We will use the following abbreviations: if $\mathbb{R}^n$ has coordinates x^i, we will use ∂_i to denote the differentiation with respect to x^i, i.e. $\partial_i = \frac{\partial}{\partial x^i}$. In addition, we will use Einstein's summation notation where the sum symbol is not written if two indices appear, one as a lower index and one as an upper index, e.g. the directional derivative is

$$\mathrm{d}_x\varphi(A) = \partial_i\varphi \cdot A^i.$$

In such a situation, the summation is implicit (and thus, it is necessary to state explicitly if the summation is not intended). We believe that this highly improves readability of formulas.

1. Tangent vectors, vector fields and derivations

1.1. Partitions of unity.

Definition 1. A *support* of a function $f\colon M \to \mathbb{R}$ is the closure

$$\operatorname{supp} f = \overline{\{x \in M \mid f(x) \neq 0\}}.$$

Definition 2. Let $\mathcal{U} = \{U_i \mid i \in \mathcal{I}\}$ be an open cover of a manifold M. A *(smooth) partition of unity subordinate to* $\mathcal{U}$ is a collection of functions $\lambda_i\colon M \to [0,1]$ such that $\operatorname{supp}\lambda_i \subseteq U_i$, such that in a neighbourhood of every point $x \in M$, there is only a finite number of non-zero λ_i and such that $\sum \lambda_i = 1$.

For simplicity, we will assume M to be compact, it is however not necessary. Then the local finiteness of the λ_i in the definition of the partition of a unity can be replaced by the finiteness, i.e. only a finite number of the λ_i is nonzero.

Theorem 3. *Let M be compact. There exist a partition of unity subordinate to any open cover of M.*

Proof. We choose a finite subcover $U_1, \ldots, U_n$. We will construct the functions λ_i inductively and denote $V_i = \{x \in M \mid \lambda_i(x) \neq 0\}$. It is enough to find the λ_i in such a way that $\bigcup V_i = M$, since then $\lambda = \sum \lambda_i > 0$ and we may replace each λ_i by λ_i/λ.

Thus, suppose that the λ_j have been constructed for $j < i$. Then let λ_i be a function that is supported in U_i and is positive on a compact set $C_i = M \smallsetminus (\bigcup_{j<i} V_j \cup \bigcup_{j>i} U_j)$. We may verify inductively that $\bigcup_{j<i} V_j \cup \bigcup_{j\geq i} U_j = M$ so that $C_i \subseteq U_i$ and such a function can be chosen by the following lemma. □

Lemma 4. *Let $C \subseteq U \subseteq M$ be subsets of a manifold M such that C is compact and U open. Then there exists a function $\lambda\colon M \to [0,1]$ with $\operatorname{supp}\lambda \subseteq U$ that is positive on C.*

Proof. For each $x \in C$, we will find a function $\lambda_x\colon M \to [0,1]$ that is positive at x and with $\operatorname{supp}\lambda_x \subseteq U$. Then choose a finite subcover of C by the open sets $V_x = \{y \in M \mid \lambda_x(y) \neq 0\}$, say $C \subseteq V_{x_1} \cup \cdots \cup V_{x_n}$. Then the required function is $\lambda = \lambda_{x_1} + \cdots + \lambda_{x_n}$.

It is enough to find the functions λ_x. Let $\varphi\colon V \to \mathbb{R}^n$ be a chart centred at x (i.e. $\varphi(x) = 0$) with $V \subseteq U$. Let $\rho_\varepsilon\colon \mathbb{R} \to [0,1]$ be the function from the next lemma for ε small enough that $[-\varepsilon,\varepsilon]^n \subseteq \varphi(V)$. Then we define

$$\lambda_x(y) = \begin{cases} \rho_\varepsilon(y^1)\cdots\rho_\varepsilon(y^n) & y \in V \\ 0 & y \notin V \end{cases}$$

where $y^1, \ldots, y^n$ are the coordinates of y in the given chart, i.e. $y^i = \varphi^i(y)$. Since λ_x is smooth on V and on $M \smallsetminus \varphi^{-1}([-\varepsilon,\varepsilon]^n)$ (where it is zero), it is smooth on M. □

Lemma 5. *There exists a function $\rho_\varepsilon\colon \mathbb{R} \to [0,1]$ such that $\rho_\varepsilon(0) > 0$ and $\rho_\varepsilon(x) = 0$ for $|x| \geq \varepsilon$.*

Proof. The function

$$\lambda(x) = \begin{cases} e^{-1/x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

is smooth (that needs to be verified, see Kolář's text). We set $\rho_\varepsilon(x) = \lambda(\varepsilon + x)\lambda(\varepsilon - x)$. □

1.2. **Germs.**

Definition 6. We say that two maps $f, g\colon M \to N$ *have the same germ at* x if they agree in some neighbourhood of x. We denote the class of f with respect to this relation as $\operatorname{germ}_x f$, the *germ* of f at x, the resulting decomposition is denoted $C^\infty_x(M,N)$; in the case of smooth functions on M, we will abbreviate it to $C^\infty_x M$.

We will need that germs of maps defined in a neighbourhood of x extend to germs of globally defined maps (this clearly does not hold for maps themselves – e.g. $1/x$ does not extend from $(0,\infty)$). We will suffice with functions, thus, let $f\colon U \to \mathbb{R}$ be a function defined in a neighbourhood $U \ni x$. Let $\lambda\colon M \to \mathbb{R}$ be such that λ equals one in a neighbourhood of x and with support in U. Then the function $\lambda \cdot f$, extended outside of U by the zero function, is smooth both on U and $M \smallsetminus \operatorname{supp} f$ (where it equals zero); clearly, it represents the same germ at x as the orginial f. Thus, the restriction map $C^\infty M \to C^\infty U$ induces a map

$$C^\infty_x M \xrightarrow{\cong} C^\infty_x U$$

which we just showed to be surjective; it is injective almost by definition (if two function f, g have the same germ in U, i.e. agree in a neighbourhood of x in U then this neighbourhood is also a neighbourhood in M and, thus, the germs of f and g in M agree).

1.3. **Tangent vectors.**

Theorem 7. *Let $D\colon C^\infty M \to \mathbb{R}$ be an $\mathbb{R}$-linear map that satisfies the following "Leibniz rule at x_0"*

$$D(f \cdot g) = Df \cdot g(x_0) + f(x_0) \cdot Dg$$

(we say that D is a derivation at x_0). Then there exists a unique tangent vector $A \in T_{x_0}M$ such that $Df = Af$, the derivative of f at x in the direction of A.

Proof. First we prove the statement for $M = \mathbb{R}^n$ and $x_0 = 0$. Let $f \in C^\infty\mathbb{R}^n$ and write

$$f(x) - f(0) = [f(t\cdot x)]^1_{t=0} = \int_0^1 \tfrac{\mathrm{d}}{\mathrm{d}t} f(t \cdot x)\ \mathrm{d}t = \int_0^1 \sum \tfrac{\partial f}{\partial x^i}(t\cdot x)x^i\ \mathrm{d}t = \sum \int_0^1 \tfrac{\partial f}{\partial x^i}(t \cdot x)\ \mathrm{d}t \cdot x^i.$$

Denoting $g_i(x) = \int_0^1 \frac{\partial f}{\partial x^i}(t \cdot x)\ \mathrm{d}t$, a smooth function on $\mathbb{R}^n$ with $g_i(0) = \frac{\partial f}{\partial x^i}(0)$, we get

$$f(x) = f(0) + \sum g_i(x) x^i.$$

Now we apply D to obtain

$$Df = \underbrace{D(f(0))}_{0} + \sum (Dg_i \cdot \underbrace{x^i(0)}_{0} + g_i(0) \cdot Dx^i) = \sum \tfrac{\partial f}{\partial x^i}(0) \cdot Dx^i = d_0 f(Dx^1, \dots, Dx^n)$$

(it is easy to see from $D(1 \cdot 1) = D1 \cdot 1 + 1 \cdot D1$ that $D1 = 0$ and consequently also $D(f(0)) = f(0) \cdot D1 = 0$), i.e. Df is the derivative of f at 0 in the direction of the vector $(Dx^1, \dots, Dx^n)$.

Let now M and x_0 be general and let f be a function that equals zero near x_0. Then there exists a function λ that equals zero near x_0 and equals one on the support of f, i.e. $f(x) \neq 0 \Rightarrow \lambda(x) = 1$. Then we may write

$$Df = D(\lambda \cdot f) = D\lambda \cdot f(x_0) + \lambda(x_0) \cdot Df = 0.$$

Thus, by the additivity, D agrees on any two functions with the same germ at x_0; in other words, D gives rise to a map $D\colon C^\infty_{x_0} M \to \mathbb{R}$. From the diagram

$$\begin{tikzcd} T_{x_0}U \arrow[r, "\cong"] \arrow[d, "\cong"'] & \mathrm{Der}_{x_0}(C^\infty U) \arrow[d, "\cong"'] \\ T_{x_0}M \arrow[r] & \mathrm{Der}_{x_0}(C^\infty M) \end{tikzcd}$$

we conclude from the local result that D is indeed a derivative in the direction of a unique vector $A \in T_{x_0}M$. $\square$

Theorem 8. *Let $D\colon C^\infty M \to C^\infty M$ be an $\mathbb{R}$-linear map that satisfies the following "Leibniz rule"*

$$D(f \cdot g) = Df \cdot g + f \cdot Dg$$

(we say that D is a derivation). Then there exists a unique vector field $X \in \mathfrak{X}M$ such that $Df = Xf$, the derivative of f in the direction of X.

Proof. For any $x \in M$, the composition of D with the evaluation map at x (sending $f \mapsto f(x)$),

$$D_x\colon C^\infty M \xrightarrow{D} C^\infty M \xrightarrow{\mathrm{ev}_x} \mathbb{R},$$

is a derivation at x and, thus, there is a unique vector $X_x \in T_xM$ such that $Df(x) = D_xf = X_xf$. It remains to show that $x \mapsto X_x$ is a *smooth* vector field.

