A brief introduction to Semi-Riemannian geometry and general relativity

Hans Ringström

May 5, 2015
Contents

1 Scalar product spaces 1
  1.1 Scalar products ................................................. 1
  1.2 Orthonormal bases adapted to subspaces ...................... 3
  1.3 Causality for Lorentz scalar product spaces ................. 4

2 Semi-Riemannian manifolds 7
  2.1 Semi-Riemannian metrics ....................................... 7
  2.2 Pullback, isometries and musical isomorphisms .............. 8
  2.3 Causal notions in Lorentz geometry .......................... 10
  2.4 Warped product metrics ........................................ 11
  2.5 Existence of metrics ........................................... 11
  2.6 Riemannian distance function ................................. 12
  2.7 Relevance of the Euclidean and the Minkowski metrics .... 13

3 Levi-Civita connection 15
  3.1 The Levi-Civita connection .................................... 15
  3.2 Parallel translation ............................................ 18
  3.3 Geodesics ....................................................... 20
  3.4 Variational characterization of geodesics ................... 22

4 Curvature 25
  4.1 The curvature tensor ............................................ 25
  4.2 Calculating the curvature tensor .............................. 27
  4.3 The Ricci tensor and scalar curvature ......................... 27
  4.4 The divergence, the gradient and the Laplacian ............. 29
  4.5 Computing the covariant derivative of tensor fields ....... 29
    4.5.1 Divergence of a covariant 2-tensor field ................ 30
  4.6 An example of a curvature calculation ......................... 31
    4.6.1 Computing the connection coefficients .................. 32
    4.6.2 Calculating the components of the Ricci tensor ....... 33
  4.7 The 2-sphere and hyperbolic space .......................... 34
4.7.1 The Ricci curvature of the 2-sphere ................................. 35
4.7.2 The curvature of the upper half space model of hyperbolic space 36
Chapter 1

Scalar product spaces

A semi-Riemannian manifold \((M,g)\) is a manifold \(M\) with a metric \(g\). A smooth covariant 2-tensor field \(g\) is a metric if it induces a scalar product on \(T_pM\) for each \(p \in M\). Before proceeding to the subject of semi-Riemannian geometry, it is therefore necessary to define the notion of a scalar product on a vector space and to establish some of the basic properties of scalar products.

1.1 Scalar products

**Definition 1.** Let \(V\) be a finite dimensional real vector space and let \(g\) be a bilinear form on \(V\) (i.e., an element of \(L(V,V;\mathbb{R})\)). Then \(g\) is called a scalar product if the following conditions hold:

- \(g\) is symmetric; i.e., \(g(v,w) = g(w,v)\) for all \(v, w \in V\).
- \(g\) is non-degenerate; i.e., \(g(v,w) = 0\) for all \(w\) implies that \(v = 0\).

A vector space \(V\) with a scalar product \(g\) is called a scalar product space.

**Remark 2.** Since a scalar product space is a vector space \(V\) with a scalar product \(g\), it is natural to write it \((V,g)\). However, we sometimes, in the interest of brevity, simply write \(V\).

The two basic examples are the Euclidean scalar product and the Minkowski scalar product.

**Example 3.** The **Euclidean scalar product** on \(\mathbb{R}^n\), \(1 \leq n \in \mathbb{Z}\), here denoted \(g_{\text{Eucl}}\), is defined as follows. If \(v = (v^1, \ldots, v^n)\) and \(w = (w^1, \ldots, w^n)\) are two elements of \(\mathbb{R}^n\), then

\[
g_{\text{Eucl}}(v,w) = \sum_{i=1}^{n} v^i w^i.
\]

The vector space \(\mathbb{R}^n\) equipped with the Euclidean scalar product is called the \((n\text{-dimensional})\) **Euclidean scalar product space**. The **Minkowski scalar product** on \(\mathbb{R}^{n+1}\), \(1 \leq n \in \mathbb{Z}\), here denoted \(g_{\text{Min}}\), is defined as follows. If \(v = (v^0, v^1, \ldots, v^n)\) and \(w = (w^0, w^1, \ldots, w^n)\) are two elements of \(\mathbb{R}^{n+1}\), then

\[
g_{\text{Min}}(v,w) = -v^0 w^0 + \sum_{i=1}^{n} v^i w^i.
\]

The vector space \(\mathbb{R}^{n+1}\) equipped with the Minkowski scalar product is called the \((n+1\text{-dimensional})\) **Minkowski scalar product space**.

In order to distinguish between different scalar products, it is convenient to introduce the notion of an index.
CHAPTER 1. SCALAR PRODUCT SPACES

Definition 4. Let $(V, g)$ be a scalar product space. Then the index, say $\iota$, of $g$ is the largest integer that is the dimension of a subspace $W \subseteq V$ on which $g$ is negative definite.

As in the case of Euclidean geometry, it is in many contexts convenient to use particular bases, such as an orthonormal basis; in other words, a basis $\{e_i\}$ such that $g(e_i, e_j) = 0$ for $i \neq j$ and $g(e_i, e_i) = \pm 1$ (no summation on $i$).

Lemma 5. Let $(V, g)$ be a scalar product space. Then there is an integer $d \leq n := \dim V$ and a basis $\{e_i\}, i = 1, \ldots, n$, of $V$ such that

- $g(e_i, e_j) = 0$ if $i \neq j$.
- $g(e_i, e_i) = -1$ if $i \leq d$.
- $g(e_i, e_i) = 1$ if $i > d$.

Moreover, $d$ equals the index of $g$.

Proof. Let $\{v_i\}$ be a basis for $V$ and let $g_{ij} = g(v_i, v_j)$. If $G$ is the matrix with components $g_{ij}$, then $G$ is a symmetric matrix. There is thus an orthogonal matrix $T$ so that $TGT^t$ is diagonal. If $T_{ij}$ are the components of $T$, then the $ij$'th component of $TGT^t$ is given by

$$
\sum_{k,l} T_{ik} G_{kl} T_{jl} = \sum_{k,l} T_{ik} g(v_k, v_l) T_{jl} = g \left( \sum_k T_{ik} v_k, \sum_l T_{jl} v_l \right).
$$

Introducing the basis $\{w_i\}$ according to

$$
w_i = \sum_k T_{ik} v_k,
$$

it thus follows that $g(w_i, w_j) = 0$ if $i \neq j$. Due to the non-degeneracy of the scalar product, $g(w_i, w_i) \neq 0$. We can thus define a basis $\{E_i\}$ according to

$$
E_i = \frac{1}{\sqrt{|g(w_i, w_i)|}} w_i.
$$

Then $g(E_i, E_i) = \pm 1$. By renumbering the $E_i$, one obtains a basis with the properties stated in the lemma.

