Jacobi fields, bundles and connections

Olli Väisänen
June 18, 2020

Secondary master's thesis in mathematics

UNIVERSITY OF JYVÄSKYLÄ
Department of Mathematics and Statistics

Supervisors: Joonas Ilmavirta, Ville Kivioja

Tiivistelmä

- *Korkeakoulu ja laitos:* Jyväskylän yliopisto, matematiikan ja tilastotieteen laitos.
- *Työn tekijä:* Olli Väisänen.
- *Työn nimi:* Jacobin kentät, kimput ja konnektiot.
- *Työn laatu opinnäytteenä:* Sivuainetutkielma.
- *Sivumäärä ja liitteet:* 31. Ei liitteitä.
- *Tieteenala:* Matematiikka, geometria.
- *Valmistumiskuukausi ja -vuosi:* Kesäkuu 2020.

Tämä teksti käsittelee Jacobin kenttiä ja niiden määrittelemiseen tarvittavia rakenteita, erityisesti vektorikimppuja ja konnektioita. Lopputuloksena osoitetaan yksi yhteen-vastaavuus Jacobin kenttien ja geodeesiperheiden variaatiokenttien välillä.

Abstract This text discusses Jacobi fields and the structures needed in their definition, vector bundles and connections in particular. The one-to-one -correspondence between Jacobi fields and the variation fields of families of geodesics is proven as a final result.

Contents

1 Introduction

A manifold is roughly speaking the most general object which can still be described by coordinates, though in general this is only possible locally. More structure can then be added to the manifold in layers. Smooth manifolds are defined by requiring that the familiar concept of differentiation in $\mathbb{R}^n$ is applicable locally on the manifold. Smooth manifolds are still very abstract spaces and more familiar concepts of geometry are found by defining angles and lengths on the manifold using a metric. This results in the structure of a Riemannian manifold.

Historically one of the more interesting objects defined on a Riemannian manifolds are geodesics, curves which extremize the length between two points. In a flat geometry, these are of course just straight lines, but in even the relatively simple geometry of a sphere has a much richer structure of geodesics. Geodesics have both individual and collective properties, and an important tool in studying the latter are the Jacobi fields which describe how families of geodesics vary under the change of a variable.

The purpose of this text is to review the structure of a Riemannian geometry. Since this branch of mathematics is too large to be decently covered by a short text, I will do this by focusing on Jacobi fields and defining the relevant structures along the way, with the final aim being to prove the connection between the so called Jacobi equation and the variation fields of geodesics. In the process, I will discuss the vector bundle structure of tensor fields and use it to define the concept of a covariant derivative on a manifold. Instead of proper Riemannian manifolds, I will mostly deal with pseudo-Riemannian manifolds which pop up in many applications such as general relativity.

This text follows primarily the presentation in the textbooks by John Lee [Lee13, Lee18]. I will assume that the reader has prior knowledge on point set topology and vector analysis in $\mathbb{R}^n$. As the discussion on the basic definitions of manifolds and tensors is rather brief, it is also useful to be somewhat familiar with these concepts.

2 Basic concepts

Before I can begin discussing the main subject of curvature and Jacobi fields, I need to introduce the basic concepts and notation of differential geometry. I begin this section by defining the concept of a manifold and, from there on, discuss tangent vectors and tensors. After this I move on to pseudo-Riemannian geometry by defining the metric. Since most of these developments are standard and can be found in any textbook on the subject, I will state many results without proof.

2.1 Manifolds

The basic object studied in differential geometry is a manifold. At its most rudimentary level a manifold is any set which can be described by coordinates locally. More structure can be found by restricting the coordinate systems, and in this text I will exclusively work with smooth manifolds, for which all coordinate transformations are infinitely many times differentiable.

I will assume that this is at least somewhat familiar to the reader and will quickly give the basic definitions and results related to smooth manifolds. I will not go into much detail regarding the topology required of manifolds and the reader can consult for example [Lee13, ch. 1] for more details.

Definition 2.1. Let M be a topological space. A **coordinate chart** is a pair (U, ϕ), where U is an open subset of M and $\phi : U \to \phi(U) \subset \mathbb{R}^n$ is a homeomorphism. Two coordinate charts (U, ϕ) and (V, ψ) are compatible if $U \cap V = \emptyset$ or the **coordinate transform** $\phi \circ \psi^{-1} : \psi(U \cap V) \to \phi(U \cap V)$ is a diffeomorphism with respect to the differentiable structure of $\mathbb{R}^n$.

An atlas is a collection $\mathcal{A}$ of mutually compatible coordinate charts $\mathcal{A} = \{((U_i, \phi_i) : i \in I\}$ where I is some indexing set and the sets U_i cover M, that is $M = \bigcup_{i \in I} U_i$. The atlas is called maximal if there are no compatible charts not included in the atlas.

A smooth manifold is a second countable Hausdorff space equipped with a maximal atlas $\mathcal{A}$.

The integer n is called the dimension of the coordinate chart and it can be proven that all overlapping compatible coordinate charts have the same dimension [Lee13, theorem 1.2]. The dimension of a connected smooth manifold is therefore constant and this constant is called the **dimension of the manifold**.

In this text, unless otherwise stated, M will always refer to a smooth manifold, p to a point on it and n to its dimension. Since all manifolds discussed are smooth, I will also simply refer to smooth manifolds as manifolds. In this text a coordinate chart on M always refers to an element of the atlas $\mathcal{A}$. When it is necessary to refer to individual coordinate functions of the chart (U, ϕ) I will denote them $x^i : U \to \mathbb{R}$ for all $i = 1, ..., n$ unless otherwise specified.

2.2 Tangent spaces and tensors

There are many objects defined on a manifold, mainly functions, tangent vectors, vector fields, tensors and tensor fields. As with manifolds, I will only discuss the smooth variants of these concepts and will often drop the smooth specifier.

The first ones to be defined are smooth functions, since they are used to define the rest of the objects. A **smooth function** on M is a map $f : M \to \mathbb{R}$ such that for any coordinate chart (U, ϕ) the composite map $f \circ \phi^{-1}$ is infinitely differentiable. The set of all smooth functions in M is denoted $C^\infty(M)$. A smooth function on an open subset U is defined analogously.

I will also need to discuss maps between manifolds. Let M and M' be manifolds. A map $f : M \to M'$ is smooth if given any two coordinate charts (U, ϕ), (U', ϕ') on M and M' respectively the composite map $\phi' \circ f \circ \phi^{-1}$ is infinitely differentiable in its domain. Smooth maps between open subsets of manifolds are defined in the same way.

A curve γ on the manifold (or its subset) is defined as usual as the map $\gamma : I \to M$ where $I \subset \mathbb{R}$ is an interval. I will assume that all curves are smooth, that is the composition $\phi \circ \gamma$ is infinitely differentiable in its domain for any coordinate chart (U, ϕ).

Next I will define tangent vectors as an equivalence class of curves on the manifold.

Definition 2.2. Let M be a manifold and $p \in M$. Two curves $\sigma_1, \sigma_2 : (-\epsilon, \epsilon) \to M$, where $\epsilon > 0$ is some real number and with $\sigma_1(0) = p = \sigma_2(0)$, are tangent at p, if in some coordinate chart around p

$$\frac{d}{dt} x^i(\sigma_1(t)) \bigg|_{t=0} = \frac{d}{dt} x^i(\sigma_2(t)) \bigg|_{t=0}$$

for all i. The equivalence class of curves

$$[\sigma] = \{\rho : (-\epsilon, \epsilon) \to M \mid \rho \text{ tangent with } \sigma \text{ at } p\}$$

is called a **tangent vector** at p. The set of all vectors at p is the **tangent space** TM_p.

Tangent vectors are well-defined as the concept of tangency of curves is independent of the chosen coordinate system. This can be proven with a straightforward calculation using the chain rule and the fact that as a coordinate transformation is a diffeomorphism, its Jacobian matrix is invertible at all points.

The concept of tangency of two curves with different domains in $\mathbb{R}$ can be defined analogously. This will not be relevant in this text and to avoid clutter I will settle for the definition presented above.

Since there is no way to add points directly in M, the addition and scalar multiplication of tangent vectors are defined in $\mathbb{R}^n$ using the coordinate charts and the results are then mapped back to M.

Proposition 2.3. *[Ish99, theorem 2.1] The tangent space TM_p is a n-dimensional vector space when addition and scalar multiplication are defined in each coordinate patch (U, ψ) as $[\sigma_1] + [\sigma_2] = [\psi^{-1} \circ (\psi \circ \sigma_1 + \psi \circ \sigma_2]$ and $\lambda[\sigma_1] = [\psi^{-1} \circ (\lambda\psi \circ \sigma_1)]$ for all $\sigma_1, \sigma_2 \in TM_p$ and $\lambda \in \mathbb{R}$.*

Note that in the definition of the vector operations, the domain $I \subset \mathbb{R}$ of the sum curve $\psi \circ \sigma_1 + \psi \circ \sigma_2$ in $\mathbb{R}^n$ must be restricted so that $(\psi \circ \sigma_1 + \psi \circ \sigma_2)(t) \in \psi(U)$ for all $t \in I$. A similar restriction has to be made for the definition of the scalar multiplication.

While the curve formulation of tangent vectors is geometrically intuitive, it is often cumbersome to work with. For the rest of this text, I will usually use the more algebraic approach of identifying vectors at p with the corresponding directional derivative operators. More formally, I will identify the space of **derivations at** p, linear maps $d : C^\infty(M) \to \mathbb{R}$ which satisfy the Leibniz rule

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

for all $f, g \in C^\infty(M)$, with the tangent space TM_p.

First I will define the action of a vector $v \in TM_p$ at some point p on a function f as the directional derivative operation along a representative curve. Denote

$$v(f) := \frac{d}{dt} f(\sigma(t))\bigg|_{t=0},$$

where $v = [\sigma]$. It is straightforward to verify that this operation is independent of the chosen representative curve and defines a derivation at the point p. I will denote this derivation with d_v.

As with the more familiar vectors of $\mathbb{R}^n$, it is often most convenient to work with a component representation of tangent vectors with respect to some basis. Since tangent vectors are also directional derivative operators, a natural basis to work with can be found by working in a coordinate patch and using the chain rule.

To this end, I will define some notation. Let f be a smooth function and (U, ϕ) a coordinate chart. Note that the function $f \circ \phi^{-1} : \phi(U) \subset \mathbb{R}^n \to \mathbb{R}$ can be differentiated as usual with respect to the differentiable structure in $\mathbb{R}^n$. As such, I will define the **partial derivatives of** f with respect to the coordinate chart (U, ϕ) as

$$\partial_i f = \frac{\partial f}{\partial x^i} := \frac{\partial}{\partial x^i} f \circ \phi^{-1}.$$

I will call the map $f \circ \phi^{-1}$ a coordinate representation of f.

