1.3 The Levi-Civita Connection

The aim of this chapter is to define on SRMFs a 'directional derivative' of a vector field (or more generally a tensor field) in the direction of another vector field . This will be done by generalising the *covariant derivative* on hypersurfaces of $\mathbb{R}^n$, see [9, Section 3.2] to general SRMFs. Recall that for a hypersurface M in $\mathbb{R}^n$ and two vector fields X, $Y \in \mathfrak{X}(M)$ the *directional derivative* $D_X Y$ of Y in direction of X is given by ([9, (3.2.3), (3.2.4)])

$$D_X Y(p) = (D_{X_p} Y)(p) = (X_p(Y^1), \ldots, X_p(Y^n)), \tag{1.3.1}$$

where Y^i ($1 \leq i \leq n$) are the components of Y. Although X and Y are supposed to be tangential to M the directional derivative $D_X Y$ need *not* be tangential. To obtain an intrinsic notion one defines on an oriented hypersurface the *covariant derivative* $\nabla_X Y$ of Y in direction of X by the tangential projection of the directional derivative, i.e., ([9, 3.2.2])

$$\nabla_X Y = (D_X Y)^{\tan} = D_X Y - \langle D_X Y, \nu \rangle \, \nu, \tag{1.3.2}$$

where ν is the Gauss map ([9, 3.1.3]) i.e., the unit normal vector field of M such that for all p in the hypersurface $\det(\nu_p, e^1, \ldots, e^{n-1}) > 0$ for all positively oriented bases $\{e^1, \ldots, e^{n-1}\}$ of $T_p M$.
This construction clearly uses the structure of the ambient Euclidean space, which in case of a general SRMF is no longer available. Hence we will rather follow a different route and define the covariant derivative as an operation that maps a pair of vector fields to another vector field and has a list of characterising properties. Of course these properties are nothing else but the corresponding properties of the covariant derivative on hypersurfaces, that is we turn the analog of [9, 3.2.4] into a definition.

1.3.1 Definition (Connection). *A* (linear) connection *on a $\mathcal{C}^\infty$-manifold M is a map*

$$\nabla : \ \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M), \qquad (X, Y) \mapsto \nabla_X Y \tag{1.3.3}$$

such that the following properties hold

(∇1) $\nabla_X Y$ *is $\mathcal{C}^\infty(M)$-linear in X*
(i.e., $\nabla_{X_1 + f X_2} Y = \nabla_{X_1} Y + f \nabla_{X_2} Y \quad \forall f \in \mathcal{C}^\infty(M), X_1, X_2 \in \mathfrak{X}(M)$),

(∇2) $\nabla_X Y$ *is $\mathbb{R}$-linear in Y*
(i.e., $\nabla_X (Y_1 + a Y_2) = \nabla_X Y_1 + a \nabla_X Y_2 \quad \forall a \in \mathbb{R}, Y_1, Y_2 \in \mathfrak{X}(M)$),

(∇3) $\nabla_X (fY) = X(f) Y + f \nabla_X Y$ *for all $f \in \mathcal{C}^\infty(M)$ (Leibniz rule).*

We call $\nabla_X Y$ the covariant derivative *of Y in direction X w.r.t. the connection ∇.*

1.3.2 Remark (Properties of ∇).

(i) Property (∇1) implies that for fixed Y the map $X \mapsto \nabla_X Y$ is a tensor field. This fact needs some explanation. First recall that by [9, 2.6.19] tensor fields are precisely $\mathcal{C}^\infty(M)$-multilinear maps that take one forms and vector fields to smooth functions, more precisely $\mathcal{T}^r_s(M) \cong L^{r+s}_{\mathcal{C}^\infty(M)}(\Omega^1(M) \times \cdots \times \mathfrak{X}(M), \mathcal{C}^\infty(M))$. Now for $Y \in \mathfrak{X}(M)$ fixed, $A = X \mapsto \nabla_X Y$ is a $\mathcal{C}^\infty(M)$-multininear map $A : \mathfrak{X}(M) \to \mathfrak{X}(M)$ which naturally is identified with the mapping

$$\bar{A} : \Omega^1(M) \times \mathfrak{X}(M) \to \mathcal{C}^\infty(M), \qquad \bar{A}(\omega, X) = \omega(A(X)) \tag{1.3.4}$$

which is $\mathcal{C}^\infty(M)$-multilinear by (∇1), hence a $(1,1)$ tensor field.

Hence we can speak of $(\nabla_X Y)(p)$ for any p in M and moreover given $v \in T_pM$ we can define $\nabla_v Y := (\nabla_X Y)(p)$, where X is any vector field with $X_p = v$.

(ii) On the other hand the mapping $Y \to \nabla_X Y$ for fixed X is *not* a tensor field since (∇3) merely demands $\mathbb{R}$-linearity.

In the following our aim is to show that on any SRMF there is exactly one connection which is compatible with the metric. However, we need a supplementary statement, which is of substantial interest of its own. In any vector space V with scalar product g we have an identification of vectors in V with covectors in V^* via

$$V \ni v \mapsto v^\flat \in V^* \quad \text{where} \quad v^\flat(w) := \langle v, w \rangle \quad (w \in V). \tag{1.3.5}$$

Indeed this mapping is injective by nondegeneracy of g and hence an isomorphism. We will now show that this construction extends to SRMFs providing a identification of vector fields and one forms.

1.3.3 Theorem (Musical isomorphism). *Let M be a SRMF. For any $X \in \mathfrak{X}(M)$ define $X^\flat \in \Omega^1(M)$ via*

$$X^\flat(Y) := \langle X, Y \rangle \qquad \forall Y \in \mathfrak{X}(M). \tag{1.3.6}$$

Then the mapping $X \mapsto X^\flat$ is a $\mathcal{C}^\infty(M)$-linear isomorphism from $\mathfrak{X}(M)$ to $\Omega^1(M)$.

Proof. First $X^\flat : \mathfrak{X}(M) \to \mathcal{C}^\infty(M)$ is obviously $\mathcal{C}^\infty(M)$-linear, hence in $\Omega^1(M)$, cf. [9, 2.6.19]. Also the mapping $\phi : X \mapsto X^\flat$ is $\mathcal{C}^\infty(M)$-linear and we show that it is an isomorphism.

ϕ is injective: Let $\phi(X) = 0$, i.e., $\langle X, Y \rangle = 0$ for all $Y \in \mathfrak{X}(M)$ implying $\langle X_p, Y_p \rangle = 0$ for all $Y \in \mathfrak{X}(M)$, $p \in M$. Now let $v \in T_pM$ and choose a vector field $Y \in \mathfrak{X}(M)$ with $Y_p = v$. But then by nondegeneracy of $g(p)$ we obtain

$$\langle X_p, v \rangle = 0 \quad \Rightarrow \quad X_p = 0, \tag{1.3.7}$$

and since p was arbitrary we infer X=0.

ϕ is surjective: Let $\omega \in \Omega^1(M)$. We will construct $X \in \mathfrak{X}(M)$ such that $\phi(X) = \omega$. We do so in three steps.

(1) The local case: Let $(\varphi = (x^1,\dots,x^n), U)$ be a chart and $\omega|_U = \omega_i dx^i$. We set $X|_U := g^{ij}\omega_i \frac{\partial}{\partial x^j} \in \mathfrak{X}(U)$. Since (g^{ij}) is the inverse matrix of (g_{ij}) we have

$$\langle X|_U, \frac{\partial}{\partial x^k}\rangle = g^{ij}\,\omega_i\,\langle \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\rangle = \omega_i\, g^{ij}\, g_{jk} = \omega_i\,\delta^i_k = \omega_k = \omega|_U(\frac{\partial}{\partial x^k}), \tag{1.3.8}$$

and by $\mathcal{C}^\infty(M)$-linearity we obtain $X^\flat|_U = \omega|_U$.