Locally, we have seen that the coordinates of X_x are obtained as $(X_x)^i = D_xx^i$. Choose a coordinate chart around $x_0 \in M$ and let $\lambda\colon M \to \mathbb{R}$ be a function that equals one in a neighbourhood $U \ni x_0$ and with support inside the coordinate chart. Then, for all $x \in U$, we get

$$(X_x)^i = D_xx^i = D_x(\lambda \cdot x^i) = D(\lambda \cdot x^i)(x)$$

and, since $\lambda \cdot x^i$ is smooth, so is $D(\lambda \cdot x^i)$. □

1.4. **Duality between tangent and cotangent spaces.** We consider the collection P_x of all paths through x, i.e. maps $\gamma\colon \mathbb{R} \to M$ satisfying $\gamma(0) = x$ and the collection F_x of all functions vanishing at x, i.e. maps $f\colon M \to \mathbb{R}$ satisfying $f(x) = 0$. Then we have the following mapping

$$\begin{aligned} F_x \times P_x &\to \mathbb{R} \\ (f, \gamma) &\mapsto \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} f \circ \gamma(t). \end{aligned}$$

Now we consider the relations on both F_x and P_x that identify those functions/paths that give rise to the same pairing. In local coordinates, we have

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} f \circ \gamma(t) = \sum \tfrac{\partial f}{\partial x^i}(x) \tfrac{\mathrm{d}\gamma^i}{\mathrm{d}t}(0) = \begin{pmatrix} \tfrac{\partial f}{\partial x^1}(x) & \cdots & \tfrac{\partial f}{\partial x^n}(x) \end{pmatrix} \begin{pmatrix} \tfrac{\mathrm{d}\gamma^1}{\mathrm{d}t}(0) \\ \vdots \\ \tfrac{\mathrm{d}\gamma^n}{\mathrm{d}t}(0) \end{pmatrix}$$

and thus, we see that this relation means having the same derivatives. We denote the class of f by d_xf, the *differential* of f at x, and the corresponding decomposition T_x^*M. We denote the class of γ by $\dot\gamma(0) = \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \gamma(t)$, the *tangent vector* to γ at 0, and the corresponding decomposition T_xM. Thus, we obtain an induced mapping

$$\begin{aligned} T_x^*M \times T_xM &\to \mathbb{R}, \\ (\mathrm{d}_xf, \dot\gamma(0)) &\mapsto \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} f \circ \gamma(t) = \mathrm{d}_xf(\dot\gamma(0)), \end{aligned}$$

or simply $(\varphi, A) \mapsto \varphi(A)$. Using the local coordinate expression, we see that T_xM and T_x^*M are dual to each other via this pairing. Even without employing coordinates, T_x^*M has an obvious vector space structure and, thus, we may endow T_xM with a vector space structure using the pairing:

$$T_xM \xrightarrow{\cong} \mathrm{Lin}(T_x^*M, \mathbb{R}) = (T_x^*M)^*$$

Concretely, this means that the sum $A + B$ of two vectors (represented by paths) is a vector that differentiates functions as $\mathrm{d}_xf(A + B) = \mathrm{d}_xf(A) + \mathrm{d}_xf(B)$. Locally, it is easy to add the representing paths, but there is no obvious coordinate-free addition of paths in M (put differently, any local chart gives a way of adding the representing paths and, while different, they produce the same tangent vector).

Later, we will also need the tangent vector $\dot\gamma(t_0)$ of a path γ at a general time t_0. This is most easily defined as $\dot\gamma(t_0) = \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \gamma(t_0 + t)$, i.e. by reparametrization. Let us look at a different point of view: the path γ has an associated tangent map $\gamma_*\colon T\mathbb{R} \to TM$ and there is a canonical vector field $\tfrac{\mathrm{d}}{\mathrm{d}t}$ on $T\mathbb{R}$ (with value at t_0 denoted by $\tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0}$) and $\dot\gamma(t_0) = \gamma_*(\tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0})$ (this is obvious because $\tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0}$ is the tangent vector at 0 of the path $t \mapsto t_0 + t$, the same reparametrization map as above).

We will also need a differential $\mathrm{d}_x f$ of a function f with $f(x) \neq 0$. Again, this is defined via "reparametrization" as $\mathrm{d}_x f = \mathrm{d}_x(f - f(x))$.

2. Lie bracket

We define the Lie bracket through derivations.

Definition 9. Let X and Y be two vector fields. Then it is easy to see that $f \mapsto XYf - YXf$ is a derivation and the corresponding vector field is denoted $[X, Y]$ and called the *Lie bracket* of the vector fields X and Y.

It is easy to derive a coordinate formula

$$\begin{aligned} XYf - YXf &= \sum_{i,j} X^j \partial_j (Y^i \partial_i f) - \sum_{i,j} Y^j \partial_j (X^i \partial_i f) \\ &= \sum_{i,j} (X^j \partial_j Y^i \partial_i f + X^j Y^i \partial_j \partial_i f) - \sum_{i,j} (X^j \partial_j Y^i \partial_i f + Y^j X^i \partial_j \partial_i f) \\ &= \sum_{i,j} (X^j \partial_j Y^i - Y^j \partial_j X^i) \partial_i f \end{aligned}$$

so that $[X, Y] = \sum_{i,j} (X^j \partial_j Y^i - Y^j \partial_j X^i) \partial_i$.

Lemma 10. *Vector fields X and Y are f-related if and only if*

$$f(\mathrm{Fl}_t^X(x)) = \mathrm{Fl}_t^Y(f(x))$$

In other words f transfers the flow lines of X into the flow lines of Y. We will use this property quite often.

Proof. Differentiating the given equality, we get

$$f_* X(x) = \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} f(\mathrm{Fl}_t^X(x)) = \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \mathrm{Fl}_t^Y(f(x)) = Y(f(x)),$$

which is precisely the definition of the f-relatedness.

In the opposite direction, given that X and Y are f-related, we wish to prove the equality from the statement, i.e. we want to prove that $\gamma(t) = f(\mathrm{Fl}_t^X(x))$ is an integral curve of Y through $f(x)$. Since clearly $\gamma(0) = f(x)$, we need only check that it satisfies the differential equation of an integral curve:

$$\dot{\gamma}(t_0) = \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0} f(\mathrm{Fl}_t^X(x)) = f_* X(\mathrm{Fl}_{t_0}^X(x)) = Y(f(\mathrm{Fl}_{t_0}^X(x)) = Y(\gamma(t_0))$$

(the third equality uses the f-relatedness). $\square$

Definition 11. Let X, Y be two vector fields on a manifold M. Then we denote

$$(\mathrm{Fl}_t^X)^* Y(x) = (\mathrm{Fl}_{-t}^X)_* Y(\mathrm{Fl}_t^X(x)) \in T_x M$$

the "pullback" of Y along the flow Fl_t^X of X. For each $x \in M$ it is defined for t small.

We will need a useful property for the proof of the next proposition. It is based on an observation that for a function $\varphi(s, t)$ of two variables, with values in a vector space, we have

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0} \varphi(t, t) = \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0} \varphi(t, t_0) + \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0} \varphi(t_0, t).$$

Now let X be a time-dependent vector field, i.e. a map $X \colon \mathbb{R} \times M \to TM$ such that $X(t, x) \in T_x M$. Let $f \colon \mathbb{R} \times M \to \mathbb{R}$ be a "time-dependent" function on M (just a function on $\mathbb{R} \times M$). Write X_t and f_t for the vector field and function obtained by plugging in a specific value of t. Then we may form the directional derivative $X_t f_t$ and

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0} X_t f_t = (\tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0} X_t) f_{t_0} + X_{t_0}(\tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0} f_t). \tag{$*$}$$

(Locally, we have $X_t f_t(x) = X(t, x)^i \frac{\partial f}{\partial x^i}(t, x)$ and we apply the previous observation.)

Proposition 12. $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0} (\mathrm{Fl}_t^X)^* Y(x) = (\mathrm{Fl}_{t_0}^X)^* [X, Y](x)$.

Proof. First assume that $t_0 = 0$ and let $f : M \to \mathbb{R}$ be a smooth function. We differentiate f in the direction of the left hand side:

$$\begin{aligned}
\Big(\tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} (\mathrm{Fl}^X_t)^* Y(x)\Big) f &\overset{(*)}{=} \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \Big((\mathrm{Fl}^X_t)^* Y(x) f\Big) \\
&= \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \Big((\mathrm{Fl}^X_{-t})_* Y(\mathrm{Fl}^X_t(x)) f\Big) \\
&= \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \Big(Y(\mathrm{Fl}^X_t(x))(f \circ \mathrm{Fl}^X_{-t})\Big) \\
&\overset{(*)}{=} \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} Y(x)(f \circ \mathrm{Fl}^X_{-t}) + \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} Y(\mathrm{Fl}^X_t(x))(f) \\
&\overset{(*)}{=} Y(x)(\tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} f \circ \mathrm{Fl}^X_{-t}) + \tfrac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} (Yf)(\mathrm{Fl}^X_t(x)) \\
&= Y(x)(-Xf) + X(x)(Yf) \\
&= -(YXf)(x) + (XYf)(x) = \Big([X,Y](x)\Big) f
\end{aligned}$$

(the steps labeled by $(*)$ involve the observation made before the proposition, the first in the opposite direction).

For a general t_0, we have $(\mathrm{Fl}^X_t)^* Y(x) = (\mathrm{Fl}^X_{t_0})^* (\mathrm{Fl}^X_{t-t_0})^* Y(x)$. Since $(\mathrm{Fl}^X_{t_0})^*$ is a linear map we can interchange with $\frac{\mathrm{d}}{\mathrm{d}t}$. □

Corollary 13. *The following conditions are equivalent:*

- $[X, Y] = 0$,
- $(\mathrm{Fl}^X_t)^* Y = Y$, *i.e.* Y *is* Fl^X_t*-related with itself for all* t,
- $\mathrm{Fl}^X_t \mathrm{Fl}^Y_s(x) = \mathrm{Fl}^Y_s \mathrm{Fl}^X_t(x)$, *i.e. the flow lines commute.*

In general we have $\mathrm{Fl}^Y_{-s} \mathrm{Fl}^X_{-t} \mathrm{Fl}^Y_s \mathrm{Fl}^X_t(x) = x + st[X,Y](x) + o(s,t)^2$.

Proof. The equivalence of the first three conditions follows immediately from the previous proposition – the second condition states that $(\mathrm{Fl}^X_t)^* Y$ is a constant function of t, i.e. that the derivative is zero and this is essentially the first condition. At the same time, the second condition is equivalent to Y being Fl^X_t-related to itself and this is equivalent to Fl^X_t preserving the integral curves of Y, which is precisely the third condition.

Differentiating the commutator of the flows twice, we get

$$\tfrac{\partial}{\partial t}\big|_{t=0} \tfrac{\partial}{\partial s}\big|_{s=0} \mathrm{Fl}^Y_{-s} \mathrm{Fl}^X_{-t} \mathrm{Fl}^Y_s \mathrm{Fl}^X_t(x) = \tfrac{\partial}{\partial t}\big|_{t=0} \Big(-Y(x) + (\mathrm{Fl}^X_t)^* Y(x)\Big) = [X,Y](x)$$

The remaining derivatives of order at most two are clearly zero. □

3. Distributions

Theorem 14. *If vector fields* $X_1, \dots, X_n$ *are linearly independent and satisfy* $[X_i, X_j] = 0$ *then, in a neighbourhood of any point* x*, there exists a coordinate chart in which* $X_i = \partial_i$.

Proof. We define a map

$$\varphi \colon \mathbb{R}^n \to M, \quad (t^1, \dots, t^n) \mapsto \mathrm{Fl}^{X_1}_{t^1} \cdots \mathrm{Fl}^{X_n}_{t^n}(x)$$

(it is defined in a neighbourhood of 0). Since we may interchange the flows (since $[X_i, X_j] = 0$), the partial derivatives equal

$$\partial_i \varphi(t) = \partial_i \mathrm{Fl}^{X_i}_{t^i} \mathrm{Fl}^{X_1}_{t^1} \cdots \widehat{\mathrm{Fl}^{X_i}_{t^i}} \cdots \mathrm{Fl}^{X_n}_{t^n}(x) = X_i(\mathrm{Fl}^{X_i}_{t^i} \mathrm{Fl}^{X_1}_{t^1} \cdots \widehat{\mathrm{Fl}^{X_i}_{t^i}} \cdots \mathrm{Fl}^{X_n}_{t^n}(x)) = X_i(\varphi(t))$$

and φ is a local diffeomorphism. We may thus use its inverse φ^{-1} as a coordinate chart on M and the above shows that, in this chart, we have $\partial_i = X_i$. □

A distribution $\mathcal{S}$ on M is called *involutive* if for every two vector fields $X, Y \in \mathcal{S}$ their bracket $[X, Y]$ also lies in $\mathcal{S}$.