For future use I will also define the Jacobian matrix or the differential of a smooth map between manifolds $f : M \to M'$ with respect to the coordinate charts (U, ϕ) and (U', ϕ') on M and M' respectively as the Jacobian matrix of the map $\phi' \circ f \circ \phi^{-1}$. The rank of the map at a point is the rank of its differential at that point. The differential and its rank are discussed fully in [Lee13, chapters 3,4].

I can now express the action of a vector on a function using the chain rule on a coordinate representation.

Theorem 2.4. *[Ish99, p. 82] Let (ϕ, U) be a coordinate chart on manifold M. The action of a vector v on function f at a point $p \in U$ can be expressed as*

$$v(f) = \sum_i v^i \partial_i f$$

for some unique coefficients v^i, $i = 1, \ldots, n$. The coefficients are unique in each coordinate system and are called the ***vector components of** v **in the coordinate basis**.*

Since sums like $v(f) = \sum_i f^i \partial_i f$ crop up everywhere in this text, I will later always use the Einstein summation convention. Under this convention I will drop the explicit summation sign and assume all repeated indices are implicitly summed over. For example, the above could be written as $v(f) = \sum_i f^i \partial_i f = v^i \partial_i f$.

This result can be used to prove the one-to-one correspondence between tangent vectors and derivations at the point p.

Theorem 2.5. *[Ish99, theorem 2.2] Let D_p be the space of all derivations at $p \in M$. The map $TM_p \to D_p$, $v \mapsto d_v$, is a linear isomorphism.*

In the following I will make no distinction between the tangent vector and the corresponding directional derivative operation.

It will also be necessary to discuss vectors at many different points of M, which gives rise to the concept of a vector field.

Definition 2.6. The map $v : M \to \bigcup_{p \in M} TM_p$, $v(p) = v_p \in TM_p$ for all p is a **smooth vector field** if it is smooth in the sense that $v(f) : p \mapsto v_p(f)$ is a smooth function for every $f \in C^\infty(M)$. The set of all vector fields on M will be denoted $VF(M)$.

Like tangent vectors are identified with derivations, I will identify vector fields with the corresponding fields of derivations, and its action on a smooth function f will be denoted $v(f) \in C^\infty(M)$. The definition of a vector field will be elaborated further in section 3 where vector fields will be identified with sections of the tangent bundle.

It can be verified that in general the map $C^\infty(M) \to C^\infty(M)$, $f \mapsto Y(X(f))$, where X and Y are vector fields, does not satisfy the Leibnitz rule. As such it does not define a derivation or a vector field.

However, it is possible to define an operation between X and Y which produces a third vector field, namely the Lie bracket. Let X, Y be vector fields on M. The **Lie bracket** of the vector fields $[X, Y]$ is defined as

$$[X, Y]_p f = X_p(Y(f)) - Y_p(X(f))$$

for all points $p \in M$ and smooth functions f. The proof that a Lie bracket is a vector field can be found in [Lee13, lemma 8.25].

Vector fields are not the only objects defined on M. I will also need the concept of covectors or duals to these vectors. The following definitions and theorems, as well as their proofs, match almost exactly the ones on vectors and vector fields.

Definition 2.7. Let $p \in M$. The set $TM_p^* = \{\omega : TM_p \to \mathbb{R} \mid \omega \text{ linear }\}$ is called the **cotangent space** at p and its elements are called **covectors**. An assignment $p \mapsto \omega_p \in TM_p^*$ is called a **1-form** on M if it is smooth in the sense that that $\omega(v)$ is a smooth function for all vector fields v.

In terms of linear algebra, the cotangent space is just the dual space of the tangent space. As covectors are linear transformations, their sums and products are natural to define as $(\omega_1+\omega_2)(v) = \omega_1(v)+\omega_2(v)$ and $(\lambda\omega_1)(v) = \lambda\omega_1(v)$ for any $\omega_1, \omega_2 \in TM_p^*$, $v \in TM_p$ and $\lambda \in \mathbb{R}$. With these definitions TM_p^* is a n-dimensional vector space [Lee13, chapter 11].

A more general class of objects, tensors, can be constructed using tangent vectors and their duals.

Definition 2.8. The multilinear map $T_p : \left(\bigtimes_{i=1}^{m} TM_p\right) \times \left(\bigtimes_{i=1}^{r} TM_p^*\right) \to \mathbb{R}$ is called a **(r,m)-tensor** at p. Denote the space of all such tensors at p with $TM_p^{(r,m)}$. The smooth assignment $p \mapsto T_p \in TM_p^{(r,m)}$ is called a **tensor field** on M.

In this definition the smoothness is in the sense that the map $p \mapsto T_p((v_1)_p, \ldots, \omega_p^1, \ldots)$ is a smooth function for any vector fields v_i and 1-forms ω^j, where $i = 1, \ldots, m$ and $j = 1, \ldots, r$. I will in the following always assume all tensor fields to be smooth.

From this definition it is immediately clear that a covector is a $(0,1)$-tensor. The action of a vector v on covector ω can be defined by $v(\omega) := \omega(v)$, and as such a tangent vector can be regarded as a $(1,0)$-tensor. $(0,0)$-tensors are be defined to be smooth functions on M.

Similarly to the tangent and cotangent bundles, the union of all (r,m)-tensor spaces is the (r,m)-**tensor bundle** $TM^{(r,m)}$ which will be discussed more in section 3.

As with individual tangent vectors it is useful to define a component representation of vector fields. In this case the components are functions on M instead of numbers. Even though this is rather simple to do starting from the coordinate bases defined earlier, the coordinate basis is only one of the possible bases on TM_p. As such I can be more general and simply choose a basis for each tangent space.

However, the choice is not entirely unrestricted. I want the component functions to be smooth and as such it would be desirable that the basis vectors vary smoothly between nearby points on M. In general, this is not possible to achieve over the entire manifold, and therefore I define the smooth set of bases over some open set $U \subset M$.

Definition 2.9. Let $e = \{e_i\}_{i=1}^n$ be a set of vector fields on an open set $U \subset M$, such that at each point p the vectors $\{(e_i)_p\}_{i=1}^n$ form a basis in TM_p. This set is called a **frame** (on TM) in U.

Using these definitions, I can show that the earlier defined coordinate bases form a frame on the coordinate patch.

Proposition 2.10. *Let (ϕ, U) be a coordinate patch. Define vectors $\{(\partial/\partial \boldsymbol{x^i})_p\}$ at $p \in U$ by*

$$(\partial/\partial \boldsymbol{x^i})_p f = \partial_i f.$$

for all functions f. These vectors form a basis in TM_p. The set of vector fields $\{\partial/\partial \boldsymbol{x^i}\}_{i=1}^n$ forms a frame on U.

Proof. Since f is by definition a smooth function, all its partial derivatives are as well, and thus $\partial/\partial \boldsymbol{x^i}$ are smooth vector fields on M. By theorem 2.4 the partial derivatives form a basis on each tangent space, which completes the proof. □

Any vector field v can be expanded in terms of a frame $e = \{e_i\}_{i=1}^n$ on $U \subset M$ as $v^i e_i$ for some smooth functions $v^i : U \to \mathbb{R}$. These functions are called the **components of v with respect to the frame e**.

I will next outline a proof for the above statement. Let $p \in U$. Since the vectors $\{(e_i)_p\}_{i=1}^n$ form a basis in TM_p, the tangent vector v_p can be expressed as $v_p = v^i(p)(e_i)_p$ for some maps $v^i : M \to \mathbb{R}$. It is then left to show that the maps v^i are smooth.

This is easiest to do starting with the coordinate frame. Suppose the component map v^i for some i is not smooth. The action of v on the coordinate function x^i is $v(x^i) = v^j \partial_j x^i = v^i$ which was assumed not to be smooth. This is a contradiction as the coordinate functions are smooth by definition and v is a vector field.

A very similar idea works on a general frame as well. Consider a point $p \in U$ and let $e_i = e_i^j \partial_j$ be the frame fields, where e_i^j are the components of the frame fields in terms of the coordinate frame. Then the requirement $e_i(f) = 1$ and $e_j(f) = 0$ for all $i \neq j$ can be represented in some coordinate chart containing p as a system of linear ordinary differential equations (ODEs). By the existence and uniqueness theorems for ODEs, there is a smooth solution f to the system defined in some neighborhood of p [Lee18, theorem 4.32]. Then $v(f) = v^j e_j(f) = v^i$ by construction and v^i must be a smooth map.

As the cotangent spaces are vector spaces as well, it is useful to define bases and frames for them as well in an analogous way. The most convenient frames are compatible with the vector frame used, and in this text I will use these special frames exclusively.

Proposition 2.11. *[Lee13, p. 278] Let* $\{(e_i)_p\}_{i=1}^n$ *form a basis in* TM_p*. There is a unique basis of covectors in* TM_p^**, denoted* $\{(\sigma^k)_p\}_{k=1}^n$*, that fulfills* $(\sigma^k)_p((e_i)_p) = \delta_i^k$*, where* δ_i^k *is the Kronecker delta. This basis is called the dual basis. If* $\{e_i\}_{i=1}^n$ *form a frame on* TM *in* U*, the dual bases form a frame for* TM^* *in* U *called the* ***dual frame****.*

The component representation for 1-forms can be defined analogously to the tangent vector case.

Giving a component expression is often the most convenient way to define a vector field or a 1-form. In principle this can be done separately for each frame. However, in order for the vector field or 1-form to be well defined the different component expressions have to agree in the overlaps of the frame domains. I will now discuss this in the case of vector fields.

More rigorously, let $e = \{e_1, ..., e_n\}$ and $e'\{e'_1, ..., e'_n\}$ be frames with domains U and U' with $U \cap U' \neq \emptyset$. At each point $p \in U \cap U'$ the bases are related with a basis change matrix $A(p)$. I will denote the elements of the matrix $A(p)$ with $A_j^i(p)$. Then $(e_j)_p = A_j^i(p)(e'_i)_p$ for all p. I will, here and later, use the shorthand notation for the previous equation $e_j = A_j^i e'_i$ when specifying the point in question is not essential. All components of A must be smooth functions, as A_j^i is the j:th component map of e_j in terms of the frame e'. Thus A is a smooth map $A : U \cap U' \to GL(n)$.

