Differential Geometry 2

Fall Term 2011/12

Andreas Čap

Institut für Mathematik, Universität Wien, Nordbergstrasse 15, A–1090 Wien

E-mail address: `Andreas.Cap@esi.ac.at`

Contents

Preface v

Chapter 1. Symplectic geometry and contact geometry 1
- Initial considerations – distributions and integrability 1
- Symplectic geometry 7
- Contact geometry 15

Chapter 2. Riemannian metrics and the Levi-Civita connection 23
- The homogeneous model 24
- The orthonormal frame bundle 27
- The Levi–Civita connection 30

Chapter 3. Basic Riemannian geometry 37
- The covariant derivative 37
- Parallel transport and geodesics 48
- The Riemann curvature 59

Bibliography 65

Preface

CHAPTER 1

Symplectic geometry and contact geometry

In this section, we want to build the bridge between analysis on manifolds and differential geometry. It is also intended to complement our later discussion of Riemannian geometry. Riemannian metrics are the simplest example of a geometric structure of finite type. They have local invariants (e.g. curvature) and the group of automorphisms of such a structure is small. These facts are (at least vaguely) familiar from the case of hypersurfaces in Euclidean spaces discussed in the first course on differential geometry. The structures we consider in this chapter are the simplest examples of the opposite end of the spectrum. They do not admit local invariants, so any two structures are locally isomorphic, and there are many diffeomorphisms preserving the structure. Still they are very interesting in view of their close connections to classical mechanics and the geometric theory of partial differential equations.

Initial considerations – distributions and integrability

A simple idea how to define a geometric structure on a manifold M would be to choose some distinguished object, e.g. a vector field or a differential form.

1.1. Distinguished vector fields. We first want to show that there is not much about single vector fields, unless they have a zero. Locally around a point in which they are non–zero, all vector fields look the same up to diffeomorphism:

Proposition 1.1. *Let M be a smooth manifold and let $\xi \in \mathfrak{X}(M)$ be a vector field. If $x \in M$ is a point such that $\xi(x) \neq 0$, then there is a chart (U, u) for M with $x \in U$ such that $\xi|_U = \frac{\partial}{\partial u^1}$.*

Proof. This is an easy consequence of the existence of the flow. First we take a chart $(\tilde{U}, \tilde{u})$ with $x \in \tilde{U}$, $\tilde{u}(x) = 0$ and $\tilde{u}(\tilde{U}) = \mathbb{R}^n$. Composing $\tilde{u}$ with a linear isomorphism, we may in addition assume that $T_x\tilde{u} \cdot \xi(x) = e_1$. Thus it suffices to prove the result in the case $M = \mathbb{R}^n$, $x = 0$ and $\xi(0) = \frac{\partial}{\partial x^1}(0)$.

Now recall that the flow of ξ is defined on an open neighborhood $\mathcal{D}(\xi)$ of $\mathbb{R}^n \times \{0\}$ in $\mathbb{R}^n \times \mathbb{R}$. Hence the set $W := \{(y^1, \dots, y^n) \in \mathbb{R}^n : (0, y^2, \dots, y^n, y^1) \in \mathcal{D}(\xi)$ is open and we can define a smooth map $\varphi : W \to \mathbb{R}^n$ by

$$\varphi(y_1, \dots, y_n) := \mathrm{Fl}^{\xi}_{y_1}(0, y_2, \dots, y_n).$$

For any $y \in W$ we have

$$(*) \qquad T_y\varphi \cdot e_1 = \tfrac{d}{dt}|_{t=0} \mathrm{Fl}^{\xi}_{y_1+t}(\tilde{u}^{-1}(0, y_2, \dots, y_n)) = \tfrac{d}{dt}|_{t=0} \mathrm{Fl}^{\xi}_t(\varphi(y)) = \xi(\varphi(y)).$$

On the other hand, the definition immediately implies that $\varphi(0) = 0$ and $T_0\varphi \cdot e_i = e_i$ for all $i > 1$, so $T_0\varphi = \mathrm{id}$. Hence shrinking W, we may assume that φ is a diffeomorphism from W onto an open neighborhood U of 0 in $\mathbb{R}^n$. Putting $u = \varphi^{-1}$, we get a chart (U, u) for $\mathbb{R}^n$ around 0 and $(*)$ exactly means that $\xi|_U = \frac{\partial}{\partial u^1}$. □

If one studies vector fields, one thus only has to look at neighborhoods of a zero. Of course $\xi(x) = 0$ is equivalent to $\mathrm{Fl}^{\xi}_t(x) = x$ for all t, so this study is the study of fixed points of dynamical systems, and ideas from dynamics play an important role in this.

1.2. Distinguished 1–forms. Let us next consider the case of a distinguished one form $\alpha \in \Omega^1(M)$, which is already considerably more varied than the case of a vector field. This becomes easily visible if one reinterprets the result for vector fields in Proposition 1.1 slightly: Recalling that a chart map is just a diffeomorphism onto an open subset of $\mathbb{R}^n$, we can simply say that for $x \in M$ such that $\xi(x) \neq 0$, there is a diffeomorphism f from an open neighborhood U of x in M onto an open subset of $\mathbb{R}^n$, such that $\xi|_U = f^*\frac{\partial}{\partial x^1}$. Applying this twice we can immediately conclude that for two manifolds M, N of the same dimension, two vector fields $\xi \in \mathfrak{X}(M)$ and $\eta \in \mathfrak{X}(N)$ and points $x \in M$ such that $\xi(x) \neq 0$ and $y \in N$ with $\eta(y) \neq 0$, there is a diffeomorphism from an open neighborhood U of x in M onto an open neighborhood of y in N such that $\xi|_U = f^*\eta$. So locally around a point in which they are non–zero, any two vector fields look the same up to diffeomorphism.

The situation can not be as simple for one–forms due to the existence of the exterior derivative. Recall that for a one–form $\alpha \in \Omega^1(M)$, the exterior derivative $d\alpha \in \Omega^2(M)$ is characterized by

$$d\alpha(\xi, \eta) = \xi \cdot \alpha(\eta) - \eta \cdot \alpha(\xi) - \alpha([\xi, \eta])$$

for $\xi, \eta \in \mathfrak{X}(M)$. Naturality of the exterior derivative says that $f^*(d\alpha) = d(f^*\alpha)$. Having given two one–forms $\alpha \in \Omega^1(M)$ and $\beta \in \Omega^1(N)$ on manifolds of the same dimension and points $x \in M$ and $y \in N$ such that both $\alpha(x)$ and $\beta(y)$ are nonzero, we not only have to take into account the linear maps $\alpha(x) : T_xM \to \mathbb{R}$ and $\beta(y) : T_yN \to \mathbb{R}$ (which of course look the same up to linear isomorphism). For a locally defined diffeomorphism f with $f^*\beta = \alpha$ and $f(x) = y$ to exist, also the skew symmetric bilinear maps $d\alpha(x) : T_xM \times T_xM \to \mathbb{R}$ and $d\beta(y) : T_yN \times T_yN \to \mathbb{R}$ must be compatible via the linear isomorphism $T_xf : T_xM \to T_yM$.

Before we study skew symmetric bilinear maps in more detail, let us turn to the simplest possible case. An analog of Proposition 1.1 in the realm of one–forms should say when a one–form locally looks like $dx^1 \in \Omega^1(\mathbb{R}^n)$ up to diffeomorphism. Since $d(dx^1) = 0$, this can be possible only for *closed* one–forms, i.e. for $\alpha \in \Omega^1(M)$ such that $d\alpha = 0$. For closed forms, the analog of Proposition 1.1 actually is true. Before we show this, we prove one of the fundamental results on the exterior derivative, the Lemma of Poincaré:

Lemma 1.2 (Lemma of Poincaré). *Any closed differential form on a smooth manifold M is locally exact. More precisely, if $\omega \in \Omega^k(M)$ a smooth k–form such that $d\omega = 0$, then for each point $x \in M$, there is an open subset $U \subset M$ with $x \in U$ and a $(k-1)$–form $\varphi \in \Omega^{k-1}(U)$ such that $\omega|_U = d\varphi$.*

Proof. It suffices to deal with the case $M = \mathbb{R}^n$, since locally around any point we can find a chart which is a diffeomorphism onto $\mathbb{R}^n$.

Consider the multiplication map $\alpha : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ defined by $\alpha(t, x) =: \alpha_t(x) = tx$. Clearly $\alpha_1 = \mathrm{id}$, so $(\alpha_1)^*\omega = \omega$ and $(\alpha_0)^*\omega = 0$, and we can write ω as

$$\omega = (\alpha_1)^*\omega - (\alpha_0)^*\omega = \int_0^1 \tfrac{d}{dt}(\alpha_t)^*\omega dt,$$

where we integrate a smooth curve in the vector space $\Omega^k(\mathbb{R}^n)$. For the *Euler vector field* $E \in \mathfrak{X}(\mathbb{R}^n)$ defined by $E(x) = x$, the flow is evidently given by $\mathrm{Fl}^E_t(x) = \alpha(e^t, x)$, so we have $\alpha_t = \mathrm{Fl}^E_{\log t}$. Using this, we compute

$$\tfrac{d}{dt}\alpha_t^*\omega = \tfrac{d}{dt}(\mathrm{Fl}^E_{\log t})^*\omega = \tfrac{1}{t}(\mathrm{Fl}^E_{\log t})^*\mathcal{L}_E\omega = \tfrac{1}{t}\alpha_t^*\mathcal{L}_E\omega,$$

where $\mathcal{L}_E$ is the Lie derivative along E. Now $\mathcal{L}_E\omega = di_E\omega + i_E d\omega$ for the insertion operator i_E, and by assumption $d\omega = 0$. Since pullbacks commute with the exterior derivative, we end up with $\frac{d}{dt}\alpha_t^*\omega = d\frac{1}{t}\alpha_t^* i_E\omega$.

Next, we observe that $T_x\alpha_t = t\,\mathrm{id}$, so

$$\tfrac{1}{t}\alpha_t^* i_E\omega(x)(X_2,\dots,X_k) = \tfrac{1}{t}\omega(tx)(tx, tX_2,\dots,tX_k) = \omega(tx)(x, tX_2,\dots,tX_k),$$

so this is defined for all t and we can view it as a smooth curve of $(k-1)$–forms. But since d defines a continuous linear operator $\Omega^{k-1}(\mathbb{R}^n) \to \Omega^k(\mathbb{R}^n)$, it commutes with the integral of a smooth curve. Thus, defining $\varphi := \int_0^1 \frac{1}{t}\alpha_t^* i_E\omega dt$, we obtain

$$d\varphi = \int_0^1 d\tfrac{1}{t}\alpha_t^* i_E\omega dt = \int_0^1 \tfrac{d}{dt}\alpha_t^*\omega = \omega.$$

□

Proposition 1.2. *Let M be a smooth manifold and let $\alpha \in \Omega^1(M)$ be a closed one–form. If $x \in M$ is a point such that $\alpha(x) \neq 0$, then there is a chart (U, u) for M with $x \in U$ such that $\alpha|_U = du^1$.*

Proof. We may start with an arbitrary chart $(\tilde{U}, \tilde{u})$ such that $\tilde{u}(\tilde{U})$ is a ball in $\mathbb{R}^n$. By the lemma of Poincaré, $d\alpha = 0$ implies that there is a smooth function $f : \tilde{U} \to \mathbb{R}$ such that $\alpha|_{\tilde{U}} = df$. Since $\alpha(x) \neq 0$, we can renumber the coordinates in $\mathbb{R}^n$ in such a way that $df(x)(\frac{\partial}{\partial u^1}) \neq 0$. Considering the smooth map $\varphi : \tilde{u}(\tilde{U}) \to \mathbb{R}^n$ defined by $\varphi(y^1,\dots,y^n) = ((f \circ \tilde{u}^{-1})(y), y^2,\dots,y^n)$ we see that $\det(T_y\varphi) = \frac{\partial f}{\partial u^1}(y) \neq 0$. Hence φ is a diffeomorphism locally around $\tilde{u}(x)$ and $u := \varphi \circ \tilde{u}$ is a chart map with $f = u^1$. □

For later use, we discuss a reinterpretation of this result, using the concept of integral submanifolds.

Definition 1.2. (1) Let M be a smooth manifold of dimension n. A k–dimensional submanifold is a subset $N \subset M$ such that for each $x \in N$ there is a chart (U, u) for M with $x \in U$ such that $u(N \cap U) = u(U) \cap \mathbb{R}^k$.

(2) Let $\mathcal{I} \subset \Omega^*(M)$ be a family of differential forms on a smooth manifold M. An *integral element for* $\mathcal{I}$ at a point $x \in M$ is a linear subspace $L \subset T_xM$ such that $\alpha(x)|_L = 0$ for all $\alpha \in \mathcal{I}$.

(3) For $\mathcal{I}$ as in (2), an *integral submanifold* for $\mathcal{I}$ is a submanifold $N \subset M$ such that $\alpha|_N = 0$ for all $\alpha \in \mathcal{I}$. Otherwise put, for each $y \in N$, T_yN has to be an integral element for $\mathcal{I}$.

Remark 1.2. (1) In the special case $M = \mathbb{R}^n$, the first part of this definition is exactly the notion of a local trivialization which is used as one of the equivalent definitions of submanifolds of $\mathbb{R}^n$.

(2) If $\mathcal{I}$ consists of a single differential form α, then we talk about integral elements and integral submanifolds for α.

(3) Finding the integral elements of a family $\mathcal{I}$ is usually a question of linear algebra, since one has just to find the common kernel of a family of multilinear maps. The question whether a given integral element comes from an integral submanifold is much more subtle and interesting.

Let us return to the case of a single one–form $\alpha \in \Omega^1(M)$. If $\alpha(x) \neq 0$, then the integral elements at x are exactly the linear subspaces of the hyperplane $\mathrm{Ker}(\alpha(x)) \subset T_xM$, so in particular this hyperplane itself is the unique integral element of dimension $n-1$. Now we can rephrase Proposition 1.2 in terms of integral submanifolds:

Corollary 1.2. *Let M be a smooth manifold of dimension n and let $\alpha \in \Omega^1(M)$ be a closed one–form. Suppose that $x \in M$ is a point such that $\alpha(x) \neq 0$. Then there is a chart (U, u) in M such that $u(U) = (-\epsilon, \epsilon) \times V$ for some $\epsilon > 0$ and some open subset $V \subset \mathbb{R}^{n-1}$ such that for each $t \in (-\epsilon, \epsilon)$ the subset $u^{-1}(\{t\} \times V) \subset M$ is an integral submanifold for α of dimension $n - 1$.*

Notice that integral elements and submanifolds do not really depend on α but only on the kernels of the maps $\alpha(x)$. The one–forms β for which these kernels are the same as for α are exactly the forms $f\alpha$ for non–vanishing smooth functions f. Notice that $d(f\alpha) = df \wedge \alpha + f d\alpha$, so if α is closed we get $d\beta = \frac{1}{f} df \wedge \beta$. Conversely, if $d\beta = \frac{1}{f} df \wedge \beta$, then one immediately verifies that $\frac{1}{f}\beta$ is closed. Hence Corollary 1.2 extends to forms with this property.

1.3. Distributions and Involutivity. The considerations in 1.2 have an obvious analog in higher codimension: for a smooth, nowhere–vanishing one–form $\alpha \in \Omega^1(M)$, we can consider the family $E_x := \operatorname{Ker}(\alpha(x)) \subset T_xM$ as a smooth family of hyperplanes (i.e. linear subspaces of codimension one) in the tangent spaces of M. The corresponding notion for higher codimension is (maybe up to the notion of smoothness) easy to guess, for completeness, we also introduce the notion of involutivity and integrability:

Definition 1.3. (1) For a smooth manifold M of dimension n, a *distribution* E of rank k on M is given by a k–dimensional subspace $E_x \subset T_xM$ for each $x \in M$. (This should not be confused with distributions in the sense of generalized functions.)

(2) A (smooth) *section* of distribution $E \subset TM$ is a vector field $\xi \in \mathfrak{X}(M)$ such that $\xi(x) \in E_x$ for all $x \in M$. A *local section* of E on an open subset U is a local vector field $\xi \in \mathfrak{X}(U)$ such that $\xi(x) \in E_x$ for all $x \in U$.

(3) A *local frame* for the distribution E is a family $\{\xi_1, \dots, \xi_k\}$ of local smooth sections of E (defined on the same open subset $U \subset M$) such that for each $x \in U$ the vectors $\xi_1(x), \dots, \xi_k(x)$ are a basis for the linear subspace $E_x \subset T_xM$.

(4) The distribution $E \subset TM$ is called *smooth* if it for each point $x \in M$ there is a local frame for E defined on an open neighborhood of $x \in M$. Smooth distributions are also called *vector subbundles* of TM.

(5) A distribution $E \subset TM$ is called *involutive* if for any two local sections ξ and η of E also the Lie bracket $[\xi, \eta]$ is a section of E.

(6) The distribution $E \subset TM$ is called *integrable* if for each $x \in M$ there is a smooth submanifold $N \subset M$ which contains x such that for each $y \in N$ we have $T_yN = E_y \subset T_yM$. Such a submanifold is called an *integral submanifold* for E.

Concerning smoothness and involutivity, there are nice alternative characterizations:

Proposition 1.3. *Let M be a smooth manifold of dimension n and $E \subset TM$ a distribution of rank k.*
(1) For the distribution E the following conditions are equivalent:

- (a) *E is smooth*
- (b) *For any point $x \in M$ there is an open neighborhood U of $x \in M$ and there are one–forms $\alpha_1, \dots, \alpha_{n-k} \in \Omega^1(U)$ such that*

$$E_y = \{\xi \in T_yM : \alpha_i(\xi) = 0 \quad \forall i = 1, \dots, n-k\}.$$

- (c) *There is a family $\mathcal{I} \subset \Omega^1(M)$ of smooth one–forms such that*

$$E_x = \{\xi \in T_xM : \alpha(\xi) = 0 \quad \forall \alpha \in \mathcal{I}\}.$$

(2) If E is smooth, then following conditions are equivalent:

(a) *E is involutive*

(b) *For any point $x \in M$ there is an open neighborhood U of $x \in M$ and a local frame $\{\xi_1, \dots, \xi_k\}$ for E defined on U such that for each $i, j = 1, \dots, k$ the vector field $[\xi_i, \xi_j]$ is a section of E.*

(c) *For any point $x \in M$ there is an open neighborhood U of x in M and there are forms $\alpha_1, \dots, \alpha_{n-k} \in \Omega^1(U)$ as in part (b) of (1) such that the restriction of $d\alpha_i$ to E vanishes for all $i = 1, \dots, n-k$.*

(d) *For any differential form $\alpha \in \Omega^*(M)$ whose restriction to E vanishes, also the restriction of $d\alpha$ to E vanishes.*

Proof. (1): $(a) \Rightarrow (b)$: For $x \in M$, condition (a) implies the existence of an open neighborhood U of $x \in M$ and of local sections $\xi_1, \dots, \xi_k$ of E defined on U which form a local frame for E. Possibly shrinking U, we may assume that it is the domain of a chart (U, u) for M, and we denote by $\partial_i := \frac{\partial}{\partial u^i}$ the corresponding coordinate vector fields. Renumbering the coordinates if necessary, we may assume that putting $\xi_i := \partial_i$ for $i = k+1, \dots, n$, the vectors $\xi_1(x), \dots, \xi_n(x)$ form a basis for T_xM. Then the fields $\xi_1, \dots, \xi_n$ must be linearly independent locally around x, so possibly shrinking U once more, we may assume that $\xi_1(y), \dots, \xi_n(y)$ is a basis of T_yM for all $y \in U$.

Now for each $y \in U$ and each $i = 1, \dots, n$, we can define a linear map $\alpha_i(y) : T_yU \to \mathbb{R}$ by putting $\alpha_i(y)(\xi_i(y)) = 1$ and $\alpha_i(y)(\xi_j(y)) = 0$ for $j \neq i$. Given a smooth vector field $\xi \in \mathfrak{X}(U)$, there are uniquely determined smooth functions $f_1, \dots, f_n$ such that $\xi = \sum_{i=1}^n f_i \xi_i$, which implies that $\alpha_i(\xi) = f_i$, so $\alpha_i \in \Omega^1(U)$ for all $i = 1, \dots, n$. Evidently, the $n-k$ one–forms $\alpha_{k+1}, \dots, \alpha_n$ satisfy the conditions in (b).

$(b) \Rightarrow (a)$: This is proved very similarly: Having given $\alpha_1, \dots, \alpha_{n-k} \in \Omega^1(U)$ one may assume that U is the domain of a chart (U, u) whose coordinates are numbered in such a way that $\alpha_1(y), \dots, \alpha_{n-k}(y), du^{n-k+1}(y), \dots, du^n(y)$ form a basis for $(T_yM)^*$ for each $y \in U$. The elements of the dual basis of T_yM fit together to define vector fields $\xi_1, \dots, \xi_n$ on U, and the last k of these form a local frame for E defined on U.

$(b) \Rightarrow (c)$: Let $\mathcal{I} := \{\alpha \in \Omega^1(M) : \alpha(x)|_{E_x} = 0 \quad \forall x \in M\}$, and for $x \in M$ put $F_x := \{\xi \in T_xM : \alpha(x)(\xi) = 0 \quad \forall \alpha \in \mathcal{I}\}$. Then F_x is the intersection of the kernels of the linear maps $\alpha(x) : T_xM \to \mathbb{R}$ for $\alpha \in \mathcal{I}$ and thus a linear subspace of T_xM. By construction $E_x \subset F_x$ and it suffices to prove the converse inclusion for each x.

Given x, there is an open neighborhood U of x in M and there are $\alpha_1, \dots, \alpha_{n-k} \in \Omega^1(U)$ satisfying the conditions of (b). Choose an open neighborhood V of x in M such that $\bar{V} \subset U$ and a smooth function $f : M \to [0,1]$ which has support contained in U and is identically one on V. For each i, we can extend $f\alpha_i$ by zero to a globally defined one–form $\tilde{\alpha}_i \in \Omega^1(M)$ and by construction $\tilde{\alpha}_i \in \mathcal{I}$ for $i = 1, \dots, k$. But then for $\xi \in F_x$, we have $\tilde{\alpha}_i(x)(\xi) = \alpha_i(x)(\xi) = 0$ and hence $\xi \in E_x$.

$(c) \Rightarrow (b)$: We have given $\mathcal{I} \subset \Omega^1(M)$ and by assumption $E_x = \{\xi \in T_xM : \alpha(x)(\xi) = 0 \quad \forall \alpha \in \mathcal{I}\}$ has dimension k for all $x \in M$. Let us now fix a point $x \in M$. Then we can find forms $\alpha_1, \dots, \alpha_{n-k} \in \mathcal{I}$ such that $E_x = \{\xi \in T_xM : \alpha_i(x)(\xi) = 0 \quad \forall i = 1, \dots, n-k\}$. (If $k < n$ there must be a form α_1 such that $\alpha_1(x) \neq 0$. Then one inductively finds further forms, since until $n-k$ forms are found, E_x must be strictly smaller than the intersection of the kernels in x of the forms found so far, so there must be a form in $\mathcal{I}$ whose value in x is non–zero on this intersection.)

Since E_x has dimension k, the functionals $\alpha_1(x), \dots, \alpha_{n-k}(x)$ are linearly independent. This implies that $\alpha_1(y), \dots, \alpha_{n-k}(y)$ are linearly independent for all y in some open neighborhood U of x. Thus $\{\xi \in T_yM : \alpha_i(y)(\xi) = 0 \quad \forall i = 1, \dots, n-k\}$ is a linear subspace of T_yM which has dimension k and contains E_y by construction, so it must coincide with E_y. Hence the restrictions of the α_i to U satisfy the conditions of (b).

In part (2), the inclusions $(a) \Rightarrow (b)$ and $(d) \Rightarrow (c)$ are obvious. To prove $(b) \Rightarrow (d)$, we assume that $\alpha \in \Omega^\ell(M)$ is a differential form such that $\alpha_y(\eta_1, \dots, \eta_\ell)$ vanishes for any $y \in M$ provided that all η_i lie in $E_y \subset T_yM$. Fixing a point $x \in M$, we can take a local frame $\xi_1, \dots, \xi_k$ of E defined on an open neighborhood U of x in M which satisfies the conditions of (b). Then a vector $\eta_i \in E_x \subset T_xM$ can be written as a linear combination of the $\xi_j(x)$. It thus suffices to show that $d\alpha(\xi_{i_0}, \dots, \xi_{i_\ell})$ vanishes for any choice of $1 \leq i_0 < i_1 < \dots < i_\ell \leq k$. But this follows immediately from the formula for the exterior derivative and the fact that the bracket of any two of the ξ is again a section of E.

$(c) \Rightarrow (a)$: Let ξ and η be two smooth sections of E and consider their Lie bracket $[\xi, \eta]$. For a point $x \in M$ take an open neighborhood U of x in M and forms $\alpha_1, \dots, \alpha_{n-k} \in \Omega^1(U)$ as in (c). Then by assumption $\alpha_i(\xi) = \alpha_i(\eta) = 0$ so the formula for the exterior derivative implies that $0 = d\alpha_i(\xi, \eta) = -\alpha_i([\xi, \eta])$. Since this holds for all $i = 1, \dots, n-k$, we conclude that $[\xi, \eta](x) \in E_x \subset T_xM$. □

1.4. The Frobenius theorem. If $\alpha \in \Omega^1(M)$ is nowhere vanishing then by part (1) of Proposition 1.3, $E_x := \mathrm{Ker}(\alpha(x)) \subset T_xM$ defines a smooth distribution E of rank $n-1$. In 1.2 we assumed that $d\alpha = 0$, which by part (2) of that Proposition implies that the distribution E is involutive. Then Corollary 1.2 shows that the distribution E is actually integrable.

Indeed, integrable distributions always have to be involutive. Suppose that E is an distribution, N is an integral manifold for E, $\alpha \in \Omega^*(M)$ is a form whose restriction to E vanishes. Denoting by $i : N \to M$ the inclusion, we get $i^*\alpha = 0$ and hence $i^*d\alpha = di^*\alpha = 0$. For $x \in N$ this means that $d\alpha(x)$ vanishes when evaluated on elements of $T_xN = E_x$, so the claim follows. Generalizing Corollary 1.2, the Frobenius theorem states that involutivity implies integrability:

Theorem 1.4 (Frobenius)**.** *Let M be a smooth manifold of dimension n and let $E \subset TM$ be a smooth involutive distribution of rank k. Then for each $x \in M$, there exists a local chart (U, u) for M with $x \in U$ such that $u(U) = V \times W \subset \mathbb{R}^n$ for open subsets $V \subset \mathbb{R}^k$ and $W \subset \mathbb{R}^{n-k}$ and for each $a \in W$ the subset $u^{-1}(V \times \{a\}) \subset M$ is an integral manifold for the distribution E. In particular, any involutive distribution is integrable.*

Proof. By Proposition 1.2, we can choose a local frame $\xi_1, \dots, \xi_k$ for E defined on an open neighborhood $\tilde{U}$ of x in M, and possibly shrinking $\tilde{U}$ we may assume that it is the domain of a chart $(\tilde{U}, \tilde{u})$ for M. Putting $\partial_i := \frac{\partial}{\partial \tilde{u}^i}$, we consider the local coordinate expressions for the vector fields ξ_j. This defines smooth functions $f^i_j : \tilde{U} \to \mathbb{R}$ for $i = 1, \dots, n$ and $j = 1, \dots, k$ such that $\xi_j = \sum_i f^i_j \partial_i$. Consider the $n \times k$–matrix $(f^i_j(y))$ for $y \in \tilde{U}$. Since the vectors $\xi_i(x)$ are linearly independent, the matrix $(f^i_j(x))$ has rank k and hence has k–linearly independent rows. Renumbering the coordinates, we may assume that the first k rows are linearly independent. Then the top $k \times k$–submatrix has nonzero determinant in x. By continuity, this is true locally around x, and possibly shrinking $\tilde{U}$, we may assume that it holds on all of $\tilde{U}$.

For $y \in \tilde{U}$ let $(g^i_j(y))$ the $k \times k$–matrix which is inverse to the first k rows of $(f^i_j(y))$. Since the inversion in $GL(k, \mathbb{R})$ is smooth, each of the g^i_j defines a smooth function on V. For $i = 1, \dots, k$ we now put $\eta_i := \sum_j g^j_i \xi_j$. These are local smooth sections of E and since the matrix $(g^i_j(y))$ is always invertible, their values span E_y for each $y \in \tilde{U}$.

Expanding η_i in the basis ∂_ℓ we obtain

$$(*) \qquad \eta_i = \sum_j g_i^j \xi_j = \sum_{j,\ell} g_i^j f_j^\ell \partial_\ell = \partial_i + \sum_{\ell > k} h_i^\ell \partial_\ell,$$

for certain smooth functions h_i^ℓ on $\tilde{U}$.

We claim that the sections η_i of E satisfy $[\eta_i, \eta_j] = 0$ for all i, j. Since the distribution E is involutive, we know that $[\eta_i, \eta_j]$ is a section of E, so there must be smooth functions c_{ij}^ℓ such that $[\eta_i, \eta_j] = \sum_\ell c_{ij}^\ell \eta_\ell$. Applying equation $(*)$ to the right hand side, we see that $[\eta_i, \eta_j]$ is the sum of $\sum_{\ell=1}^k c_{ij}^\ell \partial_\ell$ and some linear combination of the ∂_ℓ for $\ell > k$. On the other hand, inserting $(*)$ for η_i and η_j, we see that that $[\eta_i, \eta_j]$ must be a linear combination of the ∂_ℓ for $\ell > k$ only. This is only possible if all c_{ij}^ℓ vanish identically.

The vectors $T_x\tilde{u} \cdot \eta_1(x), \dots, T_x\tilde{u} \cdot \eta_k(x)$ generate a k–dimensional subspace in $\mathbb{R}^n$ and changing $\tilde{u}$ by some linear isomorphism, we may assume that this is $\mathbb{R}^k \subset \mathbb{R}^n$. Now we can find open neighborhoods V in $\mathbb{R}^k$ and $W \in \mathbb{R}^{n-k}$ of zero such that

$$\varphi(t^1, \dots, t^k, a) := (\mathrm{Fl}^{\eta_1}_{t^1} \circ \dots \circ \mathrm{Fl}^{\eta_k}_{t^k})(\tilde{u}^{-1}(a))$$

makes sense for all $t = (t^1, \dots, t^k) \in V$ and all $a \in W$. Since the vector fields η_i have pairwise vanishing Lie bracket, their flows commute, so it makes no difference in which order we apply the flows. For sufficiently small s, we can use $\mathrm{Fl}^{\eta_i}_{t^i+s} = \mathrm{Fl}^{\eta_i}_s \circ \mathrm{Fl}^{\eta_i}_{t^i}$ and this commuting property to write

$$\varphi(t^1, \dots, t^i + s, \dots, t^k, a) = \mathrm{Fl}^{\eta_i}_s(\varphi(t^1, \dots, t^i, \dots, t^k, a))$$

for $i = 1, \dots, k$ and differentiating this with respect to s at $s = 0$, we conclude that $\frac{\partial \varphi}{\partial t^i}(t, a) = \eta_i(\varphi(t, a))$ for $i = 1, \dots, k$. On the other hand, $\varphi(0, a) = \tilde{u}^{-1}(a)$, which by construction easily implies that $T_{(0,0)}\varphi$ is invertible. Possibly shrinking V and W, we may assume that φ is a diffeomorphism onto an open neighborhood U of x in M. But then we have already seen above that $\varphi^{-1}(V \times \{a\})$ is an integral submanifold, so putting $u := \varphi^{-1}$ we obtain a chart with the required properties. □

Symplectic geometry

1.5. Skew symmetric bilinear forms. After this detour on integrability of distributions, let us take up the questions on distinguished one–forms form 1.2. As indicated there, the first step is to understand skew symmetric bilinear maps on real vector spaces in a basis–independent way.

Let $b : V \times V \to \mathbb{R}$ be a bilinear form on a finite dimensional real vector space V such that $b(w, v) = -b(v, w)$. Let us first define the *nullspace* of b by $\mathcal{N}(b) := \{v \in V : b(v, w) = 0 \quad \forall w \in V\}$. This is evidently a linear subspace of V, and the number $\dim(V) - \dim(\mathcal{N}(b))$ is called the *rank* of b. The form b is called *non–degenerate* if $\mathcal{N}(b) = \{0\}$.

More generally, given a linear subspace $W \subset V$, we define $W^\circ \subset V$ by $W^\circ := \{v \in V : b(v, w) = 0 \quad \forall w \in W\}$, so $\mathcal{N}(b) = V^\circ$. Notice that the bilinear form b can be evidently restricted to any linear subspace $W \subset V$.

Lemma 1.5. *Let $b : V \times V \to \mathbb{R}$ be a non–degenerate skew symmetric bilinear form on a finite dimensional real vector space. Then for a linear subspace $W \subset V$ the following are equivalent:*

1. *The restriction of b to W is non–degenerate.*
2. *The restriction of b to W° is non–degenerate.*
3. $W \cap W^\circ = \{0\}$

(4) $V = W \oplus W^\circ$

Proof. The bilinear form b gives rise to a linear map $V \to V^*$, which sends $v \in V$ to the linear map $v^\# : V \to \mathbb{R}$ defined by $v^\#(w) := b(v, w)$. The fact that b is non–degenerate is equivalent to the fact that this map is injective, and thus a linear isomorphism since V is finite dimensional. Now by definition, $W^\circ \subset V$ is the annihilator of the image of W in V^*. Consequently, $\dim(W^\circ) = \dim(V) - \dim(W)$, which immediately implies that (3) and (4) are equivalent. Now the nullspace of $b|_W$ by definition is $W \cap W^\circ$. This immediately implies that (1) and (3) are equivalent. Finally, the above observation on dimensions shows that $\dim(W) = \dim((W^\circ)^\circ)$. Since we evidently have $W \subset (W^\circ)^\circ$ the two spaces have to coincide. Together with the above, this implies the equivalence of (2) and (3). $\square$

Proposition 1.5. *Let V be an n–dimensional real vector space and $b : V \times V \to \mathbb{R}$ a skew–symmetric bilinear form. Then the rank of b is an even number. If this rank is $2k$, then there is a basis $\{v_1, \dots, v_k, w_1, \dots, w_k, u_1, \dots, u_{n-2k}\}$ for V such that $b(v_i, w_i) = -b(w_i, v_i) = 1$ for $i = 1, \dots, k$ while b evaluates to zero on all other combination of basis elements.*

Proof. Let us first proof the result in the case the b is non–degenerate. Choose a non–zero vector $v_1 \in V$. By non–degeneracy, there is a vector $w_1 \in V$ such that $b(v_1, w_1) \neq 0$ and replacing w_1 by a multiple, we can assume that $b(v_1, w_1) = 1$. Let $W \subset V$ be the linear subspace spanned by v_1 and w_1. Since $b(v_1, v_1) = 0$ by skew symmetry, we conclude that $b(v_1, w_1) = 1$ implies that v_1 and w_1 are linearly independent, so W has dimension 2. Any vector $w \in W$ can be written as $\lambda v_1 + \mu w_1$ for $\lambda, \mu \in \mathbb{R}$, and clearly $b(w, v_1) = -\mu$ and $b(w, w_1) = \lambda$. For non–zero w, one of these numbers is nonzero, so $b|_W$ is non–degenerate. Thus $V = W \oplus W^\circ$ and $b|_{W^\circ}$ is non–degenerate by the lemma.

Now we can apply the same construction to W° to find elements v_2 and w_2 such that $b(v_2, w_2) = 1$, and continue this recursively. If $\dim(V)$ is even, then we end up with a basis $\{v_1, \dots, v_k, w_1, \dots w_k\}$ with the required properties. If $\dim(V)$ is odd, we end up with a one–dimensional space endowed with a non–degenerate skew symmetric bilinear form. This is a contradiction, since by skew symmetry any such form has to be zero. This completes the proof in the non–degenerate case.

In the general case, consider the quotient space $\underline{V} := V/\mathcal{N}(b)$. A moment of thought shows that b induces a skew symmetric bilinear form $\underline{b}$ on $\underline{V}$, defined by

$$\underline{b}(v + \mathcal{N}(b), w + \mathcal{N}(b)) := b(v, w).$$

By definition, $v + \mathcal{N}(b)$ lies in the nullspace of $\underline{b}$ if and only if $v \in \mathcal{N}(b)$, so $\underline{b}$ is non–degenerate. From above, we conclude that $\dim(\underline{V})$ (which equals the rank of b) is even, say equal to $2k$, and there is a basis $\{\underline{v}_1, \dots, \underline{v}_k, \underline{w}_1, \dots, \underline{w}_k\}$ such that $\underline{b}(\underline{v}_i, \underline{w}_i) = -\underline{b}(\underline{w}_i, \underline{v}_i) = 1$ for $i = 1, \dots, k$ while all other combinations evaluate to zero under $\underline{b}$. Now choose preimages $v_i, w_i \in V$ of these basis elements and a basis $\{u_1, \dots, u_{n-2k}\}$ for $\mathcal{N}(b)$. Then by construction, these elements have the required behavior with respect to b, and it is an easy exercise to verify that they form a basis for V. $\square$

In particular, we see that non–degenerate skew symmetric bilinear forms exist only on vector spaces of even dimension, but in these dimension they are uniquely determined up to isomorphism. A vector space V endowed with a non–degenerate skew symmetric bilinear form $b : V \times V \to \mathbb{R}$ is called a *symplectic vector space*. For a subspace $W \subset V$ we have the *annihilator* or *symplectic orthogonal* space W^0 from above and by non–degeneracy $\dim(W^0) = \dim(V) - \dim(W)$. Now a subspace $W \subset V$ is called *isotropic* if $W \subset W^0$, *coisotropic* if $W^0 \subset W$ and *Lagrangean* if $W^0 = W$. From this definition, it follows immediately that if $\dim(V) = 2n$, then $\dim(W)$ has to be $\leq n$ if W is isotropic, $\geq n$ is W is coisotropic and $= n$ if W is Lagrangean. In particular, Lagrangean subspaces are maximal isotropic subspaces and minimal coisotropic subspaces.

