Spaces and geometric objects

We begin by recalling the notion of a manifold and some of the structures associated with it that shall be used throughout this course. Our presentation is fairly standard but borrows from [4], [13] and [2]. Our definition of manifold is a slightly modified, more modern form of the Veblen-Whitehead axioms [12], which initially assumes a manifold to be a set, the topological structure being induced by the presence of coordinate systems (see also [1, Ch. 5]).

1. Manifolds. Throughout this course, we will concern ourselves with spaces that locally look like Euclidean space and permit us to freely use familiar tools from calculus — manifolds. We briefly recall their definition.

Let M be a point set. A (smooth) n-dimensional atlas on M is a collection of bijections

$$\{\varphi_\alpha : U_\alpha \to \varphi_\alpha(U_\alpha) \subset \mathbb{R}^n\}_{\alpha \in A} \tag{1.1}$$

called *charts* or *coordinate systems*[1] such that $M = \bigcup_{\alpha\in A} U_\alpha$, $(\forall \alpha, \beta)$ $\varphi_\alpha(U_\alpha \cap U_\beta) \subset \mathbb{R}^n$ is open, and the *coordinate transformations*

$$\varphi_\alpha \circ \varphi_\beta^{-1} : \varphi_\beta(U_\alpha \cap U_\beta) \to \varphi_\alpha(U_\alpha \cap U_\beta)$$

are *smooth.* For brevity, we shall write $\{\varphi_\alpha|_{U_\alpha}\}_{\alpha\in A}$ in place of (1.1) and refer to a chart as $\varphi_\alpha|_{U_\alpha}$ whenever we need to emphasise both the mapping and its domain.

We shall say that two n-dimensional atlases $\mathcal{A}_1, \mathcal{A}_2$ on M are (smoothly) *equivalent*, written $\mathcal{A}_1 \sim \mathcal{A}_2$, if $\mathcal{A}_1 \cup \mathcal{A}_2$ is also an n-dimensional atlas. It is not difficult to check that $\sim$ is an equivalence relation. We shall write $[\mathcal{A}]$ for the equivalence class of a given atlas $\mathcal{A}$ and define the *maximal atlas* containing $\mathcal{A}$ as

$$\mathcal{A}^{\max} := \bigcup_{\mathcal{B}\in[\mathcal{A}]} \mathcal{B}.$$

Definition 1.1. A *smooth manifold* of dimension n is a set M together with the equivalence class of an n-dimensional atlas $\mathcal{A}$ such that $\mathcal{A} = \{\varphi_i|_{U_i}\}$ is countable and satisfies the *Hausdorff axiom*: For all $p, q \in M$, there exist $i, j \in I$ and disjoint sets $p \in N(p) \subset U_i$, $q \in N(q) \subset U_j$ such that $\varphi_i(N(p))$ and $\varphi_j(N(q))$ are open.

Notation 1.2. We will sometimes write $\dim M$ for the dimension of a smooth manifold M and also M^n in place of M to emphasise that it is n-dimensional.

Let M be an n-dimensional manifold with countable atlas $\mathcal{A} = \{\varphi|_{U_i}\}_{i\in I}$. If we set

$$\tau_M := \{U \subset M : (\forall i)\ \varphi_i(U_i \cap U) \text{ open}\},$$

it follows immediately that (M, τ_M) is a *Hausdorff space.* Moreover, the countability of $\mathcal{A}$ and second countability of $\mathbb{R}^n$ imply that M is second countable, thus also *Lindelöf.* In particular, combining all of these properties with the existence of compact neighbourhoods furnished by coordinate systems, we conclude that M is normal and *paracompact.* These topological conditions ensure the validity of the following lemma, which will come in handy.

[1]The inverse of a chart will be referred to as a *local parametrisation.*

Lemma 1.3 (partitions of unity). Let $\{U_\alpha\}_{\alpha\in A}$ be an open cover of (M, τ_M). Then there exists a countable family of smooth functions $\{\phi_i : M \to [0,\infty[\}_{i\in I}$ such that

- for each $p \in M$, there exists an open set $U_p \ni p$ such that $\phi_i|_{U_p} \equiv 0$ for all but finitely many i;
- $\sum_i \phi_i = 1$; and
- for each i, there exists an $\alpha \in A$ such that $\operatorname{supp} \phi_i \Subset U_\alpha$.

The natural topology on M gives us a notion of continuity of mappings between manifolds. We now make precise the notion of differentiability.

Definition 1.4. Let $(M_1, [\mathcal{A}_1])$ and $(M_2, [\mathcal{A}_2])$ be smooth manifolds and $U \subset M_1$ open. A mapping $F : U \to M_2$ is said to be *k-times continuously differentiable* (or of class C^k) if for all coordinate systems $\varphi|_{U_1} \in \mathcal{A}_1$ and $\psi|_{U_2} \in \mathcal{A}_2$, $\varphi(U \cap U_1 \cap F^{-1}(U_2))$ is open and the *coordinate representation* of F

$$\psi \circ F \circ \varphi^{-1} : \varphi(U_1 \cap F^{-1}(U_2)) \to \psi(U_2)$$

is k-times continuously differentiable. If U is not open, we say F is k-times continuously differentiable if there exists a C^k function $\widetilde{F} : \widetilde{U} \to M_2$ with $\widetilde{U} \supset U$ open and $\widetilde{F}\big|_U = F$. If F is an invertible C^k mapping with C^k inverse, we call F a (C^k) *diffeomorphism*.

Remark 1.5. The condition on the domain of a coordinate representative ensures that F is continuous. In particular, for any $V \subset M_2$ open, $F^{-1}(V) = \bigcup_{U_2} F^{-1}(V \cap U_2)$ is open. △

We now turn our attention to some examples of manifolds that we shall continually revisit.

Example 1.6 (n-dimensional Euclidean space). The set $\mathbb{R}^n$ together with the equivalence class of $\mathcal{A} = \{\mathrm{id}_{\mathbb{R}^n} : \mathbb{R}^n \to \mathbb{R}^n\}$ defines a smooth n-manifold. More generally, any finite-dimensional $\mathbb{R}$-vector space $V \simeq \mathbb{R}^n$ is naturally a smooth manifold: Fixing a basis $\{b_i\}_{i=1}^n$ of V, we obtain an admissible atlas $\{\varphi_b : V \to \mathbb{R}^n\}$ with $\varphi_b(\sum_i v^i b_i) = \sum_i v^i e_i$.

Example 1.7 (n-dimensional sphere). Let $S^n = \{p \in \mathbb{R}^{n+1} : \sum_{i=1}^n (p^i)^2 = 1\}$ denote the n-dimensional unit sphere and for $\sigma = \pm$ let $P_\sigma = \sigma e_{n+1}$ denote its poles. We obtain the so-called *stereographic projection* of $P \in S^n \backslash \{P_\sigma\}$ onto $\mathbb{R}^n$ by drawing the ray $\overleftrightarrow{P_\sigma P}$ and finding its point of intersection with $\mathbb{R}^n \times \{0\} \simeq \mathbb{R}^n$. This yields the bijections

$$\begin{aligned} \varphi_\sigma : S^n \backslash \{P_\sigma\} &\to \mathbb{R}^n \\ p &\mapsto \frac{\sigma}{\sigma - p^{n+1}}(p^1, \ldots, p^n) \end{aligned}$$

with inverses given by $\varphi_\sigma^{-1}(x) = \sum_{i=1}^n \frac{2x^i}{|x|^2+1} e_i + \sigma\left(\frac{|x|^2-1}{|x|^2+1}\right) e_{n+1}$. It is clear that the coördinate transformation

$$\begin{aligned} \varphi_+ \circ \varphi_-^{-1} : \mathbb{R}^n \backslash \{0\} &\to \mathbb{R}^n \backslash \{0\} \\ x &\mapsto \frac{x}{|x|^2} \end{aligned}$$

is analytic. Moreover, this atlas is countable and satisfies the Hausdorff axiom: $P_\sigma \in \varphi_{-\sigma}^{-1}(B(0,1))$ and $\varphi_+^{-1}(B(0,1)) \cap \varphi_-^{-1}(B(0,1)) = \varnothing$.

Exercise 1.8. Let $S^n_\pm = \{p \in S^n : \pm p^{n+1} \geq 0\}$ denote the (closed) upper/lower hemisphere of S^n. Verify that $\varphi_{-\sigma}(S^n_\sigma) = \overline{B(0,1)}$ for $\sigma \in \{+,-\}$. In particular, $S^n = S^n_+ \cup S^n_-$ is the union of two compact sets, thus also compact.

Example 1.9 (n-dimensional projective space). Let $\mathbb{R}P^n := (\mathbb{R}^n \backslash \{0\})/\!\sim$ where $x \sim y$ iff $x = \lambda y$ for some $\lambda \neq 0$. For each $[x^1, \ldots, x^{n+1}] \in \mathbb{R}P^n$, we must have that at least one $x^i \neq 0$ so that $[x^1, \ldots, x^n] = [x^1/x^i, \ldots, x^{i-1}/x^i, 1, x^{i+1}/x^i, \ldots, x^n/x^i]$. This naturally gives rise to local parametrisations

$$F_i : \mathbb{R}^n \to U_i := F_i(\mathbb{R}^n) \subset \mathbb{R}P^n$$
$$F_i(x^1, \ldots, x^n) = [x^1, \ldots, x^{i-1}, 1, x^i, \ldots, x^n],$$

for each $i \in \{1, \ldots, n+1\}$. The collection $\{F_i^{-1}\}_{i=1}^{n+1}$ forms an atlas and the resulting topology may be checked to be both Hausdorff and second countable.

Example 1.10 (open submanifolds). If M^n is a C^k-manifold with atlas $\mathcal{A}$ and $U \in \tau_M$, then the set U together with the atlas $\mathcal{A}|_U := \{\varphi|_{U \cap U_\alpha} : \varphi_\alpha|_{U_\alpha} \in \mathcal{A}\}$ (check!) yields a C^k-manifold known as an *open submanifold* of M.

Example 1.11 (real general linear group). Let $M_{n\times n}(\mathbb{R})$ be the set of $n \times n$ matrices with entries in $\mathbb{R}$. Using the canonical isomorphism $M_{n\times n}(\mathbb{R}) \simeq \mathbb{R}^{n^2}$, we may view $M_{n\times n}(\mathbb{R})$ as a C^ω n^2-dimensional manifold. Let $\mathrm{GL}(n, \mathbb{R})$ denote the set of all invertible $n \times n$ matrices with entries in $\mathbb{R}$. Since

$$\mathrm{GL}(n, \mathbb{R}) = \{A \in M_{n\times n}(\mathbb{R}) : \det(A) \neq 0\}$$

and the determinant is a continuous function, we see that $\mathrm{GL}(n, \mathbb{R})$ is an open submanifold of $M_{n\times n}(\mathbb{R})$. It is called the (real) *general linear group.*

Example 1.12 (Lie groups). A group G equipped with the structure of a smooth manifold such that the multiplication map $\cdot : G \times G \to G$ and inversion map $(\cdot)^{-1} : G \to G$ are smooth is called a *Lie group.* The general linear group $\mathrm{GL}(n, \mathbb{R})$ is one such Lie group.

Exercise 1.13 (product manifolds). If M_1 and M_2 are smooth manifolds with countable atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ of dimension n_1 and n_2 respectively, then $\mathcal{A}_1 \otimes \mathcal{A}_2 := \{\varphi_{ij} := \varphi_i \times \psi_j : \varphi_i \in \mathcal{A}_1, \psi_j \in \mathcal{A}_2\}$ defines an $(n_1 + n_2)$-dimensional smooth atlas on $M_1 \times M_2$, where for $\varphi_i : U_i \to \varphi_i(U_i)$ and $\psi_j : V_j \to \psi_j(V_j)$ we define

$$\varphi_i \times \psi_j : U_i \times V_j \to \varphi_i(U_i) \times \psi_j(V_j)$$
$$(p_1, p_2) \mapsto (\varphi_i(p_1), \psi_j(p_2)).$$

The resulting $(n_1 + n_2)$-dimensional smooth manifold is called the *product* of M_1 and M_2. Note that the topology induced by $\mathcal{A}_1 \otimes \mathcal{A}_2$ coïncides with the product topology $\tau_{M_1 \times M_2}$ and if $\varphi \in \mathcal{A}_1^{\max}$ and $\psi \in \mathcal{A}_2^{\max}$, then $\varphi \times \psi \in (\mathcal{A}_1 \otimes \mathcal{A}_2)^{\max}$.

The following lemma will be used in some constructions of manifolds.

Lemma 1.14. Let $\{M_i\}_{i\in I}$ be a countable collection of n-dimensional smooth manifolds and S a point set together with a countable cover $\{W_i\}_{i\in I}$ and bijections

$$\psi_i : W_i \to M_i$$

such that for all i, j, $\psi_i(W_i \cap W_j)$ is open and

$$\psi_i \circ \psi_j^{-1} : \psi_j(W_i \cap W_j) \to \psi_i(W_i \cap W_j)$$

is smooth. Furthermore suppose that the following Hausdorff-type axiom is satisfied: For all $p, q \in S$, there exist i, j and disjoint sets $N(p) \subset W_i$ and $N(q) \subset W_j$ with $p \in N(p)$, $q \in N(q)$, and the sets $\psi_i(N(p))$ and $\psi_j(N(q))$ are open. Then there exists a unique equivalence class $[\mathcal{A}]$ of C^k-atlases such that $(S, [\mathcal{A}])$ is a C^k-manifold and the $\{\psi_i\}$ are C^k diffeomorphisms.

Proof. For each M_i, we have a countable C^k atlas $\{(\overset{i}{\varphi}_j, \overset{i}{U}_j)\}_{j\in\mathbb{N}}$. Now, if S is equipped with a C^k atlas with respect to which the $\{\psi_i\}$ are diffeomorphisms, the mappings

$$\mathcal{A} := \{\overset{i}{\varphi}_j \circ \psi_i : W_i \cap \psi_i^{-1}(\overset{i}{U}_j) \to \overset{i}{\varphi}_j(\overset{i}{U}_j) \subset \mathbb{R}^n\}_{i,j\in I}$$

must also be C^k diffeomorphisms, thus constitute a C^k-compatible atlas, since the sets $\{W_i \cap \psi_i^{-1}(\overset{i}{U}_j)\}_{i,j\in\mathbb{N}}$ cover S. Therefore, it suffices to show that $\mathcal{A}$ is a C^k atlas such that $(S, [\mathcal{A}])$ a manifold.

To simplify notation, set $\psi_{ij} := \overset{i}{\varphi}_j \circ \psi_i$. For $i_1, i_2, j_1, j_2 \in \mathbb{N}$, we have that

$$\begin{aligned}&\psi_{i_1 j_1}(W_{i_1} \cap \psi_{i_1}^{-1}(\overset{i_1}{U}_{j_1}) \cap W_{i_2} \cap \psi_{i_2}^{-1}(\overset{i_2}{U}_{j_2}))\\ &\quad = \overset{i_1}{\varphi}_{j_1}(\overset{i_1}{U}_{j_1} \cap \psi_{i_1}(W_{i_1} \cap W_{i_2}) \cap (\psi_{i_1} \circ \psi_{i_2}^{-1})(\psi_{i_2}(W_{i_1} \cap W_{i_2}) \cap \overset{i_2}{U}_{j_2}))\end{aligned}$$

is open since $\overset{i_1}{\varphi}_{j_1}$ is a homeomorphism, and

$$\psi_{i_1 j_1} \circ \psi_{i_2 j_2}^{-1} = \overset{i_1}{\varphi}_{j_1} \circ (\psi_{i_1} \circ \psi_{i_2}^{-1}) \circ \overset{i_2}{\varphi}{}_{j_2}^{-1}$$

is C^k, being a composition of C^k mappings. Thus, $\mathcal{A}$ defines a C^k atlas. To show that S satisfies the Hausdorff axiom, we verify that the topology τ_S is such that

$$\tau_S = \{V \subset S : (\forall i)\ \psi_i(V \cap W_i) \text{ open in } M_i\}.$$

Now, if $V \in \tau_S$, then for all i, j, $(\overset{i}{\varphi}_j \circ \psi_i)(V \cap W_i \cap \psi_i^{-1}(\overset{i}{U}_j))$ is open in $\mathbb{R}^n$. Since the $\{\overset{i}{\varphi}_j\}$ are continuous, this implies that for all i, j $\psi_i(V \cap W_i \cap \psi_i^{-1}(\overset{i}{U}_j))$ is open in M_i, hence so is the union $\psi_i(V \cap W_i)$ over j. Conversely, if $\psi_i(V \cap W_i)$ is open in M_i for all i, then so is $\psi_i(V \cap W_i \cap \psi_i^{-1}(\overset{i}{U}_j))$, thus also $(\overset{i}{\varphi}_j \circ \psi_i)(V \cap W_i \cap \psi_i^{-1}(\overset{i}{U}_j))$ in $\mathbb{R}^n$. The last condition on the $\{\psi_i\}$ now implies that τ_S is Hausdorff. □

In the sequel, we shall refer to an n-dimensional manifold by its underlying point set, the atlas only being made explicit if necessary, and assume it to be equipped with the topology constructured from its atlas as above.

2. The tangent bundle. We now turn our attention to the infinitesimal structure of a smooth manifold in the form of the *tangent bundle*, i.e. the union of the tangent spaces at all points of a manifold, whose definition we now recall.

Let M be an smooth n-dimensional manifold with maximal atlas $\{\varphi_\alpha\}_{\alpha\in A}$. For $\alpha, \beta \in A$, we introduce the shorthand notation

$$\frac{\partial \overset{\alpha}{x}}{\partial \overset{\beta}{x}}(p) := D(\varphi_\alpha \circ \varphi_\beta^{-1})(\varphi_\beta(p)).$$

and define its matrix $(\frac{\partial \overset{\alpha}{x}{}^j}{\partial \overset{\beta}{x}{}^i}(p))_{i,j=1}^n$ such that

$$\frac{\partial \overset{\alpha}{x}}{\partial \overset{\beta}{x}}(p) e_i = \sum_{j=1}^n \frac{\partial \overset{\alpha}{x}{}^j}{\partial \overset{\beta}{x}{}^i}(p) e_j.$$

We form the set

$$TM := \left(\bigcup_{\alpha \in A} \{\alpha\} \times U_\alpha \times \mathbb{R}^n \right) /\!\sim,$$

where the equivalence relation (check!) $\sim$ is defined such that

$$(\alpha, p, \overset{\alpha}{v}) \sim (\beta, q, \overset{\beta}{v}) \Leftrightarrow p = q \text{ and } \overset{\alpha}{v} = \frac{\partial \overset{\alpha}{x}}{\partial \overset{\beta}{x}}(p)\overset{\beta}{v}.$$

We obtain a natural *projection* $\pi : TM \to M$ defined such that $\pi([\alpha, p, \overset{\alpha}{v}]) = p$, and each *fibre* $T_pM := \pi^{-1}(\{p\})$ carries the structure of an $\mathbb{R}$-vector space with the assignments

$$\begin{aligned} \underset{1}{v} + \underset{2}{v} &:= [\alpha, p, \overset{\alpha}{\underset{1}{v}} + \overset{\alpha}{\underset{2}{v}}] \\ 0_p &:= [\alpha, p, 0] \\ c \cdot \underset{1}{v} &:= [\alpha, p, c \cdot \overset{\alpha}{\underset{1}{v}}]. \end{aligned} \tag{1.2}$$

Moreover, for each $\alpha \in A$, we obtain a natural *parametrisation*

$$\begin{aligned} \Phi_\alpha : U_\alpha \times \mathbb{R}^n &\to \pi^{-1}(U_\alpha) \\ (p, v) &\mapsto [\alpha, p, v]. \end{aligned}$$

Since M is Lindelöf and

$$(\Phi_\alpha^{-1} \circ \Phi_\beta)(p, v) = \Phi_\alpha^{-1}([\beta, p, v]) = \Phi_\alpha^{-1}([\alpha, p, \frac{\partial \overset{\alpha}{x}}{\partial \overset{\beta}{x}}(p)v]) = (p, \frac{\partial \overset{\alpha}{x}}{\partial \overset{\beta}{x}}(p)v) \tag{1.3}$$

with $(p, v) \in (U_\alpha \cap U_\beta) \times \mathbb{R}^n$, we may make use of countably many $\{\Phi_\alpha^{-1}\}$ to apply Lemma 1.14 to conclude that TM admits a smooth $2n$-dimensional atlas such that the mappings $\{\Phi_\alpha\}_{\alpha \in A}$ are smooth, thus also the projection, which may locally be written as $\pi|_{\pi^{-1}(U_\alpha)} = \mathrm{pr}_1 \circ \Phi_\alpha^{-1}$.

Definition 2.1. The smooth manifold TM is called the *tangent bundle* of M. For each $p \in M$, the fibre T_pM is the *tangent space at* $p \in M$ and its elements are called *tangent* or *contravariant vectors* at p.

Remark 2.2. For each $p \in U_\alpha$, the mapping $\Phi_\alpha(x, \cdot) : \mathbb{R}^n \to T_pM$ is an *isomorphism*. Hence, every contravariant vector may be written in the form

$$v = \sum_{i=1}^n \overset{\alpha}{v}{}^i \Phi_\alpha(p, e_i).$$

We will denote the *coordinate basis* $\{\Phi_\alpha(p, e_i)\}_{i=1}^n$ for T_pM by $\overset{\alpha}{\partial}_i\big|_p$ or $\frac{\partial}{\partial x^i}\big|_p$ whenever $\overset{\alpha}{\varphi} = (x^1, \ldots, x^n)$ and omit the overset α whenever we are dealing with a single coordinate system $\varphi|_U$. Note that by (1.3), whenever $p \in U_\alpha \cap U_\beta$, these bases are related *covariantly* according to the equation

$$\overset{\alpha}{\partial}_i\Big|_p = \sum_{j=1}^n \frac{\partial \overset{\beta}{x}{}^j}{\partial \overset{\alpha}{x}{}^i}(p) \cdot \overset{\beta}{\partial}_j\Big|_p. \tag{1.4}$$

△

The definition we have given above of a contravariant vector is in accordance with the classical definition involving the transformation behaviour of a tangent to a curve at a point: If $c : I \ni t_0 \to M$ is a differentiable curve with $p \in U$, $\varphi|_U$ a chart about p, then the assignment

$$\dot{c}(t_0) := \sum_{i=1}^{n} \frac{\mathrm{d}c^i}{\mathrm{d}t}(t_0) \cdot \partial_i|_p \in T_pM \tag{1.5}$$

is *independent* of the representative chart φ, where $c^i = x^i \circ c$. Conversely, given $v = \sum_i v^i \, \partial_i|_p \in T_pM$, we may find a $c : \,]t_0 - \varepsilon, t_0 + \varepsilon[\, \to M$ with $c(t_0) = p$ and $\dot{c}(t_0) = v$, namely $c(t) = \varphi_\alpha^{-1}(\varphi_\alpha(p) + (t - t_0)\sum_i v^i e_i)$. We shall use the various notations for the tangent to a curve that we are accustomed to from calculus, e.g.

$$\frac{\mathrm{d}c}{\mathrm{d}t}(t_0) := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} c := \dot{c}(t_0) := \dot{c}(t_0).$$

Moreover, we shall replace t with whatever variable the argument of c is represented by.

Remark 2.3. Note that if two differentiable curves c_1 and c_2 are *tangent* to one another at t_0 with respect to one coordinate system, i.e. $c_1^i(t_0) = c_2^i(t_0)$ and $\dot{c}_1^i(t_0) = \dot{c}_2^i(t_0)$, then $\dot{c}_1(t_0) = \dot{c}_2(t_0)$. This is immediate from the definition given above. △

Another rôle served by contravariant vectors is that of *directional differentiation*: For $v = \sum_i v^i \, \partial_i|_p \in T_pM$ and c as in the preceding paragraph, we define the mapping $\partial_v : C^\infty(M, \mathbb{R}) \to \mathbb{R}$ such that

$$f \mapsto \partial_v f := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} (f \circ c) = \sum_{i=1}^{n} v^i \partial_i (f \circ \varphi^{-1})(\varphi(p)).$$

It may be verified that this definition does not depend on the choice of φ or c. Moreover, the mapping $D := \partial_v$ is $\mathbb{R}$-linear and satisfies the following *Leibniz rule*: For all $f, g \in C^\infty(M)$,

$$D(fg) = f(p)Dg + g(p)Df. \tag{1.6}$$

Let $\mathrm{Der}_p(C^\infty(M))$ denote the $\mathbb{R}$-vector space of those $\mathbb{R}$-linear mappings $D : C^\infty(M) \to \mathbb{R}$ satisfying (1.6). It follows from Taylor's formula (cf. [4, Theorem 1.7.1]) that the mapping

$$\begin{aligned} T_pM &\to \mathrm{Der}_p(C^\infty(M)) \\ v &\mapsto \partial_v \end{aligned}$$

is an *isomorphism*.

Notation 2.4. We shall simply write $\partial_i|_p$ in place of $\partial_{\partial_i|_p}$ and also let $\partial_i f(p) := \partial_i|_p \cdot f$, which may easily be checked to define a smooth function on U. We shall also use the notation $\frac{\partial}{\partial x^i}$ in place of ∂_i and place indices over ∂ or x to indicate the the coordinate system from which the coordinate basis is derived. Likewise, we shall use diacritics to distinguish between coordinate systems. For example, given another coordinate system $\widetilde{\varphi}|_{\widetilde{U}}$, we write

$$\widetilde{\partial}_i f(p) = \frac{\partial f}{\partial \widetilde{x}^i}(p) = \partial_i (f \circ \widetilde{\varphi}^{-1})(\widetilde{\varphi}(p)).$$

Besides directional differentiation, we also obtain a natural analogue of the differential of a mapping between spaces: Suppose N is an m-dimensional smooth manifold and let $f : M \longrightarrow N$ be a differentiable mapping. Given a differentiable curve $c :]{-\varepsilon}, \varepsilon[\longrightarrow M$ with $c(0) = p$ and $\dot{c}(0) = v$, composition with f yields a curve $f \circ c :]{-\varepsilon}, \varepsilon[\longrightarrow N$ with $(f \circ c)(0) = f(p)$ whose tangent at 0 is given by

$$\mathrm{d}_p f(v) := \frac{\mathrm{d}(f \circ c)}{\mathrm{d}t}(0) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \partial_j f^i(p) \cdot v^j \right) \frac{\partial}{\partial y^i}\bigg|_{f(p)} \tag{1.7}$$

where $f^i := y^i \circ f$ and $(y^1, \ldots, y^m)$ is a chart on a neighbourhood of $f(p)$. Since the left-hand side does not depend on the chosen charts and the right-hand side does not depend on c, we deduce that the right-hand expression depends only on p, v and f and not on the choice of coördinate systems or curve. Moreover, this expression is *linear* in v so that it defines a linear mapping $\mathrm{d}_p f : T_pM \longrightarrow T_{f(p)}N$ called the *differential of f at $p \in M$*.[2] Note that the chain rule takes the following form: If $f : M \longrightarrow N$ and $g : N \longrightarrow O$ are differentiable mappings of smooth manifolds, then for each $p \in M$,

$$\mathrm{d}_p(g \circ f) = \mathrm{d}_{f(p)}g \circ \mathrm{d}_p f.$$

Exercise 2.5. Let V be an $\mathbb{R}$-vector space. Define the mapping

$$\begin{aligned} 𝄞 : V \times V &\longrightarrow TV \\ (p, v) &\longmapsto 𝄞_p v := \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (p + tv). \end{aligned}$$

It is not hard to see that $𝄞_p : V \longrightarrow T_pV$ is a vector space isomorphism and $𝄞$ is a *diffeomorphism*. In the same manner, if $U \subset V$ is open, we obtain a diffeomorphism $𝄞 : U \times V \longrightarrow TU$ defined the same way with the same properties.[3]

Remark 2.6. Note that tangents to curves, the directional derivative and the differential are related under the identification in Exercise 2.5 applied to $V = \mathbb{R}$ as follows (notation as in the definitions above):

$$\begin{aligned} \dot{c}(t_0) &= \mathrm{d}_{t_0} c(𝄞_{t_0} e) \\ \partial_v f &= 𝄞_{f(v)}^{-1} \mathrm{d}_p f(v). \end{aligned}$$

Moreover, we may relate the usual derivative of a function between (open subsets of) vector spaces to the differential introduced above as follows: Suppose $F : M_1 \longrightarrow M_2$ is a differentiable mapping, where M_1 and M_2 are open subsets of vector spaces V_1 and V_2 and write $\mathrm{D}_pF : V_1 \longrightarrow V_2$ for the Fréchet derivative. We compute, utilising Remark 2.3, that for $p \in M_1$ and $v \in V_1$,

$$𝄞_{F(p)} \mathrm{D}_pF(v) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} (F(p) + t\mathrm{D}_pF(v)) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} F(p + tv) = \mathrm{d}_pF(𝄞_p v).$$

△

[2] Also denoted f_* or $f_{*,p}$.

[3] Since U is open, the curve $t \mapsto p + tv$ is contained in U for sufficiently small t so that this definition makes sense.

Remark 2.7. Given a differentiable mapping $F : N \to M$ and $v \in T_q N$, we shall use the *partial derivative* notation

$$\partial_v F := \mathrm{d}_q F(v)$$

and if $v = \partial_i|_q$ with respect to a coordinate system $(y^1, \ldots, y^m)$ in a neighbourhood of q, we set

$$\frac{\partial F}{\partial y^i}(q) := \mathrm{d}_q F(\partial_i|_q).$$

△

Exercise 2.8. Let M_1 and M_2 be smooth manifolds of dimension n_1 and n_2 respectively. Write $\mathrm{pr}_i : M_1 \times M_2 \to M_i$ for the projection onto M_i and $\overset{i}{\iota}_{p_j} : M_i \to M_1 \times M_2$ for the *inclusion opposite* p_j, i.e. $\overset{1}{\iota}_{p_2}(p_1) = \overset{2}{\iota}_{p_1}(p_2) = (p_1, p_2)$. These mappings are smooth, and for each $(p_1, p_2) \in M_1 \times M_2$, we obtain an isomorphism of vector spaces

$$\begin{aligned} \psi_{(p_1,p_2)} : T_{p_1} M_1 \oplus T_{p_2} M_2 &\to T_{(p_1,p_2)}(M_1 \times M_2) \\ (v_1, v_2) &\mapsto \mathrm{d}_{p_1} \overset{1}{\iota}_{p_2}(v_1) + \mathrm{d}_{p_2} \overset{2}{\iota}_{p_1}(v_2) \end{aligned}$$

with inverse given by

$$\begin{aligned} T_{(p_1,p_2)}(M_1 \times M_2) &\to T_{p_1} M_1 \oplus T_{p_2} M_2 \\ V &\mapsto (\mathrm{dpr}_1(V), \mathrm{dpr}_2(V)). \end{aligned}$$

3. The cotangent bundle. Associated with the tangent bundle of M are certain bundles obtained by means of fibrewise algebraic operations, viz. duality and tensor product. We now turn our attention to these constructions.

For each $p \in M$, we write $T_p^* M := (T_p M)^*$ for the *dual space* of $T_p M$. Associated with any coordinate basis $\{\overset{\alpha}{\partial}_i\big|_p\}_{i=1}^n$ of $T_p M$ is a *dual basis*[4] $\{\mathrm{d}\overset{\alpha}{x}{}^i\big|_p\}_{i=1}^n$ of $T_p^* M$ defined such that for all i, j,

$$\left(\mathrm{d}\overset{\alpha}{x}{}^i\Big|_p, \overset{\alpha}{\partial}_j\Big|_p\right) = \delta^i_j = \begin{cases} 1, & i = j \\ 0, & \text{otherwise} \end{cases};$$

by (1.4), whenever $p \in U_\alpha \cap U_\beta$, these dual bases are related *contravariantly* by

$$\mathrm{d}\overset{\alpha}{x}{}^i\Big|_p = \sum_{j=1}^n \left(\mathrm{d}\overset{\alpha}{x}{}^i\Big|_p, \overset{\beta}{\partial}_j\Big|_p\right) \mathrm{d}\overset{\beta}{x}{}^j\Big|_p = \sum_{j=1}^n \frac{\partial \overset{\alpha}{x}{}^i}{\partial \overset{\beta}{x}{}^j}(p) \cdot \mathrm{d}\overset{\beta}{x}{}^j\Big|_p. \tag{1.8}$$

We now form the (disjoint) union of these dual spaces

$$T^* M := \bigcup_{p \in M} T_p^* M$$

from which we obtain a natural projection[5]

$$\begin{aligned} \pi : T^* M &\to M \\ T_p^* M \ni \nu &\mapsto p \end{aligned}$$

[4] We will later omit the overset α when dealing with a single coordinate system.

[5] We will denote all of our projections by π since context will always make clear which one we mean!

as well as local parametrisations $\{{}_*\Phi_\alpha : U_\alpha \times \mathbb{R}_n \to \pi^{-1}(U_\alpha)\}_{\alpha\in A}$ defined such that

$$ {}_*\Phi_\alpha(p, \sum_{i=1}^n v_i e^i) = \sum_{i=1}^n v_i \, \mathrm{d}\overset{\alpha}{x}{}^i\Big|_p . $$

As with TM, we may appeal to Lemma 1.14 to conclude that T^*M admits a smooth $2n$-dimensional atlas such that the $\{{}_*\Phi_\alpha\}$ (as well as π) are smooth.

Definition 3.1. The smooth manifold T^*M is called the *cotangent bundle* of M. For each $p \in M$, the fibre $\pi^{-1}(\{p\}) = T^*_pM$ is the *cotangent space* at $p \in M$ and its elements are called *cotangent* or *covariant vectors* at p.

