FMA 52068 EP
First part: Introduction to Mathematical General Relativity

Arthur Touati

January 31, 2025

Contents

These lectures are based on the following books:

- *Introduction to Smooth Manifolds*, John M. Lee,
- *Semi-Riemannian Geometry, With Applications to Relativity*, Barrett O'Neill,
- *General Relativity*, Robert M. Wald.

Chapter 1

Differential Geometry in a Nutshell

This first chapter gives a light introduction to the theory of smooth manifold. The general idea is to transport from the usual Euclidean space $\mathbb{R}^n$ the tools of differential calculus and linear algebra to more general spaces, by relying on their local nature. The emphasis is put on the notion of tensors and their manipulation.

1.1 Smooth manifolds

Roughly speaking, a (smooth) manifold is a topological space that looks locally like $\mathbb{R}^n$. The notion of atlas makes this idea rigorous.

Definition 1.1. *Let $\mathcal{M}$ be a topological space and $n \geq 1$. An atlas of dimension n is a family $\big((U_i, \varphi_i)\big)_{i \in I}$ such that*

(i) each U_i is an open subset of $\mathcal{M}$ and

$$\mathcal{M} = \bigcup_{i \in I} U_i,$$

(ii) for each $i \in I$ there exists an open subset $V_i \subset \mathbb{R}^n$ such that

$$\varphi_i : U_i \longrightarrow V_i$$

is an homoemorphism.

Since $\mathcal{M}$ is already a topological space, the notion of continuity required in the second item of Definition 1.1 is well-defined. Moreover, recall that this second item simply requires that $\varphi_i : U_i \longrightarrow V_i$ is a continuous bijection with a continuous inverse. For a given $i \in I$, (U_i, φ_i) is called a local chart on $\mathcal{M}$. It allows to define local coordinates on $\mathcal{M}$. Indeed, if π^j denotes the projection on the j-th coordinate in $\mathbb{R}^n$ (i.e $\pi^j(y) = y^j$ for $y = (y^1, \dots, y^n) \in \mathbb{R}^n$) then we define

$$x^j := \pi^j \circ \varphi_i, \tag{1.1}$$

which is a $\mathbb{R}$-valued function defined on U_i. The functions x^j are the coordinate functions associated to the chart (U_i, φ_i). We can now define the notion of smooth manifold.

Definition 1.2. *Let $n \geq 1$. A topological space $\mathcal{M}$ is called a smooth manifold of dimension n if*

(i) $\mathcal{M}$ is Hausdorff, i.e two distinct points always have disjoint open neighborhoods,

(ii) $\mathcal{M}$ admits a countable atlas $\big((U_i, \varphi_i)\big)_{i \in I}$ of dimension n such that the transition maps

$$\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \longrightarrow \varphi_i(U_i \cap U_j)$$

are smooth diffeomorphisms.

Note that $\mathbb{R}^n$ is obviously a smooth manifold of dimension n (it suffices to choose the identity function as local charts). The requirement that the topology is Hausdorff in Definition 1.2 avoids wild and unwanted behaviours. Moreover, the transition maps $\varphi_i \circ \varphi_j^{-1}$ are defined between open subsets of $\mathbb{R}^n$ so their smoothness is well-defined. The same comment applies for the following definition.

Definition 1.3. *Let $\mathcal{M}$ and $\tilde{\mathcal{M}}$ be two smooth manifolds of dimension n and $\tilde{n}$ with $\left((U_i, \varphi_i)\right)_{i\in I}$ and $\left(\left(\tilde{U}_i, \tilde{\varphi}_i\right)\right)_{i\in \tilde{I}}$ their respective atlases.*

- *A continuous scalar function $f : \mathcal{M} \longrightarrow \mathbb{R}$ is said to be smooth if*

$$f \circ \varphi_i^{-1} : \varphi_i(U_i) \longrightarrow \mathbb{R}$$

is smooth for all $i \in I$. The set of such functions is denoted $C^\infty(\mathcal{M})$.

- *A continuous function $f : \mathcal{M} \longrightarrow \tilde{\mathcal{M}}$ is said to be smooth if*

$$\tilde{\varphi}_j \circ f \circ \varphi_i^{-1} : \varphi_i\left(U_i \cap f^{-1}\left(\tilde{U}_j\right)\right) \longrightarrow \tilde{\varphi}_j\left(\tilde{U}_j\right)$$

is smooth for all $(i, j) \in I \times \tilde{I}$.

According to Definition 1.3, checking the smoothness of a function $f : \mathcal{M} \longrightarrow \tilde{\mathcal{M}}$ *a priori* requires to check the smoothness of $\tilde{\varphi}_j \circ f \circ \varphi_i^{-1}$ for all possible choice of i and j. However, thanks to the smoothness of the transition maps for both atlases, f is smooth if and only if for all $p \in \mathcal{M}$ there exists local charts which are neighborhoods of p and $f(p)$ and such that $\tilde{\varphi}_j \circ f \circ \varphi_i^{-1}$ is indeed smooth. Remark that the coordinate functions x^j defined in (1.1) and associated to a local chart (U, φ) are smooth scalar functions on U.

1.2 Vector fields and 1-forms

1.2.1 Tangent vectors and vector fields

Now that manifolds are defined, we can define interesting objects on these structures. The first objects we need are tangent vectors and vector fields. We assume given $\mathcal{M}$ a smooth manifold of dimension $n \geq 1$.

Definition 1.4. *Let $p \in \mathcal{M}$.*

- *A tangent vector X_p at the point p is a map $X_p : C^\infty(\mathcal{M}) \longrightarrow \mathbb{R}$ which is $\mathbb{R}$-linear and satisfies the Leibniz rule*

$$X_p(fg) = f(p)X_p(g) + g(p)X_p(f)$$

for all $f, g \in C^\infty(\mathcal{M})$.

- *We denote $T_p\mathcal{M}$ the set of tangent vectors at the point p and call it the tangent space to $\mathcal{M}$ at p.*

It is obvious that $T_p\mathcal{M}$ has the structure of a vector space over $\mathbb{R}$. In the next lemma, we construct useful bump functions (recall that the support of a function f is the closure of $\{p \in \mathcal{M} \mid f(p) \neq 0\}$).

Lemma 1.1. *Let $p \in \mathcal{M}$ and U a neighborhood of p. There exists $\chi \in C^\infty(\mathcal{M})$ such that $0 \leq \chi \leq 1$, $\chi = 1$ on some neighborhood of p and* $\operatorname{supp}(\chi) \subset U$.

Proof. See Exercise 3.2. □

Even though tangent vectors act on $C^\infty(\mathcal{M})$ (i.e functions defined on the whole manifold), they are local objects, as the next lemma shows.

Lemma 1.2. *Let $p \in \mathcal{M}$ and $X_p \in T_p\mathcal{M}$.*

(i) If $f, g \in C^\infty(\mathcal{M})$ are equal on a neighborhood of p, then $X_p(f) = X_p(g)$.

(ii) If $f \in C^\infty(\mathcal{M})$ is constant on a neighborhood of p, then $X_p(f) = 0$.

Proof. For the first point of the lemma, set $h = f - g$, which vanishes on a neighborhood of p. Let χ be a bump function around p adapted to this neighborhood (see Lemma 1.1), we have $\chi h = 0$ on $\mathcal{M}$. Since by linearity we have $X_p(0) = 2X_p(0)$ and thus $X_p(0) = 0$, we obtain $X_p(\chi h) = 0$. However, the Leibniz rule gives $0 = \chi(p)X_p(h) + h(p)X_p(\chi)$. Since $\chi(p) = 1$ and $h(p) = 0$ we obtain $X_p(h) = 0$, and thus $X_p(f) = X_p(g)$ by linearity. For the second point of the lemma, the first point allows us to assume that f is equal to some constant $C \in \mathbb{R}$ on the whole manifold. By linearity we have $X_p(f) = CX_p(1)$ and by the Leibniz rule $X_p(1) = 2X_p(1)$ so that $X_p(1) = 0$. □

This lemma shows that tangent vectors, acting as derivations at a point, are local objects. In particular, it allows us to define $X_p(f)$ for f only defined on a neighborhood p, such as the coordinate functions x^j associated to a local chart. In the next definition, we define the most important tangent vectors.

Definition 1.5. *If (U, φ) is such a local chart with associated coordinate functions $(x^i)_{i=1,\dots,n}$ and $p \in U$, we define the map $\partial_{x^i|_p} : C^\infty(\mathcal{M}) \longrightarrow \mathbb{R}$ by*

$$\partial_{x^i|_p}(f) := \partial_i \left(f \circ \varphi^{-1}\right)(\varphi(p)), \tag{1.2}$$

where the ∂_i on the RHS denotes the usual partial derivatives with respect to the i-th coordinate in $\mathbb{R}^n$.

Using the Leibniz rule for the usual partial derivatives in $\mathbb{R}^n$, one can check that $\partial_{x^i|_p}$ belongs to $T_p\mathcal{M}$, but we can say much more than this.

Proposition 1.1. *If (U, φ) is a local chart, then $\left(\partial_{x^i|_p}\right)_{i=1,\dots,n}$ is a basis of $T_p\mathcal{M}$ for all $p \in U$. Moreover, we have*

$$X_p = \sum_{i=1}^n X_p(x^i)\partial_{x^i|_p},$$

for all $X_p \in T_p\mathcal{M}$.

Proof. We first prove that $\left(\partial_{x^i|_p}\right)_{i=1,\dots,n}$ is linearly independent. For this we compute

$$\partial_{x^i|_p}(x^j) = \partial_i \left(x^j \circ \varphi^{-1}\right)(\varphi(p)) = \partial_i \pi^j(\varphi(p)) = \delta_i^j.$$

Therefore if there exists some number a_i such that $\sum_{i=1}^n a_i \partial_{x^i|_p} = 0$, evaluating this sum at x^j gives $a_j = 0$. This proves the linear independence of $\left(\partial_{x^i|_p}\right)_{i=1,\dots,n}$. Moreover, if $X_p \in T_p\mathcal{M}$ then define

$$D = X_p - \sum_{i=1}^n X_p(x^i)\partial_{x^i|_p}.$$

Since $T_p\mathcal{M}$ is a vector space we have $D \in T_p\mathcal{M}$ and by definition $D(x^i) = 0$ for all $i = 1, \dots, n$. Now, let $h \in C^\infty(\mathcal{M})$. Applying Taylor's formula in $\mathbb{R}^n$ to $h \circ \varphi^{-1}$ at $\varphi(p)$, one can show the existence of smooth functions $h^{[i]} \in C^\infty(\mathcal{M})$ such that in a neighborhood of p we have

$$h = h(p) + \sum_{i=1}^n \left(x^i - x^i(p)\right) h^{[i]}.$$

Using the axioms of Definition 1.4 we obtain

$$D(h) = D(h(p)) + \sum_{i=1}^n \left(\left(x^i(p) - x^i(p)\right) D\left(h^{[i]}\right) + h^{[i]}(p) D\left(x^i - x^i(p)\right)\right) = 0,$$

where we have used $D(x^i) = 0$ and the second part of Lemma 1.2. This shows that $D = 0$ and thus that $\left(\partial_{x^i|_p}\right)_{i=1,\dots,n}$ spans $T_p\mathcal{M}$. □

Definition 1.6. *The tangent bundle of $\mathcal{M}$ is defined to be*

$$T\mathcal{M} = \bigsqcup_{p\in\mathcal{M}} T_p\mathcal{M}.$$

We define $\pi : T\mathcal{M} \longrightarrow \mathcal{M}$ the natural map associating p to any element of $T_p\mathcal{M}$.

The next proposition shows that the tangent bundle is much more than simply the collection of all tangent spaces.

Proposition 1.2. *The tangent bundle $T\mathcal{M}$ is a smooth manifold of dimension $2n$ and the projection π is smooth.*

Proof. Given a countable atlas $\big((U_i,\varphi_i)\big)_{i\in I}$ on $\mathcal{M}$, we define an atlas on $T\mathcal{M}$ by first considering $\tilde{U}_i = \pi^{-1}(U_i)$. Thanks to Proposition 1.1, for all $X \in \tilde{U}_i$ there exist a unique $(\beta^k)_{k=1,\dots,n} \in \mathbb{R}^n$ and a unique $p \in U_i$ such that $X = \sum_{k=1}^n \beta^k \partial_{x^k|_p}$ (where the $(x^k)_{k=1,\dots,n}$ are associated to φ_i), and we can define

$$\tilde{\varphi}_i(X) = \left(\varphi_i(p), \beta^1, \dots, \beta^n\right).$$

We first note that $\tilde{\varphi}_i\left(\tilde{U}_i\right) = \varphi_i(U_i) \times \mathbb{R}^n$ is an open subset of $\mathbb{R}^{2n}$, and that $\tilde{\varphi}_i$ is a bijection. Moreover, we can consider the sets $\tilde{\varphi}_i^{-1}(V)$ for V any open subset of $\mathbb{R}^{2n}$ as a basis for a topology of $T\mathcal{M}$, and one can check that $T\mathcal{M}$ is Hausdorff for this choice. It remains to prove that the transition maps associated to the atlas $\left(\left(\tilde{U}_i,\tilde{\varphi}_i\right)\right)_{i\in\tilde{I}}$ as well as π are smooth. If $\left(\tilde{U}_i,\tilde{\varphi}_i\right)$ and $\left(\tilde{U}_j,\tilde{\varphi}_j\right)$ are two local charts then

$$\tilde{\varphi}_i\left(\tilde{U}_i \cap \tilde{U}_j\right) = \varphi_i(U_i \cap U_j) \times \mathbb{R}^n, \quad \tilde{\varphi}_j\left(\tilde{U}_i \cap \tilde{U}_j\right) = \varphi_j(U_i \cap U_j) \times \mathbb{R}^n,$$

and the transition map $\tilde{\varphi}_j \circ \tilde{\varphi}_i^{-1} : \varphi_i(U_i \cap U_j) \times \mathbb{R}^n \longrightarrow \varphi_j(U_i \cap U_j) \times \mathbb{R}^n$ satisfies

$$\tilde{\varphi}_j \circ \tilde{\varphi}_i^{-1}\left(z, \beta^1, \dots, \beta^n\right) = \tilde{\varphi}_j\left(\sum_{k=1}^n \beta^k \partial_{x^k|_{\varphi_i^{-1}(z)}}\right)$$

where the $(x^k)_{k=1,\dots,n}$ are associated to φ_i. If we denote by $(y^1,\dots,y^n)$ the coordinates functions associated to φ_j, we need to relate ∂_{x^k} and ∂_{y^ℓ}, which is done in Exercise 1.6:

$$\partial_{x^k|_{\varphi_i^{-1}(z)}} = \sum_{\ell=1}^n \partial_k\left(x^\ell \circ \varphi_i^{-1}\right)(z)\partial_{y^\ell|_{\varphi_i^{-1}(z)}}.$$

Therefore we have

$$\begin{aligned}
\tilde{\varphi}_j \circ \tilde{\varphi}_i^{-1}\left(z, \beta^1, \dots, \beta^n\right) &= \tilde{\varphi}_j\left(\sum_{k,\ell=1}^n \beta^k \partial_k\left(x^\ell \circ \varphi_i^{-1}\right)(z)\partial_{y^\ell|_{\varphi_i^{-1}(z)}}\right) \\
&= \left(\varphi_j(\varphi_i^{-1}(z)), \sum_k^n \beta^k \partial_k\left(x^1 \circ \varphi_i^{-1}\right)(z), \dots, \sum_k^n \beta^k \partial_k\left(x^n \circ \varphi_i^{-1}\right)(z)\right),
\end{aligned}$$

which is clearly smooth. Moreover, we have $\varphi_i \circ \pi \circ \tilde{\varphi}_j^{-1}\left(z, \beta^1, \dots, \beta^n\right) = \varphi_i \circ \varphi_j^{-1}(z)$ which is smooth. This concludes the proof. □

Definition 1.7. *A vector field is a smooth map $X : \mathcal{M} \longrightarrow T\mathcal{M}$ such that $\pi \circ X = \mathrm{Id}_{\mathcal{M}}$. We denote by $\Gamma(\mathcal{M})$ the set of vector fields on $\mathcal{M}$.*

Concretely, a vector field is a collection of arrows at each point of the manifold and which depend smoothly on the point (see Exercise 3.3 for a more geometric definition of tangent vectors). If $X \in \Gamma(\mathcal{M})$, then we denote $X(p)$ by X_p, and the condition $\pi \circ X = \mathrm{Id}_{\mathcal{M}}$ rewrites $X_p \in T_p\mathcal{M}$ for all $p \in \mathcal{M}$. Proposition 1.1 shows in particular that in a local chart (U, φ) we have

$$X = \sum_{i=1}^{n} X^i \partial_{x^i},$$

where x^i are the coordinate functions associated to φ and X^i are the smooth functions defined by $X^i(p) = X_p(x^i)$ and are called the components of X in the coordinate system $(x^i)_{i=1,\dots,n}$. If $f \in C^\infty(\mathcal{M})$ and $X \in \Gamma(\mathcal{M})$, we define $X(f)$ to be the function on $\mathcal{M}$ defined by $X(f)(p) = X_p(f)$ and we have $X(f) \in C^\infty(\mathcal{M})$. Moreover, the Leibniz rule holds

$$X(fg) = fX(g) + gX(f). \tag{1.3}$$

Therefore, there is a natural identification between $\Gamma(\mathcal{M})$ and the set of derivations of $C^\infty(\mathcal{M})$, i.e the $\mathbb{R}$-linear maps from $C^\infty(\mathcal{M})$ to itself satisfying the Leibniz rule. This identification allows to define the Lie bracket $[X, Y]$ of two vector fields by its action on smooth function

$$[X, Y](f) = X(Y(f)) - Y(X(f)).$$

This indeed defines a vector field since it can be shown to satisfy the Leibniz rule (1.3). See Exercise 1.4 for important properties of the Lie bracket.

1.2.2 1-forms

Recall that each tangent space $T_p\mathcal{M}$ is vector space of dimension n.

Definition 1.8. *We define the cotangent space $T_p^*\mathcal{M}$ to be the dual of $T_p\mathcal{M}$. We define the cotangent bundle $T^*\mathcal{M}$ to be*

$$T^*\mathcal{M} = \bigsqcup_{p\in\mathcal{M}} T_p^*\mathcal{M}.$$

Note that the usual property of the dual of a vector space implies that $T_p^*\mathcal{M}$ is a vector space of dimension n. Moreover, if x^i are coordinate functions associated to a local chart (U, φ) then

$$\mathrm{d}x^i_{|_p}(X_p) = X_p(x^i)$$

defines an element $\mathrm{d}x^i_{|_p}$ of $T_p^*\mathcal{M}$. The first computation in the proof of Proposition 1.1 rewrites

$$\mathrm{d}x^i_{|_p}\left(\partial_{x^j|_p}\right) = \partial_{x^j|_p}(x^i) = \delta^i_j.$$

This shows that $\left(\mathrm{d}x^i_{|_p}\right)_{i=1,\dots,n}$ is a basis of $T_p^*\mathcal{M}$ which is dual to the basis $\left(\partial_{x^i|_p}\right)_{i=1,\dots,n}$ of $T_p\mathcal{M}$ for all $p \in U$. By mimicking the proof of Proposition 1.2, we can give $T^*\mathcal{M}$ a natural structure of manifold of dimension $2n$ such that the natural projection $\pi^* : T^*\mathcal{M} \longrightarrow \mathcal{M}$ is smooth.

Definition 1.9. *A 1-form is a smooth map $\omega : \mathcal{M} \longrightarrow T^*\mathcal{M}$ such that $\pi^* \circ \omega = \mathrm{Id}_{\mathcal{M}}$. We denote by $\Lambda^1(\mathcal{M})$ the set of 1-forms on $\mathcal{M}$.*

As for vector fields, if $\omega \in \Lambda^1(\mathcal{M})$ we denote $\omega(p)$ by ω_p and the condition $\pi^* \circ \omega = \mathrm{Id}_{\mathcal{M}}$ becomes $\omega_p \in T_p^*\mathcal{M}$. The main purpose of 1-forms is to be dual objects to vector fields, in the sense that if $\omega \in \Lambda^1(\mathcal{M})$ and $X \in \Gamma(\mathcal{M})$ then $\omega(X)(p) = \omega_p(X_p)$ defines an element of $C^\infty(\mathcal{M})$. By duality, we can also think of vector fields as acting on 1-forms by the same formula, i.e $X(\omega) = \omega(X)$. As for vector fields, a 1-form ω can be locally expressed as

$$\omega = \sum_{i=1}^{n} \omega_i \mathrm{d}x^i,$$

where the $\omega_i = \omega\left(\partial_{x^i}\right)$ are the components of ω in the coordinate system $(x^i)_{i=1,\dots,n}$. The most important example of a 1-form is the differential of a scalar function.

Definition 1.10. *If $f \in C^\infty(\mathcal{M})$, then its differential is the 1-form locally defined by*

$$\mathrm{d}f = \sum_{i=1}^n \partial_{x^i}(f)\mathrm{d}x^i.$$

Since $\mathrm{d}x^i(X) = X^i$, one can check that the action of the differential on a vector field is given by $\mathrm{d}f(X) = X(f)$ (this could be a definition of $\mathrm{d}f$).

Remark 1.1. *As their names suggest, the tangent bundle and cotangent bundle are examples of vector bundles, which roughly speaking are a way to associate smoothly at each point of a manifold an element of a vector space. In the case of the tangent and cotangent bundle, we associate at each point tangent vectors and covectors.*

1.3 Tensors

The most important objects in mathematical general relativity are tensors, which generalize at the same time scalar functions, vector fields and 1-forms.

1.3.1 First definitions

There are several ways to define them, we choose the most direct one.

Definition 1.11. *A tensor field of type (r, s) on a smooth manifold $\mathcal{M}$ is a map*

$$T : \left(\Lambda^1(\mathcal{M})\right)^r \times (\Gamma(\mathcal{M}))^s \longrightarrow C^\infty(\mathcal{M})$$

which is $C^\infty(\mathcal{M})$-multilinear. We denote by $\mathcal{T}^r_s(\mathcal{M})$ the set of tensor fields of type (r, s).

We have already encountered a natural example of $(1, 1)$-tensor, since $T(\omega, X) = \omega(X)$ can be shown to be $C^\infty(\mathcal{M})$-multilinear. However, the map $T(\omega, X, Y) = \omega([X, Y])$ does not define a $(1, 2)$-tensor since $T(\omega, fX, Y) \neq fT(\omega, X, Y)$..