As an application of our results, we can now give a nice characterization of non–degeneracy of a skew symmetric bilinear form. In the first part of the course, we have met the wedge–product of differential forms, which was constructed in a point wise manner. Thus it makes sense for skew symmetric multilinear maps defined on a vector space. If $\varphi : V^k \to \mathbb{R}$ and $\psi : V^\ell \to \mathbb{R}$ are alternating maps, which are k–linear respectively ℓ–linear, then one defines $\varphi \wedge \psi : V^{k+\ell} \to \mathbb{R}$ by

$$(\varphi \wedge \psi)(v_1, \dots, v_{k+\ell}) := \frac{1}{k!\ell!} \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma)\varphi(v_{\sigma_1}, \dots, v_{\sigma_k})\psi(v_{\sigma_{k+1}}, \dots, v_{\sigma_{k+\ell}}).$$

This is $(k + \ell)$–linear and alternating. Moreover, the wedge product is associative and *graded commutative*, i.e. $\psi \wedge \varphi = (-1)^{k\ell}\varphi \wedge \psi$. In particular, this shows that wedging a linear functional (or a one–form) with itself, one always gets zero. But in the case of even degree, the wedge–product with itself can be non–zero. Hence one can form the square and also higher powers of a map with respect to the wedge product. Indeed, we get

Corollary 1.5. *Let V be a real vector space of dimension $2n$. Then a skew symmetric bilinear map $b : V \times V \to \mathbb{R}$ is non–degenerate if and only if $b^n = b \wedge \dots \wedge b$ (n factors) is non–zero and hence a volume–form on V.*

Proof. It follows from linear algebra that there is just one $2n$–linear, alternating map $V^{2n} \to \mathbb{R}$ up to multiples. Identifying V with $\mathbb{R}^{2n}$ by choosing a basis, the map b^n must thus correspond to some multiple of the usual determinant.

If $0 = b(v, w)$ for some fixed $0 \neq v \in V$ and all $w \in W$, then it follows immediately from the definition that $b^n(v, w_1, \dots, w_{2n-1}) = 0$ for all $w_i \in V$. In particular, we can choose the w_i in such a way that $\{v, w_1, \dots, w_{2n-1}\}$ is a basis of V. Using this basis to identify V with $\mathbb{R}^{2n}$ we conclude that $b^n = 0$ if b is degenerate.

Assuming conversely that b is non–degenerate, we take a basis $\{v_1, \dots, v_n, w_1, \dots, w_n\}$ for V as in Proposition 1.5. Denoting the dual basis by $\{v^1, \dots, v^n, w^1, \dots, w^n\}$, we can clearly write $b = \sum_{i=1}^n v^i \wedge w^i$. Associativity of the wedge product then implies that $b^2 = -2\sum_{i_1<i_2} v^{i_1} \wedge v^{i_2} \wedge w^{i_1} \wedge w^{i_2}$. Continuing inductively, one easily sees that b^n is a non–zero multiple of $v^1 \wedge \dots \wedge v^n \wedge w^1 \wedge \dots \wedge w^n$. □

1.6. Symplectic forms. Let us return to the question of distinguished one–forms from 1.2 and hence look at a manifold M of dimension n and $\alpha \in \Omega^1(M)$. Then the basic observation in 1.2 was that for $x \in M$ we not only have to consider the linear functional $\alpha(x) : T_xM \to \mathbb{R}$ but also the skew symmetric bilinear map $d\alpha(x) : T_xM \times T_xM \to \mathbb{R}$. From Proposition 1.5 we know that the basic invariant of $d\alpha(x)$ is its rank, which has to be an even number. In studying distinguished one–forms it is natural to assume that this rank is the same in all points of M. In 1.2 we have sorted out the case that the rank is zero, so now we want to look at cases in which the rank is as large as possible. There are two basic possible situations, according to whether $\dim(M)$ is even or odd. We'll start with the even dimensional case here, and take up the case of odd dimensions in 1.10 below.

If $\dim(M)$ is even, say $\dim(M) = 2n$, then we can simply require that $d\alpha(x)$ is non–degenerate for all $x \in M$. As discussed in 1.5 we can form $(d\alpha)^h = d\alpha \wedge \dots \wedge d\alpha \in \Omega^{2n}$, and in view of Corollary 1.5 the non–degeneracy condition is equivalent to requiring that $(d\alpha)^n$ is nowhere vanishing and thus a volume form on M. In particular, we see that α gives rise to an orientation on M so in particular M must be orientable for such a form to exist.

We can easily deduce a more severe restriction from this, which is the reason, why one usually uses a slightly more general concept. Namely, we can also consider the form $\beta := \alpha \wedge (d\alpha)^{n-1} \in \Omega^{2n-1}(M)$. Now from the basic properties of the exterior derivative, we immediately conclude that $0 = d(d\alpha)$ and hence $0 = d((d\alpha)^{n-1})$ and hence $(d\alpha)^n = d\beta$. Thus the volume form $(d\alpha)^n$ is always exact. But on a compact manifold, the integral of a volume form must be non–zero while the integral over an exact form must be zero by Stokes' theorem. We can bypass this problem by considering general closed two–forms instead of only the exact forms $d\alpha$. This leads to

Definition 1.6. Let M be a smooth manifold of even dimension $2n$. A *symplectic form* on M is a two–form $\omega \in \Omega^2(M)$ such that $d\omega = 0$ and $\omega(x) : T_xM \times T_xM \to \mathbb{R}$ is non–degenerate for any $x \in M$. The pair (M, ω) is then called a *symplectic manifold*.

If (M, ω) and $(\tilde{M}, \tilde{\omega})$ are symplectic manifolds, then a *symplectomorphism* between M and $\tilde{M}$ is a diffeomorphism $f : M \to \tilde{M}$ such that $f^*\tilde{\omega} = \omega$.

As above, the non–degeneracy condition is equivalent to requiring that ω^n is a volume form on M. Moreover, compatibility of the pullback with the wedge product immediately implies that any symplectomorphism is a volume preserving diffeomorphism, i.e. compatible with the volume forms.

1.7. A fundamental example. The fundamental example of a symplectic form actually is exact. To construct it, we have to recall the definition of the cotangent bundle. So let N be an arbitrary smooth manifold of dimension n. Recall that for $x \in N$ the cotangent space of N at x is the space $(T_xN)^*$ of all linear functionals $T_xN \to \mathbb{R}$. The *cotangent bundle* of N is the union of all cotangent spaces $T^*N := \cup_{x\in N}(T_xN)^*$. There is an obvious projection $p : T^*N \to N$ which sends $(T_xN)^*$ to x. It is easy to see that there is a unique topology on T^*N which makes this projection continuous and induces the standard vector spaces topology on each of the spaces $(T_xN)^*$. Then T^*N can be made into a smooth manifold of dimension $2n$ such that the projection $p : T^*N \to N$ is smooth.

Indeed, let (U, u) be a chart on N. Then we consider the open subset $T^*U := p^{-1}(U) \subset T^*N$ and the map $T^*U \to u(U) \times \mathbb{R}^{n*}$ mapping $\varphi \in (T_xN)^*$ to $(u(x), \psi \circ (T_xu)^{-1})$. One immediately verifies that this gives rise to an atlas on T^*N with smooth chart changes.

The crucial point for our purposes is that there is a *tautological one–form* on $M := T^*N$. A point $\varphi \in M$ is a linear functional $\varphi : T_xN \to \mathbb{R}$, where $x = p(\varphi) \in N$. Having given a tangent vector $\xi \in T_\varphi M$, we can form $T_\varphi p \cdot \xi \in T_{p(\varphi)}N$ and then define $\alpha(\xi) := \varphi(Tp \cdot \xi)$. Otherwise put, we have $\alpha(\varphi) = \varphi \circ T_\varphi p$, and this evidently is a linear map $T_\varphi M \to \mathbb{R}$.

The natural construction already suggests that α should be smooth. We can easily verify this directly in local coordinates, and we use the classical notion common in physics at this point. Let us start with a chart (U, q) for N, so we have local coordinates $q^1, \dots, q^n : U \to \mathbb{R}$ on the open subset $U \subset N$. Then the induced chart on T^*U has values in $u(U) \times \mathbb{R}^{n*}$ and we denote the coordinates corresponding to the second factor by $p_1, \dots, p_n$. (In physics terminology, N is the *configuration space*, T^*N the *phase space*, and the cotangent vectors are interpreted as momenta.)

Now the coordinate forms $dq^1, \dots dq^n$ form a basis for each cotangent space of N. We have defined the induced chart on T^*N in such a way that the coordinates

$p_1(\varphi), \dots, p_n(\varphi)$ of $\varphi : T_xN \to \mathbb{R}$ are exactly the coefficients one obtains when expanding φ in the basis $\{dq^1(x), \dots, dq^n(x)\}$. On the other hand, in our coordinates the projection $p : T^*N \to N$ is simply given by $(q^1, \dots, q^n, p_1, \dots, p_n) \mapsto (q^1, \dots, q^n)$. Now having given $\xi \in T_\varphi M$, we can expand it as

$$\xi^1 \tfrac{\partial}{\partial q^1} + \dots + \xi^n \tfrac{\partial}{\partial q^n} + \xi^{n+1} \tfrac{\partial}{\partial p_1} + \xi^{2n} \tfrac{\partial}{\partial p_n}$$

and then $T_\varphi p \cdot \xi = \xi^1 \frac{\partial}{\partial q^1} + \dots + \xi^n \frac{\partial}{\partial q^n}$ and $\alpha(\xi) = p_1(\varphi)\xi^1 + \dots + p_n(\varphi)\xi^n$.

This shows that $\alpha = \sum_{i=1}^n p_i dq^i$ is smooth locally and hence defines a one–form on $M = T^*N$. We can also immediately read off its exterior derivative, namely $d\alpha = \sum_{i=1}^n dp_i \wedge dq^i$. This shows that $d\alpha$ is a symplectic form on $M = T^*N$ but it is more common to use the negative and put $\omega := -d\alpha$. Moreover, over any subset of the form $p^{-1}(U)$ for some chart (U, q) of N with corresponding coordinates (q^i, p_i), the corresponding coordinate vector fields form in each point a basis of the form described in Proposition 1.5. This proves the first part of

Proposition 1.7. *Let N be any smooth n–dimensional manifold, put $M := T^*N$, let $\alpha \in \Omega^1(N)$ be the tautological one–form and $\omega = -d\alpha$. Then*

(1) ω is a symplectic form on M.

(2) For any point $x \in N$, the cotangent space $(T_xN)^* = p^{-1}(x) \subset M$ is a smooth submanifold of M whose tangent space in in each point is a Lagrangean subspace.

(3) Let $\psi \in \Omega^1(N)$ be a one–form and view ψ as a smooth function $N \to T^*N$. Then $\psi(N) \subset T^*N$ is a smooth submanifold and if ψ is closed, then each tangent space to $\psi(N)$ is a Lagrangean subspace.

(4) For any diffeomorphism $f : N \to N$ the induced diffeomorphism $T^*f : T^*N \to T^*N$, which maps $\varphi \in (T_xN)^*$ to $\varphi \circ (T_xf)^{-1} \in (T_{f(x)}N)^*$ is a symplectomorphism.

Proof. We have already seen (1). For (2) we take a chart (U, q) for N with $x \in U$ and the corresponding chart for M. Then $(T_xN)^*$ corresponds to the subspace $\{q(x)\} \times \mathbb{R}^{n*}$, so this is a global submanifold chart. Moreover, the tangent space to $(T_xN)^*$ in a point $\varphi \in (T_xN)^*$ is evidently spanned by the tangent vectors $\frac{\partial}{\partial p_i}(\varphi)$ and thus Lagrangean.

(3) Choosing a local chart (U, q) of N and considering the induced chart of M, the map $\psi|_U : U \to p^{-1}(U)$ corresponds to a map $u(U) \to u(U) \times \mathbb{R}^{n*}$ whose first component is the identity. But then $\psi(N)$ simply corresponds to the graph of the second component function. This graph is a smooth submanifold in $u(U) \times \mathbb{R}^{n*}$, and composing a submanifold chart for this with the diffeomorphism $p^{-1}(U) \to u(U) \times \mathbb{R}^{n*}$ provided by the induced chart, we obtain a submanifold chart for $\psi(N)$.

Having observed that $\psi(N) \subset T^*N$ is a submanifold, it is clear that $\psi : N \to \psi(N)$ is a diffeomorphism, since the restriction $p|_{\psi(N)} : \psi(N) \to N$ provides a smooth inverse. The tangent spaces of $\psi(N)$ are Lagrangean if and only if $\omega|_{\psi(N)} = 0$, which is equivalent to $0 = \psi^*\omega \in \Omega^1(N)$. Since $\omega = -d\alpha$, we get $\psi^*\omega = -d\psi^*\alpha$. Now we claim that $\psi^*\alpha = \psi \in \Omega^1(M)$ which shows that $\psi^*\omega = -d\psi$, which vanishes if ψ is closed.

To see this, observe that $\psi^*\alpha(x)(\xi) = \alpha(\psi(x))(T_x\psi \cdot \xi)$. Now $p \circ \psi = \mathrm{id}_N$, whence $T_{\psi(x)}p \cdot T_x\psi \cdot \xi = \xi$ which implies that $\psi^*\alpha(x)(\xi) = \psi(x)(\xi)$ and hence the claim.

(4) We have actually almost proved this already. Namely, let (U, q) be any chart for N and put $\tilde{U} := f^{-1}(U)$, $\tilde{q} := q \circ f$. Then $(\tilde{U}, \tilde{u})$ is also a chart for N. If we denote the induced charts for M by $(p^{-1}(U), \Phi)$ and $(p^{-1}(\tilde{U}), \tilde{\Phi})$, then it immediately follows from our constructions that $\tilde{\Phi} = \Phi \circ (T^*f|_{p^{-1}(\tilde{U})})$. This means that in our chosen charts, T^*f simply maps $(\tilde{q}^1, \dots, \tilde{q}^n, \tilde{p}_1, \dots, \tilde{p}_n)$ to $(q^1, \dots, q^n, p_1, \dots, p_n)$ and thus clearly pulls back

$\sum_i dq^i \wedge dp_i$ to $\sum d\tilde{q}^i \wedge d\tilde{p}_i$. But these are exactly the coordinate expressions of ω with respect to the two charts. □

Submanifolds of a symplectic manifold, whose tangent space in each point is a Lagrangean subspace are called *Lagrangean submanifolds*. Thus our proposition in particular implies that the symplectic manifold T^*N has many Lagrangean submanifolds and many symplectomorphisms.

1.8. The Darboux theorem. Looking at 1.7, one might expect that cotangent bundles form very special examples of symplectic manifolds. Initially, one could not expect that the nice choice of basis of a tangent space from Proposition 1.5 can be locally realized by coordinate vector fields. (Indeed, if one has a hypersurface in Euclidean space and one can choose local coordinates such that the coordinate vector fields are orthonormal in each point, then this chart defines an isometry to Euclidean space. In particular, for surfaces in $\mathbb{R}^3$ this implies vanishing of the Gauss curvature.) It turns out however, that locally any symplectic manifold looks like a cotangent bundle, and this is the content of the Darboux theorem. In particular this implies that there is no local symplectic geometry and that locally any symplectic manifold admits many Lagrangean submanifolds and many symplectomorphisms.

Before we can prove the Darboux theorem, we need some background on time–dependent vector fields and the associated evolution operators, which are the analogs of flows for ordinary vector fields.

Definition 1.8. (1) A *time–dependent vector field* on a smooth manifold M is a smooth map $\xi : J \times M \to TM$, where $J \subset \mathbb{R}$ is an open interval, such that $p \circ \xi = \mathrm{pr}_2$, where $p : TM \to M$ is the canonical projection and $\mathrm{pr}_2 : J \times M \to M$ is the projection onto the second factor. For $t \in J$ and $x \in M$ we will often write $\xi_t(x)$ for $\xi(t, x)$ and denote the time dependent vector field by $(\xi_t)_{t\in J}$.

(2) An *integral curve* for a time–dependent vector field $(\xi_t)_{t\in J}$ is a smooth curve $c : I \to M$ defined on a sub–interval $I \subset J$ such that $c'(t) = \xi(t, c(t))$ for all $t \in I$.

The study of time–dependent vector fields can be easily reduced to the study of ordinary vector fields. To do this, one associates to a time dependent vector field $(\xi_t)_{t\in J}$ a vector field $\tilde{\xi} \in \mathfrak{X}(J \times M)$ as follows. Recall that $T(J \times M)$ can be naturally identified with $TJ \times TM$ via the tangent maps of the two projections. Moreover, on J there is a canonical vector field $\frac{\partial}{\partial s}$, since $J \subset \mathbb{R}$ is an open interval. So it is clear that $\tilde{\xi}(t, x) := (\frac{\partial}{\partial s}(t), \xi_t(x))$ defines a smooth vector field on $J \times M$, and it is easy to relate integral curves of ξ to integral curves of $\tilde{\xi}$. The standard theory of flows gives us a an open neighborhood $\mathcal{D}(\tilde{\xi})$ of $\{0\} \times J \times M$ in $\mathbb{R} \times J \times M$ and a smooth map $\mathrm{Fl}^{\tilde{\xi}} : \mathcal{D}(\tilde{\xi}) \to J \times M$ such that $s \mapsto \mathrm{Fl}^{\tilde{\xi}}_s(t, x)$ is the maximal integral curve of $\tilde{\xi}$ starting at (t, x).

Now we define $\mathcal{D}(\xi) := \{(s, t, x) \in J \times J \times M : (s-t, t, x) \in \mathcal{D}(\tilde{\xi})\}$ and the *evolution operator* $\Phi^\xi : \mathcal{D}(\xi) \to M$ by $\Phi^\xi(s, t, x) = \mathrm{pr}_2(\mathrm{Fl}^{\tilde{\xi}}_{s-t}(t, x))$. Clearly, $\mathcal{D}(\xi)$ is an open neighborhood of $\Delta_J \times M \subset J \times J \times M$, where $\Delta_J := \{(t, t) : t \in J\}$ is the diagonal, and Φ^ξ is a smooth map. Moreover, it follows immediately that $s \mapsto \Phi^\xi_{s,t} := \Phi^\xi(s, t, x)$ is a maximal integral curve for the time–dependent vector field mapping t to (t, x). From the flow property one immediately deduces that $\Phi^\xi_{s,t} = \Phi^\xi_{s,r} \circ \Phi^\xi_{r,t}$ whenever both sides are defined. Similarly to the case of flows one can also deduce existence of one side from existence of the other under some conditions.

Using this, we now prove:

Theorem 1.8 (Darboux). *Let (M, ω) be a symplectic manifold. Then for any point $x \in M$ there is a local chart (U, u) around x with corresponding local coordinates $\{q^1, \dots, q^n, p_1, \dots, p_n\}$ such that $\omega|_U = \sum_{i=1}^n dq^i \wedge dp_i$. In particular, one can view the chart map $u : U \to u(U)$ as a symplectomorphism to an open subset of $T^*\mathbb{R}^n$.*

Sketch of proof. We prove the statement by the so called Moser trick, making use (without proof) of the fact the some identities for vector fields and differential forms carry over to the time–dependent case, see [**Mi**, section 31.11] for proofs. First choose any chart $(\tilde{U}, \tilde{u})$ with $x \in \tilde{U}$, $\tilde{u}(x) = 0$ and $\tilde{u}(\tilde{U}) = \mathbb{R}^{2n}$. Pulling back ω to $\mathbb{R}^{2n}$ via $\tilde{u}^{-1}$ we obtain a symplectic form on $\mathbb{R}^{2n}$. The value of this at 0 is a non–degenerate bilinear form on $T_0^*\mathbb{R}^{2n} = \mathbb{R}^{2n}$. If necessary doing a linear coordinate change, we may assume by Proposition 1.5 that denoting the coordinates on $\mathbb{R}^{2n}$ by $q^1, \dots, q^n, p_1, \dots, p_n$, this bilinear form is given by $\sum_{i=1}^n dq^i(0) \wedge dp_i(0)$.

Now we denote the pullback of ω by ω_0 and put $\omega_1 := \sum_{i=1}^n dq^i \wedge dp_i$. So ω_0 and ω_1 are symplectic forms on $\mathbb{R}^{2n}$ which agree in $0 \in \mathbb{R}^n$, and to complete the proof it suffices to show that these two forms are locally symplectomorphic. Now we define $\omega_t := \omega_0 + t(\omega_1 - \omega_0)$ for $t \in J := (-\epsilon, 1 + \epsilon)$ for some small $\epsilon > 0$. We can view $\omega_i(y)$ as a linear isomorphism $\mathbb{R}^{2n} \to \mathbb{R}^{2n*}$ for all $y \in \mathbb{R}^{2n}$ for $i = 1, 2$. Locally around 0, the linear maps $\omega_1(y) - \omega_0(y)$ have small norm, so $\omega_t(y)$ is a linear isomorphism for all t. Restricting to an appropriate open neighborhood V of 0, we may thus assume that ω_t is a symplectic form on V for all $t \in J$. Of course, we have $d(\omega_1 - \omega_0) = 0$. By the Poincaré Lemma 1.2, we can choose V in such a way that this implies $\omega_1 - \omega_0 = d\psi$ for some $\psi \in \Omega^1(V)$ and subtracting an appropriate constant (in the coordinates on V), we may assume that $\psi(0) = 0$.

Now Moser's trick is to construct a time–dependent vector field η_t whose evolution operator induces a family f_t of diffeomorphisms fixing 0 such that $f_0 = \mathrm{id}$ and $(f_t)^*\omega_t = \omega_0$ for all t. Having found this, f_1 will solve our problem.

Now for $y \in V$, one has $\psi(y) \in T_y^*V$ and, for each t, the non–degenerate bilinear form $\omega_t(y)$ on T_yV. Hence there exists a unique element $\eta_t(y) \in T_yV$ such that $-\psi(y) = \omega_t(y)(\eta_t(y), \_)$. Clearly, this defines a smooth time dependent vector field $(\eta_t)_{t \in J}$ on V, and by definition (with an obvious meaning for insertion operators), we have $i_{\eta_t}\omega_t = -\psi$ for all t. Now from the properties of the flow it follows that we can shrink V further in such a way that the evolution operator Φ_t^η is defined and $f_t(x) := \Phi_{t,0}^\eta(x)$ is a diffeomorphism for all $t \in [0, 1]$. Further $\Phi_{t+s,0}^\eta = \Phi_{t+s,t}^\eta \circ \Phi_{t,0}^\eta$ which implies that $(f_{t+s})^*\omega_{t+s} = (f_t)^*(\Phi_{t+s,t}^\eta)^*\omega_{t+s}$. Differentiating this with respect to s at $s = 0$, we obtain

$$\tfrac{\partial}{\partial t}(f_t)^*\omega_t = (f_t)^*(\mathcal{L}_{\eta_t}\omega_t + \tfrac{\partial}{\partial t}\omega_t),$$

where $\mathcal{L}$ denotes the Lie derivative. Expanding this as $di_{\eta_t}\omega_t + i_{\eta_t}d\omega_t = -d\psi + 0$ and using $\frac{\partial}{\partial t}\omega_t = \omega_1 - \omega_0 = d\psi$, we see that $\frac{\partial}{\partial t}(f_t)^*\omega_t = 0$, so this is constant and equal to $(f_0)^*\omega_0 = \omega_0$. □

1.9. Hamiltonian vector fields and classical mechanics. Let (M, ω) be a symplectic manifold. Then for each point $x \in M$, the value $\omega(x) : T_xM \times T_xM \to \mathbb{R}$ is a non–degenerate skew symmetric bilinear form. In particular, it induces a linear isomorphism $T_xM \to T_x^*M$ given by mapping $\xi_x \in T_xM$ to $\omega(x)(\xi_x, \_) : T_xM \to \mathbb{R}$. Likewise, for a vector field $\xi \in \mathfrak{X}(M)$, one can form $i_\xi\omega \in \Omega^1(M)$ and it follows easily that this induces a linear isomorphism $\mathfrak{X}(M) \to \Omega^1(M)$. (The only non–trivial part is to verify surjectivity. This amounts to showing that given a one–form $\alpha \in \Omega^1(M)$ the uniquely determined tangent vectors ξ_x such that $\omega(x)(\xi_x, \_) = \alpha(x)$ depend smoothly on x which is easy.)

In particular, given a smooth function $f : M \to \mathbb{R}$, there is a unique vector field $H_f \in \mathfrak{X}(M)$ such that $i_{H_f}\omega = df$. This is called the *Hamiltonian vector field* generated by f or the *symplectic gradient* of f. Starting from the function f, one can then look at integral curves for H_f so this gives rise to a dynamical system on M. The first nice observation is that $di_{H_f}\omega = ddf = 0$ which together with $d\omega = 0$ implies that $\mathcal{L}_{H_f}\omega = 0$. Since $\frac{d}{dt}(\mathrm{Fl}_t^{H_f})^*\omega = (\mathrm{Fl}_t^{H_f})^*\mathcal{L}_{H_f}\omega$, this shows that $(\mathrm{Fl}_t^{H_f})^*\omega = \omega$. So the flow of any Hamiltonian vector field consists of symplectomorphisms.

There is another structure which can be immediately obtained from this construction. Given $f, g \in C^\infty(M, \mathbb{R})$ one can define a third smooth function $\{f, g\} : M \to \mathbb{R}$ by $\{f, g\} := -\omega(H_f, H_g)$, i.e. by inserting the Hamiltonian vector fields into the symplectic form. This is called the *Poisson bracket* of f and g. From the definition, it follows immediately that $\{f, g\} = dg(H_f) = -df(H_g) = -\{g, f\}$. We can also write $\{f, g\} = i_{H_f}dg = \mathcal{L}_{H_f}g$ since $i_{H_f}g = 0$. Using the calculus on differential forms from [**DG1**, 4.9] we can further compute

$$d\{f, g\} = d\mathcal{L}_{H_f}g = \mathcal{L}_{H_f}dg = \mathcal{L}_{H_f}i_{H_g}\omega = (\mathcal{L}_{H_f}i_{H_g} - i_{H_g}\mathcal{L}_{H_f})\omega = i_{[H_f,H_g]}\omega.$$

This means that $H_{\{f,g\}} = [H_f, H_g]$. In particular, this shows that the Poisson bracket satisfies the Jacobi identity $\{f, \{g, h\}\} = \{\{f, g\}, h\} + \{g, \{f, h\}\}$ and thus makes $C^\infty(M, \mathbb{R})$ into a Lie algebra. Moreover, the map $f \mapsto H_f$ defines a homomorphism of Lie algebras from $(C^\infty(M, \mathbb{R}), \{\ ,\ \})$ to the Lie algebra of vector fields on M.

The importance of all that comes from classical mechanics. Suppose that we look at an open subset $U \subset \mathbb{R}^n$ and let us write the coordinates as $q^1, \dots, q^n$. Suppose further that we have a force field on this open subset which (in physics terminology) can be described by a potential. This just means that we have a smooth function $V : U \to \mathbb{R}$ such that the force acting on a particle in position q is given by the negative of the gradient of V in q, so it equals $(-\frac{\partial V}{\partial q^1}(q), \dots, -\frac{\partial V}{\partial q^n}(q))$. Now one passes to the phase space T^*U which is an open subset in $T^*\mathbb{R}^n$. Now suppose we want to study the movement of a point particle of mass m in the force field under consideration. This is governed by Newton's law, i.e. force equals acceleration times mass. Given a curve $q(t) = (q^1(t), \dots, q^n(t))$ in U, we can define a natural lift to a curve in T^*U, by requiring that the component in $T^*_{q(t)}U$ is the momentum of the particle in the point $q(t)$. This means that $p_i(t) = mq_i'(t)$. The components of the acceleration are then $q_i''(t) = \frac{p_i'(t)}{m}$. Thus Newton's law simply says that $p_i'(t) = -\frac{\partial V}{\partial q^n}(q(t))$. Thus the curve of our particle in phase space should be a solution of the system of first order ordinary differential equations given by

$$q_i'(t) = \tfrac{p_i(t)}{m} \qquad p_i'(t) = -\tfrac{\partial V}{\partial q^n}(q(t)).$$

Now there is a natural energy function in this situation, namely $E(q^i, p_i) := \sum \frac{p_i^2}{2m} + V(q)$ (kinetic energy plus potential energy). Now we immediately get

$$dE = \textstyle\sum_{i=1}^n \frac{p_i}{m}dp_i + \sum_{i=1}^n \frac{\partial V}{\partial q^i}dq^i.$$

Since the canonical symplectic form on T^*U is given by $\sum_i dq^i \wedge dp_i$, we conclude that

$$H_E = \textstyle\sum_{i=1}^n \frac{p_i}{m}\frac{\partial}{\partial q^i} - \sum_{i=1}^n \frac{\partial V}{\partial q^i}\frac{\partial}{\partial p_i}.$$

Thus the equations of motion deduced above are exactly the equations describing the flow lines of the Hamiltonian vector field generated by the energy function.

In view of this, it has become common to view the study of flows of Hamiltonian vector fields on symplectic manifolds as the basis of classical mechanics. Some of the things we have done so far have an immediate relevance for this study. For example, suppose that $f : M \to \mathbb{R}$ is a smooth function which *Poisson–commutes* with E, i.e. has the property that $\{E, f\} = 0$. Then we see from above that $0 = df(H_E) = H_E \cdot f$, which implies that f is constant along the flow lines of H_E, and thus defines a constant of motion of the mechanical system determined by E. This restricts the motion of the systems to the level sets of f which are smooth hypersurfaces in M locally around points in which $df \neq 0$. In particular, this applies to the function E itself.

Of particular interest is the situation that for a system on a manifold M of dimension $2n$ one finds n constants of motion $E = f_1, f_2, \dots, f_n$ such that $\{f_i, f_j\} = 0$ for all i, j and such that the open subset of M on which $df_1, \dots, df_n$ are linearly independent is dense. In this case, the system is called *integrable* and it turns out that the structure of such systems can be analyzed quite nicely, see [**Mi**, section 32].

Contact geometry

1.10. Contact forms and contact structures. As promised in 1.6 we now move to the study of one–forms α with $d\alpha$ as non–degenerate as possible in the case of odd dimensions. So lets assume that M is a smooth manifold of odd dimension, say $\dim(M) = 2n+1$. Having given a nowhere vanishing one–form $\alpha \in \Omega^1(M)$ we have, at each point $x \in M$ the subspace $\ker(\alpha(x)) \subset T_xM$ which has dimension $2n$. Hence the natural assumption of non–degeneracy we can use is that the restriction of the skew symmetric bilinear form $d\alpha(x)$ to this subspace is non-degenerate. Similarly as in the case of symplectic forms, this is equivalent to the fact that the wedge product of n copies of $d\alpha(x)$ is non–zero on $\ker(\alpha(x))$. From this, it is easy to see that our original condition on α is equivalent to the fact that $\alpha \wedge (d\alpha)^n \in \Omega^{2n+1}(M)$ is nowhere vanishing. This motivates the first part of the following definition:

Definition 1.10. Let M be a smooth manifold of odd dimension $2n+1$.

(1) A *contact form* on M is a one–form $\alpha \in \Omega^1(M)$ such that $\alpha \wedge (d\alpha)^n \in \Omega^{2n+1}(M)$ is nowhere vanishing.

(2) A *contact structure* on M is a smooth distribution $H \subset TM$ of rank $2n$ such that for any point $x \in M$ there exists an open neighborhood U of x in M and a contact form $\alpha \in \Omega^1(U)$ such that $H_y = \ker(\alpha(y))$ for all $y \in U$.

Remark 1.10. (1) A manifold must be orientable in order to admit a contact form.

(2) If $\alpha \in \Omega^1(M)$ is a contact form on M, then of course $H_x := \ker(\alpha(x))$ is a contact structure on M. Passing from a contact form to a contact structure does not only mean giving up the requirement of global existence of a contact form (and it turns out that contact structures can exist on non–orientable manifolds). The additional thing is that one can evidently rescale contact forms. If $\alpha \in \Omega^1(M)$ is a contact form and $f : M \to \mathbb{R}$ is a nowhere vanishing smooth function, then for the form $\beta := f\alpha$ we have $\ker(\beta(x)) = \ker(\alpha(x))$ and $d\beta(x) = f(x)d\alpha(x) + df(x) \wedge \alpha(x)$ and the second summand vanishes on $\ker(\beta(x)) = \ker(\alpha(x))$. Consequently, β is a contact form, too, which induces the same contact structure. Conversely, a contact structure locally defines a contact form up to multiplication by a nowhere vanishing function.

1.11. Jets. Similarly to the canonical symplectic structure on a cotangent bundle, there are also several constructions of manifolds with canonical contact forms and contact structures starting from an arbitrary smooth manifold. A particularly interesting version of this is the space of one–jets of real valued functions. This actually is a very special case of a general concept that we will develop (at least roughly) next. The aim of the concept of jets is to provide an abstract (i.e. coordinate–free) version of Taylor developments to a certain order for smooth functions between manifolds. This also leads to a geometric approach to the study of differential equations. We will only use 1–jets explicitly in the sequel, but higher jets are an important concept in differential geometry. Thus we will formulate results in general but sometimes only prove them for 1–jets in detail and sketch how to extend to higher jets.

To formulate the basic equivalence relation, observe first that for an open interval $I \subset \mathbb{R}$ and a smooth curve $c : I \to M$ in some manifold, one may view the derivative as a smooth curve $c' : I \to TM$. This allows to form $c'' : I \to TTM$ and iteratively $c^{(k)} : I \to T^kM = T \dots TM$. In particular, for any $t \in I$, $c^{(k)}(t)$ has a well defined meaning independent of any choices. For convenience, we put $c^{(0)}(t) = c(t)$.

Definition 1.11. (1) Let $I \subset \mathbb{R}$ be an open interval, $t_0 \in I$ a point and M a smooth manifold. Two smooth curves $c_1, c_2 : I \to M$ have *kth order contact* in t_0 if and only if $c_1^{(i)}(t_0) = c_2^{(i)}(t_0)$ for $i = 0, \dots, k$.

(2) Let M and N be smooth manifolds $x \in M$ a point, and f, g two N–valued smooth functions defined on some open neighborhoods of x in M. One says that f and g *have the same k–jet in x* if and only if for any open interval $I \subset \mathbb{R}$ containing 0 and any smooth curve $c : I \to M$ with $c(0) = x$, the curves $f \circ c$ and $g \circ c$ have kth order contact in 0. One writes $j^k_x f = j^k_x g$ in this case.

Both relations we have defined here visibly are equivalence relations and they are manifestly independent of coordinates or any other choices. One may view $j^k_x f$ as representing the equivalence class of f. The set of all such equivalence classes is then denoted by $J^k(M, N)$ and called the set of all *k–jets of smooth maps from M to N*. The chain rule immediately implies that two curves $c_1, c_2 : I \to M$ have kth order contact in t_0 if and only if they have the same k–jet in t_0 when viewed as smooth maps $I \to M$. What one should actually keep in mind about is the following

Lemma 1.11. *Two smooth maps $f, g : M \to N$ have the same k–jet in $x \in M$ if and only if $f(x) = g(x)$ and for some (or equivalently any) charts (U, u) for M with $x \in U$ and (V, v) for N with $f(x) = g(x) \in V$, the local coordinate representations $v \circ f \circ u^{-1}$ and $v \circ g \circ u^{-1}$ have the same derivatives in $u(x)$ up to order k.*

Sketch of proof. First we observe that two curves $c_1, c_2 : I \to M$ have kth order contact in $t_0 \in I$ if and only if $c_1(t_0) = c_2(t_0)$ and for a chart (U, u) containing this point, one has $(u \circ c_1)^{(i)}(t_0) = (u \circ c_2)^{(i)}(t_0)$ for $1 \leq i \leq k$. For $k = 1$, this is obvious since $(u \circ c_1)'(t) = T_{c_1(t_0)}u \cdot T_{t_0}c_1 \cdot 1$. The general result then follows inductively by viewing the derivatives as curves in iterated tangent bundles. It then follows from the chain rule, that the derivatives will be the same in any chart containing $c_1(t_0) = c_2(t_0)$.

Now let us prove the general result for $k = 1$. If I contains zero and $c : I \to M$ is a smooth curve with $c(0) = x$, then locally around zero, we can write $f \circ c$ as $(f \circ u^{-1}) \circ (u \circ c)$. Now the chain rule implies that

$$(v \circ f \circ c)'(0) = D(v \circ f \circ u^{-1})(u(x))((u \circ c)'(0)).$$

If $v \circ f \circ u^{-1}$ and $v \circ g \circ u^{-1}$ have the same derivative in $u(x)$, this clearly implies that $(v \circ f \circ c)'(0) = (v \circ g \circ c)'(0)$ for any curve c with $c(0) = x$. Since $f(x) = g(x)$ our observations on curves above imply that $j^1_x f = j^1_x g$.

Conversely, if $j^1_x f = j^1_x g$ we can read the argument backwards to conclude that $D(v \circ f \circ u^{-1})(u(x))$ and $D(v \circ g \circ u^{-1})(u(x))$ agree on all tangent vectors of the form $(u \circ c)'(0) = T_x u \cdot c'(0)$ for curves c as above. But since u is a diffeomorphism $T_x u$ is a linear isomorphism, so any tangent vector can be written in this form. Finally, the chain rule again implies that if the derivatives agree in one chart, then they agree in any chart.

To prove the result for higher k, one has to argue inductively and use that chain rule for $(v \circ f \circ c)^{(k)}(t_0)$. This shows that one can write $(v \circ f \circ c)^{(k)}(0)$ as a sum of $D^{(k)} f(u(x))((u \circ c)'(0), \dots, (u \circ c)'(0))$ and of terms involving only lower derivatives of f. Assuming inductively that the derivatives in $u(x)$ of f and g up to order $k-1$ agree, one concludes that $(v \circ f \circ c)^{(k)}(t_0) = (v \circ g \circ c)^{(k)}(t_0)$ for all c is equivalent to the fact that $D^{(k)} f(u(x))$ and $D^{(k)} g(u(x))$ agree provided that one inserts k copies of the same tangent vector. But by symmetry of the kth derivative (i.e. by polarization), this is equivalent to $D^{(k)} f(u(x)) = D^{(k)} g(u(x))$. The rest then follows by the chain rule as before. □

This result immediately allows us to obtain an explicit description of $J^k(U, V)$ where $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are open subsets. For $r \geq 1$ let us denote by $L^r_s(\mathbb{R}^n, \mathbb{R}^m)$ the vector space of r–linear symmetric maps $\mathbb{R}^n \times \ldots \times \mathbb{R}^n \to \mathbb{R}^m$ (with r factors $\mathbb{R}^n$). Note that for $r = 1$, we simply obtain the space of linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$. An element $\varphi \in L^r_s(\mathbb{R}^n, \mathbb{R}^m)$ also defines a smooth map $\mathbb{R}^n \to \mathbb{R}^m$ by mapping $x \in \mathbb{R}^n$ to $\varphi(x, \dots, x)$, so this is a homogeneous polynomial of degree r. Now we can define a map

$$J^k(U, V) \to U \times V \times \oplus_{r=1}^k L^r_s(\mathbb{R}^n, \mathbb{R}^m)$$

by sending $j^k_u f$ to $(u, f(u), Df(u), \dots, D^{(k)} f(u))$. By the lemma above, this map is injective. To see that it is also surjective, given an element $(u, v, \varphi_1, \dots, \varphi_k)$ in the right hand side, define $f : \mathbb{R}^n \to \mathbb{R}^m$ by $f(x) := v + \varphi_1(x - u) + \cdots + \varphi_k(x - u, \dots, x - u)$. Then this maps u to v satisfies $D^{(i)} f(u) = \varphi_i$ for $i = 1, \dots, k$. Hence it will map a neighborhood of x to V, and using a partition of unity one can change it in such a way that it maps all of U (and even all of $\mathbb{R}^n$) to V without changing the local behavior around x. The k–jet of the resulting map in u is then mapped to $(u, v, \varphi_1, \dots, \varphi_k)$. Hence we have obtained an identification of $J^k(U, V)$ with an open subset in a finite dimensional vector space.