One familiar covariant vector is the *differential of a function* $f : M \to \mathbb{R}$ at a point $p \in M$, which is defined as the unique $\mathrm{d}f(p) \in T^*_pM$ such that

$$ (\mathrm{d}f(p), v) = \partial_v f . $$

In local coordinates, $\mathrm{d}f(p) = \partial_i f(p) \cdot \mathrm{d}x^i\big|_p$. The notation here is consistent in the sense that

$$ \mathrm{d}x^i(p) = \mathrm{d}x^i\big|_p . $$

A mapping we shall have occasion to use is the *codifferential* $\delta_p f : T^*_{f(p)}N \to T^*_pM$ of a mapping $f : M \to N$ at $p \in M$, defined as the *adjoint* of $\mathrm{d}_p f$, i.e. such that for all $v_2 \in T^*_{f(p)}N$ and $v \in T_pM$,

$$ (\delta_p f(v_2), v) = (v_2, \mathrm{d}_p f(v)). $$

In local coordinates $(x^1, \ldots, x^n)$ about p and $(y^1, \ldots, y^m)$ about $f(p)$,

$$ \delta_p f(\sum_{i=1}^m v_i \, \mathrm{d}y^i\big|_{f(p)}) = \sum_{i=1}^m \sum_{j=1}^n v_i(\mathrm{d}y^i\big|_{f(p)}, \mathrm{d}_p f(\partial_j\big|_p)) \, \mathrm{d}x^j\big|_p = \sum_{i,j} v_i \cdot \partial_j f^i(p) \cdot \mathrm{d}x^j\big|_p . $$

Note that the chain rule for $\mathrm{d}_p f$ implies that whenever $f : M \to N$ and $g : N \to O$ are differentiable mappings of smooth manifolds, we have that $\delta_p(g \circ f) = \delta_p f \circ \delta_{f(p)} g$ for all $p \in M$.

4. Tensor bundles. We now turn our attention to tensors more generally. Note that as in the previous section, we may apply algebraic operations fo the tangent spaces T_pM, take their disjoint union to obtain a new manifold of vector spaces over M with a projection and natural coordinate bases. Such manifolds will be referred to as *tensor bundles*.

One example is the following: For $r, s \in \{0, 1, \ldots\}$, let

$$ (T^r_s M)_p := \underbrace{T_pM \otimes \cdots \otimes T_pM}_{r \text{ times}} \otimes \underbrace{T^*_pM \otimes \cdots \otimes T^*_pM}_{s \text{ times}} $$

and form their disjoint union $T^r_s M := \bigcup_{p\in M} (T^r_s M)_p$.[6] This is called the *bundle of* (r, s) tensors, and an element of $(T^r_s M)_p$ is called an (r, s)*-tensor at* p. For every coordinate system $\varphi|_U$ and $p \in U$, we obtain a natural coordinate basis for T_pM,[7]

$$ \{\, \partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s}\big|_p : i_1, \ldots, i_r, j_1, \ldots, j_s \in \{1, \ldots, n\}\}, $$

[6] If $r = s = 0$, we set $(T^0_0 M)_p := \mathbb{R}$.

[7] We will tack the base point p on to the end of the basis rather than repeating it throughout!

which gives rise to a Φ-map as with the tangent and cotangent bundles, as well as a π-map, making $T^r_s M$ a smooth manifold with respect to which these mappings are smooth. Note that an arbitrary element of $(T^r_s M)_p$ will be written as

$$\sum_{\substack{i_1,\dots,i_r\\ j_1,\dots,j_s}} T^{i_1\dots i_r}{}_{j_1\dots j_s}\ \partial_{i_1}\otimes\cdots\otimes\partial_{i_r}\otimes \mathrm{d}x^{j_1}\otimes\cdots\otimes \mathrm{d}x^{j_s}\big|_p .$$

The $T^{i_1\dots i_r}{}_{j_1\dots j_s}$ are referred to as the *components* of T. Note that if $\widetilde{\varphi}|_{\widetilde{U}}$ is any other coordinate system and $p \in U \cap \widetilde{U}$, then the components of T are related by

$$\widetilde{T}^{i_1\dots i_r}{}_{j_1\dots j_s} = \sum_{\substack{k_1,\dots,k_r\\ l_1,\dots,l_s}} T^{k_1\dots k_r}{}_{l_1\dots l_s}\ \frac{\partial\widetilde{x}^{k_1}}{\partial x^{i_1}}\cdot\dots\cdot\frac{\partial\widetilde{x}^{k_r}}{\partial x^{i_r}}\cdot\frac{\partial x^{l_1}}{\partial\widetilde{x}^{j_1}}\cdot\dots\cdot\frac{\partial x^{l_s}}{\partial\widetilde{x}^{j_s}}$$

We say that T is *r-times contravariant* and *s-times covariant.* If $s = 0$ (resp. $r = 0$), we say that T is a *contravariant* (resp. *covariant*) tensor at p. Note that we could have instead chosen an *arbitrary* basis $\{\varepsilon_i|_p\}_{i=1}^n$ for T_pM and its corresponding dual basis $\{\varepsilon^i\big|_p\}_{i=1}^n$ to express an (r, s)-tensor at p in terms of the tensor product basis

$$\{\varepsilon_{i_1}\otimes\cdots\otimes\varepsilon_{i_r}\otimes\varepsilon^{j_1}\otimes\cdots\otimes\varepsilon^{j_s}\big|_p\}.$$

We will use the same notation for the components of T in this case as well. In general, we will use an *arbitrary basis* unless stated otherwise. By convention, an empty product will be interpreted as the scalar 1 so that if we are dealing with the case $(r, s) = (0, 0)$, we shall interpret $\varepsilon_{i_1}\otimes\cdots\otimes\varepsilon_{i_r}\otimes\varepsilon^{j_1}\otimes\cdots\otimes\varepsilon^{j_s}\big|_p = 1$.

In a lot of cases, it is convenient to refer to a tensor in terms of its components, i.e. by giving a symbol together with free indices in the right places to indicate its type. For example, given a base point p, we would interpret T_{ij} as a $(0, 2)$-tensor, or $T^{ij}{}_{kl}$ as a $(2, 2)$-tensor. With this in mind, we now indicate some (multi)linear operations on tensors (besides addition and scalar multiplication!):

- *Tensor product*: Given $S \in (T^{r_1}_{s_1}M)_p$ and $T \in (T^{r_2}_{s_2}M)_p$, we obtain a tensor $(S \otimes T) \in (T^{r_1+r_2}_{s_1+s_2}M)_p$ such that

  $$(S\otimes T)^{i_1\dots i_{r_1+r_2}}{}_{j_1\dots j_{s_1+s_2}} = S^{i_1\dots i_{r_1}}{}_{j_1\dots j_{s_1}}\cdot T^{i_{r_1+1}\dots i_{r_1+r_2}}{}_{j_{s_1}\dots j_{s_1+s_2}} .$$

  Note that this definition coïncides with the usual tensor product

  $$(T^{r_1}_{s_1}M)_p\times(T^{r_2}_{s_2}M)_p\longrightarrow(T^{r_1}_{s_1}M)_p\otimes(T^{r_2}_{s_2}M)_p$$

  followed by the isomorphism $(T^{r_1}_{s_1}M)_p\otimes(T^{r_2}_{s_2}M)_p \simeq (T^{r_1+r_2}_{s_1+s_2}M)_p$ that appropriately rearranges the entries of the tensor product. In particular, for a contravariant S and covariant T, $S \otimes T$ is the usual tensor product, but $T \otimes S = S \otimes T$ according to our definition.

- *Contraction* or *trace*: Given $T \in (T^r_s M)_p$, we obtain a tensor $\operatorname{tr} T \in (T^{r-1}_{s-1}M)_p$ by equating and summing over a pair of indices, one covariant and one contravariant, e.g.

  $$(\operatorname{tr} T)^{i_1\dots i_{r-1}}{}_{j_1\dots j_{s-1}} = \sum_{i=1}^n T^{ii_1\dots i_{r-1}}{}_{ij_1\dots j_{s-1}} .$$

  We may of course repeat the process and contract more pairs of indices to obtain a tensor of lower order. Note in particular that if we contract a $(1, 1)$-tensor, we obtain a scalar. We will always make explicit *which* indices are being contracted.

- *Interior* or *inner products*: Given two tensors, we may define a product by combining the tensor product and contraction, viz. we form the tensor product and then contract over pairs of indices, where each pair consists of one covariant index from one of the tensors and one contravariant from the other. For example, if $v \in T_pM$ and $T \in (T^r_s M)_p$, we obtain tensors $v \llcorner T, T \lrcorner v \in (T^r_{s-1}M)_p$ such that

$$(v \llcorner T)^{i_1 \dots i_r}{}_{j_1 \dots j_{s-1}} = \sum_{i=1}^{n} v^i T^{i_1 \dots i_r}{}_{i j_1 \dots j_{s-1}}$$
$$(T \lrcorner v)^{i_1 \dots i_r}{}_{j_1 \dots j_{s-1}} = \sum_{i=1}^{n} T^{i_1 \dots i_r}{}_{j_1 \dots j_{s-1} i} v^i$$

Likewise, for $\alpha \in T^*_p M$ and T as before, we obtain tensors $\alpha \lrcorner T, T \llcorner \alpha \in (T^{r-1}_s M)_p$ such that

$$(\alpha \lrcorner T)^{i_1 \dots i_{r-1}}{}_{j_1 \dots j_s} = \sum_{i=1}^{n} \alpha_i T^{i i_1 \dots i_{r-1}}{}_{j_1 \dots j_s}$$
$$(T \llcorner \alpha)^{i_1 \dots i_{r-1}}{}_{j_1 \dots j_s} = \sum_{i=1}^{n} T^{i_1 \dots i_{r-1} i}{}_{j_1 \dots j_s} \alpha_i .$$

Note that if we take an interior product of an (r, s)-tensor S with an (s, r)-tensor T, we obtain a scalar (S, T):

$$(S, T) = \sum_{\substack{i_1, \dots, i_r \\ j_1, \dots, j_s}} S^{i_1 \dots i_r}{}_{j_1 \dots j_s} \cdot T^{j_1 \dots j_s}{}_{i_1 \dots i_r} .$$

- *Symmetrising and anti-symmetrising*: Given a tensor $T \in (T^r_s M)_p$, we have a natural action of the symmetric group $\mathfrak{S}_r$ (resp. $\mathfrak{S}_s$) on the contravariant (resp. covariant) part of T: If $\sigma \in \mathfrak{S}_r$ (resp. $\sigma \in \mathfrak{S}_s$), we obtain a new tensor $\overset{\sigma}{T}$ (resp. $\underset{\sigma}{T}$) of the same type such that

$$\overset{\sigma}{T}{}^{i_1 \dots i_r}{}_{j_1 \dots j_s} := T^{i_{\sigma(1)} \dots i_{\sigma(r)}}{}_{j_1 \dots j_s} \qquad (\text{resp. } \underset{\sigma}{T}{}^{i_1 \dots i_r}{}_{j_1 \dots j_s} := T^{i_1 \dots i_r}{}_{j_{\sigma(1)} \dots j_{\sigma(s)}}).$$

If T is covariant (resp. contravariant), we call T *symmetric* if $\overset{\sigma}{T} = T$ (resp. $\underset{\sigma}{T} = T$), and *anti-symmetric* (or *skew-symmetric*) if $\overset{\sigma}{T} = (-1)^\sigma T$ (resp. $\underset{\sigma}{T} = (-1)^\sigma T$). If T is a $(0, s)$ tensor, the tensors

$$\frac{1}{s!} \sum_{\sigma \in \mathfrak{S}_s} \underset{\sigma}{T} \qquad \text{and} \qquad \frac{1}{s!} \sum_{\sigma \in \mathfrak{S}_s} (-1)^\sigma \underset{\sigma}{T}$$

are symmetric and anti-symmetric respectively, called the *symmetric* and *anti-symmetric parts* of T; we make similar definitions for contravariant tensors. Note that in the case where $s = 2$, symmetry is equivalent to the condition $T_{ij} = T_{ji}$ and anti-symmetry to the condition $T_{ij} = -T_{ji}$; every such tensor may be uniquely decomposed according to

$$T_{ij} = \underbrace{\frac{1}{2}(T_{ij} + T_{ji})}_{\text{symmetric}} + \underbrace{\frac{1}{2}(T_{ij} - T_{ji})}_{\text{anti-symmetric}} .$$

Remark 4.1. The above operations allow us to interpret certain tensors in terms of familiar objects. For example, if g is a $(0, 2)$-tensor at p, then we may view it as a *bilinear* (or *quadratic*) form on T_pM:

$$T_pM \times T_pM \ni (v, w) \mapsto (g, v \otimes w) = g_{ij} v^i w^j.$$

Similarly, a $(0, s)$-tensor may be viewed as an s-linear form on T_pM and an $(r, 0)$-tensor as an r-linear form on T_p^*M. More generally, an (r, s)-tensor may be viewed variously as a multilinear mapping into $\mathbb{R}$ taking as input r elements of T_p^*M and s elements of T_pM, or as a multilinear mapping taking s elements of T_pM into $(T_0^rM)_p$. In particular, a $(1, 1)$-tensor T may be viewed as a mapping $T_pM \longrightarrow T_pM$: Taking the interior product of $v \in T_pM$ with T yields the contravariant vector

$$(v \llcorner T)^i = T^i{}_j v^j.$$

One such example of this is the *Kronecker delta*, which we may write more suggestively as $\delta^i{}_j$; it corresponds to none other than the identity map:

$$v \llcorner \delta = v.$$

△

5. Tensor fields. Rather than only considering tensors at a single point, we will be interested in mappings that take a point in M to a tensor space corresponding to that point. Moreover, we shall be interested in such mappings defined only on a subset of M, e.g. along a curve. This motivates the following definition.

Definition 5.1. A mapping $T : M \longrightarrow T_s^rM$ is said to be a (smooth) (r, s)-*tensor field* if T is smooth and for all $p \in M$, $T(p) \in (T_s^rM)_p$, i.e. $\pi \circ T = \mathrm{id}_M$. If T is only defined on an open subset U of M, but satisfies the other conditions, we shall call it a *local tensor field.* Finally, given a mapping $f : S \longrightarrow M$ with S a manifold or interval, if instead $T : S \longrightarrow T_s^rM$ is such that $T(q) \in (T_s^rM)_{f(q)}$, we shall say it is a *tensor field along* f.

Remark 5.2. We shall usually assume that tensor fields are smooth. When considering a tensor field along a mapping f, we shall explicitly state how regular T is, since f might not actually be smooth. △

Remark 5.3. Since $(T_0^0M)_p = \mathbb{R}$ for all p, we shall view $(0, 0)$-tensors as functions. △

Example 5.4 (vector fields)**.** We call a $(1, 0)$-tensor field a *vector field.* Note that coordinate bases give rise to local vector fields $\{\partial_i := (p \mapsto \partial_i|_p)\}_{i=1}^n$. More generally, given a collection of (smooth) local vector fields $\{\varepsilon_i\}_{i=1}^n$ over an open set U such that for each $p \in U$, $T_pM = \mathrm{span}\{\varepsilon_i(p)\}$, we call this collection a *local frame* for TM.

Example 5.5 (covector fields)**.** We call a $(0, 1)$-tensor field a *covector field* or *one-form.* As in the case of the local basis for TM, the dual coordinate bases on T^*M give rise to local covector fields $\{\mathrm{d}x^i := (p \mapsto \mathrm{d}x^i\big|_p)\}_{i=1}^n$. Moreover, given a local frame $\{\varepsilon_i\}_{i=1}^n$ for TM over U, we obtain a *dual coframe* $\{\varepsilon^i\}_{i=1}^n$ for T^*M over U such that for each $p \in U$, $\{\varepsilon^i(p)\}$ is the dual basis to $\{\varepsilon_i(p)\}_{i=1}^n$.

Example 5.6 (tangent to a curve)**.** Given a differentiable curve $c : I \longrightarrow M$, the mapping $t \mapsto \dot{c}(t)$ is a vector field along c.

Remark 5.7. The algebraic operations of §4 may be applied pointwise to tensor fields, and if the tensor fields involved in the algebraic process are smooth, so is the resulting tensor. △

Apart from the algebraic operations on tensors, we also have natural operations induced by differentiable mappings. Let $f : U \subset M \to N$ be a differentiable mapping. We have the following operations:

- *Pull-back*: This operation allows us to pull functions and more generally (local) *covariant* tensor fields on N back by f to M. We start with the simplest case: If $\phi : V \subset N \to \mathbb{R}$ is any function, then $f^*\phi := \phi \circ f : f^{-1}(V) \subset M \to \mathbb{R}$ is the *pull-back* of ϕ by f. If instead $\phi : V \to T^*N$ is a local one-form, we define the *pull-back* of ϕ by f to be the local one-form $f^*\phi : f^{-1}(V) \to T^*M$ such that

$$p \mapsto \delta_p f(\phi(f(p))) = \partial_i f^\alpha(p) \cdot \phi_\alpha(f(p))\mathrm{d}x^i.$$

  More generally, for a covariant tensor field $T : V \to T^0_s N$, we may form the pull-back in a similar manner by acting $\delta_p f$ on each entry of the tensor product, i.e. the *pull-back* of T by f is the local tensor field $f^*T : f^{-1}(V) \to T^0_s M$ such that

$$p \mapsto (\delta_p f \otimes \ldots \delta_p f)(T(f(p))) = (\partial_{i_1} f^{\alpha_1} \ldots \partial_{i_s} f^{\alpha_s})(p) \cdot T_{\alpha_1 \ldots \alpha_s}(f(p))\, \mathrm{d}x^{i_1} \otimes \cdots \otimes \mathrm{d}x^{i_s}\big|_p .$$

  Note that f^*T is smooth whenever both f and T are.

- *Push-forward*: This operation allows us to push (local) *contravariant* tensor fields on M *forward* to N to yield a tensor field along f. Firstly, if $X : U \subset M \to TM$ is a vector field, we define the *push-forward* of X along f to be the vector field $f_*X : U \to TN$ along f such that

$$p \mapsto \mathrm{d}_p f(X(p)) = X^i(p) \partial_i f^\alpha(p)\, \partial_\alpha|_{f(p)} .$$

  Analogously to the pull-back, we define the *push-forward* of a contravariant tensor $T : U \subset M \to T^r_0 M$ by f by acting $\mathrm{d}_p f$ on each entry of the tensor product, i.e. we define it to be the tensor field $f_*T : U \to T^r_0 N$ along f such that

$$p \mapsto (\mathrm{d}_p f \otimes \cdots \otimes \mathrm{d}_p f)(T(p)) = \left(T^{i_1 \ldots i_r} \cdot \partial_{i_1} f^{\alpha_1} \ldots \partial_{i_r} f^{\alpha_r}\right)(p)\, \partial_{\alpha_1} \otimes \cdots \otimes \partial_{\alpha_r}\big|_{f(p)} .$$

  Again, if both f and T are smooth, so is f_*T.

The above operations are valid for any differentiable mapping. If instead $f : U \subset M \to U' \subset N$ is a *diffeomorphism*, then we may actually define the *pull-back* of an arbitrary tensor field $T : V \subset N \to T^r_s N$ as the tensor field $f^*T : f^{-1}(V) \to T^r_0 M$ such that

$$p \mapsto (\mathrm{d}_{f(p)} f^{-1} \otimes \cdots \otimes \mathrm{d}_{f(p)} f^{-1} \otimes \delta_p f \otimes \cdots \otimes \delta_p f) T(f(p));$$

in local coordinates,

$$(f^*T)^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} = \partial_{k_1}(f^{-1})^{i_1} \circ f \ldots \partial_{k_r}(f^{-1})^{i_r} \circ f \cdot \partial_{j_1} f^{l_1} \ldots \partial_{j_s} f^{l_s} \cdot T^{k_1 \ldots k_r}{}_{l_1 \ldots l_s} .$$

Remark 5.8. In the sequel, we will usually refer to tensor fields simply as *tensors*. △

6. Vector fields and flows. As before, let M be a smooth manifold. A (smooth) mapping $X : M \longrightarrow TM$ is said to be a *vector field* on M if for each $p \in M$, $X(p) \in T_pM$. Recall that given a tangent vector $v \in T_pM$, we may always find a differentiable curve passing through p with tangent v. We now turn our attention to the possibility of *integrating* a vector field so as to obtain a curve whose tangents coïncide with a vector field X in a sense we shall now make precise.

Definition 6.1. An *integral curve* of a vector field X on M is a differentiable mapping $c : I \subset \mathbb{R} \longrightarrow M$, I an open interval, such that for all $t \in I$,

$$\dot{c}(t) = X(c(t)). \tag{1.9}$$

An integral curve $c : I \longrightarrow M$ is said to *pass through a point* $p_0 \in M$ if there exists a $t_0 \in I$ such that $c(t_0) = p_0$.

Let X be a vector field on M and suppose $c : I \longrightarrow M$ is an integral curve passing through $p_0 \in M$ (with $c(t_0) = p_0$). In order to locally describe (1.9), fix a coordinate system $\varphi : U \longrightarrow \Omega \subset \mathbb{R}^n$ with $p_0 \in U$. Writing $x_0 := \varphi(p_0)$, $f^i(x) := (X^i \circ \varphi^{-1})(x)$ and $x(t) := \varphi(c(t))$ for $t \in]t_0 - \delta_1, \delta_2[$ for sufficiently small $\delta_1, \delta_2 > 0$, the system (1.9) together with the condition $c(t_0) = p_0$ is equivalent to the system of ODE

$$\begin{aligned} x^i(t) &= f^i(x(t)) \\ x(0) &= x_0 \end{aligned}$$

for $t \in]t_0 - \delta_1, t_0 + \delta_2[$.

Example 6.2. Let $M = \mathbb{R}^2$ and $X(x, y) = -y\partial_x + x\partial_y$. An integral curve $t \mapsto (x(t), y(t))$ passing through (x_0, y_0) at $t = 0$ must satisfy the system of equations

$$\begin{aligned} \dot{x} &= -y \\ \dot{y} &= x \end{aligned}$$

together with the initial condition $(x, y)(0) = (x_0, y_0)$. This system may be solved by differentiating the first equation to obtain $\ddot{x} = -x$. After obtaining the general solution and substituting back in the first equation to obtain an expression for y, we find that

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} x_0 \\ y(0) \end{pmatrix}.$$

Therefore, we may find an integral curve passing through any given $(x_0, y_0) \in \mathbb{R}^2$, and this integral curve is defined on all of $\mathbb{R}$.

Example 6.3. Let $M = \mathbb{R}$ and consider the vector field $X(x) = x^2\partial$. An integral curve $t \mapsto x(t)$ passing through x_0 at $t = 0$ must solve the ODE $\dot{x} = x^2$ subject to the initial condition $x(0) = x_0$. We see that if $x_0 = 0$, then $t \mapsto 0$ is an integral 'curve' passing through 0 defined on all of $\mathbb{R}$. On the other hand, for $x_0 \neq 0$, we may solve this equation to obtain

$$x(t) = \frac{x_0}{1 - x_0 t}.$$

In particular, this curve is only well-defined on $\mathbb{R} \backslash \{\frac{1}{x_0}\}$.

In addition to the preceding example, note that we may just as well have a 'nice' vector field as in Example 6.2 and still not have an integral curve that may be indefinitely extended (consider Example 6.2 on $\mathbb{R}^2 \backslash \{(x, 0) : x < 0\}$).

We now recall a classical theorem on the existence, uniqueness and regularity of solutions of such systems. Its proof may be found for instance in [3, Ch. 6].

Theorem 6.4. Let $\Omega \subset \mathbb{R}^n$ be open with $B(x_0, r) \Subset \Omega$, $t_0 \in \mathbb{R}$ and $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ a smooth mapping. There exists a smooth mapping

$$x : [t_0 - \frac{r}{2\sup_{B(x_0,r)} |f|}, t_0 + \frac{r}{2\sup_{B(x_0,r)} |f|}] \times \overline{B(x_0, \frac{r}{2})} \to \Omega$$

such that for every $y \in \overline{B(x_0, \frac{r}{2})}$, $t \mapsto x(t; y)$ is a solution to the system

$$\begin{aligned} \dot{x}(t) &= f(x(t)) \\ x(t_0) &= y \end{aligned} \tag{1.10}$$

on its domain of definition. Morover, this solution has the following *local uniqueness* property: If $\widetilde{x} : \widetilde{I} \ni t_0 \to \Omega$ is any differentiable on an interval $\widetilde{I}$ solving (1.10), then $\widetilde{x}(t) = x(t)$ for all $t \in I \cap \widetilde{I}$.

In light of this theorem and the above reduction to local coordinates, we conclude that integral curves of vector fields always exist locally. We now show that there exists a unique *maximal* integral curve passing through any given point.

Theorem 6.5. For any $p_0 \in M$, there exists an open interval $I \ni 0$ and an integral curve $\gamma : I \to M$ passing through p_0 at time 0 that is unique and maximal in the following sense: If $c : J \to M$ is any other integral curve passing through p_0 at time 0, then $J \subset I$ and $c \equiv \gamma|_J$.

Proof. By Theorem 6.4, we know that for each $p_0 \in M$, there exists at least one integral curve $\gamma_0 : I_0 \to M$ with $\gamma_0(0) = p_0$. Suppose $\gamma_1 : I_1 \to M$ is any other integral curve with $\gamma_1(0) = p_0$ and let $J = \{t \in I_0 \cap I_1 : \gamma_0(t) = \gamma_1(t)\}$. By local uniqueness above, it is clear that this set is nonempty. Moreover, by continuity, it is also (sequentially) closed. Fixing $t_0 \in J$ and passing to local coordinates, since $\gamma_0^i(t_0) = \gamma_1^i(t_0)$, it follows from local uniqueness that γ_0^i and γ_1^i, therefore also γ_0 and γ_1, agree on a subinterval of $I_0 \cap I_1$ containing t_0. Since $I_0 \cap I_1$ is connected, we see that $\gamma_1 \equiv \gamma_2$ on $I_0 \cap I_1$, i.e. local uniqueness carries over to integral curves *not necessarily completely lying in a coordinate neighbourhood.*

Let I_{p_0} be the union of all open intervals containing 0 on which there exists an integral curve passing through p_0 at time 0. We see that $I_0 \subset I_{p_0}$ so that this interval is indeed nonempty. Now define $\gamma : I_{p_0} \to M$ such that for $t \in I$, $\gamma(t) = c(t)$, where $c(t)$ is the value of an integral curve defined on an interval containing 0 and t and passing through p_0 at time 0. By the preceding paragraph, γ is well-defined and possesses the desired uniqueness and maximality properties. □

For each $p \in M$, we shall write I_p for the domain of the unique maximal integral curve of X passing through p at time 0 and $t \mapsto \phi_t(p)$ for this integral curve, which yields a map $\phi : Q \subset \mathbb{R} \times M \to M$ with $Q = \bigcup_{x \in M} I_x \times \{x\}$. We shall now study this map in more detail. The following lemma shows that the $\{\phi_t\}$ possess a certain group property.

Lemma 6.6. The mapping ϕ satisfies $\phi_0 \equiv \mathrm{id}_M$ Moreover, whenever $t \in I_p$ and $s \in I_{\varphi_t(p)}$, $s + t \in I_p$ and

$$\phi_s(\phi_t(p)) = \phi_{s+t}(p).$$

Finally, for any $t \in I_p$, $-t \in I_{\phi_t(p)}$ and $(\phi_t \circ \phi_{-t})(p) = (\phi_{-t} \circ \phi_t)(p) = p$.

Proof. The first claim follows from the fact that $t \mapsto \phi_t(p)$ passes through p at $t = 0$. As for the second claim, consider the curves

$$\underset{1}{\phi} := \left(I_{\phi_t(p)} \ni s \mapsto \phi_s(\phi_t(p)) \in M\right)$$
$$\underset{2}{\phi} := \left(I_p - t \ni s \mapsto \phi_{s+t}(p)\right).$$

Both of these curves are integral curves of X passing through $\phi_t(p)$ at $s = 0$ so that Theorem 6.5 they must coïncide on the intersection of their domains. Since the intervals I_p and $I_{\phi_t(p)} + t$ both contain t, their union $\widetilde{I}$ is also an interval. Now, the curve $\widetilde{x} : \widetilde{I} \to M$ defined such that

$$s \mapsto \begin{cases} \phi_s(p), & s \in I_p \\ \phi_{s-t}(\phi_t(p)), & s \in t + I_{\phi_t(p)} \end{cases}$$

is an integral curve passing through p at $t = 0$ so that by Theorem 6.5 again,

$$I_{\phi_t(p)} + t \subset \widetilde{I} \subset I_p \subset \widetilde{I}$$

so that $\widetilde{x} \equiv \phi.(p)$. Evaluating both of these functions at $s + t \in I_{\phi_t(p)} + t$ establishes the second claim.

For the last claim, consider the curve $s \mapsto \phi_{t+s}(p)$, which includes $-t$ and 0 in its interval of definition. Since it is an integral curve passing through $\phi_t(p)$ at $s = 0$, we may again appeal to Theorem 6.5 to conclude that $-t \in I_{\phi_t(p)}$ so that by the second claim, $\phi_{-t}(\phi_t(p)) = p$. □

Theorem 6.7. For any $K \Subset M$, then there is a $\delta_K > 0$ such that $K \times]-\delta_K, \delta_K[\subset Q$. Moreover, Q is open and ϕ is smooth.

Proof. As remarked earlier, Theorem 6.4 implies that for each $p \in K$ we have the existence of an open neighbourhood $U_p \subset M$ of p and a $\delta_p > 0$ such that for each $q \in U_p$, $I_q \supset \left]-\delta_p, \delta_p\right[$. K may be covered by finitely many such sets $\{U_{p_i}\}_{i=1}^N$ so that $p \in K \Rightarrow p \in U_{p_i}$ for some i. Setting $\delta_K = \inf\{\delta_{p_i}\}$, we immediately obtain that $I_x \supset]-\delta_K, \delta_K[$, establishing the first claim.

We now show that Q is open. Firstly, since we have the existence of an open $U_p \ni p$ and $\left]-\delta_p, \delta_p\right[\subset \mathbb{R}$ for each $p \in M$ such that for any $q \in U_p$, $\left]-\delta_p, \delta_p\right[\subset I_q$, we immediately have that $\left]-\delta_p, \delta_p\right[\times U_p \subset Q$, i.e. $\{0\} \times M$ lies in the interior of Q. We now fix $(p_0, t_0) \in Q$ with $t_0 > 0$ and fix $\delta_0 \in]0, t_0[$ such that $t_0 + \delta_0 \in I_{p_0}$. Since the set $K := \phi.(p_0)([0, t_0 + \delta_0])$ is compact, we may run through the argument of the preceding proof to obtain a neighbourhood V of K and a $\delta_1 > 0$ such that $I_q \supset]-\delta_1, \delta_1[$. Let $N > 0$ be large enough so that $s_0 := \frac{t_0+\delta_0}{N} < \frac{\delta_1}{2}$ and define

$$U_0 := \underbrace{\phi_{s_0}\big|_V^{-1} \dots \phi_{s_0}\big|_V^{-1}}_{N-1 \text{ times}} V.$$

U_0 is open by continuity, and since $p_0 \in V$, $\phi_{s_0}(p_0) \in V$, ..., $\phi_{(N-1)s_0}(p_0) \in V$, we have that $p_0 \in U_0$. We now claim that $p \in U_0 \Rightarrow t_0 + \delta_0 \in I_p$. To see this, note that

$p \in U_0 \Rightarrow \phi_{s_0}^i(p) \in V$ for $i \in \{0, \ldots, N-1\}$ so that appealing to Lemma 6.6,

$$\begin{aligned} s_0 \in I_{\phi_{s_0}^{N-1}(p)} \wedge s_0 \in I_{\phi_{s_0}^{N-2}(p)} \\ \Rightarrow 2s_0 \in I_{\phi_{s_0}^{N-2}(p)} \wedge s_0 \in I_{\phi_{s_0}^{N-3}(p)} \\ \Rightarrow \cdots \Rightarrow Ns_0 = t_0 + \delta_0 \in I_p. \end{aligned}$$

Therefore, $]t_0 - \delta_0, t_0 + \delta_0[\subset I_p$ for all $p \in U_0$, i.e. $]t_0 - \delta, t_0 + \delta_0[\times U_0 \subset Q$. A similar argument shows that the same holds for $t_0 < 0$ and a suitably chosen δ_0 so that Q is open.

To show that ϕ is smooth, we again fix $(t_0, x_0) \in Q$ with $t_0 > 0$, the case of $t_0 < 0$ following analogously (and the case of $t_0 = 0$ following immediately from Theorem 6.4). Retaining the notation of the preceding paragraph, we may find an i such that $t_0 \in]is_0, (i+2)s_0[$. In particular, for $s \in]is_0, (i+2)s_0[$ and $q \in U_0$, we may write

$$\phi_s(q) = \phi_{s-is_0}(\phi_{is_0}(q)).$$

Since $s - is_0 \in]0, \delta_1[$ and both $q, \phi_{is_0}(q) \in V$, the right-hand side is a smooth function of y and q so that ϕ is smooth in a neighbourhood of (t_0, x_0). □

Motivated by the preceding considerations, we introduce the following notion.

Definition 6.8. A smooth mapping $f : Q_f \subset \mathbb{R} \times M \to M$ on an open set Q_f of the form

$$Q_f = \bigcup_{p \in M} J_p \times \{p\}$$

with $J_p \ni 0$ an open interval is said to be a *local flow* if $f(0, \cdot) = \mathrm{id}_M$ and the equality

$$f(t, f(s,p)) = f(t+s, p)$$

holds whenever both sides are defined. If $Q_f = \mathbb{R} \times M$, we say that f is a (global) *flow*.

Remark 6.9. It is customary to write $f(t,p)$ as $f_t(p)$ and refer to the family of mappings $\{f_t\}$ as a *local flow*. By the openness of Q_f, we may always find a neighbourhood $N(p)$ of any $p \in M$ and an interval J'_p containing 0 such that f_t is defined on $N(p)$ for all $t \in J'_p$. △

Lemma 6.6 and Theorem 6.7 together imply that the mapping ϕ obtained above is a local flow so that a smooth vector field always gives rise to a local flow. It turns out we may also go the other way.

Exercise 6.10. Given a local flow $\{f_t\}$, the assignment

$$X_f(p) := \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f_t(p)$$

defines a smooth vector field $X_f : M \to TM$. Let $\phi : Q \to M$ denote the local flow generated by X_f. Show that $Q_f \subset Q$ *and* $\phi|_{Q_f} \equiv f$.

Historically, vector fields have been referred to as *infinitesimal transformations*. The reasoning for this is made clear by the following lemma, which states that locally, the flow generated by a vector field is a diffeomorphism.

Lemma 6.11. Let X be a vector field on M and $\phi : Q \to M$ its local flow. Set

$$Q(t) := \{p \in M : (t,p) \in Q\}.$$

The set $Q(t)$ is open, $\phi_t(Q(t)) = Q(-t)$ and the mapping $\phi_t : Q(t) \to Q(-t)$ is a diffeomorphism.

