),\dot\gamma(t)) \tag{3.18}$$

is minimized or maximized by some curve γ iff the ***Euler–Lagrange equations*** hold:

$$\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot\gamma^i} - \frac{\partial \mathscr{L}}{\partial \gamma^i} = 0. \tag{3.19}$$

Short of giving an introduction to the calculus of variations, here is a heuristic derivation of (3.19). Let $\gamma_s(t)$ a family of curves indexed by s, such that endpoints are fixed, that is,

$$\gamma_s(a) = \gamma(a); \qquad\qquad \gamma_s(b) = \gamma(b). \tag{3.20}$$

[143] Recall our standing assumption that all maps, including curves and metrics, are smooth. Uniqueness and variational properties of geodesics change completely if the metric is just C^1 (Hartman & Wintner, 1951; Hartman, 1983). On the other hand, most of the smooth theory is already valid in the Hölder class $C^{2,1}$ (Minguzzi, 2015a).

[144] See Jost (2002), pp. 14–15. We do not prove this since it is practically irrelevant for the Lorentzian case.

The extremality condition that defines the variational problem is

$$dS(\gamma_s)/ds = 0. \tag{3.21}$$

On repeatedly using the chain rule and a partial integration, eq. (3.21) with (3.20) gives

$$\begin{aligned}\frac{dS(\gamma_s)}{ds} &= \int_a^b dt \left(\frac{\partial \mathscr{L}}{\partial \gamma_s^i}\frac{\partial \gamma_s^i}{\partial s} + \frac{\partial \mathscr{L}}{\partial \dot{\gamma}_s^i}\frac{\partial \dot{\gamma}_s^i}{\partial s}\right) = \int_a^b dt \left(\frac{\partial \mathscr{L}}{\partial \gamma_s^i}\frac{\partial \gamma_s^i}{\partial s} + \frac{\partial \mathscr{L}}{\partial \dot{\gamma}_s^i}\frac{\partial}{\partial t}\frac{\partial \gamma_s^i}{\partial s}\right)\\ &= \int_a^b dt \left(\frac{\partial \mathscr{L}}{\partial \gamma_s^i} - \frac{\partial}{\partial t}\frac{\partial \mathscr{L}}{\partial \dot{\gamma}_s^i}\right)\frac{\partial \gamma_s^i}{\partial s} + \bigg|_b^a \frac{\partial \mathscr{L}}{\partial \dot{\gamma}_s^i}\frac{d\gamma_s^i}{ds}.\end{aligned} \tag{3.22}$$

Then (3.20) gives $d\gamma_s(a)/ds = d\gamma_s(b)/ds = 0$, so that, for arbitrary γ_s and hence arbitrary $\partial\gamma_s/\partial s$, eq. (3.21) implies (3.19), in which s is dropped and hence $\partial/\partial t$ becomes d/dt.

The Euler–Lagrange equations for the length functional (3.16) are not very nice, but they can be simplified if a preferred ("affine") parametrization is used. To motivate this, instead of the length (3.16), we now start from the (kinetic) ***energy*** of our curve γ, defined as

$$E(\gamma) := \int_a^b dt\, g_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) = \int_a^b dt\, \|\dot{\gamma}(t)\|^2. \tag{3.23}$$

For the energy (3.23), the Euler–Lagrange equations (3.19) give the ***geodesic equation***

$$\ddot{\gamma}^i(t) + \Gamma^i_{jk}(\gamma(t))\dot{\gamma}^j(t)\dot{\gamma}^k(t) = 0, \tag{3.24}$$

or briefly $\ddot{\gamma}^i + \Gamma^i_{jk}\dot{\gamma}^j\dot{\gamma}^k = 0$, where $\ddot{\gamma} = d^2\gamma/dt^2$, and the ***Christoffel symbols*** are given by

$$\Gamma^i_{jk} := \tfrac{1}{2} g^{im}(g_{mj,k} + g_{mk,j} - g_{jk,m}), \tag{3.25}$$

where we have introduced another useful notational convention from GR:

$$\tau^{j_1\cdots j_l}_{i_1\cdots i_k,j} = \partial_j \tau^{j_1\cdots j_l}_{i_1\cdots i_k}. \tag{3.26}$$

Warning: the Christoffel symbols do *not* form the components of a would-be tensor "$\Gamma \in \mathfrak{X}^{(2,1)}(M)$": physicists see this from their incorrect behaviour under coordinate transformations, whereas mathematicians note that Γ fails the tensoriality test, cf. Proposition 2.7. We will see, however, that the Γ-*symbols* do combine into the Riemann *tensor*!