Now we can easily use this, to make $J^k(M, N)$ into a manifold. There is an evident projection $\pi = (\pi_M, \pi_N) : J^k(M, N) \to M \times N$ defined by $\pi(j^k_x f) = (x, f(x))$. Now take charts (U, u) for M and (V, v) for N and consider $J^k(U, V) = \pi^{-1}(U \times V) \subset J^k(M, N)$. Then for $j^k_x f \in J^k(U, V)$, we have $x \in U$ and $f(x) \in V$, and thus we can consider $j^k_{u(x)}(v \circ f \circ u^{-1}) \in J^k(u(U), v(V))$. Lemma 1.11 shows that this defines a bijection $J^k(U, V) \to J^k(u(U), v(V))$, which as we know from above is an open subset in a finite dimensional vector space. Now one can put a topology on the set $J^k(M, N)$ by defining a subset to be open if for any two charts as above its intersection with $J^k(U, V)$ has open image in $J^k(u(U), v(V))$. One then shows that this topology is separable and metrizable. Then the bijections $J^k(U, V) \to J^k(u(U), v(V))$ defined above are homeomorphisms so we may use them as charts. Finally, one easily verifies that the chart changes for these charts are made up form chart changes on M and N and their derivatives up to order k and thus are smooth. Hence $J^k(M, N)$ becomes a smooth manifold and the projection $\pi : J^k(M, N) \to M \times N$ is a smooth map.

There is a simple generalization of the map π. For $\ell < k$, there is an obvious projection $\pi^k_\ell : J^k(M, N) \to J^\ell(M, N)$ defined by $\pi^k_\ell(j^k_x f) := j^\ell_x f$. In appropriate charts, this map is given by forgetting some components and keeping the others unchanged, so it is smooth, too.

Finally, observe that for an open subset $U \subset M$ and a smooth function $f : U \to N$ we obtain a natural function $j^k f : U \to J^k(M, N)$, defined by $j^k f(x) := j^k_x f$ for all $x \in u$. In the charts constructed above, this function is just given by the local coordinate representation of f and its derivatives up to order k, so j^kf is smooth. By construction, we get $\pi_M \circ j^kf = \mathrm{id}_U$ (so j^kf is a *smooth section* of $\pi_M : J^k(M,N) \to M$) as well as $\pi_N \circ j^kf = f$.

1.12. Jets and partial differential equations. We now specialize our considerations to the case $k = 1$ and $N = \mathbb{R}$, so we look at the space $J^1(M,\mathbb{R})$ of 1–jets of real valued functions on a smooth manifold M. In this case, we can simply take the identity map as a chart on $\mathbb{R}$. Hence we start with a chart (U,u) for M and consider the induced chart

$$J^1(M,\mathbb{R}) \supset J^1(U,\mathbb{R}) \to J^1(u(U),\mathbb{R}) \cong u(U) \times \mathbb{R} \times L(\mathbb{R}^n,\mathbb{R})$$

where $n = \dim(M)$ which maps j^1_xf to $(u(x), f(x), D(f \circ u^{-1})(u(x)))$. This immediately gives rise to local coordinates on $J^1(M,\mathbb{R})$. We have $u(x) = (u^1(x),\dots,u^n(x))$ and we denote the standard coordinate on $\mathbb{R}$ by v. On $L(\mathbb{R}^n,\mathbb{R})$ we use coordinates p_i with respect to the standard basis, so $(p_1,\dots,p_n)$ denotes the functional $(a^1,\dots,a^n) \mapsto \sum a^ip_i$. Thus we obtain local coordinates of the form $(u^1,\dots,u^n,v,p_1,\dots,p_n)$. For a smooth map $f : U \to \mathbb{R}$ we see immediately what the map $j^1f : U \to J^1(M.\mathbb{R})$ looks like in these local coordinates. The local coordinate representation maps $u(U)$ to $u(U) \times \mathbb{R} \times \mathbb{R}^n$ and it evidently has the form

$$y \mapsto (y, (f \circ u^{-1})(y), \tfrac{\partial f}{\partial u^1}(u^{-1}(y)),\dots,\tfrac{\partial f}{\partial u^n}(u^{-1}(y))).$$

So in general, can think about the coordinates p_i representing formal partial derivatives.

This immediately gives a connection to partial differential equations. A partial differential equation of order k on real valued functions on a manifold M is defined as an equation on a function $f \in C^\infty(M,\mathbb{R})$ which can be written as a PDE of order k in local coordinates. (The chain rule implies that while these PDE may look completely different for different sets of coordinated, the fact that the equation can be written as a PDE of order k is independent of the choice of coordinates.)

Suppose we have given a PDE of order 1. The we take a chart (U,u) for M and write out the equation as a first order PDE in the local coordinate expression of a function f, say

$$F(y, (f \circ u^{-1})(y), \tfrac{\partial (f\circ u^{-1})}{\partial y_1}(y),\dots,\tfrac{\partial (f\circ u^{-1})}{\partial y_n}(y)) = 0.$$

Making the (harmless) assumption that F is smooth and regular (i.e. dF is nowhere vanishing), $F(u^1,\dots,u^n,v,p_1,\dots,p_n) = 0$ defines a smooth hypersurface in $J^1(U,\mathbb{R})$. The tangent spaces of this hypersurface are the kernels of dF. Now in order to really define have an equation, F should not only depend on the variables u^i, so one usually assumes that in each point $\beta \in J^1(U,\mathbb{R})$, at least on of the partial derivatives $\frac{\partial F}{\partial v}(\beta)$ and $\frac{\partial F}{\partial p_i}(\beta)$ is non–zero. Otherwise put, not all the vectors $\frac{\partial}{\partial v}(\beta)$ and $\frac{\partial}{\partial p_i}(\beta)$ should lie in $\ker(dF(\beta))$.

This admits a nice coordinate independent formulation: Recall that we have a canonical projection $\pi_M : J^1(M,\mathbb{R}) \to M$, so we can consider its tangent mappings $T_\beta\pi_M$ which evidently is surjective in each point, so π_M is a submersion. The kernel of $T_\beta\pi_M$ thus is an $n+1$–dimensional subspace of $T_\beta J^1(M,\mathbb{R})$, which in coordinates (u^i,v,p_i) as above, is evidently spanned by the vectors $\frac{\partial}{\partial p_i}(\beta)$ and $\frac{\partial}{\partial v}(\beta)$. Hence the assumption on F from above can be phrased as the fact that $\ker(dF(\beta))$ never contains $\ker(T_\beta\pi_M)$. From dimensional considerations it then follows that $\ker(dF(\beta)) \cap \ker(T_\beta\pi_M)$ has dimension n and $T_\beta J^1(M,\mathbb{R})$ is spanned by $\ker(dF(\beta))$ and $\ker(T_\beta\pi_M)$.

To have a globally defined PDE on all of M, one has to assume that the local hypersurfaces constructed in different coordinate charts fit together to define a hypersurface in $J^1(M,\mathbb{R})$. But then one can turn things around and simply *define* a first order PDE on a smooth manifold M as a smooth hypersurface $\Sigma \subset J^1(M,\mathbb{R})$. As above, one assumes that for each $\beta \in \Sigma$, the tangent space $T_\beta\Sigma$ does not contain $\ker(T_\beta\pi_M)$. By dimensional considerations this implies that the restriction of $T_\beta\pi$ to $T_\beta\Sigma$ is still surjective, so π_M restricts to a submersion on Σ. To have the PDE defined on all of M, one in addition assumes that $\pi_M : \Sigma \to M$ is surjective. From the above description, it is also clear what solutions of the equation are in this picture, so we define:

Definition 1.12. (1) A *first order PDE* on real valued functions on a smooth manifold M is a smooth hypersurface $\Sigma \subset J^1(M,\mathbb{R})$ such that the canonical projection $\pi_M : J^1(M,\mathbb{R}) \to M$ restricts to a surjective submersion $\pi_M : \Sigma \to M$.

(2) A (local) solution of the PDE defined by a hypersurface $\Sigma \subset J^1(M,\mathbb{R})$ is a smooth function $f : M \to \mathbb{R}$ (respectively defined on a open subset $U \subset M$) such that the corresponding section j^1f satisfies $j^1f(M) \subset \Sigma$ (respectively $j^1f(U) \subset \Sigma$).

The advantage of this approach is of course that one obtains a picture of a PDE which is independent of a choice of coordinates. We shall soon see that we can also dispense with smooth functions and view solutions of a PDE as special submanifolds of Σ, which further eliminates the role of coordinates. Of course, one can similarly discuss higher order PDEs by looking at higher jets, other targets than $\mathbb{R}$, and at systems of PDE by looking at submanifold of higher codimension rather than hypersurfaces. This is the basis for the geometric theory of PDEs.

1.13. The canonical contact form on one–jets. Now we can mimic the construction of the canonical one–form on a cotangent bundle to obtain a canonical one form on the manifold $J^1(M,\mathbb{R})$. Let $\beta = j^1_xf \in J^1(M,\mathbb{R})$ be a point and let $\xi \in T_\beta J^1(M,\mathbb{R})$ be a tangent vector. Via the two natural projections, we can form $T_\beta\pi_M \cdot \xi \in T_xM$ and $T_\beta\pi_{\mathbb{R}} \cdot \xi \in T_{f(x)}\mathbb{R} = \mathbb{R}$. On the other hand, we have noted that T_xf depends only on j^1_xf, and in this case, the tangent map is just $df(x)$. Thus we define the canonical one form on N by

$$\alpha(j^1_xf)(\xi) := T_\beta\pi_{\mathbb{R}} \cdot \xi - df(x)(T_\beta\pi_M \cdot \xi) \in \mathbb{R}.$$

Choosing a chart (U,u) for M and the induced coordinates (u^i,v,p_i) on $J^1(U,\mathbb{R}) \subset J^1(M,\mathbb{R})$, we can immediately write out this form explicitly in terms of the resulting coordinate vector fields. Evidently, $T\pi_M$ maps the coordinate fields $\frac{\partial}{\partial u^i}$ to their counterparts on M and annihilates the other coordinate fields. Likewise $T\pi_{\mathbb{R}}$ annihilates all coordinate fields except $\frac{\partial}{\partial v}$, which is mapped to the canonical field $\frac{\partial}{\partial s}$ on $\mathbb{R}$. Finally, the coordinates of $df(x)$ with respect to the standard basis are simply the p_i. This immediately implies that

$$\alpha|_{J^1(U,\mathbb{R})} = dv - \textstyle\sum_{i=1}^n p_i du^i.$$

This shows that $\ker(\alpha(j^1_xf))$ is spanned by the values of the vector fields $\frac{\partial}{\partial p^i}$ for $i = 1,\dots,n$ and $\frac{\partial}{\partial u^i} + p_i\frac{\partial}{\partial v}$ for $i = 1,\dots,n$. Since these fields are evidently linearly independent in each point, we have constructed a frame for $\ker(\alpha)$ over $J^1(U,\mathbb{R})$. On the other hand, from the above formula we see that the restriction of $d\alpha$ to $J^1(U,\mathbb{R})$ is given by $\sum_{i=1}^n du^i \wedge dp_i$. This immediately shows that the restriction of $d\alpha$ to $\ker(\alpha)$ is always non–degenerate (and the elements of our frame restrict to a basis in standard form on each of these spaces). In particular, α is a contact form on $J^1(M,\mathbb{R})$.

To formulate an analog of Proposition 1.7 in the contact case, we need one more notion connected to integral submanifolds. Suppose the M is a smooth manifold of dimension $2n+1$ endowed with a contact form $\alpha \in \Omega^1(M)$. Recall from Definition 1.2 that an integral submanifold for α is a smooth submanifold $N \subset M$ such that α restricts to zero on N. Equivalently, this means that $T_xN \subset \ker(\alpha(x)) \subset T_xM$ for all $x \in N$. But if $\alpha|_N = 0$, then $d\alpha|_N = 0$, so we conclude that each T_xN must be an isotropic subspace for the non–degenerate bilinear form $d\alpha(x)$ on $\ker(\alpha(x))$. This immediately implies that $\dim(N) \leq n$.

Definition 1.13. Let M be a smooth manifold endowed with a contact form $\alpha \in \Omega^1(M)$. A *Legendrean submanifold* for (M, α) is an n–dimensional integral submanifold for α.

It is easy to see that this condition makes also sense if M is just endowed with a contact structure. Indeed, if $N \subset M$ is Legendrean with respect to a chosen local contact form inducing the contact structure, then the same is true for any such contact form.

Proposition 1.13. *Let M be any smooth manifold, and consider $J^1(M,\mathbb{R})$ endowed with its canonical contact form α. Let $\pi : J^1(M,\mathbb{R}) \to M\times\mathbb{R}$ be the canonical projection.*

(1) For any point $(x,t) \in M \times R$, the preimage $\pi^{-1}(x,t) \subset J^1(M,\mathbb{R})$ is a Legendrean submanifold.

(2) For an open subset $V \subset M$ and a smooth function $f : U \to \mathbb{R}$, the image $N := j^1f(V) \subset J^1(M,\mathbb{R})$ is a smooth Legendrean submanifold. Moreover, for each $\beta \in N$ we have $T_\beta N \cap \ker(T_\beta\pi_M) = \{0\}$.

(3) Suppose that $N \subset J^1(M,\mathbb{R})$ is a Legendrean submanifold and that $\beta \in N$ is a point such that $T_\beta N \cap \ker(T_\beta\pi_M) = \{0\}$. Then there is an open neighborhood V of $x := \pi_M(\beta)$ and a smooth function $f : V \to \mathbb{R}$ such that $j^1f(V)$ coincides with an open neighborhood of β in N.

(4) Let $\varphi : M \to M$ be a diffeomorphism, and define $\Phi : J^1(M,\mathbb{R}) \to J^1(M,\mathbb{R})$ be defined by $\Phi(j^1_xf) = j^1_{\varphi(x)}(f \circ \varphi^{-1})$. Then Φ is a diffeomorphism with $\Phi^*\alpha = \alpha$.

Proof. Most of the proof is parallel to the proof of Proposition 1.7, so we are rather brief here.

(1) Take a chart (U,u) for M with $x \in U$. Then in the induced chart for $J^1(M,\mathbb{R})$ the subset $\pi^{-1}(x,t)$ just corresponds to $\{x\} \times \{t\} \times \mathbb{R}^{n*}$, so it is a submanifold. Its tangent spaces are spanned by the coordinate fields $\frac{\partial}{\partial p_i}$ which obviously insert trivially into α.

(2) Take a chart (U,u) for M with $U \subset V$. Then in the induced chart for $J^1(M,\mathbb{R})$, the subset $j^1f(V) \cap J^1(U,\mathbb{R}) = j^1f(U)$ corresponds to $\{(u^i, (f\circ u^{-1})(u^i), \frac{\partial f}{\partial u^i}(u^{-1}(u^i)))\}$. This is the graph of a smooth function and thus a smooth submanifold, so $j^1f(U)$ is a submanifold in $J^1(M,\mathbb{R})$.

Knowing that $j^1f(V) \subset J^1(M,\mathbb{R})$ is a submanifold, it is clear that $j^1f : V \to j^1f(V)$ is a diffeomorphism, since $p|_{j^1f(V)}$ is a smooth inverse. To prove that $j^1f(V)$ is Legendrean i.e. that $\alpha|_{j^1f(V)} = 0$, it thus suffices to show that $0 = (j^1f)^*\alpha \in \Omega^1(V)$. Now by definition

$$(j^1f)^*\alpha(x)(\xi) = \alpha(j^1_xf)(T_xj^1f\cdot\xi) = T_{j^1_xf}\pi_{\mathbb{R}} \cdot T_xj^1f\cdot\xi - df(x)(T_{j^1_xf}\pi_M \cdot T_xj^1f\cdot\xi).$$

But we know that $\pi_{\mathbb{R}} \circ j^1f = f$ and $\pi_M \circ j^1f = \mathrm{id}_M$. Thus $T_{j^1_xf}\pi_{\mathbb{R}} \cdot T_xj^1f\cdot\xi = T_xf\cdot\xi$ and $T_{j^1_xf}\pi_M \cdot T_xj^1f\cdot\xi = \xi$, and the result follows.

(3) The projection π_M restricts to a smooth map $\psi : N \to M$. Now by assumption $T_\beta\pi_M : T_\beta N \to T_xM$ has trivial kernel, so it must be a linear isomorphism since both

spaces have the same dimension. Thus there are open neighborhoods W of β in $J^1(M,\mathbb{R})$ and V of x in M such that $\pi_M|_W : W \to V$ is a diffeomorphism. Now define $f : V \to \mathbb{R}$ as $f := \pi_{\mathbb{R}} \circ (\pi_M|_W)^{-1}$. Clearly, this is a smooth map, and to complete the proof, it suffices to see that $(\pi_M|_W)^{-1} : V \to J^1(M,\mathbb{R})$ coincides with $j^1 f$. We can do this in a local chart (U,u) with $U \subset V$. But there, we already know that $(\pi_M|_W)^{-1}$ must have the form

$$u = (u^1,\dots,u^n) \mapsto (u^1,\dots,u^n,(f\circ u^{-1})(u),p_1(u),\dots,p_n(u)),$$

for some functions p_i, so in particular $v = (f \circ u^{-1})(u)$. Differentiating this, we see that on N, we have $dv = \sum \frac{\partial(f\circ u^{-1})}{\partial u^i} du^i$. On the other hand, since N is Legendrean, we have $\alpha|_N = 0$ and hence $dv = \sum p_i du^i$ holds on N. But the fact that π_M restricts to a local diffeomorphism on N also implies that the forms $du^1(x),\dots,du^n(x)$ restrict to a basis of $L(T_xN,\mathbb{R})$ for each $x \in N$. Thus the two formulae express the vector $dv(x)$ in this basis, so they imply that $p_i = \frac{\partial(f\circ u^{-1})}{\partial u^i}$ for $i = 1,\dots,n$, and thus $(\pi_M|_W)^{-1} = j^1 f|_V$.

(4) This follows similarly as in the proof of Proposition 1.7 from the fact that the canonical contact form has the same expression with respect to all charts induced from local charts on M. □

Remark 1.13. (1) With part (3) of the proposition we have completed the setup for a geometric approach to partial differential equations of first order. We already know from 1.12 that we can describe a first order PDE on a manifold M as a hypersurface $\Sigma \subset J^1(M,\mathbb{R})$. Now we also know how to interpret solutions in a geometric way. We just have to look for Legendrean submanifolds of $J^1(M,\mathbb{R})$ which are contained in the hypersurface Σ (or equivalently for n–dimensional integral manifolds for $\alpha|_\Sigma$), whose tangent spaces have zero intersection with $\ker(T\pi_M)$. This can also be generalize to higher order equations and to systems of equations, by looking at analogs of the canonical contact form on spaces of higher order jets. It should also be remarked that the point of view of solutions as submanifolds has some immediate advantages. For example, some types of singularities in solutions of PDEs can be described in such a way that one obtains a nice smooth Legendrean submanifold in $J^1(M,\mathbb{R})$, but the condition on the intersection with $\ker(T\pi_M)$ is not satisfied everywhere.

(2) There is an analog of the Darboux theorem (Theorem 1.8) in the contact case, often referred to as the Pfaff theorem. This is best phrased by saying that given a contact structure $H \subset TM$ on a smooth manifold M of dimension $2n+1$, one can always find local coordinates $(u^1,\dots,u^n,v,p_1,\dots,p_n)$ such that H can be written as the kernel of the form $dv - \sum p_i du^i$, see Theorem 1.9.17 of [**IL**]. Of course, this means that the contact structure is locally isomorphic to $J^1(\mathbb{R}^n,\mathbb{R})$ with its canonical contact structure. In particular, Proposition 1.13 then implies that any contact manifold locally admits many Legendrean submanifolds and many contact–diffeomorphisms.

1.14. Application: The method of characteristics. As an application of the geometric approach to first order PDEs, we describe a classical method for solving first order PDEs, the method of characteristics. For simplicity, we restrict to the case $M = \mathbb{R}^n$, so we have global coordinates $u^1,\dots,u^n,v,p_1,\dots,p_n$ on $J^1(\mathbb{R}^n,\mathbb{R})$ for which the canonical contact form is given by $dv - \sum p_i du^i$. In view of the discussion in 1.12, we initially write the hypersurface Σ describing our equation in the form $F(u^1,\dots,u^n,v,p_1,\dots,p_n) = 0$ for a smooth function F. Assuming that F really describes a first order equation (and not an equation of order zero), we assume that in each point at least one of the derivatives $\frac{\partial F}{\partial p_i}$ is non–zero (which also implies that our equation defines a smooth hypersurface in $J^1(\mathbb{R}^n, \mathbb{R})$). By the implicit function theorem, we can then locally solve for p_i and restricting to an open subset and renumbering coordinates if necessary, we assume that our hypersurface is given in the form

$$(*) \qquad p_n = f(u^1, \dots, u^n, v, p_1, \dots, p_{n-1})$$

for some smooth function f.

Now let α be the canonical contact form on $J^1(\mathbb{R}^n, \mathbb{R})$ and consider the restriction of α to Σ. Further, in view of the defining equation $(*)$, the vector field $\frac{\partial}{\partial p_n}$ never lies in $T_z\Sigma \subset T_z J^1(\mathbb{R}^n, \mathbb{R})$. This vector field is always contained in the contact subspace H_z, so we see that $T_z\Sigma$ can never contain H_z. Consequently, the intersection $T_z\Sigma \cap H_z$ must have dimension $2n-1$ for each point $z \in \Sigma$. Now consider the skew symmetric bilinear form $d\alpha(z)$ on this space. It must be degenerate since the space is odd dimensional, but on the other hand the rank must be at least $n-1$, since $d\alpha(z)$ is non–degenerate on H_z. Hence the Nullspace of $d\alpha(z)$ is one–dimensional and clearly, it depends smoothly on z. This is called the *characteristics of the equations.*

Hence we can locally choose a nowhere vanishing vector field $\xi \in \mathfrak{X}(\Sigma)$ which inserts trivially in the restriction of $d\alpha$ to $T\Sigma$. (This vector field is uniquely determined up to multiplication with a nowhere vanishing function.) Now choose an arbitrary smooth function $\varphi : \mathbb{R}^{n-1} \to \mathbb{R}$ and consider the $(n-1)$–dimensional submanifold $L \subset J^1(\mathbb{R}^n, \mathbb{R})$ defined by $u^n = C$ for some $C \in \mathbb{R}$, $v = \varphi(u^1, \dots, u^{n-1})$, $p_i = \frac{\partial \varphi}{\partial u^i}$ for $i = 1, \dots, n-1$ and then p_n via $(*)$.

Then on L, we have $dv = \sum_{i=1}^{n-1} \frac{\partial \varphi}{\partial u^i} du^i = \sum_{i=1}^{n-1} p_i du^i$ and $du^n = 0$, so $L \subset \Sigma$ is an integral submanifold for α (of dimension $n-1$). Further, we know that $\frac{\partial}{\partial p_n}$ and $T\Sigma \cap H$ spans all of H, so non–degeneracy of $d\alpha$ on H implies that $d\alpha(\frac{\partial}{\partial p_n}, \xi) = du^n(\xi) \neq 0$. But this implies that ξ is never tangent to the hypersurfaces given by $u^n = C$ for $C \in \mathbb{R}$. Hence we can flow away from the submanifold along ξ to obtain locally an n–dimensional submanifold $N \subset \Sigma \subset J^1(\mathbb{R}^n, \mathbb{R})$. (For sufficiently small open subsets $U \subset L$ and intervals I containing zero, the map $(x, t) \mapsto \mathrm{Fl}^\xi_t(x)$ is a diffeomorphism onto such a submanifold.) Moreover, the tangent space T_yN in the point $y := \mathrm{Fl}^\xi_t(x)$ is spanned by $\xi(y)$ and $T_x \mathrm{Fl}^\xi_t \cdot \eta$ for $\eta \in T_xL$. Now by construction $\alpha(\xi) = 0$ while

$$\alpha(y)(T_x \mathrm{Fl}^\xi_t \cdot \eta) = ((\mathrm{Fl}^\xi_t)^*(\alpha|_\Sigma))(\eta).$$

Now for $t = 0$, we know that $\alpha(\eta) = 0$ and $\frac{d}{dt}(\mathrm{Fl}^\xi_t)^*(\alpha|_\Sigma)(\eta) = (\mathrm{Fl}^\xi_t)^*(\mathcal{L}_\xi(\alpha))(\eta)$. Finally $\mathcal{L}_\xi \alpha = i_\xi d\alpha + d(i_\xi \alpha)$, and since we work on Σ, both these terms vanish. Thus N indeed is an integral submanifold for α and hence a Legendrean submanifold. It is also easy to see that locally around $x \in L$, the projection π_M restricts to a submersion on N, so this gives a solution of the PDE.

CHAPTER 2

Riemannian metrics and the Levi-Civita connection

In this chapter we start the discussion of Riemannian geometry. We will follow a modern version of the approach of E. Cartan, which emphasizes the point of view that Riemannian manifolds are "curved analogs" of Euclidean space. The advantage of this approach is that on the one hand it leads to a very conceptual description with nice formal properties. At the same time, one can easily convert this picture into very explicit form using orthonormal frames and coframes. Thus this approach is also known under the name "moving frames" which has to be handled with care, however, since it is also used for several other related techniques. We will start with some elementary considerations.

2.1. Riemannian metrics. The concept of a Riemannian metric is already known from the discussion of hypersurfaces in the first course on differential geometry.

Definition 2.1. (1) Let M be a smooth manifold. A *Riemannian metric* on M is a smooth $(0,2)$–tensor field g on M such that for each $x \in M$ the value $g_x : T_xM \times T_xM \to \mathbb{R}$ is a positive definite inner product.

(2) A *Riemannian manifold* (M, g) is a smooth manifold M together with a Riemannian metric g on M.

(3) Let (M, g^M) and (N, g^N) be Riemannian manifolds. An isometry $f : M \to N$ is a smooth map f such that $f^*g^N = g^M$. Otherwise put, for any point $x \in M$, the tangent map $T_xf : T_xM \to T_{f(x)}N$ is orthogonal with respect to the inner products g^M_x and $g^N_{f(x)}$.

From the definition, we see immediately what Riemannian metrics look like in local coordinates. Given a chart (U, u) on M, we can write $g|_U$ as $\sum_{i,j} g_{ij}du^i \otimes du^j$, where $g_{ij} = g(\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^j})$ is a smooth function for each $i, j = 1, \dots, n$. Of course, we have $g_{ji} = g_{ij}$ for all i and j by definition. For a point $x \in U$, we can consider the matrix $(g_{ij}(x))$ and this is exactly the matrix associated to the inner product g_x on the vector space T_xM with respect to the basis $\{\frac{\partial}{\partial u^i}(x)\}$ as usual in linear algebra. In particular, the matrix $(g_{ij}(x))$ is symmetric and positive definite. So we can alternatively view the collection g_{ij} of smooth functions on U as a function from U to the set of positive definite matrices of size n.

The symmetric $n \times n$–matrices form a vector space of dimension $n(n+1)/2$. It is well known from linear algebra, that positive matrices can be characterized as those symmetric matrices for which the determinants of all principal minors are positive. This implies that positive definite symmetric matrices form an open subset $\mathcal{S}_+(n)$ in the space of all symmetric matrices. Consequently, there are many smooth functions $g : \mathbb{R}^n \to \mathcal{S}_+(n)$. Any such function clearly defines a Riemannian metric on $\mathbb{R}^n$ (and hence on any open subset of $\mathbb{R}^n$) by simply endowing the tangent space $T_x\mathbb{R}^n = \mathbb{R}^n$ with the inner product defined by the matrix $g(x)$ for each $x \in \mathbb{R}^n$.

Given a smooth manifold M of dimension n and a chart (U, u) for M, there exists a Riemannian metric on the open subset $u(U) \subset \mathbb{R}^n$. Via the chart map u, we can pull this back to a tensor field g^U defined on the open subset $U \subset M$. Doing this for the elements U_i of an atlas for M and choosing a partition of unity f_i subordinate to this atlas, we can then form $g := \sum_i f_i g^{U_i}$ and this clearly is a smooth $(0,2)$–tensor field defined on all of M. Since each g^{U_i} is symmetric, g is obviously symmetric, too. Finally, given $x \in M$ and $0 \neq \xi \in T_xM$ we have $g_x(\xi,\xi) = \sum_i f_i(x) g_x^{U_i}(\xi,\xi)$. Now by construction, $f_i(x) \geq 0$ and at least one $f_i(x)$ must be positive. On the other hand, if $f_i(x) > 0$, then $g_x^{U_i}(\xi,\xi) > 0$, so we see that $g_x(\xi,\xi) > 0$. Thus g defines a Riemannian metric on M, and we see that any smooth manifold admits (many) Riemannian metrics.

If $f : M \to N$ is an isometry between Riemannian manifolds, then by definition $g^N_{f(x)}(T_xf \cdot \xi, T_xf \cdot \xi) = g^M_x(\xi,\xi)$ so this has to be positive if ξ is non–zero. In particular, any tangent map T_xf must be injective, so f is an immersion. If M and N have the same dimension (which is the case of main interest) then f automatically is a local diffeomorphism.

Having given a Riemannian metric g on a manifold M, one can immediately carry over the standard definition for the arc length of a smooth curve in $\mathbb{R}^n$. For an interval $I \subset \mathbb{R}$, a smooth curve $c : I \to M$ and $a, b \in I$, one defines $L_a^b(c) := \int_a^b \sqrt{g_{c(t)}(c'(t), c'(t))}dt$. This is evidently positive, unless c is constant. Further, it is invariant under reparametrizations of c and $L_a^b(c) = L_a^{b'}(c) + L_{b'}^b(c)$ for $a < b' < b \in I$. The concept of arc length clearly extends to piece–wise smooth curves.

This can then be used to define a *distance function* on M by defining $d(x,y)$ for $x, y \in M$ to be the infimum of the length of all piece–wise smooth curves starting at x and ending in y. Of course, $d(x,y) \geq 0$ and since we can run through curves in the opposite direction, we get $d(y,x) = d(x,y)$. Finally, since a piece–wise smooth curve from y to z can be joined to a piece–wise smooth curve from x to y to obtain a piece–wise smooth curve from x to z, we conclude that $d(x,z) \leq d(x,y) + d(y,z)$. A bit of thought is needed to see that $d(x,y) = 0$ is only possible for $x = y$ and hence (M,d) is a metric space in the sense of topology. Since we will prove stronger results implying this in the sequel, we only sketch the argument and leave the details as an exercise.

Fix a point $x \in M$ and choose a chart (U,u) for M such that $x \in U$, $u(x) = 0$ and $u(U)$ contains the closed unit ball of $\mathbb{R}^n$. Let B and S be the preimages of this unit ball respectively the unit sphere under u. First one observes that if $y \notin B$, then there is a point $z \in S$ such that $d(x,y) \geq d(x,z)$, since any curve emanating from x which leaves B has to pass through S. So it suffices to show that for $y \in B$, $d(x,y) = 0$ implies $y = x$, and this can be done in the chart. Let $g : u(U) \to \mathcal{S}_+(n)$ be the function describing the metric. The the function $(y,X) \mapsto g(y)(X,X)$ is continuous on $u(U) \times \mathbb{R}^n$ and hence attains a maximum and a minimum on the compact set $\bar{B}_1(0) \times S^{n-1}$, and the minimum has to be positive. Hence there are positive constants $A, B \in \mathbb{R}$ such that $A^2\|X\|^2 \leq g(y)(X,X) \leq B^2\|X\|^2$ holds for all $y \in \bar{B}_1(0)$ and all $X \in \mathbb{R}^n$. But this shows that for a smooth curve $c : [a,b] \to \bar{B}_1(0)$ which has length ℓ with respect to the standard inner product on $\mathbb{R}^n$, we have $A\ell \leq L_a^b(u \circ c) \leq B\ell$, and this implies the result.

The homogeneous model

2.2. We start the systematic study of Riemannian manifolds by looking at the simplest example, namely Euclidean space $\mathbb{R}^n$ with the Riemannian metric induced by the standard inner product on $T_x\mathbb{R}^n = \mathbb{R}^n$ for each $x \in \mathbb{R}^n$. Note that using the identity map as a chart, this metric is described by the constant function mapping $\mathbb{R}^n$ to the identity matrix in $\mathcal{S}_+(n)$. For this reason, we will denote this metric by δ, since in local coordinates it corresponds to the Kronecker delta δ_{ij}.

Now recall that a *Euclidean motion* of $\mathbb{R}^n$ is a map of the form $f(x) = Ax + b$, where $A \in O(n)$ is an orthogonal matrix and $b \in \mathbb{R}^n$ is a fixed vector. For this map, we evidently have $Df(x) = A$ for all x, so f is an isometry of the Riemannian manifold $(\mathbb{R}^n, \delta)$. Conversely, we can show

Proposition 2.2. *Let $U, V \subset \mathbb{R}^n$ be connected open subsets and suppose that $f : U \to V$ is an isometry between the Riemannian manifolds (U, δ) and (V, δ). Then f is the restriction to U of a Euclidean motion.*

Proof. By assumption, we have $\langle Df(x)(X), Df(x)(Y)\rangle = \langle X, Y\rangle$ for all $x \in U$ and $X, Y \in \mathbb{R}^n$. Now we have $\frac{d}{dt}|_{t=0} Df(x + tZ)(X) = D^2f(x)(Z, X)$ and this is symmetric in Z and X. On the other hand we know that $\langle Df(x + tZ)(X), Df(x + tZ)(Y)\rangle$ is constant and differentiating this at $t = 0$, we obtain

$$0 = \langle D^2f(x)(Z, X), Df(x)(Y)\rangle + \langle Df(x)(X), D^2f(x)(Z, Y)\rangle$$

Otherwise put, the map $(Z, X, Y) \mapsto \langle D^2f(x)(Z, X), Df(x)(Y)\rangle$ is symmetric in the first two arguments and skew symmetric in the last two arguments.

Claim: If $t : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is trilinear, symmetric in the first two and skew symmetric in the last two arguments, then $t = 0$.

Proof: For $X, Y, Z \in \mathbb{R}^n$ we compute:

$$\begin{aligned} t(X, Y, Z) = -t(X, Z, Y) = &- t(Z, X, Y) = t(Z, Y, X) \\ &= t(Y, Z, X) = -t(Y, X, Z) = -t(X, Y, Z). \end{aligned}$$

Now this shows that $\langle D^2f(x)(Z, X), Df(x)(Y)\rangle = 0$ for all Z, X, and Y, and since $Df(x)$ is a linear isomorphism, $D^2f(x)(Z, X) = 0$ for all Z and X. Thus $Df(x) = A$ for some fixed map A, which must be orthogonal by assumption. But then for any smooth curve $c : [0, 1] \to U$, we can compute

$$\begin{aligned} f(c(1)) =& f(c(0)) + \int_0^1 (f \circ c)'(t)dt = f(c(0)) + \int_0^1 (Ac'(t))dt \\ =& f(c(0)) + A(\int_0^1 c'(t)dt) = f(c(0)) + A((c(1) - c(0)) = Ac(1) + b, \end{aligned}$$

where $b = f(c(0)) - Ac(0)$. Since U is connected, we can fix $c(0) = u_0$ in U and then write any $x \in U$ as $c(1)$ for some c with $c(0) = u_0$, which completes the proof. □

Now we can get a nicer picture of Euclidean motions by realizing them as a matrix group. The idea here is rather easy. Consider $\mathbb{R}^n$ as the affine hyperplane $x_{n+1} = 1$ in $\mathbb{R}^{n+1}$. The action of an invertible matrix in $GL(n + 1, \mathbb{R})$ will preserve this hyperplane if and only if it has block form $\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}$ for $A \in GL(n, \mathbb{R})$ and $b \in \mathbb{R}^n$. Moreover, this matrix will map the point $\binom{x}{1}$ to $\binom{Ax+b}{1}$. Thus we can view the group $\mathrm{Euc}(n)$ of Euclidean motions of $\mathbb{R}^n$ as the subgroup of all matrices in $GL(n+1, \mathbb{R})$ which have the form $\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}$ with $A \in O(n)$ and $b \in \mathbb{R}^n$. We will also write (A, b) as a shorthand for this matrix. In this notation, the multiplication reads as $(A_1, b_1)(A_2, b_2) = (A_1A_2, b_1 + A_1b_2)$. Hence we see that we may identify $\mathrm{Euc}(n)$ (as a space) with the product $O(n) \times \mathbb{R}^n$.

Lemma 2.2. *(1) The group $O(n)$ is a smooth submanifold of dimension $\frac{n(n-1)}{2}$ of the vector space $M_n(\mathbb{R})$ of real $n \times n$–matrices.*

(2) For $A \in O(n)$ the tangent space $T_A O(n)$ is given by $\{X \in M_n(\mathbb{R}) : AX^t A = -X\}$. In particular, the tangent space at the unit matrix $\mathbb{I}$ is the space $\mathfrak{o}(n)$ of skew symmetric matrices.

(3) For any $X \in \mathfrak{o}(n)$ the map $L_X(A) := AX$ defines a vector field $L_X \in \mathfrak{X}(O(n))$. For each point $A \in O(n)$ the map $X \mapsto L_X(A)$ defines a linear isomorphism $\mathfrak{o}(n) \to T_A O(n)$.

(4) Putting $[X, Y] := XY - YX$ for $X, Y \in \mathfrak{o}(n)$, the vector fields from (3) satisfy $[L_x, L_Y] = L_{[X,Y]}$.

Proof. The space $\mathcal{S}(n)$ of symmetric matrices of size $n \times n$ is a real vector space of dimension $\frac{n(n+1)}{2}$. Clearly, $f(A) := A^t A - \mathbb{I}$ defines a smooth function $M_n(\mathbb{R}) \to \mathcal{S}(n)$ such that $O(n) = f^{-1}(\{0\})$. To see that this function is regular, observe that

$$Df(A)(B) = \tfrac{d}{ds}|_{s=0} f(A + sB) = B^t A + A^t B.$$

Now if $A \in O(n)$ and $C \in \mathcal{S}(n)$ is arbitrary, then putting $B = \frac{1}{2}AC$, we get $Df(A)(B) = \frac{1}{2}(C^t + C) = C$. This proves (1) and since $Df(A)(B) = 0$ is visibly equivalent to $AB^t A = -B$ for $A \in O(n)$ we obtain (2).

(3) Evidently, for fixed $X \in \mathfrak{o}(n)$, $A \mapsto AX$ defines a smooth map $M_n(\mathbb{R}) \to M_n(\mathbb{R})$, so this restricts to a smooth map on the submanifold $O(n)$. For $A \in O(n)$ one immediately verifies that $AX \in T_A O(n)$, so L_X indeed is a vector field. Moreover, for fixed $A \in O(n)$, $X \mapsto AX$ is a linear map $\mathfrak{o}(n) \to T_A(O(n))$ with inverse $Y \mapsto A^{-1}Y$, so we get the second claim.