Proof. The openness of $Q(t)$ follows immediately from the openness of Q. Now, if $p \in Q(t) \Leftrightarrow t \in I_p$, then Lemma 6.6 implies that $-t \in \phi_t(p)$, i.e. $\phi_t(p) \in Q(-t)$. Going the other way, if $q \in Q(-t) \Leftrightarrow -t \in I_q$, then Lemma 6.6 implies that $t \in I_{\phi_{-t}(q)}$, i.e. $p := \phi_{-t}(q) \in Q(t)$ so that $q = \phi_t(p) \in \phi_t(Q(t))$. The diffeomorphism property now follows from the smoothness of ϕ_t for any t together with Lemma 6.6. □

Remark 6.12. Note that $Q(t)$ might be empty for large enough t (consider $M =]-1,1[$ and $X(x) = \partial$). In general, we have that $Q(0) = M$, $M = \bigcup_{t>0} Q(t) = \bigcup_{t<0)} Q(t)$, and for $0 < t_1 < t_2$ (resp. $t_1 < t_2 < 0$),

$$Q(t_1) \supset Q(t_2) \qquad \text{(resp. } Q(t_1) \subset Q(t_2)\text{)}.$$

In other words, given any $p \in M$, if we restrict our attention to small enough t, the flow of a vector field will give rise to a diffeomorphism defined in an open neighbourhood of p. We shall use this fact implicitly in the sequel. △

We now turn our attention to a useful case where the flow generated by a vector field is a global one.

Lemma 6.13. Let $X : M \to TM$ be a vector field with *compact support*, i.e. suppose that $\operatorname{supp} X = \overline{\{p \in M : X(p) \neq 0\}}$ is compact. Then the flow ϕ generated by X is defined on all of $\mathbb{R} \times M$.

Proof. We retain the notation $Q = \bigcup_{p \in M} I_p \times \{p\}$ introduced earlier. First note that if $p \in M \backslash K$, $X(p) = 0$ so that by the uniqueness part of Theorem 6.5, we must have that $I_p = \mathbb{R}$ and $\phi_t(p) = p$ for all $t \in \mathbb{R}$. Now suppose $p \in K$. By Theorem 6.7, we may find a $\delta > 0$ such that $]-\delta, \delta[\subset I_p$ so that in fact for *any* $p \in M$, $I_p \supset]-\delta, \delta[$. We proceed by contradiction. Suppose $\beta := \sup I_p < \infty$ and consider the integral curve $s \mapsto \phi_s(\widetilde{p})$ passing through $\widetilde{p} := \phi_{\beta - \frac{\delta}{2}}(p)$ at $s = 0$. Since $]-\delta, \delta[\subset I_{\widetilde{p}}$ and $I_p - (\beta - \frac{\delta}{2}) \ni s \mapsto \phi_{s + (\beta - \frac{\delta}{2})}(p)$ is another integral curve of X passing through $\widetilde{p}$ at $s = 0$, we have that they must coïncide in a neighbourhood of 0. In particular, the mapping

$$I_p \cup \left] \beta - \frac{\delta}{2}, \beta + \frac{\delta}{2} \right[\ni t \mapsto \begin{cases} \phi_t(p), & t \in I_p \\ \phi_{t - (\beta - \frac{\delta}{2})}(\widetilde{p}), & \text{otherwise} \end{cases}$$

is an integral curve of X passing through p at $t = 0$ on an interval larger than I_p. This contradicts the maximality of I_p so that we must have $\beta = \infty$. We may argue similarly with $\alpha = \inf I_p$ to show that $\alpha = -\infty$. □

Remark 6.14. A vector field X whose induced flow is global is said to be *complete*. In particular, Lemma 6.13 implies that *all* vector fields on compact manifolds are complete. △

Example 6.15. If $M = \mathbb{R}^n$, then any vector field X with $|\partial_i X^j| \leq C$ or $|X^j| \leq C$ is complete. This is immediate from Theorem 6.4.

Example 6.16 (left-invariant vector fields on Lie groups). Let G be a Lie group and for fixed $g \in G$, let $\lambda_g : G \longrightarrow G$ denote multiplication from the left. A vector field $X : G \longrightarrow TG$ is said to be *left invariant* if for all $p \in G$, $\mathrm{d}_p\lambda_g X(p) = X(g \cdot p)$. We claim that X is complete.

Write e for the identity of G, ϕ for the flow of X and fix $\varepsilon > 0$ such that $]-\varepsilon, \varepsilon[\subset I_e$. We argue by contradiction, first supposing that $\beta := \sup I_p < \infty$ for some $p \in G$. Note that the curve

$$]-\varepsilon, \varepsilon[\ni s \longmapsto \lambda_{\phi_{\beta-\frac{\varepsilon}{2}}(p)}(\phi_s(e)) \in G$$

passes through $\phi_{\beta-\frac{\varepsilon}{2}}(p)$ at $s = 0$ and is an integral curve of X, since

$$\frac{\mathrm{d}}{\mathrm{d}s}\lambda_{\phi_{\beta-\frac{\varepsilon}{2}}(p)}(\phi_s(e)) = \mathrm{d}\lambda_{\phi_{\beta-\frac{\varepsilon}{2}}(p)}(X(\phi_s(e))) = X(\lambda_{\phi_{\beta-\frac{\varepsilon}{2}}(p)}(\phi_s(e)))$$

by left invariance. Therefore, this curve must coïncide with $s \mapsto \phi_{s-(\beta-\frac{\varepsilon}{2})}(p)$ in an interval about 0. The mapping

$$I_p \cup \left]\beta - \frac{\varepsilon}{2}, \beta + \frac{\varepsilon}{2}\right[\ni t \longmapsto \begin{cases} \phi_t(p), & t \in I_p \\ \lambda_{\phi_{\beta-\frac{\varepsilon}{2}}(p)}(\phi_{t-(\beta-\frac{\varepsilon}{2})}(e)), & \text{otherwise} \end{cases}$$

then yields the desired contradiction. We may argue similarly for $\alpha := \inf I_p$.

Theorem 6.17 (inverse function theorem). Let $f : M \longrightarrow N$ be a differentiable mapping between n-dimensional manifolds and suppose $\mathrm{d}_p f : T_pM \longrightarrow T_{f(p)}N$ is an isomorphism. Then there exists a neighbourhood $N(p) \subset M$ of p such that $f|_{N(p)} N(p) :\longrightarrow f(N(p))$ is a diffeomorphism.

Proof. Let $\varphi|_U$ be a chart about p and $\psi|_V$ a chart about $f(p)$. Then the condition that $\mathrm{d}_p f$ be an isomorphism is equivalent to the condition that

$$\det\left(\partial_j(\psi^i(\circ f \circ \varphi^{-1})(\varphi(p))\right) \neq 0.$$

By the classical inverse function theorem, there exist neighbourhoods U_1 of $\varphi(p)$ and U_2 of $\psi(f(p))$ contained in the respective coordinate systems such that $\psi \circ f \circ \varphi^{-1}\big|_{U_1} : U_1 \longrightarrow U_2$ is a diffeomorphism. The claim now follows by taking $N(p) = \varphi^{-1}(U_1)$. □

Lemma 6.18. Let $X : M \longrightarrow TM$ be a vector field and $p \in M$ such that $X(p) \neq 0$. Then there exists a coordinate system $\varphi|_U$ in a neighbourhood U of p such that on U,

$$X = \partial_1.$$

Proof. Fix a local parametrisation $F :]-\delta_0, \delta_0[^n \longrightarrow U \subset M$ with $F(0) = p$. We may assume without loss of generality that the vectors $X(p), \partial_2 F(0), \ldots, \partial_n F(0)$ are linearly independent and that $\overline{U}$ is compact. Let ϕ be the local flow generated by X. By Theorem 6.7, we may find a $\delta_1 < \delta_0$ such that $]-\delta_1, \delta_1[\subset I_p$ for $p \in U$. Now consider the mapping

$$\begin{aligned} \widetilde{F} :]-\delta_1, \delta_1[^n &\longrightarrow M \\ (t^1, \ldots, t^n) &\longmapsto \phi_{t^1}(F(0, t^2, \ldots, t^n)). \end{aligned}$$

Note that $\widetilde{F}(0) = p$, $\frac{\partial \widetilde{F}}{\partial t^1} = X \circ \widetilde{F}$ and for all $i > 1$, $\frac{\partial \widetilde{F}}{\partial t^i}(0) = \frac{\partial F}{\partial t^i}(0)$. Therefore, $\mathrm{d}_0\widetilde{F} : T_0\mathbb{R}^n \longrightarrow T_pM$ is an isomorphism so that we may find a $\delta_2 > 0$ so that $\widetilde{F}\Big|_{]-\delta_2,\delta_2[^n}$ is a diffeomorphism into its image. Therefore, $\widetilde{F}$ is a local parametrisation of M in a neighbourhood of p with respect to which $X = \frac{\partial \widetilde{F}}{\partial t^1}$ in this neighbourhood. The desired chart is given by its inverse. □

7. The Lie derivative. In this section we introduce a notion of differentiation that naturally arises from the flow of a vector field. Let $X : U \to TM$ be a (local) vector field and $\{X_t : Q(t) \to Q(-t)\}$ its local flow considered for $0 < t < t_0$ with $p \in Q(t_0) \cap U$.

Given a basis $\{b_i\}$ for T_pM, we naturally obtain a basis $\{b_i(t) := \mathrm{d}_pX_t(b_i)\}$ for $T_{X_t(p)}M$ such that $t \mapsto b_i(t)$ is a smooth vector field along the integral curve $t \mapsto X_t(p)$. It is not hard to see that if $\{b^i\} \subset T_p^*M$ is the dual basis to $\{b_i\}$, then for each t, $\{b^i(t) := \delta_{X_t(p)}X_{-t}(b^i)\} \subset T^*_{X_t(p)}M$ is dual to $\{b_i(t)\}$, and $t \mapsto b^i(t)$ is a covector field along $t \mapsto X_t(p)$. Altogether, we have obtained a *moving frame* together with a dual coframe along the integral curve $t \mapsto X_t(p)$ for sufficiently small t. We may of course expand a (local) tensor field $T : U \to T^r_sM$ in terms of this basis:

$$T(X_t(p)) = T^{i_1\dots i_r}{}_{j_1\dots j_s}(t) \cdot b_{i_1} \otimes \dots \otimes b_{i_r} \otimes b^{j_1} \otimes \dots \otimes b^{j_s}(t). \tag{1.11}$$

We now define the tensor field $\pounds_X T : U \to T^r_sM$ by differentiating the coefficients of this expansion, i.e.

$$\pounds_X T(p) := \dot{T}^{i_1\dots i_r}{}_{j_1\dots j_s}(0) \cdot b_{i_1} \otimes \dots \otimes b_{i_r} \otimes b^{j_1} \otimes \dots \otimes b^{j_s}. \tag{1.12}$$

Lemma 7.1. The definition of $\pounds_X T$ given by (1.12) is independent of the choice of basis $\{b_i\}_{i=1}^n$ and may be written as

$$\pounds_X T(p) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (X_t^* T)(p). \tag{1.13}$$

In local coordinates,

$$\begin{aligned}(\pounds_X T)^{i_1\dots i_r}{}_{j_1\dots j_s} \\ = \partial_X T^{i_1\dots i_r}{}_{j_1\dots j_s} - \sum_{l=1}^{r} \partial_k X^{i_l} \cdot T^{i_1\dots i_{l-1}k i_{l+1}\dots i_r}{}_{j_1\dots j_s} + \sum_{l=1}^{s} \partial_{j_l} X^k \cdot T^{i_1\dots i_r}{}_{j_1\dots j_{l-1}k j_{l+1}\dots j_s}.\end{aligned} \tag{1.14}$$

Proof. Note that the components of $T(X_t(p))$ in (1.11) may be written as

$$\begin{aligned}&T^{i_1\dots i_r}{}_{j_1\dots j_s}(t)\\ &= (T(X_t(p)), b_{j_1} \otimes \dots \otimes b_{j_s} \otimes b^{i_1} \otimes \dots \otimes b^{i_r}(t))\\ &= (T(X_t(p)), \mathrm{d}_pX_t(b_{j_1}) \otimes \dots \otimes \mathrm{d}_pX_t(b_{j_s}) \otimes \delta_{X_t(p)}X_{-t}(b^{i_1}) \otimes \dots \otimes \delta_{X_t(p)}X_{-t}(b^{i_r}))\\ &= ((\mathrm{d}_{X_t(p)}X_{-t} \otimes \dots \otimes \mathrm{d}_{X_t(p)}X_{-t} \otimes \delta_pX_t \otimes \dots \otimes \delta_pX_t)T(X_t(p)), b_{j_1} \otimes \dots \otimes b_{j_s} \otimes b^{i_1} \otimes \dots \otimes b^{i_r})\\ &= (X_t^*T(p), b_{j_1} \otimes \dots \otimes b_{j_s} \otimes b^{i_1} \otimes \dots \otimes b^{i_r}),\end{aligned}$$

where we have used the fact that $X_t^{-1} = X_{-t}$. The equality (1.13) now immediately follows.

To derive the local expression, fix a local coordinate system about some point p. Restricting our attention to sufficiently small t so that $X_t(p)$ lies in this coordinate system, we note that

$$\begin{aligned}&(X_t^*T)^{i_1\dots i_r}{}_{j_1\dots j_s}(p)\\ &\quad = \partial_{k_1}X^{i_1}_{-t}(X_t(p)) \dots \partial_{k_r}X^{i_r}_{-t}(X_t(p)) \cdot \partial_{j_1}X^{l_1}_t(p) \dots \partial_{j_s}X^{l_s}_t(p) \cdot T^{k_1\dots k_r}{}_{l_1\dots l_s}(X_t(p)).\end{aligned} \tag{1.15}$$

Noting that $\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} X^i_t(p) = X^i(p)$ and $X_0 = \mathrm{id}_M \Rightarrow \partial_i X^j_0 = \delta^j_i$, we compute that

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \partial_k X^j_{-t}(X_t(p)) = -\partial_k X^j(p).$$

Applying the product rule to the above coordinate expression then establishes the claim. □

Definition 7.2. The tensor field $\pounds_X T$ is called the *Lie derivative* of T in direction X.

Remark 7.3. In the case of a $(0,0)$-tensor field f, i.e. a smooth function, the above definitions reduce to $\pounds_X f = \partial_X f$. Therefore, the Lie derivative is in some sense an extension of directional differentiation. △

Remark 7.4. Note that

$$\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_t X_t^* T(p) = \frac{\mathrm{d}}{\mathrm{d}\sigma}\bigg|_{\sigma=0} (X_{t+\sigma}^* T)(p) = \frac{\mathrm{d}}{\mathrm{d}\sigma}\bigg|_{\sigma=0} (X_t^*(X_\sigma^* T))(p) = (\mathrm{d}_{X_t(p)} X_{-t} \otimes \delta_p X_t) \pounds_X T(X_t(p))$$

so that $I_p \ni t \mapsto (X_t^* T)(p)$ is constant for each $p \in U$ iff $\pounds_X T \equiv 0$. Therefore, the Lie derivative $\pounds_X T$ vanishes iff T is invariant under the flow of X. △

The local formula (1.14) hints at the true nature of the Lie derivative. As may be gleaned from this formula, $\pounds_X T(p)$ depends not only on the components of T at p and its directional derivative in direction $X(p)$, but also on the first derivative of X and p. In fact, we have the following nice characterisation of $\pounds_X$ acting on vector fields:

Theorem 7.5. Let $X, Y : U \to TM$ be vector fields. We have that

$$\pounds_X Y = [X, Y],$$

where $[X, Y]$ is the unique vector field satisfying

$$\partial_{[X,Y]} f = \partial_X \partial_Y f - \partial_Y \partial_X f$$

for $f \in C^\infty(U)$. Moreover, $[X, Y] \equiv 0$ iff the flows $\{X_t\}$ and $\{Y_s\}$ generated by X and Y commute wherever they are defined, i.e. $X_t \circ Y_s = Y_s \circ X_t$.

Proof. Note that by (1.15), we have in local coordinates

$$(\pounds_X Y)^i = \sum_{j=1}^n X^j \partial_j Y^i - Y^j \partial_j X^i.$$

On the other hand, it is a straightforward computation to check that the right-hand side is none other than $[X, Y]^i$.

Now consider for fixed t the mapping

$$Q \ni (s, p) \mapsto \widetilde{Y}(s, p) := (X_{-t} \circ Y_s \circ X_t)(p),$$

where $Q := \cup_{q \in M} J_q \times \{q\}$ is the maximal domain of definition of $\widetilde{Y}$ such that $J_q \ni 0$ is an open interval. Now, it is not hard to show that it is open. Moreover, it has all the characteristics of a local flow, except it may be that $\mathrm{pr}_2(Q) \neq M$. Nevertheless, whenever $(0, p) \in Q$, we may compute that

$$\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} \widetilde{Y}(s, p) = (X_t^* Y)(p) = Y(p),$$

since $\pounds_X Y \equiv 0$. Arguing as in Exercise 6.10 then implies that $Y_s(p) = X_{-t}(Y_s(X_t(p)))$, which establishes the claim. □

Remark 7.6. A (real or complex) vector space A together with a bilinear mapping

$$[\cdot, \cdot] : A \times A \to A$$

is said to be a *Lie algebra* if for all $X, Y, Z \in A$, we have

- $[X, Y] = -[Y, X]$ (anti-symmetry); and
- $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$ (Jacobi identity).

The space of all vector fields $M \to TM$ together with the Lie derivative (or bracket) $[X, Y] = \pounds_X Y$ is a Lie algebra in the above sense. △

We gather some useful properties of the Lie derivative in the following proposition. The proof of this proposition is a straightforward application of any one of the definitions of the Lie derivative.

Proposition 7.7. The Lie derivative enjoys the following properties for any given (local) vector field X and tensor fields S and T:

- $X \mapsto \pounds_X T$ is $\mathbb{R}$-linear.
- If $f : U \subset M \to \mathbb{R}$ is smooth, then
$$\pounds_X(fT) = \partial_X f \cdot T + f \cdot \pounds_X T.$$
- $\pounds_X(S \otimes T) = \pounds_X S \otimes T + S \otimes \pounds_X T$.
- $\pounds_X$ commutes with traces and the action of the symmetric group.
- If S and T are of the same type, then $\partial_X(S, T) = (\pounds_X S, T) + (S, \pounds_X T)$.

8. Submanifolds. Let S and M be smooth manifolds of dimension s and n respectively. A smooth mapping $F : S \to M$ is said to be an *immersion* if for each $s \in S$, $\mathrm{d}_s F$ is injective. If F is an injective immersion, it will be called an *embedding*. Just as in the classical case, F is viewed as an admissible parametrisation of a geometrically significant subset of our space, in this case M (cf. [8, §24]).

Definition 8.1. A *submanifold* of M is a subset of the form $F(S)$, where $F : S \to M$ is an embedding and S is smooth manifold. If in addition $F : S \to F(S)$ is a *topological embedding*, i.e. for all $U \subset S$ open there exists a $V \subset M$ open with $F(U) = V \cap F(S)$, the submanifold $F(S)$ will be called an *embedded* submanifold.

Remark 8.2. In order for a mapping $F(S)$ to be a submanifold in the above sense, we must have $\dim S \leq \dim M$. The number $\dim M - \dim S$ is referred to as the *codimension* of $F(S)$. Some noteworthy cases are the following:

- If $\dim S = 1$, we call $F(S)$ a (regular) *curve*. The image of any compact connected subset of S under F is an *arc* of $F(S)$. In general, we shall usually refer to curves, arcs of curves *and* the parametrisations thereof simply as *curves*. If a curve is given as a parametrisation $K \to M$ with $K \subset \mathbb{R}$ a closed interval, we assume that F may be smoothly extended to an open interval so as to parametrise a regular curve.
- If $\dim S = 2$, we call $F(S)$ a *surface*.
- If $\dim S = \dim M - 1$, we call $F(S)$ a *hypersurface*.

△

Remark 8.3. If $F(S)$ is a submanifold of M, the fact that $\mathrm{d}_s F : T_s S \to T_{F(s)} M$ is injective gives rise to an isomorphism $T_s S \simeq \operatorname{im} \mathrm{d}_s F$. In particular, a vector $v = \sum_{i=1}^n v^i \, \partial_i|_s \in T_s S$ is identified with $\mathrm{d}_s F(v) = \sum_{i=1}^n v^i \partial_i F(s) \in T_{F(s)} M$. This should be compared with the definition of the tangent space of a (hyper)surface in Euclidean space [8, §26]. △

Exercise 8.4. Suppose $\dim S = \dim M$ and $F : S \to M$ is an embedding. Show that F is a diffeomorphism onto its image so that $F(S)$ is an open submanifold of M. In particular, such submanifolds are necessarily embedded.

Exercise 8.5. Suppose S is a compact manifold and $F : S \to M$ an embedding. Show that $F(S)$ is an embedded submanifold, and F is *proper* in the following sense: Whenever $K \subset M$ is compact, $F^{-1}(K)$ is compact.

In light of Exercise 8.4, we shall restrict our attention to the case where $\dim S < \dim M$.

Example 8.6 (circle). We consider the realisation of the circle as the quotient of the additive group $\mathbb{R}$ by the (normal) subgroup $\mathbb{Z}$ of integers:

$$T^1 := \mathbb{R}/\mathbb{Z} = \{x + \mathbb{Z} : x \in \mathbb{R}\}.$$

Let $\pi : \mathbb{R} \to T^1$ denote the canonical projection $x \mapsto x + \mathbb{Z}$. It is readily checked that for any $a, b \in \mathbb{R}$ with $0 < b - a \leq 1$, $\pi|_{]a,b[}$ is injective, and $T^1 = \pi(]a, b[) \cup \pi(\left]a + \frac{1}{2}, b + \frac{1}{2}\right[)$. In particular, the collection

$$\{\pi|_{]0,1[}^{-1}, \pi|_{]\frac{1}{2},\frac{3}{2}[}^{-1}\}$$

forms a 1-dimensional smooth atlas making T^1 a 1-dimensional manifold; this may be readily seen by noting that $\pi(]0, 1[) = T^1 \setminus \{\pi(0)\}$, $\pi(\left]\frac{1}{2}, \frac{3}{2}\right[) = T^1 \setminus \{\pi(\frac{1}{2})\}$, which are mapped by both mappings in our collection to two disjoint intervals, and

$$(\pi|_{]0,1[}^{-1} \circ \pi|_{]\frac{1}{2},\frac{3}{2}[})(x) = x - \lfloor x \rfloor,$$

which is smooth on the aforementioned subintervals. The Hausdorff axiom need only be checked for $\pi(0) = \pi(1)$ and $\pi(\frac{1}{2})$: $\pi(\left]\frac{3}{4}, \frac{5}{4}\right[) \cap \pi(\left]\frac{1}{4}, \frac{3}{4}\right[) = \varnothing$. Therefore, T^1 equipped with the above smooth atlas is a 1-dimensional manifold. Moreover, since $\pi : \mathbb{R} \to T^1$ is smooth and $T^1 = \pi([0, 1])$, we see that T^1 is compact. Owing to the periodicity of $\sin$ and $\cos$, we have the mapping

$$\begin{aligned} F : T^1 &\to \mathbb{R}^2 \\ \pi(x) &\mapsto (\cos(2\pi x), \sin(2\pi x)). \end{aligned}$$

It may be readily checked using the above charts that F is a smooth embedding. Since T^1 is compact, it follows from Exercise 8.5 that $F(T^1) = S^1$ is an embedded submanifold.

Remark 8.7. The smooth atlas on S^1 induced by that of T^1, i.e.

$$\{\pi|_{]0,1[}^{-1} \circ F^{-1}, \pi|_{]\frac{1}{2},\frac{3}{2}[}^{-1} \circ F^{-1}\}$$

may be shown to be equivalent to the one introduced in Exercise 1.7. △

Example 8.8 (products of embeddings). Suppose that S_1 and S_2 are smooth manifolds of dimension m_1 and m_2 respectively and M_1 and M_2 smooth manifolds of dimension n_1 and n_2. If $F_1 : S_1 \to M_1$ and $F_2 : S_2 \to M_2$ are (topological) embeddings, then the product map

$$\begin{aligned} F := F_1 \times F_2 : S_1 \times S_2 &\to M_1 \times M_2 \\ (q_1, q_2) &\mapsto (F_1(q_1), F_2(q_2)) \end{aligned}$$

is also a (topological) embedding: Since the manifold topologies of $S_1 \times S_2$ and $M_1 \times M_2$ coïncide with the respective product topologies, it is clear that F is a continuous injection (and a homeomorphism if F_1 and F_2 are topological embeddings). The fact that it is a smooth follows from the fact that products of smooth maps are smooth, since

$$((\varphi_1 \times \varphi_2) \circ F \circ (\psi_1 \times \psi_2)^{-1}) = (\varphi_1 \circ F_1 \circ \psi_1^{-1}) \times (\varphi_2 \circ F_2 \circ \psi_2^{-1})$$

is smooth for local coordinate systems φ_1 and φ_2 on M_1 and M_2 and local coordinate systems ψ_1 and ψ_2 on S_1 and S_2. Finally, to check that F is an immersion, we use the natural identifications of Exercise 2.8: If $v = \mathrm{d}_{q_1} \overset{1}{\iota}_{q_2}(v_1) + \mathrm{d}_{q_2} \overset{2}{\iota}_{q_1}(v_2) \in T_{(q_1,q_2)} S_1 \times S_2$ is such that $\mathrm{d}_{(q_1,q_2)} F(v) = 0$, then since $\mathrm{pr}_i \circ F = F_i \circ \mathrm{pr}_i$, the chain rule implies that

$$0 = \mathrm{d}_{F(q_1,q_2)} \mathrm{pr}_i(\mathrm{d}_{(q_1,q_2)} F(v)) = \mathrm{d}_{q_i} F_i(\mathrm{d}_{(q_1,q_2)} \mathrm{pr}_i(v)) = \mathrm{d}_{q_i} F_i(v_i)$$

so that $v_1 = 0$ and $v_2 = 0$ since F_1 and F_2 are immersions.

Example 8.9 (n-torus)**.** The *n-torus* is defined as

$$T^n := \underbrace{T^1 \times \cdots \times T^1}_{n \text{ times}}.$$

Since the product manifold topology coïncides with the product topology, T^n is also a compact manifold, and we have a natural smooth projection

$$\begin{aligned} \pi_n : \mathbb{R}^n &\longrightarrow T^n \\ (x^1, \ldots, x^n) &\longmapsto (\pi(x^1), \ldots, \pi(x^n)). \end{aligned}$$

Taking the product of the embedding of Example 8.6 as in Example 8.8, we obtain a (topological) embedding $T^n \to \mathbb{R}^{2n}$.

Example 8.10 (sphere)**.** Consider the inclusion $\iota : S^n \hookrightarrow \mathbb{R}^{n+1}$, i.e. $p \mapsto p$. Using the stereographic projection charts of Example 1.7, it may be shown that this mapping makes S^n an *embedded* submanifold of $\mathbb{R}^{n+1}$.

The following example illustrates that a submanifold need not be an embedded submanifold.

Example 8.11 (figure 8)**.** Consider the mapping $x : \mathbb{R} \to \mathbb{R}^2$ such that $F(t) = (\frac{3t}{1+t^4}, \frac{3t^3}{1+t^4})$. It is clear that F is smooth and straightforward to check that it is an embedding ($F'(t)$ never vanishes). Therefore, its image Fig-8 $:= F(\mathbb{R})$ is a submanifold of $\mathbb{R}^2$. However, it is *not* an embedded submanifold: There does not exist an open $U \subset \mathbb{R}^2$ such that $F(]-1,1[) = U \cap$ Fig-8, for if there did, then we would have $B(0,\varepsilon) \cap$ Fig-8 $\subset F(]-1,1[)$ for some $\varepsilon > 0$, but $F(t) \to 0$ as $t \to \infty$ so that there exists an $N(\varepsilon) > 0$ such that $F(t) \in B(0,\varepsilon) \cap$ Fig-8 for all $t > N(\varepsilon)$, but $F(t) \notin F(]-1,1[)$!

We first note that a submanifold may be viewed locally in its domain as forming a *slice* in the target manifold in the following sense:

Theorem 8.12. Let S be a smooth m-manifold and $F : S \to M$ an embedding. For each $s \in S$, there exists a neighbourhood U_s of s and a chart $\psi_s|_{V_s}$ on a neighbourhood V_s of $F(s)$ such that the following hold:

- $\psi_s(F(s)) = 0$;
- $F(U_s) = \psi_s^{-1}(\mathbb{R}^m \times \{0\})$; and

- $\underline{\psi}_s := \mathrm{pr}_{\mathbb{R}^m} \circ \psi_s \circ F : U_s \to \mathrm{pr}_{\mathbb{R}^m}(\psi_s(F(U_s)))$ is a chart on U_s.

The theorem is a consequence of the following application of the inverse function theorem.

Lemma 8.13. Let $F : U \to \mathbb{R}^{m+k}$ be a smooth mapping with $0 \in U$ and $\mathrm{rank}(\partial_j F^i)(0) = m$. There exists a diffeomorphism $G : W \subset \mathbb{R}^{m+k} \to U' \times V \subset \mathbb{R}^{m+k}$ with $0 \in U' \subset U$, $W \supset F(U')$ and $V \subset \mathbb{R}^k$ such that for all $x \in U'$, $(G \circ F)(x) = (x, 0)$, where.

Proof. Let $j_1, \ldots, j_k \in \{1, \ldots, m+k\}$ be chosen such that the vectors

$$\{\partial_1 F(0), \ldots, \partial_m F(0), e_{j_1}, \ldots, e_{j_k}\}$$

are $\mathbb{R}$-linearly independent and define

$$\begin{aligned} \widetilde{F} : U \times \mathbb{R}^k &\to \mathbb{R}^{m+k} \\ (x^1, \ldots, x^{m+k}) &\mapsto F(x^1, \ldots, x^m) + \sum_{i=1}^{k} x^{m+i} e_{j_i}. \end{aligned}$$

It follows immediately that $\det(\partial_j \widetilde{F}^i)(0) \neq 0$ so that by the inverse function theorem, we may find open sets $U' \subset U$, $V \subset \mathbb{R}^k$ and $W \subset \mathbb{R}^{m+k}$ with $0 \in U' \times V$ and $F(0) \in W$, and a smooth mapping $G : W \to U' \times V$ such that $G = \widetilde{F}\Big|_{U' \times V}^{-1}$. In particular, for all $x \in U'$,

$$G(F(x)) = G(\widetilde{F}(x, 0)) = (x, 0). \qquad \square$$

Proof of Theorem 8.12. Fix a coordinate system $\varphi|_U$ about s and $\psi|_V$ about $F(s)$ such that $\varphi(s) = 0$ and $\psi(F(s)) = 0$. Then the mapping

$$\psi \circ F \circ \varphi^{-1}\big|_{\varphi(F^{-1}(V) \cap U)} : \varphi(F^{-1}(V) \cap U) \subset \mathbb{R}^m \to V \subset \mathbb{R}^n$$

satisfies the hypotheses of Lemma 8.13 so that there are open neighbourhoods $U' \subset F^{-1}(V) \cap U$, $V' \subset \mathbb{R}^{n-m}$ and $W \subset \mathbb{R}^n$ of the origin (in each respective space) as well as a diffeomorphism $\Phi : W \to U' \times V'$ such that $(\psi \circ F \circ \varphi^{-1})(U') = \Phi^{-1}(U' \times \{0\})$ and for $(x^1, \ldots, x^n) \in U'$,

$$\Phi \circ \psi \circ F \circ \varphi^{-1}(x^1, \ldots, x^m) = (x^1, \ldots, x^m, 0),$$

or $\varphi|_{\varphi^{-1}(U')} = \mathrm{pr}_{\mathbb{R}^m} \circ \Phi \circ \psi \circ F\big|_{\varphi^{-1}(U')}$. The theorem now follows by setting $U_s := \varphi^{-1}(U')$, $V_s := \psi^{-1}(W) \cap V$ and $\psi_s := \Phi \circ \psi|_{V_s}$. $\square$

Note that Theorem 8.12 does *not* guarantee that we obtain a single m-dimensional slice if we localise in the target, i.e. we do not necessarily have that

$$F(S) \cap V = \psi^{-1}(\mathbb{R}^m \times \{0\}) \tag{1.16}$$

for a suitable chart $\psi|_V$ about $F(s)$ for $s \in S$. However, if our submanifold is *embedded*, then any $U \subset S$ open may be written in the form $F(U) = V \cap F(S)$ for $V \subset M$ open, whence (1.16) does actually hold in this case. It turns out that the condition (1.16) is in fact equivalent to F being a topological embedding. This is a consequence of the following theorem.

Theorem 8.14. Let S be a set and $F : S \to M$ an injective mapping with the following property:

- There exists a natural number $m < n$ such that for all $s \in S$, there is a chart $\psi|_{V_s}$ about $F(s)$ such that $F(S) \cap V_s = \psi_s^{-1}(\mathbb{R}^m \times \{0\})$.

Then S admits a unique smooth m-dimensional manifold structure with respect to which the mappings

$$\{\underline{\psi}_s := (\psi_s^1, \ldots, \psi_s^m) \circ F\big|_{F^{-1}(V_s)}\}_{s \in S}$$

are diffeomorphisms and the mapping $F : S \to F(S)$ is a smooth *topological* embedding.