To derive (3.24) for (3.23), i.e., for $\mathscr{L}(\gamma(t), \dot{\gamma}(t)) = g_{ij}(\gamma(t))\dot{\gamma}^i(t)\dot{\gamma}^j(t)$, one uses

$$\frac{\partial \mathscr{L}}{\partial \gamma^i} = g_{jk,i}\dot{\gamma}^j\dot{\gamma}^k; \tag{3.27}$$

$$\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{\gamma}^i} = 2\frac{d}{dt} g_{ij}\dot{\gamma}^j = 2(g_{ij,k}\dot{\gamma}^k\dot{\gamma}^j + g_{ij}\ddot{\gamma}^j) = (g_{ij,k} + g_{ik,j})\dot{\gamma}^k\dot{\gamma}^j + 2g_{ij}\ddot{\gamma}^j. \tag{3.28}$$

Whereas solutions of (3.24) extremize the *energy* for any parametrization, for the *length* (3.16), the Euler–Lagrange equations only take the form (3.24) iff $\|\dot{\gamma}(t)\|$ is constant, in which case the parametrization of the curve $\gamma\colon [a,b] \to M$ is said to be ***affine***. In particular, if $\|\dot{\gamma}(t)\| = 1$ for all $t \in I$, then we say that γ is parametrized by ***arc length***.

Definition 3.1 *A geodesic is a curve $\gamma : I \to M$ (with $I \subset \mathbb{R}$ connected) that satisfies* (3.24).

On this definition, geodesics still extremize length, but eq. (3.24) *implies* that $\|\dot{\gamma}(t)\|$ is constant, as can be shown by computing $d(\|\dot{\gamma}(t)\|^2)/dt$ from (3.17). This time-derivative equals

$$\frac{d\|\dot{\gamma}(t)\|^2}{dt} = g_{ij,k}\dot{\gamma}^i\dot{\gamma}^j\dot{\gamma}^k + 2g_{ij}\ddot{\gamma}^i\dot{\gamma}^j. \tag{3.29}$$

Eliminating $\ddot{\gamma}^i$ via (3.24) then leads to a cancellation making the right-hand side zero; a neater calculation will be given after (3.66). *The definition of a geodesic therefore depends on the parametrization of* γ: a reparametrized geodesic may no longer satisfy (3.24), except when the reparametrization is *affine*, i.e. $s = at + b$. However, one has the following useful criterion.

Proposition 3.2 *Some curve $\gamma : [a,b] \to M$ can be reparametrized so as to* become *a geodesic iff the right-hand side of* (3.24) *equals $f \cdot \dot{\gamma}^i$ for some function $f(t)$ defined along γ.*

Proof. If some curve $t \mapsto \gamma(t)$ satisfies (3.24), then $t \mapsto \gamma(s(t))$ satisfies (3.24) with right-hand side $\ddot{s}\dot{\gamma}^i$, and conversely one can solve $f(t) = \ddot{s}(t)$ for s and switch to $\gamma \circ s^{-1}$. □

Such a (poorly parametrized) curve that is "almost" a geodesic is sometimes called a ***pregeodesic***. In $M = \mathbb{R}^n$ with flat metric (i.e. $g_{ij} = \delta_{ij}$) geodesics are straight lines that form *shortest* paths between two given points. This is also true in e.g. hyperbolic space, and it is always true for sufficiently short geodesics. On the sphere (where geodesics are great circles) one has two geodesics between two generic points; but only one has minimal length. These lengths coincide iff the two points are polar opposites, in which case one has infinitely many geodesics. See §5.5.

In the intuitive idea of geodesics the focus is on endpoints, whereas in defining geodesics as solutions to the ODE (3.24) the focus is on the initial point $\gamma(0)$ and initial velocity $\dot{\gamma}(0)$. The solution to (3.24) is uniquely defined by these data, except for I. But like any solution to an ODE, γ has some maximal domain of definition $I \subset \mathbb{R}$, and this domain may or may not equal $\mathbb{R}$.

Definition 3.3 *If all geodesics $\gamma : I \to \mathbb{R}$ with given initial point $\gamma(0)$ and initial velocity $\dot{\gamma}(0)$ can be defined on the maximum interval $I = \mathbb{R}$, we say that (M,g) is* **geodesically complete**.

For example, $\mathbb{R}^n$, the sphere S^n, and hyperbolic space H^n are geodesically complete (cf. §4.4). In the Riemannian case this is equivalent to a purely topological property. For $x, y \in M$ define

$$d(x,y) := \inf\{L(\gamma) \mid \gamma : [a,b] \to M, \gamma(a) = x, \gamma(b) = y\}. \tag{3.30}$$

It is easy to show that this defines a metric in the topological sense, i.e. a symmetric function $d : M \times M \to [0,\infty)$ that satisfies $d(x,y) = 0$ iff $x = y$ and $d(x,y) \leq d(x,z) + d(z,y)$. In other words, a Riemannian manifold (M,g) is also a metric space (M,d). For the latter, one has the usual notion of completeness in the sense that any Cauchy sequence converges.