(2) The change of charts works: We show that for any chart $(\psi = (y^1,\dots,y^n), V)$ with $U\cap V \neq \emptyset$ we have $X_U|_{U\cap V} = X_V|_{U\cap V}$. More precisely with $\omega|_V = \bar\omega_j dy^j$ and $g|_V = \bar g_{ij} dy^i \otimes dy^j$ we show that $g^{ij}\omega_i\frac{\partial}{\partial x^j} = \bar g^{ij}\bar\omega_i\frac{\partial}{\partial y^j}$.

To begin with recall that $dx^j = \frac{\partial x^j}{\partial y^i}dy^i$ ([9, 2.7.27(ii)]) and so

$$\omega|_{U\cap V} = \omega_j dx^j = \omega_j \frac{\partial x^j}{\partial y^i} dy^i = \bar\omega_i dy^i, \quad \text{implying} \quad \bar\omega_i = \omega_m \frac{\partial x^m}{\partial y^i}.$$

Moreover by [9, 2.4.15] we have $\frac{\partial}{\partial y^i} = \frac{\partial x^k}{\partial y^i}\frac{\partial}{\partial x^k}$ which gives

$$\bar g_{ij} = g\Big(\frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j}\Big) = g\Big(\frac{\partial x^k}{\partial y^i}\frac{\partial}{\partial x^k}, \frac{\partial x^l}{\partial y^j}\frac{\partial}{\partial x^l}\Big) = \frac{\partial x^k}{\partial y^i}\frac{\partial x^l}{\partial y^j}\, g\Big(\frac{\partial}{\partial y^k}, \frac{\partial}{\partial y^l}\Big) = \frac{\partial x^k}{\partial y^i}\frac{\partial x^l}{\partial y^j}\, g_{kl},$$

and so by setting $A = (a_{ki}) = (\frac{\partial x^k}{\partial y^i})$ we obtain

$$(\bar g_{ij}) = A^t (g_{ij}) A \quad \text{hence} \quad (\bar g^{ij}) = A^{-1} g^{ij} (A^{-1})^t \quad \text{and so} \quad \bar g^{ij} = \frac{\partial y^i}{\partial x^k}\, g^{kl}\, \frac{\partial y^j}{\partial x^l}.$$

Finally we obtain

$$\bar g^{ij}\,\bar\omega_i\,\frac{\partial}{\partial y^j} = \frac{\partial y^i}{\partial x^k}\, g^{kl}\, \frac{\partial y^j}{\partial x^l}\,\omega_m\,\frac{\partial x^m}{\partial y^i}\,\frac{\partial x^n}{\partial y^j}\,\frac{\partial}{\partial x^n} = g^{kl}\,\delta^m_k\,\omega_m\,\delta^n_l\,\frac{\partial}{\partial x^n} = g^{mn}\,\omega_m\,\frac{\partial}{\partial x^n}.$$

(3) Globalisation: By (2) $X(p) := X|_U(p)$ (where U is any chart neighbourhood of p) defines a vector field on M . Now choose a cover $\mathcal{U} = \{U_i |\, i\in I\}$ of M by chart neighbourhoods and a subordinate partition of unity $(\chi_i)_i$ such that $\mathrm{supp}(\chi_i) \subseteq U_i$ (cf. [9, 2.3.10]). For any $Y \in \mathfrak{X}(M)$ we then have

$$\begin{aligned}\langle X, Y\rangle &= \langle X, \sum_i \chi_i Y\rangle = \sum_i \langle X, \chi_i Y\rangle = \sum_i \langle X|_{U_i}, \chi_i, Y\rangle \\ &= \sum_i \omega|_{U_i}(\chi_i Y) = \sum_i \omega(\chi_i Y) = \omega(\sum_i \chi_i Y) = \omega(Y),\end{aligned} \tag{1.3.9}$$

and we are done. □

Hence in semi-Riemannian geometry we can always identify vectors and vector fields with covectors and one forms, respectively: X and $\phi(X) = X^\flat$ contain the same information and are called *metrically equivalent*. One also writes $\omega^\sharp = \phi^{-1}(\omega)$ and this notation is the source of the name 'musical isomorphism'. Especially in the physics literature this isomorphism is often encoded in the notation. If $X = X^i\partial_i$ is a (local) vector field then one denotes the metrically equivalent one form by $X^\flat = X_i dx^i$ and we clearly have $X_i = g_{ij}X^j$ and $X^i = g^{ij}X_j$. One also calls these operations the raising and lowering of indices. The musical isomorphism naturally extends to higher order tensors.

The next result is crucial for all the following. It is sometimes called the *fundamental Lemma of semi-Riemannian geometry.*

1.3.4 Theorem (Levi Civita connection). *Let (M, g) be a SRMF. Then there exists one and only one connection ∇ on M such that (besides the defining properties $(\nabla 1)-(\nabla 3)$ of 1.3.1) we have for all $X, Y, Z \in \mathfrak{X}(M)$*

$(\nabla 4)$ $[X, Y] = \nabla_X Y - \nabla_Y X$ *(torsion free condition)*

$(\nabla 5)$ $Z\langle X, Y\rangle = \langle\nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle$ *(metric property).*

The map ∇ is called the Levi-Civita connection *of (M, g) and it is uniquely determined by the so-called* Koszul-formula

$$\begin{aligned} 2\langle\nabla_X Y, Z\rangle =& X\langle Y, Z\rangle + Y\langle Z, X\rangle - Z\langle X, Y\rangle \\ &- \langle X, [Y, Z]\rangle + \langle Y, [Z, X]\rangle + \langle Z, [X, Y]\rangle. \end{aligned} \tag{1.3.10}$$

Proof. *Uniqueness:* If ∇ is a connection with the additional properties $(\nabla 4)$, $(\nabla 5)$ then the Koszul-formula (1.3.10) holds: Indeed denoting the right hand side of (1.3.10) by $F(X, Y, Z)$ we find

$$\begin{aligned} F(X, Y, Z) =& \langle\nabla_X Y, Z\rangle + \cancel{\langle Y, \nabla_X Z\rangle} + \cancel{\langle\nabla_Y Z, X\rangle} + \cancel{\langle Z, \nabla_Y X\rangle} - \cancel{\langle\nabla_Z X, Y\rangle} - \cancel{\langle X, \nabla_Z Y\rangle} \\ &- \cancel{\langle X, \nabla_Y Z\rangle} + \cancel{\langle X, \nabla_Z Y\rangle} + \cancel{\langle Y, \nabla_Z X\rangle} - \cancel{\langle Y, \nabla_X Z\rangle} + \langle Z, \nabla_X Y\rangle - \cancel{\langle Z, \nabla_Y X\rangle} \\ =& 2\langle\nabla_X Y, Z\rangle. \end{aligned}$$

Now by injectivity of ϕ in theorem 1.3.3, $\nabla_X Y$ is uniquely determined.