(4) From the first course on differential geometry, we know that we can compute $[L_X, L_Y]$ as $L_X \cdot L_Y - L_Y \cdot L_X$, where we view the vector field being differentiated as a function with values in $M_n(\mathbb{R})$. Then $L_X \cdot L_Y(A)$ can be computed as

$$\tfrac{d}{ds}|_{s=0}(A + sAX)Y = AXY,$$

which immediately implies the result. □

Now we have an obvious projection $p : \mathrm{Euc}(n) \to \mathbb{R}^n$ defined by $f \mapsto f(0)$, so explicitly this is given by $(A, b) \mapsto b$. Since $\mathrm{Euc}(n)$ acts transitively on $\mathbb{R}^n$ and the stabilizer of 0 is the subgroup $O(n) \subset \mathrm{Euc}(n)$, this map induces a bijection between the set $\mathrm{Euc}(n)/O(n)$ of left cosets and $\mathbb{R}^n$. Identifying $\mathrm{Euc}(n)$ with $O(n) \times \mathbb{R}^n$ as a space, p is just the second projection and hence is smooth. Moreover $T_{(A,b)} \mathrm{Euc}(n)$ can be naturally identified with $T_A O(n) \times T_b \mathbb{R}^n = T_A O(n) \times \mathbb{R}^n$. In particular, the subspace $V_{(A,b)} \mathrm{Euc}(n) \subset T_{(A,b)} \mathrm{Euc}(n)$ defined as $V_{(A,b)} \mathrm{Euc}(n) := \ker(T_{(A,b)} p)$ is simply $T_A O(n) = T_A O(n) \times \{0\} \subset T_A O(n) \times \mathbb{R}^n$.

Next, for $i = 1, \dots, n$ we define $\theta^i \in \Omega^1(\mathrm{Euc}(n))$ by letting $\theta^i((A, b))(X, v)$ be the ith component of $A^{-1}v$. Likewise, for $i, j = 1, \dots, n$, we define $\gamma_i^j \in \Omega^1(\mathrm{Euc}(n))$ by letting $\gamma_j^i(A, b)(X, v)$ be the component in the ith row and jth column of the matrix $A^{-1}X$. From the proof of Lemma 2.2 above, we know that $A^{-1}X \in \mathfrak{o}(n)$ so $\gamma_i^j = -\gamma_j^i$ (and in particular $\gamma_i^i = 0$). Note that we can also view the vector $\theta = (\theta^1, \dots, \theta^n)$ of 1–forms as a 1–form with values in $\mathbb{R}^n$ and the matrix $\gamma = (\gamma_j^i)$ of 1–forms as a 1–form with values in $\mathfrak{o}(n)$. In this language, we can identify $V_{(A,b)} \mathrm{Euc}(n)$ with $\ker(\theta(A, b))$ and observe that $\theta(A, b)$ descends to a linear isomorphism $T_b \mathbb{R}^n \to \mathbb{R}^n$. Now these forms satisfy natural differential equations:

Theorem 2.2. *(1) For each $i = 1, \dots, n$ we have $d\theta^i = -\sum_j \gamma_j^i \wedge \theta^j$.*

(2) For all $i, j = 1, \dots, n$ we have $d\gamma_j^i = -\sum_k \gamma_k^i \wedge \gamma_j^k$.

PROOF. For $(X,v) \in \mathfrak{o}(n) \times \mathbb{R}^n$ and $(A,b) \in \mathrm{Euc}(n)$, we define $L_{(X,v)}(A,b) := (AX, Av) \in T_{(A,b)}\mathrm{Euc}(n)$. Clearly, this defines a smooth vector field on $\mathrm{Euc}(n)$ and by definition $\theta(L_{(X,v)}) = v$ and $\gamma(L_{(X,v)}) = X$. In particular, $(X,v) \mapsto L_{(X,v)}(A,b)$ defines a linear isomorphism $\mathfrak{o}(n) \times \mathbb{R}^n \to T_{(A,b)}\mathrm{Euc}(n)$. Consequently, it suffices to verify the claimed equations on two vector fields of this type. Moreover, using part (4) of Lemma 2.2 for the first component and a similar argument as in the proof of this part for the second component, one sees that $[L_{(X_1,v_1)}, L_{(X_2,v_2)}] = L_{(X_1X_2 - X_2X_1, X_1v_2 - X_2v_1)}$.

(1) We use the formula $d\theta^i(\xi,\eta) = \xi \cdot \theta^i(\eta) - \eta \cdot \theta^i(\xi) - \theta^i([\xi,\eta])$. Since $\theta(L_{(X,v)})$ is constant, this is true for each θ^i. Thus we see that

$$d\theta^i(L_{(X_1,v_1)}, L_{(X_2,v_2)}) = -\theta^i([L_{(X_1,v_1)}, L_{(X_2,v_2)}]),$$

so this is just the ith component of $X_1v_2 - X_2v_1$. But since the ith component of Xv is $\sum_j X^i_j v^j$, this can be written as

$$\sum_j \left(\gamma^i_j(L_{(X_1,v_1)})\theta^j(L_{(X_2,v_2)}) - \gamma^i_j(L_{(X_2,v_2)})\theta^j(L_{(X_1,v_1)})\right),$$

which completes the proof of (1), and (2) is done in the same way. □

Remark 2.2. What we have done here has a strong (Lie–) group theoretic background. The vector fields L_X in Lemma 2.2 and $L_{(X,v)}$ in Theorem 2.2 are *left–invariant* vector fields on the Lie groups $O(n)$ and $\mathrm{Euc}(n)$. The forms θ and γ are the components of the left Maurer Cartan form for the group $\mathrm{Euc}(n)$ and Theorem 2.2 is just the component–wise interpretation of the the Maurer–Cartan equation.

The orthonormal frame bundle

2.3. For an n–dimensional Riemannian manifold (M,g) we next want to obtain a description which formally looks very similar to the description of $\mathbb{R}^n$ as a homogeneous space of $\mathrm{Euc}(n)$ from 2.2. As a motivation, consider the $\mathbb{R}^n$–valued form θ from 2.2. We have observed that for $(A,b) \in \mathrm{Euc}(n)$ the value $\theta(A,b)$ induces a linear isomorphism $T_b\mathbb{R}^n \to \mathbb{R}^n$. One easily verifies that fixing b these linear isomorphisms exactly exhaust all orthogonal linear isomorphisms between $T_b\mathbb{R}^n$ and a fixed copy of $\mathbb{R}^n$.

For (M,g) as above and $x \in M$, we therefore let $\mathcal{P}_xM$ be the set of all linear isomorphisms $\varphi : \mathbb{R}^n \to T_xM$ which are orthogonal with respect to the standard inner product on $\mathbb{R}^n$ and the inner product g_x on T_xM. Then we let $\mathcal{P}M$ be the disjoint union of the spaces $\mathcal{P}_xM$ and we denote by $p : \mathcal{P}M \to M$ the map sending $\mathcal{P}_xM$ to x. Before we can prove that $\mathcal{P}M$ is a smooth manifold, we need one more notion.

Definition 2.3. Let (M,g) be an n–dimensional Riemannian manifold and let $U \subset M$ be an open subset.

(1) A *local orthonormal frame* for M over U is a family $\{\xi_1,\dots,\xi_n\}$ of vector fields $\xi_i \in \mathfrak{X}(U)$ such that for each $x \in U$ the tangent vectors $\xi_1(x),\dots,\xi_n(x)$ form an orthonormal basis for the inner product space (T_xM, g_x).

(2) A *local orthonormal coframe* for M over U is a family $\{\sigma^1,\dots,\sigma^n\}$ of one–forms $\sigma^i \in \Omega^1(U)$ such that for each $x \in M$ and tangent vectors $\xi,\eta \in T_xM$ we get $g_x(\xi,\eta) = \sum_{i=1}^n \sigma^i_x(\xi)\sigma^i_x(\eta)$.

The motivation for these notions is that one cannot hope to obtain coordinates adapted to a general Riemannian metric. Indeed, if one has a chart (U,u) such that the coordinate vector fields $\frac{\partial}{\partial u^i}$ are orthonormal in each point, then the chart map defines an isometry from $(U, g|_U)$ to $u(U)$ endowed with the Riemannian metric inherited from $\mathbb{R}^n$. From the first course on differential geometry, we know that in the case of surfaces in $\mathbb{R}^3$, existence of such an isometry implies vanishing of the Gauss–curvature. So in particular, coordinates like that cannot exist on the unit sphere $S^2 \subset \mathbb{R}^3$. However, we shall see soon that local orthonormal frames always exist. Orthonormal coframes are just the dual version of orthonormal frames, so they always exist, too.

Proposition 2.3. *Let (M,g) be an n–dimensional Riemannian manifold and consider $p : \mathcal{P}M \to M$ as defined above.*

(1) If (U,u) is a chart on M, then there is a local orthonormal frame for M over U.

(2) If $U \subset M$ is open, then a local orthonormal frame for M over U induces a bijection $\Phi : \mathcal{P}M \supset p^{-1}(U) \to U \times O(n)$ such that $pr_1 \circ \Phi = p$. Taking U to be a chart in M and combining with charts on $O(n)$, these bijections give rise to charts for $\mathcal{P}M$. These make $\mathcal{P}M$ into a smooth manifold such that $p : \mathcal{P}M \to M$ is smooth and a surjective submersion.

(3) Composing elements of $\mathcal{P}M$ with elements from $O(n)$ defines a smooth map $r : \mathcal{P}M \times O(n) \to \mathcal{P}M$. Putting $r^A(\varphi) := r(\varphi, A)$, one has $r^{\mathbb{I}} = \mathrm{id}$ and $r^{AB} = r^B \circ r^A$, so this is a right action of $O(n)$ on $\mathcal{P}M$. The orbits of this action are the fibers of $p : \mathcal{P}M \to M$.

Proof. (1) We just have to observe that the Gram–Schmidt orthonormalization procedure can be carried out in a smooth way. So we let $\partial_i := \frac{\partial}{\partial u^i}$ for $i = 1, \dots, n$ be the coordinate vector fields determined by (U,u). Then $g(\partial_1, \partial_1)$ is a positive smooth function on U, and so $\xi_1 := \frac{1}{g(\partial_1,\partial_1)}\partial_1$ defines a smooth vector field on U such that $g(\xi_1, \xi_1) = 1$. Then $g(\xi_1, \partial_2) : U \to \mathbb{R}$ is smooth, so $\tilde{\xi}_2 := \partial_2 - g(\xi_1, \partial_2)\xi_1$ is a smooth vector field on U such that $g(\xi_1, \xi_2) = 0$. Since pointwise, this is just standard Gram–Schmidt, we see that $\tilde{\xi}_2$ is nowhere vanishing on U, so $\xi_2 := \frac{1}{g(\tilde{\xi}_2,\tilde{\xi}_2)}\tilde{\xi}_2$ is well defined. Then one continues inductively as in the standard Gram–Schmidt procedure to obtain the result.

(2) Let $\{\xi_1, \dots, \xi_n\}$ be a local orthonormal frame on U. For $x \in U$ define $\sigma(x) : \mathbb{R}^n \to T_xM$ by $\sigma(x)(v^1, \dots, v^n) := \sum_{i=1}^n v^i \xi_i(x)$. This sends the standard basis of $\mathbb{R}^n$ to the orthonormal basis $\{\xi_i(x)\}$ of T_xM and hence is an orthogonal mapping. For $x \in U$ and $\varphi_x \in \mathcal{P}_xM$, the map $\sigma(x)^{-1} \circ \varphi_x : \mathbb{R}^n \to \mathbb{R}^n$ is an orthogonal linear map and hence an element of $O(n)$. So we can define $\Phi(\varphi_x) := (x, \sigma(x)^{-1} \circ \varphi_x)$ to obtain a map $\Phi : p^{-1}(U) \to U \times O(n)$ such that $p \circ \Phi = pr_1$. Conversely, for $A \in O(n)$ the map $\sigma(x) \circ A : \mathbb{R}^n \to T_xM$ is orthogonal so we can define $\Psi : U \times O(n) \to p^{-1}(U)$ by $\Psi(x, A) := \sigma(x) \circ A$. One immediately verifies that this is inverse to Φ, whence Φ is bijective.

If $\{\eta_i\}$ is a different orthonormal frame for M over U, then $a_i^j := g(\eta_i, \xi_j)$ defines a smooth function $U \to \mathbb{R}$ for $i, j = 1, \dots, n$. Moreover, $\eta_i = \sum_{j=1}^n a_i^j \xi_j$ for all i and since both the bases $\{\eta_i(x)\}$ and $\{\xi_j(x)\}$ are orthonormal with respect to g_x, the matrix $A(x) := (a_i^j(x))$ is orthogonal for all $x \in U$. If $\sigma : U \to p^{-1}(U)$ is associated to $\{\xi_j\}$ as above and $\tau : U \to p^{-1}(U)$ is the corresponding map for η, then for $v = (v^1, \dots, v^n) \in \mathbb{R}^n$, we get

$$\tau(x)(v) = \textstyle\sum_i v^i \eta_i(x) = \sum_{i,j} v^i a_i^j(x) \xi_j(x) = \sigma(x)(A(x)v),$$

where the upper index of a_i^j determines the row and the lower one determines the column. But this means that the bijections Φ and Ψ associated to the two frames are related as $\Psi(\varphi_x) = (\mathrm{id} \times \ell_{(A(x))^{-1}})(\Phi(\varphi_x))$, where for $B \in O(n)$ the map $\ell_B : O(n) \to O(n)$ is given by $\ell_B(C) = BC$. Since ℓ_B is the restriction of a linear map $M_n(\mathbb{R}) \to M_n(\mathbb{R})$ it is smooth, so Ψ is obtained by composing a diffeomorphism with Φ.

In particular, for any subset $W \subset \mathcal{P}M$, the image $\Phi(W \cap p^{-1}(U)) \subset U \times O(n)$ will be open if and only if the same is true for $\Psi(W \cap p^{-1}(U))$. Now we declare a subset of $\mathcal{P}M$ to be open if for any chart $U \in M$ and any local orthonormal frame with corresponding bijection Φ, the image $\Phi(W \cap p^{-1}(U))$ is open. It is easy to see that it suffices to verify this for the elements of an atlas, and restricting to a countable atlas, one concludes that this topology is separable and metrisable.

Finally, let (U, u) and Φ be as above and take a chart (V, v) for $O(n)$. Then $W := \Phi^{-1}(U \times V) \subset \mathcal{P}M$ is open and we can use $w := (u \times v) \circ \Phi : W \to u(U) \times v(V) \subset \mathbb{R}^{n(n+1)/2}$ as a chart for $\mathcal{P}M$. From what we have done so far, one easily concludes that the chart changes between two such charts are smooth, so we can use them to make $\mathcal{P}M$ into a manifold. In the chart (U, u) for M and the chart (W, w) for $\mathcal{P}M$, the map p simply corresponds to the first projection $u(U) \times v(V) \to u(U)$, so it is smooth.

(3) Since for an orthogonal map $\varphi_x : \mathbb{R}^n \to T_xM$ we have $\Phi(\varphi_x) = (x, \sigma(x)^{-1} \circ \varphi_x)$, the orthogonal map $\varphi_x \circ A$ for $A \in O(n)$ is mapped to $(x, \sigma(x)^{-1} \circ \varphi_x \circ A)$. Hence under Φ the map r simply corresponds to $\mathrm{id} \times \mu : U \times O(n) \times O(n) \to O(n)$ where μ is the matrix multiplication. Since μ is the restriction of a bilinear map to a submanifold, it is smooth. Hence also $r : \mathcal{P}M \times O(n) \to \mathcal{P}M$ is smooth. The two claimed properties of r just express the facts that $C\mathbb{I} = C$ and $C(AB) = (CA)B$ for all $C \in O(n)$. Finally if $\varphi_x, \psi_x : \mathbb{R}^n \to T_xM$ are orthogonal, then also $(\varphi_x)^{-1} \circ \psi_x : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal and thus defines an element $A \in O(n)$ such that $\psi_x = \varphi_x \circ A$. From this we immediately conclude that for any fixed φ_x, the map $A \mapsto \varphi_x \circ A$ defines a bijection between $O(n)$ and $\mathcal{P}_xM = p^{-1}(x)$. □

2.4. The soldering form. Similarly to the case of the cotangent bundle discussed in 1.7 and of the bundle of one–jets discussed in 1.13, the orthonormal frame bundle carries a tautological one–form. This can be either viewed as a form with values in $\mathbb{R}^n$ or as an n–tuple of ordinary one–forms.

A point $\varphi \in \mathcal{P}M$ by definition is a linear isomorphism $\mathbb{R}^n \to T_{p(\varphi)}M$ which is orthogonal for the standard inner product on $\mathbb{R}^n$ and the inner product g_x on T_xM. Given a tangent vector $\xi \in T_\varphi\mathcal{P}M$, we can first form $T_\varphi p \cdot \xi \in T_{p(\varphi)}M$ and then apply φ^{-1} to obtain an element in $\mathbb{R}^n$.

Definition 2.4. For $i = 1, \dots, n$, $\varphi \in \mathcal{P}M$ and $\xi \in T_\varphi\mathcal{P}M$, we define $\theta^i(\varphi)(\xi)$ as the ith component of $\varphi^{-1}(T_\varphi p \cdot \xi)$.

We will soon show that the θ^i are one–forms, and we will also view $\theta := (\theta^1, \dots, \theta^n)$ as a one–form on $\mathcal{P}M$ with values in $\mathbb{R}^n$. This is called the *soldering form* on $\mathcal{P}M$. The main properties of this form are the following:

Proposition 2.4. *(1) Each θ^i is a smooth one–form on $\mathcal{P}M$. For each $\varphi \in \mathcal{P}M$, the joint kernel $\{\xi \in T_\varphi\mathcal{P}M : \theta^i(\varphi)(\xi) = 0, i = 1, \dots, n\}$ coincides with $\ker(T_\varphi p)$.*

(2) For $A = (a^i_j) \in \mathcal{O}(n)$ let (b^i_j) be the inverse matrix A^{-1}. Then $(r^A)^*\theta^i = \sum_j b^i_j \theta^j$.

(3) Let $U \subset M$ be an open subset. Then the following three sets are in bijective correspondence:

(a) *The set of smooth maps $\sigma : U \to \mathcal{P}M$ such that $p \circ \sigma = \mathrm{id}_U$.*
(b) *The set of all local orthonormal frames for M defined over U.*
(c) *The set of all local orthonormal coframes for M defined over U.*

Here the bijection between (a) and (c) is implemented by mapping σ to $(\sigma^*\theta^1, \dots, \sigma^*\theta^n)$ while the bijection between (b) and (c) is given by pointwise passage to the dual basis.

PROOF. By definition, $\theta^i(\varphi)(\xi)$ vanishes for all i if and only if $\varphi^{-1}(T_\varphi p \cdot \xi) = 0$. Since φ is a linear isomorphism, this is equivalent to vanishing of $T_\varphi p \cdot \xi$, which proves the second part of (1).

Next, we evidently have $p \circ r^A = p$ for any $A \in O(n)$, and thus $T_{r^A(\varphi)}p \cdot T_{r^A} \cdot \xi = T_\varphi p \cdot \xi$ for all $\xi \in T_\varphi \mathcal{P}M$. Since $r^A(\varphi) = \varphi \circ A$, we conclude that $\theta^i(r^A(\varphi))(T_\varphi r^A \cdot \xi)$ is the ith component of $A^{-1} \circ \varphi^{-1}(T_\varphi p \cdot \xi)$. Since the components of $\varphi^{-1}(T_\varphi p \cdot \xi)$ are just $\theta^j(\varphi)(\xi)$, (2) follows.

Next, recall from Proposition 2.3 that for an open subset $U \subset M$, a local orthonormal frame $\{\xi_i\}$ defined on U gives rise to a smooth map $\sigma : U \to \mathcal{P}M$ such that $p \circ \sigma = \mathrm{id}_U$. This has to property that $\sigma(x) : \mathbb{R}^n \to T_x M$ maps the standard basis vector e_i to $\xi_i(x)$ for $i = 1, \dots, n$. Moreover, for any $\xi \in T_x M$ we get $T_{\sigma(u)}p \cdot T_x\sigma \cdot \xi = \xi$ from $p \circ \sigma = \mathrm{id}_U$. This implies that $\theta^i(\sigma(x))(T_x\sigma \cdot \xi)$ is the ith component of $\sigma(x)^{-1}(\xi)$, so in particular, this maps $\xi_j(x)$ to 1 for $j = i$ and to 0 for $j \neq i$. This implies that the elements $\sigma^*\theta^i(x)$ form a basis for $(T_x M)^*$, which is dual to the basis $\{\xi_i(x)\}$ for $T_x M$. Since this basis is orthonormal, we compute for $\xi = \sum_i a^i \xi_i(x)$ and $\eta = \sum_j b^j \xi_j(x)$

$$g_x(\xi, \eta) = \textstyle\sum_i a^i b^i = \sum_i \sigma^*\theta^i(x)(\xi)\sigma^*\theta^i(x)(\eta).$$

From Proposition 2.3 we further know that the frame $\{\xi_i\}$ gives rise to a diffeomorphism $\Psi : U \times O(n) \to p^{-1}(U)$. To prove smoothness of the forms θ^i it thus suffices to prove that $\Psi^*\theta^i$ is smooth for each i. Now we can identify $T(U \times O(n))$ with $TU \times TO(n)$, and since $p \circ \Psi = pr_1$, $Tp \circ T\Psi \cdot (\xi, \eta) = \xi$. In particular, the local frame ξ_i over U can be completed by a local frame for $O(n)$ to a local frame for $U \times O(n)$. Since the frame elements from $O(n)$ are killed by $Tp \circ T\Psi$, it suffices to prove that $(\Psi^*\theta^i)(\xi_j, 0)$ is a smooth function for all i and j. But by definition $(\Psi^*\theta^i)(x, A)(\xi_j, 0)$ is the ith component of $\Psi(x, A)^{-1}(\xi_j) = A^{-1} \circ \sigma(x)^{-1}(\xi_j(x))$. From above, we know that $\sigma(x)^{-1}(\xi_j(x)) = e_j$. Applying A^{-1} to this, we get the jth column of A^{-1}, so the ith entry of this is just one of the matrix entries of A^{-1}. Expressing this entry as a quotient of determinants using Cramer's rule, we conclude that it depends smoothly on A, which completes the proof of (1).

Knowing that the θ^i are smooth one–forms on $\mathcal{P}M$, the above computations show that for the map σ constructed from the frame $\{\xi_i\}$ the forms $\sigma^i := \sigma^*\theta^i \in \Omega^1(U)$ form an orthonormal coframe over U, which is dual to $\{\xi_j\}$. Conversely, having given a local coframe σ^i over U, we can pointwise form the dual basis, which fits together to define smooth vector fields ξ_j over U, which by construction constitute a local orthonormal frame. Then the above computations show that the the resulting smooth map $\sigma : U \to \mathcal{P}M$ with $p \circ \sigma = \mathrm{id}_U$ has the property that $\sigma^*\theta^i = \sigma^i$ for all i. This establishes the bijections claimed in (3). □

The Levi–Civita connection

2.5. Connection forms. Having found an analog of the forms θ^j on the orthonormal frame bundle of a general Riemannian manifold, it is natural to ask whether also the forms γ^i_j from 2.2 have a natural analog in this general setting. It turns out that this is the case, but maybe in a slightly unexpected way, namely via the differential equation from part (1) of Theorem 2.2. Before we can show this, we need a bit of preparation.

Let (M, g) be a Riemannian manifold of dimension n and let $\mathcal{P}M$ be its orthonormal frame bundle. For a point $x \in M$ and a point $\varphi \in \mathcal{P}M$ lying over x, we know that the joint kernel of the values $\theta^i(x)$ is the so–called *vertical subspace* $\ker(T_\varphi p) \subset T_\varphi \mathcal{P}M$. On the other hand, we can look at the smooth map $r_\varphi : O(n) \to \mathcal{P}M$ defined by $r_\varphi(A) := r(\varphi, A) = \varphi \circ A$. Via the natural charts from 2.3 with values in $U \times O(n)$, the map r_φ corresponds to the map $A \mapsto (x, BA)$ for some fixed B, while p corresponds to the projection on the first factor. Thus we see that r_φ is a diffeomorphism from $O(n)$ onto the fiber $p^{-1}(x)$. Consequently, $T_{\mathbb{I}} r_\varphi : \mathfrak{o}(n) \to T_\varphi \mathcal{P}M$ is an injective linear map which by construction has values in $\ker(T_\varphi p)$. By dimensional considerations r_φ must be a linear isomorphism, and for $X \in \mathfrak{o}(n)$, we put $\zeta_X(\varphi) := T_{\mathbb{I}} r_\varphi \cdot X$.

Definition 2.5. A *connection form* on $\mathcal{P}M$ is a family $\gamma^i_j \in \Omega^1(\mathcal{P}M)$ of one–forms such that $\gamma^j_i = -\gamma^i_j$ and

- (i) For $A = (a^i_j) \in O(n)$ with $A^{-1} = (b^i_j)$ we have $(r^A)^*\gamma^i_j = \sum_{k,\ell} b^i_k a^\ell_j \gamma^k_\ell$.
- (ii) For any $\varphi \in \mathcal{P}M$ and any $X \in \mathfrak{o}(n)$, $\gamma^i_j(\zeta_X(\varphi))$ is the component x^i_j of the matrix X.

The conditions in the definition are most naturally expressed via viewing $\gamma = (\gamma^i_j)$ as a matrix valued one–form rather than a matrix of one–forms. Then $\gamma^j_i = -\gamma^i_j$ says that the values are in $\mathfrak{o}(n) \subset M_n(\mathbb{R})$. The condition (ii) then says that the restriction of $\gamma(\varphi)$ to $\ker(T_\varphi p) \subset T_\varphi \mathcal{P}M$ should be the inverse of the canonical isomorphism $T_{\mathbb{I}} r_\varphi$. To interpret (i), we observe that for $A \in O(n)$ and $X \in \mathfrak{o}(n)$, we have $AXA^{-1} = AXA^t \in \mathfrak{o}(n)$. Using this, condition (i) just says that $(r^A)^*\gamma = A^{-1}\gamma A$.

Lemma 2.5. *Let (M, g) be a Riemannian manifold with orthonormal frame bundle $\mathcal{P}M$.*

(1) There exists a connection form γ^i_j on $\mathcal{P}M$.

(2) Let γ^i_j be a fixed connection form on $\mathcal{P}M$ and suppose that for $i, j, k = 1, \dots, n$ we take smooth functions ${\Psi_k}^i_j : \mathcal{P}M \to \mathbb{R}$ such that ${\Psi_k}^j_i = -{\Psi_k}^i_j$ and such that for each matrix $A = (a^s_r) \in O(n)$ with inverse $A^{-1} = (b^s_r)$ we have ${\Psi_k}^i_j \circ r^A = \sum_{q,r,s} b^i_q a^r_j a^s_k {\Psi_s}^q_r$. Then $\hat{\gamma}^i_j = \gamma^i_j + \sum_k {\Psi_k}^i_j \theta^k$ is also a connection form on $\mathcal{P}M$ and all connection forms are obtained in this way.

Proof. (1) Let $U \subset M$ be open, suppose that $\{\xi_1, \dots, \xi_n\}$ is a local orthonormal frame for M defined on U, and let $\Phi : p^{-1}(U) \to U \times O(n)$ be the induced diffeomorphism from Proposition 2.3. Then $T_{(x,A)}(U \times O(n)) = T_x U \times T_A O(n)$ and from part (2) of Lemma 2.2, we know that $T_A O(n) = \{Y \in M_n\mathbb{R} : AY^tA = -Y\}$. Now we define $\tilde{\gamma}^i_j(x, A)(\xi, Y) = c^i_j$, where $(c^i_j) = A^{-1}Y$. This is evidently a smooth one–form on $U \times O(n)$ so $\gamma^i_j := \Phi^*\tilde{\gamma}^i_j \in \Omega^1(p^{-1}(U))$. From 2.2 we know that $A^{-1}Y$ is skew symmetric for any $Y \in T_A O(n)$, so $\gamma^j_i = -\gamma^i_j$.

The smooth map $(x, B) \mapsto (x, BA)$ for fixed $A \in O(n)$ evidently has derivative $(\xi, Y) \mapsto (\xi, YA)$, since multiplication from the right by A is the restriction of a linear map. Now by definition $\tilde{\gamma}^i_j(x, BA)(\xi, YA)$ is the component c^i_j of the matrix $(c^i_j) = (BA)^{-1}YA = A^{-1}(B^{-1}Y)A$. This immediately implies that the $\tilde{\gamma}^i_j$ satisfy condition (i). On the other hand, fix (x, B) and consider the map $A \mapsto (x, BA)$. Since multiplication from the left by B is the restriction of a linear map, the derivative in $\mathbb{I}$ of this map is given by $X \mapsto (0, BX)$ for $X \in \mathfrak{o}(n)$. But this shows that the γ^i_j also satisfy condition (i), and hence define a connection form on $p^{-1}(U) = \mathcal{P}U$.

Now we can consider a covering of M by open subsets U_α over which there are local orthonormal frames, and hence local connection forms. Then one takes a partition of unity subordinate to this covering and uses this to glue together these local connection forms to a globally defined one–forms. A moment of thought shows that the defining properties of connection forms are preserved in this process.

(2) From the definition of $\hat{\gamma}^i_j$ it is evident that $\hat{\gamma}^j_i = -\hat{\gamma}^i_j$ and since $\theta^k(\zeta_X) = 0$ for all $X \in \mathfrak{o}(n)$ and all k, we see that $\hat{\gamma}^i_j(\zeta_X) = \gamma^i_j(\zeta_X)$, so $\hat{\gamma}^i_j$ satisfies property (ii) from

Definition 2.5. Finally, $(r^A)^*\hat{\gamma}^i_j = (r^A)^*\gamma^i_j + \sum_k (\Psi_k{}^i_j \circ r^A)(r^A)^*\theta^k$. By our assumption on the functions $\Psi_k{}^i_j$ and part (2) of Proposition 2.4, the last term is given by

$$\textstyle\sum_{k,q,r,s,t} b^i_q a^r_j a^s_k b^k_t \Psi_s{}^q_r \theta^t = \sum_{q,r,s} b^i_q a^r_j \Psi_s{}^q_r \theta^s,$$

which shows that property (i) is satisfied, too.

It remains to show that any connection form on $\mathcal{P}M$ can be obtained in this way starting from a fixed connection form γ^i_j. So let us assume that $\hat{\gamma}^i_j$ is any connection form, and consider the difference $\Phi^i_j := \hat{\gamma}^i_j - \gamma^i_j \in \Omega^1(\mathcal{P}M)$. Of course we have $\Phi^j_i = -\Phi^i_j$, and since both summands have property (i) from Definition 2.5, we see that for $A = (a^r_s) \in O(n)$ with $A^{-1} = (b^r_s)$ we get $(r^A)^*\Phi^i_j = \sum_{r,s} b^i_r a^s_j \Phi^r_s$.

The fact that both γ and $\hat{\gamma}$ have property (ii) from definition 2.5 implies that $\Phi^i_j(\zeta_X) = 0$ for all $i, j = 1, \dots, n$ and all $X \in \mathfrak{o}(n)$. This implies that for each $\varphi \in \mathcal{P}M$ the linear map $\Phi^i_j(\varphi) : T_\varphi\mathcal{P}M \to \mathbb{R}$ vanishes on the subspace $\ker(T_\varphi p)$. But by part (1) of Proposition 2.4, $\theta(\varphi) = (\theta^1(\varphi), \dots, \theta^n(\varphi))$ induces a linear isomorphism $T_\varphi\mathcal{P}M/\ker(T_\varphi p) \to \mathbb{R}^n$. But this shows that for fixed i and j, there are n uniquely determined real numbers, which we denote by $\Psi_1{}^i_j(\varphi), \dots, \Psi_n{}^i_j(\varphi)$ such that $\Phi^i_j(\varphi) = \sum_k \Psi_k{}^i_j(\varphi)\theta^k(\varphi)$.

This then defines functions $\Psi_k{}^i_j : \mathcal{P}M \to \mathbb{R}$, and a moment of thought shows that these functions are smooth. From the construction it is evident that $\Psi_k{}^j_i = -\Psi_k{}^i_j$ for all i, j, and k, so to complete the proof, it suffices to compute $\Psi_k{}^i_j \circ r^A$. But by construction we have

$$\textstyle\sum_{q,r} b^i_q a^r_j \Phi^q_r = (r^A)^*\Phi^i_j = \sum_k (\Psi_k{}^i_j \circ r^A)(r^A)^*\theta^k = \sum_{k,s} (\Psi_k{}^i_j \circ r^A) b^k_s \theta^s,$$

and the left hand side can be expanded as $\sum_{q,r,s} b^i_q a^r_j \Psi_s{}^q_r \theta^s$. Evaluating in a point, the θ^s become linearly independent linear functionals on a finite dimensional space while all other objects on each side just represent a real factor, so we conclude that $\sum_{q,r} b^i_q a^r_j \Psi_s{}^q_r = \sum_k (\Psi_k{}^i_j \circ r^A) b^k_s$. Multiplying both sides by a^s_t and summing over s, we obtain the claimed expression for $\Psi_k{}^i_j \circ r^A$. □

From the proof we conclude that for $X \in \mathfrak{o}(n)$ the tangent vector $\zeta_X(\varphi)$ corresponds to $(0, BX) \in T_{(x,B)}U \times O(n)$ in a natural chart. This immediately implies that $\varphi \mapsto \zeta_X(\varphi)$ is a smooth mapping and thus defines a vector field $\zeta_X \in \mathfrak{X}(\mathcal{P}M)$. This is called the *fundamental vector field* generated by $X \in \mathfrak{o}(n)$. The defining property (ii) of a connection form $\gamma = (\gamma^i_j)$ can be nicely phrased as $\gamma(\zeta_X) = X$.

Given $X \in \mathfrak{o}(n)$, one can form the matrix exponential $\exp(tX) = \sum_{k=0}^{\infty} \frac{t^k}{k!} X^k$ for each $t \in \mathbb{R}$. One can verify directly that $\exp(tX) \in O(n)$ for all t and that $\exp((t+s)X) = \exp(tX)\exp(sX)$ for all $t, s \in \mathbb{R}$. The latter equation easily implies that for any $B \in O(n)$, the curve $c(t) := B \exp tX$ is the unique solution of the differential equation $c'(t) = c(t)X$ with initial condition $c(0) = B$. This in turn shows that the flow of the fundamental vector field ζ_X is given by $\mathrm{Fl}^{\zeta_X}_t = r^{\exp(tX)}$.

Now consider the soldering form $\theta = (\theta^j)$. From part (2) of Proposition 2.4 we then see that

$$(\mathrm{Fl}^{\zeta_X}_t)^*\theta^i = \textstyle\sum_j a^i_j(-t)\theta^j,$$

where $(a^i_j(t)) := \exp(tX)$. Differentiating this with respect to t at $t = 0$, we get $\mathcal{L}_{\zeta_X}\theta^i = -\sum_j (a^i_j)'(0)\theta^j$, and of course $((a^i_j)'(0)) = X$. Now we can expand the left hand side of this equation as $i_{\zeta_X} d\theta^i + i_{\zeta_X}\theta^i$ and the since $\theta^i(\zeta_X) = 0$, the second term vanishes.

Now suppose that γ^i_j is an arbitrary connection form on $\mathcal{P}M$. Then for each i, we can consider $d\theta^i + \sum_j \gamma^i_j \wedge \theta^j \in \Omega^2(\mathcal{P}M)$. Inserting the vector field ζ_X into this two–form, and using that $\theta^j(\zeta_X) = 0$ for all j, we get $i_{\zeta_X} d\theta^i + \sum_j \tilde{\gamma}^i_j(\zeta_X)\theta^j$. Since $\gamma^i_j(\zeta_X)$ equals the appropriate entry of X,we see from above that this vanishes.

Now fix a point $\varphi \in \mathcal{P}M$. Then the tangent vectors $\zeta_X(\varphi)$ for $X \in \mathfrak{o}(n)$ form the vertical subspace $\ker(T_\varphi p) \subset T_\varphi \mathcal{P}M$. Hence the bilinear map

$$d\theta^i(\varphi) + \textstyle\sum_j \gamma^i_j(\varphi) \wedge \theta^j(\varphi) : T_\varphi \mathcal{P}M \times T_\varphi \mathcal{P}M \to \mathbb{R}$$

vanishes if one of its entries is from $\ker(T_\varphi p)$. As in the proof of Proposition 2.5, this implies that there are uniquely determined real numbers $T^i_{jk}(\varphi)$ such that $T^i_{kj}(\varphi) = -T^i_{jk}(\varphi)$ and such that

$$d\theta^i(\varphi) + \textstyle\sum_j \gamma^i_j(\varphi) \wedge \theta^j(\varphi) = \sum_{jk} T^i_{jk}(\varphi)\theta^j(\varphi) \wedge \theta^k(\varphi).$$

These numbers fit together to define smooth functions $T^i_{jk} : \mathcal{P}M \to \mathbb{R}$ such that

$$d\theta^i + \textstyle\sum_j \gamma^i_j \wedge \theta^j = \sum_{j,k} T^i_{jk}\theta^j \wedge \theta^k. \tag{1}$$

The functions T^i_{jk} are called the *torsion–coefficients* of the connection form γ^i_j.

It is also easy to compute $T^i_{jk} \circ r^A$ for a matrix $A = (a^r_s) \in O(n)$ with inverse $A^{-1} = (b^r_s)$: Pulling back the right hand side of (1) along r^A, we obtain

$$\sum_{j,k} (T^i_{jk} \circ r^A)(r^A)^*\theta^j \wedge (r^A)^*\theta^k = \sum_{j,k,q,r} (T^i_{jk} \circ r^A) b^j_q b^k_r \theta^q \wedge \theta^r.$$

For the first term in the left hand side, we compute

$$(r^A)^* d\theta^i = d(r^A)^*\theta^i = d(\textstyle\sum_s b^i_s \theta^s) = \sum_q b^i_q d\theta^q,$$

while for the other term the analogous behavior follows from property (i) in Definition 2.5 and part (2) of Proposition 2.4. Expanding the result via (1), we obtain $\sum_{q,r,s} b^i_s T^s_{qr} \theta^q \wedge \theta^r$. Now taking into account that both $\sum_s b^i_s T^s_{qr}$ and $\sum_{j,k} (T^i_{jk} \circ r^A) b^j_q b^k_r$ are skew symmetric, we can evaluate in a point. There the elements $\theta^q(\varphi) \wedge \theta^r(\varphi)$ for $q < r$ are all linearly independent, and as in the proof of Proposition 2.5, we conclude that

$$T^i_{jk} \circ r^A = \sum_{q,r,s} b^i_q a^r_j a^s_k T^q_{rs}.$$

2.6. The Levi–Civita connection form. We are now ready to prove the existence of a canonical connection form on $\mathcal{P}M$. The key to this is computing the effect of a change of connection on the torsion coefficients. To get a clearer picture of what is going on, we briefly discuss the meaning of the functions ${\Psi_k}^i_j$ from Lemma 2.5 and of the torsion coefficients T^i_{jk} before proceeding. We have already observed that it is most natural to view the soldering form $\theta = (\theta^1, \dots, \theta^n)$ as a one–form with values in $\mathbb{R}^n$ and a connection form $\gamma = (\gamma^i_j)$ as a one–form with values in $\mathfrak{o}(n)$. Thus it is clear that for each $\varphi \in \mathcal{P}M$ the values ${\Psi_k}^i_j(\varphi)$ are most naturally viewed as describing a linear function $\mathbb{R}^n \to \mathfrak{o}(n)$ with the indices i and j denoting the matrix entries in $\mathfrak{o}(n)$. Likewise, the values $T^i_{jk}(\varphi)$ describe a skew symmetric bilinear map $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$.