Let v be a vector field defined in $U \cap U'$ and denote its coordinates with respect to the frames with v^i and v'^i. Then $v'^i e'_i = v = v^j e_j = v^j A_j^i e'_i$, which immediately implies that $v'^i = A_j^i v^j$. Since A is a basis change at each point, it can also be inverted to give $v^i = (A^{-1})_j^i v'^j$.

The component representation for tensors follow immediately from linearity and the component representations for vectors and covectors.

Proposition 2.12. *Let* $T \in TM_p^{(r,m)}$*,* $\{v_i\}_{i=1}^m \subset TM_p$ *and* $\{\omega^k\}_{k=1}^r \subset TM_p^*$*. Let* e *be a frame and* σ *its dual in an open* $U \subset M$*. The tensor* T *acts on these vectors and covectors as*

$$T(v_1, ..., v_m, \omega^1, ..., \omega^r) = T_{i_1 ... i_m}{}^{k_1 ... k_r} v_1^{i_1} ... v_m^{i_m} \omega_{k_1}^1 ... \omega_{k_r}^r,$$

where the ***tensor components*** *with respect to frame e are defined as*

$$T_{i_1 ... i_m}{}^{k_1 ... k_r} = T(e_{i_1}, ..., e_{i_m}, \sigma^{k_1}, ..., \sigma^{k_r}).$$

The upped indices are called covariant and the lower contravariant.

Proof. The vectors and 1-forms can be represented as $v_i = v_i^a e_a$, $\omega^k = \omega_b^k \sigma^b$ for all i, k. Expanding every argument of T gives

$$T(v_1, ..., v_r, \omega^1, ..., \omega^m) = T(v_1^{i_1} e_{i_1}, ..., \omega_{k_m}^m \sigma^{k_m}).$$

As T is linear in all its arguments the above expression simplifies into the sum

$$T(v_1,...,v_r,\omega^1,...,\omega^m) = v_1^{i_1}...v_r^{i_r}\omega_1^{k_1}...\omega_{k_m}^m T(e_{i_1},...,\sigma^{k_m}).$$

This proves the proposition. □

The transformation properties of 1-form and tensor components can be derived the same way as for vector fields. In particular if e and e' are frames, $\sigma = \{\sigma^1,...,\sigma^n\}$ is a frame dual to e and $\sigma' = \{\sigma'^1,...,\sigma'^n\}$ is defined so that $\sigma^j = (A^{-1})^j_i\sigma'^i$, it is useful to note that

$$\sigma'^i(e'_j) = A^i_k\sigma^k((A^{-1})^l_j e_l) = A^i_k(A^{-1})^l_j\delta^k_l = A^i_k(A^{-1})^k_j = \delta^i_j.$$

Thus, σ' is the dual frame to e'. This implies that dual frames transform under the inverse matrix compared with the corresponding frames.

I will also present some definitions which will initially serve only as a shorthand notation. The disjoint union of all tangent spaces is called the **tangent bundle** and it is denoted TM. Similarly the disjoint unions of all cotangent spaces and (r,m)-tensor spaces are the **cotangent bundle** TM^* and the (r,m)-tensor bundle $TM^{(r,m)}$ respectively. These spaces have a vector bundle structure, which will be discussed in section 3, in addition to which they are manifolds.

Theorem 2.13. *Tangent, cotangent and tensor bundles are manifolds.*

Proof. I will prove this only for the tangent bundle, since the proofs for the cotangent and tensor bundles are nearly identical. It is sufficient to construct a topology and a maximal atlas for the tangent bundle.

Let $\{(U_\alpha,\phi_\alpha)\}_{\alpha\in I}$ for some indexing set I be a set of coordinate patches which cover the manifold M. Since the tangent bundle is a disjoint union of the tangent spaces, for every $v \in TM$ there is an unique $x \in M$ such that $v \in TM_x$. Denote the point associated with a vector v with $\mathrm{pr}_1(v)$.

Define $V_\alpha = \bigcup_{x\in U_\alpha} TM_x$ and the map $\psi_\alpha : V_\alpha \to \mathbb{R}^{2n}$,

$$\psi_\alpha(v) = (x^1,...,x^n,v^1,...,v^n) = (\vec{x},\vec{v})$$

for all $\alpha \in I$, where $\phi_\alpha(\mathrm{pr}_1(v)) = (x^1,...,x^n)$ and $(v^1,...,v^n)$ are the components of v with respect to the coordinate frame of (U_α,ϕ_α). I will use the vector arrow to denote vectors in $\mathbb{R}^n$. It is simple to verify that ψ_α is an injection.

Define $\mathcal{B} = \{\psi_\alpha^{-1}(W) : W \subset \mathbb{R}^{2n} \text{ open }, \alpha \in I\}$. I will show that this is a basis of a topology. It is clear that the sets in $\mathcal{B}$ cover TM, and as such it is sufficient to show that given any $B,B' \in \mathcal{B}$ and $p \in B\cap B'$ there is $B'' \subset \mathcal{B}$ such that $p \in B''$.

By assumption there exist $\alpha,\beta \in I$ and open $A,A' \subset \mathbb{R}^{2n}$ such that $B = \psi_\alpha^{-1}(A)$ and $B' = \psi_\beta^{-1}(A')$. I will show that $B\cap B' \in \mathcal{B}$. Denote $\tilde{V}_\alpha = \psi_\alpha(V_\alpha)$ and note that the composite map $\psi_\beta \circ \psi_\alpha^{-1} : \tilde{V}_\alpha \cap \tilde{V}_\beta \to \tilde{V}_\alpha \cap \tilde{V}_\beta$ acts as $(\vec{x}, \vec{v}_\alpha) \mapsto (\vec{x}, \vec{v}_\beta)$, where $\vec{v}_\alpha$ and $\vec{v}_\beta$ are the components of v with respect to the different coordinate frames. Since the components of vectors transform smoothly under a change of frame, $\psi_\beta \circ \psi_\alpha^{-1}$ is a diffeomorphism and $\psi_\beta \circ \psi_\alpha^{-1}(A')$ is open. It can be verified that

$$B \cap B' = \psi_\alpha^{-1}(A \cap (\psi_\beta \circ \psi_\alpha^{-1})^{-1}(A')),$$

and so $B \cap B' \in \mathcal{B}$ [Lee13, p. 22]. $\mathcal{B}$ is thus a basis of a topology.

The maps $\psi_\alpha : V_\alpha \to \tilde{V}_\alpha$ are homeomorphisms under the topology generated by $\mathcal{B}$ by construction. The fact that this topology is also second countable and Hausdorff is proven in [Lee13, p. 22].

As the coordinate changes $\psi_\beta \circ \psi_\alpha^{-1}$ are smooth, $\mathcal{A}' = \{(V_\alpha, \psi_\alpha) : \alpha \in I\}$ is a set of mutually compatible coordinate charts. There then exists the maximal atlas $\mathcal{A} = \{\text{All coordinate charts compatible with } \mathcal{A}'\}$ and TM equipped with the atlas $\mathcal{A}$ is a ($2n$-dimensional) manifold. □

Before moving on, I must define some operations on tensors. The first of these is the tensor product with which it is possible to combine two tensors into one of higher rank. To be more precise, the tensor product is the operation $\otimes : TM^{(r_1,m_1)} \times TM^{(r_2,m_2)} \to TM^{(r_1+r_2,m_1+m_2)}$ with

$$\begin{aligned}(T_1 \otimes T_2)(v_1, \ldots, v_{m_1+m_2}, \omega^1, \ldots, \omega^{r_1+r_2}) = T_1(v_1, \ldots, v_{m_1}, \omega^1, \ldots, \omega^{r_1}) \\ \times T_2(v_{m_1+1}, \ldots, v_{m_1+m_2}, \omega^{r_1+1}, \ldots, \omega^{r_1+r_2}).\end{aligned}$$

From the definition it is immediately clear that the tensor product is linear and associative but it is not commutative.

In turn, the rank of the tensor can be lowered by contracting a pair of its indices. To avoid clutter, I will define a contraction using an example. Let $T \in TM^{(3,2)}$. A **contraction** with respect to 2nd contravariant index and 1st covariant index of T is defined as the $(2,1)$-tensor for which

$$\mathrm{tr}_1^2 T(v, \omega_1, \omega_2) = \sum_l T(e_l, v, \omega_1, \sigma^l, \omega_2)$$

for all vector fields v and 1-forms ω. In the definition some frame e is chosen for every open set and σ is its dual frame. This definition can be proven to be independent of the chosen frame by expanding e_l and σ^l in terms of some other frame and its dual. As the transformation matrices for the frame fields and their duals are inverses of each other they cancel leaving the contraction invariant. The definition of a contraction is extended to other pairs of indices and tensors of different ranks in the natural way.

A multiple contraction is defined in a similar way. To use the same example, the contraction of T:s 1st and 2nd contravariant indeces with 3rd and 2nd covariant indeces respectively is the $(1,0)$-tensor

$$\mathrm{tr}^{12}_{32} T(\omega) = \sum_{lk} T(e_l, e_k, \omega, \sigma^k, \sigma_l)$$

for any 1-form ω. As with a single contraction this definition generalizes immediately to other tensors and sets of indices. When the exact set of indices being contracted is not relevant, I will denote a contraction simply as $\mathrm{tr}T$.

The main utility of contractions is that the actions of tensors on vectors and covectors can be expressed using them and tensor products. This will be useful later in proofs involving the covariant derivatives. To avoid complications with indices, I will again present the lemma using an example of a $(3,2)$-tensor.

Lemma 2.14. *The action of a $(3,2)$-tensor field T on vector fields v_1, v_2 and 1-forms $\omega^1, \omega^2, \omega^3$ can be expressed as a contraction*

$$T(v_1, v_2, \omega_1, \omega_2, \omega_3) = tr^{12345}_{34512}(T \otimes v_1 \otimes v_2 \otimes \omega^1 \otimes \omega^2 \otimes \omega^3).$$

This result generalizes to tensors of other ranks.