Existence: For fixed X, Y the mapping $Z \mapsto F(X, Y, Z)$ is $\mathcal{C}^\infty(M)$-linear as follows by a straight forward calculation using [9, 2.5.15(iv)]. Hence $Z \mapsto F(X, Y, Z) \in \Omega^1(M)$ and by 1.3.3 there is a (uniquely defined) vector field which we call $\nabla_X Y$ such that $2\langle\nabla_X Y, Z\rangle = F(X, Y, Z)$ for all $Z \in \mathfrak{X}(M)$. Now $\nabla_X Y$ by definition obeys the Koszul-formula and it remains to show that the properties $(\nabla 1)$–$(\nabla 5)$ hold.

$(\nabla 1)$ $\nabla_{X_1+X_2} Y = \nabla_{X_1} Y + \nabla_{X_2} Y$ follows from the fact that $F(X_1+X_2, Y, Z) = F(X_1, Y, Z) + F(X_2, Y, Z)$. Now let $f \in \mathcal{C}^\infty(M)$ then we have by [9, 2.5.15(iv)]

$$2\langle\nabla_{fX} Y - f\nabla_X Y, Z\rangle = F(X, fY, Z) - fF(X, Y, Z) = \ldots = 0, \tag{1.3.11}$$

where we have left the straight forward calculation to the reader. Hence by another appeal to theorem 1.3.3 we have $\nabla_{fX} Y = f\nabla_X Y$.

(∇2) follows since obviously $Y \mapsto F(X,Y,Z)$ is $\mathbb{R}$-linear.

(∇3) Again by [9, 2.5.15(iv)] we find

$$\begin{aligned} 2\langle \nabla_X fY, Z\rangle &= F(X, fY, Z) \\ &= X(f)\langle Y, Z\rangle - \cancel{Z(f)\langle X, Y\rangle} + \cancel{Z(f)\langle X, Y\rangle} + X(f)\langle Z, Y\rangle + fF(X,Y,Z) \\ &= 2\,\langle X(f)Y + f\nabla_X Y, Z\rangle, \end{aligned} \tag{1.3.12}$$

and the claim again follows by 1.3.3.

(∇4) We calculate

$$\begin{aligned} 2\langle \nabla_X Y - \nabla_Y X, Z\rangle &= F(X,Y,Z) - F(Y,X,Z) \\ &= \ldots = \langle Z, [X,Y]\rangle - \langle Z, [Y,X]\rangle = 2\langle [X,Y], Z\rangle \end{aligned} \tag{1.3.13}$$

and onother appeal to 1.3.3 gives the statement.

(∇5) We calculate

$$2\Big(\langle \nabla_Z X, Y\rangle + \langle X, \nabla_Z Y\rangle\Big) = F(Z,X,Y) + F(Z,Y,X) = \ldots = 2Z\big(\langle X, Y\rangle\big). \qquad \square$$

1.3.5 Remark. In the case of M being an oriented hypersurface of $\mathbb{R}^n$ the covariant derivative is given by (1.3.2). By [9, 3.2.4, 3.2.5] ∇ satisfies (∇1)–(∇5) and hence is the Levi-Civita connection of M (with the induced metric).

Next we make sure that ∇ is local in both slots, a result of utter importance.

1.3.6 Lemma (Localisation of ∇). *Let $U \subseteq M$ be open and let $X, Y, X_1, X_2, Y_1, Y_2 \in \mathfrak{X}(M)$. Then we have*

(i) If $X_1|_U = X_2|_U$ then $\big(\nabla_{X_1} Y\big)\big|_U = \big(\nabla_{X_2} Y\big)\big|_U$, and

(ii) If $Y_1|_U = Y_2|_U$ then $\big(\nabla_X Y_1\big)\big|_U = \big(\nabla_X Y_2\big)\big|_U$.

Proof.

(i) By remark 1.3.2(i): $X \mapsto \nabla_X Y$ is a tensor field hence we even have that $X_1|_p = X_2|_p$ at any point $p \in M$ implying $(\nabla_{X_1} Y)|_p = (\nabla_{X_2} Y)|_p$.

(ii) It suffices to show that $Y|_U = 0$ implies $(\nabla_X Y)|_U = 0$. So let $p \in U$ and $\chi \in \mathcal{C}^\infty(M)$ with $\mathrm{supp}(\chi) \subseteq U$ and $\chi \equiv 1$ on a neighbourhood of p. By (∇3) we then have

$$0 = (\nabla_X \underbrace{\chi Y}_{=0})|_p = \underbrace{X(\chi)}_{=0}|_p Y_p + \underbrace{\chi(p)}_{=1}(\nabla_X Y)|_p \text{ and so } (\nabla_X Y)|_U = 0. \tag{1.3.14}$$

$\square$

1.3.7 Remark. Lemma 1.3.6 allows us to restrict ∇ to $\mathfrak{X}(U) \times \mathfrak{X}(U)$: Let $X, Y \in \mathfrak{X}(U)$ and $V \subseteq \bar{V} \subseteq U$ (cf. [9, 2.3.12]) and extend X, Y by vector fields $\tilde{X}, \tilde{Y} \in \mathfrak{X}(M)$ such that $\tilde{X}|_V = X|_V$ and $\tilde{Y}|_V = Y|_V$. (This can be easily done using a partition of unity subordinate to the cover $U, M \setminus \bar{V}$, cf. [9, 2.3.14].) Now we may set $(\nabla_X Y)|_V := (\nabla_{\tilde{X}} \tilde{Y})|_V$ since by 1.3.6 this definition is independent of the choice of the extensions $\tilde{X}$, $\tilde{Y}$. Moreover we may write U as the union of such V's and so $\nabla_X Y$ is a well-defined element of $\mathfrak{X}(U)$. In particular, this allows to insert the local basis vector fields ∂_i into ∇, which will be extensively used in the following.

1.3.8 Definition (Christoffel symbols). *Let $(\varphi = (x^1, \dots, x^n), U)$ be a chart of the SRMF M. The* Christoffel symbols *(of the second kind) with respect to φ are the $\mathcal{C}^\infty$-functions $\Gamma^i{}_{jk} : U \to \mathbb{R}$ defined by*

$$\nabla_{\partial_i} \partial_j =: \Gamma^k_{ij} \partial_k \qquad (1 \leq i, j \leq n). \tag{1.3.15}$$

Since $[\partial_i, \partial_j] = 0$, property ($\nabla$4) immediately gives the symmetry of the Christoffel symbols in the lower pair of indices, i.e., $\Gamma^k_{ij} = \Gamma^k_{ji}$. Observe that Γ is not a tensor and so the Christoffel symbols do not exhibit the usual transformation behaviour of a tensor field under the change of charts. The next statement, in particular, shows how to calculate the Christoffel symbols from the metric.

1.3.9 Proposition (Christoffel symbols explicitly). *Let $(\varphi = (x^1, \dots, x^n), U)$ be a chart of the SRMF (M, g) and let $Z = Z^i \partial_i \in \mathfrak{X}(U)$. Then we have*

(i) $\Gamma^k_{ij} =: \dfrac{1}{2}\, g^{km} \Gamma_{kij} = \dfrac{1}{2}\, g^{km} \left(\dfrac{\partial g_{jm}}{\partial x^i} + \dfrac{\partial g_{im}}{\partial x^j} - \dfrac{\partial g_{ij}}{\partial x^m} \right)$,

(ii) $\nabla_{\partial_i} Z^j \partial_j = \left(\dfrac{\partial Z^k}{\partial x^i} + \Gamma^k_{ij} Z^j \right) \partial_k$.

The $\mathcal{C}^\infty(M)$-functions Γ_{kij} are sometimes called the Christoffel symbols of the first kind.
Proof.