Notice that both on the space $L(\mathbb{R}^n, \mathfrak{o}(n))$ of linear maps and on the space $L^2_{alt}(\mathbb{R}^n, \mathbb{R}^n)$ of skew symmetric bilinear maps, there is a natural action of the group $O(n)$. For $A \in O(n)$ and $\Psi : \mathbb{R}^n \to \mathfrak{o}(n)$, and $T : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ these are given by $(A \cdot \Psi)(v) :=$

$A\Psi(A^{-1}v)A^{-1}$ and $(A \cdot T)(v,w) := A(T(A^{-1}v, A^{-1}w)))$, respectively. In this language, the formulae in Lemma 2.5 respectively in the end of section 2.5 simply read as $\Psi \circ r^A = A^{-1} \cdot \Psi$ and $T \circ r^A = A^{-1} \cdot T$.

Lemma 2.6. *Let (M,g) be a Riemannian manifold with orthonormal frame bundle $\mathcal{P}M$.*

(1) Let γ^i_j be a connection form on $\mathcal{P}M$, take functions ${\Psi_k}^i_j : \mathcal{P}M \to \mathbb{R}$ as in Lemma 2.5 and consider the connection form $\hat{\gamma}^i_j = \gamma^i_j + \sum_k {\Psi_k}^i_j \theta^k$. Then the torsion coefficients of the two connections are related as $\hat{T}^i_{jk} = T^i_{jk} + \frac{1}{2}({\Psi_j}^i_k - {\Psi_k}^i_j)$.

(2) The map $\partial : L(\mathbb{R}^n, \mathfrak{o}(n)) \to L^2_{alt}(\mathbb{R}^n, \mathbb{R}^n)$ defined by $\partial(\Psi)(v,w) = \Psi(v)w - \Psi(w)v$ is a linear isomorphism and $\partial(A \cdot \Psi) = A \cdot (\partial(\Psi))$ for all $A \in O(n)$.

Proof. (1) By definition, we have $d\theta^i + \sum_j \hat{\gamma}^i_j \wedge \theta^j = \sum_{k,\ell} \hat{T}^i_{k\ell}\theta^k \wedge \theta^\ell$, and inserting for $\hat{\gamma}^i_j$ in the left hand side, we obtain

$$d\theta^i + \sum_j \gamma^i_j \wedge \theta^j + \sum_{j,k} {\Psi_k}^i_j \theta^k \wedge \theta^j = \sum_{j,k} (T^i_{jk} - {\Psi_k}^i_j)\theta^j \wedge \theta^k.$$

Taking into account that $\hat{T}^i_{jk}$ must be skew symmetric in the two lower indices, we conclude that

$$\hat{T}^i_{jk} = \tfrac{1}{2}\left(T^i_{jk} - {\Psi_k}^i_j - T^i_{kj} + {\Psi_j}^i_k\right),$$

and since the $T^i_{kj} = -T^i_{jk}$, the result follows.

(2) We know that $\dim(\mathfrak{o}(n)) = \frac{n(n-1)}{2}$, so the space $L(\mathbb{R}^n, \mathfrak{o}(n))$ has dimension $\frac{n^2(n-1)}{2}$. Likewise, to specify a skew symmetric bilinear map $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, one has to specify the $\frac{n(n-1)}{2}$ values it takes on (e_i, e_j) for $i < j$. Thus also the space $L^2_{alt}(\mathbb{R}^n, \mathbb{R}^n)$ has dimension $\frac{n^2(n-1)}{2}$. Moreover,

$$\begin{aligned} \partial(A \cdot \Psi)(v,w) &= (A \cdot \Psi)(v)(w) - (A \cdot \Psi)(w)(v) \\ &= A\Psi(A^{-1}v)A^{-1}w - A\Psi(A^{-1}w)A^{-1}v = A(\partial(\Psi)(A^{-1}v, A^{-1}w)), \end{aligned}$$

so the last statement follows. To complete the proof, it therefore suffices to show that ∂ is injective. But if $\Psi : \mathbb{R}^n \to \mathfrak{o}(n)$ is a linear map such that $\partial(\Psi) = 0$ then we have $\Psi(v)w = \Psi(w)v$ for all $v, w \in \mathbb{R}^n$. But then consider the map $t : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by $t(u,v,w) := \langle \Psi(u)v, w\rangle$. Then by assumption $t(u,v,w) = t(v,u,w)$ and since $\Psi(u) \in \mathfrak{o}(n)$ we get $t(u,w,v) = \langle \Psi(u)w, v\rangle = -\langle w, \Psi(u)v\rangle = -t(u,v,w)$. But from the proof of Proposition 2.2 we know that $t = 0$, which immediately implies $\Psi = 0$. □

Theorem 2.6. *Let (M,g) be a Riemannian manifold. Then there exists a unique connection form γ^i_j on the orthonormal frame bundle $\mathcal{P}M$ whose connection coefficient vanish identically, i.e. such that $d\theta^i + \sum_j \gamma^i_j \wedge \theta^j = 0$. This is called the Levi–Civita connection form of M.*

Proof. By Proposition 2.5 there exists connection forms on $\mathcal{P}M$, we let $\hat{\gamma}^i_j$ be any such form and take $\hat{T}^i_{jk}$ be its connection coefficients. View $\hat{T}^i_{jk}$ as defining a smooth function $\mathcal{P}M \to L^2_{alt}(\mathbb{R}^n, \mathbb{R}^n)$ and compose $(\partial)^{-1} : L^2_{alt}(\mathbb{R}^n, \mathbb{R}^n) \to L(\mathbb{R}^n, \mathfrak{o}(n))$ with $(-2\hat{T}^i_{jk})$ to obtain a collection of smooth functions ${\Psi_k}^i_j$ on $\mathcal{P}M$ such that $\frac{1}{2}({\Psi_j}^i_k - {\Psi_k}^i_j) = -\hat{T}^i_{jk}$.

Since $\partial(A^{-1} \cdot \Psi) = A^{-1} \cdot \partial(\Psi)$ and $T \circ r^A = A^{-1} \cdot T$ we conclude that $\Psi \circ r^A = A^{-1} \cdot \Psi$. Thus Proposition 2.5 shows that $\gamma^i_j = \hat{\gamma}^i_j + \sum_k {\Psi_k}^i_j \theta^k$ is a connection form on $\mathcal{P}M$, which by part (1) of Lemma 2.5 has vanishing torsion coefficients, so existence is proved.

For uniqueness we just have to observe that by Proposition 2.5, any other connection form $\tilde{\gamma}^i_j$ can be written as $\gamma^i_j + \sum_k \tilde{\Psi}_k{}^i_j \theta^k$. But by part (1) of Lemma 2.6, the torsion coefficients of this connection form are given by $\tilde{T}^i_{jk} = \frac{1}{2}(\tilde{\Psi}_j{}^i_k - \tilde{\Psi}_k{}^i_j)$, which by part (2) of that Lemma vanishes only if all $\tilde{\Psi}_k{}^i_j$ vanish. □

We will discuss in detail in the next chapter how this canonical connection form can be used. Before we do that, we discuss how these fairly abstract considerations can be made explicit.

2.7. The Levi–Civita connection in a frame. The basic strategy to convert considerations on the orthonormal frame bundle into explicit formulae follows easily from Proposition 2.4. This says that a local orthonormal frame respective a local orthonormal coframe is equivalent to a local section of the orthonormal frame bundle $\mathcal{P}M$. We also know already that we can effectively construct local orthonormal frames and coframes using the Gram–Schmidt procedure. Now having given a local section of $\mathcal{P}M$, we can pull back functions and differential forms on $\mathcal{P}M$ to locally defined objects of the same type on M. In particular, the components γ^i_i of the Levi–Civita connection form can be pulled back to one–forms on the domain of our local section, and we can derive a nice characterization for these forms.

Proposition 2.7. *Let (M, g) be a Riemannian manifold of dimension n and let $U \subset M$ be open subset. Suppose that there is a local orthonormal coframe $\{\sigma^1, \dots, \sigma^n\}$ defined on U and let $\sigma : U \to \mathcal{P}M$ be the corresponding smooth section.*

Then the pullback $\omega^i_j := \sigma^*\gamma^i_j \in \Omega^1(U)$ of the Levi–Civita connection form is uniquely characterized by the facts that $\omega^j_i = -\omega^i_j$ for all i, j and that

$$0 = d\sigma^i + \textstyle\sum_j \omega^i_j \wedge \sigma^j.$$

Proof. From Proposition 2.4, we know that the section $\sigma : U \to \mathcal{P}M$ is characterized by $\sigma^*\theta^i = \sigma^i$ for $i = 1, \dots, n$. Putting $\omega^i_j := \sigma^*\gamma^i_j$ for all i, j, we clearly have $\omega^j_i = -\omega^i_j$. Moreover, on $\mathcal{P}M$ we have $0 = d\theta^i + \sum_j \gamma^i_j \wedge \theta^j$ and applying σ^* to this equation we conclude that $0 = d\sigma^i + \sum_j \omega^i_j \wedge \sigma^j$.

So it remains to see that the ω^i_j are uniquely determined by these properties. Since the forms σ^k form a local coframe, there are unique smooth functions $\psi_k{}^i_j$ such that $\omega^i_j = \sum_k \psi_k{}^i_j \sigma^k$. Using this, we see that $\sum_j \omega^i_j \wedge \sigma^j = \sum_{j,k} \psi_k{}^i_j \sigma^k \wedge \sigma^j$, which by skew symmetry of the wedge product can be written as $\sum_{j,k} \frac{1}{2}(\psi_j{}^i_k - \psi_k{}^i_j)\sigma^j \wedge \sigma^k$. Hence part (2) of Lemma 2.6 implies that the map $(\omega^i_j(x)) \mapsto \sum_j (\omega^i_j(x) \wedge \sigma^j(x))$ is injective on elements satisfying $\omega^j_i(x) = -\omega^i_j(x)$. □

Examples: (1) Flat space: In flat Euclidean space $\mathbb{R}^n$ the standard coordinate vector fields ∂_i form a global orthonormal frame. The dual coframe is simply given by $\sigma^i = dx^i$ for $i = 1, \dots, n$. Since $d\sigma^i = 0$ for all i, we conclude that the connection forms ω^i_j vanish identically in this frame.

(2) The sphere: Let us consider the unit sphere $S^n := \{x \in \mathbb{R}^{n+1} : \langle x, x \rangle = 1\}$ with the induced Riemannian metric. To get simple formulae, we use a particularly nice chart, the *stereographic projection*. Let $N = e_{n+1}$ be the north pole, put $U := S^n \setminus \{N\}$ and define $u : U \to \mathbb{R}^n$ by

$$u(x) = u(x^1, \dots, x^{n+1}) = \tfrac{1}{1-x^{n+1}}(x^1, \dots, x^n)$$

(To interpret this geometrically, one views $\mathbb{R}^n$ as the affine hyperplane through $-N$ which is orthogonal to N and one maps each point $x \in S^n$ to the intersection of the ray from N through x with that affine hyperplane.) One immediately verifies that the map

$$(y^1, \dots, y^n) \mapsto \tfrac{1}{\langle y,y\rangle+1}(2y^1, \dots, 2y^n, \langle y, y\rangle - 1)$$

is inverse to U. The ith partial derivative of this mapping is given by

$$\tfrac{-2y^i}{(1+\langle y,y\rangle)^2}(2y, \langle y, y\rangle - 1) + \tfrac{1}{1+\langle y,y\rangle}(2e_i, 2y^i),$$

which shows that

$$\tfrac{\partial}{\partial u^i} = \tfrac{1}{(1+\langle y,y\rangle)^2}\left(\textstyle\sum_{j=1}^n -4y^i y^j \tfrac{\partial}{\partial x^j} + 2(1+\langle y,y\rangle)\tfrac{\partial}{\partial x^i} + 4y^i \tfrac{\partial}{\partial x^{n+1}}\right).$$

Using that the vectors $\frac{\partial}{\partial x^j}$ are orthonormal, one immediately verifies that $\frac{\partial}{\partial u^i}$ is orthogonal to $\frac{\partial}{\partial u^k}$ for $i \neq k$. Moreover, $g(\frac{\partial}{\partial u^i}, \frac{\partial}{\partial u^i})$ is given by $\frac{4}{(1+\langle y,y\rangle)^2}$. This implies that an orthonormal frame is given by $\xi_i := f(u)\frac{\partial}{\partial u^i}$, where $f(u) = f(u^1, \dots, u^n) = \frac{1+\sum(u^j)^2}{2}$. Of course, the corresponding coframe is given by $\sigma^i = \frac{1}{f}du^i$. Consequently, $d\sigma^i = -\frac{1}{f^2}df \wedge du^i$ and since $df = \sum_j u^j du^j$ this can be written as $\sum_j \frac{u^j}{f^2} du^i \wedge du^j = \sum_j u^j \sigma^i \wedge \sigma^j$. This can be written as $-\sum_j \omega^i_j \wedge \sigma^j$ for $\omega^i_j = u^i \sigma^j - u^j \sigma^i = \frac{u^i}{f}du^j - \frac{u^j}{f}du^i$, which evidently satisfies $\omega^j_i = -\omega^i_j$ and thus gives the Levi–Civita connection in our frame.

CHAPTER 3

Basic Riemannian geometry

In this chapter we show how the Levi–Civita connection forms on the orthonormal frame bundle can be used to define various operations on the underlying manifold M. By construction, these operations are intrinsically associated to the Riemannian metric and can be used to construct invariants. The fundamental operation of this type is the covariant derivative on tensor fields. The covariant derivative can then be used to define geodesics and the exponential mapping. It can also be used to define the Riemann curvature tensor, which is the fundamental invariant of a Riemannian manifold.

The covariant derivative

3.1. The horizontal lift. Using the Levi–Civita connection form, we can canonically lift tangent vectors and vector fields from M to $\mathcal{P}M$. In general, a *lift* of a tangent vector $\xi \in T_xM$ to a point $\varphi \in p^{-1}(x) \subset \mathcal{P}M$ is a tangent vector $\tilde{\xi} \in T_\varphi \mathcal{P}M$ such that $T_\varphi p \cdot \tilde{\xi} = \xi$. A lift of a vector field $\xi \in \mathfrak{X}(M)$ is a vector field $\tilde{\xi} \in \mathfrak{X}(\mathcal{P}M)$ such that $Tp \circ \tilde{\xi} = \xi \circ p$. A vector field $\tilde{\xi} \in \mathfrak{X}(\mathcal{P}M)$ is called *projectable* if it is the lift of some vector field $\xi \in \mathfrak{X}(M)$ (which then is uniquely determined by $\tilde{\xi}$).

Proposition 3.1. *Let (M, g) be a Riemannian manifold with frame bundle $\mathcal{P}M$ and let (γ^i_j) be the Levi–Civita connection form on $\mathcal{P}M$.*

(1) For a point $x \in M$ and a tangent vector $\xi \in T_xM$ and $\varphi \in \mathcal{P}M$ with $p(\varphi) = x$, there is a unique lift $\xi^h \in T_\varphi(\mathcal{P}M)$ of ξ such that $\gamma^i_j(\xi^h) = 0$ for all i, j.

(2) If $\xi \in \mathfrak{X}(U)$ is a local vector field on M, then the horizontal lifts in points in $p^{-1}(U)$ fit together to define a smooth vector field $\xi^h \in \mathfrak{X}(U)$. For any $A \in O(n)$, we have $(r^A)^*\xi^h = \xi^h$.

Proof. (1) Since $T_\varphi p : T_\varphi \mathcal{P}M \to T_xM$ is surjective, we can find an element $\tilde{\xi} \in T_\varphi \mathcal{P}M$ such that $T_\varphi p \cdot \tilde{\xi} = \xi$. Taking $(\gamma^i_j(\varphi)(\tilde{\xi})) =: X \in \mathfrak{o}(n)$ then $\xi^h := \tilde{\xi} - \zeta_X(\varphi)$ has the two required properties.

On the other hand, if $T_\varphi p \cdot \hat{\xi} = \xi$, then $\xi^h - \hat{\xi} \in \ker(T_\varphi p)$ and hence $\xi^h - \hat{\xi} = \zeta_Y(\varphi)$, where $Y = (\gamma^i_j(\varphi)(\xi^h - \hat{\xi})) = (-\gamma^i_j(\varphi)(\hat{\xi}))$. If $\gamma^i_j(\hat{\xi}) = 0$, then this shows that $\hat{\xi} = \xi^h$.

(2) Since this is a local question, we may assume that M admits a local orthonormal frame defined on U. Then we consider the induced diffeomorphism $\Phi : p^{-1}(U) \to U \times O(n)$ as in Proposition 2.4. Since $T(U \times O(n)) = TU \times TO(n)$ we can start from $\xi \in \mathfrak{X}(U)$, form $(x, A) \mapsto (\xi(x), 0) \in \mathfrak{X}(U \times O(n))$ and pull this back by Φ to obtain a vector field $\tilde{\xi} \in \mathfrak{X}(p^{-1}(U))$. Evidently, this satisfies $T_\varphi p \cdot \tilde{\xi}(\varphi) = \xi(p(\varphi))$ for all $\varphi \in p^{-1}(U)$, so it is a lift of ξ. On the other hand, we know from 2.5 that $(T\Phi \cdot \zeta_X)(x, A) = (0, AX)$ holds for each $X \in \mathfrak{o}(n)$. So applying the construction of (1) in each point, we see that $T\Phi \cdot \xi^h$ is given by $(x, A) \mapsto (\xi(x), -A(\gamma^i_j(\tilde{\xi})(\Phi^{-1}(x, A))))$, so this is smooth, too.

For $A \in O(n)$ we have $p \circ r^A = p$ so $Tp \circ Tr^A = Tp$. This means that $T_\varphi r^A \cdot \xi^h(\varphi) \in T_{r^A(\varphi)}\mathcal{P}M$ is a lift of $\xi(p(\varphi))$. Moreover, condition (i) in the definition of a connection form immediately implies that $\gamma^i_j(T_\varphi r^A \cdot \xi^h(\varphi)) = 0$ for all i and j. Thus $T_\varphi r^A \cdot \xi^h(\varphi) = \xi^h(r^A(\varphi))$ which exactly means that $(r^A)^*\xi^h = \xi^h$. $\square$

3.2. The covariant derivative of vector fields. Via the horizontal lift, we can differentiate smooth function on $\mathcal{P}M$ (which may have values in some finite dimensional vector space) in a direction determined by a vector field on M. To get geometric operations out of this idea, we have to interpret geometric objects on M in terms of (vector valued) functions on $\mathcal{P}M$. The crucial step is doing this for vector fields.

Theorem 3.2. *Let (M, g) be a Riemannian manifold of dimension n with orthonormal frame bundle $\mathcal{P}M$, and let $\theta = (\theta^1, \dots, \theta^n)$ be the soldering form on $\mathcal{P}M$.*

(1) The space $\mathfrak{X}(M)$ of vector fields on M is in bijective correspondence with the space

$$C^\infty(\mathcal{P}M, \mathbb{R}^n)^{O(n)} := \{F : \mathcal{P}M \to \mathbb{R}^n : F(r^A(\varphi)) = A^{-1}F(\varphi) \quad \forall A \in O(n)\}$$

of $O(n)$–equivariant smooth functions $\mathcal{P}M \to \mathbb{R}^n$. Explicitly, a vector field $\xi \in \mathfrak{X}(M)$ corresponds to the function $\theta(\tilde{\xi})$, where $\tilde{\xi} \in \mathfrak{X}(\mathcal{P}M)$ is any lift of ξ.

(2) Let $\xi \in \mathfrak{X}(M)$ be a vector field with horizontal lift $\xi^h \in \mathfrak{X}(\mathcal{P}M)$, and let $F \in C^\infty(\mathcal{P}M, \mathbb{R}^n)^{O(n)}$ be an equivariant smooth function. Then the smooth function $\xi^h \cdot F : \mathcal{P}M \to \mathbb{R}^n$ is equivariant, too.

Proof. (1) Since a point $\varphi \in \mathcal{P}M$ is an orthogonal linear isomorphism $\mathbb{R}^n \to T_xM$, where $x = p(\varphi)$, an element $F(\varphi) \in \mathbb{R}^n$ gives rise to the tangent vector $\varphi(F(\varphi)) \in T_xM$. Any other point $\hat{\varphi}$ with $p(\hat{\varphi}) = x$ is of the form $\hat{\varphi} = r^A(\varphi) = \varphi \circ A$ for some $A \in O(n)$. This determines the same tangent vector if and only if $AF(r^A(\varphi)) = F(\varphi)$ i.e. iff $F(r^A(\varphi)) = A^{-1}F(\varphi)$. Hence we see that an equivariant function F gives rise to a unique tangent vector at each point of M.

If we conversely have given a tangent vector $\xi(x) \in T_xM$ for each $x \in M$, we can associate to it the function $F(\varphi) := \varphi^{-1}(\xi(p(\varphi))) \in \mathbb{R}^n$, which is evidently equivariant. Hence it suffices to prove that smoothness of the map $x \mapsto \xi(x)$ is equivalent to smoothness of the function $F : \mathcal{P}M \to \mathbb{R}^n$.

Starting with a smooth vector field $\xi \in \mathfrak{X}(M)$, we know from Proposition 3.1 that the horizontal lift ξ^h is a smooth vector field, too. But by definition of the soldering form, the function $F : \mathcal{P}M \to \mathbb{R}$ associated to ξ can be written as $\theta(\xi^h) = (\theta^1(\xi^h), \dots, \theta^n(\xi^h))$ and thus is smooth. The same clearly applies to any smooth lift of ξ.

Conversely, let $F : \mathcal{P}M \to \mathbb{R}^n$ be smooth and equivariant. We can show that the corresponding map $x \mapsto \xi(x)$ is smooth locally, so let us restrict to an open subset $U \subset M$, for which there is an orthonormal frame $\{\xi_1, \dots, \xi_n\}$ defined on U. Denoting by $\sigma : U \to \mathcal{P}M$ the corresponding local section from 2.4, we have $\sigma(x)(e_i) = \xi_i(x)$ for $i = 1, \dots, n$. But this implies that $\xi(x)$ can be written as $\sigma(x)(F(\sigma(x)))$ for all $x \in M$. Writing $F \circ \sigma = (f^1, \dots, f^n)$, we conclude that $\xi(x) = \sum_i f^i(x)\xi_i(x)$. But since $F \circ \sigma : U \to \mathbb{R}^n$ is smooth, this shows that $\xi|_U = \sum_i f^i \xi_i$ is smooth, which completes the proof of (1).

(2) From the definition of the pullback of vector fields, we see that $(\xi^h \cdot F) \circ r^A = (r^A)^*\xi^h \cdot (F \circ r^A)$. But $(r^A)^*\xi^h = \xi^h$ and writing $F = (F^1, \dots, F^n)$, equivariancy reads as $F^i \circ r^A = \sum_j b^i_j F^j$, where $A^{-1} = (b^i_j)$. Differentiating in direction ξ^h, we get $\xi^h \cdot (F^i \circ r^A) = \sum_j b^i_j(\xi^h \cdot F^j)$, which implies the result. $\square$

Definition 3.2. Let (M, g) be a Riemannian manifold with orthonormal frame bundle $\mathcal{P}M$, let $\xi, \eta \in \mathfrak{X}(M)$ be vector fields and let $F : \mathcal{P}M \to \mathbb{R}^n$ be the equivariant smooth function corresponding to η. Then the vector field corresponding to $\xi^h \cdot F : \mathcal{P}M \to \mathbb{R}^n$ is called the *covariant derivative of* η *in direction* ξ and is denoted by $\nabla_\xi \eta$.

Now we can easily clarify the basic properties of this operation:

Corollary 3.2. *Let* (M, g) *be a Riemannian manifold and let* $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ *be the covariant derivative. Then for* $\xi, \eta, \zeta \in \mathfrak{X}(M)$ *and* $f \in C^\infty(M, \mathbb{R})$ *we get:*

(1) ∇ *is bilinear over* $\mathbb{R}$.

(2) $\nabla_{f\xi}\eta = f\nabla_\xi \eta$ *and* $\nabla_\xi f\eta = f\nabla_\xi\eta + (\xi \cdot f)\eta$.

(3) $\xi \cdot g(\eta, \zeta) = g(\nabla_\xi \eta, \zeta) + g(\eta, \nabla_\xi \zeta)$.

(4) $\nabla_\xi \eta - \nabla_\eta \xi = [\xi, \eta]$.

(5) Let $(\tilde{M}, \tilde{g})$ *be another Riemannian manifold with* $\dim(\tilde{M}) = \dim(M)$ *and with covariant derivative* $\tilde{\nabla}$. *Then for an isometry* $I : \tilde{M} \to M$ *we have* $I^*(\nabla_\xi \eta) = \tilde{\nabla}_{I^*\xi} I^*\eta$.

Proof. The vector field $\xi^h + \eta^h$ evidently lifts $\xi + \eta$ and satisfies $\gamma^i_j(\xi^h + \eta^h) = 0$ for all i, j, so $\xi^h + \eta^h = (\xi + \eta)^h$. Likewise, one verifies that $(f \circ p)\xi^h = (f\xi)^h$. On the other hand, the bijection $\mathfrak{X}(M) \to C^\infty(\mathcal{P}M, \mathbb{R})^{O(n)}$ is evidently linear and if η corresponds to F, then $f\eta$ corresponds to $(f \circ p)F$. From this, (1) follows immediately, and the first part of (2) is clear since $((f \circ p)\xi^h) \cdot F = (f \circ p)(\xi^h \cdot F)$. For the second part of (2), we have $\xi^h \cdot ((f \circ p)F) = (\xi^h \cdot (f \circ p))F + (f \circ p)\xi^h \cdot F$. But $\xi^h \cdot (f \circ p)(\varphi) = (T_\varphi p \cdot \xi^h) \cdot f = \xi(p(\varphi)) \cdot f$, and hence the first summand corresponds to $(\xi \cdot f)\eta$, which completes the proof of (2).

(3) Suppose that η and ζ correspond to $F, G : \mathcal{P}M \to \mathbb{R}^n$ and let F^i and G^i be the components of these functions. Then $F(\varphi) = \varphi^{-1}(\eta(p(\varphi)))$ and likewise for G, and since φ is orthogonal we conclude that $\sum_i F^i(\varphi)G^i(\varphi) = g(\eta, \zeta)(p(\varphi))$. Hence $g(\eta, \zeta) \circ p = \sum_i F^i G^i$ and thus

$$(\xi \cdot g(\eta, \zeta)) \circ p = \xi^h \cdot \textstyle\sum_i F^i G^i = \sum_i (\xi^h \cdot F^i)G^i + \sum_i F^i(\xi^h \cdot G^i).$$

As above, we see that the last two terms are given by $g(\nabla_\xi \eta, \zeta) \circ p$ and $g(\eta, \nabla_\xi \zeta) \circ p$, respectively. Since p is surjective, we get (3).

(4) The defining property of the Levi–Civita connection form reads as $0 = d\theta^i + \sum_j \gamma^i_j \wedge \theta^j$. The horizontal lifts ξ^h and η^h of ξ and η by definition insert both trivially into each γ^i_j, so we conclude that $d\theta^i(\xi^h, \eta^h) = 0$ for all i. Inserting the definition of the exterior derivative, we obtain

$$\theta^i([\xi^h, \eta^h]) = \xi^h \cdot \theta^i(\eta^h) - \eta^h \cdot \theta^i(\xi^h).$$

Since $[\xi^h, \eta^h]$ is a lift of $[\xi, \eta]$, the left hand side describes the ith component of the equivariant function $\mathcal{P}M \to \mathbb{R}^n$ corresponding to $[\xi, \eta]$. But the right hand side evidently is the ith component of the function corresponding to $\nabla_\xi \eta - \nabla_\eta \xi$, which implies the result.

(5) This simply expresses the fact that the operation has been constructed in a natural way and although the proof may look a bit technical, it is actually straightforward. First, we can extend an isometry $I : \tilde{M} \to M$ to a smooth map $\mathcal{P}I : \mathcal{P}\tilde{M} \to \mathcal{P}M$. Indeed, a point $\tilde{\varphi} \in \mathcal{P}\tilde{M}$ is an orthogonal linear isomorphism $\mathbb{R}^n \to T_{\tilde{x}}\tilde{M}$, where $\tilde{x} = \tilde{p}(\tilde{\varphi})$. Likewise, $T_{\tilde{x}}I : T_{\tilde{x}}\tilde{M} \to T_{I(\tilde{x})}M$ is an orthogonal linear isomorphism, and we define $\mathcal{P}I(\tilde{\varphi}) := T_{\tilde{x}}I \circ \tilde{\varphi} \in \mathcal{P}M$. This immediately implies that $p \circ \mathcal{P}I = I \circ \tilde{p}$.

To see that $\mathcal{P}I$ is smooth, let $U \subset M$ be an open subset such that there is a local orthonormal frame $\{\xi_1, \dots, \xi_n\}$ for M defined on U. Consider $\tilde{U} := I^{-1}(U) \subset \tilde{M}$ and $I^*\xi_i \in \mathfrak{X}(\tilde{U})$. One immediately verifies that these vector fields constitute a local orthonormal frame for $\tilde{M}$. Moreover for the corresponding isomorphisms $\Phi : p^{-1}(U) \to U \times O(n)$ and $\tilde{\Phi} : \tilde{p}^{-1}(\tilde{U}) \to \tilde{U} \times O(n)$ as in 2.4, one easily verifies that

$$\mathcal{P}I = \Phi^{-1} \circ (I \times \mathrm{id}_{O(n)}) \circ \tilde{\Phi},$$

which implies smoothness of $\mathcal{P}I$. From the definition, it also follows immediately that $\mathcal{P}I \circ r^A = r^A \circ \mathcal{P}I$ for all $A \in O(n)$.

For the soldering forms θ^i on $\mathcal{P}M$ and $\tilde{\theta}^i$ on $\mathcal{P}\tilde{M}$, consider $(\mathcal{P}I)^*\theta^i$. This maps $\xi \in T_{\tilde{\varphi}}\mathcal{P}\tilde{M}$ to

$$\theta^i(\mathcal{P}I(\tilde{\varphi}))(T_{\tilde{\varphi}}\mathcal{P}I \cdot \xi) = (\mathcal{P}I(\tilde{\varphi}))^{-1}(Tp \cdot T\mathcal{P}I \cdot \xi).$$

Since $\mathcal{P}(\tilde{\varphi}) = T_{\tilde{x}}I \circ \tilde{\varphi}$ and $Tp \circ T\mathcal{P}I = TI \circ T\tilde{p}$, we conclude that this equals the ith component of $\tilde{\varphi}^{-1}(T_{\tilde{\varphi}}\tilde{p} \cdot \xi)$ and hence $\tilde{\theta}^i(\tilde{\varphi})(\xi)$. Thus we see that $(\mathcal{P}I)^*\theta^i = \tilde{\theta}^i$ for all i.

Next, let (γ^i_j) be the Levi–Civita connection form on $\mathcal{P}M$ and consider the one–forms $(\mathcal{P}I)^*\gamma^i_j$ on $\mathcal{P}\tilde{M}$. One easily verifies that they define a connection form. Moreover,

$$d\tilde{\theta}^i + \textstyle\sum_j((\mathcal{P}I)^*\gamma^i_j) \wedge \tilde{\theta}^j = (\mathcal{P}I)^* \left(d\theta^i + \sum_j \gamma^i_j \wedge \theta^j\right) = 0.$$

By Theorem 2.7, this implies that $(\mathcal{P}I)^*\gamma^i_j = \tilde{\gamma}^i_j$, so $\mathcal{P}I$ is compatible with the Levi–Civita connection forms. For $\xi \in \mathfrak{X}(M)$, this easily implies that $(\mathcal{P}I)^*\xi^h = (I^*\xi)^h$ and for $\eta \in \mathfrak{X}(M)$ corresponding to $F : \mathcal{P}M \to \mathbb{R}^n$, the pullback $(I^*\eta)$ corresponds to the function $F \circ \mathcal{P}I$. Now the result follows directly from the fact that $((\mathcal{P}I)^*\xi^h) \cdot (F \circ \mathcal{P}I) = (\xi^h \cdot F) \circ \mathcal{P}I$. □

Remark 3.2. The proofs of parts (1)–(3) of this Corollary are different from the proofs of part (4) and (5) and this is for good reason. In fact, given any connection form $(\hat{\gamma}^i_j)$ on the orthonormal frame bundle $\mathcal{P}M$ of a Riemannian manifold (M, g), one can define an analog of the horizontal lift operation (where now "horizontal" is interpreted with respect to $\hat{\gamma}$). This in turn can be used to define an operation $\hat{\nabla}$ on vector fields. (Notice that the isomorphism in part (1) of Theorem 3.2 depends only on the soldering form θ and not on γ.) Then one can argue as in the proof of Corollary 3.2 to show that $\hat{\nabla}$ has properties (1)–(3). This is usually expressed by saying that $\hat{\gamma}$ (respectively $\hat{\nabla}$) is a *metric connection*.

Property (4) is actually an equivalent version of the defining property of the Levi–Civita connection from Theorem 2.6, and it is this property which makes the Levi–Civita connection unique. This in turn is the reason why property (5) holds.

3.3. Interpreting tensor fields on the frame bundle. To extend the covariant derivative to arbitrary tensor fields, we can proceed similarly as in 3.2, provided that we find an interpretation of tensor fields as certain vector valued functions on $\mathcal{P}M$. For technical reasons, we discuss the case of one–forms separately.

Consider the space $\mathbb{R}^{n*} = L(\mathbb{R}^n, \mathbb{R})$, the dual space to $\mathbb{R}^n$. There is a natural action of $GL(n, \mathbb{R})$ on this space defined by $(A \cdot \lambda)(X) = \lambda(A^{-1}X)$ for $\lambda \in \mathbb{R}^{n*}$ and $X \in \mathbb{R}^n$. The motivation for the inverse is on the one hand that it makes sure that $(A \cdot \lambda)(A \cdot X) = \lambda(X)$. On the other hand, to get $AB \cdot \lambda = A \cdot (B \cdot \lambda)$, the inverse is necessary, too. Of course, we can restrict this to an action of the subgroup $O(n)$.

Lemma 3.3. *Let (M, g) be a Riemannian manifold with orthonormal frame bundle $\mathcal{P}M$. Then there is a bijective correspondence between the space $\Omega^1(M)$ of one–forms on M and the space*

$$C^\infty(\mathcal{P}M, \mathbb{R}^{n*})^{O(n)} = \{F \in C^\infty(\mathcal{P}M, \mathbb{R}^{n*}) : F(r^A(\varphi)) = A^{-1} \cdot F(\varphi) \quad \forall A \in O(n)\}.$$

Viewing vector fields as equivariant $\mathbb{R}^n$–valued smooth functions on $\mathcal{P}M$, inserting a vector field into a one–form corresponds to the point–wise application of linear functionals to vectors.

Proof. Let $\alpha \in \Omega^1(M)$ be a one–form. Given $\varphi \in \mathcal{P}_xM$, which is an orthogonal linear isomorphism $\varphi : \mathbb{R}^n \to T_xM$, where $x = p(\varphi)$, we put $F(\varphi) := \alpha_x \circ \varphi : \mathbb{R}^n \to \mathbb{R}$. Since $r^A(\varphi) = \varphi \circ A$ it follows from the definition that $F(r^A(\varphi)) = F(\varphi) \circ A = A^{-1} \cdot F(\varphi)$. To see that F is smooth, consider an open subset $U \subset M$ such that there exists a local orthonormal frame $\{\xi_i\}$ for M defined on U, and let $\sigma : U \to \mathcal{P}M$ be the corresponding smooth section. Then by definition $\sigma(x) \in \mathcal{P}_xM$ maps $e_i \in \mathbb{R}^n$ to $\xi_i(x)$ for all i. Looking at $F \circ \sigma : U \to \mathbb{R}^{n*}$ we conclude that $F(\sigma(x))(e_i) = \alpha(\xi_i)(x)$, which shows that $F \circ \sigma$ is smooth. Using that $(x, A) \mapsto \sigma(x) \circ A$ defines a diffeomorphism $U \times O(n) \to p^{-1}(U)$ and that $F(\sigma(x) \circ A) = F(\sigma(x)) \circ A$, we conclude that F is smooth on $p^{-1}(U) \subset \mathcal{P}M$.

Conversely, given φ and $F(\varphi) \in \mathbb{R}^{n*}$, we can define a linear map $\alpha_x : T_{p(\varphi)}M \to \mathbb{R}$ as $F(\varphi) \circ \varphi^{-1}$. Any other point lying over x is of the form $r^A(\varphi) = \varphi \circ A$ for $A \in O(n)$, and by definition $(r^A(\varphi), F(r^A(\varphi)))$ leads to the same functional if and only if $F(r^A(\varphi))) = A \circ F(\varphi) = A^{-1} \cdot F(\varphi)$. Hence an equivariant function induces a well defined linear functional on each tangent space. Now a vector field $\xi \in \mathfrak{X}(M)$ corresponds by Theorem 3.2 to a smooth equivariant function $G : \mathcal{P}M \to \mathbb{R}^n$ characterized by $G(\varphi) = \varphi^{-1}(\xi(p(\varphi)))$. But this means that $\alpha_x(\xi(x)) = F(\varphi)(G(\varphi))$ for all $\varphi \in \mathcal{P}_xM$. This shows that $x \mapsto \alpha_x$ is a one–form on M as well as the last claim in the lemma. □

Having this at hand, it is rather easy to pass to tensor fields. Consider the space $L^\ell_k(\mathbb{R}^n, \mathbb{R})$ of all $k + \ell$–linear maps $(\mathbb{R}^n)^k \times (\mathbb{R}^{n*})^\ell \to \mathbb{R}$. Then this carries a natural action of $GL(n, \mathbb{R})$ defined by

$$(A \cdot t)(X_1, \dots, X_k, \lambda_1, \dots, \lambda_\ell) = t(A^{-1}X_1, \dots, A^{-1}X_k, \lambda_1 \circ A, \dots, \lambda_\ell \circ A).$$

As before, the inverses are chosen in such a way that $AB \cdot t = A \cdot (B \cdot t)$. Using this, we can now formulate

Proposition 3.3. *Let (M, g) be a Riemannian manifold with orthonormal frame bundle $\mathcal{P}M$. Then there is a bijective correspondence between the space $\mathcal{T}^\ell_k(M)$ of $\binom{\ell}{k}$–tensor fields on M and the space $C^\infty(\mathcal{P}M, L^\ell_k(\mathbb{R}^n, \mathbb{R}))^{O(n)}$ of all smooth functions $F : \mathcal{P}M \to L^\ell_k(\mathbb{R}^n, \mathbb{R})$ such that $F(r^A(\varphi)) = A^{-1} \cdot F(\varphi)$ for all $A \in O(n)$.*

Viewing vector fields and one–forms as equivariant smooth functions on $\mathcal{P}M$ with values in $\mathbb{R}^n$ and $\mathbb{R}^{n*}$, respectively, inserting vector fields and one–forms into a tensor field corresponds to the point–wise insertion of elements of $\mathbb{R}^n$ and $\mathbb{R}^{n*}$ into multilinear maps.