If $g$ is definite, the last statement of the lemma is trivial. Let us therefore assume that $0 < d < n$. Clearly, the index $\iota$ of $g$ satisfies $\iota \geq d$. In order to prove the opposite inequality, let $W$ be a subspace of $V$ such that $g$ is negative definite on $W$ and such that $\dim W = \iota$. Let $N$ be the subspace of $V$ spanned by $\{e_i\}, i = 1, \ldots, d$, and $\varphi : W \to N$ be the map defined by

$$
\varphi(w) = -\sum_{i=1}^d g(w, e_i) e_i.
$$

If $\varphi$ is injective, the desired conclusion follows. Moreover,

$$
w = -\sum_{i=1}^d g(w, e_i) e_i + \sum_{i=d+1}^n g(w, e_i) e_i; \quad (1.1)
$$

this equality is a consequence of the fact that if we take the scalar product of $e_i$ with the left hand side minus the right hand side, then the result is zero for all $i$ (so that non-degeneracy implies that (1.1) holds). If $\varphi(w) = 0$, we thus have

$$
w = \sum_{i=d+1}^n g(w, e_i) e_i.$$
Compute 
\[ g(w, w) = \sum_{i,j=1}^{d+1} g(w, e_i)g(w, e_j)g(e_i, e_j) = \sum_{i=1}^{d+1} g(w, e_i)^2 \geq 0. \]

Since \( g \) is negative definite on \( W \), this implies that \( w = 0 \). Thus \( \varphi \) is injective, and the lemma follows.

Let \( g \) and \( h \) be a scalar products on \( V \) and \( W \) respectively. A linear map \( T : V \to W \) is said to preserve scalar products if \( h(Tv_1, Tv_2) = g(v_1, v_2) \). If \( T \) preserves scalar products, then it is injective (exercise). A linear isomorphism \( T : V \to W \) that preserves scalar products is called a linear isometry.

**Lemma 6.** Scalar product spaces \( V \) and \( W \) have the same dimension and index if and only if there exists a linear isometry from \( V \) to \( W \).

**Exercise 7.** Prove Lemma 6.

### 1.2 Orthonormal bases adapted to subspaces

Two important special cases of the notion of a scalar product space are the following.

**Definition 8.** A scalar product with index 0 is called a Riemannian scalar product and a vector space with a Riemannian scalar product is called a Riemannian scalar product space. A scalar product with index 1 is called a Lorentz scalar product and a vector space with a Lorentz scalar product is called a Lorentz scalar product space.

If \( V \) is an \( n \)-dimensional Riemannian scalar product space, then there is a linear isometry from \( V \) to the \( n \)-dimensional Euclidean scalar product space. If \( V \) is an \( n + 1 \)-dimensional Lorentz scalar product space, then there is a linear isometry from \( V \) to the \( n + 1 \)-dimensional Minkowski scalar product space. Due to this fact, and the fact that the reader is assumed to be familiar with Euclidean geometry, we here focus on the Lorentz setting.

In order to understand Lorentz scalar product spaces better, it is convenient to make a few more observations of a linear algebra nature. To begin with, if \( (V, g) \) is a scalar product space and \( W \) is a subspace of \( V \), then \( W^\perp = \{ v \in V : g(v, w) = 0 \, \forall w \in W \} \).

In contrast with the Riemannian setting, \( W + W^\perp \neq V \) in general.

**Exercise 9.** Give an example of a Lorentz scalar product space \( (V, g) \) and a subspace \( W \) of \( V \) such that \( W + W^\perp \neq V \).

On the other hand, we have the following result.

**Lemma 10.** Let \( W \) be a subspace of a scalar product space \( V \). Then

1. \( \dim W + \dim W^\perp = \dim V \).
2. \( (W^\perp)^\perp = W \).

**Exercise 11.** Prove Lemma 10.

Another useful observation is the following.

**Exercise 12.** Let \( W \) be a subspace of a scalar product space \( V \). Then

\[ \dim(W + W^\perp) + \dim(W \cap W^\perp) = \dim W + \dim W^\perp. \quad (1.2) \]
Due to Lemma 10 and (1.2), it is clear that

\[ V \subset W \]

\[ \text{if and only if} \]

\[ W \subset V \]

Lemma 14. Let \( W \) be a subspace of a scalar product space \( V \). Then \( W \) is said to be non-degenerate if \( g|_W \) is non-degenerate.

We then have the following observation.

**Lemma 14.** Let \( W \) be a subspace of a scalar product space \( V \). Then \( W \) is non-degenerate if and only if \( V = W + W^\perp \).

**Proof.** Due to Lemma 10 and (1.2), it is clear that \( W + W^\perp = V \) if and only if \( W \cap W^\perp = \{0\} \). However, \( W \cap W^\perp = \{0\} \) is equivalent to \( W \) being non-degenerate. \( \square \)

One important consequence of this observation is the following.

**Corollary 15.** Let \( W_1 \) be a subspace of a scalar product space \( (V, g) \). If \( W_1 \) is non-degenerate, then \( W_2 = W_1^\perp \) is also non-degenerate. Thus \( W_i, i = 1, 2, \) are scalar product spaces with indices \( \nu_i \); the scalar product on \( W_i \) is given by \( g_i = g|_{W_i} \). If \( \nu_i \) is the index of \( V \), then \( \nu = \nu_1 + \nu_2 \). Moreover, there is an orthonormal basis \( \{e_i\}, i = 1, \ldots, n \), of \( V \) which is adapted to \( W_1 \) and \( W_2 \) in the sense that \( \{e_i\}, i = 1, \ldots, d \), is a basis for \( W_1 \) and \( \{e_i\}, i = d + 1, \ldots, n \), is a basis for \( W_2 \).

**Proof.** Since \( W_1 \) is non-degenerate and \( W_2^\perp = W_1 \) (according to Lemma 10), Lemma 14 implies that

\[ V = W_1 + W_2 = W_2^\perp + W_2. \]

Applying Lemma 14 again implies that \( W_2 \) is non-degenerate. Defining \( g_i \) as in the statement of the corollary, it is clear that \( (W_i, g_i), i = 1, 2, \) are scalar product spaces. Due to Lemma 5, we know that each of these scalar product spaces have an orthonormal basis. Let \( \{e_i\}, i = 1, \ldots, d \), be an orthonormal basis for \( W_1 \) and \( \{e_i\}, i = d + 1, \ldots, n \), be an orthonormal basis for \( W_2 \). Then \( \{e_i\}, i = 1, \ldots, n \), is an orthonormal basis of \( V \). Since \( \nu_1 \) equals the number of elements of \( \{e_i\}, i = 1, \ldots, d \), with squared norm equal to \(-1\), and similarly for \( \nu_2 \) and \( \nu \), it is clear that \( \nu = \nu_1 + \nu_2 \). \( \square \)

### 1.3 Causality for Lorentz scalar product spaces

One important notion in Lorentz scalar product spaces is that of causality, or causal character of a vector.