Theorem 15 (Frobenius theorem)**.** *If* $\mathcal{S}$ *is involutive then for every* $x \in M$ *there exists a local coordinate system* $t^1, \dots, t^n$ *in a neighbourhood* U *of* x *such that the vector fields* $\partial_1, \dots, \partial_k$ *form a basis of the distribution* $\mathcal{S}$ *on* U*. In particular* $\mathcal{S}$ *is integrable.*

Proof. Let $X_1, \dots, X_k$ be local vector fields which span the distribution $\mathcal{S}$ near x and choose vector fields $X_{k+1}, \dots, X_n$ so that $(X_1, \dots, X_n)$ form a basis near x. We then define a map

$$\begin{aligned} \varphi : \mathbb{R}^n \supseteq U &\longrightarrow M \\ (t^1, \dots, t^n) &\longmapsto \mathrm{Fl}^{X_1}_{t^1} \cdots \mathrm{Fl}^{X_n}_{t^n}(x) \end{aligned}$$

The partial derivatives at the origin clearly consist of the vectors $X_i(x)$ and thus φ is a local diffeomorphism – its inverse will form our coordinate system.

Let us compute the partial derivative with respect to t^i for $i \le k$ at a general point.

$$\partial_i \varphi(t^1, \dots, t^n) = \left(\mathrm{Fl}^{X_1}_{t^1}\right)_* \cdots \left(\mathrm{Fl}^{X_{i-1}}_{t^{i-1}}\right)_* X_i \left(\mathrm{Fl}^{X_i}_{t^i} \cdots \mathrm{Fl}^{X_n}_{t^n}(x)\right)$$

To conclude the proof it is therefore enough to show that for any Y belonging to $\mathcal{S}$ the pullbacks $(\mathrm{Fl}^Y_t)^* X_i$ also belong to $\mathcal{S}$ (then the same will be true for pullbacks $(\mathrm{Fl}^Y_t)^* X$ with $X \in \mathcal{S}$ by linearity, and we apply the claim to X_i, $\left(\mathrm{Fl}^{X_{i-1}}_{t^{i-1}}\right)_* X_i$, etc.) Denote this pullback by

$$Y_i(t) = (\mathrm{Fl}^Y_t)^* X_i(x) \in T_x M$$

and write $[Y, X_i] = a_i^j X_j$. By Lemma 12 the paths $Y_i(t)$ satisfy the following system of differential equations

$$\tfrac{\mathrm{d}}{\mathrm{d}t} Y_i(t) = (\mathrm{Fl}^Y_t)^* [Y, X_i] = a_i^j (\mathrm{Fl}^Y_t(x)) Y_j(t)$$

We have $Y_i(0) = X_i(x) \in \mathcal{S}(x)$ and, since the system is linear, we must have $Y_i(t) \in \mathcal{S}(x)$ for all t (namely, there exists a solution of the system $\tfrac{\mathrm{d}}{\mathrm{d}t} Z_i(t) = a_i^j (\mathrm{Fl}^Y_t(x)) Z_j(t)$ with $Z_i \in \mathcal{S}(x)$ and with $Z_i(0) = X_i(x)$. By uniqueness, we must have $Y_i(t) = Z_i(t)$ and, thus, $Y_i(t) \in \mathcal{S}(x)$.) □

Theorem 16 (Frobenius theorem through 1-forms)**.** *Let $\omega \colon TM \to V$ be a smooth map that is linear on each $T_x M$ (we say that ω is a V-valued 1-form) and surjective. Then $\ker \omega$ is a distribution. It is integrable if and only if $\omega(X) = 0$, $\omega(Y) = 0 \Rightarrow \mathrm{d}\omega(X, Y) = 0$.*

This uses the exterior differential of the next section.

Proof. In local coordinates on M and in a basis of V, the 1-form ω is given by a matrix of maximal rank. We may assume that the left most square block is regular in a neighbourhood of a given point and use the Gauss elimination to make this matrix $(E \mid A)$. Then $\ker \omega$ is given by $(x^{n-k+1}, \dots, x^n)^T = -A(x^1, \dots, x^{n-k})$, proving that it is a (smooth) distribution. Now $\mathrm{d}\omega(X, Y) = X\omega(Y) - Y\omega(X) - \omega([X, Y])$ gives easily the result. □

4. Exterior differential

4.1. **Invariant differentiation of forms.** Let us study the invariance of higher derivatives under the change of coordinates, i.e. let ω be a k-linear form with components $\omega_{i_1 \cdots i_k}$,

$$\omega = \omega_{i_1 \cdots i_k} \mathrm{d}x^{i_1} \otimes \cdots \otimes \mathrm{d}x^{i_k}.$$

Differentiating all components in the direction of a vector X_0 gives another k-linear form $D_{X_0}\omega$ with

$$D_{X_0}\omega = \partial_{i_0} \omega_{i_1 \cdots i_k} (X_0)^{i_0} \mathrm{d}x^{i_1} \otimes \cdots \otimes \mathrm{d}x^{i_k}$$

which we interpret as a $(k+1)$-linear form $D\omega(X_0, X_1, \dots, X_k) = D_{X_0}\omega(X_1, \dots, X_k)$; its coordinates are $(D\omega)_{i_0 i_1 \cdots i_k} = \partial_{i_0} \omega_{i_1 \cdots i_k}$. Now we apply the change of coordinates f to get

$$f^*(D\omega)_{j_0 j_1 \cdots j_k} = (\partial_{i_0} \omega_{i_1 \cdots i_k} \circ f) \cdot \partial_{j_0} f^{i_0} \cdot \partial_{j_1} f^{i_1} \cdot \cdots \cdot \partial_{j_k} f^{i_k}.$$

On the other hand $D(f^*\omega)$ equals

$$\begin{aligned} D(f^*\omega)_{j_0 j_1 \cdots j_k} &= \partial_{j_0}((\omega_{i_1 \cdots i_k} \circ f) \cdot \partial_{j_1} f^{i_1} \cdot \cdots \cdot \partial_{j_k} f^{i_k}) \\ &= \partial_{j_0}((\omega_{i_1 \cdots i_k} \circ f)) \cdot \partial_{j_1} f^{i_1} \cdot \cdots \cdot \partial_{j_k} f^{i_k} \\ &\quad + (\omega_{i_1 \cdots i_k} \circ f) \cdot \partial_{j_1} f^{i_1} \cdot \cdots \cdot \partial^2_{j_0 j_k} f^{i_k} \cdot \cdots \cdot \partial_{j_k} f^{i_k} \end{aligned}$$

The first term equals to $f^*(D\omega)_{j_0 j_1 \cdots j_k}$ by the chain rule. Now the point is that to get $f^*(D\omega) = D(f^*\omega)$ (i.e. to get a differential that does not depend on coordinates), we have to get rid of the second term involving the second derivative $\partial^2_{j_0 j_k} f^{i_k}$. It clearly disappears after antisymmetrization. Some representation theory would be required to get that no other part is invariant and we will not attempt to do this here.

Thus, we get an invariant differentiation operator – the *exterior differential* – on antisymmetric forms by antisymmetrizing $D_{X_0}\omega(X_1, \ldots, X_k)$; for technical reasons, we multiply the antisymmetrization by $\frac{(k+1)!}{1!k!}$, since the form already was antisymmetric in the variables $X_1, \ldots, X_k$ and obtain

$$\mathrm{d}\omega(X_0, \ldots, X_k) = \sum_i (-1)^i D_{X_i}\omega(X_0, \ldots, \widehat{X_i}, \ldots, X_k)$$

Denoting by $\phi \wedge \psi = \frac{(|\phi|+|\psi|)!}{|\phi|!|\psi|!}\operatorname{Alt}(\phi \otimes \psi)$, we obtain for

$$\omega = \sum_{i_1 < \cdots < i_k} \omega_{i_1 \cdots i_k} \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$$

a coordinate expression

$$\mathrm{d}\omega = \sum_{i_0} \sum_{i_1 < \cdots < i_k} \partial_{i_0}\omega_{i_1 \cdots i_k} \mathrm{d}x^{i_0} \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}.$$

4.2. **Coordinate-free formula for the exterior differential.** The differentiation operator D_{X_0} satisfies the Leibniz rule

$$\begin{aligned} D_{X_0}(\omega_{i_1 \cdots i_k}(X_1)^{i_1} \cdots (X_k)^{i_k}) = {} & D_{X_0}\omega_{i_1 \cdots i_k} \cdot (X_1)^{i_1} \cdots (X_k)^{i_k} \\ & + \sum \omega_{i_1 \cdots i_k}(X_1)^{i_1} \cdots D_{X_0}(X_j)^{i_j} \cdots (X_k)^{i_k} \end{aligned}$$

which (after subtracting the sum from the right hand side) translates to

$$D_{X_0}\omega(X_1, \ldots, X_k) = D_{X_0}(\omega(X_1, \ldots, X_k)) - \sum \omega(X_1, \ldots, X_{j-1}, D_{X_0}X_j, X_{j+1}, \ldots, X_k).$$

Here the first D_{X_0} on the right is the directional derivative of the function $\omega(X_1, \ldots, X_k)$. The second appearance is, however, very different and we have $[X, Y] = D_X Y - D_Y X$. The exterior differential then equals

$$\begin{aligned} \mathrm{d}\omega(X_0, \ldots, X_k) &= \sum_i (-1)^i D_{X_i}\omega(X_0, \ldots, \widehat{X_i}, \ldots, X_k) \\ &= \sum_i (-1)^i X_i\omega(X_0, \ldots, \widehat{X_i}, \ldots, X_k) \\ &\quad - \sum_{i<j} (-1)^i \omega(X_0, \ldots, \widehat{X_i}, \ldots, X_{j-1}, D_{X_i}X_j, X_{j+1}, \ldots, X_k) \\ &\quad - \sum_{i>j} (-1)^i \omega(X_0, \ldots, X_{j-1}, D_{X_i}X_j, X_{j+1}, \ldots, \widehat{X_i}, \ldots, X_k) \end{aligned}$$

where we split the antisymmetrization of the second term according to whether $i < j$ or $i > j$. Next we move the term $D_{X_i}X_j$ onto the first spot (here the sign differs for the two possibilities):

$$\begin{aligned} &= \sum_i (-1)^i X_i\omega(X_0, \ldots, \widehat{X_i}, \ldots, X_k) \\ &\quad + \sum_{i<j} (-1)^{i+j} \omega(D_{X_i}X_j, X_0, \ldots, \widehat{X_i}, \ldots, \widehat{X_j}, \ldots, X_k) \\ &\quad - \sum_{i>j} (-1)^{i+j} \omega(D_{X_i}X_j, X_0, \ldots, \widehat{X_j}, \ldots, \widehat{X_i}, \ldots, X_k) \end{aligned}$$

and finally swap the indices i, j in the last sum and subtract, using $D_{X_i}X_j - D_{X_j}X_i = [X_i, X_j]$, to obtain the final formula

$$\begin{aligned}&= \sum_i (-1)^i X_i \omega(X_0, \dots, \widehat{X_i}, \dots, X_k) \\ &\quad + \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_0, \dots, \widehat{X_i}, \dots, \widehat{X_j}, \dots, X_k)\end{aligned}$$

5. Integration of forms

5.1. **Integral.** Let ω be an n-form on an open subset $V \subseteq \mathbb{R}^n$, say with a compact support. Writing

$$\omega(x) = a(x) \cdot \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n,$$

we define

$$\int_V \omega = \int \cdots \int_V a(x) \; \mathrm{d}x^1 \cdots \mathrm{d}x^n.$$

The important property of this integral is that for a diffeomorphism $\varphi \colon W \to V$ with *positive Jacobian*, we get

$$\int_W \varphi^* \omega = \int_V \omega.$$

This follows from the theorem about the transformation of the integral. Clearly, the above defined integral is additive in ω.