(i) Set $X = \partial_i$, $Y = \partial_j$ and $Z = \partial_m$ in the Koszul formula (1.3.10). Since all Lie-brackets vanish we obtain

$$2\langle \nabla_{\partial_i} \partial_j, \partial_m \rangle = \partial_i g_{jm} + \partial_j g_{im} - \partial_m g_{ij}, \tag{1.3.16}$$

which upon multiplying with g^{km} gives the result.

(ii) follows immediately from (∇3) and 1.3.6. □

1.3.10 Lemma (The connection of flat space). *For $X, Y \in \mathfrak{X}(\mathbb{R}^n_r)$ with $Y = (Y^1, \ldots, Y^n) = Y^i \partial_i$ let*

$$\nabla_X Y = X(Y^i)\partial_i. \tag{1.3.17}$$

Then ∇ is the Levi-Civita connection on $\mathbb{R}^n_r$ and in natural coordinates (i.e., using id *as a global chart) we have*

(i) $g_{ij} = \delta_{ij}\varepsilon_j$ (with $\varepsilon_j = -1$ for $1 \leq j \leq r$ and $\varepsilon_j = +1$ for $r < j \leq n$),

(ii) $\Gamma^i_{jk} = 0$ for all $1 \leq i, j, k \leq n$.

Proof. Recall that in the terminology of [9, Sec. 3.2] we have $\nabla_X Y = D_X Y = p \mapsto DY(p)\, X_p$ which coincides with (1.3.17). The validity of (∇1)–(∇5) has been checked in [9, 3.2.4,5] and hence ∇ is the Levi-Civita connection. Moreover we have

(i) $g_{ij} = \langle \partial_i, \partial_j \rangle = \langle e_i, e_j \rangle = \varepsilon_i \delta_{ij}$, and

(ii) $\Gamma^i_{jk} = 0$ by (i) and 1.3.9(i). □

Next we consider vector fields with vanishing covariant derivatives.

1.3.11 Definition (Parallel vector field). *A vector field X on a SRMF M is called* parallel *if $\nabla_Y X = 0$ for all $Y \in \mathfrak{X}(M)$.*

1.3.12 Example. The coordinate vector fields in $\mathbb{R}^n_r$ are parallel: Let $Y = Y^j \partial_j$ then by 1.3.10(ii) $\nabla_Y \partial_i = Y^j \nabla_{\partial_j} \partial_i = 0$. More generally on $\mathbb{R}^n_r$ the constant vector fields are precisely the parallel ones, since

$$\nabla_Y X = 0\ \forall Y \ \Leftrightarrow\ DX(p)Y(p) = 0\ \forall Y\ \forall p \ \Leftrightarrow\ DX = 0 \ \Leftrightarrow\ X = \text{const.} \tag{1.3.18}$$

In light of this example the notion of a parallel vector field generalises the notion of a constant vector field. We now present an explicit example.

1.3.13 Example (Cylindrical coordinates on $\mathbb{R}^3$). Let (r, φ, z) be cylindrical coordinates on $\mathbb{R}^3$, i.e., $(x, y, z) = (r\cos\varphi, r\sin\varphi, z)$, see figure 1.4. This clearly is a chart on $\mathbb{R}^3 \setminus \{x \geq 0,\ y = 0\}$. Its inverse $(r, \varphi, z) \mapsto (r\cos\varphi, r\sin\varphi, z)$ is a parametrisation, hence we have (cf. [9, below 2.4.11] or directly [9, 2.4.15])

$$\partial_r = \cos\varphi\partial_x + \sin\varphi\partial_y,$$
$$\partial_\varphi = rX \quad \text{with } X = -\sin\varphi\partial_x + \cos\varphi\partial_y,$$
$$\partial_z = \partial_z.$$

Setting $y^1 = r$, $y^2 = \varphi$, $y^3 = z$ we obtain

$$g_{11} = \langle\partial_r, \partial_r\rangle = 1,$$
$$g_{22} = \langle\partial_\varphi, \partial_\varphi\rangle = r^2(\cos^2\varphi + \sin^2\varphi) = r^2,$$
$$g_{33} = \langle\partial_z, \partial_z\rangle = 1,$$
$$g_{ij} = 0 \quad \text{for all } i \neq j.$$

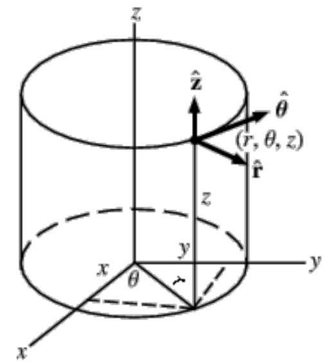

Figure 1.4: Cylindrical coordinates
♣ `fix notation` ♣

So we have

$$(g_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (g^{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.3.19}$$

and hence

$$g = g_{ij}dx^i \otimes dx^j = dr \otimes dr + r^2 d\varphi \otimes d\varphi + dz \otimes dz =: dr^2 + r^2 d\varphi^2 + dz^2.$$

By (1.3.19) $\{\partial_r, \partial_\varphi, \partial_z\}$ is orthogonal and hence (r, φ, z) is an orthogonal coordinate system. For the Christoffel symbols we find (by 1.3.9(i))

$$\Gamma^1_{\ 22} = \frac{1}{2}g^{1l}\Gamma_{l22} = \frac{1}{2}g^{11}\Gamma_{122} = \frac{1}{2}\,1\,(\underbrace{g_{12,2}}_{=0} + \underbrace{g_{21,2}}_{=0} - g_{22,1}) = \frac{-1}{2}\,2r = -r,$$
$$\Gamma^2_{\ 12} = \Gamma^2_{\ 21} = \frac{1}{2}g^{2l}\Gamma_{l21} = \frac{1}{2}g^{22}\Gamma_{221} = \frac{1}{2}\frac{1}{r^2}\,(g_{22,1} + \underbrace{g_{12,2}}_{=0} - \underbrace{g_{21,2}}_{=0}) = \frac{1}{2r^2}\,2r = \frac{1}{r},$$

and all other $\Gamma^i_{\ jk} = 0$. Hence we have $\nabla_{\partial_i}\partial_j = 0$ for all i, j with the exception of

$$\nabla_{\partial_\varphi}\partial_\varphi = -r\partial_r\,, \quad \text{and } \nabla_{\partial_\varphi}\partial_r = \nabla_{\partial_r}\partial_\varphi = \frac{1}{r}\partial_\varphi = X\,.$$

By figure 1.4 we see that ∂_r and ∂_φ are parallel if one moves in the z-direction. We hence expect that $\nabla_{\partial_z}\partial_\varphi = 0 = \nabla_{\partial_z}\partial_r$ which also results from our calculations. Moreover ∂_z is parallel since it is a coordinate vector field in the natural basis of $\mathbb{R}^3$, cf. 1.3.12.

Our next aim is to extend the covariant derivative to tensor fields of general rank. We will start with a slight detour introducing the notion of a *tensor derivation* and its basic properties and then use this machinery to extend the covariant derivative to the space $\mathcal{T}^r_s(M)$ of tensor fields of rank (r, s).