**Definition 16.** Let \( (V, g) \) be a Lorentz scalar product space. Then a vector \( v \in V \) is said to be

1. **timelike** if \( g(v, v) < 0 \),
2. **spacelike** if \( g(v, v) > 0 \) or \( v = 0 \),
3. **lightlike or null** if \( g(v, v) = 0 \) and \( v \neq 0 \).

The classification of a vector \( v \in V \) according to the above is called the **causal character** of the vector \( v \).

The importance of this terminology stems from its connection to the notion of causality in physics. According to special relativity, no information can travel faster than light. Assuming \( \gamma \) to be a curve in the Minkowski scalar product space (\( \gamma \) should be thought of as the trajectory of a physical object; a particle, a spacecraft, light etc.), the speed of the corresponding object relative to that of light is characterized by the causal character of \( \gamma \) with respect to the Minkowski scalar product.
If $\dot{\gamma}$ is timelike, the speed is strictly less than that of light, if $\dot{\gamma}$ is lightlike, the speed equals that of light.

In Minkowski space, if $v = (v^0, \bar{v}) \in \mathbb{R}^{n+1}$, where $\bar{v} \in \mathbb{R}^n$, then
\[ g(v, v) = -(v^0)^2 + |\bar{v}|^2, \]
where $|\bar{v}|$ denotes the usual norm of an element $\bar{v} \in \mathbb{R}^n$. Thus $v$ is timelike if $|v^0| > |\bar{v}|$, lightlike if $|v^0| = |\bar{v}| \neq 0$ and spacelike if $|v^0| < |\bar{v}|$ or $v = 0$. The set of timelike vectors consists of two components; the vectors with $v^0 > |\bar{v}|$ and the vectors with $-v^0 > |\bar{v}|$. Choosing one of these components corresponds to a choice of so-called time orientation (a choice of what is the future and what is the past). Below we justify these statements and make the notion of a time orientation more precise. However, to begin with, it is convenient to introduce some additional terminology.

**Definition 17.** Let $(V, g)$ be a scalar product space and $W \subseteq V$ be a subspace. Then $W$ is said to be **spacelike** if $g|_W$ is positive definite; i.e., if $g|_W$ is nondegenerate of index 0. Moreover, $W$ is said to be **lightlike** if $g|_W$ is degenerate. Finally, $W$ is said to be **timelike** if $g|_W$ is nondegenerate of index 1.

It is of interest to note the following consequence of Corollary 15.

**Lemma 18.** Let $(V, g)$ be a Lorentz scalar product space and $W \subseteq V$ be a subspace. Then $W$ is timelike if and only if $W^\perp$ is spacelike.

**Remark 19.** The words timelike and spacelike can be interchanged in the statement.

Let $(V, g)$ be a Lorentz scalar product space. If $u \in V$ is a timelike vector, the **timecone** of $V$ containing $u$, denoted $C(u)$, is defined by
\[ C(u) = \{ v \in V : g(v, v) < 0, \ g(v, u) < 0 \}. \]

The **opposite timecone** is defined to be $C(-u)$. Note that $C(-u) = -C(u)$. If $v \in V$ is timelike, then $v$ has to belong to $C(u)$ or $C(-u)$. The reason for this is that $(\mathbb{R}u)^\perp$ is spacelike; cf. Lemma 18. The following observation will be of importance in the discussion of the existence of Lorentz metrics.

**Lemma 20.** Let $(V, g)$ be a Lorentz scalar product space and $v, w \in V$ be timelike vectors. Then $v$ and $w$ are in the same timecone if and only if $g(v, w) < 0$.

**Proof.** Consider a timecone $C(u)$ (where we, without loss of generality, can assume that $u$ is a unit timelike vector). Due to Corollary 15, there is an orthonormal basis $\{e_\alpha\}$, $\alpha = 0, \ldots, n$, of $V$ such that $e_0 = u$. Then $v \in C(u)$ if and only if $v^0 > 0$, where $v = v^\alpha e_\alpha$. Note also that if $x = x^\alpha e_\alpha$, then $x$ is timelike if and only if $|x^0| > |\bar{x}|$, where $\bar{x} = (x^1, \ldots, x^n)$ and $|\bar{x}|$ denotes the ordinary Euclidean norm of $\bar{x} \in \mathbb{R}^n$.

Let $v$ and $w$ be timelike and define $v^\alpha$, $w^\alpha$, $\bar{v}$ and $\bar{w}$ in analogy with the above. Compute
\[ g(v, w) = -v^0 w^0 + \bar{v} \cdot \bar{w}, \quad (1.3) \]
where $\cdot$ denotes the ordinary dot product on $\mathbb{R}^n$. Since $v$ and $w$ are timelike, $|v^0| > |\bar{v}|$ and $|w^0| > |\bar{w}|$, so that
\[ |\bar{v} \cdot \bar{w}| \leq |\bar{v}| |\bar{w}| < |v^0 w^0|. \]

Thus the first term on the right hand side of (1.3) is bigger in absolute value than the second term. In particular, $g(v, w) < 0$ if and only if $v^0$ and $w^0$ have the same sign.

Assume that $v$ and $w$ are in the same timecone; say $C(u)$. Then $v^0, w^0 > 0$, so that $g(v, w) < 0$ by the above. Assume that $g(v, w) < 0$ and fix a timelike unit vector $u$. Then $v^0$ and $w^0$ have the same sign by the above. If both are positive, $v, w \in C(u)$. If both are negative, $v, w \in C(-u)$. In particular, $v, w$ are in the same timecone. The lemma follows. \qed
As a consequence of Lemma 20, timecones are convex; in fact, if $0 \leq a, b \in \mathbb{R}$ are not both zero and $v, w \in V$ are in the same timecone, then $av + bw$ is timelike and in the same timecone as $v$ and $w$. In particular, it is clear that the timelike vectors can be divided into two components. A choice of *time orientation* of a Lorentz scalar product space is a choice of timecone, say $C(u)$. A Lorentz scalar product space with a time orientation is called a *time oriented Lorentz scalar product space*. Given a choice of time orientation, the timelike vectors belonging to the corresponding timecone are said to be *future oriented*. Let $v$ be a null vector and $C(u)$ be a timecone. Then $g(v, u) \neq 0$. If $g(v, u) < 0$, then $v$ is said to be future oriented, and if $g(v, u) > 0$, then $v$ is said to be past oriented.
Chapter 2

Semi-Riemannian manifolds

The main purpose of the present chapter is to define the notion of a semi-Riemannian manifold and to describe some of the basic properties of such manifolds.

2.1 Semi-Riemannian metrics

To begin with, we need to define the notion of a metric.

**Definition 21.** Let \( M \) be a smooth manifold and \( g \) be a smooth covariant 2-tensor field on \( M \). Then \( g \) is called a **metric** on \( M \) if the following holds:

- \( g \) induces a scalar product on \( T_pM \) for each \( p \in M \).
- the index \( \iota \) of the scalar product induced on \( T_pM \) by \( g \) is independent of \( p \).