Theorem 3.4 **(Hopf-Rinow)** *A Riemannian manifold (M,g) is geodesically complete iff the corresponding metric space (M,d) defined by* (3.30) *is complete. In that case, any two points x,y can be joined by a geodesic of minimum length (compared with all curves from x to y).*

Since this theorem has no analogue in the Lorentzian case we will not prove it. We do note that any compact Riemannian manifold is complete. On the other hand, examples of incomplete Riemannian manifolds are provided by open bounded sets $\Omega \subset \mathbb{R}^n$ with flat metric inherited from $\mathbb{R}^n$, or $\mathbb{R}^n$ itself with one or more points or regions omitted. Such examples also show that in the incomplete case the infimum in (3.30) may not be attained. Many Lorentzian manifolds of interest to GR are geodesically *in*complete in a nontrivial (i.e. inextendible) sense; see chapter 6.

3.3 Linear connections

The definition of a geodesic as a curve γ whose tangent vector $\dot{\gamma}$ satisfies (3.24) along the curve for which $\gamma(t)$ is defined was inspired by the Riemannian case, but it will be taken as the definition of a geodesic on a semi-Riemannian manifold, too. In support, we now give a geometric perspective on the Christoffel symbols Γ^i_{jk} and hence on the geodesic equation (3.24).

Definition 3.5 *A* **linear connection** *on M (which is the same thing as a connection on the tangent bundle TM, see below), or, equivalently, a* **covariant derivative** *on $\mathfrak{X}(M)$, is a map*

$$X \mapsto \nabla_X : \mathfrak{X}(M) \to \mathfrak{X}(M), \tag{3.31}$$

where X itself is a vector field on M (i.e. $X \in \mathfrak{X}(M)$), such that:

1. *The map $X \mapsto \nabla_X$ is $\mathbb{R}$-linear as well as $C^\infty(M)$-linear, i.e.*

$$\nabla_{fX} Y = f \nabla_X Y \qquad \forall f \in C^\infty(M); \tag{3.32}$$

2. *The map $Y \mapsto \nabla_X Y$ is $\mathbb{R}$-linear but not $C^\infty(M)$-linear: it satisfies the* **Leibniz rule**

$$\nabla_X (fY) = (Xf)Y + f\nabla_X Y \qquad \forall f \in C^\infty(M). \tag{3.33}$$

This definition also makes sense on any open $U \in \mathscr{O}(M)$, and in fact if $x \in U$, then $\nabla_X Y(x)$ only depends on the value of X at x and the restriction of Y to U; this follows from (3.32) - (3.33) and the definition of a manifold. Hence we may compute covariant derivatives locally. Recall that a local frame (e_a) for $\mathfrak{X}(U)$ consists of n maps $e_a : U \to TM$ such that at each $x \in U$ the vectors $e_a(x) \in T_xM$ form a basis of T_xM $(a = 1, \ldots, n)$. The corresponding dual basis (ω^a) for $\Omega(U)$ then consists of the $\omega^a(x) \in T^*_xM$ that satisfy $\omega^a(e_b) = \delta^a_b$. The given connection ∇ is then completely characterized by its ***connection coefficients*** ω^c_{ab}, defined (at each x) by

$$\nabla_{e_a} e_b = \omega^c_{ab} e_c. \tag{3.34}$$

Indeed, from (2.68) - (2.69) we may write $X = X^a e_a$, where $X^a = \omega^a(X) \in C^\infty(U)$, so

$$\begin{aligned} \nabla_X Y &= \nabla_{X^a e_a}(Y^b e_b) = X^a \nabla_{e_a}(Y^b e_b) \\ &= X^a (e_a(Y^b) \cdot e_b + Y^b \nabla_{e_a} e_b) \\ &= X^a (e_a(Y^c) + Y^b \omega^c_{ab}) e_c. \end{aligned} \tag{3.35}$$

We write $\nabla_X Y^a$ for $(\nabla_X Y)^a$, so that $\nabla_X Y = (\nabla_X Y^a) e_a$. We therefore have

$$\nabla_X Y^a = X(Y^a) + \omega^a_{bc} X^b Y^c, \tag{3.36}$$

where $X(Y^a)$ is the action of the vector field X on the *function* $Y^a \in C^\infty(U)$. In terms of a coordinate basis $(e_\mu = \partial_\mu)$, $(\omega^\nu = dx^\nu)$, writing $\nabla_\mu := \nabla_{\partial_\mu}$ the above relations imply

$$\omega^\rho_{\mu\nu} = dx^\rho(\nabla_\mu \partial_\nu); \tag{3.37}$$

$$\nabla_X Y^\rho = X^\mu(\partial_\mu Y^\rho + \omega^\rho_{\mu\nu} Y^\nu); \tag{3.38}$$

$$\nabla_\mu Y^\rho = \partial_\mu Y^\rho + \omega^\rho_{\mu\nu} Y^\nu. \tag{3.39}$$