Interlude: Tensor derivations

In this brief interlude we introduce some basic operations on tensor fields which will be essential in the following. We recall (for more information on tensor fields see e.g. [9, Sec. 2.6]) that a tensor field $A \in \mathcal{T}^r_s(M) = \Gamma(M, T^r_s(M))$ is a (smooth) section of the (r,s)-tensor bundle $T^r_s(M)$ of M. That is to say that for any point $p \in M$, the value of the tensor field $A(p)$ is a multilinear map

$$A(p): \quad \underbrace{T_pM^* \times \cdots \times T_pM^*}_{r \text{ times}} \times \underbrace{T_pM \times \cdots \times T_pM}_{s \text{ times}} \to \mathbb{R}. \tag{1.3.20}$$

Locally in a chart $(\psi = (x^1, \ldots, x^n), V)$ we have

$$A|_V = A^{i_1 \ldots i_r}_{j_1 \ldots j_s} \, \partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_s}, \tag{1.3.21}$$

where the coefficient functions are given for $q \in V$ by

$$A^{i_1 \ldots i_r}_{j_1 \ldots j_s}(q) = A(q)(dx^{i_1}|_q, \ldots, dx^{i_r}|_q, \partial_{j_1}|_q, \ldots, \partial_{j_s}|_q). \tag{1.3.22}$$

The space $\mathcal{T}^r_s(M)$ of can be identified with the space

$$L^{r+s}_{\mathcal{C}^\infty(M)}(\underbrace{\Omega^1(M) \times \cdots \times \Omega^1(M)}_{r\text{-times}} \times \underbrace{\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)}_{s\text{-times}}, \mathcal{C}^\infty(M)) \tag{1.3.23}$$

of $\mathcal{C}^\infty(M)$-multilinear maps form one-forms and vector fields to smooth functions. Recall also the special cases $\mathcal{T}^0_0(M) = \mathcal{C}^\infty(M)$, $\mathcal{T}^1_0(M) = \mathfrak{X}(M)$, and $\mathcal{T}^0_1(M) = \Omega^1(M)$.

Additionally will also frequently deal with the following situation, which generalises the one of 1.3.2(i): If $A : \mathfrak{X}(M)^s \to \mathfrak{X}(M)$ is a $\mathcal{C}^\infty(M)$-multilinear mapping then we define

$$\begin{aligned} \bar{A} &: \Omega^1(M) \times \mathfrak{X}(M)^s \to \mathcal{C}^\infty(M) \\ &\bar{A}(\omega, X_1, \ldots, X_s) := \omega(A(X_1, \ldots, X_s)). \end{aligned} \tag{1.3.24}$$

Clearly $\bar{A}$ is $\mathcal{C}^\infty(M)$-multilinear and hence a $(1,s)$-tensor field and we will frequently and tacitly identify $\bar{A}$ and A.

We start by introducing a basic operation on tensor fields that shrinks their rank from (r,s) to $(r-1, s-1)$. The general definition is based on the following special case.

1.3.14 Lemma ($(1,1)$-contraction). *There is a unique $\mathcal{C}^\infty(M)$-linear map $\mathcal{C} : \mathcal{T}^1_1(M) \to \mathcal{C}^\infty(M)$ called the $(1,1)$-contraction such that*

$$\mathcal{C}(X, \omega) = \omega(X) \qquad \textit{for all } X \in \mathfrak{X}(M) \textit{ and } \omega \in \Omega^1(M). \tag{1.3.25}$$

Proof. Since $\mathcal{C}$ is to be $\mathcal{C}^\infty(M)$-linear it is a pointwise operation, cf. [9, 2.6.19] and we start by giving a local definition. For the natural basis fields of a chart $(\varphi = (x^1, \ldots, x^n), V)$ we necessarily have $\mathcal{C}(\partial_j, dx^i) = dx^i(\partial_j) = \delta^i_j$ and so for $\mathcal{T}^1_1 \ni A = \sum A^i_j \partial_i \otimes dx^j$ we are forced to define

$$\mathcal{C}(A) = \sum_i A^i_i = \sum_i A(dx^i, \partial_j). \tag{1.3.26}$$

It remains to show that the definition is independent of the chosen chart. Let $(\psi = (y^1, \dots, y^n), V)$ be another chart then we have using [9, 2.7.27(iii)] as well as the summation convention

$$A(dy^m, \partial_m) = A\left(\frac{\partial y^m}{\partial x^i}\, dx^i, \frac{\partial x^j}{\partial y^m} \partial x_j\right) = \underbrace{\frac{\partial y^m}{\partial x^i} \frac{\partial x^j}{\partial y^m}}_{\delta^j_i} A(dx^i, \partial x_j) = A(dx^i, \partial x_i). \tag{1.3.27}$$

□

To define the contraction for general rank tensors let $A \in \mathcal{T}^r_s(M)$, fix $1 \le i \le r$, $1 \le j \le s$ and let $\omega^1, \dots, \omega^{r-1} \in \Omega^1(M)$ and $X_1, \dots, X_{s-1} \in \mathfrak{X}(M)$. Then the map

$$\Omega(M) \times \mathfrak{X}(M) \ni (\omega, X) \mapsto A(\omega^1, \dots, \underset{i}{\omega}, \dots, \omega^{r-1}, X_1, \dots, \underset{j}{X}, \dots, X_{s-1}) \tag{1.3.28}$$

is a $(1,1)$-tensor. We now apply the $(1,1)$-contraction $\mathcal{C}$ of 1.3.14 to (1.3.28) to obtain a $\mathcal{C}^\infty(M)$-function denoted by

$$(\mathcal{C}^i_j A)(\omega^1, \dots, \omega^{n-1}, X_1, \dots, X_{s-1}). \tag{1.3.29}$$

Obviously $\mathcal{C}^i_j A$ is $\mathcal{C}^\infty(M)$-linear in all its slots, hence it is a tensor field in $\mathcal{T}^{r-1}_{s-1}(M)$ which we call the *(i,j)-contraction of A*. We illustrate this concept by the following examples.

1.3.15 Examples (Contraction).

(i) Let $A \in \mathcal{T}^2_3(M)$ then $\mathcal{C}^1_3 A \in \mathcal{T}^1_2$ is given by

$$\mathcal{C}^1_3 A(\omega, X, Y) = \mathcal{C}\big(A(\,.\,, \omega, X, Y, .)\big) \tag{1.3.30}$$

which locally takes the form

$$(\mathcal{C}^1_3 A)^k_{ij} = (\mathcal{C}^1_3(A)(dx^k, \partial_i, \partial_j) = \mathcal{C}\big(A(\,.\,, dx^k, \partial_i, \partial_j, .)\big) = A(dx^m, dx^k, \partial_i, \partial_j, \partial_m) = A^{mk}_{ijm},$$

where of course we again have applied the summation convention.

(ii) More generally the components of $\mathcal{C}^k_l A$ of $A \in \mathcal{T}^r_s(M)$ in local coordinates take the form $A^{i_1 \dots \overset{k}{m} \dots i_r}_{j_1 \dots \underset{l}{m} \dots j_s}$.

Now we may define the notion of a tensor derivation announced above as map on tensor fields that satisfies a product rule and commutes with contractions.

1.3.16 Definition (Tensor derivation). *A* tensor derivation $\mathcal{D}$ *on a smooth manifold* M *is a family of* $\mathbb{R}$*-linear maps*

$$\mathcal{D} = \mathcal{D}^r_s : \ \mathcal{T}^r_s(M) \to \mathcal{T}^r_s(M) \quad (r, s \geq 0) \tag{1.3.31}$$

such that for any pair A, B *of tensor fields we have*

(i) $\mathcal{D}(A \otimes B) = \mathcal{D}A \otimes B + A \otimes \mathcal{D}B$

(ii) $\mathcal{D}(\mathcal{C}A) = \mathcal{C}(\mathcal{D}A))$ *for any contraction* $\mathcal{C}$.