The constant index \( \iota \) is called the **index of the metric** \( g \).

**Definition 22.** A **semi-Riemannian manifold** is a smooth manifold \( M \) together with a metric \( g \) on \( M \).

Two important special cases are Riemannian and Lorentz manifolds.

**Definition 23.** Let \( (M, g) \) be a semi-Riemannian manifold. If the index of \( g \) is 0, the metric is called **Riemannian**, and \( (M, g) \) is called a **Riemannian manifold**. If the index equals 1, the metric is called a **Lorentz metric**, and \( (M, g) \) is called a **Lorentz manifold**.

Let \( (M, g) \) be a semi-Riemannian manifold. If \( (x^i) \) are local coordinates, the corresponding components of \( g \) are given by

\[
g_{ij} = g(\partial_{x^i}, \partial_{x^j}),
\]

and \( g \) can be written

\[
g = g_{ij} dx^i \otimes dx^j.
\]

Since \( g_{ij} \) are the components of a non-degenerate matrix, there is a matrix with components \( g^{ij} \) such that

\[
g^{ij} g_{jk} = \delta^i_k.
\]

Note that the functions \( g^{ij} \) are smooth, whenever they are defined. Moreover, \( g^{ij} = g^{ji} \). In fact, \( g^{ij} \) are the components of a smooth, symmetric contravariant 2-tensor field. As will become clear, this construction is of central importance in many contexts.

Again, the basic examples of metrics are the Euclidean metric and the Minkowski metric.
**Definition 24.** Let \((x^i), i = 1, \ldots, n\), be the standard coordinates on \(\mathbb{R}^n\). Then the *Euclidean metric* on \(\mathbb{R}^n\), denoted \(g_E\), is defined as follows. Let \((v^1, \ldots, v^n), (w^1, \ldots, w^n) \in \mathbb{R}^n\) and

\[
v = v^i \frac{\partial}{\partial x^i} \bigg|_p \in T_p \mathbb{R}^n, \quad w = w^i \frac{\partial}{\partial x^i} \bigg|_p \in T_p \mathbb{R}^n.
\]

Then

\[
g_E(v, w) = \sum_{i=1}^{n} v^i w^i.
\]

Let \((x^\alpha), \alpha = 0, \ldots, n\), be the standard coordinates on \(\mathbb{R}^{n+1}\). Then the *Minkowski metric* on \(\mathbb{R}^{n+1}\), denoted \(g_M\), is defined as follows. Let \((v^0, \ldots, v^n), (w^0, \ldots, w^n) \in \mathbb{R}^{n+1}\) and

\[
v = v^\alpha \frac{\partial}{\partial x^\alpha} \bigg|_p \in T_p \mathbb{R}^{n+1}, \quad w = w^\alpha \frac{\partial}{\partial x^\alpha} \bigg|_p \in T_p \mathbb{R}^{n+1}.
\]

Then

\[
g_M(v, w) = -v^0 w^0 + \sum_{i=1}^{n} v^i w^i.
\]

In Section 2.7 we discuss the relevance of these metrics.

### 2.2 Pullback, isometries and musical isomorphisms

Let \(M\) and \(N\) be smooth manifolds and \(h\) be a semi-Riemannian metric on \(N\). If \(F : M \to N\) is a smooth map, \(F^* h\) is smooth symmetric covariant 2-tensor field. However, it is not always a semi-Riemannian metric. If \(h\) is a Riemannian metric, then \(F^* h\) is a Riemannian metric if and only if \(F\) is a smooth immersion; cf. [2, Proposition 13.9, p 331]. However, if \(h\) is a Lorentz metric, \(F^* h\) need not be a Lorentz metric even if \(F\) is a smooth immersion (on the other hand, it is necessary for \(F\) to be a smooth immersion in order for \(F^* h\) to be a Lorentz metric).

**Exercise 25.** Give an example of a smooth manifold \(M\), a Lorentz manifold \((N, h)\) and a smooth immersion \(F : M \to N\) such that \(F^* h\) is not a semi-Riemannian metric on \(M\).

Due to this complication, the definition of a semi-Riemannian submanifold is slightly different from that of a Riemannian submanifold; cf. [2, p. 333].

**Definition 26.** Let \(S\) be a submanifold of a semi-Riemannian manifold \((M, g)\) with inclusion \(\iota : S \to M\). If \(\iota^* g\) is a metric on \(S\), then \(S\), equipped with this metric, is called a *semi-Riemannian submanifold* of \((M, g)\). Moreover, the metric \(\iota^* g\) is called the *induced metric* on \(S\).

Two fundamental examples are the following.

**Example 27.** Let \(S^n \subset \mathbb{R}^{n+1}\) denote the \(n\)-sphere and \(\iota_{S^n} : S^n \to \mathbb{R}^{n+1}\) the corresponding inclusion. Then the *round metric* on \(S^n\), \(g_{S^n}\), is defined by \(g_{S^n} = \iota_{S^n}^* g_{E}\); cf. Definition 24. Let \(H^n\) denote the set of \(x \in \mathbb{R}^{n+1}\) such that \(g_{\text{Min}}(x, x) = -1\) and let \(\iota_{H^n} : H^n \to \mathbb{R}^{n+1}\) denote the corresponding inclusion. Then the *hyperbolic metric* on \(H^n\), \(g_{H^n}\), is defined by \(g_{H^n} = \iota_{H^n}^* g_M\); cf. Definition 24.

**Remark 28.** Both \((S^n, g_{S^n})\) and \((H^n, g_{H^n})\) are Riemannian manifolds (we shall not demonstrate this fact in these notes; the interested reader is referred to, e.g., [1, Chapter 4] for a more detailed discussion). Note that there is a certain symmetry in the definitions: \(S^n\) is the set of \(x \in \mathbb{R}^{n+1}\) such that \(g_{\text{Eucl}}(x, x) = 1\) and \(H^n\) is the set of \(x \in \mathbb{R}^{n+1}\) such that \(g_{\text{Min}}(x, x) = -1\).
Another important example is obtained by considering a submanifold, say \( S \), of \( \mathbb{R}^n \). If \( \iota : S \rightarrow \mathbb{R}^n \) is the corresponding inclusion, then \( \iota^* g_{E} \) is a Riemannian metric on \( S \) (the Riemannian metric induced by the Euclidean metric). If \( S \) is oriented, there is also a way to define a Euclidean notion of volume of \( S \) (in specific cases, it may of course be more natural to speak of length or area). In order to justify this observation, note, first of all, that on an oriented Riemannian manifold \((M, g)\), there is a (uniquely defined) Riemannian volume form, say \( \omega_g \); cf. [2, Proposition 15.29]. The Riemannian volume of \((M, g)\) is then given by

\[
\text{Vol}(M, g) = \int_M \omega_g,
\]

assuming that this integral makes sense. If \( S \) is an oriented submanifold of \( \mathbb{R}^n \), the volume of \( S \) is then defined to be the volume of the oriented Riemannian manifold \((S, \iota^* g_{E})\).