Let M be an oriented manifold of dimension n. We assume for simplicity that M is compact. Let ω be an n-form on M. We would like to define $\int_M \omega$. Consider the maximal *orientation-preserving* atlas on M and choose a partition of unity λ_i so that $\operatorname{supp} \lambda_i$ is a subset of a domain U_i of a chart $\varphi_i \colon U_i \to V_i$. Then we define

$$\int_M \omega = \sum_i \int_{V_i} (\varphi_i^{-1})^* (\lambda_i \omega)$$

We note that $\lambda_i \omega$ has a compact support inside U_i and, thus, the pullback $(\varphi_i^{-1})^*(\lambda_i \omega)$ has a compoact support inside V_i. Thus, the integral exists and is finite. It remains to show that it does not depend on the choice of the partition λ_i.

Thus, let μ_i be another partition. Then we get

$$\sum_i \int_{V_i} (\varphi_i^{-1})^* (\lambda_i \omega) = \sum_{i,j} \int_{V_i} (\varphi_i^{-1})^* (\lambda_i \mu_j \omega)$$

Denoting $\theta = \lambda_i \mu_j \omega$, a compactly supported n-form inside $U_i \cap U_j$, whose images we denote $V_{ij} = \varphi_i(U_i \cap U_j)$ and $V_{ji} = \varphi_j(U_i \cap U_j)$, we have

$$\int_{V_i} (\varphi_i^{-1})^* \theta = \int_{V_{ij}} (\varphi_i^{-1})^* \theta = \int_{V_{ij}} \varphi_{ij}^* ((\varphi_j^{-1})^* \theta) = \int_{V_{ji}} (\varphi_j^{-1})^* \theta = \int_{V_j} (\varphi_j^{-1})^* \theta$$

by the invariance of the integral with respect to the diffeomorphism $\varphi_{ij} = \varphi_i \varphi_j^{-1} \colon V_{ij} \to V_{ji}$.

5.2. **Manifolds with boundary.** The main idea here is: in exactly the same manner in which manifolds are built from the Euclidean space $\mathbb{R}^n$, manifolds with boundary are built from the Euclidean halfspace $\mathbb{H}^{n+1} = \{(x^0, x^1, \dots, x^n) \in \mathbb{R}^{n+1} \mid x^0 \leq 0\}$. It is however important that we allow tangent vectors at the boundary hyperplane to be all vectors from $\mathbb{R}^{n+1}$, i.e.

$$T\mathbb{H}^{n+1} = \mathbb{H}^{n+1} \times \mathbb{R}^{n+1}.$$

Thus, the geometric definition using paths is inappropriate. Derivations work well if we interpret $\partial_0 f(x)$ for a boundary point x to be the one-sided partial derivative.

Formally, a map between open subsets of the half-spaces is said to be smooth, if all partial derivatives exist (one-sided where needed) and are continuous. A diffeomorphism between open subsets of $\mathbb{H}^{n+1}$ preserves the boundary points, since at an interior point, any (local) diffeomorphism has a local inverse and as such maps to an interior point.

With this notion, we define a (smooth) *manifold with boundary* W as a topological space, Hausdorff and with countable basis of topology, equipped with a maximal atlas consisting of homeomorphisms $\varphi\colon U \to V$ with V an open subset of $\mathbb{H}^{n+1}$ and with all change of coordinate maps smooth in the above sense. We define the *boundary* of W to be the set ∂W of points that correspond to the boundary points in a chart.

The standard bases $(e_0, e_1, \ldots, e_n)$ of $\mathbb{H}^{n+1}$ and $(e_1, \ldots, e_n)$ of $\mathbb{R}^n$ are considered positive. We say that $\partial\mathbb{H}^{n+1}$ is oriented via its outward normal: The outward normal is by definition e_0 (or any combination $x^0e_0 + x^1e_1 + \cdots x^ne_n$ with $x^0 > 0$) and a basis $(u_1, \ldots, u_n)$ is then a positive basis of $\partial\mathbb{H}^{n+1}$ according to this principle if and only if $(e_0, u_1, \ldots, u_n)$ is a positive basis of $\mathbb{H}^{n+1}$. This gives a way of orienting a boundary ∂W of any oriented manifold with boundary W. We will always consider ∂W with this induced orientation.

5.3. Stokes' theorem.

Theorem 17. *For a compact manifold with boundary W of dimension $n+1$ and an n-form ω on W, we have*

$$\int_{\partial W} \omega = \int_W \mathrm{d}\omega.$$

(The left hand side is really the integral of the pullback $\iota^*\omega$ along the inclusion $\iota\colon \partial W \to W$.)

Proof. Denoting $\partial\varphi_i$ the restriction of φ_i to the boundaries, $\partial\varphi_i\colon \partial U_i \to \partial V_i$, it is enough to compare the contributions

$$\int_{\partial V_i} (\partial\varphi_i^{-1})^*\iota^*(\lambda_i\omega) \quad \text{and} \quad \int_{V_i} (\varphi_i^{-1})^*\mathrm{d}(\lambda_i\omega).$$

Denoting $j\colon \partial V_i \to V_i$ the inclusion, the left hand side equals $\int_{\partial V_i} j^*(\varphi_i^{-1})^*(\lambda_i\omega)$, while the right hand side equals $\int_{V_i} \mathrm{d}(\varphi_i^{-1})^*(\lambda_i\omega)$. Thus, denoting $\theta = (\varphi_i^{-1})^*(\lambda_i\omega)$, we want to show that

$$\int_{\partial V_i} j^*\theta = \int_{V_i} \mathrm{d}\theta.$$

Since θ is an n-form on $V_i \subseteq \mathbb{H}^{n+1}$, we may write

$$\theta = \sum_i a_i \cdot \mathrm{d}x^0 \wedge \cdots \wedge \widehat{\mathrm{d}x^i} \wedge \cdots \wedge \mathrm{d}x^n.$$

Since $j(x^1, \ldots, x^n) = (0, x^1, \ldots, x^n)$, we get $j^*\mathrm{d}x^0 = 0$ and $j^*\mathrm{d}x^i = \mathrm{d}x^i$, for $i > 0$. Thus,

$$j^*\theta = a_0 j \cdot \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$$

and the integral on the left is

$$\int_{\partial V_i} j^*\theta = \int \cdots \int_{V_i} a_0(0, x^1, \ldots, x^n)\ \mathrm{d}x^1 \cdots \mathrm{d}x^n.$$

Now we simplify the integral on the right, i.e. we compute

$$\begin{aligned}
\mathrm{d}\theta &= \sum_i \partial_i a_i \cdot \mathrm{d}x^i \wedge \mathrm{d}x^0 \wedge \cdots \wedge \widehat{\mathrm{d}x^i} \wedge \cdots \wedge \mathrm{d}x^n \\
&= \sum_i (-1)^i \partial_i a_i \cdot \mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^n.
\end{aligned}$$

Now the integral simplifies to

$$\begin{aligned}
\int_{V_i} \mathrm{d}\omega &= \sum_i (-1)^i \int \cdots \int_{V_i} \partial_i a_i\ \mathrm{d}x^0 \cdots \mathrm{d}x^n \\
&= \sum_i (-1)^i \int \cdots \int_{V_i} \partial_i a_i\ \mathrm{d}x^i \mathrm{d}x^0 \cdots \widehat{\mathrm{d}x^i} \cdots \mathrm{d}x^n
\end{aligned}$$

For $i > 0$, we get $\int_{-\infty}^{\infty} \partial_i a_i \cdot \mathrm{d}x^i = a_i|_{x^i=\infty} - a_i|_{x^i=-\infty} = 0 - 0 = 0$, while $\int_{-\infty}^{0} \partial_i a_i \cdot \mathrm{d}x^0 = a_0|_{x^0=0} - a_0|_{x^0=-\infty} = a_0(0, x^1, \dots, x^n)$. Thus, the integral also equals

$$\int_{V_i} \mathrm{d}\omega = \int \cdots \int_{V_i} a_0(0, x^1, \dots, x^n)\ \mathrm{d}x^1 \cdots \mathrm{d}x^n. \qquad \square$$

Remark. It is interesting to see what we would get if we integrated over a cube instead. Then the "boundary conditions" $a_i|_{x^i=\pm\infty} = 0$ would be replaced by the non-zero restrictions to the faces of the cube and the resulting formula would be

$$\int_{I^{n+1}} \mathrm{d}\omega = \sum_i (-1)^i \int_{\partial_i^+ I^{n+1}} \omega - \sum_i (-1)^i \int_{\partial_i^- I^{n+1}} \omega$$

where the $\partial_i^\varepsilon I^{n+1}$ denotes the subset $\{(x^0, \dots, x^n) \in I^{n+1} \mid x^i = \varepsilon\}$. The signs reflect the orientations these faces so that the right hand side actually equals $\int_{\partial I^{n+1}} \omega$.

5.4. **Cohomology in top dimension.** In order to distinguish compact manifolds without boundary from those with boundary, we call then closed.

Theorem 18. *For any closed oriented Riemannian manifold M of dimension n, $H^n(M) \neq 0$.*

Proof. Since every n-form on M is closed, it is enough to find one that is not exact. We know that oriented Euclidean spaces admit a canonical volume form specified by the requirement $\mathrm{Vol}(e_1, \dots, e_n) = 1$ for any positive orthonormal basis $(e_1, \dots, e_n)$. In this way, we obtain a volume form $\mathrm{Vol} \in \Omega^n M$. In any chart compatible with the orientation,

$$\mathrm{Vol} = a \cdot \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n$$

with $a = \mathrm{Vol}(\partial_1, \dots, \partial_n) > 0$. Thus, $\int_M \mathrm{Vol}$ is an integral of a positive function and as such must be also positive. On the other hand, if $\mathrm{Vol} = \mathrm{d}\theta$, we would get

$$\int_M \mathrm{Vol} = \int_M \mathrm{d}\theta = \int_{\partial M} \theta = 0,$$

since $\partial M = \emptyset$. $\square$

5.5. **Homotopy invariance.** We would like to show that $H^k \mathbb{R}^n = 0$ for $k > 0$. This will follow from the following "homotopy invariance" property.