Proof. I will prove this for the example of a $(3,2)$-tensor, but the calculation for a tensor of any other rank is identical. It is sufficient to prove the result for each coordinate patch separately. Let U be a coordinate patch and e a frame on it. By the definitions of trace and tensor product

$$\begin{aligned}
&\mathrm{tr}^{12345}_{34512}(T \otimes v_1 \otimes v_2 \otimes \omega^1 \otimes \omega^2 \otimes \omega^3) \\
&= T(e_k, e_l, \sigma^a, \sigma^b, \sigma^c) v_1(\sigma^k) v_2(\sigma^2) \omega^1(e_a) \omega^2(e_b) \omega^3(e_c) \\
&= T_{kl}{}^{abc} v_1^k v_2^l \omega^1_a \omega^2_b \omega^3_c \\
&= T(v_1, v_2, \omega^1, \omega^2, \omega^3).
\end{aligned}$$

□

As the notation for a contraction used in 2.14 is very cumbersome, it is usually easier to use the component notation and explicitly write out the sums using the Einstein summation convention. However, later in this text the indices being contracted will be clear from the context. As I will not have to write the indices explicitly, the above notation will be sufficient.

2.3 The metric

In order to define the geometry of a manifold, it should be possible to measure lengths and angles of its tangent vectors. This requires each tangent space to be an inner product space. However, it turns out that a traditional inner product structure is somewhat too restrictive for many applications, and as such I will settle for the following definition of a metric.

Definition 2.15. Let M be a smooth manifold. A (0,2)-tensor field g on M is a **Riemannian metric** if it fulfills the following conditions for each $p \in M$:

1. Symmetry: $g_p(v, w) = g_p(w, v)$ for all $v, w \in TM_p$.
2. Positive definiteness: $g_p(v, v) \geq 0$ for all $v \in TM_p$, and $g_p(v, v) = 0$ if and only if $v = 0$.

A (0,2)-tensor field g is a **Pseudo-Riemannian metric** if it is symmetric and non-degenerate, that is for all $p \in M$ there is no $v \in TM_p$ for which $g_p(v, w) = 0$ for all $w \in TM_p$. The pair (M, g) is called a **(pseudo-)Riemannian manifold**.

Note that only the full Riemannian structure defines an inner product in each tangent space. The geometric interpretation for a Riemannian manifold is then much simpler than that of the pseudo-Riemannian case. In particular, since in that case the lengths of tangent vectors are positive definite, it is possible to define the lengths of curves in the usual way by integrating the length of the tangent vector.

Given a curve $\gamma : [a, b] \to M$ on a Riemannian manifold M, its length can be defined as

$$\ell(\gamma) = \int_a^b \sqrt{g(\dot{\gamma}(t), \dot{\gamma}(t))} dt.$$

In addition, the metric can be used to give any Riemannian manifold the metric space structure[Lee18, ch. 6].

In this text I will refer to pseudo-Riemannian metrics simply as a metric, since in most cases discussed here the full Riemannian structure is not needed.

3 Bundles

In addition to their manifold structure, tangent, cotangent and tensor bundles have also the structure of a vector bundle. This structure can be used to give one unified definition for both vector and tensor fields and later it will also be used to define more specialized objects such as vector fields along a curve.

In this section, I will first define a fibre bundle and its section. Afterwards I will further define a particular kind of a fibre bundle, a vector bundle, and discuss the pullback-operation on such bundles.

3.1 Fibre bundles and sections

A bundle is a space which can locally be represented as a product space, but which may still globally have a non-trivial topology. More rigorously

Definition 3.1. Let E and M be topological spaces and $\pi : E \to M$ a surjective and continuous map. The tuple (E, M, π) a **bundle** with **base space** M and **total space** E. The map π is called the **projection** and for $p \in M$ the preimage $\pi^{-1}(p)$ is called the **fibre** at p.

The bundle is often referred to as the bundle $\pi : E \to M$.

An intuitive picture of a bundle is a space $\pi^{-1}(p)$ attached to each point p in a continuous way. Since the definition of a bundle is very general, I will further restrict all bundles discussed here to be smooth, that is both E and M are assumed to be smooth manifolds, and the projection π is assumed to be a smooth function. I will assume this for all bundles considered in this text unless otherwise stated.

Despite the generality of the definition, bundles can be used to define an useful generalization of a smooth function on M.

Definition 3.2. A smooth map $s : M \to E$ is a (smooth) **section** on the bundle (E, M, π), if it fulfills the condition $\pi \circ s = id$ where id is the identity map.

A function between manifolds $f : X \to Y$ can be thought of as a section on the bundle $(X \times Y, M, \pi)$, where $\pi(x, y) = x$ for all $(x, y) \in X \times Y$. Bundles with less trivial topology can be used to define more complex objects. I will use the notation $\Gamma(E)$ for the space of all sections on the bundle $\pi : E \to M$.

The bundle used as the example is also a special kind of a bundle, whose fibres are all isomorphic with each other. This kind of a bundle is called a fibre bundle. As with general bundles I will assume that fibre bundles are smooth.

Definition 3.3. Let (E, M, π) be a bundle and F a manifold. The tuple (E, M, π, F) is a **fibre bundle** if for every point $p \in M$ there is a open neighborhood U and a diffeomorphism $\phi : U \times F \to \pi^{-1}(U)$ such that $(\pi \circ \phi)(x, v) = x$ for all $x \in U$ and $v \in F$. The set F is called the **fibre** of the bundle.

As with bundles, I will often refer to a fibre bundle with its projection. The maps $\phi : U \times F \to \pi^{-1}(U)$ are called local trivializations.

3.2 Vector bundles

It is not unreasonable to guess that the tangent bundle on a manifold has indeed a bundle structure where the fibre at each point is the tangent space. I will prove this later. However, since in this case each fibre is a vector space, the resulting bundle has more structure than a general fibre bundle. The same structure is also present in other tensor bundles, and it is useful to define a more restricted kind of a bundle, a vector bundle.

Definition 3.4. A bundle (E, M, π) is a (k-dimensional) **vector bundle** if for every $p \in M$ there is an open neighborhood U and a diffeomorphism $\phi : U \times \mathbb{R}^k \to \pi^{-1}(U)$ for which $(\pi \circ \phi)(x, v) = x$ for all $x \in U$ and the fibre $\pi^{-1}(x)$ has linear structure such that the map $\mathbb{R}^k \to \pi^{-1}(x) : v \mapsto \phi(x, v)$ is a linear isomorphism for all x.

Note that a vector bundle is a fibre bundle with the fibre $F = \mathbb{R}^k$. I can now give a more rigorous definition to the tangent bundle, since previously the concept was used merely as a shorthand.

Definition 3.5. Let $E = \bigcup_{p \in M} TM_p$ be the union of all tangent spaces on M equipped with the topology constructed in the proof to theorem 2.13. The **tangent bundle** on M is the bundle (E, M, π), where the projection is defined as $\pi(v_p) = p$ for all $v_p \in TM_p \subset E$. Denote the tangent bundle TM.

Theorem 3.6. *The tangent bundle on M is a vector bundle.*

Proof. Since the sets $TM_p \subset E$ are disjoint, π is well defined and it is a projection by construction. Clearly $\pi^{-1}(p) = TM_p$ for all $p \in M$. Let $p \in M$. By the definition of a manifold there is an (open) coordinate chart (U, φ) for which $p \in U$.

Any $v_x \in \pi^{-1}(U)$ can be expressed as $v_x = (x^1, ..., x^n, v^1, ..., v^n)$, where x^i are the coordinates of $x \in U$ and v^i are the components of v_x with respect to the coordinate frame. Define $\phi : U \times \mathbb{R}^n \to \pi^{-1}(U)$ with

$$\phi(x, (v^1, ..., v^n)) = (x, v^i(\partial_i)_x)$$

where ∂_i are the coordinate frame fields on U. As all its component maps are smooth, ϕ must be a smooth as well. The inverse of ϕ is straightforward to both construct and verify to be smooth in the same way. As such, ϕ is a diffeomorphism. The map ϕ also fulfills the condition $\pi(\phi(x, v)) = x$. It is then left to show that the map $\vec{v} \mapsto \phi(x, \vec{v})$ is a linear isomorphism.

Since the addition and scalar multiplication in TM_x are simply addition and multiplication of the components, $\vec{v} \mapsto \phi(x, \vec{v})$ is linear. As ϕ is a bijective, the map is a linear isomorphism. $\square$

The most important reason of giving the tangent bundle a vector bundle structure is that this allows a definition of vector fields that is more readily generalized. Analogously to the example of a function $f : X \to Y$ in the previous subsection, vector fields can be identified with smooth sections on the tangent bundle.

Proposition 3.7. *$VF(M)$ is linearly isomorphic to $\Gamma(TM)$, when the addition and scalar multiplication are defined on $\Gamma(TM)$ are defned as $(s_1 + s_2)(p) = s_1(p) + s_2(p)$ and $(cs_1)(p) = cs_1(p)$ for all $s_1, s_2 \in \Gamma(TM)$ and $c \in \mathbb{R}$.*

Proof. Since $s_1(p) + s_2(p)$, $cs_1(p) \in TM_p = \pi^{-1}(p)$ for all $s_1, s_2 \in \Gamma(TM)$, $c \in \mathbb{R}$ and $p \in M$ the operations are well defined. There is a trivial identification $\psi : VF(M) \to \Gamma(TM)$ with

$$\psi(v) = (p \mapsto (p, v_p)) \in \Gamma(TM)$$

for all vector fields v. It is straightforward to verify that ψ is both bijective and linear. □

The previous results on the tangent bundle can be generalized to all tensor bundles in order to give a definition of tensor fields in terms of bundles.

Definition 3.8. Let $E = \bigcup_{p \in M} TM_p^{(r,m)}$ equipped with its topology as a manifold. The **(r,m)-tensor bundle** on M is defined as the bundle (E, M, π) with the projection $\pi(T_p) = p$ for all $T_p \in TM_p^{(r,m)}$.

The tensor bundles are $r+m$-dimensional vector bundles and, similarly to the case with vector fields, the space of all (r,m)-tensor fields is linearly isomorphic with $\Gamma(TM^{(r,m)})$ when the vector operations are defined in the natural way. The proofs for this will be omitted, as they are almost identical to the case of a tangent bundle.

Many of the tools used with tensor bundles generalize to all vector bundles. One such a tool I will need later are frames on the vector bundle which are defined analogously to the tensor bundle case. A frame on a k-dimensional vector bundle is a k-tuple $(e_1, ..., e_k)$ of sections such that the tuple of vectors $(e_1(p), ..., e_k(p))$ forms a basis for $\mathbb{R}^k$ at every point $p \in M$. The components of a section $s : M \to E$ with respect to this frame are defined as the unique smooth maps $s^i : M \to \mathbb{R}$ for which $s = s^i e_i$. These definitions are discussed in more detail in [Bal99, p. 6].

The transformation properties derived for the components of a vector in section 2.2 generalize immediately to all sections. I will use the same notation for the frame change matrices.