The product rule in the special case $f \in \mathcal{C}^\infty(M) = \mathcal{T}^0_0(M)$ and $A \in \mathcal{T}^r_s(M)$ takes the form

$$\mathcal{D}(f \otimes A) = \mathcal{D}(fA) = (\mathcal{D}f)A + f\mathcal{D}A. \tag{1.3.32}$$

Moreover for $r = 0 = s$ the tensor derivation $\mathcal{D}^0_0$ is a derivation on $\mathcal{C}^\infty(M)$ (cf. [9, 2.5.12]) and so by [9, 2.5.13] there exists a unique vector field $X \in \mathfrak{X}(M)$ such that

$$\mathcal{D}f = X(f) \quad \text{for all } f \in \mathcal{C}^\infty(M). \tag{1.3.33}$$

Despite the fact that tensor derivations are *not* $\mathcal{C}^\infty(M)$-linear and hence not pointwise defined[2] (cf. [9, 2.6.19]) they are local operators in the following sense.

1.3.17 Proposition (Tensor derivations are local). *Let* $\mathcal{D}$ *be a tensor derivation on* M *and let* $U \subseteq M$ *be open. Then there exists a unique tensor derivation* $\mathcal{D}_U$ *on* U*, called the* restriction of $\mathcal{D}$ to U *statisfying*

$$\mathcal{D}_U(A|_U) = (\mathcal{D}A)|_U \tag{1.3.34}$$

for all tensor fields A *on* M*.*

Proof. Let $B \in \mathcal{T}^r_s(U)$ and $p \in U$. Choose a cut-off function $\chi \in \mathcal{C}^\infty_0(U)$ with $\chi \equiv 1$ in a neighbourhood of p. Then $\chi B \in \mathcal{T}^r_s(M)$ and we define

$$(\mathcal{D}_U B)(p) := \mathcal{D}(\chi B)(p). \tag{1.3.35}$$

We have to check that this definition is valid and leads to the asserted properties.

(1) The definition is independent of χ: choose another cut-off function $\tilde{\chi}$ at p and set $f = \chi - \tilde{\chi}$. Then choosing a function $\varphi \in \mathcal{C}^\infty_0(U)$ with $\varphi \equiv 1$ on $\operatorname{supp}(f)$ we obtain

$$\mathcal{D}(fB)(p) = \mathcal{D}(f\varphi B)(p) = \mathcal{D}(f)|_p(\varphi B)(p) + \underbrace{f(p)}_{=0}\mathcal{D}(\varphi B)(p) = 0, \tag{1.3.36}$$

since we have with a vector field X as in (1.3.33) that $\mathcal{D}f(p) = X(f)(p) = 0$ by the fact that $f \equiv 0$ near the point p.

[2] Recall from analysis that taking a derivative of a function is *not* a pointwise operation: It depends on the values of the function in a neighbourhood.

(2) $\mathcal{D}_U B \in \mathcal{T}^r_s(U)$ since for all $V \subseteq U$ open we have $\mathcal{D}_U B|V = \mathcal{D}(\chi B)|_V$ by definition if $\chi \equiv 1$ on V. Now observe that $\chi B \in \mathcal{T}^r_s(M)$.

(3) Clearly $\mathcal{D}_U$ is a tensor derivation on U since $\mathcal{D}$ is is a tensor derivation on M.

(4) $\mathcal{D}_U$ has the restriction property (1.3.34) since if $B \in \mathcal{T}^r_s(M)$ we find for all $p \in U$ that $\mathcal{D}_U(B|_U)(p) = \mathcal{D}(\chi B|_U)(p) = \mathcal{D}(\chi B)(p)$ and $\mathcal{D}(\chi B)(p) = \mathcal{D}(B)(p)$ since $\mathcal{D}((1-\chi)B)(p) = 0$ by the same argument as used in (1.3.36).

(5) Finally $\mathcal{D}_U$ is uniquely determined: Let $\tilde{\mathcal{D}}_u$ be another tensor derivation that satisfies (1.3.34) then for $B \in \mathcal{T}^r_s(U)$ we again have $\tilde{\mathcal{D}}_u((1-\chi)B)(p) = 0$ and so by (4)

$$\tilde{\mathcal{D}}_U(B)(p) = \tilde{\mathcal{D}}_U(\chi B)(p) = \mathcal{D}(\chi B)(p) = \mathcal{D}_U(B)(p)$$

for all $p \in U$. □

We next state and prove a product rule for tensor derivations.

1.3.18 Proposition (Product rule). *Let $\mathcal{D}$ be a tensor derivation on M. Then we have for $A \in \mathcal{T}^r_s(M)$, $\omega^1, \dots, \omega^r \in \Omega(M)$, and $X_1, \dots, X_s \in \mathfrak{X}(M)$*

$$\begin{aligned}\mathcal{D}\Big(A(\omega^1, \dots, \omega^r, X_1, \dots, X_s)\Big) =& (\mathcal{D}A)(\omega^1, \dots, \omega^r, X_1, \dots, X_s) \\ &+ \sum_{i=1}^{r} A(\omega^1, \dots, \mathcal{D}\omega^i, \dots, \omega^r, X_1, \dots, X_s) \\ &+ \sum_{j=1}^{s} A(\omega^1, \dots, \omega^r, X_1, \dots, \mathcal{D}X_j, \dots, X_s).\end{aligned} \tag{1.3.37}$$

Proof. We only show the case $r = 1 = s$ since the general case follows in complete analogy. We have $A(\omega, X) = \bar{\mathcal{C}}(A \otimes \omega \otimes X)$ where $\bar{\mathcal{C}}$ is a composition of two contractions. Indeed in local coordinates $A \otimes \omega \otimes X$ has components $A^i_j \omega_k X^l$ and $A(\omega, X) = A(\omega_i dx^i, X^j \partial_j) = \omega_i X^j A(dx^i, \partial_j) = A^i_j \omega_i X^j$ and the claim follows from 1.3.15(ii).
By 1.3.16(i)–(ii) we hence have

$$\begin{aligned}\mathcal{D}\big(A(\omega, X)\big) &= \mathcal{D}\big(\bar{\mathcal{C}}(A \otimes \omega \otimes X)\big) = \bar{\mathcal{C}}\mathcal{D}(A \otimes \omega \otimes X) \\ &= \bar{\mathcal{C}}(\mathcal{D}A \otimes \omega \otimes X) + \bar{\mathcal{C}}(A \otimes \mathcal{D}\omega \otimes X) + \bar{\mathcal{C}}(A \otimes \omega \otimes \mathcal{D}X) \\ &= \mathcal{D}A(\omega, X) + A(\mathcal{D}\omega, X) + A(\omega, \mathcal{D}X).\end{aligned} \tag{1.3.38}$$

□

The product rule (1.3.37) can obviously be solved for the term involving $\mathcal{D}A$ resulting in a formula for the tensor derivation of a general tensor field A in terms of $\mathcal{D}$ only acting on functions, vector fields, and one-forms. Moreover for a one form ω we have by (1.3.37)

$$(\mathcal{D}\omega)(X) = \mathcal{D}(\omega(X)) - \omega(\mathcal{D}X) \tag{1.3.39}$$

and so the action of a tensor derivation is determined by its action on functions and vector fields alone, a fact which we state as follows.