A fundamental notion in semi-Riemannian geometry is that of an isometry.

**Definition 29.** Let \((M, g)\) and \((N, h)\) be semi-Riemannian manifolds and \( F : M \rightarrow N \) be a smooth map. Then \( F \) is called an isometry if \( F \) is a diffeomorphism such that \( F^* h = g \).

**Remark 30.** If \( F \) is an isometry, than so is \( F^{-1} \). Moreover, the composition of two isometries is an isometry. Finally, the identity map on \( M \) is an isometry. As a consequence of these observations, the set of isometries of a semi-Riemannian manifold is a group, referred to as the group of isometries.

It will be useful to keep in mind that a semi-Riemannian metric induces an isomorphism between the sections of the tangent bundle and the sections of the cotangent bundle.

**Lemma 31.** Let \((M, g)\) be a semi-Riemannian manifold. If \( X \in \mathfrak{X}(M) \), then \( X^b \) is defined to be the one-form given by

\[
X^b(Y) = g(X, Y)
\]

for all \( Y \in \mathfrak{X}(M) \). The map taking \( X \) to \( X^b \) is an isomorphism between \( \mathfrak{X}(M) \) and \( \mathfrak{X}^*(M) \). Moreover, this isomorphism is linear over the functions. In particular, given a one-form \( \eta \), there is thus a unique \( X \in \mathfrak{X}(M) \) such that \( X^b = \eta \). The vectorfield \( X \) is denoted \( \eta^X \).

**Remark 32.** Here \( \mathfrak{X}^*(M) \) denotes the smooth sections of the cotangent bundle; i.e., the one-forms.

**Proof.** Note that, given \( X \in \mathfrak{X}(M) \), it is clear that \( X^b \) is linear over \( C^\infty(M) \). Due to the tensor characterization lemma, [2, Lemma 12.24, p. 318], is thus clear that \( X^b \in \mathfrak{X}^*(M) \). In addition, it is clear that the map taking \( X \) to \( X^b \) is linear over \( C^\infty(M) \).

In order to prove injectivity of the map, assume that \( X^b = 0 \). Then \( g(X, Y) = 0 \) for all \( Y \in \mathfrak{X}(M) \). In particular, given \( p \in M \), \( g(X_p, v) = 0 \) for all \( v \in T_p M \). Due to the non-degeneracy of the metric, this implies that \( X_p = 0 \) for all \( p \in M \). Thus \( X = 0 \) and the map is injective.

In order to prove surjectivity, let \( \eta \in \mathfrak{X}^*(M) \). To begin with, let us try to find a vectorfield \( X \) on a coordinate neighbourhood \( U \) such that \( X^b = \eta \) on \( U \). If \( \eta_i \) are the components of \( \eta \) with respect to local coordinates, then we can define a vectorfield on \( U \) by

\[
X = g^{ij} \eta_j \frac{\partial}{\partial x^i}.
\]

In this expression, \( g^{ij} \) are the components of the inverse of the matrix with components \( g_{ij} \). Then \( X \) is a smooth vectorfield on \( U \). Moreover,

\[
X^b(Y) = g(X, Y) = g_{ik} X^i Y^k = g_{ik} g^{ij} \eta_j Y^k = \delta^i_k \eta_j Y^k = \eta_j Y^j = \eta(Y).
\]

Thus \( X^b = \eta \). Due to the uniqueness, the local vectorfields can be combined to give an \( X \in \mathfrak{X}(M) \) such that \( X^b = \eta \). This proves surjectivity. \( \square \)
Let \((M, g)\) be a Lorentz manifold. A \textit{time orientation} of \((M, g)\) is a choice of time orientation of each scalar product space \((T_pM, g_p)\), \(p \in M\), such that the following holds. For each \(p \in M\), there is an open neighbourhood \(U\) of \(p\) and a smooth vector field \(X\) on \(U\) such that \(X_p\) is future oriented for all \(q \in U\). A Lorentz metric \(g\) on a manifold \(M\) is said to be \textit{time orientable} if \((M, g)\) has a time orientation. A Lorentz manifold \((M, g)\) is said to be \textit{time orientable} if \((M, g)\) has a time orientation. A Lorentz manifold with a time orientation is called a \textit{time oriented Lorentz manifold}.

**Remark 35.** Here \(g_p\) denotes the scalar product induced on \(T_pM\) by \(g\). The requirement that there be a local vector field with the properties stated in the definition is there to ensure the “continuity” of the choice of time orientation.

A choice of time orientation for a Lorentz manifold corresponds to a choice of which time direction corresponds to the future and which time direction corresponds to the past. In physics, time oriented Lorentz manifolds are of greater interest than non-time oriented ones. For this reason, the following terminology is sometimes introduced.

**Definition 36.** A time oriented Lorentz manifold is called a \textit{spacetime}.

Let us now introduce some of the notions of causality that we shall use.

**Definition 37.** Let \((M, g)\) be a Lorentz manifold. A vector \(v \in T_pM\) is said to be \textit{timelike}, \textit{spacelike} or \textit{lightlike} if it is timelike, spacelike or lightlike, respectively, with respect to the scalar product \(g_p\) induced on \(T_pM\) by \(g\). A vector field \(X\) on \(M\) is said to be \textit{timelike}, \textit{spacelike} or \textit{lightlike} if \(X_p\) is timelike, spacelike or lightlike, respectively, for all \(p \in M\). A smooth curve \(\gamma : I \to M\) (where \(I\) is an open interval) is said to be \textit{timelike}, \textit{spacelike} or \textit{lightlike} if \(\gamma(t)\) is timelike, spacelike
or lightlike, respectively, for all $t \in I$. A submanifold $S$ of $M$ is said to be spacelike if $S$ is a semi-Riemannian submanifold of $M$ such that the induced metric is Riemannian. A tangent vector which is either timelike or lightlike is said to be causal. The terminology concerning vectorfields and curves is analogous.

In case $(M, g)$ is a spacetime, it is also possible to speak of future directed timelike vectors etc. Note, however, that a causal curve is said to be future directed if and only if $\dot{\gamma}(t)$ is future oriented for all $t$ in the domain of definition of $\gamma$. Our requirements concerning vector fields is similar.

### 2.4 Warped product metrics

One construction which is very important in the context of general relativity is that of a so-called warped product metric.