Linear connections formalize Levi-Civita's notion of ***parallel transport***. It follows from (3.36) or (3.38) that $\nabla_X Y$ only depends on the values of Y along the flow lines of X, for

$$\nabla_X Y^a(x) = \frac{d}{dt} Y^a(\psi_t(x))_{|t=0} + \omega^a_{bc}(x) X^b(x) Y^c(x), \tag{3.40}$$

where ψ is the flow of X. Conversely, given some curve $\gamma : I \to M$ with tangent vectors $\dot\gamma$ (*defined along γ only!*), the covariant derivative $\nabla_{\dot\gamma} Y$ of Y along γ is well defined for any vector field Y defined near or even on $\gamma(I)$ alone; for in (local) coordinates we have

$$\begin{aligned} \nabla_{\dot\gamma} Y^\rho_{\gamma(t)} &= \dot\gamma^\mu(t)(\partial_\mu Y^\rho_{\gamma(t)} + \omega^\rho_{\mu\nu}(\gamma(t)) Y^\nu_{\gamma(t)}) \\ &= \frac{d}{dt} Y^\rho_{\gamma(t)} + \omega^\rho_{\mu\nu}(\gamma(t)) \frac{d\gamma^\mu(t)}{dt} Y^\nu_{\gamma(t)}, \end{aligned} \tag{3.41}$$

where $\gamma^\mu : I \to \mathbb{R}$ are the coordinates of the curve (in some given chart), as before.

Definition 3.6 *A (necessarily unique) vector field $t \mapsto Y_{\gamma(t)} \in T_{\gamma(t)}M$ defined along a given curve γ is the* parallel-transport *of some initial vector $Y \in T_{\gamma(0)}M$ along γ if Y satisfies*

$$\nabla_{\dot\gamma} Y = 0. \tag{3.42}$$

This generalizes the Euclidean practice of freely moving vectors in $\mathbb{R}^n$ from place to place, to arbitrary (semi) Riemannian manifolds. The price one pays is that such motions can only be carried out once a linear connection has been defined. The ***flat connection*** on $\mathbb{R}^n$ (with flat metric $g = \delta$), defined in the standard coordinates by $\omega^\rho_{\mu\nu} = 0$ gives $\nabla_\mu = \partial_\mu$ and hence $Y_{\gamma(t)} = Y_{\gamma(0)} = Y$ for all t. Hence "freely moving vectors" in $\mathbb{R}^n$ is *relative to this flat connection*.

Like the Christoffel symbols, the connection coefficients do not form the components of a tensor (the relation between the two will be clarified shortly). However, various tensors may be defined via the connection. For now, we just define the ***torsion*** $\tau_\nabla \in \mathfrak{X}^{(2,1)}(M)$ of ∇ by

$$\tau_\nabla(X, Y, \theta) := \theta(\nabla_X Y - \nabla_Y X - [X, Y]). \tag{3.43}$$

A simple computation shows that this expression is $C^\infty(M)$-linear in each entry, so Proposition 2.7 shows τ is indeed a tensor of the said kind. In the coordinate basis (∂_μ), we have

$$\tau^\rho_{\mu\nu} = \omega^\rho_{\mu\nu} - \omega^\rho_{\nu\mu}, \tag{3.44}$$

since $[\partial_\mu, \partial_\nu] = 0$. Hence the connection ∇ is ***torsion-free*** iff any of the following hold:

$$\omega^\rho_{\mu\nu} = \omega^\rho_{\nu\mu}; \tag{3.45}$$

$$\nabla_\mu \partial_\nu = \nabla_\nu \partial_\mu; \tag{3.46}$$

$$\nabla_X Y - \nabla_Y X = [X, Y]. \tag{3.47}$$

We are now in a position to restate and generalize Definition 3.1:

Definition 3.7 *Given some linear connection ∇ on M, a* **geodesic** *in M is a curve γ for which*

$$\nabla_{\dot\gamma} \dot\gamma = 0. \tag{3.48}$$

That is, the tangent vector $\dot{\gamma}$ to γ is parallel transported along γ. As before, this definition requires a specific parametrization of γ, which is unique up to affine transformations of t. One has a similar situation as in the metric case for detecting "wrongly parametrized" geodesics:

Proposition 3.8 *Some curve $\gamma : [a,b] \to M$ can be reparametrized so as to* become *a geodesic iff the right-hand side of (3.48) equals $f\dot{\gamma}$, for some function $f(t)$ defined along γ.*

The proof is analogous to Proposition 3.24. Using (local) coordinates, eq. (3.48) may be brought into a form that is strikingly similar to (3.24). Since according to (3.41) with $Y \rightsquigarrow \dot{\gamma}$ the expression $\dot{\gamma}^{\mu}\partial_{\mu}\dot{\gamma}^{\rho}$ is just $d^2\dot{\gamma}^{\rho}/dt^2 \equiv \ddot{\gamma}^{\rho}$, we obtain

$$\ddot{\gamma}^{\rho} + \omega^{\rho}_{\mu\nu}\dot{\gamma}^{\mu}\dot{\gamma}^{\nu} = 0, \tag{3.49}$$

from which it is obvious that geodesics are insensitive to the torsion (3.44) of the connection. Eq. (3.49) looks like the geodesic equation (3.24), with the difference that in (3.49) the coefficients $\omega^{\rho}_{\mu\nu}$ are defined by (3.37) in terms of an arbitrary linear connection ∇, whereas those in (3.24) are the Christoffel symbols (3.25) defined by the metric. Their relationship is as follows.