Theorem 19. *Let $h\colon [-1,1] \times M \to N$ be a smooth map and denote $h_t = h(t, -)$. Then for any closed k-form ω, we get $[h_{-1}^*\omega] = [h_1^*\omega] \in H^k M$.*

Proof. The idea of the proof is simple. Any k-form is determined by its integrals along k-dimensional cubes embedded in M. This is so because any embedding $[-1,1]^k \to M$ that maps $\partial_i(0)$ to $A_i \in T_x M$ restricts to the cube $[-\varepsilon, \varepsilon]^k$ to an embedding i_ε such that $\int (i_\varepsilon)^*\omega \sim (2\varepsilon)^k \omega(A_1, \dots, A_k)$ (equality holds in $\lim_{\varepsilon \to 0}$).

Now for an embedding $i\colon [-1,1]^k \to M$, we get an associated embedding $\mathrm{id} \times i\colon [-1,1]^{k+1} \to [-1,1] \times M$. Denote by $j_t\colon [-1,1]^k \to [-1,1]^{k+1}$ the embedding given by $j_t(t^1, \dots, t^k) = (t, t^1, \dots, t^k)$. Then $h_t^*\omega = j_t^* h^*\omega$ and both $j_{\pm 1}$ are embeddings as part of the boundary. Thus, the Stokes' theorem relates

$$\begin{aligned} \int_{[-1,1]\times[-1,1]^k} \mathrm{d}(h^*\omega) &= \int_{\partial([-1,1]\times[-1,1]^k)} h^*\omega \\ &= \int_{[-1,1]^k} h_1^*\omega - \int_{[-1,1]^k} h_{-1}^*\omega - \int_{[-1,1]\times\partial[-1,1]^k} h^*\omega \end{aligned} \qquad (*)$$

(the first two terms correspond to $\partial[-1,1] \times [-1,1]^k$). Writing

$$\mathrm{d}(h^*\omega) = a \cdot \mathrm{d}t \wedge \mathrm{d}t^1 \wedge \cdots \wedge \mathrm{d}t^k,$$

the integral on the left can be computed using Fubini's theorem as

$$\int_{[-1,1]\times[-1,1]^k} \mathrm{d}(h^*\omega) = \int_{[-1,1]^k} \left(\int_{[-1,1]} a(t, t^1, \dots, t^k)\ \mathrm{d}t \right) \mathrm{d}t^1 \cdots \mathrm{d}t^k.$$

This can be rephrased in terms of an operator $K\colon \Omega^{k+1}(I\times M)\to\Omega^k M$, given by the integral

$$K(\eta)(x)=\begin{cases}\int_{[-1,1]}\theta(t,x)\ \mathrm{d}t & \eta=\mathrm{d}t\wedge\theta\\ 0 & \eta(\partial_t,-,\dots,-)=0\end{cases}$$

as

$$\int_{[-1,1]\times[-1,1]^k}\mathrm{d}(h^*\omega)=\int_{[-1,1]^k}K\mathrm{d}(h^*\omega)=\int_{[-1,1]^k}K(h^*\mathrm{d}\omega).$$

The remaining boundary term in $(*)$ is then

$$\int_{[-1,1]\times\partial[-1,1]^k}h^*\omega=\int_{\partial[-1,1]^k}K(h^*\omega)=\int_{[-1,1]^k}\mathrm{d}K(h^*\omega),$$

again by the Stokes' theorem. Thus, we have finally obtained

$$\int_{[-1,1]^k}h_1^*\omega-\int_{[-1,1]^k}h_{-1}^*\omega=\int_{[-1,1]^k}(\mathrm{d}K(h^*\omega)+K(h^*\mathrm{d}\omega)))$$

or, in other words, $h_1^*\omega-h_{-1}^*\omega=\mathrm{d}K(h^*\omega)+K(h^*\mathrm{d}\omega)$. This implies rather easily the result, since, for ω closed, the first term on the right vanishes and, thus, the difference on the left is exact, i.e. the two terms represent the same cohomology class. □

In the situation from the above proof, we say that two chain maps (maps that commute with differentials, such as pullback maps h_t^*) are *chain homotopic* if there exists a collection of maps η (these are the compositions Kh^* in the proof) such that

$$g-f=\mathrm{d}\eta+\eta\mathrm{d}.$$

Then, f and g induce the same map in cohomology.

Corollary 20. *$H^k\mathbb{R}^n=0$ for $k>0$.*

Proof. There is a homotopy $\mathrm{id}\sim 0$ between the identity and the constant map onto the zero. Then for any closed k-form ω we have $[\omega]=[\mathrm{id}^*\,\omega]=[0^*\omega]=[0]$. □

6. Riemannian geometry

6.1. Preliminary results.

Lemma 21. *For every map*

$$F\colon \mathfrak{X}M\times\dots\times\mathfrak{X}M\to C^\infty M$$

that is $C^\infty M$-linear in each variable there exists a unique tensor field ω of type $(0,k)$ such that $F(X_1,\dots,X_k)=\omega(X_1,\dots,X_k)$.

Proof. We first prove that $F_x=\mathrm{ev}_x F$ is local; we will assume $k=1$ here for simplicity. Thus, let $X\in\mathfrak{X}M$ be zero in a neighbourhood of $x\in M$. Then there exists a function λ such that λ is zero near x and $X=\lambda\cdot X$. Then $F_x(X)=F_x(\lambda\cdot X)=\lambda(x)\cdot F_x(X)=0$. This allows one to define F_x on germs of vector fields and, consequently, a local version $F_x\colon \mathfrak{X}U\to\mathbb{R}$ as in the case of derivations.

Now, for general k, the map F is $C^\infty M$-linear in each variable and thus local in each variable. Since locally $X_j=X^{i_j}\partial_{i_j}$, we obtain

$$F(X_1,\dots,X_k)(x)=F_x(X_1,\dots,X_k)=F_x(\partial_{i_1},\dots,\partial_{i_k})X^{i_1}(x)\cdots X^{i_k}(x),$$

i.e. we have $\omega=F(\partial_{i_1},\dots,\partial_{i_k})\mathrm{d}x^{i_1}\otimes\dots\otimes\mathrm{d}x^{i_k}$. □

A similar result holds for maps $F\colon\mathfrak{X}M\times\dots\times\mathfrak{X}M\to\mathfrak{X}M$ – such maps are given by tensor fields of type $(1,k)$; the proof is the same.

A slightly generalization of the first part of the proof of the previous lemma is the following (for simplicity, we state only unary version):

Lemma 22. *Suppose that $F\colon\mathfrak{X}M\to C^\infty M$ is $\mathbb{R}$-linear and satisfies $F_x(fX)=0$ for each f that is zero in a neighbourhood of x. Then there exists a unique map $F\colon\mathfrak{X}U\to C^\infty U$ that passes to the same map $F_x\colon\mathfrak{X}_xM\to C_x^\infty M$.*

6.2. **Covariant derivative in vector spaces.** First, we describe the local covariant derivative, i.e. the covariant derivative in a vector space E, or its open subspace $U \subseteq E$. We will speak uniformly about vector fields, tensor fields and other fields as associations $F\colon x \mapsto F(x) \in \Phi(T_xU)$ – for vector fields, $\Phi(V) = V$ itself; for k-forms, $\Phi(V) = (\Lambda^k V)^*$; for functions, $\Phi(V) = \mathbb{R}$. To simplify the notation, we define $\Phi(TU) = \bigsqcup_{x\in U} \Phi(T_xU)$, so that F is a map $U \to \Phi(TU)$ with the property $F(x) \in \Phi(T_xU)$. Since we have a canonical identification $T_xU \cong E$, we may reinterpret these as maps $f\colon U \to \Phi(E)$.[1] Let now $A \in T_xU$ be a tangent vector. Then the directional derivative $Af = \mathrm{d}_x f(A) \in \Phi(E)$ can be translated back to an element of $\Phi(T_xU)$, denoted by D_AF. For a vector field $X \in \mathfrak{X}U$, we then get another field D_XF, given by $D_XF(x) = D_{X(x)}F$.

$$\begin{tikzcd}
\{\text{fields } F\colon U \to \Phi(TU)\} \arrow[r, leftrightarrow, "\cong"] \arrow[d, "D_X"'] & \{\text{maps } f\colon U \to \Phi(E)\} \arrow[d, "X"] \\
\{\text{fields } F\colon U \to \Phi(TU)\} \arrow[r, leftrightarrow, "\cong"] & \{\text{maps } f\colon U \to \Phi(E)\}
\end{tikzcd}$$

As in the case of the differential in calculus, we also consider the total derivative of F at a point x, an element $D_xF \in \hom(T_xM, \Phi(T_xM)) \cong \Phi(T_xM) \otimes T_x^*M$, given by $D_xF(A) = D_AF$. Finally, denoting again $\Phi(TM) \otimes T^*M = \bigsqcup_{x\in U} \Phi(T_xM) \otimes T_x^*M$, this total derivative defines a field $DF\colon U \to \Phi(TM) \otimes T^*M$.

In the special case of vector fields, we have $D_XY - D_YX = [X,Y]$; this will be important later. In general, $D_XD_YF - D_YD_XF = D_{[X,Y]}F$. We will also see later that this derivative is really connected with the vector space structure on E.

We will now speak briefly about parallel transport and geodesics – these are trivial in a vector space. First we observe that $D_{\dot\gamma(t)}F$ corresponds to

$$df(\dot\gamma(t)) = \tfrac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t))$$

and this depends only on the values of F along γ. Thus, we may consider a field $G(t)$ along $\gamma(t)$, i.e. $G\colon \mathbb{R} \to \Phi(TU)$ such that $G(t) \in \Phi(T_{\gamma(t)}U)$; again, this corresponds to a map $g\colon \mathbb{R} \to \Phi(E)$. Assuming that there is a field F such that $G(t) = F(\gamma(t))$, we also have $g(t) = f(\gamma(t))$. We may thus define $D_{\dot\gamma}G \overset{\text{def}}{=} D_{\dot\gamma}F$ since we have already seen that the latter corresponds simply to $\mathrm{d}f(\dot\gamma(t)) = \dot g(t)$ and thus depends only on G. As a special case, we obtain $D_{\dot\gamma(t)}\dot\gamma(t) = \ddot\gamma(t)$. In terms of this derivative, we may define the concept of a field F transporting parallelly along a path as $D_{\dot\gamma}F = 0$ – this simply means that f is constant along γ. The covariant derivative may be defined easily using this transport: in order to define $D_{\dot\gamma(0)}F$, we transport each $F(t)$ along γ to $\gamma(0)$ and thus obtain a path $\mathbb{R} \to \Phi(T_{\gamma(0)}M)$ and $D_{\dot\gamma(0)}F$ is the usual derivative of this path. The advantage of this approach is that it avoids going through $\Phi(E)$, of course at the cost of introducing the parallel transport. We will generalize this to Riemannian manifolds later.

We will now explain some simple rules that hold for computing with the local covariant derivative D. The first one is that for an actual function $f\colon U \to \mathbb{R}$ and its corresponding field $F\colon U \to \bigsqcup_{x\in U} \mathbb{R} = U \times \mathbb{R}$, the derivative is the usual directinal derivative, i.e. D_AF corresponds to Af (most of the times, these can be considered equal).