Proof. This is very similar to the proof of part (1) of Theorem 3.2 and of Lemma 3.3. Given a tensor field $t \in \mathcal{T}^\ell_k(M)$, one defines $F(\varphi) \in L^\ell_k(\mathbb{R}^n, \mathbb{R})$ by

$$F(\varphi)(X_1, \dots, X_k, \lambda_1, \dots, \lambda_\ell) := t_x(\varphi(X_1), \dots, \varphi(X_k), \lambda_1 \circ \varphi^{-1}, \dots, \lambda_\ell \circ \varphi^{-1}),$$

and verifies directly that this is equivariant. To prove smoothness, one restricts to an open subset U for which there is an orthonormal frame defined on U. Denoting by $\sigma : U \to \mathcal{P}M$ the corresponding smooth section, the component functions $F \circ \sigma$ are obtained by plugging elements of the frame and of the dual coframe into t, from which one deduces that F is smooth.

Conversely, $\varphi \in \mathcal{P}_xM$ and $F(\varphi) \in L^\ell_k(\mathbb{R}^n, \mathbb{R})$ give rise to a multilinear map t_x and $F(r^A(\varphi)) = A^{-1} \cdot F(\varphi)$ makes sure that all points in $\mathcal{P}_xM$ lead to the same map t_x. From the definition, it follows easily that inserting values of vector fields and one–forms into t_x corresponds to inserting the values of the corresponding functions into $F(\varphi)$, which shows that smoothness of F implies smoothness of t. □

Remark 3.3. (1) The different interpretations of $\binom{\ell}{k}$–tensor fields mentioned in section 4.3 of [**DG1**] have an obvious analog here. For example, the isomorphism $\mathcal{T}^1_0(M) \cong \mathfrak{X}(M)$ simply corresponds to the canonical isomorphism $L(\mathbb{R}^{n*}, \mathbb{R}) \cong \mathbb{R}^n$ which is compatible with the actions of $O(n)$ on the two spaces.

One can identify the space of bilinear maps $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ with $L^1_2(\mathbb{R}^n, \mathbb{R})$ by sending $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ to the trilinear map $(X, Y, \lambda) \mapsto \lambda(b(X, Y))$. This corresponds to viewing the values of a $\binom{1}{2}$–tensor field as bilinear maps $T_xM \times T_xM \to T_xM$ rather than as trilinear maps $T_xM \times T_xM \times T^*_xM \to \mathbb{R}$. This extends to other types of tensor fields.

(2) Similar to the coordinate representation of a smooth tensor field, one can expand it locally in terms of an orthonormal frame and the associated coframe. Take a local orthonormal frame $\{\xi_1, \dots, \xi_n\}$ defined on U, the associated coframe $\{\sigma^1, \dots, \sigma^n\}$, and a tensor field $t \in \mathcal{T}^\ell_k(M)$. Then for any choice $i_1, \dots, i_\ell, j_1, \dots, j_k \in \{1, \dots, n\}$ of indices, we get a $\binom{\ell}{k}$–tensor field $\sigma^{j_1} \otimes \dots \otimes \sigma^{j_k} \otimes \xi_{i_1} \otimes \dots \otimes \xi_{i_\ell}$ and a smooth function

$$t^{i_1,\dots,i_\ell}_{j_1,\dots,j_k} := t(\xi_{j_1}, \dots, \xi_{j_k}, \sigma^{i_1}, \dots, \sigma^{i_\ell})$$

defined on U such that

$$t|_U = \sum_{i_1,\dots,j_\ell} t^{i_1,\dots,i_\ell}_{j_1,\dots,j_k} \sigma^{j_1} \otimes \dots \otimes \sigma^{j_k} \otimes \xi_{i_1} \otimes \dots \otimes \xi_{i_\ell}.$$

From the definitions we easily conclude that in terms of the section $\sigma : U \to \mathcal{P}M$ associated to the frame and the equivariant function $F : \mathcal{P}M \to L^\ell_k(\mathbb{R}^n, \mathbb{R})$ the functions $t^{i_1,\dots,i_\ell}_{j_1,\dots,j_k}$ are simply the component functions of $F \circ \sigma$.

Example 3.3. (1) The Riemannian metric g is a $\binom{0}{2}$–tensor field. For any orthonormal frame $\{\xi_i\}$, we of course have $g(\xi_i, \xi_j) = \delta_{ij}$, which implies that g corresponds to the constant function $\langle\ ,\ \rangle \in L^0_2(\mathbb{R}^n, \mathbb{R})$.

(2) Take functions ${\Psi_k}^i_j$ as in Lemma 2.5, and use them to define a smooth map $\Psi : \mathcal{P}M \to L^1_2(\mathbb{R}^n, \mathbb{R})$ by $\Psi(\varphi)(X, Y, \lambda) := \sum_{i,j,k} {\Psi_k}^i_j X^k Y^j \lambda_i$. Then the compatibility condition of Φ with r^A from Lemma 2.5 exactly says that $\Psi(r^A(\varphi)) = A^{-1} \cdot \Psi(\varphi)$, whence Ψ determines a $\binom{1}{2}$–tensor field on M. The skew symmetry condition from Lemma 2.5 can be best expressed by considering $\Psi(\xi,\ )$ for $\xi \in \mathfrak{X}(M)$ as associating to each $x \in M$ an endomorphism of T_xM. Then the statement just says that this endomorphism is skew symmetric with respect to g_x.

To understand this geometrically, recall that the functions ${\Psi_k}^i_j$ describe the difference between two sets of connections forms. As we have noted in 3.2, any connection form gives rise to an operator which has properties (1)–(3) of Corollary 3.2. One can verify directly, that the functions ${\Psi_k}^i_j$ describe the difference between the two operators coming from the connection forms. It is not surprising that this difference is a tensor field. Let us denote the operators by ∇ and $\hat{\nabla}$ and consider $\Psi : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ defined by $\Psi(\xi, \eta) := \hat{\nabla}_\xi \eta - \nabla_\xi \eta$. Since both operators have properties (1) and (2) from Corollary 3.2, this is linear over smooth functions in both variables and thus defines a tensor field $\Psi \in \mathcal{T}^1_2(M)$. If one in addition assumes that both operations also satisfy property (3) from Corollary 3.2, then this implies $0 = g(\Psi(\xi, \eta), \zeta) + g(\eta, \Psi(\xi, \zeta))$, which explains the skew symmetry condition on Ψ.

(3) Consider the functions T^i_{jk} associated to a choice of connection form (γ^i_j) on $\mathcal{P}M$ in 2.5. We can use them to define a smooth map T from $\mathcal{P}M$ to the space of bilinear maps $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ by $T(\varphi)(X,Y) := \sum_{i,j,k} T^i_{jk} X^j Y^k e_i$, where $\{e_i\}$ is the standard basis of $\mathbb{R}^n$. Then the compatibility condition with r^A verified in the end of 2.5 reads as $T(r^A(\varphi))(X,Y) = A^{-1}T(\varphi)(AX,AY)$, i.e. $T(r^A(\varphi)) = A^{-1}\cdot T(\varphi)$. Hence T defines a $\binom{1}{2}$–tensor field on M, called the *torsion* of the connection form (γ^i_j), and a connection with vanishing torsion is called *torsion–free*. In this language, Theorem 2.6 says that there is a unique torsion–free metric connection on $\mathcal{P}M$ and this is the Levi–Civita connection.

Again, this can be nicely interpreted in terms of the associated operation on vector fields. Indeed, if $\nabla : \mathfrak{X}(M)\times\mathfrak{X}(M)\to\mathfrak{X}(M)$ is an operator which has properties (1) and (2) of Corollary 3.2, then one defines $T : \mathfrak{X}(M)\times\mathfrak{X}(M)\to\mathfrak{X}(M)$ by $T(\xi,\eta) := \nabla_\xi\eta - \nabla_\eta\xi - [\xi,\eta]$. By definition, this is skew symmetric and from the compatibility of the Lie bracket with multiplication of a vector field by a smooth function, it follows immediately that T is bilinear over smooth functions. Thus it defines a tensor field, and one can verify directly that this corresponds to the functions T^i_{jk} considered above.

3.4. The covariant derivative of tensor fields. Extending the covariant derivative to tensor fields is now straightforward. To efficiently compute with covariant derivatives, the main information we need is compatibility with natural operations.

Lemma 3.4. *If $F : \mathcal{P}M \to L^\ell_k(\mathbb{R}^n,\mathbb{R})$ is an $O(n)$–equivariant smooth function, then for any vector field $\xi \in \mathfrak{X}(M)$ with horizontal lift $\xi^h \in \mathfrak{X}(\mathcal{P}M)$, the function $\xi^h\cdot F : \mathcal{P}M \to L^\ell_k(\mathbb{R}^n,\mathbb{R})$ is $O(n)$–equivariant, too.*

Proof. We can split F into components via

$$F(\varphi)(X_1,\dots,X_k,\lambda_1,\dots,\lambda_\ell) = \sum_{i_1,\dots,i_k,j_1,\dots,j_\ell} F^{j_1,\dots,j_\ell}_{i_1,\dots,i_k}(\varphi)(X_1)^{i_1}\dots(X_k)^{i_k}\lambda_1(e_{j_1})\dots\lambda_\ell(e_{j_\ell})$$

thus defining smooth functions $F^{j_1,\dots,j_\ell}_{i_1,\dots,i_k} : \mathcal{P}M \to \mathbb{R}$ which are given by

$$F^{j_1,\dots,j_\ell}_{i_1,\dots,i_k}(\varphi) = F(\varphi)(e_{i_1},\dots,e_{i_k},e^{j_1},\dots,e^{j_k}),$$

where $\{e_1,\dots,e_n\}$ is the standard basis of $\mathbb{R}^n$ and $\{e^1,\dots,e^n\}$ is the dual basis of $\mathbb{R}^{n*}$. In this language the equivariancy condition reads as

$$F^{j_1,\dots,j_\ell}_{i_1,\dots,i_k}\circ r^A = \sum_{r_1,\dots,r_k,s_1,\dots,s_\ell} a^{r_1}_{i_1}\dots a^{r_k}_{i_k} b^{j_1}_{s_1}\dots b^{j_\ell}_{s_\ell} F^{s_1,\dots,s_\ell}_{r_1,\dots,r_k}$$

where $A = (a^i_j)\in O(n)$ with inverse $A^{-1} = (b^i_j)$. Having this at hand, the proof can be completed in exactly the same way as for part (1) of Theorem 3.2. □

Definition 3.4. Let (M,g) be a Riemannian manifold with orthonormal frame bundle $\mathcal{P}M$, let $\xi\in\mathfrak{X}(M)$ be a vector field and $t \in \mathcal{T}^\ell_k(M)$ a tensor field on M and let $F : \mathcal{P}M \to L^\ell_k(\mathbb{R}^n,\mathbb{R})$ be the equivariant smooth function corresponding to t. Then the $\binom{\ell}{k}$–tensor field field corresponding to $\xi^h\cdot F : \mathcal{P}M \to L^\ell_k(\mathbb{R}^n,\mathbb{R})$ is called the *covariant derivative of t in direction ξ* and is denoted by $\nabla_\xi t$.

To efficiently compute with the covariant derivative, we have to prove compatibility with certain natural operations. On the one hand, recall that there is the *tensor product* of tensor fields, which is defined point–wise. Given $t_1 \in \mathcal{T}^{\ell_1}_{k_1}(M)$ and $t_2 \in \mathcal{T}^{\ell_2}_{k_2}(M)$ we have $t_1\otimes t_2 \in \mathcal{T}^{\ell_1+\ell 2}_{k_1+k_2}(M)$. Likewise, we can define a tensor product $\psi_1\otimes\psi_2 \in L^{\ell_1+\ell_2}_{k_1+k_2}(\mathbb{R}^n,\mathbb{R})$ for $\psi_i \in L^{\ell_i}_{k_i}(\mathbb{R}^n,\mathbb{R})$ by distributing the arguments to the two mappings and then multiplying their values. It is evident, the in the picture of equivariant functions the tensor product of tensor fields corresponds to the point–wise tensor production of the associated equivariant functions.

For the second operation, suppose that we have a linear map $\Phi : L^\ell_k(\mathbb{R}^n, \mathbb{R}) \to L^{\ell'}_{k'}(\mathbb{R}^n, \mathbb{R})$ which is $O(n)$*–equivariant* in the sense that $\Phi(A \cdot \psi) = A \cdot \Phi(\psi)$. Given a tensor field $t \in \mathcal{T}^\ell_k(M)$ we can take the corresponding equivariant function $F : \mathcal{P}M \to L^\ell_k(\mathbb{R}^n, \mathbb{R})$. Then it is clear that also $\Phi \circ F : \mathcal{P}M \to L^{\ell'}_{k'}(\mathbb{R}^n, \mathbb{R})$ is $O(n)$–equivariant and hence corresponds to a tensor field $\Phi(t) \in \mathcal{T}^{\ell'}_{k'}(M)$.

Proposition 3.4. *Let (M, g) be a Riemannian manifold with orthonormal frame bundle $\mathcal{P}M$. Then we have*

(1) For tensor fields t_1 and t_2 on M and any $\xi \in \mathfrak{X}(M)$, we have

$$\nabla_\xi(t_1 \otimes t_2) = (\nabla_\xi t_1) \otimes t_2 + t_1 \otimes (\nabla_\xi t_2).$$

(2) If Φ is the operator on tensor fields induced by some $O(n)$–equivariant map (denoted by the same symbol), then for any $\xi \in \mathfrak{X}(M)$ we have $\Phi(\nabla_\xi t) = \nabla_\xi(\Phi(t))$.

Proof. (1) If t_i corresponds to $F_i : \mathcal{P}M \to L^{\ell_i}_{k_i}(\mathbb{R}^n, \mathbb{R})$ for $i = 1, 2$, then we have noted already that $t_1 \otimes t_2$ corresponds to the point–wise tensor product, i.e. to the function $F(\varphi) = F_1(\varphi) \otimes F_2(\varphi)$. Since the tensor product on spaces of multilinear maps is bilinear, we see that $(\xi^h \cdot F)(\varphi) = (\xi^h \cdot F_1)(\varphi) \otimes F_2(\varphi) + F_1(\varphi) \otimes (\xi^h \cdot F_2)(\varphi)$, which implies the result.

(2) Since Φ is just a linear map between finite dimensional vector spaces, we get $\xi^h \cdot (\Phi \circ F) = \Phi \circ (\xi^h \cdot F)$, which implies the result. □

Example 3.4. (1) Let us first consider the covariant derivative on $\mathcal{T}^0_0(M) = C^\infty(M, \mathbb{R})$. We just have to observe that for a smooth map $F : \mathcal{P}M \to \mathbb{R}$, being $O(n)$–equivariant simply means that $F \circ r^A = F$ for all $A \in O(n)$. But this means that F is constant along the fibers of $p : \mathcal{P}M \to M$, so $F = f \circ p$ for a smooth function $f : M \to \mathbb{R}$. But then $\xi^h \cdot F = \xi^h \cdot (f \circ p) = (Tp \cdot \xi^h \cdot f) \circ p = (\xi \cdot f) \circ p$. This shows that $\nabla_\xi f = \xi \cdot f$ for $f \in C^\infty(M, \mathbb{R})$.

(2) Consider the trace map $\mathrm{tr} : L(\mathbb{R}^n, \mathbb{R}^n) \to \mathbb{R}$. We can identify $L(\mathbb{R}^n, \mathbb{R}^n)$ with $L^1_1(\mathbb{R}^n, \mathbb{R})$. This associates to $f : \mathbb{R}^n \to \mathbb{R}^n$ the bilinear map $(v, \lambda) \mapsto \lambda(f(v))$. Under this identification, the natural action on $L^1_1(\mathbb{R}^n, \mathbb{R})$ corresponds to the obvious action $A \cdot f = A \circ f \circ A^{-1}$. Since $\mathrm{tr}(A \circ f \circ A^{-1}) = \mathrm{tr}(f)$, we can view tr as an $O(n)$–equivariant linear map $L^1_1(\mathbb{R}^n, \mathbb{R}) \to \mathbb{R}$, where $O(n)$ acts trivially on $\mathbb{R}$, i.e. each $A \in O(n)$ acts as the identity. Now given a tensor field $t \in \mathcal{T}^1_1(M)$, we can view t as a associating to each $x \in M$ a linear map from T_xM to itself. Forming the trace in each point, we get a smooth function $\mathrm{tr}(t) : M \to \mathbb{R}$ and from part (2) of Proposition 3.4 and Example (1) we see that $\mathrm{tr}(\nabla_\xi t) = \xi \cdot \mathrm{tr}(t)$ for any $\xi \in \mathfrak{X}(M)$.

A more interesting application of this is the following: Take a vector field $\eta \in \mathfrak{X}(M)$ and a one–form $\alpha \in \Omega^1(M)$ and form $\eta \otimes \alpha \in \mathcal{T}^1_1(M)$. Now as an endomorphism this acts via $\zeta \mapsto \alpha(\zeta)\eta$ which implies that $\mathrm{tr}(\eta \otimes \alpha) = \alpha(\eta) \in C^\infty(M, \mathbb{R})$. From above, we see that $\mathrm{tr}(\nabla_\xi(\eta \otimes \alpha)) = \xi \cdot (\alpha(\eta))$ for all $\xi \in \mathfrak{X}(M)$. But by part (1) of the proposition, we get

$$\nabla_\xi(\eta \otimes \alpha) = (\nabla_\xi \eta) \otimes \alpha + \eta \otimes (\nabla_\xi \alpha),$$

and taking the trace, we get $\alpha(\nabla_\xi \eta) + (\nabla_\xi \alpha)(\eta)$. Putting this together, we obtain the formula

$$(\nabla_\xi \alpha)(\eta) = \xi \cdot (\alpha(\eta)) - \alpha(\nabla_\xi \eta),$$

which completely describes the covariant derivative of one–forms in terms of the covariant derivative of vector fields.

(3) We can describe the covariant derivative of general $\binom{0}{k}$–tensor fields in a very similar way. There is an obvious evaluation map $\mathrm{ev} : L^0_k(\mathbb{R}^n, \mathbb{R}) \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ (with k copies of $\mathbb{R}^n$) defined by $\mathrm{ev}(\psi, X_1, \dots, X_k) = \psi(X_1, \dots, X_k)$ and this is $(k+1)$–linear. One easily verifies that this is equivariant for the component–wise action of $O(n)$ on the left hand side and the trivial action on the right hand side.

For a tensor field $t \in \mathcal{T}^0_k(M)$ and $\eta_1, \dots, \eta_k \in \mathfrak{X}(M)$ applying this map pointwise to the corresponding equivariant functions corresponds to

$$(t, \eta_1, \dots, \eta_k) \mapsto t(\eta_1, \dots, \eta_k) \in C^\infty(M, \mathbb{R}).$$

Proceeding as in (2) we conclude that for any $\xi \in \mathfrak{X}(M)$ we get

$$(\nabla_\xi t)(\eta_1, \dots, \eta_k) = \xi \cdot (t(\eta_1, \dots, \eta_k)) - \textstyle\sum_{i=1}^n t(\eta_1, \dots, \nabla_\xi \eta_i, \dots, \eta_k).$$

Applying this to the Riemannian metric $g \in \mathcal{T}^0_2(M)$ we conclude that the fact that $\nabla g = 0$ (which is obvious, since g corresponds to a constant function $\mathcal{P}M \to L^0_2(\mathbb{R}^n, \mathbb{R})$) can be equivalently phrased as condition (3) from Corollary 3.2.

(4) What we have done so far did not really depend on the fact that we consider matrices $A \in O(n)$ only, since both tr and ev are actually compatible with the natural actions of $GL(n, \mathbb{R})$. To get something specific to the orthogonal group, consider the map $\Phi : \mathbb{R}^n \to \mathbb{R}^{n*}$ which maps $v \in \mathbb{R}^n$ to the linear functional $w \mapsto \langle\ , v\rangle$. For $A \in O(n)$ also the matrix A^{-1} is orthogonal, so $\langle w, Av\rangle = \langle A^{-1}w, v\rangle$, so $\Phi(Av) = A^{-1} \circ \Phi(v)$ which exactly says that Φ is $O(n)$–equivariant.

Passing to tensor fields, we see that for each $\eta \in \mathfrak{X}(M)$ we have $\Phi(\eta) \in \Omega^1(M)$, which is explicitly given by $\Phi(\eta)(\zeta) = g(\zeta, \eta)$. Since $\Phi : \mathbb{R}^n \to \mathbb{R}^{n*}$ is a linear isomorphism, also $\Phi : \mathfrak{X}(M) \to \Omega^1(M)$ is a linear isomorphism. By part (2) of Proposition 3.4 this isomorphism is compatible with the covariant derivative.

Similarly, we can find $O(n)$–equivariant linear isomorphisms between $L^\ell_k(\mathbb{R}^n, \mathbb{R})$ and $L^{\ell'}_{k'}(\mathbb{R}^n, \mathbb{R})$ provided that $k + \ell = k' + \ell'$. The induced isomorphisms between spaces of tensor fields are again compatible with the covariant derivative.

(5) The observation on the metric in the end of (3) can be easily generalized. Suppose that we have an $O(n)$–equivariant function $F : \mathcal{P}M \to L^\ell_k(\mathbb{R}^n, \mathbb{R})$ which happens to be constant or locally constant. Then of course $\xi^h \cdot F = 0$ for any $\xi \in \mathfrak{X}(M)$ so the tensor field $t \in \mathcal{T}^\ell_k(M)$ corresponding to F is *parallel*, i.e. it satisfies $\nabla_\xi t = 0$ for any ξ.

Now of course if $F : \mathcal{P}M \to L^\ell_k(\mathbb{R}^n, \mathbb{R})$ is locally constant and $c : [0, 1] \to O(n)$ is a smooth curve with $c(0) = \mathbb{I}$, then for any $\varphi \in \mathcal{P}M$ we must have $F(r^{c(t)}\varphi) = F(\varphi)$ for all t, so in particular $c(1)^{-1} \cdot F(\varphi) = F(\varphi)$ for any such curve. Now it turns out that one can reach any point in $SO(n) = \{A \in O(n) : \det(A) = 1\}$ by such a curve. Hence in order for a locally constant function to be $O(n)$–equivariant, its values must at least be *$SO(n)$–invariant*, i.e. fixed by the action of any element of $SO(n)$.

In particular, let us assume that $\psi \in L^\ell_k(\mathbb{R}^n, \mathbb{R})$ is $O(n)$–invariant, i.e. $A \cdot \psi = \psi$ for any $A \in O(n)$. Then for any Riemannian manifold M, the constant map ψ is $O(n)$–equivariant on $\mathcal{P}M$ and thus defines a canonical parallel tensor field on M.

There is a very interesting example for which we only have $SO(n)$–invariance. Consider the determinant mapping $\det \in L^0_n(\mathbb{R}^n, \mathbb{R})$. By linear algebra, we have $\det(Av_1, \dots, Av_n) = \det(A)\det(v_1, \dots, v_n)$ which means that $A \cdot \det = \det(A^{-1})\det$ for any $A \in GL(n, \mathbb{R})$, so det is indeed $SO(n)$–invariant. Now suppose that (M, g) is a Riemannian manifold, which is orientable. Choosing an orientation of M, any element

$\varphi \in \mathcal{P}M$ is either orientation preserving or orientation reversing when viewed as a linear isomorphism $\mathbb{R}^n \to T_xM$. Now define $F : \mathcal{P}M \to L^0_n(\mathbb{R}^n, \mathbb{R})$ by putting $F(\varphi) = \det$ if φ is orientation preserving and $F(\varphi) = -\det$ if φ is orientation reversing. It is easy to see that F is smooth, thus defining a parallel n–form $\mathrm{vol}(g) \in \Omega^n(M)$, called the *volume form* of the metric g. Explicitly, $\mathrm{vol}(g)$ is characterized by the fact that a positively oriented orthonormal basis has unit volume.

3.5. The covariant derivative as an operator. In the proof of Corollary 3.2, we have seen that for $\xi \in \mathfrak{X}(M)$ and $f \in C^\infty(M, \mathbb{R})$, we have $(f\xi)^h = (f \circ p)\xi^h$. As there we conclude that this implies that $(f\xi)^h \cdot F = (f \circ p)\xi^h \cdot F$ holds for any equivariant function $F : \mathcal{P}M \to L^\ell_k(\mathbb{R}^n, \mathbb{R})$, whence $\nabla_{f\xi} t = f\nabla_\xi t$ for any $t \in \mathcal{T}^\ell_k(M)$. Since we already know that $\nabla_\xi t \in \mathcal{T}^\ell_k(M)$ we see that the $(k + \ell + 1)$–linear map $\nabla t : (\mathfrak{X}(M))^{k+1} \times (\Omega^1(M))^\ell \to C^\infty(M, \mathbb{R})$ defined by

$$\nabla t(\eta_0, \dots, \eta_k, \alpha_1, \dots, \alpha_\ell) := (\nabla_{\eta_0} t)(\eta_1, \dots, \eta_k, \alpha_1, \dots, \alpha_\ell)$$

is linear over $C^\infty(M, \mathbb{R})$ in each argument and thus defines a tensor field $\nabla t \in \mathcal{T}^\ell_{k+1}(M)$. Now one can take the covariant derivative of this tensor field and form $\nabla^2 t = \nabla(\nabla t) \in \mathcal{T}^\ell_{k+2}(M)$ and inductively $\nabla^r t \in \mathcal{T}^\ell_{k+r}(M)$ for all $r \in \mathbb{N}$.

In terms of this operator, we can now nicely describe how to compute the covariant derivative in terms of an orthonormal (co–)frame:

Proposition 3.5. *Let (M, g) be a Riemannian manifold of dimension n and let $U \subset M$ be open subset. Suppose that there is a local orthonormal frame $\{\xi_1, \dots, \xi_n\}$ with corresponding coframe $\{\sigma^1, \dots, \sigma^n\}$ defined on U, and let $\omega^i_j \in \Omega^1(U)$ be the unique forms such that $\omega^j_i = -\omega^i_j$ and $0 = d\sigma^i + \sum_j \omega^i_j \wedge \sigma^j$ (compare with Proposition 2.7). Then we have*

(1) $\nabla\xi_i = \sum_j \omega^j_i \otimes \xi_j \in \mathcal{T}^1_1(M)$ and $\nabla\sigma^i = -\sum_j \omega^i_j \otimes \sigma^j \in \mathcal{T}^0_2(M)$ for all $i = 1, \dots, n$.

(2) If $\eta \in \mathfrak{X}(M)$ is any vector field, and $f^1, \dots, f^n : U \to \mathbb{R}$ are the unique smooth functions such that $\eta|_U = \sum_{i=1}^n f^i\xi_i$, then

$$\nabla\eta = \sum_{i=1}^n ((df^i + \sum_{j=1}^n f^j\omega^i_j) \otimes \xi_i).$$

Proof. The orthonormal frame $\{\xi_i\}$ defined on U gives rise to a smooth section $\sigma : U \to \mathcal{P}M$ of $\mathcal{P}M$ over U. This is characterized by the fact that $\sigma^*\theta^i = \sigma^i$ or equivalently by the fact that the equivariant function $\mathcal{P}M \to \mathbb{R}^n$ corresponding to ξ_i is constantly equal e_i along the image of σ. Given a vector field $\xi \in \mathfrak{X}(M)$, we can form $T_x\sigma \cdot \xi(x) \in T_{\sigma(x)}\mathcal{P}M$ for each $x \in U$. Take the Levi–Civita connection forms $\gamma^i_j \in \Omega^1(\mathcal{P}M)$, put $X := (\gamma^i_j(T_x\sigma \cdot \xi(x))) \in \mathfrak{o}(n)$ and consider $T_x\sigma \cdot \xi(x) - \zeta_X(\sigma(x))$. Since $\zeta_X(\sigma(x))$ is vertical, this is a lift of $\xi(x)$ and since (γ^i_j) is a connection form, we have $(\gamma^i_j(-\zeta_X)) = -X$. This shows that $\xi^h(\sigma(x)) = T_x\sigma \cdot \xi(x) - \zeta_X(\sigma(x))$.

Given a smooth function $F = (F^1, \dots, F^n) : \mathcal{P}M \to \mathbb{R}^n$, we can use this to compute $\xi^h(\sigma(x)) \cdot F$. First, we have $(T_x\sigma \cdot \xi(x)) \cdot F = \xi(x) \cdot (F \circ \sigma)$. On the other hand, if we assume that F is $O(n)$–equivariant, we compute $\zeta_X \cdot F$ using ideas from 2.5. There we saw that $\mathrm{Fl}^{\zeta_X}_t = r^{\exp(tX)}$, which shows that $(\mathrm{Fl}^{\zeta_X}_t)^* F^i = \sum_j a^i_j(-t)F^j$, where $(a^i_j(t)) = \exp(tX)$. Differentiating with respect to t at $t = 0$ gives $-\sum_j X^i_j F^j$. But from Proposition 2.7 we know that $X^i_j = \gamma^i_j(T_x\sigma \cdot \xi(x)) = \omega^i_j(\xi(x))$.

Taking η as in (2) and the corresponding function F, we know that $F^i \circ \sigma = f^i$ for all $i = 1, \dots, n$ so we conclude that the ith component of the function corresponding to $\nabla_\xi\eta$ is given by $\xi(x) \cdot f^i + \sum_j \omega^i_j(\xi(x))f^j(x)$. Since this is the values of $df^i + \sum_j f^j\omega^i_j$ on $\xi(x)$ this proves (2). For $\eta = \xi_i$, the function f^i is identically one while f^j is identically zero for $j \neq i$, which implies the formula for $\nabla\xi_i$ in (1).

Finally, we have $\nabla_\xi\sigma^i = \sum_j(\nabla_\xi\sigma^i)(\xi_j)\sigma^j$. Since $\sigma^i(\xi_j)$ is always constant, we conclude $(\nabla_\xi\sigma^i)(\xi_j) = -\sigma^i(\nabla_\xi\xi_j) = -\omega^i_j(\xi)$ from above. This produces the formula for $\nabla\sigma^i$ in (1). $\square$

Before we carry this out explicitly, let us combine viewing the covariant derivative as an operator with the considerations in 3.4, which will lead to more interesting examples. Take a vector field $\eta \in \mathfrak{X}(M)$ and consider the covariant derivative $\nabla\eta \in \mathcal{T}^1_1(M)$. As discussed in 3.4 we can then form $\operatorname{tr}(\nabla\eta) =: \operatorname{div}(\eta) \in C^\infty(M,\mathbb{R})$, which is called the *divergence* of the vector field η.

Alternatively, we can interpret $\nabla\eta$ as a $\binom{0}{2}$–tensor field characterized by $\nabla\eta(\xi,\zeta) = g(\nabla_\xi\eta,\zeta)$. Now it is an obvious idea to alternate respectively symmetrize this bilinear mapping. The alternation $\operatorname{Alt}(\nabla\eta)$ is given by

$$(\xi,\zeta) \mapsto g(\nabla_\xi\eta,\zeta) - g(\nabla_\zeta\eta,\xi) = \xi\cdot g(\eta,\zeta) - \zeta\cdot g(\eta,\xi) - g(\eta,\nabla_\xi\zeta - \nabla_\zeta\xi),$$

where we have used $\nabla g = 0$. Applying part (4) of Corollary 3.2 to the last term, we see that $\operatorname{Alt}(\nabla\eta) = d\alpha$, where $\alpha \in \Omega^1(M)$ is the one–form corresponding to η, i.e. $\alpha(\xi) = g(\eta,\xi)$. So this just reproduces the exterior derivative.

The symmetrization $\operatorname{Symm}(\nabla\eta)$ is given by

$$(\xi,\zeta) \mapsto g(\nabla_\xi\eta,\zeta) + g(\nabla_\zeta\eta,\xi).$$

Using part (4) of Corollary 3.2, we can rewrite $\nabla_\xi\eta$ as $\nabla_\eta\xi - [\eta,\xi]$, and likewise for $\nabla_\zeta\eta$. Using $\nabla g = 0$, we can then write $\operatorname{Symm}(\nabla\eta)(\xi,\zeta)$ as

$$\eta\cdot g(\xi,\zeta) - g([\eta,\xi],\zeta) - g(\xi,[\eta,\zeta])$$

and it turns out that this equals $(\mathcal{L}_\eta g)(\xi,\zeta)$, so $\operatorname{Symm}(\nabla\eta)$ is the Lie derivative of g along η. The resulting differential operator is called the *Killing operator* and vector fields η such that $\operatorname{Symm}(\nabla\eta) = 0$ are called *Killing vector fields*. The importance of this concept comes from the fact that $\mathcal{L}_\eta g = 0$ is equivalent to the fact that the flow of η is a local isometry where it is defined.

We can also use this to discuss an important example of an iterated covariant derivative. Let us start with a smooth function $f : M \to \mathbb{R}$. Then we know from Example (1) of 3.4 that $\nabla_\xi f = \xi\cdot f$, which means that $\nabla f = df \in \Omega^1(M)$. From Example (4) of 3.4 we know, that we can also interpret this as a vector field. This vector field is usually called the *gradient* of f and denoted by $\operatorname{grad}(f)$. By definition, this is characterized by $g(\operatorname{grad}(f),\xi) = \xi\cdot f$ for all $\xi \in \mathfrak{X}(M)$. Following what we did before, we can now form $\nabla^2 f = \nabla(\nabla f) \in \mathcal{T}^0_2(M)$ and $\nabla\operatorname{grad}(f) \in \mathcal{T}^1_1(M)$, respectively. Forming the trace, we obtain $\operatorname{div}(\operatorname{grad}(f)) =: \Delta(f) \in C^\infty(M,\mathbb{R})$. This defines the *Laplace operator* Δ which is one of the crucial ingredients in Riemannian geometry.

Example 3.5. (1) On flat space as discussed in Example (1) of 2.7 the situation is very simple. Here the forms ω^i_j vanish identically, so each of the coordinate vector fields ∂_i is parallel. For a general vector field $\sum_i f^i\partial_i$, the covariant derivative is given by $\sum_i df^i\otimes\partial_i = \sum_{i,j}\frac{\partial f^i}{\partial x^j}dx^j\otimes\partial^i$, so this is simply the derivative of the $\mathbb{R}^n$–valued function $f = (f^1,\dots,f^n)$. Hence the parallel vector fields are exactly the linear combinations of the ∂_i with constant coefficients. The Killing equation boils down to the fact that the matrix $Df(x)$ is skew symmetric for each x, i.e. that $\frac{\partial f^i}{\partial x^j} = -\frac{\partial f^j}{\partial x^i}$. From this one easily concludes that $D^2f(x) = 0$ for any x, so $Df(x)$ equals some fixed skew symmetric matrix $X \in \mathfrak{o}(n)$ for all $x \in \mathbb{R}^n$. This easily implies that our field has the form $\sum_i(a^i + \sum_j X^i_j x^j)\partial_i$ for constants a^i and X^i_j with $X^j_i = -X^i_j$. Since one easily verifies that any such vector field is a Killing vector field, this gives a complete description.

(2) Let us take the chart (U, u) on S^n defined by stereographic projection as in Example (2) of 2.7. As there, we put $f(u) := \frac{1+\sum(u^j)^2}{2}$, and have an orthonormal frame defined by $\xi_i = f\partial_i$ with corresponding coframe given by $\sigma^i = \frac{1}{f}du^i$ with connection forms $\omega^i_j = u^i\sigma^j - u^j\sigma^i$. Using that $df = \sum_j u^j du^j$, we see from Proposition 3.5 that

$$\nabla\xi_i = \textstyle\sum_j(u^j\sigma^i - u^i\sigma^j) \otimes \xi_j = \sigma^i \otimes (\sum_j u^j\xi_j) - u^i \sum_j(\sigma^j \otimes \xi_j).$$

Viewing $\nabla\xi_i$ as a $\binom{0}{2}$–tensor, we just have to replace each ξ_j in the last formula by σ^j, and then the result can be written as $\frac{1}{f}\sigma^i \otimes df - u^i g$. Since $\sigma^i = g(\ ,\xi_i)$, this also is the formula for $\nabla\sigma^i$.

From this, one can verify directly that there are no non–zero parallel vector fields or one–forms on U. We don't do this explicitly here since it will follow from general arguments later on. Rather than that, we will describe some Killing vector fields on S^n. First $\partial_i = \frac{1}{f}\xi_i$ shows that $\nabla\partial_i = -\frac{df}{f^2} \otimes \xi_i + \frac{1}{f}\nabla\xi_i$, which, as a $\binom{0}{2}$–tensor, equals $\frac{1}{f^2}(-fu^i g - df \otimes \sigma^i + \sigma^i \otimes df)$, so $\operatorname{Symm}(\nabla\partial_i) = \frac{-u^i}{f}g$. (This says that ∂_i is a so–called *conformal Killing vector field*, which has the property that the Lie derivative of g is a multiple of g.)

Now for $j \neq i$ we see that $\nabla u^j\partial_i = du^j \otimes \partial_i + u^j\nabla\partial_i$. Viewed as a $\binom{0}{2}$–tensor, this is the sum of a multiple of $du^j \otimes du^i$, $\frac{u^j u^i}{f}g$ and something skew symmetric. But this immediately implies that $\nabla(u^j\partial_i - u^i\partial_j)$ viewed as a $\binom{0}{2}$–tensor is skew symmetric, so $u^j\partial_i - u^i\partial_j \in \mathfrak{X}(U)$ is a Killing vector field.