Theorem 3.9 **(Levi-Civita)** *Any (semi) Riemannian manifold (M,g) admits a unique linear connection ∇ (called the* **Levi-Civita connection***) that satisfies the following two properties:*

1. *The torsion τ_{∇} associated to ∇ vanishes, i.e. $\nabla_X Y - \nabla_Y X = [X,Y]$.*
2. *The connection ∇ and the metric g are related by the condition that for all $X,Y,Z \in \mathfrak{X}(M)$,*

$$X(g(Y,Z)) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z). \tag{3.50}$$

These conditions imply that the connection coefficients of ∇ are the Christoffel symbols (3.25):

$$\omega^{\rho}_{\mu\nu} = \Gamma^{\rho}_{\mu\nu}. \tag{3.51}$$

As soon as we have extended ∇ to arbitrary tensors, we will see that (3.50) comes down to

$$\nabla_X g = 0 \qquad \forall X \in \mathfrak{X}(M). \tag{3.52}$$

Also, $X(g(Y,Z))$ will be the same as $\nabla_X(g(Y,Z))$, hence some authors elegantly write (3.50) as

$$\nabla_X \langle Y,Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle. \tag{3.53}$$

Proof. Using (3.47) and (3.50), one computes

$$X(g(Y,Z)) - Z(g(X,Y)) + Y(g(Z,X)),$$

and rearranges this to obtain the so-called ***Koszul formula***, partly written in the notation (3.53):

$$\langle \nabla_X Y, Z\rangle = \tfrac{1}{2}(X\langle Y,Z\rangle + Y\langle Z,X\rangle - Z\langle X,Y\rangle - \langle X,[Y,Z]\rangle + \langle [X,Y],Z\rangle + \langle Y,[Z,X]\rangle). \tag{3.54}$$

Since g is nondegenerate this uniquely fixes $\nabla_X Y$, and in a coordinate basis this gives (3.51).

To prove existence, one easily checks (3.32) and (3.33) from (3.54). Finally, running the derivation of (3.54) from (3.47) and (3.50) backwards verifies (3.47) and (3.50). □

3.4 General connections on vector bundles

For a more general understanding of the above constructions, as well as for a clean extension of linear connections from vector fields to arbitrary tensors (which one often needs in GR), we briefly discuss connections on arbitrary vector bundles. Similar to Definition 3.10, we put:

Definition 3.10 *A* **connection on a vector bundle** $E \to M$ *is a linear map*

$$X \mapsto \nabla_X : \Gamma(E) \to \Gamma(E), \tag{3.55}$$

where $X \in \mathfrak{X}(M)$*, such that:*

1. *The map* $X \mapsto \nabla_X$ *is* $\mathbb{R}$*-linear as well as* $C^\infty(M)$*-linear in* X*, cf.* (3.32)*;*
2. *The map* $s \mapsto \nabla_X s$ *is* $\mathbb{R}$*-linear but not* $C^\infty(M)$*-linear: it satisfies the* **Leibniz rule**

$$\nabla_X(fs) = (Xf)s + f\nabla_X s \ \ (f \in C^\infty(M)). \tag{3.56}$$

A linear connection is then a connection (in the above sense) on the tangent bundle. The general story is almost the same, including the localization of $\nabla_X s(x)$ to the flow lines of X arbitrarily close to x, and hence to any $U \in \mathscr{O}(M)$, $x \in U$. In particular, define a local frame (u_a), where $\mathsf{a} = 1, \ldots, k = \dim(E_x)$, i.e. the rank of E, by the properties that (i) $u_\mathsf{a} \in \Gamma(U, E)$, i.e., the restriction of $\Gamma(E) \equiv \Gamma(M, E)$ to some $U \in \mathscr{O}(M)$; and (ii) the set $u_\mathsf{a}(x)_{a=1,\ldots \dim(E_x)}$ forms a basis of E_x for all $x \in U$. This once again yields ***connection coefficients*** defined by

$$\nabla_\mu u_\mathsf{b} = C^\mathsf{c}_{\mu\mathsf{b}} u_\mathsf{c}. \tag{3.57}$$