Proof. Note first that the $\{V_s : s \in S\}$ form an open cover of $F(S)$, where $F(S)$ is equipped with the subspace topology. Since M (thus also $F(S)$) is Lindelöf, we may pass to a countable subcover $\{V_{s_i}\}$ of $F(S)$. We now claim that the $\{\underline{\psi}_{s_i}\big|_{U_i}\}$ form a smooth m-dimensional atlas, where $U_i := F^{-1}(V_{s_i})$. The openness condition may be checked thus:

$$\begin{aligned}\underline{\psi}_{s_i}(U_i \cap U_j) &= \mathrm{pr}_{\mathbb{R}^m}(\psi_{s_i}(F(S) \cap V_{s_i} \cap V_{s_j})) \\ &= \mathrm{pr}_{\mathbb{R}^m}(\underbrace{\psi_{s_i}(F(S) \cap V_{s_i})}_{\text{open in } \mathbb{R}^m \times \{0\}} \cap \underbrace{\psi_{s_i}(V_{s_i} \cap V_{s_j})}_{\text{open in } \mathbb{R}^n}).\end{aligned}$$

On the other hand, to check the coordinate transformations, we note that for $x \in \underline{\psi}_{s_j}(U_i \cap U_j) \subset \mathbb{R}^m$, $\underline{\psi}_{s_j}(s) = x \Leftrightarrow \psi_{s_j}(F(s)) = (x, 0)$ so that $\underline{\psi}_{s_j}^{-1}(x) = F^{-1}(\psi_{s_j}^{-1}(x, 0))$. Therefore,

$$(\underline{\psi}_{s_i} \circ \underline{\psi}_{s_j}^{-1})(x) = (\mathrm{pr}_{\mathbb{R}^m} \circ \psi_{s_i} \circ \psi_{s_j}^{-1})(x),$$

which smoothly depends on x. We therefore have a smooth atlas. We now check that F is a topological embedding; this will immediately imply that the topology τ_S is Hausdorff so that S is indeed a smooth manifold, from which our smoothness claims will immediately follow. For continuity, suppose $V \subset M$ open. Then for each i,

$$\underline{\psi}_{s_i}(F^{-1}(V) \cap U_i) = (\mathrm{pr}_{\mathbb{R}^m} \circ \psi_{s_i})(F(S) \cap V_{s_i}) \cap (\mathrm{pr}_{\mathbb{R}^m} \circ \psi_{s_i})(V \cap V_{s_i})$$

is open, since it is the intersection of two open sets. Therefore, $F^{-1}(V) \in \tau_S$. As for the topological embedding property of F, if $U \in \tau_S$, we have that

$$\begin{aligned} &(\forall i) \qquad \underline{\psi}_{s_i}(U \cap U_i) \text{ open} \\ \Leftrightarrow\ &(\forall i) \qquad \mathrm{pr}_{\mathbb{R}^m}(\psi_{s_i}(F(U) \cap V_{s_i})) \text{ open} \\ \Leftrightarrow\ &(\forall i)(\exists W_i \text{ open}) \qquad \psi_{s_i}(F(U) \cap V_{s_i}) = W_i \cap (\mathbb{R}^m \times \{0\}).\end{aligned}$$

Taking the inverse image by ψ_{s_i} of both sides of the last implication, we arrive at the equality

$$F(U) \cap V_{s_i} = \underbrace{\psi_{s_i}^{-1}(W_i) \cap V_{s_i}}_{\text{open}} \cap F(S).$$

Taking the union over i then implies that $F(U) = F(S) \cap V$ for some $V \subset M$ open. □

As an alternative to describing submanifolds in terms of parametrisations, we may also describe them in terms of level sets. It turns out that such submanifolds are *always* embedded.

Theorem 8.15. Let S be an m-dimensional smooth manifold and $f : M \to S$ a smooth mapping such that

- for some $s \in S$, $f^{-1}(\{s\}) \neq \varnothing$; and
- $\mathrm{d}_p f$ is surjective for each $p \in f^{-1}(\{s\})$.

Then for each $p \in f^{-1}(\{s\})$, there is a chart $\psi_p|_{U_p}$ about p such that

$$f^{-1}(\{s\}) \cap U_p = \psi_p^{-1}(\mathbb{R}^{n-m} \times \{0\}).$$

In particular, $f^{-1}(\{s\}) \subset M$ is naturally an $(n-m)$-dimensional embedded submanifold.

Again, we require an inverse function theorem-related lemma.

Lemma 8.16. Let $U \subset \mathbb{R}^{n+k}$ be open and $f : U \to \mathbb{R}^k$ a smooth mapping such that $f^{-1}(\{0\}) \neq \varnothing$ and $\operatorname{rank}(\partial_j f^i)(x) = k$ for all $x \in f^{-1}(\{0\})$. Then for each $x_0 \in f^{-1}(\{0\})$, there exists a diffeomorphism $G : V \times W \subset \mathbb{R}^{n+k} \to U'$ with $x_0 \in U' \subset \mathbb{R}^{n+k}$, $W \subset \mathbb{R}^k$ with $0 \in W$ and $V \subset \mathbb{R}^n$ such that

$$f^{-1}(0) \cap U' = G(V \times \{0\}).$$

Proof. Note that in this case the condition on the rank of the differential of F equates to the linear independence of the covectors $\{\partial F^i(x_0)\}_{i=1}^k$. Without loss of generality, suppose that the covectors $\{e^1, \dots, e^n\} \cup \{\partial F^i(x_0)\}_{i=1}^k$ are linearly independent (otherwise replace f with $f \circ \sigma$ with σ an appropriate permutation of the $x^1, \dots, x^{n+k}$. Define the function

$$\begin{aligned} \widetilde{f} : U &\to \mathbb{R}^{n+k} \\ x &\mapsto (x^1, \dots, x^n, f(x)). \end{aligned}$$

It follows that $\det(\partial_j \widetilde{f}^i)(x_0) \neq 0$ so that by the inverse function theorem, we may find open sets $U' \subset U$, $V \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^k$ with $x_0 \in U'$ and $\widetilde{f}(x_0) \in V \times W$, and a C^l mapping $G : V \times W \to U'$ such that $G = \widetilde{f}\big|_{U'}^{-1}$. In particular, we have that

$$G^{-1}(f^{-1}(\{0\}) \cap U') = \widetilde{f}(f^{-1}(0) \cap U') = \{(v, w) \in V \times W : w = 0\} = V \times \{0\}. \qquad \square$$

Proof of Theorem 8.15. Let $\psi|_V$ be a chart about s with $\psi(s) = 0$ and $\varphi|_U$ a chart about $p \in f^{-1}(\{s\})$. Then $\psi \circ f \circ \varphi^{-1} : \varphi(U \cap f^{-1}(V)) \subset \mathbb{R}^n \to \mathbb{R}^m$ satisfies the hypotheses of Lemma 8.16 so that there exist open sets $U_p \ni p$, $V \subset \mathbb{R}^n$ and $W \subset \mathbb{R}^k$ with $0 \in W$ as well as a diffeomorphism $G : V \times W \to \varphi(U_p)$ such that

$$(\psi \circ f \circ \varphi^{-1})^{-1}(\{0\}) \cap \varphi(U_p) = G(V \times \{0\}) \Leftrightarrow f^{-1}(\{s\}) \cap U_p = (\varphi^{-1} \circ G)(V \times \{0\}).$$

Setting $\psi_p := G^{-1} \circ \varphi|_{U_p} : U_p \to V \times W$ establishes the claim. $\square$

Remark 8.17. Let f be as in Theorem 8.15 and write $\iota : f^{-1}(s) \to M$ for the inclusion mapping, which is a (topological) embedding. By Remark 8.3, we may identify $T_p f^{-1}(s)$ with $\operatorname{im} \mathrm{d}_p \iota$. Since $f \circ \iota \equiv s$, we obtain upon differentiating that $\mathrm{d}_p f \circ \mathrm{d}_p \iota \equiv 0$, i.e. $\operatorname{im} \mathrm{d}_p \iota \subset \ker \mathrm{d}_p f$; on comparing the dimensions of these vector spaces, we conclude that $\operatorname{im} \mathrm{d}_p \iota = \ker \mathrm{d}_p f$ so that $T_p f^{-1}(s) \simeq \ker \mathrm{d}_p f$. △

Given a smooth mapping $\Phi : M \to N$ between manifolds, it is clear that for a submanifold $F(S) \subset M$, the mapping $\Phi \circ F : S \to N$ yields a smooth mapping. However, suppose instead that $\Phi : N \to M$ is a smooth map such that for each $p \in N$, $\Phi(p) \in F(S)$. We then obtain a unique mapping $\underline{\Phi} : N \to S$ completing the following diagram:

$$
\begin{array}{ccc}
N & \xrightarrow{\Phi} & M \\
& \underline{\Phi} \searrow & \uparrow F \\
& & S
\end{array}
\tag{1.17}
$$

The question is, is $\underline{\Phi}$ smooth? In general, the answer is no: Consider the embedding F of Example 8.11 and the mapping $\Phi : \mathbb{R} \longrightarrow \mathbb{R}^2$ such that $t \mapsto (\frac{3t^3}{1+t^4}, \frac{3t}{1+t^4})$. In this case, $\underline{\Phi} : \mathbb{R} \longrightarrow \mathbb{R}$ takes the following form:

$$
\underline{\Phi}(t) = \begin{cases} \frac{1}{t}, & t \neq 0 \\ 0, & t = 0 \end{cases}
$$

It is clear that $\underline{\Phi}$ is not smooth, as it is not even continuous! However, as it turns out, if $\underline{\Phi}$ is continuous, then it is automatically smooth, and this is *always* the case for an embedded submanifold $F(S)$.

Lemma 8.18. Let $F(S) \subset N$ be a submanifold and $\Phi : M \longrightarrow N$ a smooth mapping such that $\Phi(M) \subset F(S)$. The mapping $\underline{\Phi}$ completing diagram (8) is smooth if and only if it is continuous. In particular, if $F(S)$ is an embedded submanifold, $\underline{\Phi}$ is always smooth.

Proof. The 'only if' part is immediate from the definition of smoothness. Suppose $\underline{\Phi}$ is continuous, i.e. for every open $V \subset S$, $\underline{\Phi}^{-1}(V) = \Phi^{-1}(F(V))$ is open in M. Picking a chart $\varphi|_U$ about $p \in M$ and an *adapted* chart $\underline{\psi}\big|_{U_s}$ about $s := \underline{\Phi}(p)$ as in Theorem (8.12), we see that the mapping

$$
\underline{\psi} \circ \underline{\Phi} \circ \varphi^{-1} : \varphi(U \cap \underline{\Phi}^{-1}(U_s)) \longrightarrow \underline{\psi}(U_s)
$$

may be written as $\underline{\psi} \circ \underline{\Phi} \circ \varphi^{-1} = \mathrm{pr}_{\mathbb{R}^m} \circ \psi \circ F \circ \underline{\Psi} \circ \varphi^{-1} = \mathrm{pr}_{\mathbb{R}^m} \circ (\psi \circ \Phi \circ \varphi^{-1})$, but the parenthetical function is smooth since Φ is smooth, hence $\underline{\Phi}$ is smooth. In the case where $F(S)$ is an embedded submanifold, for $V \subset S$ open, we have that $F(V) = W \cap F(S)$ for some $W \subset N$ open so that $\underline{\Phi}^{-1}(V) = \Phi^{-1}(F(V)) = \Phi^{-1}(W \cap F(S)) = \Phi^{-1}(W)$, which is open. □

Example 8.19 (orthogonal group)**.** The real *orthogonal group* is defined by

$$
\mathrm{O}(n, \mathbb{R}) = \{A \in \mathrm{GL}(n, \mathbb{R}) : A^T A = I\}.
$$

$\mathrm{O}(n, \mathbb{R})$ is closed under matrix multiplication, thus a subgroup of $\mathrm{GL}(n, \mathbb{R})$. We now show that it is a submanifold, and the group structure induced by that of $\mathrm{GL}(n, \mathbb{R})$ equips $\mathrm{O}(n, \mathbb{R})$ with the structure of a Lie group.

Note that the mapping $\mathrm{GL}(n, \mathbb{R}) \ni A \mapsto A^T A \in M_{n\times n}(\mathbb{R})$ actually maps into the $\frac{1}{2}n(n+1)$-dimensional vector subspace $M^{\mathrm{sym}}_{n\times n}(\mathbb{R})$ of real symmetric $n \times n$ matrices. Call this mapping f. Since $\mathrm{O}(n, \mathbb{R}) = f^{-1}(I) \ni I$, it suffices to show that $\mathrm{d}_A f : T_A \mathrm{GL}(n, \mathbb{R}) \longrightarrow T_I M^{\mathrm{sym}}_{m\times m}(\mathbb{R})$ is surjective. We make use of the treble clef isomorphism: Note that for $B \in M_{n\times n}(\mathbb{R})$, the calculational trick of Remark 2.6 yields

$$
(\text{𝄞}_I^{-1} \circ \mathrm{d}_A f \circ \text{𝄞}_A)(B) = \mathrm{D}_A f(B) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} (A + tB)(A + tB)^T = BA^T + AB^T.
$$

For fixed $X \in M^{\text{sym}}_{n\times n}(\mathbb{R})$, we see that $B = \frac{1}{2}X(A^T)^{-1}$ satisfies

$$BA^T + AB^T = \frac{1}{2}X(A^T)^{-1}A^T + \frac{1}{2}AA^{-1}X^T = \frac{1}{2}(X + X^T) = X$$

so that $\mathrm{D}_A f$ and thus also $\mathrm{d}_A f$ is surjective. Therefore, $\mathrm{O}(n, \mathbb{R})$ is an embedded submanifold of dimension $n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1)$, and the multiplication and inversion mappings are smooth by Lemma 8.18. Note that we may also use the above treble clef isomorphisms and Remark 8.17 to explicitly describe $T_I\mathrm{O}(n, \mathbb{R})$ in terms of matrices: Since $T_I\mathrm{O}(n, \mathbb{R}) \simeq \ker \mathrm{d}_I f \simeq \ker \mathrm{D}_I f$, we see that

$$\mathrm{D}_I f(B) = 0 \Leftrightarrow B + B^T = 0 \Leftrightarrow B \text{ skew symmetric}$$

so that $T_I\mathrm{O}(n, \mathbb{R}) \simeq \{B \in M_{n\times n}(\mathbb{R}) : B \text{ skew symmetric}\}$. Note that for $A \in \mathrm{O}(n, \mathbb{R})$, $A^T A = I \Rightarrow (\det A)^2 = 1$. In particular, we have the *special orthogonal group*

$$\mathrm{SO}(n, \mathbb{R}) := \{A \in \mathrm{O}(n, \mathbb{R}) : \det A = 1\} = \det|^{-1}_{\mathrm{O}(n,\mathbb{R})}(]0, \infty[),$$

which is an open submanifold of $\mathrm{O}(n, \mathbb{R})$, thus also a Lie group.

Affinely connected spaces

In the preceding chapter we introduced smooth manifolds and studied some structures naturally arising from their atlases. In this chapter, we turn our attention to structures that may be imposed on them so as to allow us to speak of such notions as *parallelism* and *distance*. Comparison with the corresponding structures in Euclidean space then gives rise to a notion of *curvature*. General references are [4, 7, 9] (modern) and [11, 5, 6, 8] (classical). For notational convenience, we introduce the following sets, given smooth manifolds M and N, an open set $U \subset M$ and a mapping $\gamma : N \longrightarrow M$:

$$\begin{aligned} \Gamma^r_s M &:= \{\text{smooth } (r,s)\text{-tensor fields } T : M \longrightarrow T^r_s M\} \\ \Gamma^r_s M\big|_U &:= \{\text{smooth } (r,s)\text{-tensor fields } T : U \longrightarrow T^r_s M\} \\ {}_\gamma\Gamma^r_s M &:= \{(r,s)\text{-tensor fields } T : N \longrightarrow T^r_s M \text{ along } \gamma\} \end{aligned}$$

We shall assume that $T \in {}_\gamma\Gamma^r_s M$ is as smooth as γ.[1]

The following lemma gives us a useful criterion for determining when a function of vector fields and covector fields arises from a tensor field (cf. [4, §3.2]).

Lemma 8.20. Let $f : (\Gamma^0_1 M)^r \times (\Gamma^1_0 M)^s \longrightarrow C^\infty(M)$ be a $C^\infty(M)$-multilinear mapping. Then there exists a unique $T \in \Gamma^r_s M$ such that for all $X_1, \ldots, X_s \in \Gamma^1_0 M$, $\alpha_1, \ldots, \alpha_r \in \Gamma^0_1 M$ and $p \in M$,

$$(T(p), (X_1 \otimes \cdots \otimes X_s \otimes \alpha_1 \otimes \cdots \otimes \alpha_r)(p)) = f(\alpha_1, \ldots, \alpha_r, X_1, \ldots, X_s)(p).$$

9. Affine connections. Recall that in Euclidean space we have a notion of directional differentiation: Given $X \in \Gamma^1_0 \mathbb{R}^n$ and $v \in T_p\mathbb{R}^n$, it is simply $\nabla_v X = v^i \partial_i X^j \partial_j$ with respect to the canonical chart. Another way of viewing this is that we have a natural mapping $\Gamma^1_0 \mathbb{R}^n \longrightarrow \Gamma^1_1 \mathbb{R}^n$ taking the vector field X^i to the tensor $X^i{}_{;j} := \frac{\partial X^i}{\partial x^j}$. We wish to adapt this notion to the case of an arbitrary manifold.

Let M be an n-dimensional manifold with atlas $\mathcal{A}$. As a first step, we take a vector field $X \in \Gamma^1_0 M$ and form the quantity $X^i{}_{;j}$ in a coordinate system $\varphi|_U = (x^1, \ldots, x^n) \in \mathcal{A}_{\max}$ and ask whether it is a tensor, i.e. given another chart $\bar\varphi|_{\bar U} = (\bar x^1, \ldots, \bar x^n)$, whether the equality

$$X^i{}_{;j}\partial_i \otimes \mathrm{d}x^j = \bar X^i{}_{;j}\bar\partial_i \otimes \mathrm{d}\bar x^j \tag{2.1}$$

holds on $U \cap \bar U$; by (1.4) and (1.8), this equality is equivalent to

$$X^i{}_{;j} = \frac{\partial x^i}{\partial \bar x^k} \cdot \frac{\partial \bar x^l}{\partial x^j} \cdot \bar X^k{}_{;l}.$$

We compute using the transformation rules of X^i and the coordinate bases that

$$X^i_{;j} = \frac{\partial X^i}{\partial x^j} = \frac{\partial \bar x^k}{\partial x^j} \cdot \frac{\partial x^i}{\partial \bar x^l} \cdot \bar X^l{}_{;k} - \frac{\partial x^i}{\partial \bar x^l}\frac{\partial x^q}{\partial \bar x^p} \cdot \frac{\partial \bar x^l}{\partial x^j \partial x^q} \cdot \bar X^p. \tag{2.2}$$

Since $\partial_k\partial_l \bar x^i \neq 0$ in general, we see that (2.1) does not hold true. However, in the case of Euclidean space, (2.2) shows us how directional differentiation appears in a *curvilinear* coordinate system. At worst, it may be expressed by the usual formula plus a term linear in the vector field. This is our point of departure.

[1] In later sections, we will consider piecewise smooth curves γ and piecewise smooth tensor fields along such curves.

Definition 9.1. An *affine connection* L on $(M, [\mathcal{A}])$ is given by a collection of n^3 (smooth) functions

$$\{L^i_{jk} : U \to \mathbb{R}\}^n_{i,j,k=1}$$

on each coordinate system $\varphi|_U$, called the *coefficients of the connection* with respect to φ, subject to the following transformation law: Whenever $\{L^i_{jk}\}$ and $\{\bar{L}^i_{jk}\}$ are the coefficients of the connection with respect to the coordinate systems $\varphi|_U = (x^1, \ldots, x^n)$ and $\bar{\varphi}|_{\bar{U}} = (\bar{x}^1, \ldots, \bar{x}^n)$ respectively, then

$$\bar{L}^i_{jk} = \frac{\partial \bar{x}^i}{\partial x^p} \cdot \frac{\partial x^q}{\partial \bar{x}^j} \cdot \frac{\partial x^r}{\partial \bar{x}^k} \cdot L^p_{qr} + \frac{\partial \bar{x}^i}{\partial x^r} \cdot \frac{\partial^2 x^r}{\partial \bar{x}^j \partial \bar{x}^k} \tag{2.3}$$

holds on $U \cap \bar{U}$. A manifold together with an affine connection is said to be *affinely connected.*

Remark 9.2. Note that if we take $\varphi = \bar{\varphi}$, then the transformation law (2.3) is satisfied identically on U. Moreover, a long but straightforward computation shows that this transformation law has the following transitive property: If $\{f^i_{jk}\}$, $\{\bar{f}^i_{jk}\}$ and $\{\tilde{f}^i_{jk}\}$ are smooth functions on the coordinate neighbourhoods U, $\bar{U}$ and $\tilde{U}$ with corresponding charts φ, $\bar{\varphi}$ and $\tilde{\varphi}$ satisfying the two equations

$$\bar{f}^i_{jk} = \frac{\partial \bar{x}^i}{\partial x^p} \cdot \frac{\partial x^q}{\partial \bar{x}^j} \cdot \frac{\partial x^r}{\partial \bar{x}^k} \cdot f^p_{qr} + \frac{\partial \bar{x}^i}{\partial x^r} \cdot \frac{\partial^2 x^r}{\partial \bar{x}^j \partial \bar{x}^k}$$
$$\tilde{f}^i_{jk} = \frac{\partial \tilde{x}^i}{\partial x^p} \cdot \frac{\partial x^q}{\partial \tilde{x}^j} \cdot \frac{\partial x^r}{\partial \tilde{x}^k} \cdot f^p_{qr} + \frac{\partial \tilde{x}^i}{\partial x^r} \cdot \frac{\partial^2 x^r}{\partial \tilde{x}^j \partial \tilde{x}^k}$$

on $\bar{U} \cap U$ and $\tilde{U} \cap U$ respectively, then we have that

$$\tilde{f}^i_{jk} = \frac{\partial \tilde{x}^i}{\partial \bar{x}^p} \cdot \frac{\partial \bar{x}^q}{\partial \tilde{x}^j} \cdot \frac{\partial \bar{x}^r}{\partial \tilde{x}^k} \cdot \bar{f}^p_{qr} + \frac{\partial \tilde{x}^i}{\partial \bar{x}^r} \cdot \frac{\partial^2 \bar{x}^r}{\partial \tilde{x}^j \partial \tilde{x}^k}$$

holds on $U \cap \tilde{U} \cap \bar{U}$. In particular, we may construct a connection on a coordinate neighbourhood U by arbitrarily defining a collection of smooth functions $\{L^i_{jk}\}$ as its coefficients with respect to $\varphi|_U$ (e.g. $L^i_{jk} \equiv 0$), then defining its coefficients $\{\bar{L}^i_{jk}\}$ in any other coordinate system $\bar{\varphi}|_{\bar{U}}$ by (2.3). The transitivity property above then guarantees that any two coefficients of this connection are related by the connection transformation law so that we do indeed have an affine connection on the open submanifold U. △

In light of the preceding remark, we see that affine connections locally exist. Since we have partitions of unity at our disposal, we may also establish global existence.

Lemma 9.3. Connections always exist.

Proof. Write $\{\varphi_\alpha|_{U_\alpha}\}_{\alpha \in A}$ for the charts of our atlas. For each $\alpha_0 \in A$, set

$${}_{\alpha_0}\overset{\alpha}{L}{}^i_{jk} = \begin{cases} 0, & \alpha = \alpha_0 \\ \frac{\partial \overset{\alpha}{x}{}^i}{\partial \overset{\alpha_0}{x}{}^r} \cdot \frac{\partial^2 \overset{\alpha_0}{x}{}^r}{\partial \overset{\alpha}{x}{}^j \partial \overset{\alpha}{x}{}^k}, & \text{otherwise} \end{cases}$$

with domain of definition $U_\alpha \cap U_{\alpha_0}$, and extend these functions to the rest of M by setting them equal to 0 elsewhere. Therefore, we have that for each α_0, the collection

$\{({}_{\alpha_0}\overset{\alpha}{L}{}^i_{jk}\big|_{U_\alpha\cap U_{\alpha_0}})^n_{i,j,k=1}\}_{\alpha\in A}$ such that (2.3) holds on $U_\alpha \cap U_\beta \cap U_{\alpha_0}$ for any $\alpha, \beta \in A$. Now let $\{\phi_l\}_{l\in I}$ be a partition of unity as in Lemma 1.3 and let $\lambda(l) \in A$ be such that $\operatorname{supp}\phi_l \subset U_{\lambda(l)}$. We now define for each coordinate system φ_α the quantities

$$\overset{\alpha}{L}{}^i_{jk} := \sum_{l\in I} \phi_l \cdot {}_{\lambda(l)}\overset{\alpha}{L}{}^i_{jk} : U_\alpha \to \mathbb{R}. \tag{2.4}$$

Now, on $U_\alpha \cap U_\beta \cap U_{\lambda(l)}$, we have from Remark 9.2 that

$${}_{\lambda(l)}\overset{\alpha}{L}{}^i_{jk} = \frac{\partial \overset{\alpha}{x}{}^i}{\partial \overset{\beta}{x}{}^p} \cdot \frac{\partial \overset{\beta}{x}{}^q}{\partial \overset{\alpha}{x}{}^j} \cdot \frac{\partial \overset{\beta}{x}{}^r}{\partial \overset{\alpha}{x}{}^k} \cdot {}_{\lambda(l)}\overset{\beta}{L}{}^p_{qr} + \frac{\partial \overset{\alpha}{x}{}^i}{\partial \overset{\beta}{x}{}^r} \cdot \frac{\partial^2 \overset{\beta}{x}{}^r}{\partial \overset{\alpha}{x}{}^j \partial \overset{\alpha}{x}{}^k}.$$

Multiplying by ϕ_l, summing over l and noting that $U_\alpha \cap U_\beta = \bigcup_l U_\alpha \cap U_\beta \cap U_{\lambda(l)}$, we see that the quantities defined by (2.4) are indeed the coefficients of a connection. □

We henceforth suppose M is equipped with an affine connection L. By the transformation law (2.2), given a vector field $X \in \Gamma^1_0 M$ and defining the quantity

$$X^i{}_{|j} := \frac{\partial X^i}{\partial x^j} + X^p L^i_{pj}$$

in each coordinate system $\varphi|_U$, we find that for any two coordinate systems φ and $\bar{\varphi}$,

$$X^i{}_{|j}\partial_i \otimes \mathrm{d}x^j = \bar{X}^i{}_{|j}\bar{\partial}_i \otimes \mathrm{d}\bar{x}^j$$

so that we obtain a well-defined $(1,1)$-tensor field

$$\nabla X := X^i{}_{|j}\partial_i \otimes \mathrm{d}x^j. \tag{2.5}$$

Lemma 9.4. The mapping

$$\nabla : \Gamma^1_0 M \to \Gamma^1_1 M$$

given by (2.5) is $\mathbb{R}$-linear and satisfies the following Leibniz rule: If $\varphi \in C^\infty(M)$ and $X \in \Gamma^1_0 M$, then

$$\nabla(\varphi X) = \varphi \nabla X + X \otimes \mathrm{d}\varphi.$$

Conversely, if $\nabla : \Gamma^1_0 M \to \Gamma^1_1 M$ is any $\mathbb{R}$-linear mapping with the above Leibniz property, then there exists a unique affine connection L with respect to which ∇ may be written in the form (2.5).

Proof. ($\Rightarrow$) The $\mathbb{R}$-linearity is clear. As for the Leibniz rule, we compute that

$$(\varphi X)^i{}_{|j} = \varphi \cdot X^i{}_{|j} + \underbrace{\partial_j\varphi \cdot X^i}_{=(X\otimes \mathrm{d}\varphi)^i{}_j}.$$

($\Leftarrow$) We first establish that ∇ is local in the following sense: If $X, Y \in \Gamma^1_0 M$ are such that $X \equiv Y$ on $U \subset M$ open, then $\nabla X \equiv \nabla Y$ on U; this will allow us to extend ∇ to an $\mathbb{R}$-linear mapping $\Gamma^1_0 M\big|_U \to \Gamma^1_1 M\big|_U$ satisfying the above Leibniz rule for $\varphi \in C^\infty(U)$, since for any $Z \in \Gamma^1_0 M\big|_U$ and $p \in U$, we can find a vector field $\widetilde{Z} \in \Gamma^1_0 M$ equal to Z on a neighbourhood $N(p) \Subset U$ of p and unambiguously define

$$\nabla Z|_{N(p)} := \nabla \widetilde{Z}\Big|_{N(p)}.$$

Since ∇ is a local operator, this is independent of the chosen extension $\widetilde{Z}$ of Z and p is arbitrary, we obtain the desired extension of ∇, whence the Leibniz rule may be established by considering φZ in place of Z. In consequence of this, we may write $X = X^i \partial_i$ and compute that

$$(\nabla X) = \nabla(X^i \partial_i) = X^i \nabla \partial_i + \partial_i \otimes \mathrm{d}X^i = (\partial_j X^i + X^p (\nabla \partial_p, \mathrm{d}x^i \otimes \partial_j))\partial_i \otimes \mathrm{d}x^j.$$

Setting $L^i_{jk} = (\nabla \partial_j, \mathrm{d}x^i \otimes \partial_k)$ in each coordinate system φ, we obtain the coefficients of a connection. The local property of ∇ is now established as follows: For each $V \Subset U$ open, let $\chi : M \to [0,1]$ be a smooth function such that $\chi|_V \equiv 1$ and $\chi|_{M \setminus U} \equiv 0$. It is clear that $\chi X \equiv \chi Y$ on M, where $X, Y \in \Gamma^1_0 M$ are such that $X \equiv Y$ on U. Hence, we have that $\nabla(\chi X) = \nabla(\chi Y)$, but by the Leibniz rule,

$$\nabla X|_V \equiv \nabla(\chi X)|_V \equiv \nabla(\chi Y)|_V = \nabla Y|_V .$$

Since $V \Subset U$ was arbitrary and may be chosen to be a (compact) neighbourhood of any $x \in U$, we have that $\nabla X|_U \equiv \nabla Y|_U$. □

Definition 9.5. A *covariant differential* (or *covariant derivative operator*) on the bundle $T^r_s M$ is an $\mathbb{R}$-linear mapping $\nabla : \Gamma^r_s M \to \Gamma^r_{s+1} M$ satisfying the following Leibniz rule: If $\varphi \in C^\infty(M)$ and $T \in \Gamma^r_s M$, then $\nabla(\varphi T) = \varphi \nabla T + T \otimes \mathrm{d}\varphi$. Given $T \in \Gamma^r_s M$ and $v \in T_x M$, the *covariant derivative of T in direction v* is given by

$$\nabla_v T := \nabla T \lrcorner v \in (T^r_s M)_x .$$

Lemma 9.4 shows that there is a one-to-one correspondence between affine connections and covariant derivative operators on TM, whence covariant derivative operators are also referred to as *connections*. Affine connections naturally give rise to covariant derivative operators on all tensor bundles $T^r_s M$. We illustrate this in the case of T^*M and defer the case of arbitrary $T^r_s M$ to the next section.

Lemma 9.6. There exists a unique covariant derivative operator $\nabla^* : \Gamma^0_1 M \to \Gamma^1_1 M$ such that for all $\omega \in \Gamma^1_0 M$ and $X, Y \in \Gamma^1_0 M$,

$$\begin{aligned} &\partial_Y(\omega, X) = (\nabla^*_Y \omega, X) + (\omega, \nabla_Y X) \\ &\Leftrightarrow \partial_Y(\omega, X) = (\nabla^* \omega, X \otimes Y) + (\nabla X, \omega \otimes Y). \end{aligned} \tag{2.6}$$

Proof. Consider the mapping

$$(\Gamma^1_0 M)^2 \ni (X, Y) \mapsto \partial_Y(\omega, X) - (\nabla X, Y \otimes \omega) \in C^\infty(M)$$

for fixed $\omega \in \Gamma^0_1 M$. It is clear that this mapping is $C^\infty(M)$-linear in Y. On the other hand, for $f \in C^\infty(M)$,

$$\begin{aligned} &\partial_Y(\omega, fX) - (\nabla(fX), Y \otimes \omega) \\ &= \partial_Y f \cdot (\omega, X) + f \cdot \partial_Y(\omega, X) - (X \otimes \mathrm{d}f, Y \otimes \omega) - f(\nabla X, Y \otimes \omega) \\ &= \partial_Y f \cdot (\omega, X) + f \cdot \partial_Y(\omega, X) - (\omega, X) \cdot \partial_Y f - f(\nabla X, Y \otimes \omega) \\ &= f \cdot ((\omega, X) - (\nabla X, Y \otimes \omega)) . \end{aligned}$$

Therefore, it is also $C^\infty(M)$-linear in X. By Lemma 8.20, there exists a unique tensor field $\nabla^* \omega \in \Gamma^1_1 M$ such that for all $X, Y \in \Gamma^1_0 M$,

$$(\nabla^* \omega, X \otimes Y) = \partial_Y(\omega, X) - (\nabla X, Y \otimes \omega).$$

Replacing ω with $f \cdot \omega$, we see that

$$\begin{aligned}(\nabla^*(f\omega), X \otimes Y) &= \partial_Y(f \cdot \omega, X) - (\nabla X, Y \otimes \omega) \\ &= \partial_Y f \cdot (\omega, X) + f \cdot (\nabla^* \omega, X \otimes Y) \\ &= (\omega \otimes \mathrm{d}f + f \nabla^* \omega, X \otimes Y)\end{aligned}$$

so that $\nabla^*(f\omega) = \omega \otimes \mathrm{d}f + f\nabla^*\omega$. $\mathbb{R}$-linearity may be similarly shown. □

Remark 9.7. We shall henceforth denote ∇^* by ∇. Taking $X = \partial_i$ and $Y = \partial_j$ in the above definition, we see that $\nabla\omega$ is given in local coordinates by

$$\omega_{i|j} := (\nabla\omega)_{ij} = \frac{\partial \omega_i}{\partial x^j} - \omega_r L^r_{ij}.$$

△

Example 9.8 (parallelisable manifolds). M is said to be *parallelisable* if there exist n vector fields $X_1, \ldots, X_n$ such that for all $x \in M$, $T_xM = \operatorname{span}\{X_1(x), \ldots, X_n(x)\}$. On such a manifold, we define a covariant differential ∇ such that

$$\nabla\left(\sum_{i=1}^n a^i X_i\right) := \sum_{i=1}^n X_i \otimes \mathrm{d}a^i.$$

This may readily be seen to be a covariant differential. The connection arising from ∇ is referred to as the canonical connection associated with the above parallelisation.

Exercise 9.9. If G is a Lie group, then it is parallelisable.

Besides giving rise to a covariant derivative operator on TM, a connection naturally gives rise to differential operators acting on vector fields along curves.

Lemma 9.10. Let $\gamma : I \longrightarrow M$ be a smooth curve. There exists a unique $\mathbb{R}$-linear mapping $\frac{\delta}{\delta t} : {}_\gamma\Gamma^1_0 M \longrightarrow {}_\gamma\Gamma^1_0 M$ such that the following properties hold:

- Whenever $\varphi : I \longrightarrow \mathbb{R}$ is differentiable and $X \in {}_\gamma\Gamma^1_0 M$, we have the following Leibniz-type rule:
$$\frac{\delta}{\delta t}(\varphi \cdot X) = \varphi' \cdot X + \varphi \cdot \frac{\delta X}{\delta t}.$$
- Whenever $X \in \Gamma^1_0 M$, $\frac{\delta}{\delta t}(X \circ \gamma) = \nabla_{\dot{\gamma}} X$.