Let us now see how tensors generalize scalar functions, vector fields and 1-forms. By convention, a $(0, 0)$-tensor is the same thing as a scalar function on the manifold. Moreover, since 1-forms act on vector fields to produce scalar functions, we can identify $(0, 1)$-tensors with 1-forms. Similarly, we can identify $(1, 0)$-tensors with vector fields since vector fields acts by duality on 1-forms. We have thus justified the following identifications:

$$\mathcal{T}^0_0(\mathcal{M}) \simeq C^\infty(\mathcal{M}), \quad \mathcal{T}^0_1(\mathcal{M}) \simeq \Lambda^1(\mathcal{M}), \quad \text{and} \quad \mathcal{T}^1_0(\mathcal{M}) \simeq \Gamma(\mathcal{M}).$$

Tensors can be multiplied in a particular sense.

Definition 1.12. *Let $T_i \in \mathcal{T}^{r_i}_{s_i}(\mathcal{M})$ for $i = 1, 2$. We define the tensor product $T_1 \otimes T_2 \in \mathcal{T}^{r_1+r_2}_{s_1+s_2}(\mathcal{M})$ by*

$$\begin{aligned} &T_1 \otimes T_2 \left(\omega_1, \ldots, \omega_{r_1+r_2}, X_1, \ldots, X_{s_1+s_2}\right) \\ &\quad = T_1 \left(\omega_1, \ldots, \omega_{r_1}, X_1, \ldots, X_{s_1}\right) \times T_2 \left(\omega_{r_1+1}, \ldots, \omega_{r_1+r_2}, X_{s_1+1}, \ldots, X_{s_1+s_2}\right). \end{aligned}$$

Note that in general we don't have $T_1 \otimes T_2 = T_2 \otimes T_1$, except when one of the tensors is a $(0, 0)$-tensor, i.e a smooth function. As vector fields or 1-forms, tensors of any type are local objects, as the next lemma shows.

Lemma 1.3. *Let $p \in \mathcal{M}$, $T \in \mathcal{T}^r_s(\mathcal{M})$. Consider 1-forms $\omega_1, \ldots, \omega_r$ and $\bar{\omega}_1, \ldots, \bar{\omega}_r$ such that $\omega_i(p) = \bar{\omega}_i(p)$ and vector fields $X_1, \ldots, X_s$ and $\bar{X}_1, \ldots, \bar{X}_s$ such that $X_i(p) = \bar{X}_i(p)$. We have*

$$T(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s)(p) = T\left(\bar{\omega}_1, \ldots, \bar{\omega}_r, \bar{X}_1, \ldots, \bar{X}_s\right)(p).$$

Proof. By multilinearity it is enough to show that $T(\omega_1,\dots,\omega_r,X_1,\dots,X_s)(p)=0$ when one of the ω_i or one of the X_i vanishes at p. Say $(\omega_1)_p=0$ and consider (U,φ) a local chart around p with coordinate functions x^i, and denote by $(\omega_1)_i$ the components of ω_1 in this coordinate system. If χ is any bump function associated to U (see Lemma 1.1), then $\chi(\omega_1)_i\in C^\infty(\mathcal{M})$ and $\chi \mathrm{d}x^i\in\Lambda^1(\mathcal{M})$ and

$$\chi^2\omega_1=\sum_{i=1}^n \chi(\omega_1)_i\chi \mathrm{d}x^i$$

holds globally on $\mathcal{M}$. Therefore, $C^\infty(\mathcal{M})$-multilinearity implies

$$\chi^2 T\left(\omega_1,\dots,\omega_r,X_1,\dots,X_s\right)=\sum_{i=1}^n \chi(\omega_1)_i T\left(\chi \mathrm{d}x^i,\dots,\omega_r,X_1,\dots,X_s\right).$$

We evaluate this equality between smooth functions at p and use $\chi(p)=1$ and $(\omega_1)_i(p)=0$ (which follows from our assumption $(\omega_1)_p=0$) to conclude the proof. □

This lemma allows us to consider tensors as fields over the manifold, i.e as assigning to each point a multilinear map $T_p:\left(T_p^*\mathcal{M}\right)^r\times\left(T_p\mathcal{M}\right)^s\longrightarrow\mathbb{R}$. This also allows us to consider the local expression of a tensor, i.e the expression of T_p for all $p\in U$ where (U,φ) is a local chart. For that, we first use the identifications between vector fields and 1-forms and $(1,0)$-tensors and $(0,1)$-tensors to define the tensor product of coordinates vector fields and coordinates 1-forms associated to (U,φ) as below:

$$\begin{aligned}\partial_{x^{i_1}}\otimes\cdots\otimes\partial_{x^{i_r}}\otimes \mathrm{d}x^{j_1}\otimes\cdots\otimes \mathrm{d}x^{j_s}&\left(\omega_1,\dots,\omega_r,X_1,\dots,X_s\right)\\&=\partial_{x^{i_1}}(\omega_1)\times\cdots\times\partial_{x^{i_r}}(\omega_r)\times \mathrm{d}x^{j_1}(X_1)\times\cdots\times \mathrm{d}x^{j_s}(X_s)\\&=(\omega_1)_{i_1}\times\cdots\times(\omega_r)_{i_r}\times(X_1)^{j_1}\times\cdot\times(X_s)^{j_s},\end{aligned}$$

where $(\omega_k)_{i_k}$ is the i_k-th component of ω_k and $(X_\ell)^{j_\ell}$ is the j_ℓ-th component of X_ℓ. Using $C^\infty(\mathcal{M})$-multilinearity we obtain

$$\begin{aligned}T\left(\omega_1,\dots,\omega_r,X_1,\dots,X_s\right)&=T\left((\omega_1)_{i_1}\mathrm{d}x^{i_1},\dots,(\omega_r)_{i_r}\mathrm{d}x^{i_r},(X_1)^{j_1}\partial_{x^{j_1}},\dots,(X_s)^{j_s}\partial_{x^{j_s}}\right)\\&=T\left(\mathrm{d}x^{i_1},\dots,\mathrm{d}x^{i_r},\partial_{x^{j_1}},\dots,\partial_{x^{j_s}}\right)(\omega_1)_{i_1}\times\cdots\times(\omega_r)_{i_r}\times(X_1)^{j_1}\times\cdots\times(X_s)^{j_s},\end{aligned}$$

where we used the famous Einstein convention for summation, i.e we sum over repeated indexes when one is up and one is down. Defining the components of T in the coordinate system $(x^i)_{i=1,\dots,n}$ by

$$T^{i_1\cdots i_r}_{j_1\cdots j_s}:=T\left(\mathrm{d}x^{i_1},\dots,\mathrm{d}x^{i_r},\partial_{x^{j_1}},\dots,\partial_{x^{j_s}}\right)$$

we have obtained the local expression

$$T=T^{i_1\cdots i_r}_{j_1\cdots j_s}\partial_{x^{i_1}}\otimes\cdots\otimes\partial_{x^{i_r}}\otimes \mathrm{d}x^{j_1}\otimes\cdots\otimes \mathrm{d}x^{j_s}.$$

As an application, let us compute the local components of $T_1\otimes T_2$ for $T_i\in\mathcal{T}^{r_i}_{s_i}(\mathcal{M})$. We have

$$\begin{aligned}(T_1\otimes T_2)^{i_1\cdots i_{r_1+r_2}}_{j_1\cdots j_{s_1+s_2}}&=T_1\otimes T_2\left(\mathrm{d}x^{i_1},\dots,\mathrm{d}x^{i_{r_1+r_2}},\partial_{x^{j_1}},\dots,\partial_{x^{j_{s_1+s_2}}}\right)\\&=T_1\left(\mathrm{d}x^{i_1},\dots,\mathrm{d}x^{i_{r_1}},\partial_{x^{j_1}},\dots,\partial_{x^{j_{s_1}}}\right)\times T_2\left(\mathrm{d}x^{i_{r_1+1}},\dots,\mathrm{d}x^{i_{r_1+r_2}},\partial_{x^{j_{s_1+1}}},\dots,\partial_{x^{j_{s_1+s_2}}}\right)\\&=(T_1)^{i_1\cdots i_{r_1}}_{j_1\cdots j_{s_1}}(T_2)^{i_{r_1+1}\cdots i_{r_1+r_2}}_{j_{s_1+1}\cdots j_{s_1+s_2}}.\end{aligned}$$

1.3.2 Contracting tensors

Contracting a tensor is a way to simplify it, i.e to go from a (r,s)-tensor to a $(r-1,s-1)$-tensor (providing $r,s\geq 1$). It can be interpreted as a trace, see Exercise 1.7 below. Contractions of tensors of arbitrary type are built on the contraction of $(1,1)$-tensors.

Lemma 1.4. *There exists a unique $C^\infty(\mathcal{M})$-linear map $C^1_1:\mathcal{T}^1_1(\mathcal{M})\longrightarrow C^\infty(\mathcal{M})$ such that $C^1_1(X\otimes\omega)=\omega(X)$ for all $X\in\Gamma(\mathcal{M})$ and $\omega\in\Lambda^1(\mathcal{M})$.*

Proof. Let C_1^1 be such a map. Let $A \in \mathcal{T}_1^1(\mathcal{M})$ and (U, φ) a local chart with associated coordinate functions $(x^i)_{i=1,\ldots,n}$. Since C_1^1 is $C^\infty(\mathcal{M})$-linear we can use the local expression of A to compute the function $C_1^1(A)$ on U. Since this local expression is $A = A^i_j \partial_{x^i} \otimes \mathrm{d}x^j$ and again because of the $C^\infty(\mathcal{M})$-linearity of C_1^1 we thus have on U:

$$C_1^1(A) = A^i_j C_1^1\left(\partial_{x^i} \otimes \mathrm{d}x^j\right) = A^i_j \mathrm{d}x^j(\partial_{x^i}) = A^i_j \delta^j_i = A^i_i,$$

where we have used $C_1^1(X \otimes \omega) = \omega(X)$. This shows the uniqueness of C_1^1, if it exists. For the existence, the above computation actually dictates what $C_1^1(A)$ should be, but we need to show that if two coordinate systems $(x^i)_{i=1,\ldots,n}$ and $(y^i)_{i=1,\ldots,n}$ overlap, then the two definitions of $C_1^1(A)$ match. For this we use the transformation rules for the coordinate vector fields and 1-forms from Exercise 1.6:

$$\begin{aligned} A(\mathrm{d}x^i, \partial_{x^i}) &= A\left(\left(\partial_k\left(x^i \circ \psi^{-1}\right) \circ \psi\right) \mathrm{d}y^k, \left(\partial_i\left(y^\ell \circ \varphi^{-1}\right) \circ \varphi\right) \partial_{y^\ell}\right) \\ &= \left(\partial_k\left(x^i \circ \psi^{-1}\right) \circ \psi\right)\left(\partial_i\left(y^\ell \circ \varphi^{-1}\right) \circ \varphi\right) A\left(\mathrm{d}y^k, \partial_{y^\ell}\right), \end{aligned}$$

where $(x^i)_{i=1,\ldots,n}$ and $(y^i)_{i=1,\ldots,n}$ are associated to φ and ψ respectively. Now if we define $f = \varphi \circ \psi^{-1}$ and $z = \psi(p)$ we have

$$\left(\partial_k\left(x^i \circ \psi^{-1}\right) \circ \psi\right)\left(\partial_i\left(y^\ell \circ \varphi^{-1}\right) \circ \varphi\right)(p) = \partial_k\left(\pi^i \circ f\right)(z) \partial_i\left(\pi^\ell \circ f^{-1}\right)(f(z)),$$

which is the (ℓ, k) coefficient of the Jacobian matrix of $f^{-1} \circ f$ at z, that is δ^ℓ_k. Therefore we have $A(\mathrm{d}x^i, \partial_{x^i}) = \delta^\ell_k A\left(\mathrm{d}y^k, \partial_{y^\ell}\right) = A\left(\mathrm{d}y^i, \partial_{y^i}\right)$. Therefore $C_1^1(A)$ is a well-defined function on $\mathcal{M}$ and its smoothness can be read on its local expression. □

An outcome of the proof of the previous lemma is the effect of the $(1,1)$ contraction C_1^1 in local coordinates. If $A \in \mathcal{T}_1^1(\mathcal{M})$ has local components $A^i_j = A\left(\mathrm{d}x^i, \partial_{x^j}\right)$ then $C_1^1(A)$ is the smooth function locally given by A^i_i (with Einstein's summation convention).

The extension to a tensor of arbitrary type is straightforward. Let $A \in \mathcal{T}_s^r(\mathcal{M})$ with $r, s \geq 1$ and choose some $1 \leq a \leq r$ and $1 \leq b \leq s$. For $\omega_1, \ldots, \omega_{r-1}$ some fixed 1-forms and $X_1, \ldots X_{s-1}$ some fixed vector fields, define $\tilde{A}^a_b[\omega_1, \ldots, \omega_{r-1}, X_1, \ldots X_{s-1}]$ by

$$\tilde{A}^a_b[\omega_1, \ldots, \omega_{r-1}, X_1, \ldots X_{s-1}](\omega, X) = A(\omega_1, \ldots, \omega_{a-1}, \omega, \omega_a, \ldots, \omega_r, X_1, \ldots, X_{b-1}, X, X_b, \ldots, X_s).$$

Since $\tilde{A}^a_b[\omega_1, \ldots, \omega_{r-1}, X_1, \ldots X_{s-1}]$ is $C^\infty(\mathcal{M})$-multilinear, it defines a $(1,1)$-tensor and we define

$$C^a_b(A)(\omega_1, \ldots, \omega_{r-1}, X_1, \ldots X_{s-1}) = C_1^1\left(\tilde{A}^a_b[\omega_1, \ldots, \omega_{r-1}, X_1, \ldots X_{s-1}]\right).$$

Since $C^a_b(A)$ is $C^\infty(\mathcal{M})$-multilinear, we have defined a $C^\infty(\mathcal{M})$-linear map $C^a_b : \mathcal{T}_s^r(\mathcal{M}) \longrightarrow \mathcal{T}_{s-1}^{r-1}(\mathcal{M})$, which is the contraction over the indices a and b. As for the $(1,1)$ contraction, we can see the effect of the (a,b) contraction in local components:

$$\begin{aligned} (C^a_b(A))^{i_1 \cdots i_{r-1}}_{j_1 \cdots j_{s-1}} &= C^a_b(A)\left(\mathrm{d}x^{i_1}, \ldots, \mathrm{d}x^{i_{r-1}}, \partial_{x^{j_1}}, \ldots, \partial_{x^{j_{s-1}}}\right) \\ &= C_1^1\left(\tilde{A}^a_b\left[\mathrm{d}x^{i_1}, \ldots, \mathrm{d}x^{i_{r-1}}, \partial_{x^{j_1}}, \ldots, \partial_{x^{j_{s-1}}}\right]\right) \\ &= \tilde{A}^a_b\left[\mathrm{d}x^{i_1}, \ldots, \mathrm{d}x^{i_{r-1}}, \partial_{x^{j_1}}, \ldots, \partial_{x^{j_{s-1}}}\right](\mathrm{d}x^k, \partial_{x^k}) \\ &= A\left(\mathrm{d}x^{i_1}, \ldots, \mathrm{d}x^{i_{a-1}}, \mathrm{d}x^k, \mathrm{d}x^{i_a}, \ldots, \mathrm{d}x^{i_{r-1}}, \partial_{x^{j_1}}, \ldots, \partial_{x^{j_{b-1}}}, \partial_{x^k}, \partial_{x^{j_a}}, \ldots, \partial_{x^{j_{s-1}}}\right) \\ &= A^{i_1 \cdots i_{a-1} k i_a \cdots i_{r-1}}_{j_1 \cdots j_{b-1} k j_b \cdots j_{s-1}}. \end{aligned}$$

We thus see that the (a,b) contraction translates in coordinates to a sum over the a-th up index and the b-th low index. The most important property of contractions is that they don't depend on the coordinate system, meaning that if $A \in \mathcal{T}_1^1(\mathcal{M})$ and if $(x^i)_{i=1,\ldots,n}$ and $(y^i)_{i=1,\ldots,n}$ are too overlapping coordinate systems then $A(\mathrm{d}x^k, \partial_{x^k}) = A(\mathrm{d}y^k, \partial_{x^k})$. Note that this would be false if we were to sum over two vector fields or two 1-forms: we need one index up and one index down to contract over them.

1.3.3 Derivation of tensors

For now, we know how to differentiate smooth functions with a vector field. It is in fact possible to differentiate any kind of tensors, with the notion of tensor derivation.

Definition 1.13. *A tensor derivation is a $\mathbb{R}$-linear map*

$$\begin{aligned} \mathcal{D} : \bigsqcup_{r,s\geq 0} \mathcal{T}^r_s(\mathcal{M}) &\longrightarrow \bigsqcup_{r,s\geq 0} \mathcal{T}^r_s(\mathcal{M}) \\ T &\longmapsto \mathcal{D}T \end{aligned}$$

such that

- *it preserves the type, i.e $\mathcal{D}\left(\mathcal{T}^r_s(\mathcal{M})\right) \subset \mathcal{T}^r_s(\mathcal{M})$,*
- *it satisfies the Leibniz rule $\mathcal{D}(A \otimes B) = \mathcal{D}A \otimes B + A \otimes \mathcal{D}B$,*
- *it commutes with any contraction, i.e $\mathcal{D}(C(A)) = C(\mathcal{D}A)$.*

Since a tensor derivation $\mathcal{D}$ preserves the type and since the tensor product for two $(0,0)$-tensors (i.e smooth functions) is just the multiplication of smooth functions, the above Leibniz rule implies the standard Leibniz rule for functions $\mathcal{D}(fg) = \mathcal{D}(f)g + f\mathcal{D}(g)$. Since derivations on functions are vector fields, there must exists $X \in \Gamma(\mathcal{M})$ such that $\mathcal{D}(f) = X(f)$ for all $f \in C^\infty(\mathcal{M})$. The following proposition gives the expression of $\mathcal{D}T$ in terms of T.

Proposition 1.3. *Let $\mathcal{D}$ be a tensor derivation and let $T \in \mathcal{T}^r_s(\mathcal{M})$. For all $\omega_i \in \Lambda^1(\mathcal{M})$ and $X_i \in \Gamma(\mathcal{M})$ we have*

$$\begin{aligned} (\mathcal{D}T)(\omega_1, \dots, \omega_r, X_1, \dots, X_s) = \mathcal{D}\left(T(\omega_1, \dots, \omega_r, X_1, \dots, X_s)\right) \\ - \sum_{k=1}^{r} T\left(\omega_1, \dots, \omega_{k-1}, \mathcal{D}\omega_k, \omega_{k+1}, \dots, \omega_r, X_1, \dots, X_s\right) \\ - \sum_{\ell=1}^{s} T\left(\omega_1, \dots, \omega_r, X_1, \dots, X_{\ell-1}, \mathcal{D}X_\ell, X_{\ell+1}, \dots, X_s\right). \end{aligned}$$

Proof. In local coordinates, we have

$$A(\omega_1, \dots, \omega_r, X_1, \dots, X_s) = A^{i_1 \cdots i_r}_{j_1 \cdots j_s} (\omega_1)_{i_1} (\omega_r)_{i_r} (X_1)^{j_1} \cdots (X_s)^{j_s}$$

so that there exists C a product of $r+s$ contractions such that

$$A(\omega_1, \dots, \omega_r, X_1, \dots, X_s) = C\left(A \otimes \omega_1 \otimes \cdots \otimes \omega_r \otimes X_1 \otimes \cdots \otimes X_s\right).$$

Therefore, using the commutation with contraction (and thus with product of contractions) and the Leibniz rule we get

$$\begin{aligned} &\mathcal{D}\left(A(\omega_1, \dots, \omega_r, X_1, \dots, X_s)\right) \\ &= C\left(\mathcal{D}\left(A \otimes \omega_1 \otimes \cdots \otimes \omega_r \otimes X_1 \otimes \cdots \otimes X_s\right)\right) \\ &= C\left(\mathcal{D}A \otimes \omega_1 \otimes \cdots \otimes \omega_r \otimes X_1 \otimes \cdots \otimes X_s\right) \\ &\quad + \sum_{k=1}^{r} C\left(A \otimes \omega_1 \otimes \cdots \otimes \omega_{k-1} \otimes \mathcal{D}\omega_k \otimes \omega_{k+1} \otimes \cdots \otimes \omega_r \otimes X_1 \otimes \cdots \otimes X_s\right) \\ &\quad + \sum_{\ell=1}^{s} C\left(A \otimes \omega_1 \otimes \cdots \otimes \omega_r \otimes X_1 \otimes \cdots \otimes X_{\ell-1} \otimes \mathcal{D}X_\ell \otimes X_{\ell+1} \otimes \cdots \otimes X_s\right) \\ &= (\mathcal{D}A)(\omega_1, \dots, \omega_r, X_1, \dots, X_s) \\ &\quad + \sum_{k=1}^{r} A\left(\omega_1, \dots, \omega_{k-1}, \mathcal{D}\omega_k, \omega_{k+1}, \dots, \omega_r, X_1, \dots, X_s\right) \\ &\quad + \sum_{\ell=1}^{s} A\left(\omega_1, \dots, \omega_r, X_1, \dots, X_{\ell-1}, \mathcal{D}X_\ell, X_{\ell+1}, \dots, X_s\right), \end{aligned}$$

which concludes the proof, after isolating the first term. $\square$

This proposition shows that in order to know the derivative of any tensor, it is sufficient to know how to differentiate smooth functions, 1-forms and vector fields. Actually, we don't need to know how to differentiate 1-forms, as next lemma shows.