The difference with the tangent bundle is that the three indices carried by C are no longer of the same type: b and c label basis vectors in E_x, whereas μ refers to the canonical coordinate base of T_xM (recall that $\nabla_\mu = \nabla_{\partial_\mu}$). Writing $s(x) = s^\mathsf{a}(x)u_\mathsf{a}(x)$, we now have

$$\nabla_\mu s^\mathsf{a} = \partial_\mu s^\mathsf{a} + C^\mathsf{a}_{\mu\mathsf{b}} s^\mathsf{b}, \tag{3.58}$$

cf. (3.39). This is often written as

$$\nabla_\mu s = \partial_\mu s + \omega_\mu s, \tag{3.59}$$

in which s is seen as a vector with components s^a relative to the given basis (u_a) and hence ω_μ is a matrix with components $C^\mathsf{a}_{\mu\mathsf{b}}$, or $s \in \Gamma(E)$ and $\omega_\mu(x) \in \mathrm{Hom}(E_x, E_x)$.[145]

A vector bundle E may be equipped with a ***metric***, i.e. nondegenerate symmetric bilinear form $g_x : E_x \times E_x \to \mathbb{R}$ defined for each $x \in M$, that is smooth in x in the sense that for any $s, t \in \Gamma(E)$ the function $g(s,t) : M \to \mathbb{R}$ defined by $x \mapsto g_x(s(x), t(x))$ is smooth. For example, a (semi) Riemannian metric on M is a metric on $E = TM$ in precisely this sense. A connection ∇ on E is then called ***metric*** if for all $s, t \in \Gamma(E)$ we have

$$X(g(s,t)) = g(\nabla_X s, t) + g(s, \nabla_X t). \tag{3.60}$$

[145] Even more abstractly, connections may be regarded as maps $\nabla : \Gamma(E) \to \Gamma(T^*M \otimes E) \equiv \Omega^1(E)$, i.e. the space of E-valued 1-forms, that satisfy $\nabla(fs) = df \otimes s + f\nabla s$; the connection with the main text is $\nabla_X s = \nabla s(X)$. In that case we may write $\nabla = d + \omega$, where $\omega \in \Omega^1(\mathrm{Hom}(E,E))$, i.e. ω is a 1-form taking values in the vector bundle $\mathrm{Hom}(E,E)$. Even more generally (for those familiar with the de Rham complex $\Omega^\bullet(M)$ and its relative $\Omega^\bullet(E)$), we may define $\nabla : \Omega^p(E) \to \Omega^{p+1}(E)$, where $p = 0, \ldots, k$ with $\Omega^0(E) \equiv \Gamma(E)$, as the unique extension of the above map $\nabla : \Omega^0(E) \to \Omega^1(E)$ that satisfies $\nabla(\alpha \otimes s) = d\alpha \otimes s + (-1)^p \alpha \wedge \nabla s$, where $\alpha \in \Omega^p(M)$ and $s \in \Gamma(E)$.

For example, the Levi-Civita connection on TM is obviously metric in this sense.

Furthermore, take $E = T^*M$, and define ∇^* in coordinates through its components by

$$\nabla^*_\mu \theta_\nu := \partial_\mu \theta_\nu - \Gamma^\rho_{\mu\nu}\theta_\rho, \tag{3.61}$$

where the $\Gamma^\rho_{\mu\nu}$ are the Christoffel symbols defined by some (semi) Riemannian metric on M, cf. (3.25). This turns out to be a connection indeed (check the axioms), whose rationale (notably of the minus sign!) is the Leibniz-type property (or product rule)

$$X(\theta(Y)) = (\nabla^*_X\theta)(Y) + \theta(\nabla_X Y), \tag{3.62}$$

which, omitting the star, may look even more elegant in the form

$$\nabla_X\langle\theta, Y\rangle = \langle\nabla_X\theta, Y\rangle + \langle\theta, \nabla_X Y\rangle, \tag{3.63}$$

where by *fiat* we have declared that on functions (such as $\langle\theta, Y\rangle \equiv \theta(Y)$) the covariant derivative ∇_X is simply X, i.e. $\nabla_X f \equiv Xf$, $f \in C^\infty(M)$. Eq. (3.62) or (3.63) might, of course, have been used to define $\nabla^* \equiv \nabla : \Omega(M) \to \Omega(M)$ in the first place, yielding (3.61). In fact, any linear connection defines a dual connection ∇^* on T^*M by (3.62).

Combining (3.39) and (3.61), we define a covariant derivative $\nabla^{(k,l)} : \mathfrak{X}^{(k,l)} \to \mathfrak{X}^{(k,l)}$ by

$$\begin{aligned}(\nabla^{(k,l)}_\mu \tau)^{\rho_1\cdots\rho_l}_{\nu_1\cdots\nu_k} \equiv \nabla^{(k,l)}_\mu \tau^{\rho_1\cdots\rho_l}_{\nu_1\cdots\nu_k} = \partial_\mu \tau^{\rho_1\cdots\rho_l}_{\nu_1\cdots\nu_k} + \Gamma^{\rho_1}_{\mu\sigma}\tau^{\sigma\cdots\rho_l}_{\nu_1\cdots\nu_k} + \cdots + \Gamma^{\rho_l}_{\mu\sigma}\tau^{\rho_1\cdots\sigma}_{\nu_1\cdots\nu_k} \\ - \Gamma^\sigma_{\mu\nu_1}\tau^{\rho_1\cdots\rho_l}_{\sigma\cdots\nu_k} - \cdots - \Gamma^\sigma_{\mu\nu_k}\tau^{\rho_1\cdots\rho_l}_{\nu_1\cdots\sigma}.\end{aligned} \tag{3.64}$$