**Definition 38.** Let $(M_i, g_i), i = 1, 2,$ be semi-Riemannian manifolds, $\pi_i : M_1 \times M_2 \to M_i$ be the projection taking $(p_1, p_2)$ to $p_i$, and $f \in C^\infty(M_1)$ be strictly positive. Then the warped product, denoted $M = M_1 \times_f M_2$, is the manifold $M = M_1 \times M_2$ with the metric

$$g = \pi_1^*g_1 + (f \circ \pi_1)^2 \pi_2^*g_2.$$  

**Exercise 39.** Prove that the warped product is a semi-Riemannian manifold.

One special case of this construction is obtained by demanding that $f = 1$. In that case, the resulting warped product is referred to as a semi-Riemannian product manifold. One basic example of a warped product is the following.

**Example 40.** Let $M_1 = I$ (where $I$ is an open interval), $M_2 = \mathbb{R}^3$, $g_1 = -dt \otimes dt$, $g_2 = g_E$ (the Euclidean metric on $\mathbb{R}^3$) and $f$ be a smooth strictly positive function on $I$. Then the resulting warped product is the manifold $M = I \times \mathbb{R}^3$ with the metric

$$g = -dt \otimes dt + f^2(t) \sum_{i=1}^3 dx^i \otimes dx^i,$$

where $t$ is the coordinate on the first factor in $I \times \mathbb{R}^3$ and $x^i, i = 1, 2, 3$, are the coordinates on the last three factors. The geometry of most models of the universe used by physicists today are of the type $(M, g)$. What varies from model to model is the function $f$.

### 2.5 Existence of metrics

In semi-Riemannian geometry, a fundamental question to ask is: given a manifold, is there a semi-Riemannian metric on it? In the Riemannian setting, this question has a simple answer.

**Proposition 41.** Every smooth manifold with or without boundary admits a Riemannian metric.

**Proof.** The proof can be found in [2, p. 329].

In the Lorentzian setting, the situation is more complicated.

**Proposition 42.** A manifold $M$, $n := \dim M \geq 2$, admits a time orientable Lorentz metric if and only if there is an $X \in \mathfrak{X}(M)$ such that $X_p \neq 0$ for all $p \in M$. 

Proof. Assume that there is a nowhere vanishing smooth vector field $X$ on $M$. Let $h$ be a Riemannian metric on $M$ (such a metric exists due to Proposition 41). By normalizing $X$ if necessary, we can assume $h(X,X) = 1$. Define $g$ according to
\[
g = -2X^\flat \otimes X^\flat + h,
\]
where $X^\flat$ is defined in Lemma 31. Then
\[
g(X,X) = -2[X^\flat(X)]^2 + 1 = -1.
\]
Given $p \in M$, let $e_2|_p, \ldots, e_n|_p \in T_pM$ be such that $e_1|_p, \ldots, e_n|_p$ is an orthonormal basis for $(T_pM, h_p)$, where $e_1|_p = X^p$. Then $\{e_i|_p\}$ is an orthonormal basis for $(T_pM, g_p)$. Moreover, it is clear that the index of $g_p$ is 1. Thus $(M, g)$ is a Lorentz manifold. Since we can define a time orientation by requiring $X^p$ to be future oriented for all $p \in M$, it is clear that $M$ admits a time orientable Lorentz metric.

Assume now that $M$ admits a time orientable Lorentz metric $g$. Fix a time orientation. Let $\{U_\alpha\}$ be an open covering of $M$ such that on each $U_\alpha$ there is a timelike vector field $X_\alpha$ which is future pointing (that such a covering exists is a consequence of the definition of a time orientation; cf. Definition 34). Let $\{\phi_\alpha\}$ be a partition of unity subordinate to the covering $\{U_\alpha\}$. Define $X$ by
\[
X = \sum_\alpha \phi_\alpha X_\alpha.
\]
Fix a $p \in M$. At this point, the sum consists of finitely many terms, so that
\[
X^p = \sum_{i=1}^k a_i X_{i,p},
\]
where $0 < a_i \in \mathbb{R}$ and $X_{i,p} \in T_pM$ are future oriented timelike vectors. Due to Lemma 20 it is then clear that $X^p$ is a future oriented timelike vector. In particular, $X$ is thus a future oriented timelike vector field. Since such a vector field is nowhere vanishing, it is clear that $M$ admits a non-zero vector field.

It is of course natural to ask what happens if we drop the condition that $(M, g)$ be time orientable. However, in that case there is a double cover which is time orientable (for those unfamiliar with covering spaces, we shall not make any use of this fact). It is important to note that the existence of a Lorentz metric is a topological restriction: not all manifolds admit Lorentz metrics. As an orientation in the subject of Lorentz geometry, it is also of interest to make the following remark (we shall not make any use of the statements made in the remark in what follows).

**Remark 43.** If $(M, g)$ is a spacetime such that $M$ is a closed manifold (in other words, $M$ is compact and without boundary), then there is a closed timelike curve in $M$. In other words, there is a future oriented timelike curve $\gamma$ in $M$ such that $\gamma(t_1) = \gamma(t_2)$ for some $t_1 < t_2$ in the domain of definition of $\gamma$. This means that it is possible to travel into the past. Since this is not very natural in physics, spacetimes $(M, g)$ such that $M$ is closed are not very natural (in contrast with the Riemannian setting). For a proof of this statement, see [1, Lemma 10, p. 407]. In general relativity, one often requires spacetimes to satisfy an additional requirement called global hyperbolicity (which we shall not define here) which involves additional conditions concerning causality. Moreover, globally hyperbolic spacetimes $(M, g)$, where $n + 1 = \dim M$, are topologically products $M = \mathbb{R} \times \Sigma$ where $\Sigma$ is an $n$-dimensional manifold.

### 2.6 Riemannian distance function

Let $(M, g)$ be a Riemannian manifold. Then it is possible to associate a distance function
\[
d : M \times M \to [0, \infty)
\]
2.7 RELEVANCE OF THE EUCLIDEAN AND THE MINKOWSKI METRICS

with $(M,g)$. Since the basic properties of the Riemannian distance function are described in [2, pp. 337–341], we shall not do so here.

2.7 Relevance of the Euclidean and the Minkowski metrics

It is of interest to make some comments concerning the relevance of the Euclidean and the Minkowski metrics. The Euclidean metric gives rise to Euclidean geometry, and the relevance of this geometry is apparent in much of mathematics. For that reason, we here focus on the Minkowski metric.

Turning to Minkowski space, it is of interest to recall the origin of the special theory of relativity (for those uninterested in physics, the remainder of this section can be skipped). In special relativity, there are frames of reference (in practice, coordinate systems) which are preferred, the so-called inertial frames. These frames should be thought of as the “non-accelerated” frames, and two inertial frames travel at “constant velocity” relative to each other. It is of interest to relate measurements made with respect to different inertial frames. Let us consider the classical and the special relativistic perspective separately.