Remark 3.5. We have noted above that for $\alpha \in \Omega^1(M)$, the covariant derivative $\nabla\alpha \in \mathcal{T}^0_2(M)$ has the property that $\operatorname{Alt}(\nabla\alpha) = d\alpha \in \Omega^2(M)$. An analog of this holds for differential forms of higher degree: For $\tau \in \Omega^k(M) \subset \mathcal{T}^0_k(M)$ and $\eta_0, \dots, \eta_k \in \mathfrak{X}(M)$ we see from example 3.4 (3) that

$$\nabla\tau(\eta_0, \dots, \eta_k) = \eta_0 \cdot \tau(\eta_1, \dots, \eta_k) - \sum_{j=1}^k \tau(\eta_1, \dots, \nabla_{\eta_0}\eta_j, \dots, \eta_k).$$

From this it follows immediately that this expression is already alternating in $\eta_1, \dots, \eta_k$. This implies that we can compute $\frac{1}{k!}\operatorname{Alt}(\nabla\tau)(\eta_0, \dots, \eta_k)$ as

$$\sum_{i=0}^k (-1)^i \nabla\tau(\eta_i, \eta_0, \dots, \eta_{i-1}, \eta_{i+1}, \dots, \eta_k).$$

The second term in the above formula (including the minus sign) can be rewritten as

$$\sum_{j=1}^k (-1)^j \tau(\nabla_{\eta_0}\eta_j, \eta_1, \dots, \widehat{\eta_j}, \dots, \eta_k).$$

Performing the alternation as indicated above, the term $\nabla_{\eta_i}\eta_j$ occurs exactly once for each $i \neq j$ and the sign is $(-1)^{i+j}$ for $i < j$ and $(-1)^{i+j+1}$ for $i > j$. Hence using $\nabla_{\eta_i}\eta_j - \nabla_{\eta_j}\eta_i = [\eta_i, \eta_j]$ we arrive at $\frac{1}{k!}\operatorname{Alt}(\nabla\tau) = d\tau$.

Parallel transport and geodesics

3.6. Parallel transport. We now discuss how the Levi–Civita connection can be used to relate the tangent spaces of M along a curve. Let $c : I \to M$ be a smooth curve and let $\xi \in \mathfrak{X}(M)$ be a vector field. Then we say that ξ is *parallel along* c if $0 = \nabla_{c'(t)}\xi \in T_{c(t)}M$ holds for all $t \in I$. Our first task is to show that this actually depends only on the restriction of ξ to the subset $c(I) \subset M$. In fact, the right concept is the following:

Definition 3.6. For a smooth curve $c : I \to M$ in a smooth manifold M, a *vector field along* c is a smooth map $\xi : I \to TM$ such that $p \circ \xi = c$, i.e. such that $\xi(t) \in T_{c(t)}M$ for all $t \in I$.

Observe in particular that c' is a vector field along c in this sense. Now one can show that for a vector field ξ along c, one obtains a well defined vector field $\nabla_{c'}\xi$ along c. One way to prove this is to choose vector fields $\tilde{\xi}, \tilde{\eta} \in \mathfrak{X}(M)$ such that $\tilde{\xi}(c(t)) = \xi(t)$ and $\tilde{\eta}(c(t)) = c'(t)$ holds for all $t \in I$ and then show that $\nabla_{\tilde{\eta}}\tilde{\xi}(c(t))$ is independent of the choices for all t. We will follow a different, more conceptual approach based on lifts of curves. A *lift* of c to $\mathcal{P}M$ is a smooth curve $\tilde{c} : I \to \mathcal{P}M$ such that $p \circ \tilde{c} = c$. Such a lift is called *horizontal* if and only if $\gamma^i_j(\tilde{c}'(t)) = 0$ holds for all $t \in I$ and all i, j.

Proposition 3.6. *Let (M, g) be a Riemannian manifold with orthonormal frame bundle $\mathcal{P}M$, let $c : I \to M$ be a smooth curve defined on $I = [a, b] \subset \mathbb{R}$.*

For any $\varphi_0 \in p^{-1}(c(a)) \subset \mathcal{P}M$ there exists a unique horizontal lift $c^h_{\varphi_0} : I \to M$ which maps a to φ_0. For $A \in O(n)$ we have $c^h_{r^A(\varphi_0)} = r^A \circ c^h_{\varphi_0}$.

Proof. Let $U \subset M$ be an open subset such that there is a local orthonormal frame for M defined on U and let $\Phi : p^{-1}(U) \to U \times O(n)$ be the induced diffeomorphism as in Proposition 2.3. If $c(I) \subset U$, then $\tilde{c} : I \to p^{-1}(U)$ is a lift of c if and only if $\Phi(\tilde{c}(t)) = (c(t), A(t))$ for some smooth curve $A : I \to O(n)$. In particular, this shows that lifts exist in this situation. Moreover, if $\tilde{c}_1, \tilde{c}_2$ are two such lifts, then there is a smooth curve $B : I \to O(n)$ such that $\tilde{c}_2(t) = r^{B(t)}(\tilde{c}_1(t))$. (This curve is given by $B(t) = A_1(t)^{-1}A_2(t)$.) Conversely, given $\tilde{c}_1$ and B, we can use this to define a lift $\tilde{c}_2$.

Differentiating the curve $t \mapsto A_1(t)B(t)$ in $O(n)$, we get $A_1'(t)B(t) + A_1(t)B'(t)$. Moreover $B'(t) \in T_{B(t)}O(n)$ and thus $B(t)^{-1}B'(t) =: X(t) \in \mathfrak{o}(n)$ and this defines a smooth curve $X : I \to \mathfrak{o}(n)$. In terms of this, our derivative can be written as $A_1'(t)B(t) + A_1(t)B(t)X(t)$. Carrying this back via Φ, we conclude that

$$\tilde{c}_2'(t) = T_{r^{B(t)}} \cdot \tilde{c}_1'(t) + \zeta_{X(t)}(\tilde{c}_1(t)).$$

Next let us suppose that $\tilde{c}_1$ is horizontal. Then $\gamma^i_j(\tilde{c}_1'(t)) = 0$ for all t, and since the γ^i_j define a connection form, we also get $\gamma^i_j(T_{r^{B(t)}} \cdot \tilde{c}_1'(t)) = 0$ for all i and j. Hence we conclude that $\gamma^i_j(\tilde{c}_2'(t))$ coincides with the i, j–component of $X(t)$, so $\tilde{c}_2'$ is also horizontal if and only if $X(t) = 0$ for all t. This is equivalent to $B'(t) = 0$, i.e. to $B(t)$ being constant. Hence (still assuming that $c(I) \subset U$), there is at most one horizontal lift with a chosen initial value and then all horizontal lifts are obtained by applying r^B for fixed $B \in O(n)$.

Starting from an arbitrary lift $\tilde{c}_1$, we can view the $\gamma^i_j(\tilde{c}_1'(t))$ as defining a smooth curve $Y(t)$ in $\mathfrak{o}(n)$. Making the ansatz $\tilde{c}_2(t) = r^{B(t)}(\tilde{c}_1(t))$, we conclude from the above formula for $\tilde{c}_2'(t)$ that the curve in $\mathfrak{o}(n)$ corresponding to $\gamma^i_j(\tilde{c}_2'(t))$ is given by $B(t)^{-1}Y(t)B(t) + B(t)^{-1}B'(t)$. Thus we conclude that in order to find a horizontal lift, we have to solve the differential equation $B'(t) = Y(t)B(t)$ on curves in $O(n)$ with the curve Y in $\mathfrak{o}(n)$ being given.

Now for $Y \in \mathfrak{o}(n)$ and $A \in O(n)$, one immediately sees that $YA \in T_AO(n)$. Hence we conclude that $E_{(t,A)} := \mathbb{R} \cdot (1, Y(t)A) \subset \mathbb{R} \times T_AO(n) \cong T_{(t,A)}(I \times O(n))$ defines a smooth one–dimensional distribution on $I \times O(n)$. By the Frobenius Theorem, there are local integral submanifolds of this distribution, and if N is an integral manifold, then by definition the tangent map of the first projection restricts to a linear isomorphism on each tangent space of M, so this is a local diffeomorphisms around each point. Inverting this, and composing with the second projection, we can interpret N as the graph of a smooth function, and by construction this solves the equation. Hence we get local unique solutions with given initial values. By compactness of $[a, b]$ these can be pieced together to a global solution. This completes the proof in the case that $c(I) \subset U$.

In general, compactness of $[a, b]$ implies that we can find finitely many subsets $U_1, \dots, U_n$ as above and points $a = a_1 < a_2 < b_1 < \dots < a_n < b_n = b$ such that $c([a_i, b_i]) \subset U_i$. Fixing $\varphi_0 \in \mathcal{P}_{c(a)}M$ we can construct a horizontal lift of $\tilde{c}_1 : [a_1, b_1] \to \mathcal{P}M$ of $c|_{[a_1,b_1]}$ with $\tilde{c}_1(a_1) = \varphi_0$. Next take $\tilde{c}_1(a_2) \in \mathcal{P}_{c(a_2)}M$ and construct a horizontal lift $\tilde{c}_2 : [a_2, b_2] \to \mathcal{P}M$ of $c|_{[a_2,b_2]}$ with this initial value. Then by construction $\tilde{c}_1|_{[a_2,b_1]}$ and $\tilde{c}_2|_{[a_2,b_1]}$ are horizontal lifts of $c|_{[a_2,b_1]}$, which have the same value in a_2 and thus have to agree. So these curves piece together to a smooth curve on $[a_1, b_2]$. Iterating this procedure, we get a horizontal lift of c with initial value φ_0. □

Now suppose that $c : [a, b] \to M$ is a smooth curve and $\xi : I \to TM$ is a vector field along c. Take a horizontal lift $\tilde{c} : I \to \mathcal{P}M$ of c, and define $F : I \to \mathbb{R}^n$ by $F(t) := (\tilde{c}(t))^{-1}(\xi(t))$, where keep in mind that $\tilde{c}(t) \in \mathcal{P}_{c(t)}M$ is an orthogonal linear isomorphism $\mathbb{R}^n \to T_{c(t)}M$. Now take the derivative $F'(t)$ and consider the tangent vector $\eta(t) := (c(t))(F'(t)) \in T_{c(t)}M$, where we interpret $c(t)$ as a linear isomorphism as above.

If $\hat{c} : I \to \mathcal{P}M$ is another horizontal lift of c, then there is a unique element $A = (a^i_j) \in O(n)$ such that $\hat{c}(a) = r^A(\tilde{c}(a))$. By Proposition 3.6 we have $\hat{c} = r^A \circ \tilde{c}$, i.e. $\hat{c}(t) = \tilde{c}(t) \circ A$ as a linear isomorphism. Starting from the horizontal lift $\hat{c}$ instead of $\tilde{c}$, we thus get the function $A^{-1} \circ F$ instead of F. Since A is linear, we get $(A^{-1} \circ F)' = A^{-1} \circ F'$ and hence $(\hat{c}(t))((A^{-1} \circ F)'(t)) = (\tilde{c}(t))(F'(t))$ so the tangent vector $\eta(t)$ is independent of the choice of the horizontal lift. Hence we can use this to get a well defined vector field $\nabla_{c'}\xi$ along c.

Let us finally fix a horizontal lift $\tilde{c}$ of c and choose vector fields $\tilde{\xi}, \tilde{\eta} \in \mathfrak{X}(M)$ such that $\tilde{\xi}(c(t)) = \xi(t)$ and $\tilde{\eta}(c(t)) = c'(t)$. Then the smooth equivariant function $\tilde{F} : \mathcal{P}M \to \mathbb{R}^n$ corresponding to $\tilde{\xi}$ by definition satisfies $\tilde{F}(\tilde{c}(t)) = F(t)$. Moreover, since $\tilde{c}$ is horizontal, $\tilde{c}'(t)$ is the horizontal lift of $T_{\tilde{c}(t)}p \cdot \tilde{c}'(t) = (p \circ \tilde{c})'(t) = c'(t) = \tilde{\eta}(c(t))$. Since $\tilde{c}'(t) \cdot \tilde{F} = (F \circ \tilde{c})'(t) = F'(t)$ this shows that $\nabla_{c'(t)}\xi(t) = \nabla_{\tilde{\eta}}\tilde{\xi}(c(t))$. This proves the first part of

Theorem 3.6. *Let (M, g) be a Riemannian manifold and $c : [a, b] \to M$ a smooth curve.*

(1) For any vector field $\xi : I \to TM$ along c, there is a well defined vector field $\nabla_{c'}\xi$ along c such that for vector fields $\tilde{\xi}, \tilde{\eta} \in \mathfrak{X}(M)$ with $\tilde{\xi}(c(t)) = \xi(t)$ and $\tilde{\eta}(c(t)) = c'(t)$ we have $\nabla_{\tilde{\eta}}\tilde{\xi}(c(t)) = \nabla_{c'}\xi(t)$.

(2) For any tangent vector $\xi_0 \in T_{c(a)}M$, there is a unique vector field $\mathrm{Pt}^{\xi_0}_c : I \to TM$ along c such that $\mathrm{Pt}^{\xi_0}(a) = \xi_0$ and $\nabla_{c'}\mathrm{Pt}^{\xi_0} = 0$.

(3) Mapping ξ_0 to $\mathrm{Pt}^{\xi_0}(b)$ defines an orthogonal linear isomorphism $\mathrm{Pt}_c : T_{c(a)}M \to T_{c(b)}M$.

(4) The parallel transport is invariant under orientation preserving reparametrizations, i.e. if $\psi : [a', b'] \to [a, b]$ is an orientation preserving diffeomorphism then $\mathrm{Pt}_{c\circ\psi} = \mathrm{Pt}_c$.

Proof. We have proved (1) already. To prove (2), choose a horizontal lift $\tilde{c}$ of c and put $X := (\tilde{c}(a))^{-1}(\xi_0) \in \mathbb{R}^n$, where we view $\tilde{c}(a) \in \mathcal{P}_aM$ as a linear isomorphism $T_{c(a)}M \to \mathbb{R}^n$. Then we define $\mathrm{Pt}^{\xi_0}_c(t) := (\tilde{c}(t))(X) \in T_{c(t)}M$. Of course, this defines a vector field along c such that $\mathrm{Pt}_c^{\xi_0}(a) = \xi_0$. Moreover, by construction $\mathrm{Pt}_c^{\xi_0}$ corresponds to the constant function X via $\tilde{c}$, so $\nabla_{c'} \mathrm{Pt}_c^{\xi_0} = 0$. Conversely, if we have a vector field $\xi : I \to TM$ along c such that $\nabla_{c'}\xi = 0$, then the function $F(t) = (\tilde{c}(t))^{-1}(\xi(t))$ must be constant and if $\xi(a) = \xi_0$ then $F(a) = X$, which proves uniqueness.

(3) From the proof of part (2) we see that we can write Pt_c as $\tilde{c}(b) \circ (\tilde{c}(a))^{-1}$ for any horizontal lift $\tilde{c}$ of c.

(4) If $\tilde{c}$ is a horizontal lift of c, then evidently $\tilde{c} \circ \psi$ is a lift of $c \circ \psi$ and since $(\tilde{c} \circ \psi)'(t) = \tilde{c}'(\psi(t))\psi'(t)$, this lift is horizontal. Now the result immediately follows from (3). □

Remark 3.6. (1) The concepts related to parallel transport generalize without essential changes to tensor fields. One defines tensor fields along a smooth curve c similarly to vector fields along c and shows that for a tensor field t along c, one has a well defined tensor field $\nabla_{c'}t$ along c, which can also be computed using extensions. Prescribing the value in the initial point of c, one can then uniquely construct a parallel tensor field along c, and use this to define parallel transport between tensor spaces.

(2) The parallel transport is the crucial ingredient for the concept of holonomy, which is one of the cornerstones of Riemannian geometry. Starting from a point $x_0 \in M$, one considers piecewise smooth closed curves starting and ending in x_0. For each such curve, one obtains an orthogonal map $\mathrm{Pt}_c : T_{x_0}M \to T_{x_0}M$ and this maps constitute a subgroup of the orthogonal group $O(T_{x_0}M)$. Choosing an orthogonal identification with $\mathbb{R}^n$, this gives a subgroup $\mathrm{Hol}_{x_0}(M) \subset O(n)$ which is well defined up to conjugation. One easily shows that, provided that M is connected, different choices of x_0 lead to conjugate subgroups in $O(n)$, so one usually talks about the *holonomy group* of M as a subgroup of $O(n)$ which is well defined up to conjugation.

The importance of the concept of holonomy comes from the fact that by works of G. deRham and M. Berger one can describe all possible holonomy groups.

3.7. Geodesics. From the theory developed so far, we now get a class of distinguished curves in any Riemannian manifold in a very simple way. Recall from 3.6 that for a smooth curve $c : I \to M$, the derivative $c' : I \to TM$ defines a vector field along c. Hence we can simply require that c' is parallel along c.

Definition 3.7. A smooth curve $c : I \to M$ in a Riemannian manifold is called a *geodesic* in M if $\nabla_{c'}c' = 0$ along c.

With the tools developed so far, we can easily prove existence of geodesics. We first observe that the soldering forms θ^i together with the Levi–Civita connection forms γ^i_j induces a trivialization of the tangent bundle $T\mathcal{P}M$ of the orthonormal frame bundle. More precisely, consider the map $T\mathcal{P}M \to \mathcal{P}M \times (\mathbb{R}^n \oplus \mathfrak{o}(n))$ defined by mapping $\xi \in T_\varphi\mathcal{P}M$ to $(\varphi, (\theta^i(\xi)), (\gamma^i_j(\xi)))$. This is evidently smooth and it maps onto the first factor. Moreover, from Proposition 2.4 we know that the joint kernel of the forms θ^i on a tangent space of $T_\varphi\mathcal{P}M$ coincides with the vertical subspace $\ker(T_p\varphi)$. From the definition of a connection form it follows that $(\gamma^i_j(\varphi))$ is injective on this subspace. This easily implies that our mapping as well as its derivative in a point is bijective. As a bijective local diffeomorphism it has to be a diffeomorphism.

This in turn induces a bijection between the space $\mathfrak{X}(\mathcal{P}M)$ of vector fields and the space $C^\infty(\mathcal{P}M, \mathbb{R}^n \oplus \mathfrak{o}(n))$. In particular, we can consider the *constant vector fields* on $\mathcal{P}M$, which correspond to constant functions. The constant vector field corresponding to $X \in \mathfrak{o}(n) \subset \mathbb{R}^n \oplus \mathfrak{o}(n)$ is simply the fundamental vector field ζ_X by definition of a connection form, so nothing new is obtained there. Hence the main new ingredient is formed by the constant vector fields $\tilde{v} \in \mathfrak{X}(\mathcal{P}M)$ corresponding to elements $v \in \mathbb{R}^n$. Observe that the definition of the soldering form immediately implies that for a point $\varphi \in \mathcal{P}M$, the orthogonal isomorphism $\varphi : \mathbb{R}^n \to T_{p(\varphi)}M$ is characterized by $\varphi(v) = T_\varphi p \cdot \tilde{v}(\varphi)$.

For later use, we also observe that the subspace $H_\varphi\mathcal{P}M \subset T_\varphi\mathcal{P}M$ formed by the vectors $\tilde{v}(\varphi)$ always has dimension n. Since they coincide with the joint kernel of the forms γ^i_j in the point φ, we conclude from Proposition 1.3 that these spaces fit together to define a smooth distribution $H\mathcal{P}M \subset T\mathcal{P}M$ of rank n. This is called the *horizontal distribution* of the Levi–Civita connection.

Lemma 3.7. *Let (M, g) be a Riemannian manifold, $x \in M$ a point and $\xi \in T_xM$ a tangent vector. Then there exists a unique maximal open interval $I_x \subset \mathbb{R}$ containing* 0 *and a unique maximal geodesic $c : I_x \to M$ such that $c(0) = x$ and $c'(0) = \xi$.*

Proof. We first prove local existence and uniqueness. Let $p : \mathcal{P}M \to M$ be the orthonormal frame bundle of M, choose a point $\varphi \in p^{-1}(x)$ and consider the horizontal lift $\xi^h \in T_\varphi\mathcal{P}M$ of ξ (see Proposition 3.1). Put $v := (\theta^1(\xi^h), \dots, \theta^n(\xi^h)) \in \mathbb{R}^n$ and consider the associated constant vector field $\tilde{v} \in \mathfrak{X}(M)$.

Suppose that $\tilde{c} : I \to \mathcal{P}M$ is an integral curve of $\tilde{v}$ defined on an interval I containing 0 such that $\tilde{c}(0) = \varphi$ and consider the smooth curve $c = p \circ \tilde{c} : I \to M$. Of course we have $c(0) = x$ and $c'(0) = Tp\cdot\tilde{c}'(0) = Tp\cdot\xi^h = \xi$. Moreover, by definition $\tilde{c}'(t) = \tilde{v}(\tilde{c}(t))$, so $\gamma^i_j(\tilde{c}'(t)) = 0$ for all i and j. This means that $\tilde{c}$ is the horizontal lift of c with initial value φ. Hence we can compute $\nabla_{c'}c'$ using this horizontal lift. But by construction, the vector field c' along c corresponds via the horizontal lift $\tilde{c}$ to the constant function v. Thus $\nabla_{c'}c' = 0$ along c, so c is a geodesic.

Any other choice for the initial point in $p^{-1}(x)$ is of the form $r^A(\varphi)$ for some $A \in O(n)$. Now consider the curve $\hat{c} := r^A \circ \tilde{c} : I \to \mathcal{P}M$. Of course $\hat{c}(0) = r^A(\varphi)$ and $\hat{c}'(t) = T_{\tilde{c}(t)}r^A \cdot \tilde{c}'(t)$. Hence $\theta^i(\hat{c}'(t)) = (r^A)^*\theta^i(\tilde{c}'(t))$ which is constant by equivariancy of the soldering form. Likewise, $\gamma^i_j(\hat{c}'(t)) = (r^A)^*\gamma^i_j(\tilde{c}'(t))$ which vanishes by equivariancy of the connection forms. Thus $\hat{c}$ is again an integral curve of a constant vector field, and since $Tp \cdot \hat{c}'(0) = Tp \circ Tr^A \cdot \tilde{c}'(0) = \xi$, this is the integral curve one gets when starting the above construction from $r^A(\varphi)$ instead of φ. But since $p \circ \hat{c} = p \circ r^A \circ \tilde{c} = p \circ \tilde{c} = c$, a different initial point leads to the same curve c.

Conversely, given a geodesic $c : I \to M$ with $c(0) = x$ and $c'(0) = \xi$, consider a horizontal lift $\tilde{c}$ of c. Then by definition $\gamma^i_j(\tilde{c}'(t)) = 0$, whence the function $I \to \mathbb{R}^n$ corresponding to the vector field c' along c can be written as $(\theta^i(\tilde{c}'(t)))$. Since $\nabla_{c'}c' = 0$, this must be constant, so $\tilde{c}$ is an integral curve of a constant vector field $\tilde{v}$ with $v \in \mathbb{R}^n$. Since $p \circ \tilde{c} = c$, we must have $Tp \cdot \tilde{c}'(0) = \xi$ so $\tilde{c}'(0)$ coincides with the horizontal lift of ξ to the point $\tilde{c}(0)$.

Having established the correspondence with integral curves of constant vector fields, we conclude that locally there is a unique geodesic with given initial point and initial direction. Then one can piece together local geodesics to obtain a unique maximal geodesic with the given initial data. □

From the construction of geodesics, we can easily deduce that they are preserved under isometries. This will lead to surprising restrictions on how many isometries may exist.

Proposition 3.7. *Let M and $\tilde{M}$ be two Riemannian manifolds of the same dimension and $\Phi : M \to \tilde{M}$ an isometry. Then for any geodesic $c : I \to M$ in M, the image $\Phi \circ c : I \to \tilde{M}$ is a geodesic, too.*

Proof. In the proof of Corollary 3.2, we have seen that Φ uniquely lifts to a local diffeomorphism $\mathcal{P}\Phi : \mathcal{P}M \to \mathcal{P}\tilde{M}$ such that $\tilde{p} \circ \mathcal{P}\Phi = \Phi \circ p$, $\mathcal{P}\Phi$ is equivariant for the principal right actions, $(\mathcal{P}\Phi)^*\tilde{\theta}^i = \theta^i$ and $(\mathcal{P}\Phi)^*\tilde{\gamma}^i_j = \gamma^i_j$. This in particular implies that $\mathcal{P}\Phi$ pulls back constant vector fields on $\mathcal{P}M$ to constant vector fields on $\mathcal{P}\tilde{M}$ with the same generator in $\mathbb{R}^n$.

Now given c, take $t \in I$, put $x = c(t)$ and $\xi := c'(t)$, choose $\varphi \in \mathcal{P}_xM$ and consider $\mathcal{P}\Phi(\varphi) \in \mathcal{P}_{\Phi(x)}\tilde{M}$. Then from the proof of Corollary 3.2 we know that $\mathcal{P}\Phi(\varphi)(T_x\Phi \cdot \xi) = \varphi(\xi) =: v \in \mathbb{R}^n$. Now let $\eta \in \mathfrak{X}(\mathcal{P}M)$ and $\tilde{\eta} \in \mathfrak{X}(\mathcal{P}\tilde{M})$ be the constant vector fields corresponding to v. Then form above we know that $(\mathcal{P}\Phi)^*\tilde{\eta} = \eta$.

From Lemma 3.7, we know that c can be locally around x written as the projection of the integral curve $\hat{c}$ of η through φ. From the construction, we conclude that $\mathcal{P}\Phi \circ \hat{c}$ is an integral curve of $\tilde{\eta}$ through $\mathcal{P}\Phi(\varphi)$. But this shows that $\tilde{p} \circ \mathcal{P}\Phi \circ \hat{c} = \Phi \circ p \circ \hat{c} = \Phi \circ c$ is a geodesic in $\tilde{M}$. □

To proceed towards an elementary characterization of geodesics, we give a description of the covariant derivative in local coordinates. Suppose that (M, g) is a Riemannian manifold and (U, u) is a chart of M. Then consider the associated coordinate vector fields $\partial_i := \frac{\partial}{\partial u^i}$. Then for $i, j, k = 1, \dots, n$, we define the *Christoffel–Symbols* $\Gamma^k_{ij} : U \to \mathbb{R}$ by $\nabla_{\partial_i}\partial_j = \sum_k \Gamma^k_{ij}\partial_k$. Observe that since $[\partial_i, \partial_j] = 0$, the Christoffel symbols are symmetric in the lower indices, i.e. $\Gamma^k_{ji} = \Gamma^k_{ij}$ for all i, j, k. Now we extend the Christoffel symbols in each point to a bilinear map $T_xM \times T_xM \to T_xM$ by putting

$$\Gamma(\textstyle\sum_i \xi^i\partial_i(x), \sum_j \eta^j\partial_j(x)) := \sum_{i,j,k} \xi^i\eta^j\Gamma^k_{ij}\partial_k(x).$$

One has to be careful at this stage that the Christoffel symbols have a complicated behavior under a change of coordinates. In particular, these bilinear maps do *not* fit together to define a tensor field. We will not need the explicit transformation law, although computing it is a nice exercise.

From the definition of the covariant derivative, it follows immediately how to express a general covariant derivative in terms of the Christoffel symbols. Indeed, for $\xi = \sum_i \xi^i\partial_i$ and $\eta = \sum_j \eta^j\partial_j$ we get

$$\nabla_\xi\eta = \textstyle\sum_{i,j} \xi^i\nabla_{\partial_i}(\eta^j\partial_j) = \sum_{i,j} \xi^i\frac{\partial\eta^j}{\partial u^i}\partial_j + \Gamma(\xi, \eta),$$

so the Christoffel symbols in general measure the difference between the covariant derivative and the component–wise directional derivative of η in direction ξ.

In these terms, we can immediately express the differential equation characterizing geodesics in local coordinates. For a smooth curve $c : I \to U$ we can expand the derivative as $c'(t) = \sum_i \xi^i\partial_i(c(t))$, and then extend the ξ^i in some way to locally defined smooth functions. Then the components of $c''(t)$ are of course given by

$$\tfrac{d}{dt}\xi^i(c(t)) = D\xi^i(c(t))(c'(t)) = \textstyle\sum_j \xi^j(c(t))\frac{\partial\xi^i}{\partial u^j}.$$

This shows that c is a geodesic if and only if $c''(t) + \Gamma(c'(t), c'(t)) = 0$ for all t. Hence in local coordinates geodesics are the solutions of a system of second order ODEs, which explains Lemma 3.7 and provides an alternative proof of this result.

Example 3.7. (1) On the flat space $\mathbb{R}^n$, the Christoffel symbols vanish identically by definition. Thus the geodesic equation reduces to $c''(t) = 0$ and geodesics are the linearly parametrized straight lines.

(2) Let us consider the chart on S^n defined by stereographic projection as in Examples 2.7 and 3.5. So on $\mathbb{R}^n$ we consider the vector fields $\xi^i = f\partial^i$, where $f(u) = \frac{1+\sum(u^j)^2}{2}$.

From Example 3.5 we get $\nabla_{\xi^i}\xi^i = \sum_{j\neq i} u^j\xi^j$. Now the curve $c(t) = \tan(t/2)e_i$ evidently satisfies $c'(t) = \frac{1}{2}(\tan^2(t/2))e_i = f(c(t))\partial_i$, so $c(t)$ is an integral curve for ξ^i. From above, we thus get $\nabla_{c'}c'(t) = 0$, so c is a geodesic. Under the chart defined by stereographic projection (see 2.7) this corresponds to the curve $x^i = \sin(t)$, $x^{n+1} = \cos(t)$, so this is simply a great circle in its natural parametrization. One can verify in a similar way that $c(t) = \tan(t/2)v$ defines a geodesic for any unit vector v, and this gives all great circles through the south pole. Note however, that straight lines which do not go through zero do not correspond to geodesics on S^n (with any parametrization).

3.8. The variational characterization of geodesics. While this aspect of geodesics is rather an aside in our approach it is the basis for the "elementary" development of Riemannian geometry and it gives intuitive insight. The basic idea is that apart from the arc length as discussed in 2.1, there is also a natural notion of energy that one can associate to a curve in a Riemannian manifold: Given a smooth curve $c : [a, b] \to M$, one defines the *energy* of c as

$$E_a^b(c) := \tfrac{1}{2}\int_a^b g(c(t))(c'(t), c'(t))dt.$$

From a physics point of view, this is the amount of work needed to move a particle of unit mass along the curve c. Now there is a natural concept of smooth families of curves, namely a smooth map $[a, b] \times (-\epsilon, \epsilon) \to M$. What we want to show is that the geodesics in M are exactly the curves which are critical points for the energy functional, assuming that the endpoints are fixed. The first step towards this is to derive a formula for the covariant derivative which is very helpful in other situations, too. This can be used to give a formula for the Christoffel symbols on the domain of a chart in terms of the functions g_{ij} describing the metric in this chart.

Proposition 3.8. *Let (M, g) be a Riemannian manifold, and let ∇ be the covariant derivative.*

(1) ("Koszul formula") For vector fields $\xi, \eta, \zeta \in \mathfrak{X}(M)$ we have

$$\begin{aligned} 2g(\nabla_\xi\eta, \zeta) = \xi \cdot g(\eta, \zeta) - \zeta \cdot g(\xi, \eta) + \eta \cdot g(\zeta, \xi) \\ + g([\xi, \eta], \zeta) + g([\zeta, \xi], \eta) - g([\eta, \zeta], \xi). \end{aligned}$$

(2) Let (U, u) be a chart on M with coordinate vector fields ∂_i, let g_{ij} be the functions describing the metric in this chart and let g^{ij} be the components of the pointwise inverse matrix. Then the Christoffel symbols in our chart are given by

$$\Gamma_{ij}^k = \sum_\ell g^{k\ell}\tfrac{1}{2}\left(\tfrac{\partial g_{\ell j}}{\partial u^i} + \tfrac{\partial g_{i\ell}}{\partial u^j} - \tfrac{\partial g_{ij}}{\partial u^\ell}\right)$$

PROOF. (1) We write out the fact that ∇ is metric from part (3) of Corollary 3.2 three times with cyclically permuted arguments and multiply the middle line by -1 to get

$$\begin{aligned} \xi \cdot g(\eta, \zeta) &= g(\nabla_\xi\eta, \zeta) + g(\eta, \nabla_\xi\zeta) \\ -\zeta \cdot g(\xi, \eta) &= -g(\nabla_\zeta\xi, \eta) - g(\xi, \nabla_\zeta\eta) \\ \eta \cdot g(\zeta, \xi) &= g(\nabla_\eta\zeta, \xi) + g(\zeta, \nabla_\eta\xi) \end{aligned}$$

Now we add up these three lines, use $\nabla_\xi\zeta - \nabla_\zeta\xi = [\xi, \zeta]$ and likewise for η and ζ. Finally, we can replace $\nabla_\xi\eta + \nabla_\eta\xi$ by $2\nabla_\xi\eta - [\xi, \eta]$ to obtain the result.

(2) Observe first that since the ∂_i form a basis for each tangent space, a tangent vector $\xi \in T_xM$ is uniquely determined by the numbers $b_\ell := g_x(\xi, \partial_\ell(x))$ for $\ell = 1, \dots, n$. If $\xi = \sum a^k \partial_k(x)$ for real numbers a^k, then $b_\ell = \sum_k a^k g_{k\ell}$ and hence $a^k = \sum_\ell g^{k\ell} b_\ell$. Now by definition $\nabla_{\partial_i}\partial_j = \sum_k \Gamma^k_{ij}\partial_k$, so we see that

$$\Gamma^k_{ij}(x) = \sum_\ell g^{k\ell} g_x(\nabla_{\partial_i}\partial_j, \partial_\ell).$$

computing the left hand side using the Koszul formula, all terms involving Lie brackets vanish, while $\partial_i \cdot g(\partial_j, \partial_k) = \frac{\partial g_{jk}}{\partial u^i}$ and similarly for the other summands. □

Remark 3.8. Note that in the proof of the Koszul formula we have not really used how the covariant derivative is defined, but only properties (3) and (4) from Corollary 3.2. So the proof actually shows that the covariant derivative is uniquely determined by these properties. The second part of the proposition then computes the covariant derivative of coordinate vector fields in terms of the functions g_{ij}. In an "elementary" approach to Riemannian geometry, one may take the observations on critical points for the energy that we are going to make next as a motivation, then use the formula in (2) to define the covariant derivative for coordinate vector fields and extend this to general vector fields via requiring properties (1) and (2) of Corollary 3.2.

To discuss critical points for the energy functional, we work in a chart. So we have given a family $c_s(t) := c(t,s)$ for a smooth map c from $[a,b] \times (-\epsilon, \epsilon)$ to some open subset V of $\mathbb{R}^n$. This set is endowed with a metric given by functions g_{ij}, and we want to compute $\frac{\partial}{\partial s}|_{s=0} E^b_a(c_s)$. Since everything depends smoothly on s, we can exchange integration with the partial derivative, so we first have to compute

$$\tfrac{\partial}{\partial s} g(c(s,t))(c'(s,t), c'(s,t)),$$

where $c'(s,t) = \frac{\partial}{\partial t} c(s,t)$. Recall that positive definite matrices form an open subset $\mathcal{S}_+(n)$ in the vector space of symmetric matrices. Viewing $g = (g_{ij})$ as a smooth function $V \to \mathcal{S}_+(n)$, our expression becomes $\langle g(c(s,t))c'(s,t), c'(s,t)\rangle$ so this comes from a trilinear map. Let us write $r(s,t) := \frac{\partial}{\partial s} c(s,t)$ so this is the variation of c. Using this, the derivative is given by

$$\tfrac{\partial}{\partial s}\langle (g \circ c)c', c'\rangle = \langle D(g \circ c)(r)c', c'\rangle + 2\langle (g \circ c)r', c'\rangle,$$

where we have used the fact that partial derivatives commute as well as symmetry of the matrix. Now if we integrate this with respect to dt, we can partially integrate in the second summand in order to get rid of the derivative of r. This leads to the *first variational formula*

$$\begin{aligned}\tfrac{\partial}{\partial s} E^b_a(c_s) =& g(c_s(b))(c_s'(b), r_s(b)) - g(c_s(a))(c_s'(a), r_s(a)) \\ &+ \int_a^b \left(\tfrac{1}{2}\langle D(g \circ c)(r)c', c'\rangle - \langle D(g \circ c)(c')r, c'\rangle - \langle (g \circ c)r, c''\rangle\right) dt\end{aligned}$$

The nice feature of this is that the value for fixed s, only depends on $r(\_, s)$ and not on its derivatives. Moreover, if we want to consider variations with fixed end points, then $r_s(a) = r_s(b) = 0$, so we do not have to consider the boundary terms. But writing out the first two terms in the integral in coordinates and using symmetry for the second one, we get

$$\textstyle\sum_{ijk} \frac{1}{2}(c')^i(c')^j r^k \left(\frac{\partial g_{ij}}{\partial u^k} - \frac{\partial g_{ik}}{\partial u^j} - \frac{\partial g_{jk}}{\partial u^i}\right) = -g(\Gamma(c', c'), r),$$

where we have used part (2) of Proposition 3.8 in the last step. This implies that the whole integrand in the first variational formula can be written as

$$-g(c_s''(t)+\Gamma(c_s'(t),c_s'(t)),r(t))dt.$$

Evidently, this vanishes at $s=0$ for any choice of r if and only if $c_0''(t)+\Gamma(c_0'(t),c_0'(t))=0$ for all t, i.e. if and only if c_0 is a geodesic. So we conclude that the geodesics are exactly critical points of the energy functional for the variations fixing the end points.

3.9. The exponential mapping. We can next use geodesics to construct canonical charts on a Riemannian manifold. There are two slightly different ways to formulate this, for either of which one has to fix a point $x\in M$. Then one can either get a canonical chart with center x that has values in the tangent space T_xM. The other option is to get a family of charts centered at x with values in $\mathbb{R}^n$, which is parametrized by the fiber $\mathcal{P}_xM$ and hence essentially by $O(n)$. We start the discussion with the first possibility, for which one usually considers the local parametrization inverse to the chart map as the main object.

Let (M,g) be a Riemannian manifold, $x\in M$ a point and $\xi\in T_xM$ be a tangent vector at the point x. Then by Lemma 3.7 there is a maximal geodesic $c:I\to M$ emanating from x in direction ξ, and if $1\in I$, then we define $\exp_x(\xi):=c(1)\in M$. Thus we obtain a map $\exp_x$ from some subset of T_xM to M, called the *exponential mapping.*

Proposition 3.9. *Let (M,g) be a Riemannian manifold and $x\in M$ a point.*

(1) There is an open neighborhood U of zero in T_xM such that $\exp_x$ defines a smooth map $U\to M$. This can be chosen in such a way that for $\xi\in U$ and $r\in[0,1]$ we also have $r\xi\in U$. Moreover, $T_0\exp_x=\mathrm{id}_{T_xM}$ so $\exp_x$ is a diffeomorphisms locally around zero.

(2) For any $\xi\in T_xM$, the geodesic emanating from x in direction ξ is locally given as $t\mapsto\exp_x(t\xi)$, so straight lines through 0 in T_xM are mapped to geodesics in M.

(3) Take $\epsilon>0$ such that $\exp_x$ restricts to a diffeomorphism on the ball of radius ϵ around 0 in T_xM. Then for $0<\rho<\epsilon$ the geodesic sphere

$$S_\rho(x):=\exp_x(\{\xi\in T_xM:g_x(\xi,\xi)=\rho^2\})$$

is a smooth submanifold in M which intersects any geodesic through x orthogonally.

(4) Let $\tilde M$ be another Riemannian manifold of the same dimension as M with exponential maps $\widetilde{\exp}$ and let $\Phi:M\to\tilde M$ be an isometry. Then for any point $x\in M$, there is an open neighborhood U of zero in T_xM such that $\Phi\circ\exp_x=\widetilde{\exp}_{\Phi(x)}\circ T_x\Phi$.