Proof. It may be shown using similar methods to Lemma 9.4 that $\frac{\delta}{\delta t}$ is *local* in the sense that whenever $X, Y \in {}_\gamma\Gamma^1_0 M$ are such that $X \equiv Y$ on $J \subset I$ open, then $\frac{\delta X}{\delta t} \equiv \frac{\delta Y}{\delta t}$ on J. Therefore, $\frac{\delta}{\delta t}$ extends to an $\mathbb{R}$-linear mapping ${}_{\gamma|_J}\Gamma^1_0 M \longrightarrow {}_{\gamma|_J}\Gamma^1_0 M$ with the above properties with respect to the curve $\gamma|_J$ for any $J \subset I$ open. By the above properties, since we may write $X(t) = X^i(t)\, \partial_i|_{\gamma(t)}$ for $t \in \gamma^{-1}(U)$, we find that

$$\frac{\delta X}{\delta t} = \frac{\mathrm{d}X^i}{\mathrm{d}t}\partial_i \circ \gamma + X^i \frac{\delta}{\delta t}(\partial_i \circ \gamma) = \left(\frac{\mathrm{d}X^i}{\mathrm{d}t} + X^p \frac{\mathrm{d}\gamma^q}{\mathrm{d}t} L^i_{pq} \circ \gamma\right) \partial_i \circ \gamma \tag{2.7}$$

on $\gamma^{-1}(U)$. This establishes uniqueness. Existence follows from verifying that the right-hand expression does not depend on the choice of local coordinate system, which is a straightforward computation. □

Definition 9.11. The operator $\frac{\delta}{\delta t} : {}_\gamma\Gamma^1_0 M \to {}_\gamma\Gamma^1_0 M$ associated with a smooth curve γ and the connection L is called the *absolute derivative* operator along γ, and for $X \in {}_\gamma\Gamma^1_0 M$ we call $\frac{\delta X}{\delta t}$ the *absolute derivative* of X.

Remark 9.12. Given a smooth one-parameter family of curves $\{\gamma(s,\cdot) : I \to M\}_{s\in J}$, which may be viewed as a curve $s \mapsto \gamma(s,t)$ or $t \mapsto \gamma(s,t)$, we obtain two absolute derivatives $\frac{\delta}{\delta t}$ and $\frac{\delta}{\delta s}$ acting on elements of ${}_\gamma\Gamma^1_0 M$; their local expressions are given by (2.7) with the derivatives appropriately replaced with partial derivatives $\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial s}$. We likewise obtain n absolute derivative operators along a mapping $\Omega \subset \mathbb{R}^n \to M$. △

Remark 9.13. Arguing as in Lemma , we obtain a unique operator $\frac{\delta}{\delta t} : {}_\gamma\Gamma^0_1 M \to {}_\gamma\Gamma^0_1 M$ such that for all $\omega \in {}_\gamma\Gamma^0_1 M$ and $X \in {}_\gamma\Gamma^1_0 M$,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\omega, X) = (\frac{\delta\omega}{\delta t}, X) + (\omega, \frac{\delta X}{\delta t}).$$

We also refer to this as the *absolute derivative operator* along γ. It is given locally by

$$\frac{\delta\omega_i}{\delta t} = \frac{\mathrm{d}\omega_i}{\mathrm{d}t} - \omega_p \frac{\mathrm{d}\gamma^q}{\mathrm{d}t} L^p_{iq} \circ \gamma.$$

△

Note that in any coordinate system $\varphi|_U$, we may decompose the coefficients of L as follows:

$$L^i_{jk} = \Gamma^i_{jk} + \Omega^i{}_{jk},$$

where $\Gamma^i_{jk} = \frac{1}{2}(L^i_{jk} + L^i_{kj})$ and $\Omega^i{}_{jk} = \frac{1}{2}(L^i_{jk} - L^i_{kj})$ are the *symmetric* and *antisymmetric* parts of L^i_{jk}. Note that if we form the same quantities in another coordinate system $\bar\varphi|_{\bar U}$, the transformation law (2.3) yields the transformation laws

$$\begin{aligned}\bar\Gamma^i_{jk} &= \frac{\partial\bar x^i}{\partial x^p} \cdot \frac{\partial x^q}{\partial \bar x^j} \cdot \frac{\partial x^r}{\partial \bar x^k} \cdot \Gamma^p_{qr} + \frac{\partial\bar x^i}{\partial x^r} \cdot \frac{\partial^2 x^r}{\partial\bar x^j \partial\bar x^k} \\ \bar\Omega^i{}_{jk} &= \frac{\partial\bar x^i}{\partial x^p} \cdot \frac{\partial x^q}{\partial \bar x^j} \cdot \frac{\partial x^r}{\partial \bar x^k} \cdot \Omega^p{}_{qr}\end{aligned}$$

on $U \cap \bar U$. We immediately see that the $\{\Gamma^i_{jk}\}$ define a connection Γ on M, whereas the $\{\Omega^i{}_{jk}\}$ define a $(1,2)$-tensor field on M, called the *torsion tensor*, which we write as

$$\Omega = \Omega^i{}_{jk} \partial_i \otimes \mathrm{d}x^j \otimes \mathrm{d}x^k.$$

Exercise 9.14. Show that for $X, Y \in \Gamma^1_0 M$,

$$\Omega \lrcorner (X \otimes Y) = \frac{1}{2}([X,Y] - \nabla_X Y - \nabla_Y X). \tag{2.8}$$

Remark 9.15. One instance where the torsion tensor arises is the following: Let $f \in C^\infty(M)$ and consider the *Hessian* of f:

$$(\nabla \mathrm{d}f)_{ij} = \frac{\partial^2 f}{\partial x^j \partial x^i} - \partial_k f \cdot L^k_{ij}.$$

In Euclidean space (where $L^k_{ij} \equiv 0$), $\nabla \mathrm{d}f$ is symmetric. More generally however,

$$(\nabla \mathrm{d}f)_{ij} - (\nabla \mathrm{d}f)_{ji} = -2\partial_k f \cdot \Omega^k{}_{ij} \neq 0.$$

△

We shall call a connection L *symmetric* (or *torsion-free*) if in every coordinate system $L^i_{jk} = \Gamma^i_{jk}$, which is equivalent to the condition $\Omega \equiv 0$. In later considerations we shall only be interested in symmetric connections.

We note the following characterising property of symmetric connections.

Lemma 9.16. A connection L if symmetric if and only if for each $p \in M$, there exists a coordinate system $\varphi|_U$ with $p \in U$ such that $L^i_{jk}(p) = 0$. Such a coordinate system is called a *normal* coordinate system.

Proof. ($\Leftarrow$) For all $p \in M$, we have that in a normal coordinate system, $L^i_{jk}(p) = 0 \Rightarrow \Omega^i{}_{jk}(p) = 0 \Rightarrow \Omega(p) = 0$ so that $\Omega \equiv 0$.

($\Rightarrow$) Suppose L is symmetric and fix a coordinate system $\varphi|_U$ about p. Consider the n functions

$$\bar{x}^i = (x^i - x^i(p)) + \frac{1}{2} L^i_{rs}(p) \cdot (x^r - x^r(p)) \cdot (x^s - x^s(p)).$$

It is straightforward to see that $\bar{x}^i(p) = 0$, $\frac{\partial \bar{x}^i}{\partial x^j}(p) = \delta^i_j$ and $\frac{\partial^2 \bar{x}^i}{\partial x^j x^k}(p) = L^i_{jk}(p)$. Due to the second condition, the inverse function theorem implies that there is a neighbourhood $\bar{U}$ of p such that

$$\bar{\varphi} := (\bar{x}^1, \dots, \bar{x}^n)\big|_{\bar{U}}$$

is a coordinate system. The transformation law (2.3) now implies that

$$L^i_{jk}(p) = \bar{L}^i_{jk}(p) + L^i_{jk}(p) \Leftrightarrow \bar{L}^i_{jk}(p) = 0. \qquad \square$$

10. Parallelism. As in the preceding section, we suppose M is an n-dimensional affinely connected manifold with connection L. On parallelisable manifolds such as $\mathbb{R}^n$, we may naturally displace or translate vectors in a parallel manner: If M is parallelisable, $\{\varepsilon_i\}$ is a global frame for TM and $v = v^i \varepsilon_i(p) \in T_pM$, then for any $q \in M$, we obtain the *parallel translate* of v to q defined by

$$V(q) = v^i \varepsilon_i(q).$$

This vector field has the defining property that $\nabla V \equiv 0$ with respect to the connection of Example 9.8. On an arbitrary affinely connected manifold, the condition $\nabla V \equiv 0$ may be locally formulated as

$$\partial_j V^i + V^k \cdot L^i_{kj} = 0,$$

which is a system of (linear) PDE. Whenever $\nabla V \equiv 0$ is satisfied, we call V a *parallel vector field.* We will see later on that the existence of parallel vector fields is connected to the notion of curvature. As a first step, we consider vector fields along curves.

Definition 10.1. Let $\gamma : I \to M$ be a smooth curve on an interval I and $V \in {}_\gamma\Gamma^1_0 M$. V is said to be *parallel along* γ (with respect to L) if

$$\frac{\delta V}{\delta t} \equiv 0. \tag{2.9}$$

Note that if we write (2.9) out in local coordinates, writing $V(t) = V^i(t)\, \partial_i|_{\gamma(t)}$, we obtain the system

$$\frac{\mathrm{d}V^i}{\mathrm{d}t} + V^j \dot{\gamma}^k L^i_{jk} \circ \gamma = 0. \tag{2.10}$$

This system is linear, but not autonomous. Nevertheless, it may be treated using the same ODE theory used to prove Theorem 6.4. If we fix $t_0 \in I$, choose our coordinate neighbourhood U above such that $\gamma(t_0) \in U$ and write $J \subset I$ for the largest subinterval of I containing t_0 such that $\gamma^{-1}(U) \supset J$, then given the initial condition $V^i(t_0) = v^i$, there exists a solution to (2.10) satisfying this initial condition and defined on *all of* J,[2] and if $W^i\big|_{J_0}$ is another solution to (2.10) with initial condition $W^i(t_0) = v^i$, then $J_0 \subset J$ and $W^i \equiv V^i\big|_{J_0}$. We may argue as in Theorem 6.5 to conclude the following existence and uniqueness result.

Theorem 10.2. Let $\gamma : I \to M$ be a smooth curve, $t_0 \in I$ and $v_0 \in T_{\gamma(t_0)}M$. There exists a unique parallel vector field $V \in {}_{\gamma}\Gamma^1_0 M$ such that $V(t_0) = v_0$.

Given a smooth curve $\gamma : I \to M$, $t_0 \in I$ and $v_0 \in T_{\gamma(t_0)}M$, we shall write

$$I \ni t \mapsto P^t_{\gamma,t_0}(v_0) \in T_{\gamma(t)}M$$

for the unique parallel vector field along γ passing through v_0 at $t = t_0$. This mapping is referred to as the *parallelism*, *parallel transport* or *parallel displacement* associated to L. We summarise its important properties.

Proposition 10.3. The parallelism associated to L has the following properties:

- *Parametrisation invariance*: If $\gamma : I \to M$ is a smooth curve and $\chi : J \to I$ is a smooth bijection, then for all $s, s_0 \in J$ and $v_0 \in T_{\gamma(\chi(s_0))}M$,

$$P^{\chi(s)}_{\gamma,\chi(s_0)}(v_0) = P^s_{\gamma\circ\chi,s_0}(v_0).$$

- *Group property*: If $\gamma : I \to M$ is a smooth curve and $t_0, t_1, t_2 \in I$ fixed, then for any $v_0 \in T_{\gamma(t_0)}M$,

$$P^{t_2}_{\gamma,t_0}(v_0) = P^{t_2}_{\gamma,t_1}(P^{t_1}_{\gamma,t_0}(v_0).$$

- *Isomorphy*: If $\gamma : I \to M$ is a smooth curve and $t_0, t_1 \in I$ fixed, then $P^{t_1}_{\gamma,t_0} : T_{\gamma(t_0)}M \to T_{\gamma(t_1)}M$ is a vector space isomorphism with inverse given by $P^{t_0}_{\gamma,t_1}$.
- *Smooth dependence*: If $\{\gamma(s,\cdot) : I \to M\}_{s\in J}$ is a smooth family of smooth curves and $v \in {}_{\gamma(\cdot,t_0)}\Gamma^1_0 M$ for some $t_0 \in I$, then the mapping

$$(s,t) \mapsto P^t_{\gamma(s,\cdot),t_0}(v(s)) \tag{2.11}$$

is smooth.

Proof. Smooth dependence may be established as follows: For fixed $\sigma \in J$, let

$$I(\sigma) := \{\tau \in I : (2.11) \text{ is smooth on a neighbourhood of } (\sigma,\tau)\}.$$

Using ODE theory, we see that $t_0 \in I(\sigma)$ so that $I(\sigma) \neq \varnothing$. Moreover, by definition, $I(\sigma)$ is open, since $\tau \in I(\sigma) \Rightarrow$ (2.11) smooth on $(]\sigma-\varepsilon,\sigma+\varepsilon[\times]\tau-\varepsilon,\tau+\varepsilon[) \cap (J\times I)$ for some $\varepsilon > 0$, it is clear that for any $\tau_0 \in \left]\tau-\frac{\varepsilon}{2},\tau+\frac{\varepsilon}{2}\right[\cap I$, (2.11) is smooth on $\left(]\sigma-\varepsilon,\sigma+\varepsilon[\times\left]\tau_0-\frac{\varepsilon}{2},\tau_0+\frac{\varepsilon}{2}\right[\right) \cap (J\times I)$ so that $\left]\tau-\frac{\varepsilon}{2},\tau+\frac{\varepsilon}{2}\right[\cap I \subset I(\sigma)$. To show closedness, let $\{\tau_n\}_{n=1}^{\infty} \subset I(\sigma)$ be such that $\tau_n \xrightarrow{n\to\infty} \tau \in I$. Let U be a coordinate neighbourhood containing $\gamma(\sigma,\tau)$. For sufficiently large $N \in \mathbb{N}$ and small $\varepsilon > 0$, we have that $\gamma(\sigma,]\tau_N-\varepsilon,\tau+\varepsilon[\cap I) \subset U$. Moreover, for $\varepsilon_1 > 0$ small, (2.11) is smooth on $(]\sigma-\varepsilon_1,\sigma+\varepsilon_1[\times]\tau_N-\varepsilon_1,\tau_N+\varepsilon_1[) \cap (J\times I)$. By the group property, $P^t_{\gamma(s,\cdot),t_0}(v(s)) =$

[2]This is a consequence of linearity.

$P^t_{\gamma(s,\cdot),\tau_N}(v_1(s))$ with $v_1(s) = P^{\tau_N}_{\gamma(s,\cdot),t_0}(v(s))$ and $(s,t) \in (]\sigma - \varepsilon_1, \sigma + \varepsilon_1[\times]\tau_N - \varepsilon, \tau + \varepsilon[) \cap (J \times I)$. Since v_1 is smooth for $s \in]\sigma - \varepsilon_1, \sigma + \varepsilon_1[\cap J$ by hypothesis, appealing again to ODE theory, we see that (2.11) is smooth on $(]\sigma - \varepsilon_1, \sigma + \varepsilon_1[\times]\tau_N - \varepsilon, \tau + \varepsilon[) \cap (J \times I)$ so that $\tau \in I(\sigma)$. □

Remark 10.4 (parallelism matrices). For a fixed curve $\gamma : I \to M$ with $t_0, t_1 \in I$, we may describe P^t_{γ,t_0} as follows: If $\varphi|_U$ is a coordinate system with $\gamma(t_0) \in U$ and $\tilde{\varphi}|_{\tilde{U}}$ one with $\gamma(t_1) \in \tilde{U}$, then for any $v = v^j \, \partial_j\big|_{\gamma(t_0)}$ and t close to t_1,

$$P^t_{\gamma,t_0}(v) = a^i_j(t,t_0) v^j \, \tilde{\partial}_j\Big|_{\gamma(t)}, \tag{2.12}$$

where $t \mapsto a^i_j(t,t_0)$ is smooth. Moreover, by isomorphy, the inverse matrix $(\alpha^i_j(t,t_0)) := (a^i_j(t,t_0))^{-1}$ exists. If in particular, t_1 is sufficiently close to t_0, we may choose $\tilde{\varphi} = \varphi$ so that $a^i_j(t_0,t_0) = \delta^i_j$. △

Remark 10.5. In light of the fact that P^t_{γ,t_0} is an isomorphism, we naturally obtain a notion of parallelism for elements of T^*M by considering the mapping

$${}^t(P^t_{\gamma,t_0})^{-1} : T^*_{\gamma(t_0)}M \to T^*_{\gamma(t)}M.$$

By Remark 10.4, we may express the matrix of this mapping as

$${}^t(P^t_{\gamma,t_0})^{-1} \, \mathrm{d}x^j\big|_{\gamma(t_0)} = \alpha^j_i(t,t_0) \, \mathrm{d}x^i\big|_{\gamma(t)}$$

so that for $\nu \in T^*_{\gamma(t_0)}M$, $t \mapsto {}^t(P^t_{\gamma,t_0})^{-1}\nu$ is a smooth 1-form along γ. More generally, we obtain a parallelism for arbitrary *tensors* by defining the mapping

$$\begin{aligned} P^t_{\gamma,t_0} : (T^r_s M)_{\gamma(t_0)} &\to (T^r_s M)_{\gamma(t)} \\ T &\mapsto (\underbrace{P^t_{\gamma,t_0} \otimes \cdots \otimes P^t_{\gamma,t_0}}_{r \text{ times}} \otimes \underbrace{{}^t(P^t_{\gamma,t_0})^{-1} \otimes \cdots \otimes {}^t(P^t_{\gamma,t_0})^{-1}}_{s \text{ times}})T. \end{aligned}$$

Using the local coordinate representation of Remark 10.4, we see that $t \mapsto P^t_{\gamma,t_0}T$ is smooth as well. Note that this induced parallelism is compatible with tensor and interior products in the sense that it distributes over them. For instance, if $S \in (T^{r_1}_{s_1}M)_{\gamma(t_0)}$ and $T \in (T^{r_2}_{s_2}M)_{\gamma(t_0)}$, then

$$P^t_{\gamma,t_0}(S \otimes T) = P^t_{\gamma,t_0}S \otimes P^t_{\gamma,t_0}T.$$

△

Remark 10.6 (Parallel frames). Let $\{b_i\}_{i=1}^n$ be a basis for $T_{\gamma(t_0)}M$ with dual basis $\{b^i\}_{i=1}^n$ for $T^*_{\gamma(t_0)}M$. Using the parallelism, we set $b_i(t) := P^t_{\gamma,t_0}b_i$ and $b^i(t) := {}^t(P^t_{\gamma,t_0})^{-1}b^i$. It is easy to see that the $\{b_i(t)\}$ span $T_{\gamma(t)}M$ and are dual to the $\{b^i(t)\}$. Moreover, these definitions imply that $b_i \in {}_\gamma\Gamma^1_0 M$ and $b^i \in {}_\gamma\Gamma^0_1 M$. We call these *parallel frames* along γ. △

The parallelism we have obtained depends crucially on our choice of connection L. On the other hand, it turns out we may actually reconstruct L, more specifically the notions of absolute and covariant differentiation, from the parallelism we obtained above.

Lemma 10.7. Let $\gamma : I \to M$ be a smooth curve and $X \in {}_\gamma\Gamma^1_0 M$. Then for any $t_0 \in I$, and parallel frame $\{b_i\}$ along γ with respect to which we may write $X(t) = X^i(t)b_i(t)$,

$$\frac{\delta X}{\delta t}(t_0) = \dot{X}^i(t_0)b_i(t_0) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=t_0} (P^t_{\gamma,t_0})^{-1}(X(\gamma(t))). \tag{2.13}$$

Proof. Exercise. □

Motivated by the characterisation (2.13) of absolute differentiation, we define the *absolute derivative* $\frac{\delta T}{\delta t}$ of a tensor $T \in {}_\gamma\Gamma^r_s M$ as follows: Fix a parallel frame $\{b_i\}$ along γ and write

$$T(t) = T^{i_1\dots i_r}{}_{j_1\dots j_s}(t) b_{i_1}(t) \otimes \cdots \otimes b_{i_r}(t) \otimes b^{j_1}(t) \otimes \cdots \otimes b^{j_s}(t).$$

We now define the absolute derivative of T such that for each $t_0 \in I$,

$$\frac{\delta T}{\delta t}(t_0) = \dot{T}^{i_1\dots i_r}{}_{j_1\dots j_s}(t_0) \cdot b_{i_1}(t_0) \otimes \cdots \otimes b_{i_r}(t_0) \otimes b^{j_1}(t_0) \otimes \cdots \otimes b^{j_s}(t_0). \tag{2.14}$$

It is clear that $\frac{\delta T}{\delta t} \in {}_\gamma\Gamma^r_s M$, and by the same argument in Lemma 10.7, we have the equality

$$\frac{\delta T}{\delta t}(t_0) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} (P^t_{\gamma,t_0})^{-1} T(t),$$

which shows that the definition (2.14) is independent of the chosen parallel frame. The following proposition gives us a local expression and a covariant differential naturally associated with this absolute derivative.

Proposition 10.8. Let $T \in {}_\gamma\Gamma^r_s M$. The absolute derivative of T is given in any local coordinate system by

$$\left(\frac{\delta T}{\delta t}\right)^{i_1\dots i_r}{}_{j_1\dots j_s} = \dot{T}^{i_1\dots i_r}{}_{j_1\dots j_s} + \sum_{l=1}^{r} T^{i_1\dots i_{l-1} c i_{l+1}\dots i_r}{}_{j_1\dots j_s} \cdot L^{i_l}_{cd} \circ \gamma \cdot \dot{\gamma}^d \tag{2.15}$$

$$- \sum_{l=1}^{s} T^{i_1\dots i_r}{}_{j_1\dots j_{l-1} c j_{l+1}\dots j_s} \cdot L^c_{j_l d} \circ \gamma \cdot \dot{\gamma}^d. \tag{2.16}$$

Moreover, there exists a unique covariant differential $\nabla : \Gamma^r_s M \longrightarrow \Gamma^r_{s+1} M$ such that for any $T \in \Gamma^r_s M$, $v \in T_p M$ and curve $\gamma : I \ni t_0 \longrightarrow M$ with $\gamma(t_0) = p$ and $\dot{\gamma}(t_0) = v$,

$$\nabla_v T = \frac{\delta(T \circ \gamma)}{\delta t}(t_0).$$

Proof. By 10.4, for t sufficiently close to t_0, we may write the parallelism matrices associated with γ in a single coordinate system so that

$$(P^t_{\gamma,t_0})^{-1} (T^{i_1\dots i_r}{}_{j_1\dots j_s}(t)\, \partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s}|_{\gamma(t)})$$
$$= T^{i_1\dots i_r}{}_{j_1\dots j_s}(t) \cdot \left(\alpha^{k_1}_{i_1} \dots \alpha^{k_r}_{i_r} \cdot a^{j_1}_{l_1} \dots a^{j_s}_{l_s}\right)(t, t_0) \cdot \partial_{k_1} \otimes \cdots \otimes \partial_{k_r} \otimes \mathrm{d}x^{l_1} \otimes \cdots \otimes \mathrm{d}x^{l_s}|_{\gamma(t_0)}.$$

Differentiating this expression at $t = t_0$ and using the fact that $a^i_j(t_0, t_0) = \alpha^i_j(t_0, t_0) = \delta^i_j$, $\dot{a}^i_j(t_0, t_0) = -L^i_{jk}(\gamma(t_0)) \cdot \dot{\gamma}^k(t_0)$ and $\dot{\alpha}^i_j(t_0, t_0) = -\dot{a}^i_j(t_0, t_0)$, the local expression above follows immediately from the product rule.

Now suppose $T \in \Gamma^r_s M$. Applying (2.16) to $T \circ \gamma$, where $\gamma(t_0) = p$ and $\dot{\gamma}(t_0) = v \in T_p M$, we see that

$$\frac{\delta(T \circ \gamma)}{\delta t}(t_0) = \underbrace{\left(T^{i_1\dots i_r}{}_{j_1\dots j_s | k}(p)\, \partial_{i_1} \otimes \cdots \otimes \partial i_r \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s} \otimes \mathrm{d}x^k|_p\right)}_{=:\nabla T(p)} \llcorner v, \tag{2.17}$$

where

$$T^{i_1\ldots i_r}{}_{j_1\ldots j_s|k} = \partial_k T^{i_1\ldots i_r}{}_{j_1\ldots j_s} + \sum_{l=1}^{r} T^{i_1\ldots i_{l-1}ci_{l+1}\ldots i_r}{}_{j_1\ldots j_s} \cdot L^{i_l}_{ck} - \sum_{l=1}^{s} T^{i_1\ldots i_r}{}_{j_1\ldots j_{l-1}cj_{l+1}\ldots j_s} \cdot L^{c}_{j_l k}.$$

Since the left-hand side of (2.17) is independent of coordinates, we see that $\nabla T(p)$ is well-defined, and from the expression above we see that ∇T is indeed smooth. We leave the verification of $\mathbb{R}$-linearity and the Leibniz rule to the reader as an exercise. □

Remark 10.9. Just as with vector fields along a curve γ, we call a tensor $T \in {}_\gamma\Gamma^r_s M$ *parallel* along γ if $\frac{\delta T}{\delta t} \equiv 0$. In particular, retaining the notation of Remark 10.6, it is straightforward to verify that

$$t \mapsto (b_{i_1} \otimes \cdots \otimes b_{i_r} \otimes b^{j_1} \otimes \cdots \otimes b^{j_s})(t)$$

is parallel along γ, and given any $T_0 \in (T^r_s M)_{\gamma(t_0)}$, $T(t) = P^t_{\gamma,t_0} T_0$ is the unique parallel tensor equal to T_0 at $t = t_0$. △

Since parallel displacement is compatible with tensor and interior products, we obtain some useful Leibniz-type rules for absolute and covariant differentiation.

Proposition 10.10. Let $\gamma : I \longrightarrow M$ be a smooth curve. The following hold:

- If $S \in {}_\gamma\Gamma^{r_1}_{s_1} M$ and $T \in {}_\gamma\Gamma^{r_2}_{s_2} M$, then
$$\frac{\delta}{\delta t}(S \otimes T) = \frac{\delta S}{\delta t} \otimes T + S \otimes \frac{\delta T}{\delta t}.$$
- If $T \in {}_\gamma\Gamma^r_s M$ and $S \in {}_\gamma\Gamma^s_r M$, then
$$\frac{\delta}{\delta t}(T, S) = (\frac{\delta T}{\delta t}, S) + (T, \frac{\delta S}{\delta t})$$
- If $T \in {}_\gamma\Gamma^r_s M$ with $s > 0$ and $v \in {}_\gamma\Gamma^1_0 M$, then
$$\frac{\delta}{\delta t}(v \llcorner T) = \frac{\delta v}{\delta t} \llcorner T + v \llcorner \frac{\delta T}{\delta t};$$
$$\frac{\delta}{\delta t}(T \lrcorner v) = T \lrcorner \frac{\delta v}{\delta t} + \frac{\delta T}{\delta t} \lrcorner v.$$

Proof. For the first claim, note that for $t_0 \in I$, $(P^t_{\gamma,t_0})^{-1} = P^{t_0}_{\gamma,t}$ so that

$$(P^t_{\gamma,t_0})^{-1}(S(t) \otimes T(t)) = (P^t_{\gamma,t_0})^{-1} S(t) \otimes (P^t_{\gamma,t_0})^{-1} T(t).$$

Differentiating both sides with respect to t at $t = t_0$ and using the chain rule, we obtain

$$\begin{aligned}
\frac{\delta}{\delta t}(S \otimes T)(t_0) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} (P^t_{\gamma,t_0})^{-1}(S(t) \otimes T(t)) \\
&= \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} (P^t_{\gamma,t_0})^{-1} S(t) \right) \otimes T(t_0) + S(t_0) \otimes \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=t_0} (P^t_{\gamma,t_0})^{-1} T(t) \right) \\
&= \left(\frac{\delta S}{\delta t} \otimes T + S \otimes \frac{\delta T}{\delta t} \right)(t_0).
\end{aligned}$$

To establish the second claim, we note that $(T(t), S(t)) = ((P^t_{\gamma,t_0})^{-1} T(t), (P^t_{\gamma,t_0})^{-1} S(t))$ and differentiate with respect to t at $t = t_0$ similarly. The third claim follows analogously. □

If in the above proposition the tensors along γ are of the form $T \circ \gamma$ with $T \in \Gamma^r_s M$, then we may use the relation $\frac{\delta}{\delta t}(T \circ \gamma) = \nabla_{\dot\gamma} T$ to deduce the following corollary:

Corollary 10.11. Let $v \in T_p M$. The following hold:

- If $S \in \Gamma^{r_1}_{s_1} M$ and $T \in \Gamma^{r_2}_{s_2} M$, then
$$\nabla_v(S \otimes T) = \nabla_v S \otimes T(p) + S(p) \otimes \nabla_v T.$$
- If $T \in \Gamma^r_s M$ and $S \in \Gamma^s_r M$, then
$$\partial_v(T, S) = (\nabla_v T, S(p)) + (T(p), \nabla_v S)$$
- If $T \in \Gamma^r_s M$ with $s > 0$ and $w \in \Gamma^1_0 M$, then
$$\begin{aligned}\nabla_v(w \llcorner T) &= \nabla_v w \llcorner T(p) + w(p) \llcorner \nabla_v T;\\ \nabla_v(T \lrcorner w) &= T(p) \lrcorner \nabla_v w + \nabla_v T \lrcorner w.\end{aligned}$$

Remark 10.12. Using the compatibility of absolute and covariant differentiation with respect to interior products, it follows that these operations also commute with traces, since for $T \in (T^r_s M)_p$ and $v_1, \dots, v_{r-1} \in T_p M$, $v^1, \dots, v^{s-1} \in T^*_p M$,

$$(\operatorname{tr} T, v_1 \otimes \cdots \otimes v_{r-1} \otimes v^1 \otimes \cdots \otimes v^{s-1}) = (T, \delta \otimes v_1 \otimes \cdots \otimes v_{r-1} \otimes v^1 \otimes \cdots \otimes v^{s-1}). \quad \triangle$$

11. Higher-order derivatives. In light of the preceding section, given a tensor field $T \in \Gamma^r_s M$ with $r + s > 0$, we may form its covariant differential $\nabla T \in \Gamma^r_{s+1} M$. For each $k \in \mathbb{N}$, the *kth covariant differential of* T is defined such that $\nabla^0 T := T$ and $\nabla^k T = \nabla(\nabla^{k-1} T) \in \Gamma^r_{s+k} M$. For a function $f \in C^\infty(M)$, we set $\nabla^2 f := \nabla \mathrm{d} f$ and for $k \geq 3$, $\nabla^k f := \nabla(\nabla^{k-1} f) \in \Gamma^0_k M$. The components of $\nabla^k T$ in local coordinates shall be denoted $T^{i_1 \dots i_r}{}_{j_1 \dots j_s | j_{s+1} \dots j_{s+k}}$.

Just as the last slot of ∇T is considered as a direction of differentiation, hence the notation $\nabla_X T = \nabla T \lrcorner X$, we shall also interpret the last k slots of $\nabla^k T$ as directions of differentiating. Therefore, for vector fields (or contravariant vectors) $X_1, \dots, X_k$, we set

$$\begin{aligned}\nabla_{X_1, \dots, X_k} T &:= \nabla^k T \lrcorner (X_k \otimes \cdots \otimes X_1)\\ &= T^{i_1 \dots i_r}{}_{j_1 \dots j_s | l_1 \dots l_k} X_k^{l_1} \dots X_1^{l_k} \partial_{i_1} \otimes \cdots \otimes \partial_{i_r} \otimes \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_s}.\end{aligned}$$

Note that the order has been reversed due to our conventions. Also, $\nabla_{X_1, \dots, X_k} T$ is *not* equal to $\nabla_{X_1} \nabla_{X_2} \dots \nabla_{X_k} T$, though we may relate these quantities to each other. In the case $k = 2$, writing $X_1 = X$ and $X_2 = Y$, we have

$$\begin{aligned}\nabla_X \nabla_Y T &= \nabla_X(\nabla T \lrcorner Y)\\ &= \nabla_X \nabla T \lrcorner Y + \nabla T \lrcorner \nabla_X Y = (\nabla^2 T \lrcorner X \lrcorner Y) + \nabla_{\nabla_X Y} T = \nabla_{X,Y} T + \nabla_{\nabla_X Y} T.\end{aligned} \tag{2.18}$$

12. Curvature. In Section 9, we saw that a peculiarity of asymmetric connections ($\Omega \not\equiv 0$) was that the Hessian $\nabla^2 f$ of a function $f \in C^\infty(M)$ was not necessarily symmetric, and the obstruction to this was the torsion tensor:

$$(\nabla^2 f, X \otimes Y - Y \otimes X) = -2(\Omega, \mathrm{d} f \otimes X \otimes Y).$$

This defect may be remedied by replacing our connection L with its symmetrisation Γ. We now turn our attention to the question of symmetry of the $\nabla_{X,Y} T$ in X and Y for a tensor T, first considering the concrete case of $(0, 1)$-tensors.

Lemma 12.1. Let $a \in \Gamma^0_1 M$. There exists a tensor $R \in \Gamma^1_3 M$ such that in any local coordinate system,

$$a_{i|jk} - a_{i|kj} = a_r \cdot R^r{}_{ijk} - 2a_{i|s} \cdot \Omega^s{}_{jk}. \tag{2.19}$$

The components of R are given by

$$R^i{}_{jkl} = \frac{\partial L^i_{jl}}{\partial x^k} - \frac{\partial L^i_{jk}}{\partial x^l} + L^r_{jl} L^i_{rk} - L^r_{jk} L^i_{rl}. \tag{2.20}$$

In particular, for contravariant vectors $v, w \in T_p M$, there holds

$$\nabla_{v,w} a - \nabla_{w,v} a = (R^i{}_{jkl} a_i w^k v^l) \mathrm{d}x^j - 2\nabla_{\Omega \lrcorner (w \otimes v)} a. \tag{2.21}$$

Proof. This is a straightforward computation. □

Definition 12.2. The tensor $R \in \Gamma^1_3 M$ with components given by (2.20) is called the *curvature tensor* of L.