1.3.19 Corollary. *If two tensor derivations $\mathcal{D}_1$ and $\mathcal{D}_2$ agree on functions $\mathcal{C}^\infty(M)$ and on vector fields $\mathfrak{X}(M)$ then they agree on all tensor fields, i.e., $\mathcal{D}_1 = \mathcal{D}_2$.*

More importantly a tensor derivation can be constructed from its action on just functions and vector fields in the following sense.

1.3.20 Theorem (Constructing tensor derivations). *Given a vector field $V \in \mathfrak{X}(M)$ and an $\mathbb{R}$-linear mapping $\delta : \mathfrak{X}(M) \to \mathfrak{X}(M)$ obeying the product rule*

$$\delta(fX) = V(f)X + f\delta(X) \quad \text{for all } f \in \mathcal{C}^\infty(M),\ \mathfrak{X} \in \mathfrak{X}(M). \tag{1.3.40}$$

Then there exists a unique tesor derivation $\mathcal{D}$ on M such that $\mathcal{D}^0_0 = V : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M)$ and $\mathcal{D}^1_0 = \delta : \mathfrak{X}(M) \to \mathfrak{X}(M)$.

Proof. Uniqueness is a consequence of 1.3.19 and we are left with constructing $\mathcal{D}$ using the product rule.
To begin with, by (1.3.39) we necessarily have for any one-form ω

$$(\mathcal{D}\omega)(X) \equiv (\mathcal{D}^0_1\omega)(X) = V(\omega(X)) - \omega(\delta(X)), \tag{1.3.41}$$

which obviously is $\mathbb{R}$-linear. Moreover, $\mathcal{D}\omega$ is $\mathcal{C}^\infty(M)$-linear hence a one-form since

$$\begin{aligned}\mathcal{D}\omega(fX) &= V(\omega(fX)) - \omega(\delta(fX)) = V(f\omega(X)) - \omega(V(f)X) - \omega(f\delta(X))\\ &= fV(\omega(X)) + \cancel{V(f)\omega(X)} - \cancel{V(f)\omega(X)} - f\omega(\delta(X))\\ &= f\big(V(\omega(X)) - \omega(\delta(X)\big) = f\mathcal{D}\omega(X).\end{aligned} \tag{1.3.42}$$

Similarly for higher ranks $r + s \geq 2$ we have to define $\mathcal{D}^r_s$ by the product rule (1.3.37): Again it is easy to see that $\mathcal{D}^r_s$ is $\mathbb{R}$-linear and that $\mathcal{D}^r_s A$ is $\mathcal{C}^\infty(M)$-multilinear hence in $\mathcal{T}^r_s(M)$.
We now have to verify (i), (ii) of definition 1.3.16. We only show $\mathcal{D}(A \otimes B) = \mathcal{D}A \otimes B + B \otimes \mathcal{D}A$ in case $A, B \in \mathcal{T}^1_1(M)$, the general case being completely analogous:

$$\begin{aligned}\big(\mathcal{D}(A \otimes B)\big)(\omega^1, \omega^2, X_1, X_2) =& V(A(\omega^1, X_1) \cdot B(\omega^2, X_2))\\ &- \Big(A(\mathcal{D}\omega^1, X_1)B(\omega^2, X_2) + A(\omega^1, X_1)B(\mathcal{D}\omega^2, X_2)\Big)\\ &- \Big(A(\omega^1, \mathcal{D}X_1)B(\omega^2, X_2) + A(\omega^1, X_1)B(\omega^2, \mathcal{D}X_2)\Big)\\ =& \Big(V\big(A(\omega^1, X_1)\big) - A(\mathcal{D}\omega^1, X_1) - A(\omega^1, \mathcal{D}X_1)\Big)B(\omega^2, X_2)\\ &+ A(\omega^1, X_1)\Big(V\big(B(\omega^2, X_2)\big) - B(\mathcal{D}\omega^2, X_2) - B(\omega^2, \mathcal{D}X_2)\Big)\\ =& (\mathcal{D}A \otimes B + A \otimes \mathcal{D}B)(\omega^1, \omega^2, X_1, X_2).\end{aligned}$$

Finally, we show that $\mathcal{D}$ commutes with contractions. We start by considering $\mathcal{C} : \mathcal{T}_1^1(M) \to \mathcal{C}^\infty(M)$. Let $A = X \otimes \omega \in \mathcal{T}_1^1(M)$, then we have by (1.3.41)

$$\mathcal{D}(\mathcal{C}(X \otimes \omega)) = \mathcal{D}(\omega(X)) = V(\omega(X)) = \omega(\delta(X)) + \mathcal{D}(\omega)(X), \tag{1.3.43}$$

which agrees with

$$\mathcal{C}(\mathcal{D}(X \otimes \omega)) = \mathcal{C}(\mathcal{D}X \otimes \omega + X \otimes \mathcal{D}\omega) = \omega(\mathcal{D}X) + (\mathcal{D}\omega)(X). \tag{1.3.44}$$

Obviously the same holds true for (finite) sums of terms of the form $\omega^i \otimes X_i$. Since $\mathcal{D}$ is local (proposition 1.3.17) and $\mathcal{C}$ is even pointwise it suffices to prove the statement in local coordinates. But there each $(1,1)$-tensor is a sum as mentioned above. The extension to the general case is now straight forward. We only explicitly check it for $A \in \mathcal{T}_2^1(M)$:

$$\begin{aligned}\big(\mathcal{D}_1^0(\mathcal{C}_2^1 A)\big)(X) = \mathcal{D}_0^0\big((\mathcal{C}_2^1 A)(X)\big) - (\mathcal{C}_2^1 A)(\mathcal{D}_0^1 X) = \mathcal{D}_0^0\big(\mathcal{C}(A(\,.\,,X,.))\big) - \mathcal{C}\big(A(\,.\,,\mathcal{D}X,.)\big)\\
&= \mathcal{C}\Big(\mathcal{D}_1^1\big(A(\,.\,,X,.)\big) - A(\,.\,,\mathcal{D}X,.)\Big)\\
&= \mathcal{C}\Big((\omega,Y) \mapsto \mathcal{D}\big(A(\omega,X,Y)\big) - A(\mathcal{D}\omega,X,Y) - A(\omega,X,\mathcal{D}Y) - A(\omega,\mathcal{D}X,Y)\Big)\\
&= \mathcal{C}\big((\omega,Y) \mapsto (\mathcal{D}A)(\omega,X,Y)\big) = \big(\mathcal{C}_2^1(\mathcal{D}A)\big)(X).\end{aligned}$$

□

As a first important example of a tensor derivation we discuss the Lie derivative.

1.3.21 Example (Lie derivative on $\mathcal{T}_s^r$). Let $X \in \mathfrak{X}(M)$. Then we define the tensor derivative L_X, called the *Lie derivative* with respect to X by setting

$$\begin{aligned} L_X(f) &= X(f) \quad \text{for all } f \in \mathcal{C}^\infty(M), \text{ and}\\ L_X(Y) &= [X,Y] \quad \text{for all vector fields } Y \in \mathfrak{X}(M).\end{aligned}$$

Indeed this definition generalises the Lie derivative or Lie bracket of vector fields to general tensors in $\mathcal{T}_s^r(M)$ since by theorem 1.3.20 we only have to check that $\delta(Y) = L_X(Y) = [X,Y]$ satisfies the product rule (1.3.40). But this follows immediately form the corresponding property of the Lie bracket, see [9, 2.5.15(iv)].