The classical perspective. In the classical perspective, the transformation laws are obtained by demanding that time is absolute. The relation between two inertial frames is then specified by fixing the relative velocity, an initial translation, and a rotation. More specifically, given two inertial frames $F$ and $F'$, there are $t_0 \in \mathbb{R}$, $v, x_0 \in \mathbb{R}^3$ and $A \in SO(3)$ such that if $(t,x) \in \mathbb{R}^4$ are the time and space coordinates of an event with respect to the inertial frame $F$ and $(t',x') \in \mathbb{R}^4$ are the coordinates of the same event with respect to the inertial frame $F'$, then

$$t' = t + t_0,$$  \hfill (2.1)

$$x' = Ax + vt + x_0.$$  \hfill (2.2)

The special relativistic perspective. The classical laws of physics transform well under changes of coordinates of the form (2.1)–(2.2). However, it turns out that Maxwell’s equations do not. This led Einstein to use a different starting point, namely that the speed of light is the same in all inertial frames. One consequence of this assumption is that time is no longer absolute. Moreover, if one wishes to compute the associated changes of coordinates when going from one inertial frame to another, they are different from (2.1)–(2.2). The group of transformations (taking the coordinates of one inertial frame to the coordinates of another frame) that arise when taking this perspective is called the group of Lorentz transformations. The main point of introducing Minkowski space is that the group of isometries of Minkowski space are exactly the group of Lorentz transformations.
Chapter 3

The Levi-Civita connection, parallel translation and geodesics

Einstein’s equations of general relativity relate the curvature of a spacetime with the matter content of the spacetime. In order to understand this equation, it is therefore important to understand the notion of curvature. This subject has a rich history, and here we only give a quite formal and brief introduction to it. One way to define curvature is to examine how a vector is changed when parallel translating it along a closed curve in the manifold. In order for this to make sense, it is of course necessary to assign a meaning to the notion of “parallel translation”. In the case of Euclidean space, the notion is perhaps intuitively clear; we simply fix the components of the vector with respect to the standard coordinate frame and then change the base point. Transporting a vector in Euclidean space (along a closed curve) in this way yields the identity map; one returns to the vector one started with. This is one way to express that the curvature of Euclidean space vanishes. Using a rather intuitive notion of parallel translation on the 2-sphere, one can convince oneself that the same is not true of the 2-sphere.

In order to proceed to a formal development of the subject, it is necessary to clarify what is meant by parallel translation. One natural way to proceed is to define an “infinitesimal” version of this notion. This leads to the definition of a so-called connection.

3.1 The Levi-Civita connection

In the end we wish to define the notion of a Levi-Civita connection, but we begin by defining what a connection is in general.

**Definition 44.** Let $M$ be a smooth manifold. A map $\nabla : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ is called a connection if

- $\nabla_X Y$ is linear over $C^\infty(M)$ in $X$,
- $\nabla_X Y$ is linear over $\mathbb{R}$ in $X$,
- $\nabla_X (fY) = X(f)Y + f\nabla_X Y$ for all $X, Y \in \mathfrak{X}(M)$ and all $f \in C^\infty(M)$.

The expression $\nabla_X Y$ is referred to as the covariant derivative of $Y$ with respect to $X$ for the connection $\nabla$.

Note that the first condition of Definition 44 means that

$$\nabla_{f_1X_1 + f_2X_2} Y = f_1 \nabla_{X_1} Y + f_2 \nabla_{X_2} Y$$
for all \( f_i \in C^\infty(M), X_i, Y \in \mathfrak{X}(M), i = 1, 2 \). The second condition of Definition 44 means that
\[
\nabla_X (a_1 Y_1 + a_2 Y_2) = a_1 \nabla_X Y_1 + a_2 \nabla_X Y_2
\]
for all \( a_i \in \mathbb{R}, X, Y_i \in \mathfrak{X}(M), i = 1, 2 \). On \( \mathbb{R}^n \) there is a natural connection.

**Definition 45.** Let \((x^i), i = 1, \ldots, n\), be the standard coordinates on \( \mathbb{R}^n \). Let \( X, Y \in \mathfrak{X}(\mathbb{R}^n) \) and define
\[
\nabla_X Y = X(Y^i) \frac{\partial}{\partial x^i},
\]
where
\[
Y = Y^i \frac{\partial}{\partial x^i}.
\]
Then \( \nabla \) is referred to as the *standard connection* on \( \mathbb{R}^n \).

**Exercise 46.** Prove that the standard connection on \( \mathbb{R}^n \) is a connection in the sense of Definition 44.

Let \((M, g)\) be a semi-Riemannian manifold. Our next goal is to argue that there is preferred connection, given the metric \( g \). However, in order to single out a preferred connection, we have to impose additional conditions. One such condition would be to require that
\[
\langle \nabla_X Y, Z \rangle = g(\nabla_X Y, Z) + g(Y, \nabla_X Z)
\]
for all \( X, Y, Z \in \mathfrak{X}(M) \). In what follows, it is going to be a bit cumbersome to use the notation \( g(X, Y) \). We therefore define \( \langle \cdot, \cdot \rangle \) by
\[
\langle X, Y \rangle = g(X, Y);
\]
we shall use this notation both for vectorfields and for individual vectors. With this notation, (3.1) can be written
\[
X \langle Y, Z \rangle = \langle \nabla_X Y, Z \rangle + \langle Y, \nabla_X Z \rangle.
\]
A connection satisfying this requirement is said to be *metric*. However, it turns out that the condition (3.2) does not determine a unique connection. In fact, we are free to add further conditions. One such condition would be to impose that \( \nabla_X Y - \nabla_Y X \) can be expressed in terms of only \( X \) and \( Y \), without any reference to the connection. Since \( \nabla_X Y - \nabla_Y X \) is antisymmetric, one such condition would be to require that
\[
\nabla_X Y - \nabla_Y X = [X, Y]
\]
for all \( X, Y \in \mathfrak{X}(M) \). A connection satisfying this criterion is said to be *torsion free*. Remarkably, it turns out that conditions (3.2) and (3.3) uniquely determine a connection, referred to as the Levi-Civita connection.