Lemma 1.5. *Let $X \in \Gamma(\mathcal{M})$ and $\mathcal{D}_0^1 : \Gamma(\mathcal{M}) \longrightarrow \Gamma(\mathcal{M})$ a map satisfying*

$$\mathcal{D}_0^1(fY) = X(f)Y + f\mathcal{D}_0^1(Y)$$

for all $f \in C^\infty(\mathcal{M})$ and $Y \in \Gamma(\mathcal{M})$. There exists a unique tensor derivation $\mathcal{D}$ such that

$$\mathcal{D}_{|_{\mathcal{T}_0^0(\mathcal{M})}} = X, \qquad \mathcal{D}_{|_{\mathcal{T}_0^1(\mathcal{M})}} = \mathcal{D}_0^1. \tag{1.4}$$

Proof. If $\mathcal{D}$ exists, then for $\omega \in \Lambda^1(\mathcal{M})$ and $Y \in \Gamma(\mathcal{M})$ we have

$$\begin{aligned} X(\omega(Y)) &= \mathcal{D}\left(C(Y \otimes \omega)\right) \\ &= C\left(\mathcal{D}(Y \otimes \omega)\right) \\ &= C\left(\mathcal{D}Y \otimes \omega + Y \otimes \mathcal{D}\omega\right) \\ &= \omega(\mathcal{D}Y) + \mathcal{D}\omega(Y), \end{aligned}$$

where we used the commutation with contractions and the Leibniz rule. Therefore for all $\omega \in \Lambda^1(\mathcal{M})$, the 1-form $\mathcal{D}\omega$ is given by

$$(\mathcal{D}\omega)(Y) = X(\omega(Y)) - \omega\left(\mathcal{D}_0^1 Y\right),$$

for all $Y \in \Gamma(\mathcal{M})$. Therefore the action on $\mathcal{T}_1^0(\mathcal{M})$ is uniquely defined, and thus thanks to Proposition 1.3, the action on all tensors is uniquely defined. This proves uniqueness. For the existence, define $\mathcal{D}_0^0 = X$ (viewed as a derivation of smooth functions), $\mathcal{D}_1^0 : \Lambda^1(\mathcal{M}) \longrightarrow \Lambda^1(\mathcal{M})$ by the above formula $(\mathcal{D}_1^0\omega)(Y) = \mathcal{D}_0^0(\omega(Y)) - \omega\left(\mathcal{D}_0^1 Y\right)$. If now $r + s \geq 2$ and $A \in \mathcal{T}_s^r(\mathcal{M})$ then define $\mathcal{D}_s^r A$ by the formula

$$\begin{aligned} \left(\mathcal{D}_s^r A\right)(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s) = {} & \mathcal{D}_0^0\left(A(\omega_1, \ldots, \omega_r, X_1, \ldots, X_s)\right) \\ & - \sum_{k=1}^{r} A\left(\omega_1, \ldots, \omega_{k-1}, \mathcal{D}_1^0\omega_k, \omega_{k+1}, \ldots, \omega_r, X_1, \ldots, X_s\right) \\ & - \sum_{\ell=1}^{s} A\left(\omega_1, \ldots, \omega_r, X_1, \ldots, X_{\ell-1}, \mathcal{D}_0^1 X_\ell, X_{\ell+1}, \ldots, X_s\right). \end{aligned}$$

By using the Leibniz rule for $\mathcal{D}_0^0$ and $\mathcal{D}_0^1$ one can show that $\mathcal{D}_1^0\omega$ is $C^\infty(\mathcal{M})$ and thus $\mathcal{D}_1^0$ that is well-defined. One can also show that $\mathcal{D}_1^0$ satisfies also the Leibniz rule $\mathcal{D}_1^0(f\omega) = f\mathcal{D}_1^0\omega + X(f)\omega$ which implies that $\mathcal{D}_s^r A$ is $C^\infty(\mathcal{M})$-multilinear and thus that $\mathcal{D}_s^r : \mathcal{T}_s^r(\mathcal{M}) \longrightarrow \mathcal{T}_s^r(\mathcal{M})$ is well-defined. We now define a candidate $\mathcal{D} : \bigsqcup_{r,s\geq 0} \mathcal{T}_s^r(\mathcal{M}) \longrightarrow \bigsqcup_{r,s\geq 0} \mathcal{T}_s^r(\mathcal{M})$ by setting $\mathcal{D}_{|_{\mathcal{T}_s^r(\mathcal{M})}} := \mathcal{D}_s^r$. The map $\mathcal{D}$ obviously preserves the type and satisfies (1.4) so it remains to show that $\mathcal{D}$ satisfies the Leibniz rule and commutes with contraction. This is left as an exercise for the reader. □

This lemma shows that a tensor derivation is entirely characterized by a vector field, i.e a way to differentiate smooth functions, and a way to differentiate vector fields (more on this in Exercise 1.9). In these notes, we will encounter two ways of differentiating vector fields (and thus two tensor derivations), see Exercise 1.8 for the first one and Chapter 2 for the second one. We highlight an important formula previously obtained: the derivative of a 1-form is given by

$$\mathcal{D}\omega(X) = \mathcal{D}(\omega(X)) - \omega(\mathcal{D}X), \tag{1.5}$$

where on the RHS, the derivative of a function and the derivative of a vector field appear.

1.4 Exercises

Exercise 1.1. *Consider the sphere*

$$\mathbb{S}^2 = \left\{(x^1, x^2, x^3) \in \mathbb{R}^3 \;\middle|\; (x^1)^2 + (x^2)^2 + (x^3)^2 = 1\right\}.$$

1. *Using the six open subsets of* $\mathbb{S}^2$

$$U_j = \left\{(x^1, x^2, x^3) \in \mathbb{S}^2 \mid x^j > 0\right\}, \qquad V_j = \left\{(x^1, x^2, x^3) \in \mathbb{S}^2 \mid x^j < 0\right\},$$

for $j = 1, 2, 3$ *and the maps from* $\mathbb{S}^2$ *to* $\mathbb{R}^2$ *defined by*

$$\varphi_1(x) = (x^2, x^3), \qquad \varphi_2(x) = (x^1, x^3), \qquad \varphi_3(x) = (x^1, x^2),$$

show that $\mathbb{S}^2$ *is a 2-dimensional smooth manifold.*

2. *Same question with the two open subsets* $U_1 = \mathbb{S}^2 \setminus \{(0,0,1)\}$ *and* $U_2 = \mathbb{S}^2 \setminus \{(0,0,-1)\}$ *and the maps*

$$\varphi_1(x) = \left(\frac{x^1}{1-x^3}, \frac{x^2}{1-x^3}\right), \qquad \varphi_2(x) = \left(\frac{x^1}{1+x^3}, \frac{x^2}{1+x^3}\right).$$

3. *Generalize to* $\mathbb{S}^n$.

Exercise 1.2. *Consider* $f : \mathbb{R} \longrightarrow \mathbb{R}$ *defined by* $f(t) = e^{\frac{1}{t^2-1}}$ *for* $|t| < 1$ *and* $f(t) = 0$ *for* $|t| \geq 1$.

1. *Show that the convolution*

$$h(t) = \frac{1}{\int_{\mathbb{R}} f} \int_{\mathbb{R}} f(t-y) \mathbf{1}_{[-2,2]}(y) \mathrm{d}y$$

defines a smooth function satisfying $0 \leq h \leq 1$, $\mathrm{supp}(h) \subset [-3,3]$ *and* $h_{|[-1,1]} = 1$.

2. *Prove Lemma 1.1.*

Exercise 1.3. *Let* $\mathcal{M}$ *be a smooth manifold,* $p \in \mathcal{M}$ *and define*

$$C_p\mathcal{M} = \{c \ \in C^\infty((-1,1), \mathcal{M}) \mid c(0) = p\}.$$

Given a local chart (U, φ) *around* p*, two curves* $c_1, c_2 \in C_p\mathcal{M}$ *are said to be equivalent if* $(\varphi \circ c_1)'(0) = (\varphi \circ c_2)'(0)$.

1. *Show that this equivalence relation* $\sim$ *is well-defined and does not depend on the local chart.*
2. *Assume that* $\varphi(p) = 0$ *and define a vector space structure on* $\widetilde{T}_p\mathcal{M} := C_p\mathcal{M}/\sim$.
3. *Define an isomorphism between* $\widetilde{T}_p\mathcal{M}$ *and* $T_p\mathcal{M}$.

Exercise 1.4. *Let* $X, Y, Z \in \Gamma(\mathcal{M})$, $a, b \in \mathbb{R}$.

1. *Prove that*

$$[aX + Y, Z] = a[X, Z] + [Y, Z],$$
$$[X, Y] = -[Y, X],$$
$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

The last property above is called the Jacobi identity.

2. *If* $f, g \in C^\infty(\mathcal{M})$*, prove that*

$$[fX, gY] = fg[X, Y] + fX(g)Y - gY(f)X.$$

3. *Compute the local components of* $[X, Y]$.

Exercise 1.5. *Let* $\mathcal{M}$ *be a smooth manifold. For* $\omega \in \Lambda^1(\mathcal{M})$*, we define* $\mathrm{d}\omega$ *by*

$$\mathbf{d}\omega : (X, Y) \in (\Gamma(\mathcal{M}))^2 \longmapsto X(\omega(Y)) - Y(\omega(X)) - \omega([X, Y]).$$

1. *For all* $\omega \in \Lambda^1(\mathcal{M})$*, show that* $\mathbf{d}\omega$ *is a* $(0,2)$*-tensor (called the exterior derivative of* ω*).*

2. *Show that* $\mathbf{d} : (\omega, X, Y) \in \Lambda^1(\mathcal{M}) \times (\Gamma(\mathcal{M}))^2 \longmapsto \mathbf{d}\omega(X,Y)$ *does not define a* $(1,2)$*-tensor.*

3. *Compute* $\mathbf{d}\mathrm{d}f$.

Exercise 1.6. *Let* $\mathcal{M}$ *be a smooth manifold and* (U,φ) *and* (V,ψ) *two local charts such that* $U \cap V \neq \emptyset$, *and denote by* $(x^i)_{i=1,\dots,n}$ *and* $(y^i)_{i=1,\dots,n}$ *their coordinates functions.*

1. *Express* ∂_{x^i} *in terms of the* ∂_{y^j}*'s and* $\mathrm{d}x^i$ *in terms of the* $\mathrm{d}y^j$*'s.*

2. *Let* $T \in \mathcal{T}^r_s(\mathcal{M})$ *be a tensor. Express its components in the coordinate system* $(x^i)_{i=1,\dots,n}$ *with respect to its components in the coordinate system* $(y^i)_{i=1,\dots,n}$.

Exercise 1.7. *Let* $A \in \mathcal{T}^1_1(\mathcal{M})$ *and* $p \in \mathcal{M}$. *Define a linear operator* $\tilde{A}_p : T_p\mathcal{M} \longrightarrow T_p\mathcal{M}$ *and show that* $C^1_1(A)(p) = \mathrm{tr}\tilde{A}_p$.

Exercise 1.8. *Let* $X \in \Gamma(\mathcal{M})$. *We define the following operations on smooth functions and vector fields:*

$$\mathcal{L}_X f := X(f), \qquad \mathcal{L}_X Y := [X,Y].$$

1. *Show that this defines a unique tensor derivation* $\mathcal{L}_X$ *(called the Lie derivative with respect to* X*).*

2. *Show that* $\mathcal{L}_{aX+Y} = a\mathcal{L}_X + \mathcal{L}_Y$ *and* $\mathcal{L}_{[X,Y]} = [\mathcal{L}_X, \mathcal{L}_Y]$.

3. *For* $f \in C^\infty(\mathcal{M})$, *show that* $\mathrm{d}\mathcal{L}_X f = \mathcal{L}_X \mathrm{d}f$ *(where* d *is the differential).*

4. *For* $\omega \in \Lambda^1(\mathcal{M})$, *show that* $\mathbf{d}\mathcal{L}_X\omega = \mathcal{L}_X\mathbf{d}\omega$ *(where* $\mathbf{d} : \mathcal{T}^0_1(\mathcal{M}) \longrightarrow \mathcal{T}^0_2(\mathcal{M})$ *is defined in Exercise 1.5).*

Exercise 1.9. *Let* $\mathfrak{D}(\mathcal{M})$ *be the vector space of tensor derivations on a smooth manifold* $\mathcal{M}$. *For* $B \in \mathcal{T}^1_1(\mathcal{M})$, *we define* $(\mathcal{D}_B)^1_0 : \Gamma(\mathcal{M}) \longrightarrow \Gamma(\mathcal{M})$ *by* $(\mathcal{D}_B)^1_0(X)(\omega) = B(\omega, X)$.

1. *Show that for every* $B \in \mathcal{T}^1_1(\mathcal{M})$ *there exists a unique* $\mathcal{D}_B \in \mathfrak{D}(\mathcal{M})$ *such that* $\mathcal{D}_{B|_{\Gamma(\mathcal{M})}} = (\mathcal{D}_B)^1_0$.

2. *Show that*

$$\mathfrak{D}(\mathcal{M}) = \{\mathcal{L}_X \mid X \in \Gamma(\mathcal{M})\} \oplus \left\{\mathcal{D}_B \;\middle|\; B \in \mathcal{T}^1_1(\mathcal{M})\right\},$$

where the Lie derivative $\mathcal{L}_X$ *is defined in Exercise 1.8.*

Chapter 2

Pseudo-Riemannian Geometry

The real start of general relativity is the definition of a metric tensor, which is by far the most important type of tensor of these lectures since all the other ones (such as the Riemann tensor) will be defined with respect to a given metric tensor.

2.1 The metric tensor

We recall some facts about bilinear algebra. If V is a n-dimensional real vector space, then a bilinear map $f : V \times V \longrightarrow \mathbb{R}$ is said to be symmetric if $f(v,w) = f(w,v)$, non-degenerate if $f(v,w) = 0$ for all w implies $v = 0$. Being non-degenerate is equivalent to the invertibility of any matrix representing f, i.e $(f(v_i, v_j))_{1\leq i,j\leq n}$ for $(v_i)_{i=1,\dots,n}$ some basis of V. If f is symmetric and non-degenerate the Gram-Schmidt algorithm ensures the existence of an orthonormal basis $(e_i)_{i=1,\dots,n}$ of V for f, i.e a basis satisfying $f(e_i, e_j) = 0$ if $i \neq j$ and $f(e_i, e_i) \in \{-1, +1\}$. The number of e_i such that $f(e_i, e_i) = -1$ can be shown to be independent of the orthonormal basis (this is called Sylvester's law). Therefore, we can associate unambiguously to f a sequence $(-, \dots, -, +, \dots, +)$ of length n representing how many basis vectors satisfy $f(e_i, e_i) = -1$ or $f(e_i, e_i) = 1$, this is the signature of f.

Definition 2.1. *Let $\mathcal{M}$ be a smooth manifold.*

- *A pseudo-Riemannian metric tensor $\mathbf{g}$ is a $(0,2)$-tensor such that for all $p \in \mathcal{M}$, $\mathbf{g}_p$ is a symmetric non-degenerate bilinear form on $T_p\mathcal{M}$ with signature independent of p.*
- *A pseudo-Riemannian manifold $(\mathcal{M}, \mathbf{g})$ is a smooth manifold endowed with a pseudo-Riemannian metric.*

Recall that thanks to the locality of tensors (proved in Lemma 1.3), it is meaningful to speak about $\mathbf{g}_p$ as a bilinear form on $T_p\mathcal{M}$. While most of the forthcoming definitions are valid for any signature, only two cases are truly interesting.

Definition 2.2. *Let $(\mathcal{M}, \mathbf{g})$ be a pseudo-Riemannian manifold.*

- *If the signature of $\mathbf{g}$ is $(+, \dots, +)$, then $(\mathcal{M}, \mathbf{g})$ is called a Riemannian manifold.*
- *If the signature of $\mathbf{g}$ is $(-, +, \dots, +)$, then $(\mathcal{M}, \mathbf{g})$ is called a Lorentzian manifold.*

Note that in the Riemannian case, the metric tensor defines a scalar product $\mathbf{g}_p$ on each $T_p\mathcal{M}$, since in addition to being symmetric and non-degenerate, $\mathbf{g}_p$ is also positive, in the sense that $\mathbf{g}_p(X_p, X_p) \geq 0$ for all $X_p \in T_p\mathcal{M}$ and $\mathbf{g}_p(X_p, X_p) = 0$ implies $X_p = 0$. In the Lorentzian case this is not the case anymore, even though we still think of the metric as measuring some kind of physical distance. In particular, $\mathbf{g}_p(X_p, X_p)$ can have an arbitrary sign and even vanish.

Definition 2.3. *Let $(\mathcal{M}, \mathbf{g})$ be a Lorentzian metric, $p \in \mathcal{M}$ and $X_p \in T_p\mathcal{M}$.*

- *If $\mathbf{g}_p(X_p, X_p) > 0$, X_p is a spacelike tangent vector.*
- *If $\mathbf{g}_p(X_p, X_p) < 0$, X_p is a timelike tangent vector.*

- *If* $\mathbf{g}_p(X_p, X_p) = 0$*,* X_p *is a null tangent vector.*

This can be extended to vector fields $X \in \Gamma(\mathcal{M})$*.*

The most important example of Riemannian manifold is simply the Euclidean space $(\mathbb{R}^n, \mathbf{g}_{eucl})$ where $\mathbf{g}_{eucl}$ is the so-called Euclidean metric

$$\mathbf{g}_{eucl} = \mathrm{d}x^1 \otimes \mathrm{d}x^1 + \cdots + \mathrm{d}x^n \otimes \mathrm{d}x^n,$$

where $(x^1, \ldots, x^n)$ are the standard Euclidean coordinates on $\mathbb{R}^n$ (note that this formula defines $\mathbf{g}_{eucl}$ on the whole space since this is a global coordinate system). The most important example of Lorentzian manifold is the Minkowski spacetime $(\mathbb{R}^{1+n}, \mathbf{m})$ where $\mathbf{m}$ is the so-called Minkowski metric

$$\mathbf{m} = -\mathrm{d}x^0 \otimes \mathrm{d}x^0 + \mathrm{d}x^1 \otimes \mathrm{d}x^1 + \cdots + \mathrm{d}x^n \otimes \mathrm{d}x^n,$$

where $(x^0, x^1, \ldots, x^n)$ are the standard Euclidean coordinates on $\mathbb{R}^{1+n}$. The Minkowski spacetime is the spacetime of special relativity, and the whole purpose of studying more general Lorentzian manifolds is to get general relativity. If $X = X^0\partial_0 + X^1\partial_1 + \cdots X^n\partial_n$ is a vector field in Minkowski spacetime then $\mathbf{m}(X, X) = -\left(X^0\right)^2 + \left(X^1\right)^2 + \cdots (X^n)^2$.

Remark 2.1. *From now on, we use greek letters instead of latin ones to denote coordinates on a manifold. This is consistent with the physical interpretation of Lorentzian geometry, where time plays a special role:* x^0 *is the time coordinate and* x^i *are the spatial coordinates, and greek letters range from 0 to* n*.*

The components of the metric tensor $\mathbf{g}$ in a coordinate local chart $(x^\alpha)_{\alpha=0,\ldots,n}$ are defined, as usual for tensors, by $\mathbf{g}_{\alpha\beta} = \mathbf{g}(\partial_{x^\alpha}, \partial_{x^\beta})$. The defining properties of a metric tensor implies that the matrix $(\mathbf{g}_{\alpha\beta})_{0\leq\alpha,\beta\leq n}$ is symmetric and invertible at each point of the manifold. We denote by $\mathbf{g}^{\alpha\beta}$ the components of its inverse, and thanks to the formula for the inverse of a matrix, we have $\mathbf{g}^{\alpha\beta} \in C^\infty(\mathcal{M})$. The fact that $(\mathbf{g}_{\alpha\beta})_{0\leq\alpha,\beta\leq n}$ and $(\mathbf{g}^{\alpha\beta})_{0\leq\alpha,\beta\leq n}$ are the inverse of one another is of course equivalent to the fact that their product is the identity matrix, i.e that

$$\mathbf{g}^{\alpha\beta}\mathbf{g}_{\beta\gamma} = \delta^\alpha_\beta,$$

where we crucially used Einstein's summation convention.

Lemma 2.1. *Let* $(\mathcal{M}, \mathbf{g})$ *be a pseudo-Riemannian manifold. There exists a unique* $(2,0)$*-tensor, denoted* $\mathbf{g}^{-1}$ *and called the inverse metric tensor, satisfying* $C^1_1\left(\mathbf{g}^{-1} \otimes \mathbf{g}\right) = \mathrm{Id}$ *where* $\mathrm{Id}(\omega, X) = \omega(X)$ *is the identity* $(1,1)$*-tensor. In a local chart we have* $\left(\mathbf{g}^{-1}\right)^{\alpha\beta} = \mathbf{g}^{\alpha\beta}$*.*

Proof. If such a $\mathbf{g}^{-1}$ exists, then the requirement $C^1_1\left(\mathbf{g}^{-1} \otimes \mathbf{g}\right) = \mathrm{Id}$ reads in coordinates $\left(\mathbf{g}^{-1}\right)^{\alpha\beta}\mathbf{g}_{\beta\gamma} = \delta^\alpha_\gamma$ so that by uniqueness of the inverse of a matrix we must have $\left(\mathbf{g}^{-1}\right)^{\alpha\beta} = \mathbf{g}^{\alpha\beta}$. For the existence, we would like to define a tensor $\mathbf{g}^{-1}$ be setting its component in a local chart to be $\mathbf{g}^{\alpha\beta}$. As this depends on the chart, we need to show that these components transform as a $(2,0)$-tensor (see Exercise 1.6). This is left as an exercise for the reader (use the transformation rule satisfied by the $(0,2)$-tensor $\mathbf{g}$). ☐

An important feature of $\mathbf{g}$ and $\mathbf{g}^{-1}$ is that they can be used to change the type of a tensor. We start starting with the easy case of $(1,0)$ and $(0,1)$-tensors, i.e vector fields and 1-forms.

Lemma 2.2. *Let* $X \in \Gamma(\mathcal{M})$ *and* $\omega \in \Lambda^1(\mathcal{M})$*.*

- *The map* $Y \in \Gamma(\mathcal{M}) \longmapsto \mathbf{g}(X, Y)$ *defines a 1-form denoted* $X^\flat$*, its components in a local chart are* $(X^\flat)_\alpha = \mathbf{g}_{\alpha\beta}X^\beta$*.*
- *The map* $\xi \in \Lambda^1(\mathcal{M}) \longmapsto \mathbf{g}^{-1}(\omega, \xi)$ *defines a vector field denoted* $\omega^\#$*, its components in a local chart are* $(\omega^\#)^\alpha = \mathbf{g}^{\alpha\beta}\omega_\beta$*.*

Moreover, the maps $\flat : X \in \Gamma(\mathcal{M}) \longmapsto X^\flat \in \Lambda^1(\mathcal{M})$ *and* $\# : \omega \in \Lambda^1(\mathcal{M}) \longmapsto \omega^\# \in \Gamma(\mathcal{M})$ *are* $C^\infty(\mathcal{M})$*-linear isomorphisms satisfying*

$$\flat \circ \# = \mathrm{Id}_{\Lambda^1(\mathcal{M})}, \qquad \# \circ \flat = \mathrm{Id}_{\Gamma(\mathcal{M})}. \tag{2.1}$$

Proof. The map $X^\flat$ defined by $X^\flat(Y) = \mathbf{g}(X,Y)$ is $C^\infty(\mathcal{M})$-linear and thus defines a $(0,1)$-tensor, that is a 1-form. The map $\omega^\#$ defined by $\omega^\#(\xi) = \mathbf{g}^{-1}(\omega,\xi)$ is $C^\infty(\mathcal{M})$-linear and thus defines a $(1,0)$-tensor, that is a vector field. Moreover, using the local expressions $X = X^\beta\partial_\beta$ and $\omega = \omega_\beta \mathrm{d}x^\beta$ we find

$$(X^\flat)_\alpha = X^\flat(\partial_\alpha) = \mathbf{g}(X,\partial_\alpha) = \mathbf{g}(\partial_\beta,\partial_\alpha)X^\beta = \mathbf{g}_{\alpha\beta}X^\beta,$$
$$(\omega^\#)^\alpha = \omega^\#(\mathrm{d}x^\alpha) = \mathbf{g}^{-1}(\omega,\mathrm{d}x^\alpha) = \mathbf{g}^{-1}(\mathrm{d}x^\beta,\mathrm{d}x^\alpha)\omega_\beta = \mathbf{g}^{\alpha\beta}\omega_\beta.$$

The $C^\infty(\mathcal{M})$-linearity of the maps $\flat$ and $\#$ is obvious, and the properties (2.1) can thus be checked in local coordinates

$$\flat\circ\#(\omega)_\alpha = \mathbf{g}_{\alpha\beta}(\omega^\#)^\beta = \mathbf{g}_{\alpha\beta}\mathbf{g}^{\beta\gamma}\omega_\gamma = \delta^\gamma_\alpha\omega_\gamma = \omega_\alpha,$$
$$\#\circ\flat(X)^\alpha = \mathbf{g}^{\alpha\beta}(X^\flat)_\beta = \mathbf{g}^{\alpha\beta}\mathbf{g}_{\beta\gamma}X^\gamma = \delta^\alpha_\gamma X^\gamma = X^\alpha,$$

where we used $\mathbf{g}_{\alpha\beta}\mathbf{g}^{\beta\gamma} = \delta^\gamma_\alpha$ twice. □

As we can see on the components expression $(X^\flat)_\alpha = \mathbf{g}_{\alpha\beta}X^\beta$ and $(\omega^\#)^\alpha = \mathbf{g}^{\alpha\beta}\omega_\beta$, the metric is used to lower indices while the inverse metric is used to raise indices, which explains the use of the musical notations $\flat$ and $\#$. Moreover, since the musical isomorphisms are isomorphisms, X and $X^\flat$ (or ω and $\omega^\#$) contain the exact same information, and are thus viewed as different manifestation of a single object: vector fields and 1-form are basically the same thing, and we go from one to the other by using either $\mathbf{g}$ or $\mathbf{g}^{-1}$.