Remark 12.3. Before describing $\nabla_{v,w} T - \nabla_{w,v} T$ for an arbitrary tensor T, we note that for $X, Y \in \Gamma^1_0 M$ and $T \in \Gamma^r_s M$, (2.18) implies that

$$\nabla_{X,Y} T - \nabla_{Y,X} T = \nabla_X \nabla_Y T - \nabla_Y \nabla_X T - \nabla_{\nabla_X Y - \nabla_Y X} T;$$

using (2.8) and rearranging, we obtain the equality

$$\nabla_X \nabla_Y T - \nabla_Y \nabla_X T - \nabla_{[X,Y]} T = \nabla_{X,Y} T - \nabla_{Y,X} T + 2\nabla_{\Omega \lrcorner (Y \otimes X)} T. \tag{2.22}$$

Define $R(X,Y)T := \nabla_X \nabla_Y T - \nabla_Y \nabla_X T - \nabla_{[X,Y]} T$. It is straightforward to check that this expression is $C^\infty(M)$-linear in each of its arguments. Using the same reasoning as in Lemma 8.20, we conclude that for any $p \in M$, $v, w \in T_p M$ and $r, s \in \mathbb{N} \cup \{0\}$ with $r + s > 0$, we obtain a well defined linear mapping

$$\begin{aligned} R(v,w) : (T^r_s M)_p &\longrightarrow (T^r_s M)_p \\ t &\longmapsto (R(X,Y)T)(p) \end{aligned}$$

where $X, Y \in \Gamma^1_0 M$ and $T \in \Gamma^r_s M$ are such that $X(p) = v$, $Y(p) = w$ and $T(p) = t$. $R(v,w)$ is called the *curvature endomorphism.* Comparing (2.22) with (2.21), we see that for $(r,s) = (0,1)$, we have

$$R(v,w)a = R^i{}_{jkl}(p) a_i w^k v^l \mathrm{d}x^j = (a \lrcorner R) \lrcorner (w \otimes v).$$

Now suppose X and Y are as above and suppose $Z \in \Gamma^1_0 M$. By definition of the Lie bracket,

$$\partial_X \partial_Y (a, Z) - \partial_Y \partial_X (a, Z) - \partial_{[X,Y]}(a, Z) = 0. \tag{2.23}$$

Using the Leibniz rule, we may explicitly compute that

$$\partial_X \partial_Y (a, Z) = (\nabla_X \nabla_Y a, Z) + (\nabla_X a, \nabla_Y Z) + (a, \nabla_X \nabla_Y Z)$$

so that (2.23) becomes $(R(X,Y)a, Z) + (a, R(X,Y)Z) = 0$, or

$$R(X,Y)Z = -(R(X,Y)\mathrm{d}x^i, Z)\partial_i = -R^i_{jkl} Z^j Y^k X^l \partial_i$$

so that $R(v,w)z = -R^i_{jkl}(p) z^j w^k v^l \ \partial_i|_p$ for $v, w, z \in T_p M$. (2.22) therefore implies that

$$\nabla_{v,w} Z - \nabla_{w,v} Z = -R^i_{jkl}(p) Z^j(p) w^k v^l \ \partial_i|_p - 2\nabla_{\Omega(p) \lrcorner (w \otimes v)} Z.$$

We therefore see that in addition to the torsion tensor, the curvature tensor (or equivalently the curvature endomorphism) is the obstruction to the Hessian of vector and covector fields being symmetric with respect to the directions of differentiation. Similar 'commutation formulæ' may be derived for arbitrary $T \in \Gamma^r_s M$. △

Exercise 12.4. Show the following:

(a) For tensors S and T at p and $v, w \in T_pM$,

$$R(v, w)(S \otimes T) = R(v, w)S \otimes T + S \otimes R(v, w)T.$$

(b) Whenever $T \in (T^r_s M)_p$,

$$R(v, w)T = \left(\sum_{i=1}^{r} R^i(v, w) + \sum_{j=1}^{s} R_j(v, w) \right) T,$$

where $R^i(v, w), R_j(v, w) : (T^r_s M)_p \to (T^r_s M)_p$ are linear mappings such that for $v_1, \ldots, v_r \in T_pM$ and $\omega^1, \ldots, \omega^s \in T^*_pM$,

$$\begin{aligned} &R^i(v, w)(v_1 \otimes \cdots \otimes v_r \otimes \omega^1 \otimes \cdots \otimes \omega^s) \\ &\quad = v_1 \otimes \cdots \otimes v_{i-1} \otimes R(v, w)v_i \otimes v_{i+1} \otimes \cdots \otimes \omega^s \end{aligned}$$

and

$$\begin{aligned} &R_j(v, w)(v_1 \otimes \cdots \otimes v_r \otimes \omega^1 \otimes \cdots \otimes \omega^s) \\ &\quad = v_1 \otimes \cdots \otimes v_r \otimes \omega^1 \otimes \cdots \otimes \omega^{j-1} \otimes R(v, w)\omega^j \otimes \omega^{j+1} \otimes \cdots \otimes \omega^s. \end{aligned}$$

(c) For $T \in (T^r_s M)_p$ and $v, w \in T_pM$,

$$\begin{aligned} (R(v, w)T)^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} = &\sum_{l=1}^{r} T^{i_1 \ldots i_{l-1} k i_{l+1} \ldots i_r}{}_{j_1 \ldots j_s} \cdot R^{i_l}{}_{kab}(p) v^a w^b \\ &- \sum_{l=1}^{s} T^{i_1 \ldots i_r}{}_{j_1 \ldots j_{l-1} k j_{l+1} \ldots j_s} \cdot R^k{}_{j_l ab}(p) v^a w^b. \end{aligned}$$

(d) If $T \in \Gamma^r_s M$, the *generalised Ricci identity* holds:

$$\begin{aligned} T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s | ab} - T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s | ba} = &-\sum_{l=1}^{r} T^{i_1 \ldots i_{l-1} k i_{l+1} \ldots i_r}{}_{j_1 \ldots j_s} \cdot R^{i_l}{}_{kab} \\ &+ \sum_{l=1}^{s} T^{i_1 \ldots i_r}{}_{j_1 \ldots j_{l-1} k j_{l+1} \ldots j_s} \cdot R^k{}_{j_l ab} \end{aligned}$$

We summarise some symmetry properties of the curvature tensor.

Proposition 12.5. The curvature tensor R possesses the following symmetries:

- Skew-symmetry: $R^i{}_{jkl} = -R^i{}_{jlk} \Leftrightarrow \forall v, w \in T_pM, T \in (T^r_s M)_p\ R(v, w)T = -R(w, v)T$.
- First Bianchi identity: If L is symmetric, then

$$\begin{aligned} &R^i{}_{jkl} + R^i{}_{klj} + R^i{}_{ljk} = 0 \\ &\Leftrightarrow \forall v, w, z \in T_pM,\ R(v, w)z + R(w, z)v + R(z, v)w = 0. \end{aligned}$$

- Second Bianchi identity:

$$\begin{aligned} &R^i{}_{jkl|m} + R^i{}_{jlm|k} + R^i{}_{jmk|l} = 0 \\ &\Leftrightarrow \forall v, w, z \in T_pM,\ \nabla_v R \lrcorner (w \otimes z) + \nabla_w R \lrcorner (z \otimes v) + \nabla_z R \lrcorner (v \otimes w) = 0 \end{aligned}$$

The curvature tensor may also be naturally interpreted in terms of parallel displacement. We first note the following commutation formulæ for absolute derivatives along families of curves.

Proposition 12.6. Suppose $\{\gamma(s,\cdot) : I \to M\}_{s\in J}$ is a smooth 1-parameter family of curves. Then the commutation formula

$$\frac{\delta}{\delta s}\frac{\partial\gamma}{\partial t} - \frac{\delta}{\delta t}\frac{\partial\gamma}{\partial s} = -2(\Omega\circ\gamma)\lrcorner(\frac{\partial\gamma}{\partial s}\otimes\frac{\partial\gamma}{\partial t})$$

holds, and if T is any tensor field along γ, then

$$\frac{\delta}{\delta s}\frac{\delta T}{\delta t} - \frac{\delta}{\delta t}\frac{\delta T}{\delta s} = R(\frac{\partial\gamma}{\partial s},\frac{\partial\gamma}{\partial t})T.$$

Proof. We show the first formula, the second following using similar techniques. We compute that

$$\frac{\delta}{\delta s}\frac{\partial\gamma}{\partial t} = \frac{\delta}{\delta s}(\frac{\partial\gamma^i}{\partial t}\partial_i\circ\gamma) = \frac{\partial^2\gamma^i}{\partial s\partial t}\partial_i\circ\gamma + \frac{\partial\gamma^i}{\partial t}\cdot\frac{\partial\gamma^j}{\partial s}\nabla_{\partial_j\circ\gamma}\partial_i$$

so that using (2.8), we see that

$$\frac{\delta}{\delta s}\frac{\partial\gamma}{\partial t} - \frac{\delta}{\delta t}\frac{\partial\gamma}{\partial s} = \frac{\partial\gamma^i}{\partial t}\cdot\frac{\partial\gamma^j}{\partial s}\left(\nabla_{\partial_j}\partial_i - \nabla_{\partial_i}\partial_j\right)\circ\gamma = -2(\Omega\circ\gamma)\lrcorner(\frac{\partial\gamma}{\partial s}\otimes\frac{\partial\gamma}{\partial t}).$$ □

The following theorem provides the link between parallelism and curvature

Theorem 12.7. Let $\{\gamma(s,\cdot) : [0,1] \to M\}_{s\in[0,1]}$ be smooth 1-parameter family of curves such that $\gamma(\cdot,0)\equiv p$ and $\gamma(\cdot,1)\equiv q$. Fix a vector $v\in T_pM$ and a covector $\nu\in T_q^*M$ and set $v(s,t) := P^t_{\gamma(s,\cdot),0}v$ and $\nu(s,t) := P^t_{\gamma(s,\cdot),1}\nu$. Then

$$(\nu, v(1,1) - v(0,1)) = \iint_{[0,1]\times[0,1]}(\nu, R(\frac{\partial\gamma}{\partial t},\frac{\partial\gamma}{\partial s})v)$$

Proof. By the fundamental theorem of calculus,

$$\begin{aligned}(\nu, v(1,1) - v(0,1)) &= (\nu(1,1), v(1,1)) - (\nu(0,1), v(0,1)) \\ &= \int_0^1 (\frac{\delta\nu}{\delta s}(s,1), v(s,1)) + (\nu(s,1), \frac{\delta v}{\delta s}(s,1))\mathrm{d}s.\end{aligned} \tag{2.24}$$

Since $\gamma(\cdot,1)\equiv q$ and $\nu(\cdot,1)\equiv\nu$ is constant, $\frac{\delta\nu}{\delta s}(s,1) = 0$. Moreover, since $\gamma(\cdot,0)\equiv p$ and $v(\cdot,0)\equiv v$ is constant, it follows that $\frac{\delta v}{\delta s}(s,0) = 0$ so that we may use the fundamental theorem of calculus again to write

$$(\nu(s,1), \frac{\delta v}{\delta s}(s,1)) = \int_0^1 \frac{\partial}{\partial t}(\nu(s,t), \frac{\delta v}{\delta s}(s,t))\mathrm{d}s.$$

Finally, since $\nu(s,\cdot)$ and $v(s,\cdot)$ are parallel along $\gamma(s,\cdot)$, we have that $\frac{\delta\nu}{\delta t}\equiv 0$ and $\frac{\delta v}{\delta t}\equiv 0$ so that the Leibniz rule and commutation formula of Proposition implies

$$\frac{\partial}{\partial t}(\nu, \frac{\delta v}{\delta s}) = (\nu, \frac{\delta}{\delta t}\frac{\delta v}{\delta s}) = (\nu, R(\frac{\partial\gamma}{\partial t},\frac{\partial\gamma}{\partial s})v).$$

Plugging all of this back into (2.24) establishes the claim. □

Remark 12.8. We may use Theorem 12.7 to deduce the effect of curvature on parallel displacement in *simply connected regions*: Suppose $\Omega \subset M$ is simply connected and $c : [0,1] \to \Omega$ is a smooth closed curve, i.e. $c(0) = c(1) = p$. By simple connectedness, for $q = c(t_0)$ fixed with $q \neq p$, we may find a one-parameter family of curves γ as in Theorem 12.7 such that $\gamma(0,t) = c(t_0 t)$ and $\gamma(1,t) = c((t_0 - 1)t + 1)$; consequently, since parallel transport is parametrisation invariant and preserves interior products, we see that for v and ν as in Theorem 12.7,

$$\begin{aligned}(\nu(1,0), P^1_{c,0}v - v) &= (\nu(1,0), P^1_{c,t_0}(P^{t_0}_{c,0}v) - v)\\ &= (P^0_{\gamma(1,\cdot),1}\nu, P^0_{\gamma(1,\cdot),1}(P^1_{\gamma(0,\cdot),0}v) - v)\\ &= (\nu, P^1_{\gamma(0,\cdot),0}v - P^1_{\gamma(1,\cdot),1}v) = -\iint_{[0,1]\times[0,1]}(\nu, R(\frac{\partial\gamma}{\partial t}, \frac{\partial\gamma}{\partial s})v).\end{aligned}$$

Alternatively, the effect of parallel displacement along an infinitesimal loop may be shown to be governed by curvature. △

Corollary 12.9. Suppose $\Omega \subset M$ is simply connected and $R \equiv 0$ on Ω. Then given any $p \in \Omega$ and a vector $v \in T_pM$, there exists a unique parallel vector field $V \in \Gamma^1_0 M$ such that $V(p) = v$. In particular, parallel displacement along a curve depends only on the endpoints.

Proof sketch. For any $q \in \Omega$, we define $V(q) := P^b_{\gamma,0}V(p)$, where $\gamma : [0,b] \to \Omega$ is any smooth curve with $\gamma(0) = p$ and $\gamma(b) = q$. By Theorem 12.7, $V(q)$ is well-defined. To show that it is smooth, note that if $q \in \varphi^{-1}B(0,\varepsilon)$ with $B(0,\varepsilon) \subset \varphi(U)$ for some $\varepsilon > 0$ and coordinate system $\varphi|_U$ in a neighbourhood of p with $\varphi(p) = 0$, then we may choose $\gamma(t) = \varphi^{-1}(t\varphi(q))$, $t \in [0,1]$, as our curve. By smooth dependence, we see that V depends smoothly on q. To establish this for arbitrary $q \in \Omega$, we consider the set

$$\Omega_0 := \{q \in \Omega : \ V(q) \text{ smooth}\}.$$

and show that it is open and closed in Ω. By connectedness, it then follows that $\Omega_0 = \Omega$. □

13. Paths and the exponential map.

Besides giving rise to a notion of parallel displacement, affine connections give rise to a natural analogue of straight lines in a manifold.

Definition 13.1. A curve $c : I \to M$ on an open interval is said to be a *path*[3] if $\dot{c}$ is parallel along c, i.e.

$$\frac{\delta \dot{c}}{\delta t} = 0. \tag{2.25}$$

Example 13.2 (Euclidean space). If $M = \mathbb{R}^n$ so that $L^i_{jk} \equiv 0$, then (2.25) is equivalent to the equation $\frac{d^2c^i}{dt^2} \equiv 0$. Given the initial conditions $c^i(t_0) = p^i$ and $\dot{c}^i(t_0) = v^i$, we see that the unique solution to this equation is $c^i(t) = p^i + v^i \cdot (t - t_0)$, i.e. paths are none other than straight lines.

[3]Other common terms are *geodesic* and *autoparallel curve*. We will reserve the term geodesic for a special class of paths in a Riemannian manifold.

Example 13.3 (Lie groups). Let G be an n-dimensional Lie group with identity element e. Given $v \in T_eG$, there is a natural left-invariant vector field $X_v \in \Gamma_0^1 G$ such that $X_v(g) = \mathrm{d}_e\lambda_g$. Fixing a basis $\{b_i\} \subset T_eG$, we naturally obtain a parallelisation of G by left-invariant vector fields $\{\beta_i := X_{b_i}\}$ and therefore a natural affine connection which we call the *left-invariant affine connection* on G (cf. Example 9.8). It is easy to check that the affine connection so constructed does not depend on the choice of basis $\{b_i\}$ for T_eG. The paths in this affinely connected space are of the form

$$c(t) = \lambda_g \exp(tv)$$

for $g \in G$ and $v \in T_eG$.

Exercise 13.4. Verify the claims of Example 13.3 in detail.

In local coordinates, we may write (2.25) as

$$\frac{\mathrm{d}^2 c^i}{\mathrm{d}t^t} + L^i_{jk} \circ c \cdot \frac{\mathrm{d}c^j}{\mathrm{d}t} \cdot \frac{\mathrm{d}c^k}{\mathrm{d}t} = 0.$$

This is a nonlinear *second-order* system of ODE. We may treat it using the methods of §6 as follows:[4] First note that any coordinate system $\varphi|_U$ on M naturally gives rise to a local coordinate system $(z^1, \dots, z^{2n}) := \varphi_{(1)} : \pi^{-1}(U) \to \varphi(U) \times \mathbb{R}^n$ on TM, where $\pi : TM \to M$ is the projection map defined in §2, and $\varphi_{(1)}$ is given by

$$\varphi_{(1)}(v) = (\varphi(\pi(v)), \mathrm{d}x^1(v), \dots, \mathrm{d}x^n(v)).$$

Altogether, the collection $\{\varphi_{(1)} : \varphi \in \mathcal{A}\}$ forms an atlas on TM. A *long* computation then shows that the vector field *on the tangent bundle of* M given by

$$X(v) = \sum_{i=1}^{n} v^i \left.\frac{\partial}{\partial z^i}\right|_v - \sum_{i=n+1}^{2n} \sum_{j,k=1}^{n} L^i_{jk}(\pi(v)) \cdot v^j v^k \left.\frac{\partial}{\partial z^i}\right|_v \tag{2.26}$$

is *well-defined*, i.e. independent of the choice of coordinate system $\varphi_{(1)}$ from the above atlas, and $X \in \Gamma_0^1 TM$. This vector field gives rise to a local flow $\phi : Q \subset \mathbb{R} \times TM \to TM$ so that for each $v \in T_pM$, $I_v \ni t \mapsto \phi(t, v)$ is the unique maximal integral curve of (2.26). We now obtain a solution to our original system by defining

$$\begin{aligned} \gamma_v : I_v &\to M \\ t &\mapsto \pi(\phi(t, v)) \end{aligned}$$

for each $v \in TM$. Indeed, since $\varphi \circ \pi \circ \varphi_{(1)}^{-1} = \mathrm{pr}_1$, we have the relation $\mathrm{d}_v\pi(X(v)) = v$, whence $\dot\gamma_v(t) = \mathrm{d}_{\phi(t,v)}\pi(\dot\phi(t, v)) = \phi(t, v)$ and $\ddot\gamma_v(t) = \dot\phi(t, v) = X(\phi(t, v))$ so that comparison with X above shows γ_v is indeed a path. Conversely, any path $c : I \to M$ is such that $\dot c : I \to TM$ is an integral curve of X. Altogether, we have the following theorem.

Theorem 13.5. Given any $v \in T_pM$, there exists a path $\gamma_v : I_v \to M$ on an open interval $I_v \subset \mathbb{R}$ with $0 \in I_v$ such that $\gamma_v(0) = p$ and $\dot\gamma_v(0) = v$. Moreover, γ_v is unique and maximal in the following sense: If $c : I \ni 0 \to M$ is any other path with $c(0) = p$ and $\dot c(0) = v$, then $I \subset I_v$ and $c \equiv \gamma_v|_I$.

[4] The trick used here is the differential geometric analogue of the reduction of a normal system of higher order ODE to a first-order system, cf. e.g. [3, Ch. 6 §5].

Remark 13.6. There are a few peculiarities of the definition of a path worth noting:

- If $c : I \to M$ is a path and $\widetilde{c} := c \circ \chi$ is a *reparametrisation*, i.e. $\chi : J \to I$ is a diffeomorphism of intervals, then (2.25) implies that $\widetilde{c}$ is a path iff $\chi'' \equiv 0$, i.e. $\chi(t) = mt + c$ for $m, c \in \mathbb{R}$. Therefore, the parametrisation of a path is essentially fixed up to *affine reparametrisations*, and there is no loss of generality in considering paths in terms of initial conditions on intervals containing 0.

- For any $\sigma \in \mathbb{R}$ and $v \in T_pM$, we have the equality

$$\gamma_{\sigma v}(t) = \gamma_v(\sigma t)$$

  for all $t \in I_{\sigma_v} = \frac{1}{\sigma} I_v$. This follows from uniqueness and maximality by considering both sides of the equality as functions of t. △

- The vector field X of (2.26) may be written as

$$X(v) = \sum_{i=1}^{n} v^i \left.\frac{\partial}{\partial z^i}\right|_v - \sum_{i=n+1}^{2n} \sum_{j,k=1}^{n} \Gamma^i_{jk}(\pi(v)) \cdot v^j v^k \left.\frac{\partial}{\partial z^i}\right|_v,$$

  where Γ^i_{jk} is the symmetric part of L^i_{jk} (see §9). In particular, a manifold equipped with a connection L has the same paths as if it were equipped with the symmetrisation Γ of L. Finally, if we have an explicit description of all paths emanating from all $p \in M$ then, subject to appropriate regularity conditions, we may recover the symmetric part of L.

In light of these properties, we shall henceforth assume that L is symmetric, writing Γ in place of L, and introduce the following definition:

Definition 13.7. The *exponential map* $\exp : Q(1) \to M$ is defined to be $\exp(v) := \gamma_v(1)$, where $Q(1) = \{v \in TM : (1, v) \in Q\}$ and Q is the maximal domain of the local flow of the vector field (2.26). For each $p \in M$, we also introduce the *pointwise* exponential map $\exp_p := \exp|_{Q(1) \cap T_pM}$. If $Q(1) = TM$, we shall call Γ *complete.*

From §6, we know that $Q(1)$ is open in TM (and nonempty since $0_p \in Q(1)$ for all $p \in M$) and $\exp$ (resp. $\exp_p$) is a smooth mapping. Moreover, by Remark 13.6, we have the equality

$$\gamma_v(t) = \exp(tv) = \exp_{\pi(v)}(tv).$$

In the case where $M = \mathbb{R}^n$ and $\Gamma^i_{jk} \equiv 0$, $\exp_p(tv) = p + t\text{𝄞}_p^{-1} v$ and $\exp$ is defined on all of TM. In general, $Q(1)$ is a strict subset of TM.

Exercise 13.8. Consider T^1 equipped with the unique connection such that $\nabla_{\frac{d}{d\theta}} \frac{d}{d\theta} = \frac{d}{d\theta}$. Show that $Q(1) = \text{𝄞}_p\,]-1, \infty[$.

We note that for each $p \in M$, $\exp_p$ naturally gives rise to a *normal* coordinate system in which paths emanating from p assume a very simple form.

Lemma 13.9. For any $p \in M$, there exists a system of coordinates $\varphi|_U$ on a neighbourhood U of p such that $\varphi(p) = 0$ and whenever $v = v^i\, \partial_i|_p$ with respect to these coordinates, $\gamma_v(t) = \varphi(t \sum_{i=1}^n v^i e_i)$. Moreover, we have the relation $\Gamma^i_{jk}(\varphi^{-1}(x)) x^j x^k = 0$ for $x \in \varphi(U)$ so that $\Gamma^i_{jk}(p) = 0$.

Proof. Introduce a basis $\{b_i\}_{i=1}^n$ for T_pM and consider the following mapping for small $\varepsilon > 0$:

$$F : B(0,\varepsilon) \subset \mathbb{R}^n \to M$$
$$x \mapsto \exp_p(\sum_{i=1}^n x^i b_i)$$

Note that for $y \in \mathbb{R}^n$, $\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} F(tv) = \sum_{i=1}^n v^i b_i$ by definition of exp. Therefore, F is a local diffeomorphism so that by the inverse function theorem, F is a diffeomorphism onto its image for sufficiently small $\varepsilon > 0$. The desired chart is now given by $\varphi := F|_{F(B(0,\varepsilon))}^{-1}$. □

Remark 13.10. Note that the proof of Lemma 13.9 implies that $\exp_p$ is a diffeomorphism in a neighbourhood of $0 \in T_pM$. △

While the above results guarantee that we may find a unique path passing through a given point and travelling in any given direction, it is unclear whether there even exists a path passing through any two given points $p, q \in M$. We shall tackle this problem locally. We first establish the following lemma.

Lemma 13.11. For any $p \in M$, there exists a *simple* neighbourhood, i.e. an open neighbourhood $U \ni p$ such that for any $p_1, p_2 \in U$, there exists at most one path p_1p_2 completely contained in U.

Proof. We introduce the mapping $\Phi : Q(1) \to M \times M$ such that $\Phi(v) = (\pi(v), \exp(v))$. Let $\varphi|_V$ be a normal coordinate system in a neighbourhood of p as in Lemma 13.9 and $\varphi_{(1)}$ the induced coordinate system on $\pi^{-1}(V) \subset TM$. We show that $\mathrm{d}_{0_p}\Phi$ is an isomorphism as follows: Note that $\widetilde{\Phi} := (\varphi \times \varphi) \circ \Phi \circ \varphi_{(1)}^{-1} = \mathrm{pr}_1 \times \widetilde{\exp}$, where $\widetilde{\exp} = \varphi \circ \exp \circ \varphi_{(1)}^{-1}$. Now, since for $i \leq n$

$$\frac{\partial}{\partial x^i}\widetilde{\exp}(0) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \varphi(\exp(\varphi_{(1)}^{-1}(te_i))) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \varphi(\exp_{\varphi^{-1}(te_i)}(0)) = e_i$$

and similarly,

$$\frac{\partial}{\partial x^{i+n}}\widetilde{\exp}(0) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \varphi(\exp(\varphi_{(1)}^{-1}(te_{i+n}))) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \varphi(\exp(tb_i)) = e_i.$$

Therefore, we have that

$$\left(\frac{\partial\widetilde{\Phi}^i}{\partial x^j}\right)_{i,j} = \begin{pmatrix} I & 0 \\ I & I \end{pmatrix}.$$

Therefore, there exists a neighbourhood $\widetilde{U}$ of 0_p, which we may assume to be of the form $\varphi_{(1)}^{-1}(V \times B(0,\delta))$ for some neighbourhood V of 0 and $\delta > 0$ small. Now let $U := \varphi^{-1}(B(0,\delta_1))$ where $\delta_1 > 0$ is small enough such that $U \times U \subset \Phi(\widetilde{U})$. If for $p_1, p_2 \in U$ there exists a path $\gamma : [0,1] \to M$ such that $\gamma(0) = p_1$ and $\gamma(1) = p_2$, then since $\gamma(t) \in U$, we have that $(p_1, \gamma(t)) \in U \times U \Rightarrow v(t) := \Phi^{-1}(p_1, \gamma(t)) \in \widetilde{U}$. But $\gamma(t) = \mathrm{pr}_2(\Phi(v(t))) \Rightarrow \dot{\gamma}(0) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \mathrm{pr}_2(\Phi(v(t)))$. By the existence and uniqueness theorem, this uniquely determines γ. Note that if instead $\gamma(0) = p_2$ and $\gamma(1) = p_1$, then considering the affine reparametrisation $t \mapsto \gamma(1-t)$ puts us in the realm of the preceding case. □

Note that this lemma says that we may choose a neighbourhood of any given point in such a way that there is at most one path joining any two points inside of this neighbourhood. The following theorem, due to Whitehead [14], says that we may actually choose this neighbourhood in such a way that there is exactly one path passing through any two if its points.

Theorem 13.12. For any $p \in M$, there exists a simple *convex* neighbourhood, i.e. an open neighbourhood $V \ni p$ such that for any $p_1, p_2 \in V$, there exists a unique path passing through p_1 and p_2.

Proof. Let U be a simple neighbourhood of p as in Lemma 13.11. We first establish the following auxiliary result:

- If $f : U \to \mathbb{R}$ is a smooth function such that $\{f < 0\}$ is connected and the quadratic form $f_{|ij}$ is positive definite on $\{f = 0\}$, then $\{f < 0\}$ is convex.

  Subproof. Set $S := \{f < 0\}$ and fix $p_0 \in S$ and let $C = \{q \in S : \exists p_0 q \text{ lying in } S\}$. Firstly, $C \neq \varnothing$ because $p_0 \in C$.

  We first show that C is closed in S: Suppose $\{q_n\} \subset C$ with $q_n \to q \in S$. Since $p_0 q_n$ lies in S and therefore U, we may write it in the form $\gamma_n(t) = \exp(tv_n)$ with $v_n = \Phi|_{\tilde{U}}^{-1}(p_0, q_n) \in T_{p_0}M$. By the continuity of $\Phi|_{\tilde{U}}^{-1}$, $v_n \xrightarrow{n\to\infty} v = \Phi|_{\tilde{U}}^{-1}(p_0, q)$. Now, the curve $\gamma(t) = \exp(tv)$ is such that $\gamma(0) = p_0 \in S$, $\gamma(1) = q \in S$ and $f(\gamma(t)) \leq 0$. However, the equality $f(\gamma(t)) = 0$ is impossible: Suppose t_0 is the smallest number such that $f(\gamma(t_0)) = 0$. Since $f(\gamma(t)) \leq 0$, we must also have $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=t_0}(f \circ \gamma) = 0$. A straightforward computation shows that for small h,

  $$f(\gamma(t_0 + h)) = \frac{1}{2} f_{|ij}\dot{\gamma}^i(t_0)\dot{\gamma}^j(t_0)h^2 + O(h^3) \text{ as } h \to 0,$$

  which implies that $f(\gamma(t_0 + h)) > 0$ for small h, but this contradicts the fact that $f(t) < 0$ for $t < t_0$. Therefore, $f \circ \gamma < 0 \Leftrightarrow \gamma(t) \in S$ for all t.

  To see that C is open in S, note that for all $t \in [0, 1]$ and $q \in C$, we have that $\exp(t\Phi^{-1}(p_0, q)) \in S$. By continuity and compactness, we have that

  $$\exp(t\Phi^{-1}(p_0, q_1)) \in S$$

  for all $t \in [0, 1]$ and q_1 sufficiently close to q. This shows openness. Since S is connected, we must have $C = S$. □

To complete the proof, we need only exhibit an f satisfying the properties above and take $V = \{f < 0\}$. For fixed $r > 0$, let $f : U \to \mathbb{R}$ be defined by

$$f(p) = \frac{1}{2}(|\varphi(p)|^2 - r^2) = \frac{1}{2}(\sum_{i=1}^{n} x^i(p)^2 - r^2),$$

where φ is the normal coordinate system of Lemma 13.9. The set $S = \{f < 0\}$ is clearly connected, being the continuous image of a ball in Euclidean space. On the other hand, $f_{|i} = x^i$ so that

$$f_{|ij} = \delta_{ij} - \sum_{k=1}^{n} x^k \Gamma_{ij}^k.$$

Note that for $v \in T_qM$ with $q \in S$,

$$f_{|ij}v^iv^j = |v|^2 - \sum_{i,j,k} x^k v^i v^j \Gamma^k_{ij} \geq |v|^2 \cdot \left(1 - r \cdot \sup_{\{f=0\}} |\Gamma|\right)$$

where $|\Gamma| := \sqrt{\sum_{i,j,k}(\Gamma^k_{ij})^2}$. Since $\Gamma^k_{ij}(p) = 0$, we can guarantee that $f_{|ij}v^iv^j > 0$ by choosing r sufficiently small. The auxiliary result above therefore applies and we have established the theorem. □

14. Jacobi fields. We now turn our attention to families of paths in an affinely connected manifold. Suppose $\{\gamma(s,\cdot) : [a,b] \to M\}_{s\in]-\varepsilon,\varepsilon[}$ is a one-parameter family of paths and consider the *infinitesimal variation* $X(t) := \frac{\partial\gamma}{\partial s}(0,t)$ of $\gamma := \gamma(0,\cdot)$. It is clear that $X \in {}_\gamma\Gamma^1_0 M$. Moreover, we compute that

$$\frac{\delta^2 X}{\delta t^2} = \left(\frac{\delta}{\delta t}\frac{\delta}{\delta t}\frac{\partial\gamma}{\partial s}\right)(0,\cdot) = \left(\frac{\delta}{\delta t}\frac{\delta}{\delta s}\frac{\partial\gamma}{\partial t}\right)(0,\cdot) = R(\dot\gamma, X)\dot\gamma.$$

Such vector fields play an important rôle in studying the behaviour of the exponential map and immediately imply a link to the curvature tensor.

Definition 14.1. A *Jacobi field* along a path γ is a vector field $J \in {}_\gamma\Gamma^1_0 M$ solving the *Jacobi equation*

$$\frac{\delta^2 J}{\delta t^2} + R(J,\dot\gamma)\dot\gamma = 0. \tag{2.27}$$

The Jacobi equation is clearly linear in J so that the set of all Jacobi fields is an $\mathbb{R}$-vector subspace of ${}_\gamma\Gamma^1_0 M$. On the other hand, it turns out that it is *finite dimensional*: Let $\{b_i\}_{i=1}^n$ be a parallel frame along γ and write $J = J^i b_i$. We see that (2.27) is equivalent to the second-order system of linear ODE

$$\ddot J^i + J^j(b^i, R(b_j,\dot\gamma)\dot\gamma) = 0.$$

We therefore obtain the following result.

Theorem 14.2. Let $\gamma : [a,b] \to M$ be a path and $J_0, J_1 \in T_{\gamma(a)}M$. There exists a unique Jacobi field $J \in {}_\gamma\Gamma^1_0 M$ such that $J(a) = J_0$ and $\frac{\delta J}{\delta t}(a) = J_1$. In particular, the space of all Jacobi fields along γ is $2n$ dimensional.

As we saw above, Jacobi fields naturally arise as infinitesimal variations of paths. If in particular we consider a family of paths $\{\gamma(s,\cdot) : [a,b] \to M\}_{s\in]-\varepsilon,\varepsilon[}$ such that $\gamma(\cdot,a) \equiv \gamma(0,a)$, we obtain a Jacobi field J such that $J(a) = 0$. Such Jacobi fields may be represented explicitly by means of the exponential map.