3.3 Pullback bundle

For later use, I will also define the pullback operation on fibre bundles. The concept of a pullback is of great use in many different fields of differential geometry although the precise definition varies from case to case. In essence, given a map between manifolds and some structure on one of them, the map can be used to pull back similar structure to the other manifold. In this case the structure pulled back is that of a fibre bundle. The presentation here follows closely that in [Bal99].

Definition 3.9. Let (E, M, π, F) be a fibre bundle and M' a manifold. Let $f : M' \to M$ be a smooth map. The **pullback bundle** f^*E is defined as the fibre bundle (E', M', π', F) where

$$E' := \{(p', e) \in M' \times E \mid f(p') = \pi(e)\}$$

and the projection is $\pi'(p', e) = p'$ for all $(p', e) \in E'$.

Note that $(\pi')^{-1}(p') = \{(p', e) \, : \, \pi(e) = f(p')\} = \{(p', e) \, : \, e \in \pi^{-1}(f(p'))\}$ for all points $p' \in M'$. It is then simple to see that the fibre $(\pi')^{-1}(p')$ is naturally identified with with the corresponding fibre $\pi^{-1}(f(p'))$. As (E, M, π, F) is a fibre bundle, $(\pi')^{-1}(p')$ is isomorphic to F. Thus, the original bundle and the pullback bundle have the same fibre F.

I will next use this to show that the pullback bundle is well defined. It is necessary to show that E' is a submanifold of $M' \times E$ and the local trivializations exist. I will start with the former and only prove it locally, that is that every point has a neighborhood which is a submanifold.

Define $F_1, F_2 : M' \times E \to M$, $F_1(p', e) = f(p')$ and $F_2(p', e) = \pi(e)$. Let (p'_0, e_0) be any point in E' and (U, ϕ) be a coordinate patch on M with $f(p'_0) \in U$ and denote $V = F_1^{-1}(U) \cap F_2^{-1}(U) \cap E'$. As F_1 and F_2 are smooth V is an open neighborhood of (p'_0, e_0) with respect to the relative topology of E'. Define $F : F_1^{-1}(U) \cap F_2^{-1}(U) \to \mathbb{R}^n$ with $F(p', e) = \phi \circ F_1(p', e) - \phi \circ F_2(p', e)$. V is then the level set $F^{-1}(\{0\})$.

Denote the dimensions of M with n. Consider the differential matrix of the map F at an arbitrary point (p', e). As F_1 is independent of e and F_2 is independent of p', the differential of F is a block matrix where the blocks are the differentials of F_1 and F_2 respectively. As such, the rank of F at $(p', e) \in F_1^{-1}(U) \cap F_2^{-1}(U)$ is larger than or equal to the rank of F_2, which is the same as the rank of the projection map π at e.

The rank of the projection $\pi : E \to M$ at e is n. This can be proven by expressing π in the neighborhood W of e in a coordinate system adapted to the local trivialization $W \simeq \pi(W) \times F$ where F is the fibre of the bundle. In this coordinate system π is the identity map in its n first coordinates and is independent of the rest, and as such the rank is n.

The rank of F cannot be larger than the dimension of its target manifold $\mathbb{R}^n$. Using this and the previous result the rank of F at (p', e) is n. As the point $(p', e) \in F_1^{-1}(U) \cap F_2^{-1}(U)$ was not specified the rank of F is constant. Since $V = F^{-1}(\{0\})$ is the level set of a smooth map of constant rank, V is a submanifold [Lee13, theorem 5.12].

Next, I will construct the local trivializations. Let $p' \in M'$. Since (E, M, π, F) is a fibre bundle there is a neighbourhood U of point $f(p')$ and a local trivialization $\phi : U \times F \to \pi^{-1}(U)$.

The open set $V = f^{-1}(U)$ is an neighbourhood of p'. I will define $\psi : V \times F \to \pi'^{-1}(V)$ as

$$\psi(x, v) = (x, \phi(f(x), v))$$

for all $x \in V$ and $v \in F$. This is well defined since $\pi(\phi(f(x), v)) = f(x)$ and $\pi'((x, \phi(f(x), v))) = x$. ψ is also smooth since its component maps are smooth and clearly $\pi'(\psi(x, v)) = x$. It is then left to show that it is bijective.

Let $(y, e) \in \pi'^{-1}(V)$, that is $y \in V$ and $e \in E$ such that $\pi(e) = f(y)$. As ϕ is a diffeomorphism, there is $v \in F)$ such that $\phi(f(y), v) = e$. Then

$$\psi(y, v) = (y, \phi(f(y), v)) = (y, e)$$

and ψ is surjective. If $y, y' \in V$ and $y \neq y'$, it is trivial that $\psi(y, v) \neq \psi(y', v')$ for all $v, v' \in F$. Let $v, v' \in F$ such that $v \neq v'$ and $y, y' \in V$. Since ϕ is injective, $\phi(f(y), v) \neq \phi(f(y'), v')$ and so $\psi(y, v) \neq \psi(y', v')$. ψ is then injective.

The inverse of ψ can be verified to act as $\psi^{-1}(x, e) = (x, \mathrm{pr}_2(\phi^{-1}(e)))$ for all $(x, e) \in \pi^{-1}(V) \subset M' \times E$. In the previous expression pr_2 is the projection to the second component in $U \times F$. With the same reasoning as with ψ, the inverse of ψ is a smooth map. Thus ψ is a diffeomorphism and the local trivialization I was looking for.

The structure of a vector bundle is also preserved in a pullback. To prove this it is sufficient to show that the local trivialization satisfies the linearity requirement. Using the same definitions as for the general fibre bundle, let $x \in V$ and denote $L : \mathbb{R}^k \to \pi^{-1}(x)$, $L(v) = \psi(x, v) = (x, \phi(f(x), v))$.

Since ϕ satisfies the linearity requirement,

$$L(v + v') = (x, \phi(f(x), v + v')) = (x, \phi(f(x), v) + \phi(f(x), v')).$$

for all $v, v' \in \mathbb{R}^k$. The second component of $(x, e) \in \pi^{-1}(x)$ is an element of $\pi^{-1}(f(x))$ which is linearly isomorphic with $\mathbb{R}^k$. As such, vector operations on $\pi^{-1}(x)$ can be defined in a natural way on the second component. With this linear structure,

$$L(v + v') = (x, \phi(f(x), v)) + (x, \phi(f(x), v')).$$

L is also compatible with the scalar multiplication, which can be verified with the same calculation as above. The map ψ then satisfies the linearity requirement and f^*E is a vector bundle.

As with the entire bundle structure, sections on the bundle can also be pulled back into sections on the pullback bundle.

Definition 3.10. Let $s : M \to E$ be a section on $\pi : E \to M$ and $f : M' \to M$. The pullback of s is the section f^*s on the pullback bundle f^*E, where

$$(f^*s)(p') = (p', s \circ f(p')) \in E'.$$

Note that for every section s' on f^*E the first component of $s'(p')$ is just p' at every point. Since the first component of the section contains no additional information it is less cumbersome to work with only the latter component.

Lemma 3.11. *A smooth map $s' : M' \to f^*E$ is a section if and only if there is a smooth map $s : M' \to E$ such that $\pi \circ s = f$ and $s'(p') = (p', s(p'))$ for all points $p' \in M'$.*

Proof. Let $s' : M' \to f^*E$ be a section and define $s(p') = \mathrm{pr}_2(s'(p')) \in E$. Then by definition of the pullback bundle $\pi(s(p')) = f(p')$.

Let $s : M' \to E$ and $s' : M' \to f^*E$ be smooth maps and $s'(p') = (p', s(p'))$. This is well defined, since $\pi(s(p')) = f(p')$. As $\pi'(s'(p')) = p'$ trivially, s' is a section. □

The map $s : M' \to E$ is called a **section of** E **along** f. I will denote the space of all sections of E along f with $\Gamma_f(E)$. Since sections on f^*E are in one-to-one correspondence with sections on E along f, I will make no distinction between the two.

The most important example of a section along a map encountered later in this text is a tensor field along a curve. Let $\gamma : I \to M$ be a smooth curve on M, where $I \subset \mathbb{R}$ is an open interval. Since γ is a smooth map between manifolds it can be used to pull back the tensor bundles on M. Sections of tensor bundles along γ are called tensor fields along the curve.

As with proper sections, sections of a vector bundle along a map can be expressed in terms of their components. Let U be an open set on M and $\{e_1, \ldots, e_k\}$ a frame on it. Note that the vectors $\{e_1 \circ f(p'), \ldots, e_k \circ f(p')\}$ form a basis for all $p' \in f^{-1}(U)$. Thus any section s on U along f can be expanded as $s = s^i(e_i \circ f)$ for some maps $s^i : M' \to \mathbb{R}$. The maps $(s^1, \ldots, s^k)$ are called the components of s with respect to the pulled back frame.

The components of sections along a map transform as usual. Let $\{e_1, \ldots, e_k\}$ and $\{e'_1, \ldots, e'_k\}$ be frames in overlapping open sets U and U' on M and A be the change of frame $e_i = A^j_i e'_j$. This immediately implies

$$(e_i \circ f)(p') = (A^j_i \circ f)(p')(e'_j \circ f)(p')$$

for all points $p' \in f^{-1}(U \cap U')$. Thus, $A \circ f$ is gives the change of frame on the pullback bundle and components sections transform as $s'^i = (A^i_j \circ f)s^j$ where s^i and s'^j are the components of a section s in the different frames.

4 Curvature

In this section, I will define and discuss the curvature of a pseudo-Riemannian manifold. For this end, I will first define a connection on the tangent bundle and after that, I will show that there is a unique torsion free metric compatible connection, namely the Levi-Civita connection. Finally, I will define the curvature tensor.

4.1 Connection and the covariant derivative

First step towards defining the curvature tensor is to define a directional derivative for vector fields and tensors in general. However, as vectors at different points are objects of different tangent spaces, they cannot be summed, and thus the usual difference quotinent definition is not usable. Instead, I begin by noting that the operation of the usual directional derivative is linear and satisfies the Leibnitz rule and demand the same of the new operation.