If M is connected and $\Phi_1,\Phi_2:M\to\tilde M$ are isometries such that $\Phi_1(x_0)=\Phi_2(x_0)$ and $T_{x_0}\Phi_1=T_{x_0}\Phi_2$ holds for one point $x_0\in M$, then $\Phi_1=\Phi_2$.

Proof. Consider the product $\mathcal{P}M\times\mathbb{R}^n$. Using $T(\mathcal{P}M\times\mathbb{R}^n)=T\mathcal{P}M\times\mathbb{R}^n$ we can define a vector field Ξ on this manifold by $\Xi(\varphi,v):=(\tilde v(\varphi),0)$. Clearly, the flow of this vector field through (φ,v) up to time t is given by $(\mathrm{Fl}_t^{\tilde v}(\varphi),v)$. Now choose an element $\varphi_0\in\mathcal{P}_xM$, i.e. an orthogonal linear isomorphism $\mathbb{R}^n\to T_xM$. By the general theory of flows, it follows that there is a number $\epsilon>0$ and and open neighborhood of $(\varphi_0,0)$ in $\mathcal{P}M\times\mathbb{R}^n$ on which the flow of ξ is defined for all t with $|t|\leq\epsilon$. Intersecting this neighborhood with $\{\varphi_0\}\times\mathbb{R}^n$ we obtain an open neighborhood $\tilde V$ of 0 in $\mathbb{R}^n$ such that for $v\in\tilde V$ the flow $\mathrm{Fl}_t^{\tilde v}(\varphi_0)$ is defined provided that $|t|\leq\epsilon$. Now observe that for $\lambda\in\mathbb{R}$ and $v\in\mathbb{R}^n$, we have $\widetilde{\lambda v}=\lambda\tilde v$ and $\mathrm{Fl}_t^{\lambda\tilde v}=\mathrm{Fl}_{\lambda t}^{\tilde v}$ wherever defined.

From the proof of Lemma 3.7, we know that for $\xi\in T_xM$, the geodesic emanating from x in direction ξ can be written as $p\circ\mathrm{Fl}_t^{\tilde v}(\varphi_0)$, where $v=(\varphi_0)^{-1}(\xi)\in\mathbb{R}^n$. This implies that putting $U := \{\xi \in T_xM : \frac{1}{\epsilon}\varphi_0(\xi) \in \tilde{V}\}$ the map $\exp_x$ is defined and smooth on U, which implies the first two claims in (1). For $\xi \in T_xM$, we of course have $t\xi \in U$ for $|t|$ sufficiently small, and putting $v = (\varphi_0)^{-1}(\xi)$ we get $\exp_x(t\xi) = \mathrm{Fl}_1^{t\tilde{v}}(\varphi_0) = \mathrm{Fl}_t^{\tilde{v}}(\varphi_0)$, which implies (2). This in turn shows that $\frac{d}{dt}|_0 \exp_x(t\xi) = \xi$, which completes the proof of (1).

(3) Any sphere is a smooth submanifold in T_xM and since $S_\rho(x)$ is the image of an appropriate sphere under a diffeomorphism, it is a submanifold, too. To prove that geodesics through x intersect the sphere orthogonally, fix ρ, consider a smooth curve $v(s)$ in T_xM such that $\|v(s)\| = \rho$ for all s, and define $c(s,t) = c_s(t) := \exp_x(tv(s))$. This is a smooth family of curves and by construction each c_s is a geodesic. Moreover, $c_s'(0) = v(s)$ and hence $g(c_s'(0), c_s'(0)) = \rho^2$ for all s. Since each c_s is a geodesic, the function $t \mapsto g(c_s'(t), c_s'(t))$ is constant. This is clear either form the fact that $c_s'(t)$ lifts to a constant vector field on $\mathcal{P}M$ or one observes that

$$\tfrac{d}{dt} g(c_s'(t), c_s'(t)) = c_s'(t) \cdot g(c_s'(t), c_s'(t)) = 2g(\nabla_{c_s'} c_s'(t), c_s'(t)) = 0.$$

Hence we see that $E_0^1(c_s) = \rho^2$ for all s and hence $\frac{\partial}{\partial s}|_0 E_0^1(c_s) = 0$. Computing this derivative via the first variational formula from 3.8 the integral vanishes since c_0 is a geodesic, so only the boundary terms remain. But the variation of the family is given by $r(s,t) = \frac{\partial}{\partial s} c(s,t) = T\exp_x \cdot tv'(s)$, so $r(0,0) = 0$ for all s, while $r(0,1) = T\exp_x \cdot v'(0)$, so we obtain $0 = g_{c_0(1)}((c_0)'(1), T\exp_x \cdot v'(0))$ for any choice of v. Of course any vector tangent to $S_\rho(x)$ can be written as $T\exp_x \cdot v'(0)$, which completes the proof.

(4) For $\xi \in T_xM$ consider the geodesic $c(t) = \exp(t\xi)$ in M. Of course, $(\Phi \circ c)'(0) = T_x\Phi \cdot \xi \in T_{\Phi(x)}\tilde{M}$. But from Proposition 3.7 we know that $\Phi \circ c$ is a geodesic in $\tilde{M}$, so $(\Phi \circ c)(t) = \widetilde{\exp}_{\Phi(x)}(tT_x\Phi \cdot \xi)$, which implies the first claim.

For the second claim, consider $A := \{x \in M : \Phi_1(x) = \Phi_2(x), T_x\Phi_1 = T_x\Phi_2\} \subset M$. This is evidently closed and by assumption it is nonempty, since $x_0 \in A$. But from what we have just proved it follows that for $x \in A$, the maps Φ_1 and Φ_2 coincide on an open neighborhood of x and then also their tangent maps have to coincide in each point of this neighborhood. But this shows that A is open and since M is connected, this implies $A = M$. □

Using this result we can now prove one of the fundamental results on geodesics, namely that two points which are close enough can be joined by a geodesic which is a shortest curve connecting the two points. Basically this is done by writing curves in polar coordinates in the parametrization given by the exponential mapping. This also proves that the distance function d from 2.1 indeed makes M into a metric space.

Corollary 3.9. *Let (M,g) a Riemannian manifold, $x \in M$ a point and $\epsilon > 0$ a number such that $\exp_x$ restricts to a diffeomorphism from $B_\epsilon(0) := \{\xi \in T_xM : g_x(\xi,\xi) < \epsilon^2\}$ onto an open neighborhood U of x in M.*

(1) Let $u : [a,b] \to [0,\epsilon)$ and $v : [a,b] \to T_xM$ be smooth functions such that $g_x(v(t),v(t)) = 1$ for all t and put $c(t) := \exp_x(u(t)v(t))$. Then the arc length of c satisfies $L_a^b(c) \geq |u(b) - u(a)|$ and equality holds if and only if u is monotonous and v is constant.

(2) For $y = \exp_x(\xi) \in U$, the geodesic $t \mapsto \exp_x(t\xi)$ is a length–minimizing curve joining x to y, and up to reparametrizations it is the unique such curve.

Proof. (1) By definition $c'(t) = T\exp_x \cdot (u'(t)v(t) + u(t)v'(t))$ and by part (3) of Proposition 3.9 the two vectors $T\exp_x \cdot (u'(t)v(t))$ and $T\exp_x \cdot (u(t)v'(t))$ are perpendicular. Along the line spanned by $v(t)$, the vector $T\exp_x \cdot v(t)$ is the speed vector of a geodesic, whence we conclude that $g(T\exp_x \cdot(u'(t)v(t)), T\exp_x \cdot(u'(t)v(t))) = |u'(t)|^2$. By Pythagoras, $g(c'(t), c'(t)) \geq |u'(t)|^2$ with equality only for $v'(t) = 0$. Hence we obtain $L_a^b(c) \geq \int_a^b |u'(t)|dt \leq |\int_a^b u'(t)dt| = |u(b) - u(a)|$ as claimed. The first inequality becomes an equality if and only if $v'(t) = 0$ for all t i.e. iff v is constant, while the second one becomes an equality if and only if $u'(t)$ has constant sign and hence u is monotonous.

(2) For $y \in S_\rho(x)$, we of course have $d(x, y) \leq \rho$, since the geodesic joining x to y has length ρ. From (1) we see that any curve joining x to y which stays in $S_\rho \cup \exp_x(B_\rho(0))$ has length at least ρ, since any such curve can be written in the form used in (1). But any curve leaving this set has to have larger length, since the part to the first intersection with $S_\rho(x)$ already has length ρ. This shows that the geodesic is a length minimizing curve. Conversely, a minimizing curve must stay in $S_\rho \cup \exp_x(B_\rho(0))$, and then the equality part of (1) says that a minimizing curve must be of the form $\exp_x(u(t)v)$ for a monotonous function u, and hence a reparametrization of the geodesic $\exp_x(tv)$. □

Remark 3.9. There are lots of further interesting facts about the exponential mapping. On the one hand, one can define exp as a map on an open subset of TM by $\exp|_{T_xM} = \exp_x$. Denoting by $q : TM \to M$ the canonical projection, one can the prove that $(q, \exp)$ defines a diffeomorphism from an open neighborhood of the zero section in TM (i.e. the set of the zero vectors of all tangent spaces) onto an open neighborhood of the diagonal in $M \times M$.

Second there is the concept of completeness, which is crucial in Riemannian geometry. A Riemannian metric is called complete if any geodesic is defined on all of $\mathbb{R}$. If M is compact, then one easily shows that $\mathcal{P}M$ is compact (since $O(n)$ is closed and bounded in $M_n(\mathbb{R})$ and hence compact). Thus any vector field on $\mathcal{P}M$ is complete, whence any Riemannian metric on M is complete.

For non–compact manifolds, the so–called *Hopf–Rinov theorem* characterizes complete Riemannian metrics as those for which the metric space (M, d) is complete in the topological sense, or equivalently by closed bounded sets in (M, d) being compact. On a complete Riemannian manifold, any two points can be joined by a geodesic which is a length minimizing curve.

3.10. Normal coordinates. The second way to construct canonical charts from geodesics is usually referred to as *Riemannian normal coordinates centered at $x \in M$*. In addition to the center $x \in M$, one has to choose an element $\varphi_0 \in \mathcal{P}_xM$, i.e. an orthogonal linear isomorphism $\mathbb{R}^n \to T_xM$. Then $\exp_x \circ\varphi_0$ defines a diffeomorphism from an open neighborhood V of 0 in $\mathbb{R}^n$ to an open neighborhood U of x in M, whose inverse defines a chart $u : U \to V$. The main properties of normal coordinates follow easily from the properties of the exponential map we have verified in 3.9.

Theorem 3.10. *Let (M, g) be a Riemannian manifold, $x \in M$, let (U, u) be a normal coordinate chart centered at x and put $V = u(U) \subset \mathbb{R}^n$. Then we have:*

(1) The straight lines $t \mapsto tv$ (defined for sufficiently small t) are geodesics through $x = u^{-1}(0)$. The geodesic spheres $S_\rho(x)$ around x which are contained in U are mapped diffeomorphically onto spheres around 0 in $\mathbb{R}^n$ of the same radius.

(2) The functions $g_{ij} : V \to \mathbb{R}$ describing the metric have the property that $g_{ij}(0) = \delta_{ij}$ and their partial derivatives vanish in zero.

Proof. (1) follows directly from the results proved in 3.9. By construction $T_0u^{-1} = T_0\exp_x \circ\varphi_0 = \varphi_0$, so this is orthogonal, whence $g_{ij}(0) = \delta_{ij}$. The fact that the straight lines through zero are geodesics implies that $\nabla_{\partial_i}\partial_i$ vanishes along the line te_i locally around 0, so in particular $\nabla_{\partial_i}\partial_i(0) = 0$. Likewise $0 = \nabla_{\partial_i+\partial_j}(\partial_i + \partial_j)(0)$ and expanding this, we conclude that $2\nabla_{\partial_i}\partial_j(0) = 0$ for all i, j. But then we can compute $\partial_i \cdot g_{jk}$ as

$$\partial_i \cdot g(\partial_j, \partial_k) = g(\nabla_{\partial_i}\partial_j, \partial_k) + g(\partial_j, \nabla_{\partial_i}\partial_k),$$

so this also vanishes in 0. □

Viewing a normal coordinate chart as approximating a general Riemannian metric by the flat metric on $\mathbb{R}^n$ it tells us that such an approximation is possible to first order in a point (i.e. the value and the first partial derivatives are the same). This may look weak at the first glance, but we shall see that one cannot do any better than that. This is due to the fact that the second partial derivatives of the functions g_{ij} in 0 are related to the Riemannian curvature at x which is an invariant of the metric g. Of course one also has additional properties like the lines through 0 being orthogonal to the spheres around zero, but they are not sufficient to deduce a higher order approximation in the center.

The Riemann curvature

3.11. The Riemann curvature is the basic invariant of a Riemannian manifold. The motivation for its definition comes from the differential equations we deduced for the homogeneous model in Theorem 2.2. Having at hand the tautological soldering form $\theta = (\theta^i)$ we have used the first of these equations to define the Levi–Civita connection forms, i.e. we have found a unique connection form $\gamma = (\gamma^i_j)$ such that $d\theta^i + \sum_j \gamma^i_j \wedge \theta^j = 0$. Now we can look at the second equation and thus consider $d\gamma^i_j + \sum_k \gamma^i_k \wedge \gamma^k_j$. We proceed similarly as in 2.5 and first show that this expression vanishes if one inserts one fundamental vector field.

We know form 2.5 that for $X = (X^i_j) \in \mathfrak{o}(n)$, the flow of the fundamental vector field ζ_X is given by $\mathrm{Fl}^{\zeta_X}_t = r^{\exp(tX)}$. Since $i_{\zeta_X}\gamma^i_j = X^i_j$ is constant, $di_{\zeta_X}\gamma^i_j = 0$, so we can compute

$$i_{\zeta_X} d\gamma^i_j = \mathcal{L}_{\zeta_X}\gamma^i_j = \tfrac{d}{dt}|_{t=0}(\mathrm{Fl}^{\zeta_X}_t)^*\gamma^i_j.$$

By definition of a connection form, we have $(r^{\exp(tX)})^*\gamma^i_j = b^i_k(t)\gamma^k_\ell a^\ell_j(t)$, where $\exp(tX) = (a^i_j(t))$ and $(b^i_j(t)) = \exp(-tX)$. Differentiating at $t = 0$, we obtain

$$\gamma^i_j(\zeta_X, \eta) = \textstyle\sum_k(-X^i_k\gamma^k_j(\eta) + \gamma^i_k(\eta)X^k_j) = -\sum_k((\gamma^i_k \wedge \gamma^k_j)(\zeta_X, \eta)),$$

which proves our claim. Still as in 2.5, we thus conclude that there are uniquely determined smooth functions $R_{k\ell}{}^i{}_j$ with $R_{\ell k}{}^i{}_j = -R_{k\ell}{}^i{}_j$ such that

$$d\gamma^i_j + \textstyle\sum_k \gamma^i_k \wedge \gamma^k_j = \sum_{k,\ell} R_{k\ell}{}^i{}_j \theta^k \wedge \theta^\ell.$$

Theorem 3.11. *(1) The functions $R_{k\ell}{}^i{}_j$ define a $\binom{1}{3}$–tensor field R on M, which, viewed as a map $\mathfrak{X}(M) \times \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$, is given by*

$$R(\xi_1, \xi_2)(\eta) = \nabla_{\xi_1}\nabla_{\xi_2}\eta - \nabla_{\xi_2}\nabla_{\xi_1}\eta - \nabla_{[\xi_1,\xi_2]}\eta.$$

This tensor field satisfies $R(\xi_2, \xi_1) = -R(\xi_1, \xi_2)$ and that each of the maps $R(\xi_1, \xi_2)$ is skew symmetric, i.e. $g(R(\xi_1, \xi_2)(\eta_1), \eta_2) = -g(R(\xi_1, \xi_2)(\eta_2), \eta_1)$.

(2) If M and $\tilde{M}$ are Riemannian manifolds of the same dimension and $\Phi : \tilde{M} \to M$ is an isometry then the curvatures R and $\tilde{R}$ are related by $\tilde{R} = \Phi^*R$.

(3) For a Riemannian manifold $(M.g)$ the following are equivalent:

(a) *The tensor field R vanishes identically.*
(b) *For vector fields $\xi, \eta \in \mathfrak{X}(M)$ with horizontal lifts $\xi^h, \eta^h \in \mathfrak{X}(\mathcal{P}M)$ we have $[\xi^h, \eta^h] = [\xi, \eta]^h$.*

(c) *The horizontal distribution defined by the Levi–Civita connection (see 3.7) is involutive.*
(d) *Each point $x \in M$ has an open neighborhood which is isometric to an open subset of $\mathbb{R}^n$. ("(M,g) is locally flat")*

PROOF. (1) It is clear from the construction that the function $R_{k\ell}{}^i{}_j$ are most naturally viewed as describing a skew symmetric map from $\mathbb{R}^n \times \mathbb{R}^n$ (corresponding to k and ℓ) to $\mathfrak{so}(n)$. Since $\mathfrak{so}(n)$ consists of linear maps $\mathbb{R}^n \to \mathbb{R}^n$, this means that they are actually trilinear maps $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$. In view of Proposition 3.3, we only have to prove $O(n)$–equivariancy to show that the functions describe a tensor field. So we take $A = (a^i_j) \in O(n)$ and let (b^i_j) be the inverse matrix and we have to compute $R_{k\ell}{}^i{}_j \circ r^A$. Again, this is closely parallel to what we have done in 2.5.

By definition of a connection form, we have $(r^A)^*\gamma^i_j = \sum_{k,\ell} b^i_k a^\ell_j \gamma^k_\ell$. Applying the exterior derivative, we get the same behavior for $d\gamma^i_j$. Likewise, compatibility of pullbacks with the wedge product implies that

$$(r^A)^*(\textstyle\sum_t \gamma^i_t \wedge \gamma^t_j) = \sum_{t,k,\ell,r,s} b^i_k a^\ell_t b^t_r a^s_j (\gamma^k_\ell \wedge \gamma^r_s) = \sum_{k,\ell,t} b^i_k a^\ell_j (\sum_r \gamma^k_t \wedge \gamma^t_\ell).$$

This describes the complete behavior of the left hand side of the defining equation for the functions $R_{ij}{}^k{}_\ell$. This shows that the pullback of the left hand side is given by

$$\textstyle\sum_{k,\ell,r,s} b^i_k a^\ell_j R_{rs}{}^k{}_\ell \theta^r \wedge \theta^s.$$

For the right hand side we get

$$\textstyle\sum_{k,\ell} (R_{ij}{}^k{}_\ell \circ r^A)(r^A)^*\theta^k \wedge (r^A)^*\theta^\ell = \sum_{k,\ell,r,s} (R_{ij}{}^k{}_\ell \circ r^A) b^k_r b^\ell_s \theta^r \wedge \theta^s.$$

Since the $\theta^r \wedge \theta^s$ form a basis for the space of skew symmetric bilinear maps on each tangent space, we conclude that for each r and s we have

$$\textstyle\sum_{k,\ell} (R_{ij}{}^k{}_\ell \circ r^A) b^k_r b^\ell_s = \sum_{k,\ell} b^i_k a^\ell_j R_{rs}{}^k{}_\ell,$$

and bringing the two $b's$ to the right hand side, we get the desired equivariancy property. Thus the functions define a $\binom{1}{3}$–tensor field R as claimed. The first skew symmetry property of R follows from the fact that $R_{ji}{}^k{}_\ell = -R_{ij}{}^k{}_\ell$ while the second follows from the fact that for all fixed i,j the matrix $(R_{ij}{}^k{}_\ell)$ lies in $\mathfrak{so}(n)$.

To prove the interpretation in terms of covariant derivatives, take $\xi_1, \xi_2, \eta \in \mathfrak{X}(M)$ and consider the horizontal lifts ξ^h_1 and $\xi^h_2 \in \mathfrak{X}(\mathcal{P}M)$ and the equivariant function $F : \mathcal{P}M \to \mathbb{R}^n$ corresponding to η. Then the equivariant function representing $\nabla_{\xi_1}\nabla_{\xi_2}\eta - \nabla_{\xi_2}\nabla_{\xi_1}\eta - \nabla_{[\xi_1,\xi_2]}\eta$ is given by

$$(*) \qquad \xi^h_1 \cdot \xi^h_2 \cdot F - \xi^h_2 \cdot \xi^h_1 \cdot F - [\xi_1,\xi_2]^h \cdot F = ([\xi^h_1,\xi^h_2] - [\xi_1,\xi_2]^h) \cdot F.$$

Moreover, since $[\xi^h_1,\xi^h_2]$ is a lift of $[\xi_1,\xi_2]$ we conclude $[\xi^h_1,\xi^h_2] - [\xi_1,\xi_2]^h = \zeta_{\gamma([\xi^h_1,\xi^h_2])}$. Denoting the components of F by F^i, we thus conclude that the ith component of $(*)$ is given by $\sum_j \gamma^i_j([\xi^h_1,\xi^h_2])F^j$.

On the other hand, since horizontal lifts insert trivially into each γ^i_j, we see that $0 = \gamma^i_k \wedge \gamma^k_j(\xi^h_1,\xi^h_2)$. Moreover, inserting the definition of the exterior derivative, we also see that $d\gamma^i_j(\xi^h_1,\xi^h_2) = -\gamma^i_j([\xi^h_1,\xi^h_2])$, and this completes the proof of (1).

(2) From Corollary 3.2 we know that for vector fields $\xi, \eta \in \mathfrak{X}(M)$, we have $\Phi^*(\nabla_\xi \eta) = \tilde{\nabla}_{\Phi^*\xi}\Phi^*\eta$. In view of the formula for R from (1), this immediately implies that

$$\Phi^*(R(\xi_1,\xi_2)(\eta)) = \tilde{R}(\Phi^*\xi_1, \Phi^*\xi_2)(\Phi^*\eta),$$

which implies the claim.

(3) (a) $\Rightarrow$ (b): From the proof of part (1) we see that vanishing of the tensor field R implies that $\gamma^i_j([\xi^h,\eta^h])=0$ for all i and j. Since $[\xi^h,\eta^h]$ is a lift of $[\xi,\eta]$ this implies (b).

(b) $\Rightarrow$ (c): Starting from a local frame ξ_i for TM, the horizontal lifts ξ_i^h form a local frame for the horizontal distribution. By (b), the Lie brackets of these sections are again sections of the horizontal distribution, which by Proposition 1.3 implies involutivity.

(c) $\Rightarrow$ (d): Given x, choose a point $\varphi\in\mathcal{P}_xM$. By condition (c) and the Frobenius theorem, there is a smooth submanifold $N\subset\mathcal{P}M$ containing φ whose tangent spaces coincide with the subspaces defining the horizontal distribution. We can restrict the projection $p:\mathcal{P}M\to M$ to obtain $p|_N:N\to M$, and by construction the tangent maps of $p|_N$ are linear isomorphisms. Hence we may assume that $p|_N$ is a diffeomorphism from N onto an open neighborhood U of x in M, so its inverse defines a smooth section $\sigma:U\to\mathcal{P}M$.

This section defines a local orthonormal frame $\{\xi_1,\dots,\xi_n\}$ for M on U, see Proposition 2.4. By construction, the tangent maps of σ have values in the horizontal distribution, so $\sigma^*\gamma^i_j=0$ for all i and j. Using Propositions 2.7 and 3.5 we conclude that $\nabla\xi_i=0$ for all i. This also shows that $[\xi_i,\xi_j]=\nabla_{\xi_i}\xi_j-\nabla_{\xi_j}\xi_i=0$. This allows us to proceed as in the proof of the Frobenius theorem (see 1.4): Putting

$$\psi(t^1,\dots,t^n):=\mathrm{Fl}^{\xi_1}_{t_1}\circ\dots\circ\mathrm{Fl}^{\xi_n}_{t_n}(x)$$

defines a smooth map from some open neighborhood of 0 in $\mathbb{R}^n$ to an open neighborhood of x contained in U. Using that the flows commute, one verifies that the partial derivatives of ψ are given by the vector fields ξ_i. This also shows that ψ^{-1} is a chart with coordinate vector fields $\partial_i=\xi_i$. Hence the tangent maps of this chart are orthogonal, so it is not only a diffeomorphism but also an isometry onto its image.

(d) $\Rightarrow$ (a) immediately follows from part (2). $\square$

This result has several immediate consequences. First of all, part (2) shows that the Riemann curvature is an invariant of Riemannian manifolds, and part (3) shows that it is a complete obstruction to local isometry with the homogeneous model.

Second, the description in terms of the covariant derivative in part (1) immediately implies that the curvature provides obstructions to the existence of parallel vector fields. Namely, suppose that $\eta\in\mathfrak{X}(M)$ is parallel, i.e. such that $\nabla_\xi\eta=0$ for all $\xi\in\mathfrak{X}(M)$. Then the formula from part (1) shows that $R(\xi_1,\xi_2)(\eta)=0$ for all $\xi_1,\xi_2\in\mathfrak{X}(M)$.

This easily generalizes to tensor fields using the description of the covariant derivative of tensor fields in terms of the covariant derivative of vector fields. For example, for $\alpha\in\Omega^1(M)$, $\nabla_\xi\alpha=0$ for all ξ is equivalent to $\xi\cdot\alpha(\eta)=\alpha(\nabla_\xi\eta)$ for all $\xi,\eta\in\mathfrak{X}(M)$, see Example (2) of 3.4. Given arbitrary vector fields $\xi_1,\xi_2\in\mathfrak{X}(M)$, the definition of the Lie bracket and the formula from part (1) then imply that then

$$0=\xi_1\cdot\xi_2\cdot\alpha(\eta)-\xi_2\cdot\xi_1\cdot\alpha(\eta)-[\xi_1,\xi_2]\cdot\alpha(\eta)=\alpha(R(\xi_1,\xi_2)(\eta)).$$

Hence we see that $\nabla\alpha=0$ implies $\alpha\circ R(\xi_1,\xi_2)=0$ for all $\xi_1,\xi_2\in\mathfrak{X}(M)$.

3.12. Computing the curvature. It is now relatively straightforward to compute the Riemann curvature in a frame. So assume that (M,g) is a Riemannian manifold and $U\subset M$ is an open subset such that there is a local orthonormal frame $\{\xi_1,\dots,\xi_n\}$ for M defined on U. Then from 2.4 we know that this frame gives rise to a smooth section $\sigma:U\to\mathcal{P}M$ and from Proposition 2.7 we know how to determine the connection forms $\omega^i_j=\sigma^*\gamma^i_j$ associated to the frame.

Proposition 3.12. *Let ω^i_j be the connections forms associated to a local orthonormal frame as in Proposition 2.7 and define $\Omega^i_j := d\omega^i_j + \sum_k \omega^i_k \wedge \omega^k_j \in \Omega^2(M)$. Then the Riemann curvature is given by $R(\eta_1,\eta_2)(\xi_i) = \sum_j \Omega^j_i(\eta_1,\eta_2)\xi_j$.*

PROOF. Using the fact that pullbacks commute with the exterior derivative as well as with the wedge product, we conclude that

$$\Omega^i_j = \sigma^*(d\gamma^i_j + \textstyle\sum_k \gamma^i_k \wedge \gamma^k_j) = \sigma^*(\sum_{k,\ell} R_{k\ell}{}^i{}_j \theta^k \wedge \theta^\ell) = \sum_{k,\ell}(R_{k\ell}{}^i{}_j \circ \sigma)\sigma^k \wedge \sigma^\ell.$$

Here $\sigma^1,\dots,\sigma^n$ is the local orthonormal coframe dual to the frame $\xi_1,\dots,\xi_n$ and we have used that $\sigma^*\theta^i = \sigma^i$ for each i. From 3.3 we know that the functions $R_{k\ell}{}^i{}_j \circ \sigma$ just describe the components of the tensor field R with respect to the local orthonormal frame defined by σ. Otherwise put, $R(\xi_a,\xi_b)(\xi_c) = \sum_d R_{ab}{}^d{}_c \xi^d$. Now for any vector field $\eta \in \mathfrak{X}(M)$, we have $\eta|_U = \sum_k \sigma^k(\eta)\xi_k$, and inserting this, the claimed formula follows. □

Example 3.12. Let us compute the curvature in the chart for S^n given by stereographic projection, see Examples 2.7, 3.5, and 3.7. From 2.7, we know that $\omega^i_j = \frac{u^i}{f}du^j - \frac{u^j}{f}du^i$, where $f(u) = \frac{1+\sum(u^j)^2}{2}$ and thus $df = \sum u^k du^k$. Now we immediately compute that

$$d\omega^i_j = -\tfrac{u^i}{f^2}df \wedge du^j + \tfrac{u^j}{f^2}df \wedge du^i + \tfrac{2}{f}du^i \wedge du^j.$$

On the other hand, we get

$$\textstyle\sum_k(\frac{u^i}{f}du^k - \frac{u^k}{f}du^i) \wedge (\frac{u^k}{f}du^j - \frac{u^j}{f}du^k) = \frac{u^i}{f^2}df \wedge du^j - \frac{u^j}{f^2}df \wedge du^i - \frac{\sum(u^k)^2}{f^2}du^i \wedge du^j.$$

Adding these two expressions up, we conclude that $\Omega^i_j = \frac{1}{f^2}du^i \wedge du^j$. Since $\xi_i = f\partial_i$ we see that $\Omega^i_j(\xi_a,\xi_b)$ equals 1 for $a=i$, $b=j$ and -1 for $a=j$ and $b=i$ and 0 for all other values of a and b. This shows that

$$R(\xi_a,\xi_b)(\xi_c) = \textstyle\sum_i \Omega^i_c(\xi_a,\xi_b) = \delta_{ca}\xi_b - \delta_{cb}\xi_a.$$

Another simple way to compute the curvature is using the formula for the covariant derivative in local coordinates. By definition, for the coordinate vector fields ∂_i, we have $\nabla_{\partial_j}\partial_k = \sum_\ell \Gamma^\ell_{jk}\partial_\ell$, where the Γ^i_{jk} are the Christoffel symbols, so these are just smooth functions. Applying ∇_{∂_i} to this, we obtain

$$\textstyle\sum_\ell \frac{\partial\Gamma^\ell_{jk}}{\partial u^i}\partial_\ell + \sum_\ell \Gamma^\ell_{jk}\nabla_{\partial_i}\partial_\ell.$$

Expanding the last term and alternating in i and j, we conclude that $R(\partial_i,\partial_j)(\partial_k)$ is given by

$$\textstyle\sum_\ell \left(\frac{\partial\Gamma^\ell_{jk}}{\partial u^i} - \frac{\partial\Gamma^\ell_{ik}}{\partial u^j} + \sum_r(\Gamma^r_{jk}\Gamma^\ell_{ri} - \Gamma^r_{ik}\Gamma^\ell_{rj})\right)\partial_\ell.$$

Using the formula for the Christoffel symbols Γ^i_{jk} from Proposition 3.8, on can convert this into a formula in terms of the first and second partial derivatives of the functions g_{ij} which describe the metric in the given coordinates. This becomes particularly simple in the origin of a normal coordinate system. Recall from 3.10 that in normal coordinates, the first partial derivatives of the functions g_{ij} vanish. In view of Proposition 3.8 this implies that the Christoffel symbols vanish at the origin. There are further simplifications because of special symmetries for the second partials of the functions g_{ij} in normal coordinates. In the end one finds that

$$R(\partial_i,\partial_j)(\partial_k)(0) = \textstyle\sum_\ell \left(\frac{\partial^2 g_{j\ell}}{\partial u^i \partial u^k} - \frac{\partial^2 g_{i\ell}}{\partial u^j \partial u^k}\right)\partial_\ell(0).$$

3.13. Extracting parts of the curvature. While it is a tensor field, the Riemann curvature is not really easy to handle in general. Basically, we can view R as a 4–linear operator on tangent spaces, i.e. look at $g(R(\xi_1,\xi_2)\eta_1,\eta_2)$ (in a point). Then we know that this expression is skew symmetric in the $\xi's$ and in the η's. There is a further fundamental symmetry of the curvature tensor called the *Bianchi identity.* We can easily prove this as follows:

Take the defining equation for the Levi–Civita connection forms, $0=d\theta^i+\sum_j\gamma^i_j\wedge\theta^j$, and apply the exterior derivative to get

$$0=\textstyle\sum_j(d\gamma^i_j\wedge\theta^j-\gamma^i_j\wedge d\theta^j)=\sum_j(d\gamma^i_j+\sum_k\gamma^i_k\wedge\gamma^k_j)\wedge\theta^j=\sum_{j,k,\ell}R_{k\ell}{}^i{}_j\theta^k\wedge\theta^\ell\wedge\theta^j.$$

Since the values of the forms $\theta^k\wedge\theta^\ell\wedge\theta^j$ in each point form a basis for the space of trilinear alternating forms on T_xM, this implies that alternating $R_{k\ell}{}^i{}_j$ over j, k, and ℓ gives zero. Since the expression is already alternating in k and ℓ, this can be equivalently phrased as

$$0=R(\xi,\eta)(\zeta)+R(\zeta,\xi)(\eta)+R(\eta,\zeta)(\xi)$$

for all $\xi,\eta,\zeta\in\mathfrak{X}(M)$. From the identities verified so far, one can further deduce that $g(R(\xi_1,\xi_2)(\eta_1),\eta_2)=g(R(\eta_1,\eta_2)(\xi_1),\xi_2)$. These identities together imply that there is no way to get from the curvature tensor to a simpler object by forming symmetric or alternating parts in any sense. (More precisely, any kind of symmetrizations or alternations either leads to zero or to a non–zero multiple of the curvature tensor itself.)

There are non–trivial traces of the curvature tensor one can form however. They can be most easily defined locally using a local orthonormal frame $\{\xi_1,\dots,\xi_n\}$. It is then easy to verify directly that the quantities defined do not depend on the choice of this frame. One defines the *Ricci curvature* $\mathrm{Ric}\in\mathcal{T}^0_2(M)$ and the *scalar curvature* $R\in C^\infty(M,\mathbb{R})$ by

$$\mathrm{Ric}(\eta_1,\eta_2):=\textstyle\sum_i g(R(\eta_1,\xi_i)(\eta_2),\xi_i)$$

and then $R:=\sum_i\mathrm{Ric}(\xi_i,\xi_i)$. From this definition it is clear that Ric is symmetric, i.e. $\mathrm{Ric}(\xi,\eta)=\mathrm{Ric}(\eta,\xi)$ for all $\xi,\eta\in\mathfrak{X}(M)$. In each point the Ricci curvature thus defines a symmetric bilinear form on T_xM, so it makes sense to study eigenvalues and ask whether this bilinear form is positive or negative definite, and so on.

An important special case of a curvature condition is the following. A Riemann metric is called an *Einstein metric* if its Ricci curvature is proportional to the metric in each point, which then implies that $\mathrm{Ric}(\xi,\eta)=\frac{1}{n}Rg(\xi,\eta)$. It turns out that in this case the scalar curvature R must be constant, so the Ricci tensor is actually a constant multiple of the metric.

Another important method of extracting parts of the curvature is inspired by the Gauss curvature of surfaces in $\mathbb{R}^3$, see [**DG1**, 3.7]. Take a Riemannian manifold (M,g), a point $x\in M$, and a plane (i.e. a two–dimensional linear subspace) $E\subset T_xM$. Then choose an orthonormal basis $\{\xi,\eta\}$ of E and consider $g_x(R_x(\xi,\eta)(\xi),\eta))$. One easily verifies directly that this number independent of the choice of the orthonormal basis, so it defines a number $K_x(E)$, called the *sectional curvature* of E. The simples possible Riemannian manifolds are the ones with *constant sectional curvature*, where this is the same for all points and all planes. More generally, one can look at Riemann manifolds for which all sectional curvatures have the same sign (e.g. "positive sectional curvature") or allow some bounds on the values of the sectional curvature.

Example 3.13. Let us compute the curvatures of the sphere using the chart given by stereographic projection. From 3.12 we know that for the orthonormal frame $\{\xi_1,\dots,\xi_n\}$ we have always considered in this example, we have $R(\eta_1,\eta_2)(\xi_i)=\sum_j\Omega^j_i(\eta_1,\eta_2)\xi_j$, where $\Omega^i_j = \frac{1}{f^2} du^i \wedge du^j = \sigma^i \wedge \sigma^j$. Here the σ^i are the elements of the coframe dual to the frame $\{\xi_i\}$ Now observe that for each vector field ζ, we have $\zeta = \sum_i \sigma^i(\zeta)\xi_i$. This implies that

$$(*) \qquad R(\eta_1,\eta_2)(\zeta) = \textstyle\sum_{i,j} \sigma^i(\zeta)\Omega^j_i(\eta_1,\eta_2)\xi_j$$

and inserting this into $g(\ ,\xi_k)$ we get

$$\sum_i \sigma^i(\zeta)\Omega^k_i(\eta_1,\eta_2) = \sum_i \sigma^i(\zeta)(\sigma^k(\eta_1)\sigma^i(\eta_2) - \sigma^i(\eta_1)\sigma^k(\eta_2)).$$

Putting $\eta_2 = \xi_k$, we just get $\sum_{i\neq k} \sigma^i(\zeta)\sigma^i(\eta_1)$, and summing over all k, we conclude that $\mathrm{Ric}(\eta_1,\zeta) = (n-1)\sum_i \sigma^i(\zeta)\sigma^i(\eta_1) = (n-1)g(\eta_1,\zeta)$ by definition of an orthonormal coframe. Thus the sphere is an Einstein manifold with scalar curvature $n(n-1) > 0$.

Likewise, formula $(*)$ shows that

$$R(\eta_1,\eta_2)(\eta_1) = \textstyle\sum_{i,j} \sigma^i(\eta_1)\Omega^j_i(\eta_1,\eta_2)\xi_j,$$

and this inner product of this with η_2 is given by

$$\textstyle\sum_{i,j} \sigma^i(\eta_1)\sigma^j(\eta_2)\Omega^j_i(\eta_1,\eta_2) = \sum_{i,j} \sigma^i(\eta_1)\sigma^j(\eta_2)(\sigma^i(\eta_1)\sigma^j(\eta_2) - \sigma^j(\eta_1)\sigma^i(\eta_2))$$

If η_1 and η_2 are orthonormal, then $\sum_i(\sigma^i(\eta_1))^2 = 1$ and likewise for η_2, while for mixed terms $\sum_i \sigma^i(\eta_1)\sigma^i(\eta_2) = 0$. This easily implies that the sphere has constant sectional curvature one.

Bibliography

[DG1] A. Čap, "Differentialgeometrie 1", lecture notes (in German) version spring term 2011, available at http://www.mat.univie.ac.at/∼cap/lectnotes.html

[IL] T. Ivey, J.M. Landsberg, "Cartan for Beginners: Differential Geometry via Moving Frames and Exterior Differential Systems", Graduate Studies in Mathematics 61, Amer. Math. Soc., 2003.

[Mi] P. Michor, "Topics in Differential Geometry", Graduate Studies in Mathematics 93, Amer. Math. Soc., 2008.