Finally we return to the Levi-Civita covariant derivative on a SRMF (M,g), cf. 1.3.4. We want to extend it from vector fields to arbitrary tensor fields using theorem 1.3.20. A brief glance at the assumptions of the latter theorem reveals that the defining properties (∇2) and (∇3) are all we need. So the following definition is valid.

1.3.22 Definition (Covariant derivative for tensors). *Let M be a SRMF and $X \in \mathfrak{X}(M)$. The* (Levi-Civita) covariant derivative ∇_X *is the uniquely determined tensor derivation on M such that* *(i)* $\nabla_X f = X(f)$ *for all* $f \in \mathcal{C}^\infty(M)$*, and*

(ii) $\nabla_X Y$ *is the Levi-Civita covariant derivative of* Y *w.r.t.* X *as given by 1.3.4.*

The covariant derivative w.r.t. a vector field X is a generalisation of the directional derivative. Similar to the case of multi-dimensional calculus in $\mathbb{R}^n$ we may collect all such directional derivatives into one differential. To do so we need to take one more technical step.

1.3.23 Lemma. *Let* $A \in \mathcal{T}^r_s(M)$*, then the mapping*

$$\mathfrak{X}(M) \ni X \mapsto \nabla_X A \in \mathcal{T}^r_s(M)$$

is $\mathcal{C}^\infty(M)$*-linear.*

Proof. We have to show that for X_1, $X_2 \in \mathfrak{X}(M)$ and $f \in \mathcal{C}^\infty(M)$ we have

$$\nabla_{X_1+fX_2} A = \nabla_{X_1} A + f\nabla_{X_2} A \quad \text{for all } A \in \mathcal{T}^r_s(M). \tag{1.3.45}$$

However, by 1.3.20 we only have to show this for $A \in \mathcal{T}^0_0(M) = \mathcal{C}^\infty(M)$ and $A \in \mathcal{T}^1_0(M) = \mathfrak{X}(M)$. But for $A \in \mathcal{C}^\infty(M)$ equation (1.3.45) holds by definition and for $A \in \mathfrak{X}(M)$ this is just property (∇1). □

1.3.24 Definition (Covariant differential). *For* $A \in \mathcal{T}^r_s(M)$ *we define the* covariant differential $\nabla A \in \mathcal{T}^r_{s+1}$ *of* A *as*

$$\nabla A(\omega^1, \dots, \omega^r, X_1, \dots, X_s, X) := (\nabla_X A)(\omega^1, \dots, \omega^r, X_1, \dots, X_s) \tag{1.3.46}$$

for all $\omega^1, \dots, \omega^r \in \Omega^1(M)$ *and* $X_1, \dots, X_s \in \mathfrak{X}(M)$*.*

1.3.25 Remark.

(i) In case $r = 0 = s$ the covariant differential is just the exterior derivative since for $f \in \mathcal{C}^\infty(M)$ and $X \in \mathfrak{X}(M)$ we have

$$(\nabla f)(X) = \nabla_X f = X(f) = df(X). \tag{1.3.47}$$

(ii) ∇A is a 'collection' all the covariant derivatives $\nabla_X A$ into one object. The fact that the covariant rank is raised by one, i.e., that $\nabla A \in \mathcal{T}^r_{s+1}(M)$ for $A \in \mathcal{T}^r_s(M)$ is the source of the name *covariant* derivative/differential.

(iii) In complete analogy with vector fields (cf. definition 1.3.11) we call $A \in \mathcal{T}^r_s(M)$ *parallel* if $\nabla_X A = 0$ for all $X \in \mathfrak{X}(M)$ which we can now simply write as $\nabla A = 0$.

(iv) The metric condition ($\nabla 5$) just says that g itself is parallel since by the product rule 1.3.18 we have for all X, Y, $Z \in \mathfrak{X}(M)$

$$(\nabla_Z g)(X,Y) = \nabla_Z(g(X,Y)) - g(\nabla_Z X, Y) - g(X, \nabla_Z, Y) \tag{1.3.48}$$

which vanishes by ($\nabla 5$).

(v) If in a local chart the tensor field $A \in \mathcal{T}^r_s(M)$ has components $A^{i_1\dots i_r}_{j_1\dots j_s}$ the components of its covariant differential $\nabla A \in \mathcal{T}^r_{s+1}(M)$ are denoted by $A^{i_1\dots i_r}_{j_1\dots j_s;k}$ and take the form

$$A^{i_1\dots i_r}_{j_1\dots j_s;k} = \frac{\partial A^{i_1\dots i_r}_{j_1\dots j_s}}{\partial x^k} + \sum_{l=1}^{r} \Gamma^{i_l}_{km} A^{i_1\dots m \dots i_r}_{j_1\dots\dots\dots j_s} - \sum_{l=1}^{s} \Gamma^{m}_{kj_l} A^{i_1\dots\dots\dots i_r}_{j_1\dots m \dots j_s}\,. \tag{1.3.49}$$

Our next topic is the notion of a covariant derivative of vector fields which are not defined on all of M but just, say on (the image of) a curve. Of course then we can only expect to be able to define a derivative of the vector field in the direction of the curve. Intuitively such a notion corresponds to the rate of change of the vector field as we go along the curve. We begin by making precise the notion of such vector fields but do not restrict ourselves to the case of curves.

1.3.26 Definition (Vector field along a mapping). *Let N, M be smooth manifolds and let $f \in \mathcal{C}^\infty(N, M)$. A* vector field along f *is a smooth mapping*

$$Z:\ N \to TM \quad \textit{such that } \pi \circ Z = f, \tag{1.3.50}$$

where $\pi : TM \to M$ is the vector bundle projection. We denote the $\mathcal{C}^\infty(N)$-module of all vector fields along f by $\mathfrak{X}(f)$.

The definition hence says that $Z(p) \in T_{f(p)}M$ for all points $p \in N$. In the special case of $N = I \subseteq \mathbb{R}$ a real interval and $f = c : I \to M$ a $\mathcal{C}^\infty$-curve we call $\mathfrak{X}(c)$ the space of *vector fields along the curve c*. In particular, in this case $t \mapsto \dot c(t) \equiv c'(t) \in \mathfrak{X}(c)$. More precisely we have (cf. [9, below (2.5.3)]) $c'(t) = T_t c(1) = T_t c(\frac{\partial}{\partial t}|_t) \in T_{c(t)}M$. Also recall for later use that for any $f \in \mathcal{C}^\infty(M)$ we have $c'(t)(f) = T_t c(\frac{d}{dt}|_t)(f) = \frac{d}{dt}|_t(f \circ c)$ and consequently in coordinates $\varphi = (x^1, \dots, x^n)$ the local expression of the velocity vector takes the form $c'(t) = c'(t)(x^i)\partial_i|_{c(t)} = \frac{d}{dt}|_t(x^i \circ c)\partial_i|_{c(t)}$. (For more details see e.g. [11, 1.17 and below].)
In case M is a SRMF we may use the Levi-Civita covariant derivative to define the derivative Z' of $Z \in \mathfrak{X}(c)$ along the curve c.

1.3.27 Proposition (Induced covariant derivative). *Let $c : I \to M$ be a smooth curve into the SRMF M. Then there exists a unique mapping $\mathfrak{X}(c) \to \mathfrak{X}(c)$*

$$Z \mapsto Z' \equiv \frac{\nabla Z}{dt} \tag{1.3.51}$$

called the induced covariant derivative *such that*