Definition 4.1. Let $\pi : E \to M$ be a smooth vector bundle over a manifold M and $\Gamma(E)$ be the set of smooth sections over E. The map $\nabla : VF(M) \times \Gamma(E) \to \Gamma(E)$, $(X,Y) \mapsto \nabla_X Y$, is a **connection** or a **covariant derivative** if it fulfills the following properties:

- For all $X_1, X_2 \in VF(M)$, $f_1, f_2 \in C^\infty(M)$ and $Y \in \Gamma(E)$

$$\nabla_{f_1 X_1 + f_2 X_2} Y = f_1 \nabla_{X_1} Y + f_2 \nabla_{X_2} Y.$$

- For all $X \in VF(M)$, $a, b \in \mathbb{R}$ and $Y_1, Y_2 \in \Gamma(E)$

$$\nabla_X (aY_1 + bY_2) = a\nabla_X Y_1 + b\nabla_X Y_2.$$

- For all $X \in VF(M)$, $Y \in \Gamma(E)$ and $f \in C^\infty(M)$

$$\nabla_X (fY) = f\nabla_X Y + X(f)Y.$$

This definition does not specify a unique connection, and without additional structure on the manifold there is no natural way to choose one connection over another. It turns out that, as is the case with tensors, a connection on the tangent bundle can be defined using a set of functions in each coordinate patch.

Lemma 4.2. *Let ∇ be a connection on the tangent bundle TM, $\{e_1, ..., e_n\}$ be a frame of vector fields on an open set $U \subset M$ and $X, Y \in VF(M)$. Then*

$$\nabla_X Y = X^i (e_i(Y^k) + \omega^k_{ij} Y^j) e_k$$

*in U, where the **connection coefficients** ω^k_{ij} are defined as $\nabla_{e_i} e_j = \omega^k_{ij} e_k$.*

Proof. In U the vector fields can be expressed in terms of their components $X = X^k e_k$, $Y = Y^k e_k$. Then

$$\begin{aligned}\nabla_X Y &= \nabla_{X^k e_k}(Y^l e_l)\\ &= X^k \nabla_{e_k}(Y^l e_l)\\ &= X^k((\nabla_{e_k} Y^l)e_l + Y^l \nabla_{e_k} e_l)\\ &= X^k((\nabla_{e_k} Y^l)e_l + Y^l \omega^i_{kl} e_i)\\ &= X^i(e_i(Y^k) + \omega^k_{ij} Y^j)e_k,\end{aligned}$$

where in the last line the labeling of the sums was changed. □

This calculation did not actually use any of the particular properties of a tangent bundle, and as such it can immediately be generalized to any vector bundle: If ∇ is a connection on a vector bundle $\pi : E \to M$, $(e_1, ..., e_k)$ is a frame on an open $U \subset M$, $X \in VF(M)$ and s is a section on E,

$$\nabla_X s = (X(s^i) + \omega^i_j(X)s^j)e_i,$$

where $\nabla_X e_j = \omega^i_j(X)e_i$. The object ω is called the connection 1-form associated with ∇ and the usual connection coefficients are $\omega^i_{jk} = \omega^i_j(e_k)$.

As seen from the above expression, a connection on the tangent bundle is defined uniquely in terms of its connection coefficients. However, since $\nabla_X s$ has to be a well defined section, the connection coefficients have to transform in a particular way when changing a frame.

As before, let $\{e_1, ..., e_k\}$ and $\{e'_1, ..., e'_k\}$ be frames in overlapping open sets U and U' on M and A be the change of frame $e_i = A^j_i e'_j$. Denote $\nabla_{e_i} e_j = \omega^k_{ij} e_k$ and $\nabla_{e'_i} e'_j = \tilde{\omega}^k_{ij} e'_k$ and $w = \nabla_X s$ for some $X \in VF(M)$ and $s \in \Gamma(E)$. Since w is a section, its components transform as

$$\begin{aligned}w'^i &= A^i_k w^k\\ &= A^i_k(X(s^k) + \omega^k_j(X)s^j)\\ &= A^i_k(X((A^{-1})^k_j s'^j) + \omega^k_j(X)(A^{-1})^j_l s'^l)\\ &= X(s'^i) + (A^i_k X((A^{-1})^k_j) + A^i_k \omega^k_l(X)(A^{-1})^l_j)s'^j.\end{aligned}$$

On the other hand w'^i can also be expressed as $w'^i = X(s'^i) + \tilde{\omega}^i_j(X)s'^j$ and comparing the two expressions results in

$$\tilde{\omega}^i_j(X) = A^i_k X((A^{-1})^k_j) + A^i_k \omega^k_l(X)(A^{-1})^l_j.$$

In terms of connection coefficients the condition reads

$$\tilde{\omega}^i_{jk} = A^i_l \partial_k((A^{-1})^l_j) + A^i_r \omega^r_{lk}(A^{-1})^l_j.$$

For more details see [Bal99, section 2.1].

In principle it would be possible to define a connection on each of the tensor bundles separately. However, once a connection is chosen for the tangent bundle, there is a unique connection on each tensor bundle compatible with the tensor operations defined in section 2.

Theorem 4.3. *Let ∇ be a connection on TM. The following requirements define an unique connection on each of the tensor bundles over M:*

- *On $TM^{(0,0)} = M \times \mathbb{R}$ the connection is defined as $\nabla_X f = X(f)$.*
- *On TM the connection agrees with ∇.*
- *The connections fulfill the generalized Leibnitz rule*

$$\nabla_X(T_1 \otimes T_2) = T_1 \otimes \nabla_X T_2 + \nabla_X T_1 \otimes T_2$$

for all tensors T_1, T_2 of any rank.

- *The connection commutes with all contractions.*

Proof. I will prove the theorem by explicitly constructing the coordinate expression for the action of the connection in each $TM^{(r,m)}$.

In $TM^{(0,0)}$ the action on f is given by $\nabla_X f = X(f)$. Since X is a derivation, this expression is clearly linear in X and fulfills the Leibnitz rule, and so it defines a connection. In $TM^{(1,0)}$ ∇ is a connection by definition.

Let $\omega \in TM^{(0,1)}$ and note that $\omega(Y) = \mathrm{tr}(\omega \otimes Y)$ for all vector fields Y. Using lemma 2.14 and the fact that the connection commutes with traces

$$\begin{aligned}
\nabla_X(\omega(Y)) &= \nabla_X \mathrm{tr}(\omega \otimes Y) \\
&= \mathrm{tr}(\nabla_X(\omega \otimes Y)) \\
&= \mathrm{tr}((\nabla_X \omega) \otimes Y + \omega \otimes (\nabla_X Y)) \\
&= \mathrm{tr}((\nabla_X \omega) \otimes Y) + \mathrm{tr}(\omega \otimes (\nabla_X Y)) \\
&= (\nabla_X \omega)(Y) + \omega(\nabla_X Y),
\end{aligned}$$

from which follows $(\nabla_X \omega)(Y) = \nabla_X(\omega(Y)) - \omega(\nabla_X Y) = X(\omega(Y)) - \omega(\nabla_X Y)$. The resulting expression is fully defined in terms of the connections on $TM^{(1,0)}$ and $TM^{(0,0)}$ and fulfills both the Leibnitz rule and linearity, and so it describes a well defined connection on $TM^{(0,1)}$.

The definition for a (r, m) tensor can be solved in the same way. Let $T \in TM^{(r,m)}$. Using lemma 2.14 and commuting the contraction and the connection gives

$$\begin{aligned}\nabla_X(T(v_1,...,\omega^m)) &= \nabla_X(\mathrm{tr}(T \otimes v_1 \otimes ... \otimes \omega^m)) \\ &= \mathrm{tr}(\nabla_X(T \otimes v_1 \otimes ... \otimes \omega^m)) \\ &= \mathrm{tr}((\nabla_X T) \otimes v_1 \otimes ... \otimes \omega^m + T \otimes \nabla_X(v_1 \otimes ... \otimes \omega^m)) \\ &= ... \\ &= \mathrm{tr}((\nabla_X T) \otimes v_1 \otimes ... \otimes \omega^m + T \otimes \nabla_X(v_1) \otimes ... \otimes \omega^m \\ &\quad + ... + T \otimes v_1 \otimes ... \otimes \nabla_X(\omega^m)) \\ &= (\nabla_X T)(v_1,...,\omega^m) + T(\nabla_X v_1,...,\omega^m) + ... + T(v_1,...,\nabla_X \omega^m).\end{aligned}$$

where the generalized Leibnitz rule was used repeatedly. From this expression I can solve

$$(\nabla_X T)(v_1,...,\omega^m) = \nabla_X(T(v_1,...,\omega^m)) - T(\nabla_X v_1,...,\omega^m) - ... - T(v_1,...,\nabla_X \omega^m).$$

This expression is well defined in terms of covariant derivatives of functions, vector fields and 1-forms. It can also be verified to fulfill the linearity and Leibnitz rule requirements of a connection, since each term fulfills them separately. □

The expression for the covariant derivative of a tensor can be written out in terms of a frame. I will do this explicitly for the case of a $(0,2)$-tensor such as a metric. Expanding both argument vector fields in terms of a frame $\{e_1,...,e_n\}$ results in

$$\begin{aligned}(\nabla_X T)(v_1, v_2) &= X(T(v_1, v_2)) - T(\nabla_X v_1, v_2) - T(v_1, \nabla_X v_2) \\ &= X^k \partial_k (T_{ij} v_1^i v_2^j) - T_{ij} X^k (\partial_k v_1^i + \omega^i_{kl} v_1^l) v_2 \\ &\quad - T_{ij} v_1^i X^k (\partial_k v_2^j + \omega^j_{kl} v_2^l) \\ &= X^k (\partial_k T_{ij} - \omega^l_{ki} T_{lj} - \omega^l_{kj} T_{il}) v_1^i v_2^j.\end{aligned}$$

The component of $\nabla_X T$ can be identified as $\partial_k T_{ij} - \omega^l_{ki} T_{lj} - \omega^l_{kj} T_{il}$.

In the rest of the text I will always use these compatible connections for the tensor bundles, and when there is no danger of confusion I will refer to them all as "the connection". The symbol ∇ will always refer to the connection relevant to the tensor it operates on.

4.2 Connection on a pullback bundle

A connection is another structure that can be pulled back with a map. First it is necessary to define a pushforward of a tangent vector.

Definition 4.4. Let M and M' be manifolds and v be a tangent vector at the point $p \in M$. Let $f : M \to M'$ be a smooth map and σ a representative curve of v. The **pushforward of** v is the tangent vector $f_* v = [f \circ \sigma]$ at $f(p) \in M'$.

Note that while individual vectors can be pushed forward between manifolds, the pushforward of a vector field is not well defined unless f is a diffeomorphism. An example of a pushforward that will be useful later on is the pushforward along a curve $\gamma : I \to M$ of a coordinate basis vector $(\partial_t)_p$ at some point $p \in I$. For any smooth function g the action of the pushforward vector is

$$\gamma_*(\partial_t)_p(g) = \frac{d}{dt}(g \circ \gamma(t))|_{t=p} = \dot{\gamma}(p)(g).$$

As such the pushforward is $\gamma_* \partial_t = \dot{\gamma}$ at all points on the curve.