2.2 The Levi-Civita connection

It will be important to be able to differentiate a vector field with respect to another one. However the standard point of view of differentiation, i.e considering the rate of change between two points, does not work here since X_p and X_q (for $X \in \Gamma(\mathcal{M})$ and $p \neq q$) live in different vector spaces so that $X_p - X_q$ is not defined. We rely instead on the notion of connection.

Definition 2.4. *Let $\mathcal{M}$ a smooth manifold. A connection D is a map*

$$D : (\Gamma(\mathcal{M}))^2 \longrightarrow \Gamma(\mathcal{M})$$
$$(X,Y) \longmapsto D_X Y$$

which is $C^\infty(\mathcal{M})$-linear with respect to its first argument, $\mathbb{R}$-linear with respect to its second argument and satisfies the following Leibniz rule

$$D_X(fY) = X(f)Y + fD_XY,$$

for $X, Y \in \Gamma(\mathcal{M})$ and $f \in C^\infty(\mathcal{M})$.

Thanks to the $C^\infty(\mathcal{M})$-linearity with respect to its first argument, the value of D_XY at p depends only on the value of X at p, whereas it depends on the value of Y on a neighborhood of Y. Moreover, the absence of $C^\infty(\mathcal{M})$-linearity with respect to the second argument shows that a connection does not define a $(1,2)$-tensor.

On a given smooth manifold there are a lot of possible connections, since locally we must have

$$D_XY = \left(X(Y^k) + X^iY^j\Gamma^k_{ij}\right)\partial_{x^k}$$

for some smooth functions Γ^k_{ij} satisfying $D_{\partial_{x^i}}\partial_{x^j} = \Gamma^k_{ij}\partial_{x^k}$. Exercise **??** shows how one can always define a connection. Crucially, the existence of a metric tensor allows us to define a canonical connection, called the Levi-Civita connection.

Theorem 2.1. *Let $(\mathcal{M},\mathbf{g})$ be a pseudo-Riemannian manifold. There exists a unique connection $\mathbf{D}$ such that* *(i)* $\mathbf{D}$ *is torsion free, i.e*

$$[X,Y] = \mathbf{D}_X Y - \mathbf{D}_Y X,$$

(ii) $\mathbf{D}$ *is compatible with* $\mathbf{g}$*, i.e*

$$X\left(\mathbf{g}(Y,Z)\right) = \mathbf{g}\left(\mathbf{D}_X Y, Z\right) + \mathbf{g}\left(Y, \mathbf{D}_X Z\right).$$

Moreover, it is characterized by the Koszul formula

$$\begin{aligned} 2\mathbf{g}(\mathbf{D}_Y Z, X) = {} & Y\left(\mathbf{g}\left(Z,X\right)\right) + Z\left(\mathbf{g}\left(Y,X\right)\right) - X\left(\mathbf{g}\left(Y,Z\right)\right) \\ & - \mathbf{g}\left(Y,[Z,X]\right) + \mathbf{g}\left(Z,[X,Y]\right) + \mathbf{g}\left(X,[Y,Z]\right). \end{aligned} \tag{2.2}$$

Proof. Since $\mathbf{g}_p$ is non-degenerate at each p, the formula (2.2) defines a unique vector field $\mathbf{D}_Y Z \in \Gamma(\mathcal{M})$. Let us show that this indeed define a connection $\mathbf{D}$ which is torsion free and compatible with the metric:

- $\mathbf{D}$ is a connection. The $\mathbb{R}$-linearity with respect to the second argument is obvious. For the Leibniz rule we compute

$$\begin{aligned} 2\mathbf{g}(\mathbf{D}_Y(fZ) - Y(f)Z - f\mathbf{D}_Y Z, X) &= Y\left(f\mathbf{g}\left(Z,X\right)\right) + fZ\left(\mathbf{g}\left(Y,X\right)\right) - X\left(f\mathbf{g}\left(Y,Z\right)\right) \\ &\quad - \mathbf{g}\left(Y,[fZ,X]\right) + f\mathbf{g}\left(Z,[X,Y]\right) + \mathbf{g}\left(X,[Y,fZ]\right) \\ &\quad - 2Y(f)\mathbf{g}(Z,X) \\ &\quad - fY\left(\mathbf{g}\left(Z,X\right)\right) - fZ\left(\mathbf{g}\left(Y,X\right)\right) + fX\left(\mathbf{g}\left(Y,Z\right)\right) \\ &\quad + f\mathbf{g}\left(Y,[Z,X]\right) - f\mathbf{g}\left(Z,[X,Y]\right) - f\mathbf{g}\left(X,[Y,Z]\right) \\ &= -Y(f)\mathbf{g}\left(Z,X\right) - X(f)\mathbf{g}\left(Y,Z\right) \\ &\quad - \mathbf{g}\left(Y,[fZ,X]\right) + \mathbf{g}\left(X,[Y,fZ]\right) \\ &\quad + f\mathbf{g}\left(Y,[Z,X]\right) - f\mathbf{g}\left(X,[Y,Z]\right) \\ &= 0, \end{aligned}$$

  where we used the Leibniz rule for vector fields and the second question of Exercise 1.4. The $C^\infty(\mathcal{M})$-linearity with respect to the first argument is proved similarily:

$$\begin{aligned} 2\mathbf{g}(\mathbf{D}_{fY} Z - f\mathbf{D}_Y Z, X) &= 2\mathbf{g}(\mathbf{D}_{fY} Z, X) - 2f\mathbf{g}(\mathbf{D}_Y Z, X) \\ &= fY\left(\mathbf{g}\left(Z,X\right)\right) + Z\left(\mathbf{g}\left(fY,X\right)\right) - X\left(\mathbf{g}\left(fY,Z\right)\right) \\ &\quad - \mathbf{g}\left(fY,[Z,X]\right) + \mathbf{g}\left(Z,[X,fY]\right) + \mathbf{g}\left(X,[fY,Z]\right) \\ &\quad - fY\left(\mathbf{g}\left(Z,X\right)\right) - fZ\left(\mathbf{g}\left(Y,X\right)\right) + fX\left(\mathbf{g}\left(Y,Z\right)\right) \\ &\quad + f\mathbf{g}\left(Y,[Z,X]\right) - f\mathbf{g}\left(Z,[X,Y]\right) - f\mathbf{g}\left(X,[Y,Z]\right) \\ &= Z(f)\mathbf{g}\left(Y,X\right) - X(f)\mathbf{g}\left(Y,Z\right) \\ &\quad + f\mathbf{g}\left(Z,[X,Y]\right) + X(f)\mathbf{g}\left(Z,Y\right) + f\mathbf{g}\left(X,[Y,Z]\right) - Z(f)\mathbf{g}\left(X,Y\right) \\ &\quad - f\mathbf{g}\left(Z,[X,Y]\right) - f\mathbf{g}\left(X,[Y,Z]\right) \\ &= 0. \end{aligned}$$

- $\mathbf{D}$ is torsion free. We compute

$$\begin{aligned} 2\mathbf{g}\left(\mathbf{D}_Y Z - \mathbf{D}_Z Y - [Y,Z], X\right) &= Y\left(\mathbf{g}\left(Z,X\right)\right) + Z\left(\mathbf{g}\left(Y,X\right)\right) - X\left(\mathbf{g}\left(Y,Z\right)\right) \\ &\quad - \mathbf{g}\left(Y,[Z,X]\right) + \mathbf{g}\left(Z,[X,Y]\right) + \mathbf{g}\left(X,[Y,Z]\right) \\ &\quad - Z\left(\mathbf{g}\left(Y,X\right)\right) - Y\left(\mathbf{g}\left(Z,X\right)\right) + X\left(\mathbf{g}\left(Y,Z\right)\right) \\ &\quad + \mathbf{g}\left(Z,[Y,X]\right) - \mathbf{g}\left(Y,[X,Z]\right) - \mathbf{g}\left(X,[Z,Y]\right) \\ &\quad - 2\mathbf{g}\left([Y,Z],X\right) \\ &= 0, \end{aligned}$$

  where we used the symmetry of $\mathbf{g}$ and the antisymmetry of the Lie bracket (see Exercise 1.4).

- $\mathbf{D}$ is compatible with $\mathbf{g}$. We compute

$$\begin{aligned}
2\left(\mathbf{g}\left(\mathbf{D}_X Y, Z\right)+\mathbf{g}\left(\mathbf{D}_X Z, Y\right)\right) &= X\left(\mathbf{g}(Y, Z)\right)+Y\left(\mathbf{g}(Z, X)\right)-Z\left(\mathbf{g}(Y, X)\right) \\
&\quad -\mathbf{g}(X,[Y, Z])+\mathbf{g}(Y,[Z, X])+\mathbf{g}(Z,[X, Y]) \\
&\quad +X\left(\mathbf{g}(Y, Z)\right)+Z\left(\mathbf{g}(Y, X)\right)-Y\left(\mathbf{g}(Z, X)\right) \\
&\quad -\mathbf{g}(X,[Z, Y])+\mathbf{g}(Z,[Y, X])+\mathbf{g}(Y,[X, Z]) \\
&= 2X\left(\mathbf{g}(Y, Z)\right)-\mathbf{g}(X,[Y, Z]+[Z, Y]) \\
&\quad +\mathbf{g}(Y,[Z, X]+[X, Z])+\mathbf{g}(Z,[X, Y]+[Y, X]) \\
&= 2X\left(\mathbf{g}(Y, Z)\right),
\end{aligned}$$

where we used the antisymmetry of the Lie bracket.

We have proved the existence of a torsion free connection which is compatible with $\mathbf{g}$. The uniqueness follows from the fact if a connection is compatible with $\mathbf{g}$ then

$$\begin{aligned}
& Y\left(\mathbf{g}(Z, X)\right)+Z\left(\mathbf{g}(Y, X)\right)-X\left(\mathbf{g}(Y, Z)\right)-\mathbf{g}(Y,[Z, X])+\mathbf{g}(Z,[X, Y])+\mathbf{g}(X,[Y, Z]) \\
& =-\mathbf{g}\left(Y,[Z, X]-\mathbf{D}_Z X+\mathbf{D}_X Z\right)+\mathbf{g}\left(Z,[X, Y]+\mathbf{D}_Y X-\mathbf{D}_X Y\right)+\mathbf{g}\left(X,[Y, Z]+\mathbf{D}_Y Z+\mathbf{D}_Z Y\right).
\end{aligned}$$

If morever $\mathbf{D}$ is assumed to be torsion free then we indeed obtain the formula (2.2) which completely characterizes $\mathbf{D}_Y Z$. □

From now on we denote the coordinate vector fields ∂_{x^α} by ∂_α.

Lemma 2.3. *If $(x^\alpha)_{\alpha=0,\ldots,n}$ is a coordinate system we define the Christoffel symbols to be the smooth scalar functions $\Gamma^\alpha_{\mu\nu}$ such that*

$$\mathbf{D}_{\partial_\mu}\partial_\nu=\Gamma^\alpha_{\mu\nu}\partial_\alpha.$$

They are given by

$$\Gamma^\alpha_{\mu\nu}=\frac{1}{2}\mathbf{g}^{\alpha\beta}\left(\partial_\mu\mathbf{g}_{\nu\beta}+\partial_\nu\mathbf{g}_{\mu\beta}-\partial_\beta\mathbf{g}_{\mu\nu}\right).$$

In particular we have $\Gamma^\alpha_{\mu\nu}=\Gamma^\alpha_{\nu\mu}$.

Proof. We start with the Koszul formula applied with ∂_μ, ∂_ν and ∂_α:

$$\begin{aligned}
2\mathbf{g}(\mathbf{D}_{\partial_\mu}\partial_\nu,\partial_\alpha) &= \partial_\mu\left(\mathbf{g}(\partial_\nu,\partial_\alpha)\right)+\partial_\nu\left(\mathbf{g}(\partial_\mu,\partial_\alpha)\right)-\partial_\alpha\left(\mathbf{g}(\partial_\mu,\partial_\nu)\right) \\
&\quad -\mathbf{g}(\partial_\mu,[\partial_\nu,\partial_\alpha])+\mathbf{g}(\partial_\nu,[\partial_\alpha,\partial_\mu])+\mathbf{g}(\partial_\alpha,[\partial_\mu,\partial_\nu]) \\
&= \partial_\mu\left(\mathbf{g}(\partial_\nu,\partial_\alpha)\right)+\partial_\nu\left(\mathbf{g}(\partial_\mu,\partial_\alpha)\right)-\partial_\alpha\left(\mathbf{g}(\partial_\mu,\partial_\nu)\right),
\end{aligned}$$

where we used the fact that the commutator of two coordinate vector fields vanishes. Using the definition of the Christoffel symbols and the metric components we thus get

$$2\Gamma^\beta_{\mu\nu}\mathbf{g}_{\alpha\beta}=\partial_\mu\mathbf{g}_{\nu\alpha}+\partial_\nu\mathbf{g}_{\mu\alpha}-\partial_\alpha\mathbf{g}_{\mu\nu}.$$

We multiply each of these equations by $\mathbf{g}^{\alpha\gamma}$ and sum over α and obtain

$$2\Gamma^\beta_{\mu\nu}\mathbf{g}_{\alpha\beta}\mathbf{g}^{\alpha\gamma}=\mathbf{g}^{\alpha\gamma}\left(\partial_\mu\mathbf{g}_{\nu\alpha}+\partial_\nu\mathbf{g}_{\mu\alpha}-\partial_\alpha\mathbf{g}_{\mu\nu}\right).$$

Using $\mathbf{g}_{\alpha\beta}\mathbf{g}^{\alpha\gamma}=\delta^\gamma_\beta$ concludes the proof. □

In the case of Minkowski spacetime $(\mathbb{R}^{3+1},\mathbf{m})$, the Christoffel symbols identically vanish in Euclidean coordinates (same for the Euclidean space). Therefore, in Euclidean coordinates $\mathbf{D}_{\partial_\alpha}\partial_\beta=0$ and the Levi-Civita connection reduces to the standard partial derivatives: if $Y=Y^\beta\partial_\beta$ then $\mathbf{D}_\alpha Y=\left(\partial_\alpha Y^\beta\right)\partial_\beta$. Despite the notation, Christoffel symbols are not the component of a tensor! See Exercise 2.2 for an example where they vanish in a coordinate system and don't vanish in another one. By using the Leibniz rule satisfied by every connection and Lemma 1.5 we can define a tensor derivation with the Levi-Civita connection.

Lemma 2.4. *Let $X\in\Gamma(\mathcal{M})$. There exists a unique tensor derivation, still denoted $\mathbf{D}_X$, such that $\mathbf{D}_X(f)=X(f)$ and $\mathbf{D}_X(Y)=\mathbf{D}_X Y$ for all $f\in C^\infty(\mathcal{M})$ and $Y\in\Gamma(\mathcal{M})$.*

This tensor derivation has very nice properties (compared for instance with the Lie derivative), see Exercise 2.4.

2.3 Curvature

As Exercise 1.8 shows, if the Lie bracket $[X,Y]$ vanishes then the commutator $[\mathcal{L}_X, \mathcal{L}_Y]$ vanishes. However, if we consider the Levi-Civita connection instead of the Lie derivative we obtain a very important object ultimately linked to the geometric idea of curvature.

2.3.1 The Riemann curvature tensor and its properties

Definition 2.5. *Let $(\mathcal{M}, \mathbf{g})$ be a pseudo-Riemannian manifold.*

- *We define the Riemann endomorphism $\mathbf{R} : \Gamma(\mathcal{M})^3 \longrightarrow \Gamma(\mathcal{M})$ by*

$$\mathbf{R}(X,Y)Z = \mathbf{D}_X\mathbf{D}_Y Z - \mathbf{D}_Y\mathbf{D}_X Z - \mathbf{D}_{[X,Y]}Z. \tag{2.3}$$

- *We define the Riemann tensor $\mathbf{Rm} : \Gamma(\mathcal{M})^4 \longrightarrow C^\infty(\mathcal{M})$ by*

$$\mathbf{Rm}(W,Z,X,Y) = \mathbf{g}(W, \mathbf{R}(X,Y)Z).$$

Lemma 2.5. *The Riemann tensor is a $(0,4)$-tensor.*

Proof. Since the $C^\infty(\mathcal{M})$-linearity of the Riemann tensor with respect to the first argument is obvious, we only need to prove the $C^\infty(\mathcal{M})$-multilinearity of the Riemann endomorphism. If $f \in C^\infty(\mathcal{M})$ we have

$$\begin{aligned}
\mathbf{R}(X,Y)fZ &= \mathbf{D}_X\mathbf{D}_Y(fZ) - \mathbf{D}_Y\mathbf{D}_X(fZ) - \mathbf{D}_{[X,Y]}(fZ) \\
&= \mathbf{D}_X\left(Y(f)Z + f\mathbf{D}_Y Z\right) - \mathbf{D}_Y\left(X(f)Z + f\mathbf{D}_X Z\right) - [X,Y](f)Z + f\mathbf{D}_{[X,Y]}Z \\
&= X(Y(f))Z + Y(f)\mathbf{D}_X Z + X(f)\mathbf{D}_Y Z + f\mathbf{D}_X\mathbf{D}_Y Z \\
&\quad - Y(X(f))Z - X(f)\mathbf{D}_Y Z - Y(f)\mathbf{D}_X Z - f\mathbf{D}_Y\mathbf{D}_X Z - [X,Y](f)Z + f\mathbf{D}_{[X,Y]}Z \\
&= f\mathbf{R}(X,Y)Z,
\end{aligned}$$

where we used the Leibniz rule and the definition of the Lie bracket. Moreover

$$\begin{aligned}
\mathbf{R}(fX,Y)Z &= \mathbf{D}_{fX}\mathbf{D}_Y Z - \mathbf{D}_Y\mathbf{D}_{fX} Z - \mathbf{D}_{[fX,Y]}Z \\
&= f\mathbf{D}_X\mathbf{D}_Y Z - \mathbf{D}_Y(f\mathbf{D}_X Z) - \mathbf{D}_{f[X,Y]-Y(f)X}Z \\
&= f\mathbf{D}_X\mathbf{D}_Y Z - Y(f)\mathbf{D}_X Z - f\mathbf{D}_Y\mathbf{D}_X Z - f\mathbf{D}_{[X,Y]}Z + Y(f)\mathbf{D}_X Z \\
&= f\mathbf{R}(X,Y)Z,
\end{aligned}$$

where we used $[fX,Y] = f[X,Y] - Y(f)X$. The last property $\mathbf{R}(X,fY)Z = f\mathbf{R}(X,Y)Z$ follows from the antisymmetry of the map $\mathbf{R}(\cdot,\cdot)Z$. ☐

The next proposition gathers algebraic and differential properties of the Riemann tensor.