The second one is that any linear and natural $\tau = \tau_V\colon \Phi(V) \to \Psi(V)$ defines, for a field $F\colon U \to \Phi(TU)$, a field $\tau F\colon U \to \Psi(TU)$, as the composition $t \mapsto \tau(F(t))$. Since τ is linear, we obtain easily

$$D_A(\tau F) = \tau D_AF.$$

Our last tool will be a general Leibniz rule. Let two fields $F\colon U \to \Phi(TU)$ and $G\colon U \to \Psi(U)$ be given. Then we may form their tensor product $F \otimes G\colon U \to \Phi(TU) \otimes \Psi(TU)$ (again the union of the $\Phi(T_xU) \otimes \Psi(T_xU)$). This corresponds to the map $f \otimes g\colon U \to \Phi(E) \otimes \Psi(E)$, whose derivative is

$$\mathrm{d}_x(f \otimes g)(A) = \mathrm{d}_x f(A) \otimes g(x) + f(x) \otimes \mathrm{d}_x g(A)$$

[1] Formally, we have a single requirement: Φ should be a functor from the category of vector spaces and their isomorphisms to itself.

(express everything in some bases of the $\Phi(E)$ and $\Psi(E)$, observe that the coordinates of $f \otimes g$ are actual products of the coordinates of f and g and thus, the standard Leibniz rule applies; alternatively, avoiding coordinates, tensor multiply the defining relations for the differential $f(x+\xi) = f(x) + \mathrm{d}_x f(\xi) + o(\xi)$ and the same for g and observe that $\mathrm{d}_x f(\xi) \otimes \mathrm{d}_x g(\xi) \in o(\xi)$). This yields immediately

$$D_A(F \otimes G) = D_A F \otimes G + F \otimes D_A G.$$

Now suppose that $F\colon U \to (T^*U)^{\otimes k}$ is a tensor field of type $(0,k)$. Then the evaluation map

$$\mathrm{ev}\colon (V^*)^{\otimes k} \otimes V^{\otimes k} \to \mathbb{R}$$

is linear and natural and yields

$$\begin{aligned} D_A(\mathrm{ev}(F \otimes X_1 \otimes \cdots \otimes X_k)) &= \mathrm{ev}(D_A(F \otimes X_1 \otimes \cdots \otimes X_k)) \\ &= \mathrm{ev}(D_A F \otimes X_1 \otimes \cdots \otimes X_k) \\ &\quad + \sum_i \mathrm{ev}(F \otimes X_1 \otimes \cdots \otimes D_A X_i \otimes \cdots \otimes X_k) \end{aligned}$$

Expanding the evaluation map, this reads

$$D_A(F(X_1, \ldots, X_k)) = (D_A F)(X_1, \ldots, X_k) + \sum_i F(X_1, \ldots, D_A X_i, \ldots, X_k)$$

The term on the left is the directional derivative. In particular, we will need shortly a formula for the derivative of a tensor field g of type $(0,2)$:

$$(D_A g)(X,Y) = Ag(X,Y) - g(D_A X, Y) - g(X, D_A Y).$$

A similar result holds for tensor fields of type $(1,k)$.

6.3. **Covariant derivative for submanifolds of Euclidean spaces.** We start with the following situation. Let $M \subseteq E$ be a submanifold. Then we have the following concepts available in M: parallel transport and covariant derivative. We start with the parallel transport which we find more intuitive. Let $\gamma\colon \mathbb{R} \to M$ be a path and $X\colon \mathbb{R} \to TM$ be a vector field along γ, i.e. we assume $X(t) \in T_{\gamma(t)}M$. We say that X *transports parallelly* along γ in M if $\dot{X}(t)$ is perpendicular to $T_{\gamma(t)}M$. Denoting by $P(x)$ the orthogonal projection $T_x E \to T_x M$, this means $P(\gamma(t))\dot{X}(t) = 0$.

Denoting $\nabla_X Y = P(D_X Y)$, the condition of the parallel transport is thus $\nabla_{\dot{\gamma}} X = 0$. Since we have $D_X Y - D_Y X = [X,Y]$ and $[X,Y]$ is tangent to M if both X and Y are (so that $[X,Y]$ is preserved by P), we obtain

$$\nabla_X Y - \nabla_Y X = [X,Y].$$

We say that the covariant derivative ∇ is symmetric. The second property that we need is that $\nabla g = 0$ where g is the metric on M (we say that g is *covariantly constant*). First, we have to introduce the covariant derivative $\nabla_X g$. This is done by postulating the formula from the previous section,

$$(\nabla_X g)(Y,Z) = X\langle Y,Z\rangle - \langle \nabla_X Y, Z\rangle - \langle Y, \nabla_X Z\rangle.$$

Now we claim that this is zero: this follows from the same rule for D (here g corresponds to the constant map $U \to E^* \otimes E^*$, taking every point to the given scalar product in E, yielding $D_X g = 0$) with the projection P applied (since Z is tangent to M, its product with $D_X Y$ is the same as with $\nabla_X Y = P D_X Y$).

6.4. **Riemannian manifolds and linear connections.**

Definition 23. A *Riemannian metric* on a smooth manifold M is a choice of a scalar product on each $T_x M$ that depends smoothly on $x \in M$. In detail, it is a tensor field of type $(0,2)$, i.e. a smooth map $g\colon M \to (T^*M)^{\otimes 2}$, that is symmetric and positive definite at each point (i.e. each $g_x \in (T^*M)^{\otimes 2}$ should be symmetric and positive definite).

Definition 24. A *Riemannian manifold* is a manifold M equipped with a Riemannian metric.

Example 25. The Euclidean space with the constant field g. Any submanifold $M \subseteq E$ of a Euclidean space E with the restriction of the scalar product on E to M.

Definition 26. A *linear connection* on a manifold M is a mapping

$$\nabla\colon \mathfrak{X}M \times \mathfrak{X}M \to \mathfrak{X}M,$$

denoted $(X, Y) \mapsto \nabla_X Y$, satisfying the conditions

$$\begin{aligned} \nabla_X(Y_1 + Y_2) &= \nabla_X Y_1 + \nabla_X Y_2, \\ \nabla_X(gY) &= Xg \cdot Y + g\nabla_X Y, \\ \nabla_{X_1+X_2} Y &= \nabla_{X_1} Y + \nabla_{X_2} Y, \\ \nabla_{fX} Y &= f\nabla_X Y. \end{aligned}$$

Example 27. The local covariant derivative D on an open subset of a vector space (all properties are trivial). The covariant derivative $\nabla_X Y = PD_X Y$ on a submanifold of a Euclidean space – the only non-trivial axiom is the second one (apply the projection P to the equality $D_x(gY) = Xg \cdot Y + gD_X Y$ and observe that the second term belongs to TM so that it is preserved by P).

Definition 28. We say that a connection ∇ is *symmetric* is $\nabla_X Y - \nabla_Y X = [X, Y]$.

Definition 29. A connection ∇ on a Riemannian manifold M is *metric* if

$$X\langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle.$$

The conditions of a connection imply that it is local (i.e. $\nabla_X Y(x)$ depends only on the germs of X and Y at x). Let now $D_X Y$ again denote the usual local covariant derivative in a given coordinate system on M. Then $\nabla_X Y - D_X Y$ is $C^\infty M$-linear in both X and Y and, thus, given by a tensor field Γ of type $(1, 2)$, i.e.

$$\nabla_X Y = D_X Y + \Gamma(X, Y).$$

The symmetry of ∇ translates to $\Gamma(X, Y) = \Gamma(Y, X)$, i.e. the symmetry of Γ, and the metricity of ∇ translates to

$$\begin{aligned} 0 &= X\langle Y, Z\rangle - \langle D_X Y + \Gamma(X, Y), Z\rangle - \langle Y, D_X Z + \Gamma(X, Z)\rangle \\ &= (D_X g)(Y, Z) - \langle \Gamma(X, Y), Z\rangle - \langle Y, \Gamma(X, Z)\rangle, \end{aligned}$$

i.e. $\langle \Gamma(X, Y), Z\rangle + \langle Y, \Gamma(X, Z)\rangle = (D_X g)(Y, Z)$. According to the following lemma, this determines $\langle \Gamma(X, Y), Z\rangle$ uniquely and, thus, also $\Gamma(X, Y)$, so that we obtain the following theorem.

Theorem 30. *There exists a unique symmetric and metric connection on a given Riemannian manifold – it is called the* Levi-Civita connection*.* □

Lemma 31. *The map* $\mathrm{sym}_{23}\colon (S^2V \otimes V)^* \to (V \otimes S^2V)^*$*, given by* $\mathrm{sym}_{23}\,\omega(X, Y, Z) = \omega(X, Y, Z) + \omega(X, Z, Y)$ *is an isomorphism.*

Proof. The spaces have the same dimensions; thus, it is enough to show that the kernel is zero. But any $\omega \in \ker \mathrm{sym}_{23}$ is symmetric in the first two and antisymmetric in the last two variables, hence zero. □

Remark. In fact, it is not difficult to show that the inverse is given by

$$((\mathrm{sym}_{23})^{-1}\theta)(X, Y, Z) = \tfrac{1}{2}(\theta(X, Y, Z) + \theta(Y, X, Z) - \theta(Z, X, Y)).$$

However, we will not make use of this formula.

6.5. **Parallel transport, geodesics.** The equation for a parallel transport is

$$0 = \nabla_{\dot\gamma} X = D_{\dot\gamma} X + \Gamma(\dot\gamma, X),$$

i.e. $\dot X = -\Gamma(\dot\gamma, X)$. This is a differential equation and, locally, a unique solution exists through each $X(0)$. However, since the solution exists globally for the zero vector, it must exist for any small vector and then for any vector since the parallel transport is clearly linear – any linear combination (with constant coefficients) of parallel vector fields is also parallel.

Another observation is that if both X and Y transport parallelly along γ then

$$\nabla_{\dot\gamma}\langle X, Y\rangle = \langle \nabla_{\dot\gamma} X, Y\rangle + \langle X, \nabla_{\dot\gamma} Y\rangle = 0$$

and the scalar product $\langle X, Y\rangle$ is constant along γ – we say that the parallel transport preserves the scalar product (in fact, this is equivalent to the metricity of ∇).

We denote by Pt_t^γ the map $T_{\gamma(t_0)}M \to T_{\gamma(t_0+t)}M$ obtained by transporting vectors parallelly along γ. We have thus proved that each Pt_t^γ is orthogonal.

A geodesic path is a path γ such that $\dot\gamma$ transports parallelly along γ. This reads

$$\ddot\gamma = -\Gamma(\dot\gamma, \dot\gamma)$$

and is a differential equation of second order. Again, locally, a unique solution exists with any given $A = \dot\gamma(0) \in T_{\gamma(0)}M$. We will temporarily denote it γ_A. Then it is pretty much clear that

$$\gamma_{sA}(t) = \gamma_A(st).$$

Thus, denoting $\exp A = \gamma_A(1)$, we obtain $\gamma_A(t) = \gamma_{tA}(1) = \exp tA$. The map $\exp\colon TM \to M$ is not defined globally; however, it is defined in a neighbourhood of the zero section of TM, since $\exp 0_x = x$. Each $\exp_x\colon T_xM \to M$ is a local diffeomorphism at 0_x (since its derivative is the identity – $\exp_{x*0_x} A = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \exp_x(tA) = A$).