Lemma 14.3. Suppose J is a Jacobi field along a path given by $\gamma(t) = \exp_p(tv)$ such that $J(0) = 0$. Then

$$J(t) = \left.\frac{\partial}{\partial s}\right|_{s=0} \exp(t(v + s\frac{\delta J}{\delta t}(0))) = t\mathrm{d}_{tv}\exp(\S_{tv}\frac{\delta J}{\delta t}(0)). \tag{2.28}$$

Proof. For each s, $t \mapsto \exp(t(v + s\frac{\delta J}{\delta t}(0)))$ is a path so that the right-hand side of (2.28) defines a Jacobi field. It is clear that the right-hand side vanishes at $t = 0$ and has absolute derivative $\frac{\delta J}{\delta t}(0)$ at $t = 0$. That this field coïncides with J now follows from Theorem 14.2. □

Jacobi fields vanishing at both endpoints of a path characterise those points where the pointwise exponential map fails to be a diffeomorphism. To make this notion precise, we introduce the following definition.

Definition 14.4. Two points $p, q = \exp_p(v) \in M$ are said to be *conjugate* if there exists a Jacobi field J along the path $[0, 1] \ni t \mapsto \exp(tv)$ such that $J(0) = 0$, $J(1) = 0$ and $J \not\equiv 0$.

Proposition 14.5. Two points $p, \exp_p(v) \in M$ are conjugate iff $\mathrm{d}_v \exp_p$ is *singular* (i.e. is of rank less than n).

Proof. We see from the definition of conjugacy and Lemma 14.3 that $p, \exp_p(v)$ are conjugate iff there is a Jacobi field J along $t \mapsto \exp(tv)$ with

$$0 = J(1) = \mathrm{d}_v \exp(\text{𝄞}_v \frac{\delta J}{\delta t}(0)).$$ □

Exercise 14.6. Suppose $p, q := \exp_p(v)$ are not conjugate. Then given any $v \in T_pM$ and $w \in T_qM$, there exists a unique Jacobi field J along $t \mapsto \exp_p(tv)$ such that $J(0) = v$ and $J(1) = w$.

Riemannian manifolds

We saw in the preceding chapter how the notion of an affine connection naturally gave rise to a notion of differentiation of tensor fields, parallel displacement of tensor fields along a curve and curvature, as well as paths, an analogue of the classical notion of a straight line. In this chapter, we turn our attention to a class of manifolds which, roughly speaking, possesses an inner product on each of its tangent space that varies smoothly in a certain sense, which in turn allows us to measure lengths of (contravariant) vectors, angles between vectors, lengths of curves and ultimately *distances between points*. Furthermore, there is a natural affine connection associated to this additional structure, naturally making a Riemannian manifold affinely connected. General references for this chapter are [7],[10], [9] (modern) and [5], [11] (classical).

15. Definition and basic structure.

Definition 15.1. A *Riemannian metric* on M is a symmetric tensor $g \in \Gamma^0_2 M$ such that for each $p \in M$, the bilinear form

$$T_pM \times T_pM \ni (v, w) \mapsto (g(p), v \otimes w) = g_{ij}(p)v^i w^j$$

is positive definite. A pair (M, g) consisting of a manifold and Riemannian metric is said to be a *Riemannian manifold*.

Remark 15.2. In light of the positive-definiteness condition, a Riemannian metric may be viewed as a smoothly varying pointwise *inner product*. We shall usually use the notation

$$\langle v, w\rangle := (g(p), v \otimes w)$$

for given contravariant vectors $v, w \in T_pM$. Consequently, a Riemannian manifold naturally yields analogues of classical notions of geometry:

- *Lengths and angles*: Given a contravariant vector $v \in T_pM$, we define the *length* of v to be
$$|v| := \sqrt{\langle v, v\rangle} = \sqrt{g_{ij}v^i v^j},$$
and given two contravariant vectors $v, w \in T_pM$, we define the *angle* between v and w to be that $\measuredangle(v, w) \in [0, 2\pi[$ such that
$$\cos \measuredangle(v, w) = \frac{\langle v, w\rangle}{|v| \cdot |w|}.$$
- *Arc length*: Suppose $\gamma : I \longrightarrow M$ is a *piecewise smooth curve* on the interval I, i.e. γ is continuous and there is a subdivision $I = \bigcup_{i=1}^N I_i$, I_i intervals whose interiors are pairwise disjoint, such that $\gamma|_{\overline{I_i}}$ is smooth. Then $\dot\gamma(t)$ is defined, smooth and bounded for all but finitely many $t \in I$. We define the *arc length* of γ to be
$$L(\gamma) := \int_I |\dot\gamma|.$$

Note that L is invariant under reparametrisations of γ: If $\chi : J \to I$ is any smooth monotone function, then $L(\gamma \circ \chi) = L(\gamma)$. In the sequel, we shall seek curves of minimal length, for which purpose we introduce the set

$$\Omega_{p,q} := \{\gamma : [0,1] \to M \text{ piecewise smooth} : \gamma(0) = p, \gamma(1) = q\}$$

for each $p, q \in M$. By parametrisation invariance, it is sufficient to consider such curves parametrised on the unit interval. △

Exercise 15.3. The *energy* functional $E : \Omega_{p,q} \to \mathbb{R}$ is defined by

$$E(\gamma) = \frac{1}{2} \int_0^1 |\dot{\gamma}|^2.$$

(a) Show that the inequality $L(\gamma) \leq \sqrt{2E(\gamma)}$ holds for all $\gamma \in \Omega_{p,q}$ with equality holding iff $|\dot{\gamma}|$ is equal to a constant (almost everywhere).

(b) Suppose that $\gamma^* \in \Omega_{p,q}$ minimises L, i.e. $L(\gamma^*) \leq L(\gamma)$ for all $\gamma \in \Omega_{p,q}$, and is such that $|\dot{\gamma}^*|$ is equal to a constant (almost everywhere). Show that γ^* also minimises E.

Remark 15.4 (orthonormal frames)**.** Analogously to the notion of orthonormal basis, we say that a local frame $\{\varepsilon_i \in \Gamma^1_0 M|_U\}_{i=1}^n$ for TM is *orthonormal* if for all $p \in U$,

$$\langle \varepsilon_i(p), \varepsilon_j(p) \rangle = 0.$$

Local orthonormal frames always exist. For instance, one could reduce a given local frame (e.g. $\{\partial_i\}_{i=1}^n$) to an orthonormal one by employing the Gram-Schmidt algorithm pointwise. Note that if $\{\varepsilon^i\}_{i=1}^n$ is the local frame dual to $\{\varepsilon_i\}$, we may locally write

$$g = \sum_{i,j} \delta_{ij} \varepsilon^i \otimes \varepsilon^j = \sum_i \varepsilon^i \otimes \varepsilon^i.$$ △

Exercise 15.5. The *unit sphere bundle* is defined to be

$$SM := \{v \in TM : |v|^2 = 1\}.$$

(a) Show that SM is an embedded submanifold of TM.

(b) Show that if M is compact, then so is SM.

(c) Suppose that M is compact and equipped with the Levi-Cività connection. Show that if $\gamma_v : I_v \to M$ is the unique maximal path with $\dot{\gamma}_v(0) = v \in T_pM$, then $I_v = \mathbb{R}$.

Remark 15.6 (musical isomorphisms)**.** Since $\langle \cdot, \cdot \rangle : T_pM \times T_pM \to \mathbb{R}$ is non-degenerate, the mapping

$$(\cdot)^\flat : T_pM \to T_p^*M$$

is an *isomorphism* with inverse denoted by $(\cdot)^\sharp : T_p^*M \to T_pM$. In local coordinates,

$$(v^i\, \partial_i|_p)^\flat = g_{ij}(p) v^j \,\mathrm{d}x^i\big|_p, \tag{3.1}$$

and writing $(g^{ij}) := (g_{ij})^{-1}$ for the inverse matrix of the matrix of components of g,

$$(v_i \,\mathrm{d}x^i\big|_p)^\sharp = g^{ij}(p) v_j\, \partial_i|_p. \tag{3.2}$$

Therefore, on a Riemannian manifold, we have a natural identification of contravariant vectors with covariant vectors, and the above expressions show that if X is a vector field and ω a covector field, then $X^\flat$ is a covector field and $\omega^\sharp$ a vector field with the same level of smoothness as X and ω respectively. The same holds for fields along curves.

As a general rule, given *any* local frame $\{\varepsilon_i\}$ for TM with dual coframe $\{\varepsilon^i\}$ for T^*M, we will locally represent the isomorphisms $\flat$ and $\sharp$ by *lowering* and *raising* indices, i.e. given a contravariant vector $v := v^i\varepsilon_i(p)$ and a covariant vector $\nu := \nu_i\varepsilon^i(p)$, we shall write $v_i := \left(v^\flat\right)_i$ and $\nu^i := \left(\nu^\sharp\right)^i$. In particular, if $\varepsilon_i = \partial_i$, the lowered and raised components are given by (3.1) and (3.2), whereas if $\{\varepsilon_i\}_{i=1}^n$ form a local orthonormal frame, then $v_i = v^i$ and $\nu^i = \nu_i$.

By taking tensor products of the musical isomorphisms with identity mappings, we can also raise and lower indices of higher-order tensors. For example, if $T \in \Gamma_2^0 M$ with components T_{ij} in a local frame, we obtain a tensor $(\sharp \otimes \mathrm{id}_{T^*M})T$ with components $T^i_{\ j} := g^{ik}T_{kj}$ in any local frame. △

Remark 15.7. Using g, we may also take traces *with respect to* g of arbitrary higher-order tensors with respect to pairs of covariant or contravariant indices: For instance, if $T \in \Gamma_2^0 M$, we define $\mathrm{tr}_{\,g} T$ to be $\mathrm{tr}\,(\sharp \otimes \mathrm{id}_{T^*M})T$. With respect to a local frame, $\mathrm{tr}_{\,g} T = g^{ij}T_{ij}$. △

Remark 15.8. In light of Remark 15.6, we naturally obtain an inner product on each cotangent space

$$\begin{aligned} \langle \cdot, \cdot \rangle : T_p^*M &\to T_p^*M \\ (\nu, \omega) &\mapsto \langle \nu, \omega \rangle := \left\langle \nu^\sharp, \omega^\sharp \right\rangle ; \end{aligned}$$

this naturally gives rise to a well-defined $(2,0)$-tensor field $g^* \in \Gamma_0^2 M$ such that for all $p \in M$ and $\nu, \omega \in T_p^*M$, $(g^*(p), \nu \otimes \omega) = \langle \nu, \omega \rangle$. If $\{\varepsilon_i\}_{i=1}^n$ is any local frame with dual $\{\varepsilon^i\}$, we may write

$$g^* = g^{ij} \cdot \varepsilon_i \otimes \varepsilon_j,$$

where $(g^{ij}) := (g_{ij})^{-1}$ and $g_{ij} = (g, \varepsilon_i \otimes \varepsilon_j)$, so that smoothness follows by taking $\varepsilon_i := \partial_i$ and comparing with Remark 15.6. More generally, g induces a unique inner product $\langle \cdot, \cdot \rangle$ on each tensor space $(T_s^r M)_p$ such that for $v_1, \ldots, v_r, w_1, \ldots w_r \in T_pM$ and $\nu^1, \ldots, \nu^s, \omega^1, \ldots, \omega^s \in T_p^*M$,

$$\begin{aligned} &\left\langle v_1 \otimes \cdots \otimes v_r \otimes \nu^1 \otimes \cdots \otimes \nu^s, w_1 \otimes \cdots \otimes w_r \otimes \omega^1 \otimes \cdots \otimes \omega^s \right\rangle \\ &\quad = \langle v_1, w_1 \rangle \ldots \langle v_r, w_r \rangle \cdot \left\langle \nu^1, \omega^1 \right\rangle \ldots \langle \nu^s, \omega^s \rangle , \end{aligned}$$

as follows from the universal mapping principle for tensor products applied twice. In particular, for $S, T \in (T_s^r M)_p$, we have in any local frame that

$$\langle S, T \rangle = S^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} T^{k_1 \ldots k_r}{}_{l_1 \ldots l_s} \left(g_{i_1 k_1} \ldots g_{i_r k_r} \cdot g^{j_1 l_1} \ldots g^{j_s l_s} \right)(p).$$

Note that if $S, T \in \Gamma_s^r M\big|_U$ for $U \subset M$ open, then $\langle S, T \rangle \in C^\infty(U)$. Moreover, the musical isomorphisms and their tensor products with combinations of id_{T_pM} and $\mathrm{id}_{T_p^*M}$ are all *isometries* with respect to these inner products. △

The next theorem tells us that Riemannian manifolds are naturally affinely connected. Consequently, all of the theory developed in Chapter 2 applies here.

Theorem 15.9 (fundamental theorem of Riemannian geometry). There exists a unique *symmetric* affine connection Γ with respect to which g is parallel, i.e.

$$\nabla g \equiv 0 \Leftrightarrow \forall X, Y, Z \in \Gamma_0^1 M \ \partial_X \langle Y, Z\rangle = \langle \nabla_X Y, Z\rangle + \langle Y, \nabla_X Z\rangle .$$

Proof sketch. There are two approaches:

- Local: The condition $\nabla g \equiv 0$ is equivalent in local coordinates to

$$0 = g_{ij|k} = \partial_k g_{ij} - \Gamma^l_{ik} g_{lj} - \Gamma^l_{jk} g_{il}.$$

 Using the symmetry of Γ^i_{jk} in j and k, we find that

$$\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij} = 2 g_{kl} \Gamma^l_{ij}.$$

 This yields $\Gamma^k_{ij} = \frac{1}{2} g^{kl}(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij})$ so that Γ is uniquely determined. Noting that the tensor transformation law for g in coordinate systems φ and $\tilde{\varphi}$ implies

$$\tilde{\partial}_k g_{ij} = \left(\tilde{\partial}_k \tilde{\partial}_i x^p \tilde{\partial}_j x^q + \tilde{\partial}_i x^p \tilde{\partial}_k \tilde{\partial}_j x^q\right) g_{pq} + \tilde{\partial}_i x^p \tilde{\partial}_j x^q \tilde{\partial}_k x^l \partial_l g_{pq},$$

 adding up the combinations of indices occurring in the expression for Γ^k_{ij} above immediately shows that the Γ is indeed an affine connection.

- Invariant: Symmetry is equivalent to the vanishing of the torsion tensor, which amounts to the relation

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

 for all vector fields $X, Y \in \Gamma_0^1 M$. Since g is parallel, we have for all $X, Y, Z \in \Gamma_0^1 M$ that

$$\langle \nabla_X Y, Z\rangle = \partial_X \langle Y, Z\rangle - \langle Y, \nabla_X Z\rangle .$$

 Now, since Γ is symmetric, $-\langle Y, \nabla_X Z\rangle = -\langle Y, \nabla_Z X\rangle - \langle Y, [X, Z]\rangle$, but then since g is parallel, $-\langle Y, \nabla_Z X\rangle = -\partial_Z \langle Y, X\rangle + \langle \nabla_Z Y, X\rangle$. Using symmetry and $\nabla g = 0$ two more times as above, we are left with an expression of the form $-\langle \nabla_X Y, Z\rangle + \varphi(X, Y, Z)$. Rearranging, we obtain

$$\begin{aligned}&\langle \nabla_X Y, Z\rangle \\ &= \frac{1}{2}\left(\partial_X \langle Y, Z\rangle + \partial_Y \langle Z, X\rangle - \partial_Z \langle X, Y\rangle + \langle [X, Y], Z\rangle + \langle [Z, X], Y\rangle - \langle [Y, Z], X\rangle\right).\end{aligned}$$

 It is straightforward to check that the right-hand side is $C^\infty(M)$ linear in Z so that $\nabla_X Y$ is uniquely determined and ∇ defines a covariant differential arising from a symmetric connection such that $\nabla g \equiv 0$. □

Remark 15.10. The connection of Theorem 15.9 is called the *Levi-Cività connection.* We shall *always* assume that a given Riemannian manifold is equipped with this connection. Note that since $\nabla \delta \equiv 0$ for any affine connection,

$$g_{ij} g^{jk} = \delta^k{}_j \Rightarrow \underbrace{g_{ij|l}}_{=0} g^{jk} + g_{ij} g^{jk}{}_{|l} = 0 \Rightarrow g^{jk}{}_{|l} \equiv 0.$$

More generally, we obtain the following Leibniz rule for a contravariant vector (field) X and tensor fields $S, T \in \Gamma^r_s M$:

$$\partial_X \langle S, T\rangle = \langle \nabla_X S, T\rangle + \langle S, \nabla_X T\rangle . \qquad \triangle$$

Remark 15.11 (Levi-Cività displacement). The condition $\nabla g \equiv 0$ implies in particular that along any (piecewise) smooth curve $\gamma : I \longrightarrow M$, $\frac{\delta}{\delta t}(g \circ \gamma) = \nabla_{\dot\gamma} g \equiv 0$; consequently, for tensor fields $S, T \in {}_\gamma\Gamma^r_s M$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle S, T\rangle = \left\langle \frac{\delta S}{\delta t}, T \right\rangle + \left\langle S, \frac{\delta T}{\delta t}\right\rangle. \qquad \triangle$$

In particular, if S and T are parallel along γ, $\langle S, T\rangle$ must be *constant*. Thus, on a Riemannian manifold equipped with the Levi-Cività connection, *parallel displacement is an isometry*, i.e.

$$\left\langle P^t_{\gamma,t_0} S, P^t_{\gamma,t_0} T\right\rangle = \langle S, T\rangle$$

for all $t_0, t \in I$ and fixed $S, T \in (T^r_s M)_{\gamma(t_0)}$.

Remark 15.12 (differential operators of note). The presense of a metric gives rise to some familiar differential operators:

- Recall that for $f \in C^\infty(M)$, we have defined the differential of f as $\mathrm{d}f = \partial_{\varepsilon_i} f \cdot \varepsilon^i \in \Gamma^0_1 M$. The *gradient* of f is given by

$$\nabla f := (\mathrm{d}f)^\sharp = g^{ij}\partial_{\varepsilon_i} f \varepsilon_j \in \Gamma^1_0 M.$$

- Given a vector field $X \in \Gamma^1_0 M$, we may define its *divergence* by

$$\operatorname{div} X := \operatorname{tr} \nabla X = (\nabla X, \varepsilon_i \otimes \varepsilon^i) \in C^\infty(M).$$

  This gives rise to the *Laplacian* of a smooth function f: $\Delta f := \operatorname{div}(\nabla f)$. For more general tensors T, we may form $\operatorname{div} T$ by taking a trace of ∇T with respect to the *last* index (corresponding to the direction of differentiation) and any other index, usually the first. For example, for $T \in \Gamma^0_2 M$ as above,

$$\operatorname{div} T = T_{ij|k} \cdot g^{ik} = T^i{}_{j|i}\,.$$

$\triangle$

The following are some prominent examples of Riemannian manifolds:

- *Euclidean space*: $(\mathbb{R}^n, g_\delta)$ with $g_\delta := \delta_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j = \sum_{i=1}^n \mathrm{d}x^i \otimes \mathrm{d}x^i$.
- *Hyperbolic space*: $(\mathbb{H}^n, g_{\mathrm{hyp}})$ with $\mathbb{H}^n := \{x \in \mathbb{R}^n : x^n > 0\}$ and $g_{\mathrm{hyp}} := \frac{\delta_{ij}}{x^n}\mathrm{d}x^i \otimes \mathrm{d}x^j$.
- *Immersed spaces*: Let (M, g) be an n-dimensional Riemannian manifold and $F : S \longrightarrow M$ an immersion. Recall that $v \in T_q S$ is naturally identified with $\mathrm{d}_q F(v) \in T_{F(q)}M$. This suggests a natural inner product on each $T_q S$:

$$(v, w) \mapsto \langle v, w\rangle_a := \left\langle \mathrm{d}_q F(v), \mathrm{d}_q F(w)\right\rangle_g = (F^* g(q), v \otimes w).$$

  Setting

$$a := F^* g = \partial_i F^\alpha \cdot \partial_j F^\beta \cdot g_{\alpha\beta} \cdot \mathrm{d}x^i \otimes \mathrm{d}x^j \in \Gamma^0_2 M,$$

  we see that a defines a Riemannian metric on S. It is variously called the *subspace metric*, the *pullback of g by F* and the *first fundamental form of F*.
- *Sphere*: $(S^n, g_{\mathrm{sphere}})$ where g_{sphere} is the first fundamental form of the canonical embedding $\iota : S^n \longrightarrow (\mathbb{R}^{n+1}, g_\delta)$. This metric may be written out more explicitly using stereographic projection or polar coordinates.

- *Lie groups*: Let G be a Lie group and $\langle\cdot,\cdot\rangle_e$ be any inner product on T_eG, which we consider as an element $\alpha \in T_e^*G \otimes T_e^*G$. Using left translation, we define $\gamma_G \in \Gamma_2^0 G$ such that

$$\gamma_G(g) := \delta_g \lambda_{g^{-1}}(\alpha).$$

It may be readily verified that γ_G is indeed a Riemannian metric.

- *Flat torus*: (T^n, g_{T^n}) where g_{T^n} is given by the preceding example, the inner product α on $T_{\pi_n(0)}T^n$ being canonically given by

$$\alpha = (\delta_0 \pi_n)^{-1} g_\delta(0).$$

If we make use of the left-invariant (coordinate) basis fields $\{\frac{\partial}{\partial\theta^i}\}_{i=1}^n \subset \Gamma_0^1 T^n$, defined as the unique left-invariant vector fields with $\frac{\partial}{\partial\theta^i}\big|_{\pi_n(0)} := \mathrm{d}_0\pi_n \left(\frac{\partial}{\partial x^i}\big|_0\right)$, and let $\{\mathrm{d}\theta^i\}_{i=1}^n$ be the corresponding dual coframe, then we may write

$$g_{T^n} = \delta_{ij} \cdot \mathrm{d}\theta^i \otimes \mathrm{d}\theta^j.$$

- *Warped product manifolds*: Suppose (M, g_M) and (N, g_N) are Riemannian manifolds of dimension n and m respectively and $\phi \in C^\infty(M)$, $\psi \in C^\infty(N)$ *positive* functions. Define $g \in \Gamma_2^0(M \times N)$ such that in each coordinate system $\varphi \times \psi =: (x^1, \ldots, x^n, y^1 \ldots y^n)$ on $M \times N$ as in Exercise 1.13,

$$g := \phi(g_M)_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j + \psi(g_N)_{\alpha\beta}\mathrm{d}y^\alpha \otimes \mathrm{d}y^\beta.$$

It is straightforward to check that g is well defined and is a Riemannian metric, since g is none other than the (weighted) sum of g_M and g_N under the identifications of Exercise 2.8.

Remark 15.13. Riemannian metrics *always* exist: In each coordinate system $\varphi|_U$, we may equip each T_pM with the Euclidean metric g_U. Extending these local metric tensors to the rest of M arbitrarily and using a partition of unity $\{\phi_k\}$ subordinate to a covering of M by coordinate systems, letting U_k be a coordinate neighbourhood with $\operatorname{supp}\phi_k \Subset U_k$, we obtain a Riemannian metric by setting

$$g := \sum_k \phi_k \cdot g_{U_k}. \qquad \triangle$$

We shall henceforth assume given a Riemannian manifold (M, g).

16. Riemannian curvature. In §12, we saw that an affine connection gives rise to a natural notion of curvature in the form of a $(1,3)$-tensor field R, as well as the so-called curvature endomorphism. The Levi-Cività connection on a Riemannian manifold possesses a few additional curvature quantities:

- *Covariant curvature tensor*: Let

$$\begin{aligned} &\mathrm{Rm} := (\flat \otimes \mathrm{id}_{T^*M} \otimes \mathrm{id}_{T^*M} \otimes \mathrm{id}_{T^*M})R \\ \Leftrightarrow\ &\forall \{v_i\}_{i=1}^4 \in T_pM\ (\mathrm{Rm}(p), v_1 \otimes \cdots \otimes v_4) = \langle R(v_3, v_4)v_2, v_1\rangle\,. \end{aligned}$$

Clearly $\mathrm{Rm} \in \Gamma_4^0 M$, and in coordinates:

$$\mathrm{Rm} = g_{ir}R^r_{jkl}\mathrm{d}x^i \otimes \mathrm{d}x^j \otimes \mathrm{d}x^k \otimes \mathrm{d}x^l =: R_{ijkl}\mathrm{d}x^i \otimes \mathrm{d}x^j \otimes \mathrm{d}x^k \otimes \mathrm{d}x^l.$$

Rm possesses the following symmetries, which may be readily deduced from those of R or by working in geodesic normal coordinates:

– $R_{ijkl} = R_{ijlk}$;
– $R_{klij} = R_{ijkl}$;
– $R_{ijkl} + R_{iklj} + R_{iljk} = 0$.

We shall refer to the tensors Rm and R variously as the *curvature* or *Riemann tensor.*

- *Sectional curvature*: Suppose $\sigma := \mathrm{span}\{v, w\} \subset T_pM$ is a 2-plane and v, w are orthonormal. The *sectional, Gauß* or *Riemannian curvature* of M associated with σ is defined by

$$K(\sigma) := \langle R(v, w)w, v\rangle = (\mathrm{Rm}(p), v \otimes w \otimes v \otimes w) = R_{ijkl}(p)v^i w^j v^k w^l.$$

It is straightforward to check that this definition does not depend on the particular choice of orthonormal vectors representing σ. Moreover, if v and w are not orthonormal, the vectors $\frac{v}{|v|}$ and $\frac{w - \left\langle w, \frac{v}{|v|}\right\rangle \frac{v}{|v|}}{\sqrt{|w|^2 - \frac{\langle w, v\rangle^2}{|v|^2}}}$ are so that by the symmetries of R,

$$K(\sigma) = \frac{\langle R(v, w)w, v\rangle}{|v|^2|w|^2 - \langle v, w\rangle^2} = \frac{R_{ijkl}v^i w^j v^k w^l}{(g_{ik}g_{jl} - g_{il}g_{jk})v^i w^j v^k w^l}.$$

A Riemannian manifold is said to be *isotropic at* $p \in M$ if $K(\sigma)$ is independent of the choice of 2-plane σ in T_pM, i.e. if $K(\sigma_p) = f(p)$ for some smooth function $f \in C^\infty(M)$ and *any* 2-plane $\sigma_p \subset T_pM$, and of *constant sectional curvature* if $K(\sigma)$ is independent of the choice of 2-plane σ in *any* tangent space, i.e. if K is equivalently a constant.

- *Ricci and scalar curvatures*: By the symmetries of the of the Riemann tensor, there is only one potentially nontrivial trace of R up to sign:

$$R^i{}_{jik} = -R^i{}_{jki}.$$

We define the *Ricci curvature tensor* Ric $:= \Gamma^0_2 M$ by

$$\mathrm{Ric} = R^i{}_{jik}\mathrm{d}x^j \otimes \mathrm{d}x^k =: R_{jk}\mathrm{d}x^j \otimes \mathrm{d}x^k.$$

Note that Ric is symmetric. We may take a further trace with respect to g to obtain the *scalar curvature* or *curvature invariant*, written scal or R:

$$\mathrm{scal} := g^{ij}R_{ij}.$$

It is clear that the covariant curvature tensor carries as much information as the ordinary curvature tensor. With a bit of work, it may also be shown that specifying all sectional curvatures of M is equivalent to specifying the curvature tensor. The Ricci tensor and scalar curvature carry as much information as R if $n = 2$; in general, they may be interpreted as appropriate averages of the curvatures of M.

Exercise 16.1. Let $p \in M$, $\{v_i\}_{i=1}^n$ an orthonormal basis for T_pM and $\sigma_{ij} := \mathrm{span}\{v_i, v_j\}$.

(a) Show that the Ricci tensor and scalar curvature satisfy the following relations:

$$(\mathrm{Ric}, v_i \otimes v_i) = \sum_{j \neq i} K(\sigma_{ij}); \qquad \mathrm{scal} = \sum_{i=1}^n \sum_{j \neq i} K(\sigma_{ij}).$$

(b) Let $\lambda > 0$ and define the *rescaled metric* $g^\lambda := \lambda g$. Show that the curvature quantities K^λ, Ric^λ and scal^λ of g^λ are related to those of g by

$$K^\lambda = \lambda^{-1}K; \qquad \mathrm{Ric}^\lambda = \mathrm{Ric}; \qquad \mathrm{scal}^\lambda = \lambda^{-1}\mathrm{scal}.$$

We turn our attention to isotropic and constant sectional curvature manifolds. The following lemma gives us a canonical form for isotropic manifolds.

Lemma 16.2. Suppose (M, g) is isotropic at $p \in M$ and $K(\sigma_p) = f(p)$ for each 2-plane $\sigma_p \subset T_pM$. Then at p the Riemann tensor takes the form

$$R_{ijkl} = f \cdot (g_{ik}g_{jl} - g_{il}g_{jk}). \tag{3.3}$$

Proof. Consider the tensor $T \in (T^0_4M)_p$ defined by

$$T_{ijkl} := R_{ijkl} - f \cdot (g_{ik}g_{jl} - g_{il}g_{jk}),$$

where all tensors on the right-hand side are evaluated at p. Note that T_{ijkl} possesses the same symmetries as R. Now, the relation $K(\sigma) = f(p)$ for $\sigma = \mathrm{span}\{v, w\}$ is equivalent to $T_{ijkl}(p)v^iw^jv^kw^l = 0$. By polarisation,

$$T_{ijkl} + T_{kjil} + T_{ilkj} + T_{klij} = 0 \Leftrightarrow T_{ijkl} + T_{kjil} = 0.$$

By the first Bianchi identity, $0 = T_{ijkl} + T_{kjil} = T_{ijkl} - T_{kilj} - T_{klji} = 2T_{ijkl} - T_{ikjl}$ so that $T_{ijkl} = \frac{1}{2}T_{ikjl}$. Swapping j and k yields $T_{ikjl} = \frac{1}{2}T_{ijkl}$ so that $T_{ijkl} \equiv 0$. ◻

Theorem 16.3 (Schur)**.** Suppose (M, g) is a connected Riemannian manifold of dimension $n \geq 3$ that is isotropic at *all* $p \in M$. Then M is of constant sectional curvature.

This theorem is a consequence of the following identity:

Lemma 16.4. The *Einstein tensor*, given by $G \in \Gamma^0_2M$ such that

$$G = \mathrm{Ric} - \frac{1}{2}\mathrm{scal}g,$$

is *divergence free*, i.e. $\mathrm{div}\, G = 0$.

Proof. Multiplying the second Bianchi identity $R^i{}_{jkl|m} + R^i{}_{jlm|k} + R^i{}_{jmk|l} = 0$ by g^{ik} and summing over i and k, we obtain the identity

$$R_{jl|m} + R^i{}_{jlm|i} - R_{jm|l} = 0.$$

Multiplying by g^{jm} and summing over j and m, we then obtain

$$(\mathrm{div}\,\mathrm{Ric})_l - \partial_l \mathrm{scal} + g^{jm}R^i{}_{jlm|i} = 0;$$

this last term may be simplified to $(\mathrm{div}\,\mathrm{Ric})_l$ using the symmetries of the Riemann tensor, whence the result follows. ◻

Proof of Theorem 16.3. In this case, $\mathrm{Ric} = (n-1)f \cdot g$ and $\mathrm{scal} = n(n-1)f$ so that the Einstein tensor takes the form

$$G = \left(n - 1 - \frac{1}{2}n(n-1)\right) fg = -\frac{1}{2}(n-1)(n-2)fg,$$

and by Lemma 16.4,

$$0 = \mathrm{div}\, G = -\frac{1}{2}(n-1)(n-2)\mathrm{d}f$$

so that f must be a constant. ◻

17. Riemannian manifolds as metric spaces. Let (M, g) be a connected Riemannian manifold. By virtue of the notion of arc length for piecewise smooth curves, we may naturally equip M with the structure of a *metric space* as follows: For $x, y \in M$, define

$$d(x, y) = \inf\{L(\gamma) : \gamma \in \Omega_{x,y}\}. \tag{3.4}$$

Lemma 17.1. The function $d : M \times M \to \mathbb{R}$ defined by (3.4) is a *metric.*

Proof. We check the relevant properties:

- *Positivity*: It is clear that we always have $d(x, y) \geq 0$ and that this lower bound is attained for $x = y$ (let $\gamma \equiv x$). Now, suppose $x \neq y$ and fix a local parametrisation $F : \Omega \to M$ such that $0 \mapsto x$ with respect to which we write g as the matrix g_{ij}. Since

$$\sum_{i,j=1}^{n} g_{ij}(F(z))v^i v^j > 0$$

for each $z \in \Omega$ and $v \in S^{n-1} \subset \mathbb{R}^n$. Therefore, by continuity and compactness, we may find a $c > 0$ such that for all $z \in B(0, \varepsilon_0) \Subset \Omega$ and $v \in S^{n-1}$, $\sum_{i,j=1}^{n} g_{ij}F(z)v^i v^j \geq c$ or, replacing v with $\frac{v}{|v|_{g_\delta}}$, we obtain

$$\sum_{i,j=1}^{n} g_{ij}(F(z))v^i v^j \geq c|v|_{g_\delta}^2 \tag{3.5}$$

for all $z \in B(0, \varepsilon_0)$ and $v \in \mathbb{R}^n$. Now, let $\varepsilon \in]0, \varepsilon_0[$ be small enough such that $y \notin F(B(0, \varepsilon))$. Then any continuous piecewise smooth curve $\gamma : [0, 1] \to M$ connecting x and y must be such $t^* := \inf\{t \in [0, 1] : \gamma(t) \in F(\partial B(0, \varepsilon))\} > 0$; in particular, by (3.5),

$$L(\gamma) \geq \int_0^{t^*} |\frac{\mathrm{d}}{\mathrm{d}t}(F^{-1} \circ \gamma)|_{g_\delta} \geq \varepsilon.$$

Since this holds for any such curve γ, we obtain $d(x, y) \geq \varepsilon > 0$.

- *Symmetry*: If $\gamma^- : [0, 1] \to M$ denotes the inverse curve to γ, then we see that $L(\gamma^-) = L(\gamma)$. Since the correspondence $\gamma \mapsto \gamma^-$ is bijective on the set of piecewise smooth curves on $[0, 1]$, we immediately obtain that $d(x, y) = d(y, x)$.