Proposition 2.1. *The Riemann tensor satisfies the following properties:*

(i) Antisymmetry and symmetry:

$$\mathbf{Rm}(W,Z,X,Y) = -\mathbf{Rm}(W,Z,Y,X), \tag{2.4}$$
$$\mathbf{Rm}(W,Z,X,Y) = -\mathbf{Rm}(Z,W,X,Y), \tag{2.5}$$
$$\mathbf{Rm}(W,Z,X,Y) = \mathbf{Rm}(X,Y,W,Z). \tag{2.6}$$

(ii) First Bianchi identity:

$$\mathbf{Rm}(W,X,Y,Z) + \mathbf{Rm}(W,Y,Z,X) + \mathbf{Rm}(W,Z,X,Y) = 0, \tag{2.7}$$

(iii) Second Bianchi identity:

$$\mathbf{D}_X\mathbf{Rm}(V,W,Y,Z) + \mathbf{D}_Y\mathbf{Rm}(V,W,Z,X) + \mathbf{D}_Z\mathbf{Rm}(V,W,X,Y) = 0. \tag{2.8}$$

Proof. The antisymmetry property (2.4) directly follows from (2.3). For (2.5), we use three times the compatibility of $\mathbf{D}$ with $\mathbf{g}$:

$$\begin{aligned}\mathbf{Rm}(W,Z,X,Y) &= \mathbf{g}(\mathbf{D}_X\mathbf{D}_Y Z,W) - \mathbf{g}(\mathbf{D}_Y\mathbf{D}_X Z,W) - \mathbf{g}(\mathbf{D}_{[X,Y]}Z,W) \\ &= X(\mathbf{g}(\mathbf{D}_Y Z,W)) - Y(\mathbf{g}(\mathbf{D}_X Z,W)) + \mathbf{g}(\mathbf{D}_X Z,\mathbf{D}_Y W) - \mathbf{g}(\mathbf{D}_X W,\mathbf{D}_Y Z) \\ &\quad - [X,Y](\mathbf{g}(Z,W)) + \mathbf{g}(\mathbf{D}_{[X,Y]}W,Z) \\ &= -X(\mathbf{g}(Z,\mathbf{D}_Y W)) + Y(\mathbf{g}(Z,\mathbf{D}_X W)) + \mathbf{g}(\mathbf{D}_X Z,\mathbf{D}_Y W) - \mathbf{g}(\mathbf{D}_X W,\mathbf{D}_Y Z) \\ &\quad + \mathbf{g}(\mathbf{D}_{[X,Y]}W,Z) \\ &= -\mathbf{g}(\mathbf{D}_X Z,\mathbf{D}_Y W) - \mathbf{g}(Z,\mathbf{D}_X\mathbf{D}_Y W) + \mathbf{g}(\mathbf{D}_Y Z,\mathbf{D}_X W) + \mathbf{g}(Z,\mathbf{D}_Y\mathbf{D}_X W) \\ &\quad + \mathbf{g}(\mathbf{D}_X Z,\mathbf{D}_Y W) - \mathbf{g}(\mathbf{D}_X W,\mathbf{D}_Y Z) + \mathbf{g}(\mathbf{D}_{[X,Y]}W,Z) \\ &= -\mathbf{Rm}(Z,W,X,Y).\end{aligned}$$

For the first Bianchi identity (2.7), we use twice the torsion free property of $\mathbf{D}$:

$$\begin{aligned}&\mathbf{R}(X,Y)Z + \mathbf{R}(Z,X)Y + \mathbf{R}(Y,Z)X \\ &\quad = \mathbf{D}_{\mathbf{D}_Y Z}X + [X,\mathbf{D}_Y Z] - \mathbf{D}_{\mathbf{D}_X Z}Y - [Y,\mathbf{D}_X Z] - \mathbf{D}_{\mathbf{D}_X Y}Z + \mathbf{D}_{\mathbf{D}_Y X}Z \\ &\qquad + \mathbf{D}_{\mathbf{D}_X Y}Z + [Z,\mathbf{D}_X Y] - \mathbf{D}_{\mathbf{D}_Z Y}X - [X,\mathbf{D}_Z Y] - \mathbf{D}_{\mathbf{D}_Z X}Y + \mathbf{D}_{\mathbf{D}_X Z}Y \\ &\qquad + \mathbf{D}_{\mathbf{D}_Z X}Y + [Y,\mathbf{D}_Z X] - \mathbf{D}_{\mathbf{D}_Y X}Z - [Z,\mathbf{D}_Y X] - \mathbf{D}_{\mathbf{D}_Y Z}X + \mathbf{D}_{\mathbf{D}_Z Y}X \\ &\quad = [X,\mathbf{D}_Y Z - \mathbf{D}_Z Y] + [Y,\mathbf{D}_Z X - \mathbf{D}_X Z] + [Z,\mathbf{D}_X Y - \mathbf{D}_Y X] \\ &\quad = [X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] \\ &\quad = 0,\end{aligned}$$

where we have used the Jacobi identity for the Lie bracket (see Exercise 1.4). For (2.6), we first deduce from (2.7) the four identities

$$\begin{aligned}\mathbf{Rm}(W,Y,X,Z) + \mathbf{Rm}(W,X,Z,Y) + \mathbf{Rm}(W,Z,Y,X) &= 0, \\ \mathbf{Rm}(X,W,Z,Y) + \mathbf{Rm}(X,Z,Y,W) + \mathbf{Rm}(X,Y,W,Z) &= 0, \\ \mathbf{Rm}(Z,X,Y,W) + \mathbf{Rm}(Z,Y,W,X) + \mathbf{Rm}(Z,W,X,Y) &= 0, \\ \mathbf{Rm}(Y,Z,W,X) + \mathbf{Rm}(Y,W,X,Z) + \mathbf{Rm}(Y,X,Z,W) &= 0.\end{aligned}$$

Adding these four identities and regrouping terms together with (2.4) and (2.5) leads to

$$0 = 2\left(\mathbf{Rm}(Z,W,X,Y) - \mathbf{Rm}(X,Y,Z,W)\right),$$

which concludes the proof of (2.6). One way to prove (2.8) is to compute by brute force, but there is a more clever one which benefits from the tensorial nature of (2.8). Indeed, thanks to this tensorial nature, it is enough to prove (2.8) in a particular local coordinates system. We consider $p \in \mathcal{M}$ and consider local normal coordinates $(x^\rho)_\rho$, which are such that the Christoffel symbols vanish in this coordinates system (such coordinates can be defined with the so-called exponential map). In particular, this implies that $(\mathbf{D}_\alpha \partial_\beta)_p = 0$ in this coordinates system. Using this and the definition of a tensor derivation we obtain at p:

$$\mathbf{D}_\alpha \mathbf{Rm}(\partial_\beta,\partial_\gamma,\partial_\mu,\partial_\nu) = \partial_\alpha\left(\mathbf{Rm}(\partial_\beta,\partial_\gamma,\partial_\mu,\partial_\nu)\right) = \mathbf{g}(\partial_\beta, \mathbf{D}_\alpha \mathbf{R}(\partial_\mu,\partial_\nu)\partial_\gamma),$$

where we also the used the compatibility with the metric. We continue using $[\partial_\mu,\partial_\nu] = 0$:

$$\mathbf{D}_\alpha \mathbf{Rm}(\partial_\beta,\partial_\gamma,\partial_\mu,\partial_\nu) = \mathbf{g}(\partial_\beta, \mathbf{D}_\alpha\mathbf{D}_\mu\mathbf{D}_\nu\partial_\gamma - \mathbf{D}_\alpha\mathbf{D}_\nu\mathbf{D}_\mu\partial_\gamma).$$

This gives:

$$\begin{aligned}&\mathbf{D}_\alpha\mathbf{Rm}(\partial_\beta,\partial_\gamma,\partial_\mu,\partial_\nu) + \mathbf{D}_\mu\mathbf{Rm}(\partial_\beta,\partial_\gamma,\partial_\nu,\partial_\alpha) + \mathbf{D}_\nu\mathbf{Rm}(\partial_\beta,\partial_\gamma,\partial_\alpha,\partial_\mu) \\ &= \mathbf{g}(\partial_\beta, (\mathbf{D}_\alpha\mathbf{D}_\mu\mathbf{D}_\nu - \mathbf{D}_\alpha\mathbf{D}_\nu\mathbf{D}_\mu + \mathbf{D}_\mu\mathbf{D}_\nu\mathbf{D}_\alpha - \mathbf{D}_\mu\mathbf{D}_\alpha\mathbf{D}_\nu + \mathbf{D}_\nu\mathbf{D}_\alpha\mathbf{D}_\mu - \mathbf{D}_\nu\mathbf{D}_\mu\mathbf{D}_\alpha)\partial_\gamma) \\ &= \mathbf{g}(\partial_\beta, ((\mathbf{D}_\alpha\mathbf{D}_\mu - \mathbf{D}_\mu\mathbf{D}_\alpha)\mathbf{D}_\nu + (\mathbf{D}_\mu\mathbf{D}_\nu - \mathbf{D}_\nu\mathbf{D}_\mu)\mathbf{D}_\alpha + (\mathbf{D}_\nu\mathbf{D}_\alpha - \mathbf{D}_\alpha\mathbf{D}_\nu)\mathbf{D}_\mu)\partial_\gamma) \\ &= \mathbf{R}(\partial_\beta, \mathbf{D}_\nu\partial_\gamma, \partial_\alpha, \partial_\mu) + \mathbf{R}(\partial_\beta, \mathbf{D}_\alpha\partial_\gamma, \partial_\mu, \partial_\nu) + \mathbf{R}(\partial_\beta, \mathbf{D}_\mu\partial_\gamma, \partial_\nu, \partial_\alpha) \\ &= 0,\end{aligned}$$

where we again used $[\partial_\mu,\partial_\nu] = 0$ to reconstruct the Riemann tensor. □

2.3.2 The Einstein vacuum equations

Because of its type, the Riemann tensor **Rm** is too complicated and we would like to simplify it by contracting it. However we have only learned how to contract tensors acting on at least one 1-form and one vector field, which strictly speaking is not the case of the Riemann tensor. Nevertheless, with the help of the inverse metric we can transform one vector field into a 1-form and then contract. To be very rigorous, we define the following musical map $\# : \mathcal{T}^0_s(\mathcal{M}) \longrightarrow \mathcal{T}^1_{s-1}(\mathcal{M})$ defined by

$$T^{\#}(\omega, X_1, \ldots, X_{s-1}) = T(\omega^{\#}, X_1, \ldots, X_{s-1}).$$

We can now define the metric contraction $C_{1a} : \mathcal{T}^0_s(\mathcal{M}) \longrightarrow \mathcal{T}^0_{s-2}(\mathcal{M})$ (for $1 < a \leq s$) by

$$C_{1a}(T) = C^1_{a-1}(T^{\#}).$$

In coordinates this gives

$$C_{1a}(T)_{j_1 \cdots j_{s-2}} = \mathbf{g}^{k\ell} T_{kj_1 \cdots j_{a-2}\ell j_{a-1} \cdots j_{s-2}}.$$

Definition 2.6. *Let $(\mathcal{M}, \mathbf{g})$ be a pseudo-Riemannian manifold. We define the Ricci tensor and the scalar curvature by*

$$\mathbf{Ric} = C_{13}(\mathbf{Rm}) \quad \textit{and} \quad \mathbf{R} = C_{12}(\mathbf{Ric}).$$

The Ricci tensor is thus a $(0,2)$-tensor. In a local chart we have $\mathbf{Ric}_{\alpha\beta} = \mathbf{g}^{\mu\nu}\mathbf{Rm}_{\mu\alpha\nu\beta}$. The symmetries of the Riemann tensor gathered in Proposition 2.1 imply that the metric contraction defining the Ricci tensor is the only interesting one since $C_{12}(\mathbf{Rm}) = 0$ and $C_{14}(\mathbf{Rm}) = -C_{13}(\mathbf{Rm})$. They also imply that the Ricci tensor is a symmetric tensor since

$$\mathbf{Ric}_{\alpha\beta} - \mathbf{Ric}_{\beta\alpha} = \mathbf{g}^{\mu\nu}\mathbf{Rm}_{\mu\alpha\nu\beta} - \mathbf{g}^{\nu\mu}\mathbf{Rm}_{\nu\beta\mu\alpha} = \mathbf{g}^{\mu\nu}\mathbf{Rm}_{\mu\alpha\nu\beta} - \mathbf{g}^{\mu\nu}\mathbf{Rm}_{\mu\alpha\nu\beta} = 0,$$

where we used (2.6) and the symmetry of $\mathbf{g}$. The scalar curvature is a $(0,0)$-tensor, i.e a smooth function, locally given by $\mathbf{R} = \mathbf{g}^{\alpha\beta}\mathbf{Ric}_{\alpha\beta}$.

Now that we have the Ricci tensor and the scalar curvature, we can finally define the Einstein equations of general relativity: if $(\mathcal{M}, \mathbf{g})$ is a Lorentzian manifold, the **Einstein equations** for $\mathbf{g}$ are

$$\mathbf{Ric} - \frac{1}{2}\mathbf{R}\mathbf{g} = \mathbf{T}, \tag{2.9}$$

where $\mathbf{T}$ is a $(0,2)$-tensor called the stress energy tensor. It has to be divergence free and it describes the energy and matter in the spacetime. Several comments are in order:

- The fact that the RHS of (2.9) is divergence free necessarily implies that the LHS is also divergence free (this will be proved below). The stress energy tensor $\mathbf{T}$ usually depends on additional fields with physical meaning, such as a scalar field ϕ, an electromagnetic field $F_{\mu\nu}$, a fluid u^μ, a density of particles $f(x,p)$... The divergence free condition then recasts the wave equation for ϕ, the Maxwell equation for F, the Euler equation for u or the Vlasov equation for f. As every divergence free condition, these equations are thus naturally interpreted as conservation laws for various physical quantities.
- The LHS of (2.9), usually called the Einstein tensor, is a divergence free symmetric $(0,2)$-tensor which depends only on zeroth, first and second order derivatives of $\mathbf{g}$ and which is linear in the second order derivatives (see next section for a proof of that). A theorem of Lovelock proves that in dimension $3+1$, the only such tensors are $\mathbf{Ric} - \frac{1}{2}\mathbf{R}\mathbf{g}$ and $\mathbf{g}$ itself. This explains why the only modification of the equations (2.9) that preserves the above mentioned properties is the addition of a term of the form $\Lambda\mathbf{g}$ to the LHS, where $\Lambda \in \mathbb{R}$ is called the cosmological constant. Another way of obtaining the Einstein tensor is by minimizing the so-called Einstein-Hilbert action $\int_{\mathcal{M}} \mathbf{R}\mathrm{dVol}_{\mathbf{g}}$.
- The tensorial nature of the various curvature related objects considered here concretely implies that if (2.9) or (2.10) hold in one coordinates system, then they hold in any coordinates system. This is the relativity principle. The equivalence principle, made famous by Einstein's thought experiment, translates mathematically to the existence of normal coordinates along a geodesic (since in such coordinates the Christoffel symbols vanish and the geodesic equation is simply $\ddot{x} = 0$).

In order to define the divergence of a tensor, we first define $\mathbf{D} : \mathcal{T}^0_s(\mathcal{M}) \longrightarrow \mathcal{T}^0_{s+1}(\mathcal{M})$ by

$$\mathbf{D}T(X_0, X_1, \ldots, X_s) = \mathbf{D}_{X_0} T(X_1, \ldots, X_s).$$

If now $T \in \mathcal{T}^0_s(\mathcal{M})$ with $s \geq 1$, we define its divergence to be the $(0, s-1)$-tensor defined by

$$\mathrm{div} T = C_{12}(\mathbf{D}T).$$

Proposition 2.2. *We have*

$$\mathrm{div}\left(\mathbf{Ric} - \frac{1}{2}\mathbf{R}\mathbf{g}\right) = 0.$$

Proof. We first note that

$$\begin{aligned} \mathrm{div}(f\mathbf{g}) &= C_{12}(\mathbf{D}(f\mathbf{g})) \\ &= C_{12}(f\mathbf{D}\mathbf{g} + \mathbf{D}f \otimes \mathbf{g}) \\ &= C_{12}(\mathrm{d}f \otimes \mathbf{g}) \\ &= \mathrm{d}f, \end{aligned}$$

where we used that $\mathbf{D}\mathbf{g} = 0$ and $\mathbf{D}f = \mathrm{d}f$. Therefore we need to prove that $\mathrm{div}\mathbf{Ric} = \frac{1}{2}\mathrm{d}\mathbf{R}$. We start from (2.8)

$$\mathbf{D}_X\mathbf{Rm}(V, W, Y, Z) + \mathbf{D}_Y\mathbf{Rm}(V, W, Z, X) + \mathbf{D}_Z\mathbf{Rm}(V, W, X, Y) = 0.$$

We use (2.4) (which can be shown to also hold for $\mathbf{D}_Z\mathbf{Rm}$) for the last term and rewrite the middle term with the definition of $\mathbf{D}$:

$$\mathbf{D}_X\mathbf{Rm}(V, W, Y, Z) + \mathbf{D}\mathbf{Rm}(Y, V, W, Z, X) - \mathbf{D}_Z\mathbf{Rm}(V, W, Y, X) = 0.$$

Since $\mathbf{D}_X\mathbf{g}^{-1} = 0$, we have $C_{1a}(\mathbf{D}_X T) = \mathbf{D}_X(C_{1a}(T))$ and therefore applying C_{13} to the above identity gives

$$\mathbf{D}_X\mathbf{Ric}(W, Z) + \mathrm{div}\mathbf{Rm}(W, Z, X) - \mathbf{D}_Z\mathbf{Ric}(W, X) = 0.$$

Using the definition of $\mathbf{D}$ for the first term this rewrites

$$\mathbf{D}\mathbf{Ric}(X, W, Z) + \mathrm{div}\mathbf{Rm}(W, Z, X) - \mathbf{D}_Z\mathbf{Ric}(W, X) = 0.$$

We now apply C_{12} to this equality (using again the commutativity with $\mathbf{D}_Z$):

$$\mathrm{div}\mathbf{Ric} + C_{13}(\mathrm{div}\mathbf{Rm}) = \mathrm{d}\mathbf{R},$$

where we used $\mathbf{D}_Z\mathbf{R} = \mathrm{d}\mathbf{R}(Z)$. Again because of $\mathbf{D}\mathbf{g}^{-1} = 0$ we can prove that $C_{13}(\mathrm{div}\mathbf{Rm}) = \mathrm{div}\mathbf{Ric}$, so that we have proved $2\mathrm{div}\mathbf{Ric} = \mathrm{d}\mathbf{R}$, which concludes the proof. □

2.3.3 Wave coordinates

If $\mathbf{T} = 0$, then the Einstein equations (2.9) simplify: contracting $\mathbf{Ric} = \frac{1}{2}\mathbf{R}\mathbf{g}$ with $\mathbf{g}^{-1}$ gives $\mathbf{R} = \frac{n}{2}\mathbf{R}$ so that $\mathbf{R} = 0$ (if $n \geq 3$). Therefore the **Einstein vacuum equations** are simply

$$\mathbf{Ric} = 0. \tag{2.10}$$

In Minkowski spacetime $(\mathbb{R}^{1+3}, \mathbf{m})$ the Riemann tensor identically vanishes and thus this spacetime is an obvious solutions of (2.10). If we think of $\mathbf{T}$ as the source term in (2.9) and thus of (2.10) as being an equation without source, we might deduce that $(\mathbb{R}^{1+3}, \mathbf{m})$ is the only solution to (2.10). This is far from being the case! We give two arguments in favor of the non-triviality of (2.10). First, shortly after the publication of Einstein's theory of general relativity, Schwarzschild discovered the so-called Schwarzschild metric

$$\mathbf{g}_{Sch} = -\left(1 - \frac{2m}{r}\right)\mathrm{d}t \otimes \mathrm{d}t + \left(1 - \frac{2m}{r}\right)^{-1}\mathrm{d}r \otimes \mathrm{d}r + r^2\mathbf{g}_{\mathbb{S}^2},$$

where $m > 0$. By considering more adapted coordinates, this metric can be shown to solve (2.10) on the manifold $\mathbb{R}_t \times \{r > 0\} \times \mathbb{S}^2$ and its Riemann tensor does not vanish (in particular implying that the Schwarzschild spacetime is not isometric to Minkowski). The physical meaning of $\mathbf{g}_{Sch}$ is the description of a black hole without rotation, which is thus a non-trivial solution of the Einstein vacuum equations. However, this argument is not entirely satisfactory since one could see $\mathbf{g}_{Sch}$ as a solution of (2.9) with a measured-valued stress-energy tensor concentrated at $r = 0$.

The best way to show that (2.10) admits plenty of regular solutions is actually to exhibit its PDE nature. In order to do so, we define the so-called wave operator associated to any Lorentzian metric $\mathbf{g}$:

$$\Box_{\mathbf{g}} f = C_{12}(\mathrm{Hess}(f)),$$

where $\mathrm{Hess}(f)$ is defined in Exercise 2.5. In local coordinates this becomes

$$\Box_{\mathbf{g}} f = \mathbf{g}^{\alpha\beta}\mathrm{Hess}(f)(\partial_\alpha, \partial_\beta) = \mathbf{g}^{\alpha\beta}\left(\partial_\alpha\partial_\beta f - \Gamma^\mu_{\alpha\beta}\partial_\mu f\right).$$

Note that in the case of Minkowski spacetime in Euclidean coordinates we recover the standard wave operator $\Box = -\partial_t^2 + \Delta$, where $\Delta = \partial^i\partial_i$. Note that in the case of a Riemannian metric the very same definition leads to the so-called Laplace-Beltrami operator $\Delta_{\mathbf{g}}$, which generalizes Δ.

Proposition 2.3. *In any coordinates system, the Ricci tensor admits the decomposition*

$$\mathbf{Ric}_{\mu\nu} = -\frac{1}{2}\Box_{\mathbf{g}}\mathbf{g}_{\mu\nu} + \frac{1}{2}\left(\mathbf{g}_{\rho\mu}\partial_\nu H^\rho + \mathbf{g}_{\rho\nu}\partial_\mu H^\rho\right) + P_{\mu\nu}(\mathbf{g})(\partial\mathbf{g}, \partial\mathbf{g}),$$

where $H^\rho := \mathbf{g}^{\alpha\beta}\Gamma^\rho_{\alpha\beta}$ and where $P_{\mu\nu}(\mathbf{g})(\partial\mathbf{g}, \partial\mathbf{g})$ denotes terms of the form $\mathbf{g}^{-1}\mathbf{g}^{-1}\partial\mathbf{g}\partial\mathbf{g}$.