Those who do not like coordinate definitions "by formula" may be reassured that $\nabla^{(k,l)}$ is the unique connection on $T^{(k,l)}M$ that, similarly to (3.63), satisfies the Leibniz rule

$$\begin{aligned}\nabla_X(\tau(X_1,\ldots,X_k,\theta^1,\ldots,\theta^l)) = (\nabla^{(k,l)}_X\tau)(X_1,\ldots,X_k,\theta^1,\ldots,\theta^l) \\ + \tau(\nabla_X X_1,\ldots,X_k,\theta^1,\ldots,\theta^l) + \cdots + \tau(X_1,\ldots,X_k,\theta^1,\ldots,\nabla^*_X\theta^l),\end{aligned} \tag{3.65}$$

where the case $k = l = 0$ is taken to mean $\nabla^{(0,0)}_X = X$ on $\mathfrak{X}^{(0,0)}(M) = C^\infty(M)$. Eq. (3.65) recovers $\nabla^{(0,1)} = \nabla$ on $\mathfrak{X}^{(0,1)}(M) = \mathfrak{X}(M)$ as well as $\nabla^{(1,0)} = \nabla^*$ on $\mathfrak{X}^{(1,0)}(M) = \Omega(M)$.

This construction of $\nabla^{(k,l)}$ works for any linear connection ∇. If the latter is the Levi-Civita connection, then (3.65) implies that its defining property (3.50) elegantly reads

$$\nabla^{(2,0)}_X g \equiv \nabla_X g = 0. \tag{3.66}$$

As in (3.66), in general one often writes ∇ for any $\nabla^{(k,l)}$, and physicists write (3.66) as

$$g_{\mu\nu;\sigma} = 0, \tag{3.67}$$

using the ***semi-colon notation***, in which $\tau^{\rho_1\cdots\rho_l}_{\nu_1\cdots\nu_k;\mu}$ means $\nabla_\mu\tau^{\rho_1\cdots\rho_l}_{\nu_1\cdots\nu_k}$, much as $\tau^{\rho_1\cdots\rho_l}_{\nu_1\cdots\nu_k,\mu}$ means $\partial_\mu\tau^{\rho_1\cdots\rho_l}_{\nu_1\cdots\nu_k}$. As an application, let us show once again that $d(\|\dot\gamma(t)\|)/dt = 0$ for geodesics γ:

$$\frac{d\|\dot\gamma(t)\|^2}{dt} = \frac{dg(\dot\gamma,\dot\gamma}{dt} = \dot\gamma(g(\dot\gamma,\dot\gamma)) = (\nabla_{\dot\gamma}g)(\dot\gamma,\dot\gamma) + g(\nabla_{\dot\gamma}\dot\gamma,\dot\gamma) + g(\dot\gamma,\nabla_{\dot\gamma}\dot\gamma),$$

where we used (3.65). Eqs. (3.66) and (3.48) then make the right-hand side $0+0+0=0$.

Alternatively, one may recall the description (2.64) of $T^{(k,l)}M$ as the tensor product of k copies of T^*M and l copies of TM. In general, given two vector bundles $E^{(1)} \to M$ and $E^{(2)} \to M$, with connections $\nabla^{(1)}$ and $\nabla^{(2)}$, there is a unique connection $\nabla^{(1\otimes 2)}$ on the vector bundle tensor product $E^{(1)} \otimes E^{(2)} = \sqcup_{x\in M} E^{(1)}_x \otimes E^{(2)}_x$ that satisfies the product rule

$$\nabla^{(1\otimes 2)}(s^{(1)} \otimes s^{(2)}) = \nabla^{(1)}(s^{(1)}) \otimes s^{(2)} + s^{(1)} \otimes \nabla^{(2)}(s^{(2)}). \tag{3.68}$$

This may be iterated to the tensor product of finitely many vector bundles, and hence (for any linear connection ∇) the connection $\nabla^{(k,l)}$ defined by (3.64) or (3.65) is just the tensor product of the individual connections on each copy of TM or T^*M present in $T^{(k,l)}M$.