It is sufficient to define the pullback connection on sections along f and I will do that in terms of their components.

Definition 4.5. Let $f : M' \to M$ be a smooth map, $\pi : E \to M$ a vector bundle and ∇ a connection on E with the connection 1-form ω. The **pullback connection** ∇^f is the connection on f^*E for which

$$\nabla^f_X s = (X(s^i) + (\omega^i_j \circ f)(f_* X)s^j)(e_i \circ f),$$

for all sections on M' along f, vector fields X on M' and any frame $\{e_i\}_{i=1}^n$ on $U \subset M$.

The section along f $\nabla^f_X s$ can be verified to be well defined using the transformation properties of ω discussed in the previous subsection [Bal99, chapter 2.3]. ∇^f is clearly linear in its arguments, since ω is also linear. As X acts as a derivation, ∇^f also satisfies the Leibniz rule. Thus, ∇^f is a well defined connection.

Note that the covariant derivative of a pullback section is the pullback of a differentiated section. If there is no danger of confusion, I will drop the superscript f from ∇^f.

An important and simple example of a pullback connection is a **connection along a curve**. I will consider that in the case of tangent vectors. Let $\gamma : I \to M$, where I is an interval, be a curve. The covariant derivative along γ of a vector field v along γ in the direction of the only coordinate basis vector on I is

$$\nabla^\gamma_{\partial_t} v = \nabla_{\partial_t} v = (\partial_t v^i + \omega^i_{jk} v^j \dot{\gamma}^k)(e_i \circ \gamma).$$

In the rest of the text I will denote $\nabla_t = \nabla_{\partial_t}$.

4.3 The Levi-Civita connection

While choosing the connection for the tangent bundle specifies the connections on the tensor bundles, there was no obvious way to choose the first reference connection. Introducing a metric on the manifold changes this. First, I need to define the torsion of a connection.

Definition 4.6. The **torsion tensor** is the tensor field T defined as

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y],$$

where X, Y are vector fields on M and $[X, Y]$ is their Lie bracket.

The components of the torsion tensor with respect to a frame can be read immediately from the definition. Let e be a frame in M and $[e_i, e_j] = \gamma^k_{ij} e_k$. The components of the torsion tensor T in this frame are given by

$$T^i_{jk} = \omega^i_{jk} - \omega^i_{kj} - \gamma^i_{jk}.$$

A connection for which the torsion tensor vanishes everywhere is said to be **torsion free**. Since partial derivatives on a smooth function commute, the Lie bracket of the coordinate frame fields vanish, which immediately leads to a useful symmetry property of a torsion free connection.

Lemma 4.7. *Let e be a coordinate frame in some coordinate patch. The torsion tensor vanishes in this patch if and only if $\omega^i_{jk} = \omega^i_{kj}$ with respect to this frame.*

Since a metric defines a (not necessarily positive definite) inner product on the tangent spaces, it is natural to demand that the covariant differentiation obeys the Leibnitz rule with regards to this inner product, that is

$$\nabla_X g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z).$$

for all vector fields X, Y, Z. Using the Leibnitz rule for the component expression of the derivative yields

$$\nabla_X g(Y, Z) = X^i((\nabla_i g_{jk}) Y^j Z^k + g_{jk}(\nabla_i Y^j) Z^k + g_{jk} Y^j (\nabla_i Z^k)),$$

where $\nabla_i = \nabla_{e_i}$. This gives the desired expression if the first term vanishes, which in turn motivates the following definition:

Definition 4.8. Let M be a pseudo-Riemannian manifold and ∇ be a connection on M. The connection is **metric compatible** if

$$\nabla_X g = 0$$

for any vector field X.

Demanding that a connection is both torsion free and metric compatible singles out one special connection, the **Levi-Civita connection**, from all the possible choices.

Theorem 4.9. *Let M be a pseudo-Riemannian manifold. There is an unique connection on M which is torsion free and metric compatible.*

Proof. I will prove the existence and uniqueness of the Levi-Civita connection by constructing an explicit expression for the connection coefficients in a coordinate patch in the coordinate frame. In the torsion free and metric compatible case I will denote the connection coefficients Γ^i_{jk}. Suppose that ∇ is a metric compatible and torsion free metric. By metric compatibility

$$
\begin{aligned}
0 = \nabla_i g_{jk} = \partial_i g_{jk} - \Gamma^l_{ij} g_{lk} - \Gamma^l_{ik} g_{lj}, \\
0 = \nabla_j g_{ki} = \partial_j g_{ki} - \Gamma^l_{jk} g_{li} - \Gamma^l_{ji} g_{lk}, \\
0 = \nabla_k g_{ij} = \partial_k g_{ij} - \Gamma^l_{ki} g_{lj} - \Gamma^l_{kj} g_{li}.
\end{aligned}
$$

Since the connection is torsion free by lemma 4.7 its coefficients are symmetric in the lower indices. The metric is symmetric as well, and so subtracting the last two equations from the first yields

$$0 = \partial_i g_{jk} - \partial_j g_{ki} - \partial_k g_{ij} + 2\Gamma^l_{jk} g_{li}.$$

The connection coefficients can be solved from this expression:

$$\Gamma^l_{jk} = \frac{1}{2} g^{li} (\partial_j g_{ik} + \partial_k g_{ij} - \partial_i g_{jk}).$$

This expression can be verified to fulfill the transformation properties of connection coefficients, and therefore the connection is well defined. As such, the torsion free and metric compatible connection exists. As I constructed the explicit connection coefficients, the connection must be unique as well. □

The connection coefficients with respect to the coordinate frame were calculated in the above proof and they will be referred to as the **Christoffel symbols** of the metric. In the rest of the text, unless otherwise specified, I will use the Levi-Civita connection.

4.4 Riemannian curvature

There is now enough tools available to define the curvature tensor on the manifold. The definition of the curvature tensor itself does not refer to any particular connection, every connection results in their own definition of curvature. However, the familiar curvature of a surface is found by choosing the Levi-Civita connection. This is explored fully in [Lee18, ch. 7], but in this text I will only need the definition itself.

Definition 4.10. Let M be a pseudo-Riemannian manifold with a connection. The **curvature tensor** R is defined by

$$R(X,Y)W = \nabla_X \nabla_Y W - \nabla_Y \nabla_X W - \nabla_{[X,Y]} W$$

for any vector fields X, Y and W. If the connection is the Levi-Civita connection, the curvature tensor is called the **Riemann tensor**.

5 Jacobi fields

5.1 Geodesics

Geodesics are the generalization of euclidean straight lines to general manifolds. The usual definition of a straight line in $\mathbb{R}^n$ is the shortest path between two points, but this approach runs soon into trouble with a general metric. The problems are especially prevalent with non-Riemannian metrics, with which the minimizing curve between two points does not exist at all.

Instead geodesics must be defined by another property of straight lines, that is in $\mathbb{R}^n$ (certain parametrizations of) straight lines are the only curves with a zero acceleration vector. The acceleration of a curve in $\mathbb{R}^n$ can be defined simply as componentwise derivative along the curve of the tangent vector. This definition generalizes to pseudo-Riemannian manifolds by taking the covariant derivative along the curve as defined in section 4.2 of the tangent vector.

Motivated by this, I first define the parallel transport of a vector along a curve.

Definition 5.1. Let $\gamma : [a,b] \to M$ be a curve with $\gamma(a) = p$ and v_0 a tangent vector at p. The **parallel transport** of v along γ, denoted $v(\lambda)$, is defined as the vector field v along curve γ which satisfies the equation

$$\nabla_\lambda v(\lambda) = 0$$

with initial condition $v(a) = v_0$, where λ is curve parameter.

The existence and uniqueness of a parallel transport vector field can be proven using the existence and uniqueness theorems on ODEs [Lee18, theorem 4.32]. The acceleration of the curve is therefore zero if and only if the curve parallel transports its own tangent vector.

Definition 5.2. A **geodesic** is a curve $\lambda \mapsto \gamma(\lambda)$ which parallel transports its tangent vector, $\nabla_\lambda \dot{\gamma} = 0$. This equation is called the geodesic equation.

Writing this equation out in a coordinate frame yields the usual component representation of the geodesic equation

$$\frac{d^2\gamma^i}{d\lambda^2} + \omega^i_{jk}\frac{d\gamma^j}{d\lambda}\frac{d\gamma^k}{d\lambda} = 0,$$

where λ is the curve parameter along λ. Note that whether a curve is a geodesic can be dependent on its parametrization.

5.2 The exponential map

Geodesics on a manifold have some collective properties which are necessary when discussing Jacobi fields later. These properties are most easily stated using the so called exponential map which I will define shortly. I will state most results in this subsection without proof and the reader may consult [Lee18] for more details.

It is a well-known result that a system of linear ODE:s has a solution for all initial values, and the solution is a smooth function of the initial values [Lee18, theorem 4.31]. Since in every coordinate patch the geodesic equation can be expressed as a system of linear ODE:s, given any point $p \in M$ and a tangent vector $v \in TM_p$, at least locally there exists a geodesic with a tangent vector v at p.

It is possible to extend this beyond the single coordinate chart. Stating this is simplified by first defining the concept of a maximal geodesics. In short, a geodesic $\gamma : I \to M$ is maximal if there is no geodesic $\tilde{\gamma} : \tilde{I} \to M$ for which $I \subsetneq \tilde{I}$ and $\gamma(t) = \tilde{\gamma}(t)$ for all $t \in I$. The main existence result for geodesics can then be stated as follows:

Theorem 5.3. *[Lee18, corollary 4.28] For every $p \in M$ and $v \in TM_p$ there is a unique maximal geodesic $\gamma_{p,v} : I \to M$, $0 \in I$, with the initial conditions $\gamma_{p,v}(0) = p$ and $\dot{\gamma}_{p,v}(0) = v$.*

This geodesic is called the geodesic with initial point p and the initial velocity v and I will denote it γ_v. I will not explicitly state the initial point when it is not necessary.

The exponential map can be defined using these geodesics, although it is first necessary to specify its domain. Let $\mathcal{E} = \{v \in TM : \gamma_v \text{ defined on } [0,1]\}$.

Definition 5.4. The **exponential map** is defined as $\exp : \mathcal{E} \to M$ with $\exp(v) = \gamma_v(1)$ for all $v \in TM$. For all $p \in M$ the restricted exponential map $\exp_p$ is the restriction of the exponential map to the set $\mathcal{E}_p = \mathcal{E} \cap TM_p$.