Proof. Since we don't need the exact expression of the semilinear terms of the form $\mathbf{g}^{-1}\mathbf{g}^{-1}\partial\mathbf{g}\partial\mathbf{g}$, we will denote them by $O\left(\left(\mathbf{g}^{-1}\partial\mathbf{g}\right)^2\right)$. In coordinates we have:

$$\begin{aligned}
\mathbf{Ric}_{\mu\nu} &= \mathbf{g}^{\alpha\beta}\mathbf{g}\left(\partial_\alpha, \mathbf{R}(\partial_\beta, \partial_\nu)\partial_\mu\right) \\
&= \mathbf{g}^{\alpha\beta}\left(\mathbf{g}\left(\partial_\alpha, \mathbf{D}_\beta\mathbf{D}_\nu\partial_\mu\right) - \mathbf{g}\left(\partial_\alpha, \mathbf{D}_\nu\mathbf{D}_\beta\partial_\mu\right)\right) \\
&= \mathbf{g}^{\alpha\beta}\left(\mathbf{g}\left(\partial_\alpha, \mathbf{D}_\beta\left(\Gamma^\rho_{\nu\mu}\partial_\rho\right)\right) - \mathbf{g}\left(\partial_\alpha, \mathbf{D}_\nu\left(\Gamma^\rho_{\beta\mu}\partial_\rho\right)\right)\right) \\
&= \mathbf{g}^{\alpha\beta}\mathbf{g}_{\alpha\rho}\partial_\beta\Gamma^\rho_{\nu\mu} - \mathbf{g}^{\alpha\beta}\mathbf{g}_{\alpha\rho}\partial_\nu\Gamma^\rho_{\beta\mu} + O\left(\left(\mathbf{g}^{-1}\partial\mathbf{g}\right)^2\right) \\
&= \partial_\beta\Gamma^\beta_{\nu\mu} - \partial_\nu\Gamma^\beta_{\beta\mu} + O\left(\left(\mathbf{g}^{-1}\partial\mathbf{g}\right)^2\right).
\end{aligned}$$

By differentiating the relation $\mathbf{g}^{\alpha\beta}\mathbf{g}_{\beta\gamma} = \delta^\alpha_\gamma$ we find that $\partial_\alpha\mathbf{g}^{\mu\nu} = -\mathbf{g}^{\rho\mu}\mathbf{g}^{\sigma\nu}\partial_\alpha\mathbf{g}_{\rho\sigma}$ which schematically reads $\partial\mathbf{g}^{-1} = \mathbf{g}^{-1}\mathbf{g}^{-1}\partial\mathbf{g}$. Therefore, when differentiating the Christoffel symbols, we schematically get $\partial\left(\mathbf{g}^{-1}\partial\mathbf{g}\right) = \mathbf{g}^{-1}\partial^2\mathbf{g} + \left(\mathbf{g}^{-1}\partial\mathbf{g}\right)^2$. Therefore we have

$$\begin{aligned}
\mathbf{Ric}_{\mu\nu} &= \frac{1}{2}\mathbf{g}^{\alpha\beta}\left(\partial_\beta\partial_\mu\mathbf{g}_{\alpha\nu} + \partial_\beta\partial_\nu\mathbf{g}_{\alpha\mu} - \partial_\beta\partial_\alpha\mathbf{g}_{\mu\nu} - \partial_\nu\partial_\mu\mathbf{g}_{\alpha\beta} - \partial_\nu\partial_\beta\mathbf{g}_{\alpha\mu} + \partial_\nu\partial_\alpha\mathbf{g}_{\mu\beta}\right) + O\left(\left(\mathbf{g}^{-1}\partial\mathbf{g}\right)^2\right) \\
&= -\frac{1}{2}\Box_{\mathbf{g}}\mathbf{g}_{\mu\nu} + \frac{1}{2}\mathbf{g}^{\alpha\beta}\left(\partial_\beta\partial_\mu\mathbf{g}_{\alpha\nu} - \partial_\nu\partial_\mu\mathbf{g}_{\alpha\beta} + \partial_\nu\partial_\alpha\mathbf{g}_{\mu\beta}\right) + O\left(\left(\mathbf{g}^{-1}\partial\mathbf{g}\right)^2\right),
\end{aligned}$$

where we used the fact that $\Box_{\mathbf{g}} f = \mathbf{g}^{\alpha\beta}\partial_\alpha\partial_\beta f + O\left(\mathbf{g}^{-1}\mathbf{g}^{-1}\partial\mathbf{g}\partial f\right)$. On the other hand we have $H^\rho = \mathbf{g}^{\alpha\beta}\mathbf{g}^{\rho\sigma}\left(\partial_\alpha\mathbf{g}_{\beta\sigma} - \frac{1}{2}\partial_\sigma\mathbf{g}_{\alpha\beta}\right)$ and thus

$$\begin{aligned}
\frac{1}{2}\left(\mathbf{g}_{\rho\mu}\partial_\nu H^\rho + \mathbf{g}_{\rho\nu}\partial_\mu H^\rho\right) &= \frac{1}{2}\mathbf{g}^{\alpha\beta}\mathbf{g}^{\rho\sigma}\left(\mathbf{g}_{\rho\mu}\partial_\nu\left(\partial_\alpha\mathbf{g}_{\beta\sigma} - \frac{1}{2}\partial_\sigma\mathbf{g}_{\alpha\beta}\right) + \mathbf{g}_{\rho\nu}\partial_\mu\left(\partial_\alpha\mathbf{g}_{\beta\sigma} - \frac{1}{2}\partial_\sigma\mathbf{g}_{\alpha\beta}\right)\right) \\
&\quad + O\left(\left(\mathbf{g}^{-1}\partial\mathbf{g}\right)^2\right) \\
&= \frac{1}{2}\mathbf{g}^{\alpha\beta}\left(\partial_\nu\partial_\alpha\mathbf{g}_{\beta\mu} + \partial_\mu\partial_\alpha\mathbf{g}_{\beta\nu} - \partial_\mu\partial_\nu\mathbf{g}_{\alpha\beta}\right) + O\left(\left(\mathbf{g}^{-1}\partial\mathbf{g}\right)^2\right),
\end{aligned}$$

which concludes the proof. □

The previous proposition shows that if one can construct coordinates $(x^\rho)_{\rho=0,\ldots,3}$ such that $H^\rho = 0$, then the Einstein vacuum equations become

$$\Box_{\mathbf{g}} \mathbf{g}_{\mu\nu} = 2P_{\mu\nu}(\mathbf{g})(\partial\mathbf{g}, \partial\mathbf{g}).$$

In the vocabulary of PDEs, this is a system of 10 coupled wave equations for the metric coefficients $\mathbf{g}_{\mu\nu}$. This system is nonlinear for two reasons:

- The RHS is a quadratic expression in first order derivatives of the metric $\partial\mathbf{g}$, such terms are called semilinear. Though we did not give its exact expression, its structure plays a crucial role in many problems, such as the long-time behaviour of solutions or the absence of shocks in the evolution.
- The wave operator on the LHS depends obviously on the solution $\mathbf{g}$, such term is called quasilinear. The existence of gravitational waves can be predicted from this wave operator, even though a proof of their physicality (i.e that they cannot be erased by a change of coordinates) requires more work.

Coordinates $(x^\rho)_{\rho=0,\ldots,3}$ such that $H^\rho = 0$ are called wave coordinates since $H^\rho = -\Box_{\mathbf{g}} x^\rho$. The fact that, with an appropriate choice of coordinates, the Einstein vacuum equations can be recast as a system of wave equations hints at a Cauchy formulation of these equations in analogy with the Cauchy formulation for the standard wave equation

$$\begin{cases} \Box u = F, \\ (u, \partial_t u)_{|\{t=0\}} = (f, h), \end{cases} \tag{2.11}$$

where F, f and h are given. Of course, the geometric nature of the Einstein vacuum equations would require several modifications of (2.11):

- The hypersurface $\{t = 0\}$ in Minkowski is replaced by the concept of a spacelike hypersurface Σ, i.e a $n-1$-dimensional submanifold of $\mathcal{M}$ with timelike normal vector field N.
- Similarly, $u_{|\{t=0\}}$ is replaced by the induced metric $\mathbf{g}_{|\Sigma}$ on the hypersurface, and $\partial_t u_{|\{t=0\}}$ is replaced by the Lie derivative $\mathcal{L}_N \mathbf{g}_{|\Sigma}$ (since $\mathbf{D}_N \mathbf{g} = 0$ we cannot use the Levi-Civita connection here).

Another consequence of the geometric nature of general relativity is that the Cauchy data $(\mathbf{g}, \mathcal{L}_N\mathbf{g})_{|\Sigma}$ cannot be freely chosen, they need to solve the so-called constraint equations. This distinguishes drastically the Einstein vacuum equations from (2.11), where (f, h) can be freely chosen. Nevertheless, understanding the existence, uniqueness and behaviour of solutions to (2.11) (and its semilinear and quasilinear generalisations) is a necessary prerequisite to the study of the Einstein vacuum equations and their solutions.

2.4 Exercises

Exercise 2.1. *Let $\mathcal{M}$ be a smooth manifold and D and $\tilde{D}$ two connections.*

1. *Show that*

$$(\omega, X, Y) \in \Lambda^1(\mathcal{M}) \times (\Gamma(\mathcal{M}))^2 \longmapsto \omega\left(D_X Y - D_Y X - [X, Y]\right)$$

defines a $(1, 2)$-tensor (called the torsion tensor of D).

2. *Show that*

$$(\omega, X, Y) \in \Lambda^1(\mathcal{M}) \times (\Gamma(\mathcal{M}))^2 \longmapsto \omega\left(D_X Y - \tilde{D}_X Y\right)$$

defines a $(1, 2)$-tensor.

Exercise 2.2. *Let $(\mathbb{R}^3, \mathbf{g}_{eucl})$ be the 3-dimensional Euclidean space. Compute the components of $\mathbf{g}_{eucl}$ and the Christoffel symbols in spherical coordinates.*

Exercise 2.3. *Let $(\mathcal{M}, \mathbf{g})$ be a pseudo-Riemannian manifold and $f \in C^\infty(\mathcal{M})$. We define another pseudo-Riemannian metric $\tilde{\mathbf{g}} = e^{2f}\mathbf{g}$. Give the expression of $\tilde{\mathbf{D}}$, the Levi-Civita connection associated to $\tilde{\mathbf{g}}$, in terms of $\mathbf{D}$, the Levi-Civita connection associated to $\mathbf{g}$.*

Exercise 2.4. *Let $(\mathcal{M}, \mathbf{g})$ be a pseudo-Riemannian manifold and $X \in \Gamma(\mathcal{M})$. Show that the tensor derivation $\mathbf{D}_X$ satisfies the following properties*

(i) $\mathbf{D}_X \mathbf{g} = 0$ *and* $\mathbf{D}_X \mathbf{g}^{-1} = 0$.

(ii) It commutes with the musical isomorphisms (hint: rewrite them with standard contractions and tensor product).

Exercise 2.5. *Let $(\mathcal{M}, \mathbf{g})$ be a pseudo-Riemannian manifold and $f \in C^\infty(\mathcal{M})$. We define the Hessian of f by* $\mathrm{Hess}(f)(X,Y) = \mathbf{D}_X \mathrm{d}f(Y)$ *for $X, Y \in \Gamma(\mathcal{M})$. Show that* $\mathrm{Hess}(f)$ *is a symmetric $(0,2)$-tensor.*

Exercise 2.6. *If $V \in \Gamma(\mathcal{M})$, we define its curl by*

$$\mathrm{curl}(V)(X,Y) = \mathbf{g}(\mathbf{D}_X V, Y) - \mathbf{g}(\mathbf{D}_Y V, X).$$

1. *Show that* $\mathrm{curl}(V) = \mathbf{d}V^\flat$ *where* $\mathbf{d}$ *is defined in Exercise 1.5.*
2. *Define* $\mathrm{grad} f = (\mathrm{d}f)^\#$ *and show that* $\mathrm{curl}(\mathrm{grad} f) = 0$.

Exercise 2.7. *If $X \in \Gamma(\mathcal{M})$, we define its divergence by* $\mathrm{div} X = \mathrm{div} X^\flat$*, where $X^\flat$ is seen as a $(0,1)$-tensor.*

1. *Compute the expression of* $\mathrm{div} X$ *in coordinates.*
2. *Show that* $\Box_{\mathbf{g}} f = \mathrm{div}(\mathrm{grad} f)$.
3. *Show the alternative expression*

$$\mathrm{div} X = \frac{1}{\sqrt{-\det \mathbf{g}}} \partial_\beta \left(\sqrt{-\det \mathbf{g}} X^\beta \right),$$

where $\det \mathbf{g}$ *denotes the determinant of the matrix* $(\mathbf{g}_{\alpha\beta})_{0 \leq \alpha,\beta \leq 3}$.

Exercise 2.8. *Let $(\mathcal{M}, \mathbf{g})$ be a pseudo-Riemannian manifold.*

1. *Let $\varphi \in C^\infty(\mathcal{M})$ such that $\Box_{\mathbf{g}} \varphi = 0$ and define the scalar field stress-energy tensor*

$$T = \mathrm{d}\varphi \otimes \mathrm{d}\varphi - \frac{1}{2} \mathbf{g}(\mathrm{grad}\varphi, \mathrm{grad}\varphi) \mathbf{g}.$$

Show that $\mathrm{div} T = 0$.

2. *Let F be an antisymmetric $(0,2)$-tensor satisfying the Maxwell vacuum equations*

$$\mathbf{D}_\alpha F_{\beta\gamma} + \mathbf{D}_\beta F_{\gamma\alpha} + \mathbf{D}_\gamma F_{\alpha\beta} = 0,$$
$$\mathrm{div} F = 0,$$

and define the electromagnetic stress-energy tensor

$$T_{\mu\nu} = \mathbf{g}^{\alpha\beta} F_{\mu\alpha} F_{\nu\beta} - \frac{1}{4} \mathbf{g}_{\mu\nu} \mathbf{g}^{\alpha\rho} \mathbf{g}^{\beta\sigma} F_{\alpha\beta} F_{\rho\sigma}.$$

Show that $\mathrm{div} T = 0$.

Chapter 3

Correction of the exercises

3.1 Exercises from Chapter 1

Exercise 3.1.

1. *Since each point on* $\mathbb{S}^2$ *has at least one non-zero coordinates, we indeed have*

$$\mathbb{S}^2 = \bigcup_{j=1}^{3} (U_j \cup V_j).$$

The maps φ_j *are homeomorphisms from* U_j *to the open disk* $\mathbb{D} := \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2<1\}$ *and from* V_j *to* $\mathbb{D}$. *For instance, the inverse of* $\varphi_1 : V_1 \longrightarrow \mathbb{D}$ *is given by* $\varphi_1^{-1}(x,y) = \left(-\sqrt{1-x^2-y^2}, x, y\right)$. *Therefore* $\big((U_1,\varphi_1),(U_2,\varphi_2),(U_3,\varphi_3),(V_1,\varphi_1),(V_2,\varphi_2),(V_3,\varphi_3)\big)$ *is an atlas for* $\mathbb{S}^2$. *To prove that it makes* $\mathbb{S}^2$ *a smooth manifold we only need to show that the transition maps are smooth where they are defined (since the Hausdorff property is obvious). Note that* $U_i \cap U_j \neq \emptyset$, $V_i \cap V_j \neq \emptyset$ *and* $U_i \cap V_j \neq \emptyset$ *if and only if* $i \neq j$. *Therefore, there are* 12 *intersections to consider and thus* 24 *transition maps to consider. We only treat three of them:*

- *Consider* (U_1,φ_1) *and* (U_2,φ_2). *We have* $U_1 \cap U_2 = \{(x^1,x^2,x^3) \in \mathbb{S}^2 \mid x^1, x^2 > 0\}$ *and if* $(x,y) \in \varphi_1(U_1 \cap U_2)$ *then*

$$\varphi_2 \circ \varphi_1^{-1}(x,y) = \varphi_2\left(\sqrt{1-x^2-y^2}, x, y\right) = \left(\sqrt{1-x^2-y^2}, y\right).$$

- *Consider* (V_2,φ_2) *and* (V_3,φ_3). *We have* $V_2 \cap V_3 = \{(x^1,x^2,x^3) \in \mathbb{S}^2 \mid x^2, x^3 < 0\}$ *and if* $(x,y) \in \varphi_2(V_2 \cap V_3)$ *then*

$$\varphi_3 \circ \varphi_2^{-1}(x,y) = \varphi_3\left(x, -\sqrt{1-x^2-y^2}, y\right) = \left(x, -\sqrt{1-x^2-y^2}\right).$$

- *Consider* (V_1,φ_1) *and* (U_3,φ_3). *We have* $V_1 \cap U_3 = \{(x^1,x^2,x^3) \in \mathbb{S}^2 \mid x^1 < 0,\ x^3 > 0\}$ *and if* $(x,y) \in \varphi_3(V_1 \cap U_3)$ *then*

$$\varphi_1 \circ \varphi_3^{-1}(x,y) = \varphi_1\left(x, y, \sqrt{1-x^2-y^2}\right) = \left(y, \sqrt{1-x^2-y^2}\right).$$

All these maps are smooth so $\mathbb{S}^2$ *with this atlas is a smooth manifold.*

2. *We obviously have* $\mathbb{S}^2 = U_1 \cup U_2$. *The map* φ_1 *is a homeomorphism from* U_1 *to* $\mathbb{R}^2$ *with inverse*

$$\varphi_1^{-1}(x,y) = \left(\frac{2x}{x^2+y^2+1}, \frac{2y}{x^2+y^2+1}, \frac{x^2+y^2-1}{x^2+y^2+1}\right).$$

The map φ_2 is a homeomorphism from U_2 to $\mathbb{R}^2$ with inverse

$$\varphi_2^{-1}(x,y) = \left(\frac{2x}{x^2+y^2+1}, \frac{2y}{x^2+y^2+1}, \frac{1-x^2-y^2}{x^2+y^2+1} \right).$$

Therefore $\big((U_1,\varphi_1),(U_2,\varphi_2)\big)$ is an atlas for $\mathbb{S}^2$. We have $U_1 \cap U_2 = \mathbb{S}^2 \setminus \{(0,0,1),(0,0,-1)\}$ and $\varphi_i(U_1\cap U_2) = \mathbb{R}^2 \setminus \{(0,0\}$ for $i=1,2$. Moreover for $(x,y) \neq (0,0)$ we have

$$\varphi_2 \circ \varphi_1^{-1}(x,y) = \varphi_1 \circ \varphi_2^{-1}(x,y) = \left(\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2} \right),$$

which is a smooth function. Therefore $\mathbb{S}^2$ with this atlas is a smooth manifold.

3. *We can generalize both examples to $\mathbb{S}^n$ for $n \geq 1$.*
   - *Define $U_j^{\pm} = \left\{ (x^1,\dots,x^{n+1}) \in \mathbb{S}^n \;\middle|\; \pm x^j > 0 \right\}$ and $\varphi_j\left((x^1,\dots,x^{n+1})\right) = (x^1,\dots,x^{j-1},x^{j+1},\dots,x^n)$. We can show that $\big((U_1^+,\varphi_1),\dots,(U_{n+1}^+,\varphi_{n+1}),(U_1^-,\varphi_1),\dots,(U_{n+1}^-,\varphi_{n+1})\big)$ is an atlas making $\mathbb{S}^n$ a smooth manifold composed of $2(n+1)$ charts.*
   - *Define $U^{\pm} = \mathbb{S}^n \setminus \{(0,\dots,0,\pm 1\}$ and*

     $$\varphi^{\pm}\left((x^1,\dots,x^n,x^{n+1})\right) = \left(\frac{x^1}{1 \mp x^{n+1}}, \dots, \frac{x^n}{1 \mp x^{n+1}} \right).$$

     We can show that $\big((U^+,\varphi^+),(U^-,\varphi^-)\big)$ is an atlas making $\mathbb{S}^n$ a smooth manifold composed of 2 charts.

Exercise 3.2.

1. *We first note that f is non-zero and smooth (which can be proved by induction on the derivatives $f^{(k)}$ of f) and that $\operatorname{supp} f = [-1,1]$. Differentiating under the integral symbol (using integrability of smooth functions on compact sets) we can also prove that h is smooth. Changing variable we rewrite h as*

   $$h(t) = \frac{1}{\int_{\mathbb{R}} f} \int_{t-2}^{t+2} f$$

   which already shows that $0 \leq h \leq 1$. Moreover if $|t| > 3$, then $[t-2,t+2] \cap [-1,1] = \emptyset$ so that $h(t) = 0$. Finally, if $|t| \leq 1$, then $[-1,1] \subset [t-2,t+2]$ and thus $\int_{t-2}^{t+2} f = \int_{\mathbb{R}} f$ so that $h(t) = 1$.

2. *We start by dilating h: for $\varepsilon > 0$ we define $h_\varepsilon(t) = h\left(\frac{t}{\varepsilon}\right)$. It is a smooth function satisfying $0 \leq h_\varepsilon \leq 1$, $\operatorname{supp}(h_\varepsilon) \subset [-3\varepsilon, 3\varepsilon]$ and $h_{|[-\varepsilon,\varepsilon]} = 1$. Now, let $p \in \mathcal{M}$ and U a neighborhood of p. Let (V,φ) a local chart around p such that $\overline{V} \subset U$. If $\varepsilon > 0$ is sufficiently small, we have $B_{4\varepsilon} := \left\{ x \in \mathbb{R}^n \;\middle|\; \|x - \varphi(p)\|_2^2 \leq 4\varepsilon \right\} \subset \varphi(V)$. On V we can define $N(q) = \sum_{i=1}^n \left(x^i(q) - x^i(p)\right)^2$ so that $\varphi^{-1}(B_\eta) = \{q \in V \mid N(q) \leq \eta\}$ for all $\eta \leq 4\varepsilon$. We set $\chi = h_\varepsilon \circ N$ on V and $\chi = 0$ elsewhere. On $\varphi^{-1}(B_\varepsilon)$ we have $N \leq \varepsilon$ so that $\chi = 1$ on $\varphi^{-1}(B_\varepsilon)$ (which is indeed a neighborhood of p), and $\operatorname{supp}\chi \subset U$. Finally the smoothness of χ follows from the smoothness of h_ε and the fact that there is only one local chart to check.*

Exercise 3.3.

1. *Since $c(0) = p$ and φ is defined in a neighborhood of p, the function $\varphi \circ c$ is well-defined on a neighborhood of 0 for every $c \in C_p\mathcal{M}$ and $(\varphi \circ c)'(0)$ is well-defined (since both φ and c are smooth). Moreover if φ and ψ are two local charts around p, we have $\varphi \circ c = \varphi \circ \psi^{-1} \circ \psi \circ c$ so that the chain rule in $\mathbb{R}^n$ implies*

   $$(\varphi \circ c)'(0) = \mathrm{d}\left(\varphi \circ \psi^{-1}\right)_{\psi(p)} (\psi \circ c)'(0),$$

   where $\mathrm{d}\left(\varphi \circ \psi^{-1}\right)_{\psi(p)}$ is the standard differential of the function $\varphi \circ \psi^{-1}$ at $\psi(p) \in \mathbb{R}^n$ (remember that $(\varphi \circ c)'(0)$ and $(\psi \circ c)'(0)$ are vectors in $\mathbb{R}^n$). However, since the transition maps $\varphi \circ \psi^{-1}$ are supposed to be smooth diffeomorphisms (where they are defined), their differential is a linear isomorphism and

   $$(\varphi \circ c_1)'(0) = (\varphi \circ c_2)'(0) \iff (\psi \circ c_1)'(0) = (\psi \circ c_2)'(0).$$

2. *For* $c \in C_p\mathcal{M}$ *we denote by* $[c]_p \in \widetilde{T}_p\mathcal{M}$ *the equivalence class of* c. *For* $c, c_1, c_2 \in C_p\mathcal{M}$ *and* $\lambda \in \mathbb{R}$ *we define*

$$\begin{aligned} [c_1]_p + [c_2]_p &:= \left[\varphi^{-1} \circ (\varphi \circ c_1 + \varphi \circ c_2)\right]_p , \\ \lambda[c]_p &:= \left[\varphi^{-1} \circ (\lambda\varphi \circ c)\right]_p . \end{aligned}$$

We need to show that the right hand sides don't depend on the representant, i.e that if $c_1 \sim \tilde{c}_1$ *and* $c_2 \sim \tilde{c}_2$ *then*

$$\varphi^{-1} \circ (\varphi \circ c_1 + \varphi \circ c_2) \sim \varphi^{-1} \circ (\varphi \circ \tilde{c}_1 + \varphi \circ \tilde{c}_2) .$$

Since $\varphi \circ \varphi^{-1} \circ (\varphi \circ c_1 + \varphi \circ c_2) = \varphi \circ c_1 + \varphi \circ c_2$, *the linearity of the derivatives of a function from* $\mathbb{R}$ *to* $\mathbb{R}^n$ *implies what we want. An identical reasoning proves that if* $c \sim \tilde{c}$ *then* $\varphi^{-1} \circ (\lambda\varphi \circ c) \sim \varphi^{-1} \circ (\lambda\varphi \circ \tilde{c})$. *Therefore the above operations are well-defined on* $\widetilde{T}_p\mathcal{M}$. *All the algebraic properties that a vector space need to satisfy are obvious, except maybe who is the zero element for* $+$: *it is the equivalence class* $0_{\widetilde{T}_p\mathcal{M}}$ *of the constant path* $c(t) = p$, *and we indeed have* $[c]_p + [0]_p = \left[\varphi^{-1} \circ (\varphi \circ c + \varphi(p))\right]_p = [c]_p$ *since* $\varphi(p) = 0$.