We will now show how the covariant derivative can be reconstructed from the parallel transport. Let e_i be an orthonormal frame at x and transport it parallelly along a path γ through x. Then we get vector fields E_i along γ and they will still be orthonormal since parallel transport preserves scalar product. Let X be a vector field along γ and express it in this orthonormal frame as $X = f^i E_i$. Then

$$\nabla_{\dot\gamma} X = \nabla_{\dot\gamma}\left(f^i E_i\right) = \nabla_{\dot\gamma} f^i \cdot E_i + f^i \cdot \underbrace{\nabla_{\dot\gamma} E_i}_{0} = \nabla_{\dot\gamma} f^i \cdot E_i.$$

In other words, expressing the vector field in a parallel orthonormal frame makes it into a function $f\colon \mathbb{R} \to \mathbb{R}^n$ and then the covariant derivative is simply the usual derivative $\frac{\mathrm{d}}{\mathrm{d}t} f$.

We remark that $f^i(t) e_i = \mathrm{Pt}_{-t}^\gamma X(t)$ so that $\nabla_{\dot\gamma(0)} X = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \mathrm{Pt}_{-t}^\gamma X(t)$. In plain words, transporting the vector field X along γ to $\gamma(0)$ produces a path of vectors in $T_{\gamma(0)}M$ and $\nabla_{\dot\gamma(0)} X$ is then the usual derivative at zero of this function $\mathbb{R} \to T_xM$ with values in a vector space. This will be important in the next section.

6.6. **Second covariant derivative.** Let us compute the second covariant derivative of a field

$$\nabla^2 F(X, Y) = \nabla\nabla F(X, Y) = \nabla_X(\nabla F(Y)) - \nabla F(\nabla_X Y) = \nabla_X \nabla_Y F - \nabla_{\nabla_X Y} F.$$

Now define the antisymmetric version of $\nabla^2 F$ as

$$\nabla^2_{\mathrm{alt}} F(X, Y) = \nabla^2 F(X, Y) - \nabla^2 F(X, Y) = \nabla_X \nabla_Y F - \nabla_Y \nabla_X F - \nabla_{\nabla_X Y - \nabla_Y X} F,$$

where the last term becomes $\nabla_{[X,Y]} F$ for a symmetric connection. There are two important special cases – that of functions where $\nabla^2_{\mathrm{alt}} f(X, Y) = \nabla_{[X,Y] - \nabla_X Y - \nabla_Y X} f$, i.e. the derivative of f in the direction of (minus) the torsion (which is zero for symmetric connections – clearly, this is equivalent to the symmetry of ∇), and that of vector fields where $\nabla^2_{\mathrm{alt}} Z(X, Y)$ is the so-called curvature.

By definition, $\nabla^2_{\mathrm{alt}} F(X, Y)$ is $C^\infty M$-linear in both X and Y. We will now show that $\nabla^2_{\mathrm{alt}} F$ is $C^\infty M$-linear in F. Thus, let h be a function and compute

$$\begin{aligned}
\nabla^2(hF)(X, Y) &= \nabla_X \nabla_Y (hF) - \nabla_{\nabla_X Y}(hF) \\
&= \nabla_X(\nabla_Y h \cdot F + h \cdot \nabla_Y F) - \nabla_{\nabla_X Y} h \cdot F - h \cdot \nabla_{\nabla_X Y} F \\
&= \nabla_X \nabla_Y h \cdot F + \nabla_Y h \cdot \nabla_X F + \nabla_X h \cdot \nabla_Y F + h \cdot \nabla_X \nabla_Y F - \nabla_{\nabla_X Y} h \cdot F - h \cdot \nabla_{\nabla_X Y} F \\
&= h \cdot \nabla^2 F(X, Y) + \nabla^2 h(X, Y) \cdot F + \nabla_Y h \cdot \nabla_X F + \nabla_X h \cdot \nabla_Y F
\end{aligned}$$

and both $\nabla^2 h$ and the remaining terms disappear after antisymmetrization, yielding

$$\nabla^2_{\mathrm{alt}}(hF)(X, Y) = h \nabla^2_{\mathrm{alt}} F(X, Y)$$

or simply $\nabla^2_{\mathrm{alt}}(hF) = h \nabla^2_{\mathrm{alt}} F$.

This means that the alternating second derivative is "algebraic" in the sense that it is a value of a field on (X, Y, F). In fact, it turns out that $TM \otimes T^*M$ has a natural action[2] on $\Phi(TM)$ and $\nabla^2_{\mathrm{alt}} F(X, Y)$ is simply the action of $R(X, Y)$ on F (here $R(X, Y)$ is the curvature tensor evaluated in two arguments, i.e. still a tensor field of type $(1, 1)$; it could be viewed as the map $Z \mapsto R(X, Y)Z$).

Since ∇^2_{alt} is zero on functions (the symmetry on ∇) and since the local covariant derivative of any field is obtained by differentiating covariantly the coordinates of the field (and these are functions!), we see that the curvature of a local covariant derivative is zero.

6.7. **Curvature.** The curvature is defined to be $R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$. It is a tensor of type $(1, 3)$ that is clearly antisymmetric in the first two variables. We have seen that the curvature of the Euclidean covariant derivative is zero. In fact, this characterizes the Euclidean connection, as the following theorem says.

Before going into the proof, we give a geometric meaning to the curvature. Let X, Y be vector fields that commute, i.e. such that $[X, Y] = 0$. Then $\nabla_X \nabla_Y Z(x)$ is obtained as the mixed partial derivative $\frac{\partial^2}{\partial s \partial t} A(0, 0)$ of the vector valued function $A(s, t) \in T_x M$ given by transporting parallelly $Z(\mathrm{Fl}^Y_t(\mathrm{Fl}^X_s(x)))$ along the flow line of Y back to $\mathrm{Fl}^X_s(x)$ and then along the flow line of X back to x. A similar formula holds for the second term. We may however define Z by first transporting $Z(0, 0) \in T_x M$ along the flow line of Y and then along the flow lines of X so that the second term actually becomes zero. Thus, we finally obtain

$$R(X, Y)Z = \frac{\partial^2}{\partial s \partial t}\Big|_{(s,t)=(0,0)} \mathrm{Pt}^{\mathrm{Fl}^X}_{-s} \, \mathrm{Pt}^{\mathrm{Fl}^Y}_{-t} \, \mathrm{Pt}^{\mathrm{Fl}^X}_{s} \, \mathrm{Pt}^{\mathrm{Fl}^Y}_{t} \, Z$$

(it would look slightly better to change both s, t to their opposites – then this becomes a commutator).

Continuing the notation of the above proof, we will show that for $R(X, Y)Z = 0$ and $[X, Y] = 0$, the parallel transports along the flow lines of X and Y commute: $0 = R(X, Y)Z = \nabla_X \nabla_Y Z$, so that $\nabla_Y Z$ transports parallelly along the flow lines of X. Since $\nabla_Y Z = 0$ for $s = 0$, it must be zero everywhere, i.e. Z also transports parallelly along the flow lines of Y. In particular, we obtain $Z(s, t)$ also by transporting $Z(0, 0)$ first parallelly along the flow line of X to get $Z(0, t)$ and then parallelly along the flow line of Y (this is what we have just proved).

Theorem 32. *The following conditions are equivalent.*

1. *The curvature is zero.*
2. *The parallel transport does not locally depend on the path.*
3. *There is an atlas in which all the Γ are zero.*
4. *There is an atlas consisting of isometries.*

Proof. We will prove (1)⇒(2)⇒(3)⇒(4)⇒(1).

(1)⇒(2): We use the fact that the parallel transports along vector fields X, Y commute when $[X, Y] = 0$. Start with a vector $Z_0 \in T_x M$ and transport it parallelly along the local vector fields $\partial_1, \ldots, \partial_n$ to obtain a vector field Z with $Z(x) = Z_0$. Since Z was obtained by parallel transport along ∂_i (any one could have been used the last), we have $\nabla_{\partial_i} Z = 0$. This holds for any i and, thus, $\partial_X Z = 0$ for any X. In particular, Z transports parallelly along any path, implying that the parallel transport of $Z(x)$ along a path from x to y is always $Z(y)$.

(2)⇒(3): Suppose that the parallel transport does not locally depend on the path. Start with a basis (e_i) of $T_x M$ and transport it locally to a neighbourhood to obtain vector fields E_i. Then $[E_i, E_j] = \nabla_{E_i} E_j - \nabla_{E_j} E_i = 0$ and there exists a chart with $E_i = \partial_i$. In particular, $\Gamma(E_i, E_j) = \nabla_{E_i} E_j - D_{E_i} E_j = 0$ and $\Gamma = 0$.

(3)⇒(4): Clearly, to obtain a local isometry, it is enough to have $D_X g = 0$ for all X (then g is constant and we may modify the chart by a linear isomorphism). But

$$0 = (\nabla_X g)(Y, Z) = (D_X g)(Y, Z) - \langle \Gamma(X, Y), Z \rangle - \langle Y, \Gamma(X, Z) \rangle = (D_X g)(Y, Z).$$

(4)⇒(1) is clear since we have $\nabla_X Y = D_X Y$ in a Euclidean space and the curvature is zero. □

[2] It is the derivative at the identity of the map $GL(T_x M) \to GL(\Phi(T_x M))$, $\alpha \mapsto \Phi(\alpha)$, given by the functor Φ.

6.8. **Remarks about covariant derivative of other fields.** The above point of view may also be applied to other fields. Let Φ be a functor from the category of vector spaces and their isomorphisms to itself[3]. By a field F of type Φ, we will understand an association $M \ni x \mapsto F(x) \in \Phi(T_xM)$ that will be smooth in the following sense: a local chart $\varphi\colon U \to \mathbb{R}^n$ gives trivializations $\varphi_{*x}\colon T_xM \xrightarrow{\cong} \mathbb{R}^n$ of each tangent space T_xM, $x \in U$, and applying the functor Φ then gives a map (the expression of the field in local coordinates)

$$\begin{aligned} U &\to \Phi(\mathbb{R}^n) \\ x &\mapsto \Phi(\varphi_{*x})(F(x)) \end{aligned}$$

that should be smooth.[4] Now we apply this in the same way as above, i.e. in the situation that γ is a path and F a field along γ. Choose a parallel frame (E_i) that gives trivializations $\alpha_t\colon T_{\gamma(t)}M \xrightarrow{\cong} \mathbb{R}^n$ (different from those above) and use them to translate the field F to a function $f\colon \mathbb{R} \to \Phi(\mathbb{R}^n)$, $t \mapsto \Phi(\alpha_t)(F(t))$. Then we take the usual derivative $\frac{\mathrm{d}}{\mathrm{d}t}f\colon \mathbb{R} \to \Phi(\mathbb{R}^n)$ and translate it back to a field $\nabla_{\dot\gamma}F$ along γ, i.e.

$$\nabla_{\dot\gamma(t)}F = \Phi(\alpha_t)^{-1}(\tfrac{\mathrm{d}}{\mathrm{d}t}f(t)).$$

A different parallel frame differs by a *constant* (and orthogonal) transformation $(E_i) = (\widetilde{E}_i)P$ which corresponds to an automorphism $P\colon \mathbb{R}^n \to \mathbb{R}^n$ satisfying $\widetilde{\alpha}_t = P\alpha_t$. This implies $\widetilde{f} = \Phi(P)f$ and, since $\Phi(P)$ is linear, we obtain $\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{f}(t) = \Phi(P)\frac{\mathrm{d}}{\mathrm{d}t}f$, yielding finally the independence of $\nabla_{\dot\gamma}F$ on the choice of the frame.