It follows from (3.62) that (for any ∇) the connection $\nabla^{(k,l)}$ *commutes with contraction*. Contracting the first upper and lower indices and writing $\sigma^{\rho_2\cdots\rho_l}_{\nu_2\cdots\nu_k} = \tau^{\nu_1\rho_2\cdots\rho_l}_{\nu_1\nu_2\cdots\nu_k}$, one has

$$(\nabla^{(k,l)}_\mu \tau)^{\nu_1\rho_2\cdots\rho_l}_{\nu_1\nu_2\cdots\nu_k} = (\nabla^{(k,l)}_\mu \sigma)^{\rho_2\cdots\rho_l}_{\nu_2\cdots\nu_k}, \tag{3.69}$$

and similarly for any other pair of upper and lower indices. In particular, this makes the physicists' notation $\tau^{\nu_1\rho_2\cdots\rho_l}_{\nu_1\nu_2\cdots\nu_k;\mu}$ unambiguous. For example, for the Ricci tensor (see §4.5) we have

$$R_{\mu\nu;\sigma} = R^\rho_{\mu\rho\nu;\sigma}. \tag{3.70}$$

If ∇ satisfies (3.52), then $\nabla^{(k,l)}$ in addition commutes with contraction in the metric sense explained before (3.15), so that e.g., using (3.67), for the Ricci scalar we have

$$R_{,\sigma} = R_{;\sigma} = (g^{\mu\nu}R_{\mu\nu})_{;\sigma} = g^{\mu\nu}{}_{;\sigma}R_{\mu\nu} + g^{\mu\nu}R_{\mu\nu;\sigma} = g^{\mu\nu}R_{\mu\nu;\sigma}. \tag{3.71}$$

Finally, $\nabla^{(k,l)}$ may be used to rewrite the formula (2.94) for the Lie derivative as

$$\begin{aligned} \mathscr{L}_X \tau^{\rho_1\cdots\rho_l}_{\nu_1\cdots\nu_k} = \nabla_X \tau^{\rho_1\cdots\rho_l}_{\nu_1\cdots\nu_k} + (\nabla_{\nu_1}X^\nu)\tau^{\rho_1\cdots\rho_l}_{\nu\cdots\nu_k} + \cdots + (\nabla_{\nu_n}X^\nu)\tau^{\rho_1\cdots\rho_l}_{\nu_1\cdots\nu} \\ - (\nabla_\rho X^{\rho_1})\tau^{\rho\cdots\rho_l}_{\nu_1\cdots\nu_k} - \cdots - (\nabla_\rho X^{\rho_l})\tau^{\rho_1\cdots\rho}_{\nu_1\cdots\nu_k}, \end{aligned} \tag{3.72}$$

since all Christoffel symbols cancel out (check!).[146] For example, using (3.52) we obtain

$$\mathscr{L}_X g_{\mu\nu} = (\nabla_\mu X^\rho)g_{\rho\nu} + (\nabla_\nu X^\rho)g_{\mu\rho} = X_{\nu;\mu} + X_{\mu;\nu}. \tag{3.73}$$

A vector field X for which $\mathscr{L}_X g = 0$ is called a ***Killing (vector) field***.[147] Eq. (3.73) gives

$$X_{\nu;\mu} + X_{\mu;\nu} = 0. \tag{3.74}$$

Flows of Killing fields are ***isometries***, that is, diffeomorphisms preserving the metric. In the notation of (2.84), this means that $\psi_t^{(2,0)} g = g$, which is usually written as $\psi_t^* g = g$. By (2.95), Killing fields always form a Lie algebra, whose associated Lie group (up to global analytic issues) is the subgroup of $\mathrm{Diff}(M)$ consisting of isometries.

In Minkowski space-time $(\mathbb{M}, \eta)$, the Christoffels symbols vanish (at least in the usual coordinates), so that $\nabla_\mu = \partial_\mu$ and $X_{\mu;\nu} = X_{\mu,\nu}$. Hence Killing fields satisfy $\partial_\mu X_\nu = -\partial_\nu X_\mu$, whose general solution is a 10-dimensional vector space (within $\mathfrak{X}(\mathbb{R}^4)$) with basis

$$\begin{aligned} X_{(\nu)} &= \partial_\nu; & X_{(\rho\sigma)} &= x_\rho\partial_\sigma - x_\sigma\partial_\rho; \\ X^\mu_{(\nu)}(x) &= \delta^\mu_\nu \ (\nu = 0,1,2,3); & X^\mu_{(\rho\sigma)}(x) &= x_\rho\delta^\mu_\sigma - x_\sigma\delta^\mu_\rho, \ (\rho,\sigma = 0,1,2,3), \end{aligned} \tag{3.75}$$

where $x_\rho = \eta_{\rho\sigma}x^\sigma$. This is the Lie algebra of the Poincaré-group (which is the subgroup of $GL_4(\mathbb{R})$ preserving the Minkowski metric η). See also Appendix A, §§A.1 - A.2.

[146] $\mathscr{L}_X$ is not a connection (as it fails to be $C^\infty(M)$-linear in X), but $\mathscr{L}_X$ and ∇_X both satisfy the Leibniz rule.

[147] Named after the German mathematician Wilhelm Killing (1847–1923), not the movie about Cambodja.