3. *For* $[c]_p \in \widetilde{T}_p\mathcal{M}$ *we define* $\Psi([c]_p) \in T_p\mathcal{M}$ *to be the derivation at* p *defined by*

$$\Psi([c]_p)(f) = (f \circ c)'(0)$$

for all $f \in C^\infty(\mathcal{M})$. *We first need to show that* $(f \circ c)'(0)$ *does not depend on the representant in the equivalence class* $[c]_p$. *We compute using again the chain rule:*

$$\begin{aligned} (f \circ c)'(0) &= (f \circ \varphi^{-1} \circ \varphi \circ c)'(0) \\ &= \partial_i(f \circ \varphi^{-1})(\varphi(p))\pi^i\left((\varphi \circ c)'(0)\right) \\ &= \partial_{x^i|_p}(f)\pi^i\left((\varphi \circ c)'(0)\right), \end{aligned}$$

where $(x^i)_{i=1,\dots,n}$ *is associated with* φ. *Therefore, if* $c_1 \sim c_2$ *then* $(f \circ c_1)'(0) = (f \circ c_2)'(0)$ *and* $\Psi : \widetilde{T}_p\mathcal{M} \longrightarrow T_p\mathcal{M}$ *is a well-defined map. The linearity of* Ψ *follows from*

$$\begin{aligned} \Psi([c_1]_p + [c_2]_p)(f) &= (f \circ \varphi^{-1} \circ (\varphi \circ c_1 + \varphi \circ c_2))'(0) \\ &= \partial_{x^i|_p}(f)\pi^i\left((\varphi \circ c_1 + \varphi \circ c_2)'(0)\right) \\ &= \partial_{x^i|_p}(f)\pi^i\left((\varphi \circ c_1)'(0)\right) + \partial_{x^i|_p}(f)\pi^i\left((\varphi \circ c_2)'(0)\right) \\ &= \Psi([c_1]_p)(f) + \Psi([c_2]_p)(f) \end{aligned}$$

and

$$\begin{aligned} \Psi(\lambda[c]_p)(f) &= (f \circ \varphi^{-1} \circ (\lambda\varphi \circ c))'(0) \\ &= \partial_{x^i|_p}(f)\pi^i\left((\lambda\varphi \circ c)'(0)\right) \\ &= \partial_{x^i|_p}(f)\lambda\pi^i\left((\varphi \circ c)'(0)\right) \\ &= \lambda\Psi([c]_p)(f), \end{aligned}$$

where we used the linearity of π^i *and of the derivatives for a function from* $\mathbb{R}$ *to* $\mathbb{R}^n$. *It remains to prove that* Ψ *is an isomorphism. A previous computation shows that*

$$\Psi([c]_p) = \pi^i\left((\varphi \circ c)'(0)\right)\partial_{x^i|_p}.$$

If $j = 1, \dots, n$ *and* $c_j(t) := \varphi^{-1}(0, \dots, 0, t, 0, \dots, 0)$ *(where the* t *is in the* j*-th slot) then* $\pi^i\left((\varphi \circ c_j)'(0)\right) = \delta^i_j$ *and* $\Psi([c_j]_p) = \partial_{x^j|_p}$. *Since* $\left(\partial_{x^j|_p}\right)_{j=1,\dots,n}$ *is a basis of* $T_p\mathcal{M}$ *and* Ψ *is linear this shows that* Ψ *is surjective. Moreover, if* $\Psi([c]_p) = 0$, *then* $(\varphi \circ c)'(0) = 0$. *However, if we denote* c_0 *the constant path equal to* p, *then* $(\varphi \circ c_0)'(0) = 0$ *and thus* $c \sim c_0$ *i.e* $[c]_p = [c_0]_p = 0_{\widetilde{T}_p\mathcal{M}}$. *This shows that* Ψ *is injective.*

Exercise 3.4.

1. *We use the definition of the Lie bracket of two vector fields as the following derivation of smooth functions:*

$$[X,Y](f)(p) = X(Y(f))(p) - Y(X(f))(p).$$

The three identities then follow from straightforward computations.

2. *We prove that the two sides of the identity have the same action on any smooth function* f*. Using the Leibniz rule we have*

$$\begin{aligned}[fX, gY](h) &= fX(gY(h)) - gY(fX(h)) \\ &= fgX(Y(h)) + fX(g)Y(h) - gfY(X(h)) - gY(f)X(h) \\ &= fg[X,Y](h) + fX(g)Y(h) - gY(f)X(h).\end{aligned}$$

3. *We use the local expression of* X *and* Y*, the first identity of the first question and the second question:*

$$\begin{aligned}[X,Y] &= \left[\sum_{i=1}^{n} X^i \partial_{x^i}, \sum_{i=1}^{n} Y^j \partial_{x^j}\right] \\ &= \sum_{i,j=1,\ldots,n} [X^i \partial_{x^i}, Y^j \partial_{x^j}] \\ &= \sum_{i,j=1,\ldots,n} \left(X^i Y^j [\partial_{x^i}, \partial_{x^j}] + X^i \partial_{x^i}(Y^j)\partial_{x^j} - Y^j \partial_{x^j}(X^i)\partial_{x^i}\right).\end{aligned}$$

However, using the definition of the ∂_{x^i} *we have*

$$\begin{aligned}\partial_{x^i}(\partial_{x^j} f)(p) &= \partial_{x^i|_p}(\partial_{x^j} f) \\ &= \partial_i \left(\partial_{x^j} f \circ \varphi^{-1}\right)(\varphi(p)) \\ &= \partial_i \partial_j (f \circ \varphi^{-1})(\varphi(p))\end{aligned}$$

so that the standard formula $\partial_i \partial_j = \partial_j \partial_i$ *in* $\mathbb{R}^n$ *implies* $[\partial_{x^i}, \partial_{x^j}] = 0$ *on* $\mathcal{M}$*. Therefore, we obtain*

$$[X,Y]^j = X(Y^j) - Y(X^j).$$

Exercise 3.5.

1. *Since* $\mathbf{d}\omega(X,Y) = -\mathbf{d}\omega(Y,X)$ *we only need to check* $C^\infty(\mathcal{M})$*-linearity with respect to the first argument. Let* $f \in C^\infty(\mathcal{M})$*, using* $\omega(fX) = f\omega(X)$ *and* $[fX,Y] = f[X,Y] - Y(f)X$ *(see the previous exercise) we obtain*

$$\begin{aligned}\mathbf{d}\omega(fX,Y) &= fX(\omega(Y)) - Y(\omega(fX)) - \omega([fX,Y]) \\ &= fX(\omega(Y)) - Y(f\omega(X)) - \omega(f[X,Y] - Y(f)X) \\ &= fX(\omega(Y)) - fY(\omega(X)) - Y(f)\omega(X) - f\omega(f[X,Y]) + Y(f)\omega(X) \\ &= f\mathbf{d}\omega(X,Y).\end{aligned}$$

2. *According to the previous question we only need to check the* $C^\infty(\mathcal{M})$*-linearity with respect to the 1-form. Using the definition of the differential of a smooth function* $\mathrm{d}f(X) = X(f)$ *we find*

$$\begin{aligned}\mathbf{d}(f\omega)(X,Y) &= X(f\omega(Y)) - Y(f\omega(X)) - f\omega([X,Y]) \\ &= \mathrm{d}f(X)\omega(Y) - \omega(X)\mathrm{d}f(Y) + f\mathbf{d}\omega(X,Y)\end{aligned}$$

so that $\mathbf{d}(f\omega) = \mathrm{d}f \otimes \omega - \omega \otimes \mathrm{d}f + f\mathbf{d}\omega$ *and* $\mathbf{d} : (\omega, X, Y) \in \Lambda^1(\mathcal{M}) \times (\Gamma(\mathcal{M}))^2 \longmapsto \mathbf{d}\omega(X,Y)$ *is not* $C^\infty(\mathcal{M})$*-linear with respect to its first argument and thus does not define a tensor.*

3. *Using* $\mathrm{d}f(X) = X(f)$ *we obtain*

$$\begin{aligned}\mathbf{d}\mathrm{d}f(X,Y) &= X(\mathrm{d}f(Y)) - Y(\mathrm{d}f(X)) - \mathrm{d}f([X,Y]) \\ &= X(Y(f)) - Y(X(f)) - [X,Y](f) \\ &= 0,\end{aligned}$$

where we used the definition of the Lie bracket. This shows that $\mathbf{d}\mathrm{d}f = 0$.

Exercise 3.6.

1. *For the coordinate vector fields, we compute their action on any* $f \in C^\infty(\mathcal{M})$. *If* $p \in \mathcal{M}$ *we have*

$$\begin{aligned}\partial_{x^i}(f)(p) &= \partial_{x^i|_p}(f) \\ &= \partial_i(f \circ \varphi^{-1})(\varphi(p)) \\ &= \partial_i(f \circ \psi^{-1} \circ \psi \circ \varphi^{-1})(\varphi(p)).\end{aligned}$$

Using the chain rule in $\mathbb{R}^n$ *and Einstein's summation convention we obtain*

$$\begin{aligned}\partial_{x^i}(f)(p) &= \sum_{\ell=1}^{n} \partial_i \left(\pi^\ell \circ \psi \circ \varphi^{-1}\right)(\varphi(p)) \partial_\ell (f \circ \psi^{-1})(\psi \circ \varphi^{-1}(\varphi(p))) \\ &= \partial_i \left(y^\ell \circ \varphi^{-1}\right)(\varphi(p)) \partial_\ell (f \circ \psi^{-1})(\psi(p)) \\ &= \partial_i \left(y^\ell \circ \varphi^{-1}\right)(\varphi(p)) \partial_{y^\ell}(f)(p)\end{aligned}$$

so that

$$\partial_{x^i} = \left(\partial_i \left(y^\ell \circ \varphi^{-1}\right) \circ \varphi\right) \partial_{y^\ell}. \tag{3.1}$$

For coordinate 1-forms, we start with the local expression of any 1-form applied to $\mathrm{d}x^i$*:*

$$\mathrm{d}x^i = \mathrm{d}x^i\left(\partial_{y^k}\right)\mathrm{d}y^k.$$

Using the symmetric version of (3.1) *we compute*

$$\begin{aligned}\mathrm{d}x^i\left(\partial_{y^k}\right) &= \left(\partial_k \left(x^\ell \circ \psi^{-1}\right) \circ \psi\right) \mathrm{d}x^i\left(\partial_{x^\ell}\right) \\ &= \left(\partial_k \left(x^\ell \circ \psi^{-1}\right) \circ \psi\right) \delta^i_\ell \\ &= \partial_k \left(x^i \circ \psi^{-1}\right) \circ \psi\end{aligned}$$

so that

$$\mathrm{d}x^i = \left(\partial_k \left(x^i \circ \psi^{-1}\right) \circ \psi\right) \mathrm{d}y^k. \tag{3.2}$$

2. *We use the following notations for the components of* T *in the two coordinate systems* $(x^i)_{i=1,\dots,n}$ *and* $(y^i)_{i=1,\dots,n}$*:*

$$T^{i_1\cdots i_r}_{j_1\cdots j_s} = T\left(\mathrm{d}x^{i_1}, \dots, \mathrm{d}x^{i_r}, \partial_{x^{j_1}}, \dots, \partial_{x^{j_s}}\right), \quad T^{\tilde{i}_1\cdots\tilde{i}_r}_{\tilde{j}_1\cdots\tilde{j}_s} = T\left(\mathrm{d}y^{\tilde{i}_1}, \dots, \mathrm{d}y^{\tilde{i}_r}, \partial_{y^{\tilde{j}_1}}, \dots, \partial_{y^{\tilde{j}_s}}\right).$$

Now using (3.1) *and* (3.2) *and the* $C^\infty(\mathcal{M})$*-multilinearity of tensors we obtain*

$$\begin{aligned}T^{i_1\cdots i_r}_{j_1\cdots j_s} &= \left(\partial_{\tilde{i}_1}\left(x^{i_1} \circ \psi^{-1}\right) \circ \psi\right) \times \cdots \times \left(\partial_{\tilde{i}_r}\left(x^{i_r} \circ \psi^{-1}\right) \circ \psi\right) \\ &\quad \times \left(\partial_{j_1}\left(y^{\tilde{j}_1} \circ \varphi^{-1}\right) \circ \varphi\right) \times \cdots \times \left(\partial_{j_s}\left(y^{\tilde{j}_s} \circ \varphi^{-1}\right) \circ \varphi\right) \\ &\quad \times T^{\tilde{i}_1\cdots\tilde{i}_r}_{\tilde{j}_1\cdots\tilde{j}_s}.\end{aligned}$$

In particular, T *vanishes in the coordinate system* $(x^i)_{i=1,\dots,n}$ *(i.e* $T^{i_1\cdots i_r}_{j_1\cdots j_s} = 0$ *for all choices of indices) if and only if it vanishes in the coordinate system* $(y^i)_{i=1,\dots,n}$.

Exercise 3.7. *Let* $(v_1,\dots,v_n)$ *be a basis of* $T_p\mathcal{M}$ *and* $(v_1^*,\dots,v_n^*)$ *be the dual basis of* $T_p^*\mathcal{M}$*, i.e the only basis of* $T_p^*\mathcal{M}$ *satisfying* $v_i^*(v_j)=\delta_{ij}$*. We define* $\tilde{A}_p$ *to be the only endormorphism of* $T_p\mathcal{M}$ *such that*

$$\tilde{A}_p(v_i)=\sum_{j=1}^n A_p(v_j^*,v_i)v_j.$$

In the basis $(v_1,\dots,v_n)$*, the matrix representing* $\tilde{A}_p$ *is* $(A_p(v_i^*,v_j))_{1\le i,j\le n}$ *and thus the trace of* $\tilde{A}_p$ *is* $\sum_{i=1}^n A_p(v_i^*,v_i)$ *which indeed matches* $C_1^1(A)(p)$ *and does not depend on the basis of* $T_p\mathcal{M}$*.*

Exercise 3.8.

1. *According to the lecture notes, it suffices to show that* $\mathcal{L}_X$ *is a derivation on vector fields. Using the second question of Exercise 1.4 we obtain*

$$\begin{aligned}\mathcal{L}_X(fY)&=[X,fY]\\&=f[X,Y]+X(f)Y\\&=(\mathcal{L}_Xf)Y+f\mathcal{L}_XY.\end{aligned}$$

2. *The set of tensor derivations is a vector space so that* $a\mathcal{L}_X+\mathcal{L}_Y$ *is again a tensor derivation. Therefore we are asked to prove that two tensor derivations are equal, and thanks to Lemma 1.5 in the lecture notes it is enough to check that they coincide on functions and vector fields, which is completely obvious. Similarly, one can show that the Lie bracket of two tensor derivations is again a tensor derivation (as we did for derivations of functions when we defined the Lie bracket of two vector fields). We have* $\mathcal{L}_{[X,Y]}f=[X,Y]f$ *and* $[\mathcal{L}_X,\mathcal{L}_Y]f=X(Y(f))-Y(X(f))$ *so* $\mathcal{L}_{[X,Y]}$ *and* $[\mathcal{L}_X,\mathcal{L}_Y]$ *coincide on functions by definition of the Lie bracket. Moreover we have* $\mathcal{L}_{[X,Y]}(Z)=[[X,Y],Z]$ *and*

$$[\mathcal{L}_X,\mathcal{L}_Y](Z)=[X,[Y,Z]]-[Y,[X,Z]]=[[X,Y],Z]$$

where we used the Jacobi identity for the Lie bracket. Therefore $\mathcal{L}_{[X,Y]}$ *and* $[\mathcal{L}_X,\mathcal{L}_Y]$ *coincide on vector fields and are thus equal thanks to Lemma 1.5.*

3. *Let* $Y\in\Gamma(\mathcal{M})$*, we compute the two sides. By definition of the differential of the smooth function* $\mathcal{L}_Xf=X(f)$ *we have*

$$(\mathrm{d}\mathcal{L}_Xf)(Y)=Y(\mathcal{L}_Xf)=Y(X(f)).$$

By definition of a tensor derivation we find

$$(\mathcal{L}_X\mathrm{d}f)(Y)=X(Y(f))-\mathrm{d}f([X,Y])=X(Y(f))-[X,Y](f).$$

Finally, by definition of the Lie bracket we get

$$(\mathcal{L}_X\mathrm{d}f)(Y)=X(Y(f))-X(Y(f))+Y(X(f))=(\mathrm{d}\mathcal{L}_Xf)(Y).$$

4. *Let* $W,Z\in\Gamma(\mathcal{M})$*, we compute*

$$\begin{aligned}\mathbf{d}\mathcal{L}_X\omega(W,Z)&=W\left(\mathcal{L}_X\omega(Z)\right)-Z\left(\mathcal{L}_X\omega(W)\right)-\mathcal{L}_X\omega([W,Z])\\&=W\left(X(\omega(Z))\right)-Z\left(X(\omega(W))\right)-W\left(\omega([X,Z])\right)+Z\left(\omega([X,W])\right)\\&\quad-X(\omega([W,Z]))+\omega([X,[W,Z]]),\end{aligned}$$

where we used the expression of $\mathcal{L}_X\omega(Z)=X(\omega(Z))-\omega([X,Z])$ *(recall the definition of the Lie derivative on functions and vector fields). We commute* W *and* X *and* Z *and* X *in the first two terms:*

$$\begin{aligned}\mathbf{d}\mathcal{L}_X\omega(W,Z)&=X\left(W(\omega(Z))\right)+[W,X](\omega(Z))-X\left(Z(\omega(W))\right)+[X,Z](\omega(W))\\&\quad-W\left(\omega([X,Z])\right)+Z\left(\omega([X,W])\right)\\&\quad-X(\omega([W,Z]))+\omega([X,[W,Z]])\\&=X\left(\mathbf{d}\omega(W,Z)\right)-[X,W](\omega(Z))+[X,Z](\omega(W))\\&\quad-W\left(\omega([X,Z])\right)+Z\left(\omega([X,W])\right)\\&\quad-\omega([W,[Z,X]])-\omega([Z,[X,W]])\end{aligned}$$

where we recognized the expression of $\mathbf{d}\omega(W,Z)$ *and also used the Jacobi identity for the Lie bracket (first question of Exercise 1.4). Using again the definition of* $\mathbf{d}$ *we get*

$$\mathbf{d}\mathcal{L}_X\omega(W,Z) = X\left(\mathbf{d}\omega(W,Z)\right) - \mathbf{d}\omega([X,W],Z) - \mathbf{d}\omega(W,[X,Z]) = \mathcal{L}_X\mathbf{d}\omega(W,Z).$$

Exercise 3.9.

1. *If such* $\mathcal{D}_B$ *exists, let us show that* $\mathcal{D}_B(f) = 0$ *for all* $f \in C^\infty(\mathcal{M})$*. Let* $f \in C^\infty(\mathcal{M})$*,* $X \in \Gamma(\mathcal{M})$*, the Leibniz rule gives on the one hand* $(\mathcal{D}_B)(fX) = f(\mathcal{D}_B)(X) + \mathcal{D}_B(f)X$*. On the other hand the condition* $\mathcal{D}_{B|_{\Gamma(\mathcal{M})}} = (\mathcal{D}_B)^1_0$ *implies that* $\mathcal{D}_{B|_{\Gamma(\mathcal{M})}}$ *is* $C^\infty(\mathcal{M})$*-linear so that* $(\mathcal{D}_B)(fX) = f(\mathcal{D}_B)(X)$*. Therefore we have* $\mathcal{D}_B(f)X = 0$ *for all* f *and* X*, which shows that* $\mathcal{D}_{B|_{C^\infty(\mathcal{M})}} = 0$*. Lemma 1.5 of the notes then shows the existence and uniqueness of* $\mathcal{D}_B \in \mathfrak{D}(\mathcal{M})$ *such that* $\mathcal{D}_{B|_{C^\infty(\mathcal{M})}} = 0$ *and* $\mathcal{D}_{B|_{\Gamma(\mathcal{M})}} = (\mathcal{D}_B)^1_0$*.*

2. *We need to show that the map* $(X,B) \in \Gamma(\mathcal{M}) \times \mathcal{T}^1_1(\mathcal{M}) \longmapsto \mathcal{L}_X + \mathcal{D}_B \in \mathfrak{D}(\mathcal{M})$ *is a linear isomorphism:*

   - *The linearity follows from the linearity of* $X \longmapsto \mathcal{L}_X$ *and* $B \longmapsto \mathcal{D}_B$*.*
   - *Assume that* $(X,B) \in \Gamma(\mathcal{M}) \times \mathcal{T}^1_1(\mathcal{M})$ *is such that* $\mathcal{L}_X + \mathcal{D}_B$ *is the zero tensor derivation. Since* $\mathcal{D}_{B|_{C^\infty(\mathcal{M})}} = 0$ *we obtain* $X(f) = 0$ *for all* $f \in C^\infty(\mathcal{M})$ *which shows that* $X = 0$ *and* $\mathcal{D}_B = 0$*. Since* $\mathcal{D}_{B|_{\Gamma(\mathcal{M})}} = (\mathcal{D}_B)^1_0$ *this implies* $(\mathcal{D}_B)^1_0 = 0$ *and thus* $B = 0$*. This shows that the map* $(X,B) \longmapsto \mathcal{L}_X + \mathcal{D}_B$ *is injective.*
   - *Let* $\mathcal{D} \in \mathfrak{D}(\mathcal{M})$*. Since derivations on functions are the same as vector fields there exists a unique* $X \in \Gamma(\mathcal{M})$ *such that* $\mathcal{D}(f) = X(f)$*. Set* $\tilde{\mathcal{D}} = \mathcal{D} - \mathcal{L}_X$*. We have* $\tilde{\mathcal{D}}_{|_{C^\infty(\mathcal{M})}} = 0$ *which implies that* $\tilde{\mathcal{D}}_{|_{\Gamma(\mathcal{M})}}$ *is* $C^\infty(\mathcal{M})$*-linear (thanks to the Leibniz rule for* $\tilde{\mathcal{D}}$*). For* $(\omega, Y) \in \Lambda^1(\mathcal{M}) \times \Gamma(\mathcal{M})$ *we define* $B(\omega,Y) = \tilde{\mathcal{D}}_{|_{\Gamma(\mathcal{M})}}(Y)(\omega)$ *which defines a* $(1,1)$*-tensor since it is indeed* $C^\infty(\mathcal{M})$*-multilinear. By construction we have* $\tilde{\mathcal{D}}_{|_{\Gamma(\mathcal{M})}} = (\mathcal{D}_B)^1_0$ *so that the uniqueness part of the previous question implies* $\tilde{\mathcal{D}} = \mathcal{D}_B$ *and* $\mathcal{D} = \mathcal{L}_X + \mathcal{D}_B$*. This shows that the map* $(X,B) \longmapsto \mathcal{L}_X + \mathcal{D}_B$ *is surjective.*

We see in fact that the Lie derivative is quite superfluous here and we could get the same result with say the covariant derivative $\mathbf{D}_X$ *from Chapter 2.*