- *Triangle inequality*: For each $\varepsilon > 0$ and fixed $x, y, z \in M$, we may find curves $\gamma_1 : I_1 \to M$ and $\gamma_2 : I_2 \to M$ connecting x and z and y and z respectively such that $L(\gamma_1) \leq d(x, z) + \frac{\varepsilon}{2}$ and $L(\gamma_2) \leq d(y, z) + \frac{\varepsilon}{2}$. Now, using the parametrisation invariance of L, we see that

$$d(x, y) \leq L(\gamma_1) + L(\gamma_2) \leq d(x, z) + d(y, z) + \varepsilon$$

for all $\varepsilon > 0$. □

Two questions naturally arise:

- Does the topology induced by d coïncide with that of M?

- Given $p, q \in M$, does there exist a $\gamma \in \Omega_{p,q}$ such that $L(\gamma) = d(p, q)$?

We shall see that the answer to the first question is yes, whereas the latter question is a bit more subtle. We first attempt to answer the second question by searching for minimisers of L in a suitable subset of $\Omega_{p,q}$. Moreover, due to the parametrisation invariance of L, it is desirable to fix a canonical parametrisation.

Lemma 17.2. Suppose $\gamma \in \Omega_{p,q}$ is *regular*, i.e. $|\dot\gamma| \neq 0$ whenever it is defined. Then

- There exists a (piecewise smooth) $\chi : [0,1] \to [0,1]$ such that $c := \gamma \circ \chi$ is parametrised by *reduced arc length*, i.e. $|\dot c|$ is equal to a constant (almost everywhere).
- The curve c is a minimiser of L amongst all regular curves in $\Omega_{p,q}$ iff it is a minimiser amongst all such curves of the *energy* $E : \Omega_{p,q} \to \mathbb{R}$ defined by

$$E(\gamma) = \frac{1}{2}\int_0^1 |\dot\gamma|^2.$$

- The inequality $E(c) \le E(\gamma)$ holds.

Proof. Consider the function $f : [0,1] \to [0,1]$ defined by

$$f(t) = \frac{L(\gamma|_{[0,t]})}{L(\gamma)}.$$

f is bijective, differentiable and satisfies $f'(t) = \frac{1}{L(\gamma)}|\dot\gamma|$. By the inverse function theorem, its inverse χ is piecewise smooth and wherever it is defined, we have that

$$|\frac{\mathrm{d}}{\mathrm{d}t}(\gamma \circ \chi)|(t) = |\gamma'(\chi(t))| \cdot \chi'(t) = L(\gamma),$$

i.e. $\gamma \circ \chi \in \Omega'_{p,q}$. The latter claims follow from Exercise 15.3. □

In light of these observations, when considering the class of piecewise smooth regular curves, we may look for minimisers of the energy E instead of arc length. As a first step, we consider a sort of 'directional derivative' of E, which motivates the following definition.

Definition 17.3. A *broken variation* of a piecewise smooth curve $c : [a,b] \to M$ is a one-parameter family of curves $\{\gamma(s,\cdot) : [a,b] \to M\}_{s\in]-\varepsilon,\varepsilon[}$ such that

- $\gamma(0,\cdot) \equiv c$;
- $s \mapsto \gamma(s,t)$ is smooth for fixed t; and
- there is a subdivision $[a,b] = \bigcup_{i=1}^N [t_{i-1}, t_N]$ with $t_0 = a < t_1 < \cdots < t_N = b$ such that $\gamma|_{]-\varepsilon,\varepsilon[\times[t_{i-1},t_i]}$ is smooth for all i.

A broken variation $\{\gamma(s,\cdot) : [a,b] \to M\}$ of $c : [a,b] \to M$ is said to be *proper* if $s \mapsto \gamma(s,a)$ and $s \mapsto \gamma(s,b)$ are constant functions.

Remark 17.4. It is straightforward to verify that given any piecewise smooth vector field $X : [a,b] \to M$ along a piecewise smooth curve $c : [a,b] \to M$ with the same break points, the family of curves $\{t \mapsto \exp_{\gamma(t)}(sX(t))\}$ defines a broken variation of c for sufficiently small s and satisfies the equation $\frac{\partial\gamma}{\partial s}(0,\cdot) \equiv X$. △

Theorem 17.5 (first variation formula)**.** If $c : [a, b] \longrightarrow M$ is a piecewise smooth regular curve and $\{\gamma(s, \cdot)\}$ a broken variation of c with $\frac{\partial \gamma}{\partial s}(0, \cdot) \equiv X \in {}_c\Gamma^1_0 M$, then

$$\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} E(\gamma(s, \cdot)) = \sum_{i=1}^{N} \left\langle X(t_i), \dot{c}(t_i^-) - \dot{c}(t_i^+) \right\rangle + \langle X(b), \dot{c}(b^-) \rangle - \left\langle X(a), \dot{c}(a^+) \right\rangle - \int_a^b \langle X, \ddot{c} \rangle .$$

Corollary 17.6. If $c \in \Omega_{p,q}$ is regular, minimises L amongst all regular curves in $\Omega_{p,q}$ and is parametrised by reduced arc length, then c is a smooth path.

Proof. By the preceding results, c must be a minimiser of E amongst all regular curves in $\Omega_{p,q}$. In particular, the first variation formula must be equal to zero for all proper variations, i.e. whenever $X \in {}_c\Gamma^1_0 M$ with $X(0) = 0$ and $X(1) = 0$,

$$\sum_{i=1}^{N} \left\langle X(t_i), \dot{c}(t_i^-) - \dot{c}(t_i^+) \right\rangle - \int_a^b \langle X, \ddot{c} \rangle = 0.$$

We now establish the claim by choosing X appropriately. First of all, for each $i \in \{1, \ldots, N\}$, we may let $X = \chi \cdot \ddot{c}$, where $\chi \in C_0^\infty(]t_{i-1}, t_i[)$, in which case all of the boundary terms drop out and we are left with

$$-\int_{t_{i-1}}^{t_i} \chi \cdot |\ddot{c}|^2 = 0.$$

Since both χ and i are arbitrary, we must have that $\ddot{c} \equiv 0$ on $[a, b] \backslash \{t_i\}_{i=1}^N$. To show that c cannot have any break points, we choose X as follows: For each $i \in \{1, \ldots, N\}$, fix an orthonormal frame $\{\varepsilon_j\}_{j=1}^n$ in a neighbourhood of $c(t_i)$ and set $X = \chi \cdot \varepsilon_j \circ c$, where $\chi \in C_0^\infty(]\frac{t_{i-1}+t_i}{2}, \frac{t_i+t_{i+1}}{2}[)$ is any function equal to 1 at t_i. Using the first variation formula again, we are left with

$$\left\langle \varepsilon_j, \dot{c}(t_i^-) - \dot{c}(t_i^+) \right\rangle = 0.$$

Since j and i were arbitrary, we must have $\dot{c}(t_i^-) = \dot{c}(t_i^+)$ for each i, i.e. c is a C^1 path. By uniqueness of paths, it immediately follows that c must be smooth. □

In light of Corollary 17.6, we turn our attention to the study of paths in a Riemannian manifold, which we henceforth refer to as *geodesics*. From §13, we know that these are uniquely defined (up to affine reparametrisations) by their initial velocities, i.e. given $v \in T_pM$, there is a unique, maximal geodesic

$$\gamma_v(t) = \exp_p(tv).$$

such that $\gamma_v(0) = p$ and $\dot{\gamma}_v(0) = v$.

Proposition 17.7. Geodesics and the exponential map enjoy the following properties:

- Every geodesic γ_v has *constant speed* (or is parametrised by *reduced arc length*), i.e. $|\dot{\gamma}_v|$ is constant.
- Every geodesic γ_v emanating from p crosses *geodesic spheres*[1] of the form $\exp_p(\partial B(0_p, R))$ orthogonally (*Gauß Lemma*). More precisely, if $v, w \in T_pM$, then

  $$\langle v, w \rangle = \left\langle \mathrm{d}_v \exp_p(\text{𝄞}_v v), \mathrm{d}_v \exp_p(\text{𝄞}_v w) \right\rangle$$

  so that if $|v| = R$ and $w :]-\varepsilon, \varepsilon[\longrightarrow \partial B(0_p, R)$ is any smooth curve with $w(0) = v$, then $\left\langle \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=1} \exp(tv), \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=0} \exp(w(s)) \right\rangle = 0$.

[1] We assume R is small enough so that $\partial B(0_p, R)$ is contained in the domain of definition of $\exp_p$.

- The exponential map gives rise to *geodesic normal coordinates*: If $\psi : (\mathbb{R}^n, \langle\cdot,\cdot\rangle) \to (T_pM, \langle\cdot,\cdot\rangle_g)$ is any isometry of inner product spaces, then for sufficiently small $\delta > 0$, the mapping

  $$F : \exp \circ \psi : B(0,\delta) \to \exp(B(0_p,\delta))$$

  is a local parametrisation whose inverse yields a normal coordinate system such that $F(0) = p$ and with respect to which $g_{ij}(p) = \delta_{ij}$.

Proof. That every geodesic has constant speed follows from the fact that $\frac{\mathrm{d}}{\mathrm{d}t}\langle\dot\gamma_v,\dot\gamma_v\rangle = 2\langle\ddot\gamma_v,\dot\gamma_v\rangle = 0$. As for the geodesic normal coordinate system, we already know from the proof of Lemma 13.9 that F gives rise to a normal coordinate system; that $g_{ij}(p) = \delta_{ij}$ follows from the fact that

$$\partial_i F(0) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} F(te_i) = \mathrm{d}_0 \exp(\text{𝄞}_{0_p}\psi(e_i)) = \psi(e_i),$$

i.e. $g_{ij}(p) = \big\langle \partial_i F(0), \partial_j F(0)\big\rangle = \big\langle \psi(e_i), \psi(e_j)\big\rangle = \delta_{ij}$.

We now turn our attention to the Gauß Lemma. Consider the family of geodesics

$$\gamma(s,t) := \exp_p(t(v+sw))$$

for $t \in [0,1]$ and small s. It is easy to see that the following equations hold:

$$\frac{\partial\gamma}{\partial t}(0,1) = \mathrm{d}_v \exp_p(\text{𝄞}_v v) \qquad \frac{\partial\gamma}{\partial s}(0,1) = \mathrm{d}_v \exp_p(\text{𝄞}_v w)$$
$$\frac{\partial\gamma}{\partial t}(0,0) = v \qquad \frac{\partial\gamma}{\partial s}(0,0) = 0$$

Now, noting that by the constant speed property

$$\left|\frac{\partial\gamma}{\partial t}(s,t)\right|^2 = \left|\frac{\partial\gamma}{\partial t}(s,0)\right|^2 = |v+sw|^2$$

for all s, we see upon differentiating with respect to s at $s = 0$ that

$$\langle v,w\rangle = \left\langle \frac{\delta}{\delta s}\left(\frac{\partial\gamma}{\partial t}\right)(0,t), \frac{\partial\gamma}{\partial t}(0,t)\right\rangle = \left\langle \frac{\delta}{\delta t}\left(\frac{\partial\gamma}{\partial s}\right)(0,t), \frac{\partial\gamma}{\partial t}(0,t)\right\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\left\langle \frac{\partial\gamma}{\partial s}(0,t), \frac{\partial\gamma}{\partial t}(0,t)\right\rangle.$$

Integrating over $[0,1]$ yields $\langle v,w\rangle = \langle \partial_s\gamma(0,1), \partial_t\gamma(0,1)\rangle$. □

We now exploit this proposition to show that geodesic balls $\exp_p(B(0,R))$ coïncide with the balls $B(p,R)$ arising from d for sufficiently small R. To this end, for each $p \in M$, we define the *injectivity radius* at p to be[2]

$$\mathrm{inj}_p := \sup\{R > 0 : \exp_p|_{B(0_p,R)} \text{ is a diffeomorphism}\}.$$

Exercise 17.8. Show that the Gauß Lemma implies that in geodesic normal coordinates about p (with local parametrisation F),

$$g_{ij}(F(x))x^j = x^i.$$

In particular, if $\rho(p,\cdot) : \exp(B(0_p, \mathrm{inj}_p)) \to \mathbb{R}$ is the *geodesic distance from p* defined by $\rho(p,q) := |\exp_p|^{-1}_{B(0_p,\mathrm{inj}_p)}(q)|$, then

$$\mathrm{d}\rho(p,F(x)) = \sum_{i=1}^n \frac{x^i}{|x}\,\mathrm{d}x^i|_{F(x)}\,; \qquad \nabla\rho(p,F(x)) = \sum_{i=1}^n \frac{x^i}{|x|}\,\partial_i|_{F(x)}\,,$$

and both quantities have norm 1.

[2] It may be shown that the injectivity of $\exp_p$ on a ball is equivalent to it being a diffeomorphism there.

Lemma 17.9. If $p \in M$ and $q = \exp_p(v) \in \exp(B(0_p, \mathrm{inj}_p))$, then $d(p,q) = |v|$. Moreover, if $c : [0,1] \longrightarrow M$ is any admissible curve with $c(0) = p$ and $c(1) = q$, then $L(c) = d(p,q)$ if and only if c is the geodesic $t \mapsto \exp(tv)$ (up to reparametrisation).

Proof. Let $c \in \Omega_{p,q}$ and set

$$t^* := \inf\{t \in [0,1] : \rho(p, c(t)) = |v|\}.$$

By the Cauchy-Schwarz inequality and the Gauß Lemma,

$$\int_0^1 |\dot c| \geq \int_0^{t^*} |\dot c| \geq \int_0^{t^*} (\mathrm{d}\rho(p,\cdot)(c(t)), \dot c(t))\mathrm{d}t = \rho(p, c(t^*)) - \rho(p, c(0)) = |v|.$$

Therefore, $d(p,q) \geq |v|$. However, $L(c) = \int_0^1 |v| = |v|$ for the geodesic $t \mapsto \exp(tv)$, whence $d(p,q) \leq |v|$ so that we have equality.

Now suppose $L(c) = |v|$ and let $t_* = \sup\{t \in [0,1] : c(t) = p\}$. Since

$$|v| = \int_0^1 |\dot c| = \int_0^{t_*} |\dot c| + \int_{t_*}^{t^*} |\dot c| + \int_{t^*}^1 |\dot c| \geq |v| + \int_0^{t_*} |\dot c| + \int_{t^*}^1 |\dot c|,$$

we see that $\dot c \equiv 0$ outside of $[t_*, t^*]$. We now show that c must be a (possibly reparametrised) geodesic by writing it out in 'geodesic polar coördinates.' We may write $c(t) = \exp_p(r(t)\sigma(t))$ with $r(t) > 0$ continuous and $|\sigma(t)| = 1$ whenever $t \in]t_*, t^*[$. Moreover, on $]t_*, t^*[$ we have

$$\dot c(t) = \frac{\dot r(t)}{r(t)} \mathrm{d}_{r(t)\sigma(t)} \exp_p r(t)\sigma(t) + \mathrm{d}_{r(t)\sigma(t)} \exp_p r(t)\dot\sigma(t).$$

By the Gauß Lemma, the two terms on the right-hand side are mutually orthogonal, whence

$$|\dot c(t)|^2 = \dot r(t)^2 + r(t)^2 |\dot\sigma(t)|^2.$$

Moreover, since $\lim_{t \searrow t_*} c(t) = p$ and $c(t^*) = \exp_p(v)$, we have that $\lim_{t\searrow t_*} r(t) = 0$ and $r(t^*) = |v|$. Therefore, applying the Cauchy-Schwarz inequality again, we note that

$$|v| = \lim_{t \searrow t_*} \int_t^{t^*} |\dot c| \geq \lim_{t\searrow t_*} \int_t^{t^*} \dot r = |v|,$$

where the inequality is an equality iff $\dot\sigma \equiv 0$ and $\dot r \geq 0$ (wherever they are defined). Since we may decompose $[t_*, t^*] = \bigcup_{i=0}^{N-1} [t_i, t_{i+1}]$ with $t_0 = t_*$, $t_N = t^*$ and $c|_{[t_i, t_{i+1}]}$ smooth, we may write

$$c|_{[t_i,t_{i+1}]}(t) = \exp(r(t)\sigma_i)$$

with $|\sigma_i| = 1$ for each $i \in \{0, \ldots, N-1\}$. By the continuity of c however, we must have $\sigma_i = \sigma_{i+1}$ for all i, but then since $r(t^*)\sigma_N = v$, we have that $\sigma_i = \frac{v}{|v|}$ for all i. Extending r to $[0,1]$ such that $r|_{[0,t_*]} \equiv 0$ and $r|_{[t^*,1]} \equiv |v|$, we finally arrive at the equality

$$c(t) = \exp\left(\frac{r(t)}{|v|} v\right)$$

with $r(t) \geq 0$ monotone and continuous such that $r(0) = 0$ and $r(1) = v$, i.e. c is a (possibly reparametrised) geodesic. □

Corollary 17.10. If $p \in M$ and $r \leq \mathrm{inj}_p$, then $B(p,r) = \exp(B(0_p, r))$ and

$$\forall q \in B(p, \mathrm{inj}_p) \qquad d(p,q) = \rho(p,q).$$

In particular, the topology induced by d coïncides with that of M, and $(p,q) \mapsto d(p,q)^2$ is smooth in a neighbourhood of the diagonal $D = \{(p,p) \in M \times M : p \in M\}$.

Proof. It suffices to show that $B(p,r) = \exp(B(0_p,r))$ for all $r < \mathrm{inj}_p$. Note that for $q = \exp_p v \in \exp(B(0_p,r))$, we have that

$$d(p,q) \leq L(\gamma_v) = |v| < r$$

so that the inclusion $\exp(B(0_p,r)) \subset B(p,r)$ always holds. It therefore suffices to show that $B(p,r) \subset \exp(B(0_p,r))$. Suppose $B(p,r) \setminus \exp(B(0_p,r))$ is nonempty and let q be in this set. Let $c \in \Omega_{p,q}$ and set

$$t^* := \inf\{t \geq 0 : \rho(p, c(t)) = r\}.$$

This infimum is finite, for if it were not, then $\rho(p, c(0)) = 0$ and the continuity of $t \mapsto \rho(p, c(t))$ would then imply that $c(1) = q \in \exp(B(0_p, r))$, which we have ruled out. Therefore, we see as in Lemma 17.9 that

$$L(c) \geq L(c|_{[0,t^*]}) \geq \rho(p, c(t^*)) - \rho(p, c(0)) = r,$$

whence $d(p,q) \geq r$, which contradicts $d(p,q) < r$. Thus, every element $q \in B(p,r)$ is of the form $q = \exp_p v$ for $v \in B(0_p, r)$ so that by Lemma 17.9

$$d(p,q) = |v| = \rho(p,q).$$

This implies the equality of geodesic and metric distances from p in a neighbourhood of p, whence the smoothness claim follows from Lemma 13.11 and the equality of topologies now follows from the fact that the topologies of $(M, \mathcal{A})$ and (M, d) are generated by the same (small) balls. □

18. Geodesics and completeness. We now turn our attention to some finer properties of geodesics, culminating in a characterisation of completeness of the Levi-Cività connection in terms of that of d. We first prove the following.

Lemma 18.1. Suppose $K \subset M$ is compact. There exists a $\delta > 0$ with the following property: If $x \in K$ and $y \in B(x, \delta)$, then there exists a length minimising geodesic $\gamma \in \Omega_{x,y}$.

Proof. By Lemma 13.9, for each $x \in K$ there exists a $\delta_x > 0$ such that whenever $x_1, x_2 \in B(x, \delta_x)$, there exists a length-minimising geodesic in Ω_{x_1,x_2} of length less than $2\delta_x$. That this geodesic is length-minimising follows from the proof, where it may be seen that δ_x is chosen in such a way that $x_2 \in B(x_1, \mathrm{inj}_{x_1})$. Since K is compact, we may find finitely many $\{p_i\}_{i=1}^N \subset K$ and $\{\delta_i := \delta_{p_i}\}_{i=1}^N$ such that $K \subset \bigcup_{i=1}^N B(p_i, \frac{1}{2}\delta_i)$. Setting $\delta := \frac{1}{2}\min\{\delta_i\}$, we see that whenever $x \in K$, we must have $x \in B(p_i, \frac{1}{2}\delta_i)$ for some i so that if $d(x,y) < \delta$, we have $d(y, p_i) < \delta_i$, i.e. $x, y \in B(p_i, \delta_i)$ so that they are connected by a length-minimising geodesic. □

Lemma 18.2 (der Hilfssatz). If $p \in M$ and $r_0 > 0$ is such that $\exp_p$ is *defined* on $B(0_p, r_0)$, then the following hold:

- For all $r \in]0, r_0[$, $\exp(B(0_p, r)) = B(p, r)$ and $\overline{B(p, r)}$ is compact.

- Whenever $q \in B(p, r)$, there exists a length-minimising geodesic $c \in \Omega_{p,q}$.

Proof. For each $r \in]0, r_0[$, let

$$C_r := \{\exp_p(v) \in \exp_p(B(0, r)) : L(\gamma_v) = d(p, \exp_p(v))\}.$$

It suffices to show that $C_r = B(p, r)$ for all $r \in]0, r_0[$, since the inclusion $C_r \subset \exp_p(B(0, r)) \subset B(p, r)$ always holds. The compactness of $\overline{B(p, r)}$ then follows from the compactness of $\overline{C_r} \subset \exp_p(\overline{B(0, r)})$. Note also that $r_1 < r_2 \Rightarrow C_{r_1} \subset C_{r_2}$, and if $C_r = B(p, r)$ holds for some r, then $C_s = B(p, s)$ for all $s < r$.

Now, let

$$J := \{r \in]0, r_0[: C_r = B(p, r)\}.$$

By Corollary 17.10, $\left]0, \mathrm{inj}_p\right] \subset J$ so that $J \neq \varnothing$. J may be seen to be open as follows: Suppose $r \in J$ so that $\overline{C_r} = \overline{B(p, r)}$ is compact. Let $\delta = \frac{1}{2}\min\{\delta_K, r, r_0 - r\}$ with δ_K as in Lemma 18.1 and $K = \overline{B(p, r)}$ and fix a minimising sequence $\{\gamma_i\}_{i=1}^{\infty} \subset \Omega_{p,q}$ with $L(\gamma_i) < d(p, q) + \frac{1}{i}$. It suffices to show that for all $\varepsilon \in]0, \delta[$, $B(p, r + \varepsilon) \backslash B(p, r) \subset C_{r+\varepsilon}$.

Since $t \mapsto d(p, \gamma_i(t))$ is continuous for each i, we may find a $t_i \in [0, 1]$ such that $d(p, \gamma_i(t_i)) = r - \varepsilon$ by the mean value theorem. Since $\overline{B(p, r - \varepsilon)} = \overline{C}_{r-\varepsilon}$ is compact, we may replace $\{\gamma_i(t_i)\}$ with a convergent subsequence $\gamma_i(t_i) \xrightarrow{i\to\infty} y$, and there must hold $d(p, y) = r - \varepsilon$. In particular, using the minimising property of $\{\gamma_i\}$ again,

$$d(p, q) > -\frac{1}{i} + L(\gamma_i|_{[0,t_i]}) + L(\gamma_i|_{[t_i,1]}) \geq -\frac{1}{i} + d(p, \gamma_i(t_i)) + d(\gamma_i(t_i), q).$$

Taking limits as $i \longrightarrow \infty$ and using the continuity of d again, we see that

$$d(p, q) \geq d(p, y) + d(y, q).$$

The triangle inequality implies that the reverse inequality holds too, i.e.

$$d(p, q) = d(p, y) + d(y, q), \tag{3.6}$$

which implies that $d(y, q) = d(p, q) - d(p, y) < r + \varepsilon - (r - \varepsilon) = 2\varepsilon < \delta_K$. By assumption, there is a length-minimising geodesic $\gamma_{d(p,y)v} \in \Omega_{p,y}$ with $v \in B(0, 1)$ and by Lemma 18.1, we may find a length-minimising geodesic $c \in \Omega_{y,q}$. The curve

$$t \mapsto \begin{cases} \gamma_v(d(p, q)t), & t \in [0, \frac{d(p,y)}{d(p,q)}[\\ c(\frac{d(p,q)t - d(p,y)}{d(p,q) - d(p,y)}), & t \in]\frac{d(p,y)}{d(p,q)}, 1] \end{cases}$$

lies in $\Omega_{p,q}$, is of constant speed (check!) and is itself length minimising by (3.6) so that it must be a smooth geodesic. Since it has initial velocity $d(p, q)v \in B(0, r + \varepsilon)$, we deduce that $q \in C_{r+\varepsilon}$, i.e. $B(p, r + \varepsilon) \backslash B(p, r) \subset C_{r+\varepsilon}$, i.e. $C_{r+\varepsilon} = B(p, r + \varepsilon)$.

To see that J is closed, note that if $\{r_i\}_{i=1}^{\infty} \subset J$ converges to some $r \in]0, r_0[$ (assumed monotone increasing), then $q \in B(p, r)$ implies $q \in B(p, r_i) = C_{r_i}$ for sufficiently large i. But then $q \in C_r$ so that $B(p, r) \subset C_r$. □

Theorem 18.3 (Hopf-Rinow). The following conditions are equivalent:

(i) (M, d) is a complete metric space.

(ii) Γ is complete, i.e. for all $p \in M$, $\exp_p$ is defined on all of T_pM.

(iii) For *some* $p \in M$, $\exp_p$ is defined on all of T_pM.

(iv) If $A \subset M$ is closed and bounded, then it is compact.

If any of the above conditions hold, then for every $p, q \in M$, there exists a length-minimising geodesic $\gamma_v \in \Omega_{p,q}$.

Proof. By Lemma 18.2, if the second condition holds, then for any $p, q \in M$, we have that $q \in B(p, r)$ for some r so that p and q may be joined by a length-minimising geodesic. It therefore suffices to show the equivalences above.

(i)⇒(ii): Let $v \in S_pM$. It suffices to show that the maximal interval of existence of the geodesic I_v is $\mathbb{R}$. Suppose $\sup I_v := T < \infty$. Whenever $s, t \nearrow T$, we have that $d(\exp(sv), \exp(tv)) \leq |s - t| \cdot |v| \to 0$ so that for any $t_i \nearrow T$, $\{\exp(t_i v)\}$ is a Cauchy sequence so that it must converge, and this limit is independent of the sequence $t_i \nearrow T$, i.e. $\{\exp(tv) : t \in [0, T[\}$ lies in a compact set K. Since $\widetilde{K} := \pi^{-1}(K) \cap SM \subset TM$ is compact, there is a $\delta > 0$ such that the geodesic flow $\{X_t\}$ starting from any point in $\widetilde{K}$ contains δ in its interval of definition. In particular, we have that $\delta \in I_{X_{T-\frac{\delta}{2}}(v)} \Rightarrow T + \frac{\delta}{2} \in I_v$, contradicting maximality.

(ii)⇒(iii): This is trivial.

(iii)⇒(iv): If A is bounded, then for all $x \in A$ and *any* $p \in M$, we have that $d(x, p) \leq C$ for some fixed $C > 0$, i.e. $A \subset \overline{B(p, C)}$. Since the latter is compact and A is closed, we conclude that A is compact.

(iv)⇒(i): Let $\{p_i\}_{i=1}^{\infty}$ be a Cauchy sequence, i.e. for all $\varepsilon > 0$ we may find an $N > 0$ so that $i, j \geq N \Rightarrow d(p_i, p_j) < \varepsilon$. Taking $\varepsilon = 1$, we see that for all $i \geq N$, $d(p_N, p_i) \leq 1$, i.e. $\{p_i\}$ lies in a closed, bounded set. Since such a set is compact by assumption, we conclude that $\{p_i\}$ possesses a subsequence $\{p_{i_k}\}$ converging to some p in M, but then $d(p_k, p) \leq d(p_{i_k}, p) + d(p_{i_k}, p_k) \xrightarrow{k\to\infty} 0$. □

Example 18.4. All compact Riemannian manifolds are complete (cf. Exercise 15.5).

Example 18.5. Hyperbolic space (H^n, g_{hyp}) is complete.

19. Geodesic deviation. We now focus our attention on *complete* Riemannian manifolds and analyse the length-minimising nature of geodesics in more detail. We start by computing the *second variation* of the energy of a (smooth) variation of geodesics.

Theorem 19.1 (Synge's formula). Suppose $c : [a, b] \to M$ is a geodesic and $\{\gamma(s, \cdot) : [a, b] \to M\}_{s \in]-\varepsilon, \varepsilon[}$ is a *smooth* variation of c with $X := \frac{\partial \gamma}{\partial s}(0, \cdot)$ and $Y := \frac{\delta}{\delta s}\frac{\partial \gamma}{\partial s}(0, \cdot)$. Then

$$\left.\frac{\mathrm{d}^2}{\mathrm{d}s^2}\right|_{s=0} E(\gamma(s, \cdot)) = [\langle Y(t), \dot{c}(t)\rangle]_{t=a}^{t=b} + \int_a^b |\dot{X}|^2 - \langle R(X, \dot{c})\dot{c}, X\rangle \,. \tag{3.7}$$

Theorem 19.2 (Bonnet-Myers). Let (M, g) be a complete n-dimensional Riemannian manifold such that for some $K_0 > 0$ the lower Ricci curvature bound $\mathrm{Ric} \geq (n-1)K_0 g$ holds, i.e. for all $v \in TM$,

$$(\mathrm{Ric}, v \otimes v) \geq (n-1)K_0|v|^2.$$

Then for all $p, q \in M$, $d(p, q) \leq \frac{\pi}{\sqrt{K_0}}$ so that M must be compact.

Remark 19.3. Note that by Exercise 16.1, this Ricci curvature bound is satisfied if for instance $K(\sigma) \geq K_0$ for all 2-planes σ. △

Proof. Let $p, q \in M$. It suffices to show that if $c \in \Omega_{p,q}$ is any smooth geodesic with $L := |\dot{c}(0)| = L(c) > \frac{\pi}{\sqrt{K_0}}$, then it cannot be energy minimising amongst (regular) curves (and is therefore not length minimising by Lemma 17.2). Since M is complete, $d(p, q) = L(c)$ for some geodesic, whence we must have $d(p, q) < \frac{\pi}{\sqrt{K_0}}$. For this it suffices to choose an appropriate variation γ of c with $\gamma(\cdot, 0) \equiv p$, $\gamma(\cdot, 1) \equiv q$ and for which the second variation (3.7) is negative.

For each $i \in \{1, \dots, n\}$, let $\{\mu_i \in {}_c\Gamma_0^1 M\}_{i=1}^{n-1}$ be parallel unit vector fields along c such that $\langle \mu_i, \dot{c} \rangle \equiv 0$ for all i and set $\eta(t) = \sin(\pi t)$. It is not difficult to see that

$$\gamma_i(s, t) := \exp_{c(t)}(s\eta(t)\mu_i(t))$$

defines a smooth variation of c for sufficiently small s with $\gamma(s, \cdot) \in \Omega_{p,q}$ for all s and, using the notation of Theorem 19.1, $X = \eta \cdot \mu_i$, $Y(0) = 0$ and $Y(1) = 0$. Therefore, (3.7) reads

$$\left.\frac{\mathrm{d}^2}{\mathrm{d}s^2}\right|_{s=0} E(\gamma_i(s, \cdot)) = \int_0^1 \dot{\eta}^2 - \eta^2 \cdot \langle R(\mu_i, \dot{c})\dot{c}, \mu_i \rangle \,.$$

Summing over i, we obtain

$$\begin{aligned}
\sum_{i=1}^{n-1} \left.\frac{\mathrm{d}^2}{\mathrm{d}s^2}\right|_{s=0} E(\gamma_i(s, \cdot)) &= \int_0^1 (n-1)\dot{\eta}^2 - \eta^2 L^2 \cdot \underbrace{\sum_{i=1}^{n-1} \left\langle R(\mu_i, \frac{\dot{c}}{|\dot{c}|})\frac{\dot{c}}{|\dot{c}|}, \mu_i \right\rangle}_{=(\mathrm{Ric}, \frac{\dot{c}}{|\dot{c}|} \otimes \frac{\dot{c}}{|\dot{c}|})} \\
&\leq (n-1) \int_0^1 \dot{\eta}^2 - \eta^2 L^2 K_0 \\
&= (n-1) \int_0^1 \pi^2 \cos^2(\pi t) - K_0 L^2 \sin^2(\pi t) \\
&= (n-1) \left(\frac{\pi^2}{2} - \frac{K_0 L^2}{2} \right) < 0
\end{aligned}$$

if $L > \frac{\pi}{\sqrt{K_0}}$, in which case one of these $n-1$ variations gives rise to a local *maximum*. □

Bibliography

[1] V. I. Arnold, *Ordinary differential equations. Transl. and ed. by Richard A. Silverman. 1st paperback ed., 2nd printing.* 1980.

[2] L. Auslander and R. E. Mackenzie, *Introduction to differentiable manifolds. 2nd corr. ed.* 1977.

[3] G. Birkhoff and G. C. Rota, *Ordinary differential equations.* 1969.

[4] R. L. Bishop and S. I. Goldberg, *Tensor analysis on manifolds.* 1968.

[5] L. P. Eisenhart, *Riemannian geometry.* 1967.

[6] L. P. Eisenhart, *Non-Riemannian geometry. Reprint of the 1927 original.* Mineola, NY : Dover Publications, reprint of the 1927 original ed., 2005.

[7] D. Gromoll, W. Klingenberg, and W. Meyer, *Riemannsche Geometrie im Großen. 2. Aufl.*, vol. 55. Springer, Cham, 1975.

[8] E. Kreyszig, *Differential Geometry.* Dover Books on Mathematics. Dover Publications, 2013.

[9] P. Petersen, *Riemannian geometry*, vol. 171 of *Graduate Texts in Mathematics.* Springer, New York, second ed., 2006.

[10] W. A. Poor, *Differential geometric structures.* McGraw-Hill Book Company Inc., New York, 1981.

[11] J. L. Synge and A. Schild, *Tensor calculus. Repr. from corr. ed. 1969 (Univ. of Toronto Press).* 1978.

[12] O. Veblen and J. H. C. Whitehead, *The foundations of differential geometry.*, vol. 29. Cambridge University Press, Cambridge, 1932.

[13] F. W. Warner, *Foundations of differentiable manifolds and Lie groups*, vol. 94 of *Graduate Texts in Mathematics.* Springer-Verlag, New York-Berlin, 1983. Corrected reprint of the 1971 edition.

[14] J. H. C. Whitehead, "Convex regions in the geometry of paths. Addendum.," *Q. J. Math., Oxf. Ser.* **4** (1933) 226–227.