There are three simple rules that allow one to compute the covariant derivative of the standard fields.

The zeroth one is that a function is a field of type $\mathbb{R}$, the constant functor taking every vector space to $\mathbb{R}$ and every isomorphism to the identity. Then, independently of the frame, we obtain $f = F$ and $\nabla_{\dot\gamma}F = \frac{\mathrm{d}}{\mathrm{d}t}F$. For a global function (as opposed to a function along γ which then happens to be just $F \circ \gamma$), we obtain $\nabla F = \mathrm{d}F$, the differential of F (i.e. $\nabla_A F = \mathrm{d}F(A) = AF$, the derivative of F in the direction of the vector A).

The first one is that linear natural transformations commute with the differentiation, i.e. if $\tau = \tau_V\colon \Phi(V) \to \Psi(V)$ is linear and natural then we define, for a field F of type Φ, a field τF of type Ψ, as the composition $t \mapsto \tau(F(t))$. We obtain easily

$$\nabla(\tau F) = \tau\nabla F.$$

The second concerns the tensor product of fields. Given functors Φ and Ψ, the tensor products $\Phi(V) \otimes \Psi(V)$ clearly form another functor denoted $\Psi \otimes \Psi$. Given fields F and G of types Φ and Ψ respectively, the association $x \mapsto F(x) \otimes G(x) \in \Phi(T_xM) \otimes \Psi(T_xM)$ is a field of type $\Phi \otimes \Psi$ and the corresponding function is $f \otimes g\colon \mathbb{R} \to \Phi(\mathbb{R}^n) \otimes \Psi(\mathbb{R}^n)$; its derivative is

$$\mathrm{d}_t(f \otimes g)(A) = \mathrm{d}_t f(A) \otimes g(t) + f(t) \otimes \mathrm{d}_t g(A)$$

(by tensoring the defining relations for the differential $f(t+\tau) = f(t) + \mathrm{d}_t f(\tau) + o(\tau)$ and the same for g, observing that $\mathrm{d}_t f(\tau) \otimes \mathrm{d}_t g(\tau) \in o(\tau)$). This yields immediately

$$\nabla(F \otimes G) = \nabla F \otimes G + F \otimes \nabla G.$$

Now suppose that ω is a field of type $\hom(\Phi^{\otimes k}, \Psi)$. Then the evaluation map

$$\mathrm{ev}_V\colon \ \hom(\Phi(V)^{\otimes k}, \Psi(V)) \otimes \Phi(V)^{\otimes k} \to \Psi(V)$$

[3]It is useful to think of this as a way of producing out of coordinates $V \xrightarrow{\cong} \mathbb{R}^n$ on V some sort of coordinates $\Phi(V) \xrightarrow{\cong} \Phi(\mathbb{R}^n)$ on $\Phi(V)$, e.g. $\Lambda^k V^* \xrightarrow{\cong} \Lambda^k(\mathbb{R}^n)^*$, $\omega \mapsto \sum_{i_1<\cdots<i_k} \omega_{i_1\cdots i_k} f^{i_k} \wedge \cdots \wedge f^{i_1}$

[4]More generally, we may equivalently ask for the expression F in any field of frames (E_i) (simply an n-tuple of vector fields that provides a basis at each point) to be smooth; such a frame does not necessarily come from a local chart (e.g. we may require an orthonormal frame and it may be impossible to obtain one from a chart).

is linear and natural and yields

$$\begin{aligned}\nabla_X(\operatorname{ev}(\omega\otimes X_1\otimes\cdots\otimes X_k)) &= \operatorname{ev}\nabla_X(\omega\otimes X_1\otimes\cdots\otimes X_k)\\ &= \operatorname{ev}(\nabla_X\omega\otimes X_1\otimes\cdots\otimes X_k)\\ &\quad+\sum_i \operatorname{ev}(\omega\otimes X_1\otimes\cdots\otimes\nabla_X X_i\otimes\cdots\otimes X_k)\end{aligned}$$

When the evaluation map is expanded, this reads

$$\nabla_X(\omega(X_1,\dots,X_k)) = (\nabla_X\omega)(X_1,\dots,X_k)+\sum_i \omega(X_1,\dots,\nabla_X X_i,\dots,X_k)$$

Let us compute as an example the second covariant derivative of any field

$$\nabla^2F(X,Y)=\nabla\nabla F(X,Y)=\nabla_X(\nabla F(Y))-\nabla F(\nabla_X Y)=\nabla_X\nabla_Y F-\nabla_{\nabla_X Y}F.$$

Now define the antisymmetric version of $\nabla^2 F$ as

$$\nabla^2_{\mathrm{alt}}F(X,Y)=\nabla^2F(X,Y)-\nabla^2F(X,Y)=\nabla_X\nabla_Y F-\nabla_Y\nabla_X F-\nabla_{\nabla_X Y-\nabla_Y X}F,$$

where the last term becomes $\nabla_{[X,Y]}F$ for a symmetric connection. There are two important special cases – that of functions where $\nabla^2_{\mathrm{alt}}f(X,Y)=\nabla_{[X,Y]-\nabla_X Y-\nabla_Y X}f$, i.e. the derivative of f in the direction of (minus) the torsion, and that of vector fields where $\nabla^2_{\mathrm{alt}}Z(X,Y)=R(X,Y)Z$, i.e. the curvature (at least when torsion is zero).

7. Spaces of constant curvature

7.1. **Sectional curvature.** First we observe that $\langle R(X,Y)Z,W\rangle$ is also anti-symmetric in the variables Z, W. This follows from

$$\begin{aligned}0&=(\nabla^2_{\mathrm{alt}}\langle Z,W\rangle)(X,Y)=\nabla_X\nabla_Y\langle Z,W\rangle-\nabla_Y\nabla_X\langle Z,W\rangle-\nabla_{\nabla_X Y-\nabla_Y X}\langle Z,W\rangle\\ &=\langle R(X,Y)Z,W\rangle+\langle Z,R(X,Y)W\rangle\end{aligned}$$

(the terms where each Z, W receives one of the ∇_X, ∇_Y cancel out).

Let p be a 2-dimensional vector subspace of T_xM with orthonormal basis (e_1,e_2). We define the *sectional curvature* $K(p)=\langle R(e_1,e_2)e_2,e_1\rangle$. Denoting $R(X,Y,Z,W)=\langle R(X,Y)W,Z\rangle$, we see that this depends only on $X\wedge Y$ and $Z\wedge W$. Thus, when replacing these by a linear combination, the whole expression gets multiplied by the product of the determinants of the transformation matrices. In the case of $K(p)$, this means the square of an orthogonal matrix and the value does not change.

7.2. **Sphere.** We compute the sectional curvature of a sphere $S^n\subseteq\mathbb{R}^{n+1}$. Thus, let $e_1,e_2\in T_{x_0}S^n$ be two orthonormal vectors. We extend them to vector fields on S^n in the following way: think of the e_i as a constant vector field on $\mathbb{R}^{n+1}$ and project it orthogonally to obtain a vector field E_i on TS^n; at a point x, this equals

$$E_i(x)=e_i-\langle x,e_i\rangle x$$

In fact, this formula prescribes a vector field on $\mathbb{R}^{n+1}$ – this is useful since we want to apply the covariant derivative of $\mathbb{R}^{n+1}$:

$$D_AE_i=-\langle A,e_i\rangle x-\langle x,e_i\rangle A$$

Projecting to TS^n, the first term becomes zero and the second term remains unchanged (since A is now assumed tangent to S^n), i.e.

$$\nabla_AE_i=-\langle x,e_i\rangle A$$

This leads to

$$[E_i,E_j]=\nabla_{E_i}E_j-\nabla_{E_j}E_i=\langle x,e_i\rangle E_j-\langle x,e_j\rangle E_i$$

and finally

$$R(E_i,E_j)E_j=\nabla_{E_i}\underbrace{\nabla_{E_j}E_j}_{-\langle x,e_j\rangle E_j}-\nabla_{E_j}\underbrace{\nabla_{E_i}E_j}_{-\langle x,e_j\rangle E_i}-\nabla_{[E_i,E_j]}E_j$$

Now the derivative of $\langle x, e_j\rangle$ in the direction of E_i equals $\langle E_i, e_j\rangle$, since the function is linear in x. Thus,

$$\begin{aligned} R(E_i,E_j)E_j = &-\langle x, e_j\rangle \cdot \underbrace{\nabla_{E_i}E_j}_{-\langle x,e_j\rangle E_i} - \langle E_i, e_j\rangle E_j \\ &+ \langle x, e_j\rangle \cdot \underbrace{\nabla_{E_j}E_i}_{-\langle x,e_i\rangle E_j} + \langle E_j, e_j\rangle E_i \\ &+ \langle x, e_j\rangle(\langle x, e_i\rangle E_j - \langle x, e_j\rangle E_i) \end{aligned}$$

(the function $\langle x, e_j\rangle$ is linear in x) so that

$$\langle R(E_i,E_j)E_j, E_i\rangle(x_0) = -\langle e_i, e_j\rangle\langle e_j, e_i\rangle + \langle e_i, e_i\rangle\langle e_j, e_j\rangle = 1 \cdot \mathrm{Vol}(e_i, e_j)^2$$

as $\langle x_0, e_j\rangle = 0$, since $e_j \in T_{x_0}S^n$.

7.3. **Hyperbolic space.** We compute the sectional curvature of a hyperbolic space $H^n \subseteq \mathbb{R}^{n+1}$ equipped with a metric $g = -\mathrm{d}x^0 \otimes \mathrm{d}x^0 + \mathrm{d}x^1 \otimes \mathrm{d}x^1 + \cdots + \mathrm{d}x^n \otimes \mathrm{d}x^n$, where

$$H^n = \{x \in \mathbb{R}^{n+1} \mid g(x,x) = -1\}.$$

Since each $x \in H^n$ generates a 1-dimensional subspace where the metric is negative definite, it is easy to see that on its orthogonal complement – and this is again T_xH^n – the metric must be positive definite (the inertia theorem). Thus, let $e_1, e_2 \in T_{x_0}H^n$ be two orthonormal vectors. We extend them to vector fields on H^n, this times the formula is

$$E_i(x) = e_i + \langle x, e_i\rangle x$$

(because $\langle x, x\rangle = -1$). Again, this prescribes a vector field on $\mathbb{R}^{n+1}$ and:

$$D_A E_i = \langle A, e_i\rangle x + \langle x, e_i\rangle A$$

Projecting to TS^n, we get

$$\nabla_A E_i = \langle x, e_i\rangle A$$

This leads to the same formula for $R(E_i, E_j)E_j$ as above, only with different signs. Since the surviving terms in $\langle R(E_i, E_j)E_j, E_i\rangle$ contain exactly one scalar product from $R(E_i, E_j)E_j$, we get

$$\langle R(E_i,E_j)E_j, E_i\rangle(x_0) = \langle e_i, e_j\rangle\langle e_j, e_i\rangle - \langle e_i, e_i\rangle\langle e_j, e_j\rangle = -1 \cdot \mathrm{Vol}(e_i, e_j)^2.$$