3.2 Exercises from Chapter 2

Exercise 3.10. *The spherical coordinates are* (r,θ,φ) *such that the Euclidean coordinates* (x^1,x^2,x^3) *are given by*

$$x^1 = r\sin\theta\cos\phi, \quad x^2 = r\sin\theta\sin\phi, \quad x^3 = r\cos\theta.$$

We want to compute $\mathbf{g}(\partial_r,\partial_r)$*,* $\mathbf{g}(\partial_r,\partial_\theta)$ *etc. (where we denote* $\mathbf{g}_{eucl}$ *simply by* $\mathbf{g}$*), therefore we would like to express the coordinate vector fields of* (r,θ,ϕ) *in terms of the coordinate vector fields of* (x^1,x^2,x^3)*. For this we use the formula* (3.1) *of Exercise 1.6:*

$$\partial_{y^i} = \left(\partial_i\left(x^\ell \circ \psi^{-1}\right) \circ \psi\right)\partial_{x^\ell}.$$

where $(y^i)_{i=1,\dots,n}$ *are associated to* ψ*. The definition of the spherical coordinates is*

$$x^1 \circ \psi^{-1}(r,\theta,\phi) = r\sin\theta\cos\phi, \quad x^2 \circ \psi^{-1}(r,\theta,\phi) = r\sin\theta\sin\phi, \quad x^3 \circ \psi^{-1}(r,\theta,\phi) = r\cos\theta,$$

which is equivalent to saying that $\psi(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta) = (r,\theta,\phi)$*. Therefore we have*

$$\begin{aligned}\partial_r &= \partial_r\left(r\sin\theta\cos\phi\right)\partial_{x^1} + \partial_r\left(r\sin\theta\sin\phi\right)\partial_{x^2} + \partial_r\left(r\cos\theta\right)\partial_{x^3}\\ &= \sin\theta\cos\phi\partial_{x^1} + \sin\theta\sin\phi\partial_{x^2} + \cos\theta\partial_{x^3},\end{aligned}$$

$$
\begin{aligned}
\partial_\theta &= \partial_\theta\left(r\sin\theta\cos\phi\right)\partial_{x^1} + \partial_\theta\left(r\sin\theta\sin\phi\right)\partial_{x^2} + \partial_\theta\left(r\cos\theta\right)\partial_{x^3} \\
&= r\cos\theta\cos\phi\partial_{x^1} + r\cos\theta\sin\phi\partial_{x^2} - r\sin\theta\partial_{x^3},
\end{aligned}
$$

$$
\begin{aligned}
\partial_\phi &= \partial_\phi\left(r\sin\theta\cos\phi\right)\partial_{x^1} + \partial_\phi\left(r\sin\theta\sin\phi\right)\partial_{x^2} + \partial_\phi\left(r\cos\theta\right)\partial_{x^3} \\
&= -r\sin\theta\sin\phi\partial_{x^1} + r\sin\theta\cos\phi\partial_{x^2}.
\end{aligned}
$$

Therefore, using $\mathbf{g}(\partial_{x^i}, \partial_{x^j}) = \delta_{ij}$ *we get for instance*

$$
\begin{aligned}
\mathbf{g}_{rr} &= \mathbf{g}(\partial_r, \partial_r) \\
&= \sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi + \cos^2\theta \\
&= 1.
\end{aligned}
$$

The other non-zero components are $\mathbf{g}_{\theta\theta} = r^2$ *and* $\mathbf{g}_{\phi\phi} = r^2\sin^2\theta$*. Moreover, one can check that the non-zero Christoffel symbols are*

$$
\begin{aligned}
&\Gamma^r_{\theta\theta} = -r, && \Gamma^r_{\phi\phi} = -r\sin^2\theta, && \Gamma^\theta_{r\theta} = \frac{1}{r}, \\
&\Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta, && \Gamma^\phi_{r\phi} = \frac{1}{r}, && \Gamma^\phi_{\theta\phi} = \frac{\cos\theta}{\sin\theta}.
\end{aligned}
$$

Exercise 3.11.

1. *The* $C^\infty(\mathcal{M})$*-linearity with respect to the 1-form is obvious, so we only need to check the* $C^\infty(\mathcal{M})$*-multilinearity of* $(X,Y) \longmapsto D_XY - D_YX - [X,Y]$*. Let* $f \in C^\infty(\mathcal{M})$*, using the Leibniz rule for* D *and a property of the Lie bracket we obtain*

$$
\begin{aligned}
D_X(fY) - D_{fY}X - [X, fY] &= X(f)Y + fD_XY - fD_YX - f[X,Y] - X(f)Y \\
&= f\left(D_XY - D_YX - [X,Y]\right).
\end{aligned}
$$

   Since $(X,Y) \longmapsto D_XY - D_YX - [X,Y]$ *is antisymmetric this also proves that* $D_{fX}Y - D_Y(fX) - [fX,Y] = f\left(D_XY - D_YX - [X,Y]\right)$*.*

2. *The* $C^\infty(\mathcal{M})$*-linearity with respect to the 1-form and the first vector field is obvious, so we only need to check the* $C^\infty(\mathcal{M})$*-linearity of* $Y \longmapsto D_XY - \tilde{D}_XY$ *for* X *fixed. Let* $f \in C^\infty(\mathcal{M})$*, using the Leibniz rule for* D *and* $\tilde{D}$ *we obtain:*

$$
D_X(fY) - \tilde{D}_X(fY) = X(f)Y + fD_XY - X(f)Y - f\tilde{D}_XY = f\left(D_XY - \tilde{D}_XY\right).
$$

Exercise 3.12. *Let* $(x^\alpha)_\alpha$ *be a local coordinate system. If we denote by* $\tilde{\Gamma}$ *the Christoffel symbols of* $\tilde{\mathbf{g}}$ *we have*

$$
\begin{aligned}
\tilde{\Gamma}^\alpha_{\mu\nu} &= \frac{1}{2}\tilde{\mathbf{g}}^{\alpha\beta}\left(\partial_\mu\tilde{\mathbf{g}}_{\nu\beta} + \partial_\nu\tilde{\mathbf{g}}_{\mu\beta} - \partial_\beta\tilde{\mathbf{g}}_{\mu\nu}\right) \\
&= \frac{1}{2}e^{-2f}\mathbf{g}^{\alpha\beta}e^{2f}\left(\partial_\mu\mathbf{g}_{\nu\beta} + \partial_\nu\mathbf{g}_{\mu\beta} - \partial_\beta\mathbf{g}_{\mu\nu}\right) + e^{-2f}\mathbf{g}^{\alpha\beta}e^{2f}\left((\partial_\mu f)\mathbf{g}_{\nu\beta} + (\partial_\nu f)\mathbf{g}_{\mu\beta} - (\partial_\beta f)\mathbf{g}_{\mu\nu}\right) \\
&= \Gamma^\alpha_{\mu\nu} + \delta^\alpha_\nu\partial_\mu f + \delta^\alpha_\mu\partial_\nu f - \mathbf{g}_{\mu\nu}\mathbf{g}^{\alpha\beta}\partial_\beta f.
\end{aligned}
$$

Therefore if $X, Y \in \Gamma(\mathcal{M})$ *we have*

$$
\begin{aligned}
\tilde{\mathbf{D}}_XY^\alpha &= X(Y^\alpha) + X^\mu Y^\nu\tilde{\Gamma}^\alpha_{\mu\nu} \\
&= \mathbf{D}_XY^\alpha + X(f)Y^\alpha + Y(f)X^\alpha - \mathbf{g}(X,Y)\mathbf{g}^{\alpha\beta}\partial_\beta f.
\end{aligned}
$$

This shows that $\tilde{\mathbf{D}}_XY = \mathbf{D}_XY + X(f)Y + Y(f)X - \mathbf{g}(X,Y)\mathrm{grad}_{\mathbf{g}}f$*.*

Exercise 3.13. *For the second part of the lemma we start with the formula for* $\mathbf{D}_X\mathbf{g}(Y,Z)$ *obtained with Proposition 1.3:*

$$\begin{aligned}\mathbf{D}_X\mathbf{g}(Y,Z) &= \mathbf{D}_X\left(\mathbf{g}(Y,Z)\right) - \mathbf{g}\left(\mathbf{D}_X(Y),Z\right) - \mathbf{g}\left(Y,\mathbf{D}_X(Z)\right)\\ &= X\left(\mathbf{g}(Y,Z)\right) - \mathbf{g}\left(\mathbf{D}_XY,Z\right) - \mathbf{g}\left(Y,\mathbf{D}_XZ\right)\end{aligned}$$

where we have used the definition of the tensor derivation $\mathbf{D}_X$*. This is precisely identically zero thanks to the compatibility of the Levi-Civita connection with* $\mathbf{g}$*. Now let us prove that* $\mathbf{D}_X\mathbf{g}^{-1}=0$*. We have* $C_1^1(\mathbf{g}^{-1}\otimes\mathbf{g}) = \mathrm{Id}$ *and one can show that* $\mathbf{D}_X\mathrm{Id}=0$ *(actually this is true for any tensor derivation) so that the commutation with standard contraction, the Leibniz rule and the fact that* $\mathbf{D}_X\mathbf{g}=0$ *gives* $C_1^1(\mathbf{D}_X\mathbf{g}^{-1}\otimes\mathbf{g})=0$*. The non-degeneracy of* $\mathbf{g}$ *then implies that* $\mathbf{D}_X\mathbf{g}^{-1}=0$*. Finally, rewriting the musical isomorphisms with contractions and tensor product we obtain*

$$\begin{aligned}\mathbf{D}_XY^\flat &= \mathbf{D}_X\left(C_1^1(Y\otimes\mathbf{g})\right)\\ &= C_1^1\left(\mathbf{D}_X(Y\otimes\mathbf{g})\right)\\ &= C_1^1\left(\mathbf{D}_XY\otimes\mathbf{g}\right)\\ &= \left(\mathbf{D}_XY\right)^\flat,\end{aligned}$$

where we used the commutation of any tensor derivation with contractions and the fact that $\mathbf{D}_X\mathbf{g}=0$*. An almost identical computation shows that* $\mathbf{D}_X\omega^\# = (\mathbf{D}_X\omega)^\#$ *(using* $\mathbf{D}_X\mathbf{g}^{-1}=0$ *this time).*

Exercise 3.14. *By definition of a tensor derivation we have*

$$\mathrm{Hess}(f)(X,Y) = X(\mathrm{d}f(Y)) - \mathrm{d}f(\mathbf{D}_XY) = X(Y(f)) - \mathbf{D}_XY(f),$$

where we also used the definition of the differential. Therefore

$$\begin{aligned}\mathrm{Hess}(f)(X,Y) - \mathrm{Hess}(f)(Y,X) &= X(Y(f)) - \mathbf{D}_XY(f) - Y(X(f)) + \mathbf{D}_YX(f)\\ &= \left([X,Y] - \mathbf{D}_XY + \mathbf{D}_YX\right)(f)\\ &= 0,\end{aligned}$$

where we used the torsion free property of the metric.

Exercise 3.15.

1. *By definition of* $\mathbf{d}$ *and of the musical isomorphisms we have*

$$\begin{aligned}\mathbf{d}V^\flat(X,Y) &= X(V^\flat(Y)) - Y(V^\flat(X)) - V^\flat([X,Y])\\ &= X(\mathbf{g}(V,Y)) - Y(\mathbf{g}(V,X)) - \mathbf{g}(V,[X,Y]).\end{aligned}$$

Using now the compatibility if $\mathbf{D}$ *with* $\mathbf{g}$ *and its torsion free property we get*

$$\begin{aligned}\mathbf{d}V^\flat(X,Y) &= \mathbf{g}(\mathbf{D}_XV,Y) - \mathbf{g}(\mathbf{D}_YV,X) + \mathbf{g}(V,\mathbf{D}_XY - \mathbf{D}_YX - [X,Y])\\ &= \mathbf{g}(\mathbf{D}_XV,Y) - \mathbf{g}(\mathbf{D}_YV,X).\end{aligned}$$

2. *Thanks to the first question we have* $\mathrm{curl}(\mathrm{grad}f) = \mathbf{d}(\mathrm{grad}f)^\flat$*. Thanks to the definition of the gradient we get* $\mathrm{curl}(\mathrm{grad}f) = \mathbf{d}\left((\mathrm{d}f)^\#\right)^\flat$*. Thanks to the properties of the musical isomorphisms we get* $\mathrm{curl}(\mathrm{grad}f) = \mathbf{d}\mathrm{d}f = 0$ *where we also used a result from Exercise 1.5.*

Exercise 3.16.

1. *By definition of the divergence of the 1-form* $X^\flat$ *we have in coordinates*

$$\begin{aligned}\mathrm{div}X &= \mathbf{g}^{\alpha\beta}\mathbf{D}_\alpha X^\flat(\partial_\beta)\\ &= \mathbf{g}^{\alpha\beta}\left(\partial_\alpha(X^\flat(\partial_\beta)) - X^\flat(\mathbf{D}_\alpha\partial_\beta)\right)\\ &= \mathbf{g}^{\alpha\beta}\left(\partial_\alpha X_\beta - \Gamma^\mu_{\alpha\beta}X_\mu\right),\end{aligned}$$

where we use the standard shortcut $X_\alpha = (X^\flat)_\alpha$*.*

2. *If* $X = \mathrm{grad} f$ *then by definition of the gradient we have* $X^\flat = \mathrm{d}f$ *and* $X_\alpha = \partial_\alpha f$ *so that* $\mathrm{div}(\mathrm{grad} f) = \mathbf{g}^{\alpha\beta}\left(\partial_\alpha\partial_\beta f - \Gamma^\mu_{\alpha\beta}\partial_\mu f\right) = \Box_{\mathbf{g}} f$.

3. *We have*

$$\frac{1}{\sqrt{-\det \mathbf{g}}}\partial_\beta\left(\sqrt{-\det \mathbf{g}}X^\beta\right) = \partial_\beta X^\beta + \frac{X^\beta}{\sqrt{-\det \mathbf{g}}}\partial_\beta\left(\sqrt{-\det \mathbf{g}}\right)$$
$$= \partial_\beta X^\beta + \frac{X^\beta}{2\det \mathbf{g}}\partial_\beta \det \mathbf{g}.$$

Using the Jacobi formula for the differential of the determinant we obtain $\partial_\beta \det \mathbf{g} = \det \mathbf{g}\mathrm{tr}\left(\mathbf{g}^{-1}\partial_\beta \mathbf{g}\right)$ *where here* $\mathbf{g}$ *denotes its matrix representation and* $\mathbf{g}^{-1}$ *its inverse. Therefore* $\partial_\beta \det \mathbf{g} = \det \mathbf{g}\mathbf{g}^{\mu\nu}\partial_\beta \mathbf{g}_{\mu\nu}$ *and*

$$\frac{1}{\sqrt{-\det \mathbf{g}}}\partial_\beta\left(\sqrt{-\det \mathbf{g}}X^\beta\right) = \partial_\beta X^\beta + \frac{X^\beta}{2}\mathbf{g}^{\mu\nu}\partial_\beta \mathbf{g}_{\mu\nu}$$
$$= \mathbf{g}^{\alpha\beta}\partial_\beta X_\alpha + X_\rho\left(\partial_\beta \mathbf{g}^{\rho\beta} + \frac{1}{2}\mathbf{g}^{\beta\rho}\mathbf{g}^{\mu\nu}\partial_\beta \mathbf{g}_{\mu\nu}\right),$$

where we used $X^\beta = \mathbf{g}^{\beta\rho}X_\rho$. *Thanks to the expression of the Christoffel symbols we have*

$$\partial_\beta \mathbf{g}^{\rho\beta} + \frac{1}{2}\mathbf{g}^{\beta\rho}\mathbf{g}^{\mu\nu}\partial_\beta \mathbf{g}_{\mu\nu} = \frac{1}{2}\mathbf{g}^{\rho\sigma}\left(-2\mathbf{g}^{\beta\alpha}\partial_\beta \mathbf{g}_{\alpha\sigma} + \mathbf{g}^{\mu\nu}\partial_\sigma \mathbf{g}_{\mu\nu}\right) = -\mathbf{g}^{\mu\nu}\Gamma^\rho_{\mu\nu},$$

so that we obtain

$$\frac{1}{\sqrt{-\det \mathbf{g}}}\partial_\beta\left(\sqrt{-\det \mathbf{g}}X^\beta\right) = \mathbf{g}^{\alpha\beta}\partial_\beta X_\alpha - \mathbf{g}^{\mu\nu}\Gamma^\rho_{\mu\nu}X_\rho = \mathrm{div}X.$$

Exercise 3.17.

1. *Since* $\mathrm{div}(f\mathbf{g}) = \mathrm{d}f$ *we need to prove that* $\mathrm{div}(\mathrm{d}\varphi\otimes\mathrm{d}\varphi) = \frac{1}{2}\mathrm{d}\left(\mathbf{g}(\mathrm{grad}\varphi,\mathrm{grad}\varphi)\right)$. *For the LHS we find*

$$\mathbf{D}(\mathrm{d}\varphi\otimes\mathrm{d}\varphi)(X,Y,Z) = \mathbf{D}_X(\mathrm{d}\varphi\otimes\mathrm{d}\varphi)(Y,Z) = \mathrm{Hess}(\varphi)(X,Y)\mathrm{d}\varphi(Z) + \mathrm{Hess}(\varphi)(X,Z)\mathrm{d}\varphi(Y)$$

so that

$$\mathrm{div}(\mathrm{d}\varphi\otimes\mathrm{d}\varphi)(Z) = C_{12}(\mathrm{Hess}(\varphi))\mathrm{d}\varphi(Z) + C_{12}(\mathrm{d}\varphi\otimes\mathrm{Hess}(\varphi))(Z) = C_{12}(\mathrm{d}\varphi\otimes\mathrm{Hess}(\varphi))(Z)$$

where we used $C_{12}(\mathrm{Hess}(\varphi)) = \Box_{\mathbf{g}}\varphi = 0$. *For the RHS we have*

$$\frac{1}{2}\mathrm{d}\left(\mathbf{g}(\mathrm{grad}\varphi,\mathrm{grad}\varphi)\right)(Z) = \frac{1}{2}Z(\mathbf{g}(\mathrm{grad}\varphi,\mathrm{grad}\varphi)) = \mathbf{g}(\mathbf{D}_Z\mathrm{grad}\varphi,\mathrm{grad}\varphi).$$

Using the musical isomorphisms we get $\mathbf{D}_Z\mathrm{grad}\varphi = \mathbf{D}_Z(\mathrm{d}\varphi)^\# = (\mathbf{D}_Z\mathrm{d}\varphi)^\#$ *and thus this rewrites*

$$\frac{1}{2}\mathrm{d}\left(\mathbf{g}(\mathrm{grad}\varphi,\mathrm{grad}\varphi)\right)(Z) = \mathbf{g}^{-1}(\mathbf{D}_Z\mathrm{d}\varphi,\mathrm{d}\varphi)$$
$$= \mathbf{g}^{\alpha\beta}\mathbf{D}_Z\mathrm{d}\varphi(\partial_\alpha)\mathrm{d}\varphi(\partial_\beta)$$
$$= C_{12}(\mathrm{d}\varphi\otimes\mathrm{Hess}(\varphi))(Z)$$

where we used the symmetry of the Hessian. This concludes the proof.

2. *We compute using the expression of the divergence of a symmetric* $(0,2)$*-tensor in coordinates and use several times* $\mathbf{D}_X\mathbf{g} = 0$ *and* $\mathbf{D}_X\mathbf{g}^{-1} = 0$ *and te Leibniz rule :*

$$\mathrm{div}T_\nu = \mathbf{g}^{\gamma\mu}\mathbf{D}_\gamma T_{\mu\nu}$$
$$= \mathbf{g}^{\gamma\mu}\mathbf{g}^{\alpha\beta}\mathbf{D}_\gamma\left(F_{\mu\alpha}F_{\nu\beta}\right) - \frac{1}{4}\mathbf{g}^{\alpha\rho}\mathbf{g}^{\beta\sigma}\mathbf{D}_\nu\left(F_{\alpha\beta}F_{\rho\sigma}\right)$$
$$= \mathbf{g}^{\gamma\mu}\mathbf{g}^{\alpha\beta}F_{\mu\alpha}\mathbf{D}_\gamma F_{\nu\beta} + \mathbf{g}^{\alpha\beta}F_{\nu\beta}\mathbf{g}^{\gamma\mu}\mathbf{D}_\gamma F_{\mu\alpha} - \frac{1}{4}\mathbf{g}^{\alpha\rho}\mathbf{g}^{\beta\sigma}F_{\alpha\beta}\mathbf{D}_\nu F_{\rho\sigma} - \frac{1}{4}\mathbf{g}^{\alpha\rho}\mathbf{g}^{\beta\sigma}F_{\rho\sigma}\mathbf{D}_\nu F_{\alpha\beta}.$$

The second term vanishes because of the second Maxwell equation, the third and fourth are the same so that

$$\begin{aligned}\mathrm{div}T_\nu &= \mathbf{g}^{\gamma\mu}\mathbf{g}^{\alpha\beta}F_{\mu\alpha}\mathbf{D}_\gamma F_{\nu\beta} - \frac{1}{2}\mathbf{g}^{\alpha\rho}\mathbf{g}^{\beta\sigma}F_{\alpha\beta}\mathbf{D}_\nu F_{\rho\sigma} \\ &= F^{\gamma\beta}\left(\mathbf{D}_\gamma F_{\nu\beta} + \frac{1}{2}\mathbf{D}_\nu F_{\beta\gamma}\right),\end{aligned}$$

where we also used the antisymmetry of F and defined $F^{\gamma\beta} := F_{\mu\alpha}\mathbf{g}^{\mu\gamma}\mathbf{g}^{\alpha\beta}$. Note that the antisymmetry of F implies $F^{\gamma\beta} = -F^{\beta\gamma}$. We now contract the first Maxwell equation with $F^{\gamma\beta}$:

$$F^{\gamma\beta}\mathbf{D}_\nu F_{\beta\gamma} + F^{\gamma\beta}\mathbf{D}_\beta F_{\gamma\nu} + F^{\gamma\beta}\mathbf{D}_\gamma F_{\nu\beta} = 0.$$

Thanks to the antisymmetry of $F_{\alpha\beta}$ and $F^{\alpha\beta}$ the middle term becomes $F^{\gamma\beta}\mathbf{D}_\beta F_{\gamma\nu} = F^{\beta\gamma}\mathbf{D}_\beta F_{\nu\gamma}$ which is the same as $F^{\gamma\beta}\mathbf{D}_\gamma F_{\nu\beta}$, i.e the third term. Therefore we have proved $F^{\gamma\beta}\mathbf{D}_\nu F_{\beta\gamma} + 2F^{\gamma\beta}\mathbf{D}_\gamma F_{\nu\beta} = 0$ which precisely implies $\mathrm{div}T = 0$.