There are many important results concerning the exponential map, but in this text I will use only a few of them.

Theorem 5.5. *[Lee18, prop. 5.19] The exponential map has the following properties:*

- *For each $v \in TM$ the geodesic with initial velocity v is given by $\gamma_v(t) = \exp(tv)$ whenever either side is defined.*
- *The exponential map is smooth.*
- *$\mathcal{E}$ is an open subset of TM and contains the image of the zero vector field.*

The first of these properties is clear from the definition of the exponential map and the second one can be proven by using the smoothness of the solution to a system of linear ODE:s.

5.3 Jacobi fields

I can now finally define the Jacobi fields. In non-rigorous terms a Jacobi field is the variation vector field of some family of geodesics, that is it encodes how the geodesic changes with the family parameter. I will give the rigorous definition next:

Definition 5.6. Let $G : [0.1] \times (-\epsilon, \epsilon) \to M$, $G(t,s) = \gamma_s(t)$, where $\epsilon > 0$ is some real number, the curve $t \mapsto \gamma_s(t)$ is a geodesic for every s and G is smooth. G is called a **one-parameter family of geodesics** and denote $\gamma = \gamma_0$. The **variation field** of this family is the vector field along γ

$$J(\lambda) = \partial_s \gamma_s(\lambda)\Big|_{s=0},$$

where λ is the curve parameter along γ. The partial denotes a tangent vector to the curve $s \mapsto \gamma_s(\lambda)$.

A family of geodesics G satisfies a number of identities that will be of use later. I will use the notation in which $(\partial_t G)(t', s')$ is the tangent vector of the curve $t \mapsto G(s', t)$ at point $t = t'$ and similarly $(\partial_s G)(t', s')$ is the tangent vector to the curve $s \mapsto G(s, t')$ at $s = s'$. I will not explicitly state the point (s', t') where the tangent vectors are evaluated.

Lemma 5.7. *A family of geodesics G satisfies the equation $\nabla_t \partial_s G = \nabla_s \partial_t G$.*

Proof. I will show this in a coordinate chart. In this coordinate chart G can be expressed as $G(s,t) = (\gamma_s^1(t), ..., \gamma_s^n(t))$. The i:th component of $\nabla_t \partial_s G$ with respect to the coordinate frame is then

$$(\nabla_t \partial_s G)^i = \partial_t(\partial_s \gamma_s^i) + \Gamma^i_{jk}(\partial_t \gamma_s^j)(\partial_s \gamma_s^k).$$

Since the Christoffel symbol is symmetric in the lower indices, the latter term is immediately symmetric with respect to changing t and s. Since partial derivatives of a smooth function commute, the first term is also symmetric, which proves the proposition. □

Lemma 5.8. $[\nabla_t, \nabla_s]V = R(\partial_t G, \partial_s G)V$ *for any vector field V along G.*

Proof. I will here give only the outline of the proof. The reader can find the full details in [Lee18, proposition 7.5].

It is sufficient to prove the lemma locally. As in the previous proof, in a coordinate chart G can be expressed as $G(s,t) = (\gamma_s^1(t), ..., \gamma_s^n(t))$. The vector field V can be expanded as $V(s,t) = V^i(s,t)\partial_i$, where ∂_i are the coordinate frame fields along G.

Calculating the action of the commutator on V using the Leibniz rule results in

$$[\nabla_t, \nabla_s]V = V^i(\nabla_s \nabla_t \partial_i - \nabla_t \nabla_s \partial_i).$$

This form already resembles the Riemann tensor. If ∇ was a connection on the manifold M instead of its pullback along G, the rest of the calculation would proceed by expanding the covariant derivatives as $\nabla_t = (\partial \gamma_s^j / \partial t)\nabla_j$ and $\nabla_s = (\partial \gamma_s^j / \partial s)\nabla_j$.

Substituting this to the commutator results in

$$\begin{aligned} V^i(\nabla_s\nabla_t\partial_i - \nabla_t\nabla_s\partial_i) &= V^i(\frac{\partial\gamma_s^k}{\partial s}\nabla_k(\frac{\partial\gamma_s^j}{\partial t}\nabla_j\partial_i) - \frac{\partial\gamma_s^k}{\partial t}\nabla_k(\frac{\partial\gamma_s^k}{\partial s}\nabla_j\partial_i)) \\ &= V^i\frac{\partial\gamma_s^j}{\partial t}\frac{\partial\gamma_s^k}{\partial s}R(\nabla_j\nabla_k\partial_i - \nabla_k\nabla_j\partial_i) \\ &= V^i\frac{\partial\gamma_s^j}{\partial t}\frac{\partial\gamma_s^k}{\partial s}R(\partial_j,\partial_k)\partial_i, \\ &= R(\partial_t G, \partial_s G)V. \end{aligned}$$

where in the second equality the Leibniz rule was used and most of the terms cancel.

This calculation can be applied to vector fields along G as well by noting that the frame fields along G can be extended to corresponding frame fields on M. After this the calculation can be carried out on M and the resulting vector field can be identified with the corresponding vector field along G. □

Using these two lemmas I can prove that a variation field of any family of geodesics satisfies the so called Jacobi equation.

Theorem 5.9. *Let J be the variation field of the family G. J satisfies the **Jacobi equation***

$$\nabla_t^2 J + R(J,\dot\gamma)\dot\gamma = 0.$$

where $\dot\gamma = (\partial_t G)(t,0)$.

Proof. Since G is a family of geodesics, it satisfies the geodesic equation

$$\nabla_t\partial_t G = \nabla_t\dot\gamma_s = 0$$

for all t and s. Taking the s-covariant derivative and commuting the outermost derivatives gives

$$0 = \nabla_s\nabla_t\partial_t G = (\nabla_t\nabla_s + [\nabla_s,\nabla_t])\partial_t G.$$

Using lemma 5.7 the first term becomes

$$\nabla_t\nabla_s\partial_t G = \nabla_t^2\partial_s G.$$

As $\nabla_t G = \dot\gamma_s$ is a vector field, lemma 5.8 gives

$$[\nabla_s,\nabla_t]\dot\gamma_s = R(\partial_s G,\partial_t G)\dot\gamma_s = R(\partial_s G,\dot\gamma_s)\dot\gamma_s.$$

Substituting the previous two results yields

$$0 = \nabla_t^2\partial_s G + R(\partial_s G,\dot\gamma_s)\dot\gamma_s$$

and evaluating this expression at $s = 0$ results in the Jacobi equation

$$\nabla_t^2 J + R(J,\dot\gamma)\dot\gamma = 0.$$ □

Vector fields which satisfy the Jacobi equation are called **Jacobi fields**, and as proven above, any variation field of a family of geodesics is a Jacobi field. The converse turns out to be true as well in many cases. In the following proof I will assume the existence and uniqueness theorems for the systems of differential equations.

Theorem 5.10. *Let $\gamma : I \to M$, where I is a compact interval in $\mathbb{R}$. Any Jacobi field on γ is the variation field of some family of geodesics.*

Proof. This proof is inspired by [Ilm20, theorem 5.8] and [Lee18, prop. 10.4].

Let J be a vector field on γ satisfying the Jacobi equation. It can be assumed without loss of generality that $0 \in I$. Let $a : (-\epsilon, \epsilon) \to M$ be a curve with $a(0) = \gamma(0)$ and $\dot{a}(0) = J(0)$. This is always possible, since $J(0)$ is an equivalence class of curves satisfying these conditions and I can just choose one of those.

Let $b(s)$ be any vector field along $a(s)$ for which $b(0) = \dot{\gamma}(0)$ and $\nabla_s b|_s = \nabla_t J(0)$. Define $G' : \mathcal{D} \to M$, $G'(t,s) = \exp(tb(s))$, where $\mathcal{D} \subset \mathbb{R}^2$ is the largest domain where G' can be defined. Note that the curve $t \mapsto G'(t,0)$ is by definition the maximal geodesic which agrees with γ in the overlap of domains and as such $I \times \{0\} \subset \mathcal{D}$. It remains to show that there is $\delta \in \mathbb{R}$ such that $I \times (-\delta, \delta) \subset \mathcal{D}$.

The domain of the exponential map $\mathcal{E}$ is an open set and as $\mathcal{D}$ is its preimage under the continuous map $(t,s) \mapsto tb(s)$, $\mathcal{D}$ must be open as well. As $\mathcal{D}$ is open, for every $p \in I$ I can choose $\delta_p = \frac{1}{2}\mathrm{dist}((p,0), \mathbb{R}^2 - \mathcal{D}) > 0$. Here dist is the Euclidean distance function $\mathrm{dist}(x, U) := \inf\{|x-y| : y \in U \subset \mathbb{R}^n\}$ for all $x \in \mathbb{R}^2$. Then $B(\delta_p, p) = \{x \in \mathbb{R}^2 : |x-p| < \delta_p\} \subset \mathcal{D}$. As I is compact and the distance function is a continuous map there is $\delta = \min\{\delta_p : p \in I\} > 0$. This implies

$$I \times (-\delta, \delta) \subset \bigcup_{p \in I} B(\delta_p, p) \subset \mathcal{D},$$

and I can define the family of geodesics G as the restriction of G', $G : I \times (-\delta, \delta) \to M$, $G(t,s) = G'(t,s)$.

G is a variation of γ as $G(t,0) = \gamma(t)$ for all $t \in I$. Denote the variation field of G by J'. Note that $J'(0) = \partial_s G(s,0) = \nabla_s b(0) = J(0)$ and using lemma (...)

$$\nabla_t J'(0) = \nabla_t \partial_s G(0,0) = \nabla_s \partial_t G(0,0).$$

Since $\partial_t G(0,0) = \dot{\gamma}(0) = b(0)$,

$$\nabla_t J'(0) = \nabla_s b(0) = \nabla_t J(0)$$

by the construction of b.

As J' is a variation field it must satisfy the Jacobi equation. Since J and J' satisfy the same system of linear second order ODEs and agree to first order at 0, the uniqueness of solutions demands $J = J'$. This completes the proof. □

References

[Bal99] Werner Ballman. Vector bundles and connections, 1999.

[Ilm20] Joonas Ilmavirta. Geometry of geodesics, 2020.

[Ish99] Chris J. Isham. *Modern Differential Geometry for Physicists*. World Scientific, second edition, 1999.

[Lee13] John Lee. *Introduction to Smooth Manifolds*. Springer, second edition, 2013.

[Lee18] John Lee. *Introduction to Riemannian Manifolds*. Springer, second edition, 2018.