**Theorem 47.** Let \((M, g)\) be a semi-Riemannian manifold. Then there is a unique connection \( \nabla \) satisfying (3.2) and (3.3) for all \( X, Y, Z \in \mathfrak{X}(M) \). It is called the Levi-Civita connection of \((M, g)\). Moreover, it is characterized by the Koszul formula:
\[
2 \langle \nabla_X Y, Z \rangle = \langle X, Y \rangle + \langle Y, Z \rangle + \langle Z, [X, Y] \rangle + \langle Y, [Z, X] \rangle + \langle Z, [X, Y] \rangle.
\]

**Proof.** Assume that \( \nabla \) is a connection satisfying (3.2) and (3.3) for all \( X, Y, Z \in \mathfrak{X}(M) \). Compute
\[
\langle \nabla_X Y, Z \rangle = \langle X, Y \rangle + \langle Y, \nabla_X Z \rangle = \langle X, Y \rangle + \langle Y, [X, Z] \rangle + \langle Y, \nabla_X Z \rangle
\]
\[
= \langle X, Y \rangle + \langle [Y, X], Z \rangle + \langle Y, \nabla_X Z \rangle
\]
\[
= \langle X, Y \rangle + \langle [Y, X], Z \rangle - \langle Z, [Y, X] \rangle + \langle Y, [Z, X] \rangle + \langle Y, \nabla_X Z \rangle + \langle Y, Z, X \rangle
\]
\[
= \langle X, Y \rangle - \langle Z, [Y, X] \rangle + \langle Y, [Z, X] \rangle + \langle Y, \nabla_X Z \rangle + \langle Y, Z, X \rangle
\]
\[
= \langle X, Y \rangle + \langle Y, Z, X \rangle - \langle Y, [X, Z] \rangle + \langle Y, [Z, X] \rangle + \langle Y, \nabla_X Z \rangle + \langle Y, Z, X \rangle - \langle Z, \nabla_Y X \rangle
\]
\[
= \langle X, Y \rangle + \langle Y, [Z, X] \rangle + \langle Y, \nabla_X Z \rangle + \langle Y, Z, X \rangle - \langle Z, \nabla_Y X \rangle.
\]
3.1. THE LEVI-CIVITA CONNECTION

where we have applied (3.2) and (3.3). In the fourth and fifth steps, we also rearranged the terms and used the antisymmetry of the Lie bracket. Note that this equation implies that (3.4) holds. This leads to the uniqueness of the Levi-Civita connection. The reason for this is the following. Assume that \( \nabla \) and \( \hat{\nabla} \) both satisfy (3.2) and (3.3). Then, due to the Koszul formula,

\[
\langle \nabla_X Y - \hat{\nabla}_X Y, Z \rangle = 0
\]

for all \( X, Y, Z \in \mathfrak{X}(M) \). Due to Lemma 31, this implies that

\[
\nabla_X Y = \hat{\nabla}_X Y
\]

for all \( X, Y \in \mathfrak{X}(M) \). In other words, there is at most one connection satisfying the conditions (3.2) and (3.3). Given \( X, Y \in \mathfrak{X}(M) \), let \( \theta_{X,Y} \) be defined by the condition that \( 2\theta_{X,Y}(Z) \) is given by the right hand side of (3.4). It can then be demonstrated that \( \theta_{X,Y} \) is linear over \( C^\infty(M) \); in other words,

\[
\theta_{X,Y}(f_1 Z_1 + f_2 Z_2) = f_1 \theta_{X,Y}(Z_1) + f_2 \theta_{X,Y}(Z_2)
\]

for all \( X, Y, Z_i \in \mathfrak{X}(M) \), \( f_i \in C^\infty(M) \), \( i = 1, 2 \) (we leave it as an exercise to verify that this is true). Due to the tensor characterization lemma, [2, Lemma 12.24, p. 318], it thus follows that \( \theta_{X,Y} \) is a one-form. By appealing to Lemma 31, we conclude that there is a smooth vectorfield \( \theta^k_{X,Y} \) such that

\[
\langle \theta^k_{X,Y}, Z \rangle = \theta_{X,Y}(Z),
\]

where the right hand side is given by the right hand side of (3.4). We define \( \nabla_X Y \) by

\[
\nabla_X Y = \theta^k_{X,Y}.
\]

Then \( \nabla \) is a function from \( \mathfrak{X}(M) \times \mathfrak{X}(M) \) to \( \mathfrak{X}(M) \). However, it is not obvious that it satisfies the conditions of Definition 44. Moreover, it is not obvious that it satisfies (3.2) and (3.3). In other words, there are five conditions we need to verify. Let us verify the first condition in the definition of a connection. Note, to this end, that \( 2(\nabla_{fX} Y, Z) \) is given by the right hand side of (3.4), with \( X \) replaced by \( fX \). However, a straightforward calculation shows that if you replace \( X \) by \( fX \) in (3.4), then you obtain \( f \) times the right hand side of (3.4). In other words,

\[
\langle \nabla_{fX} Y, Z \rangle = f \langle \nabla_X Y, Z \rangle.
\]

Due to Lemma 31, it follows that \( \nabla_{fX} = f \nabla_X Y \). We leave it as an exercise to prove that \( \nabla_X Y \) is linear in \( X \) and \( Y \) over \( \mathbb{R} \), and conclude that the first two conditions of Definition 44 are satisfied. To prove that the third condition holds, compute

\[
2(\nabla_X(fY), Z) = X(fY, Z) + fY(X, Z) - Z(X, fY) - \langle X, [fY, Z] \rangle + \langle fY, [Z, X] \rangle + \langle Z, [X, fY] \rangle
\]

\[
= X(f)(Y, Z) + fX(Y, Z) + fY(Z, X) - fZ(X, Y)
\]

\[
+ Z(f)X(Y, Z) + fY(Z, X) + X(f)(Z, Y) + f\langle Z, [X, Y] \rangle
\]

\[
= 2f(\nabla_X Y, Z) + 2X(f)(Y, Z) = 2(f \nabla_X Y + X(f)Y, Z).
\]

Appealing to Lemma 31 yields

\[
\nabla_X(fY) = f \nabla_X Y + X(f)Y.
\]

Thus the third condition of Definition 44 is satisfied. We leave it to the reader to verify that (3.2) and (3.3) are satisfied. \( \square \)

**Exercise 48.** Prove that the connection \( \nabla \) constructed in the proof of Theorem 47 satisfies the conditions (3.2) and (3.3) for all \( X, Y, Z \in \mathfrak{X}(M) \).
Let \((M, g)\) be a semi-Riemannian manifold and \(\nabla\) be the associated Levi-Civita connection. It is of interest to express \(\nabla\) with respect to local coordinates \((x^i)\). Introduce, to this end, the notation \(\Gamma^k_{ij}\) by
\[
\nabla_{\partial_i} \partial_j = \Gamma^k_{ij} \partial_k,
\]
where we use the short hand notation
\[
\partial_i = \frac{\partial}{\partial x^i}.
\]
The smooth functions \(\Gamma^k_{ij}\), defined on the domain of the coordinates, are called the Christoffel symbols. Using the notation \(g_{ij} = g(\partial_i, \partial_j)\), let us compute
\[
g_{kl} \Gamma^k_{ij} = \langle \partial_i, \Gamma^k_{ij} \partial_k \rangle = \langle \nabla_{\partial_i} \partial_k, \partial_l \rangle = \frac{1}{2} \left( \partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij} \right),
\]
where we have used the Koszul formula, (3.4), in the last step. Multiplying this equality with \(g^{ml}\) and summing over \(l\) yields
\[
\Gamma^m_{ij} = \frac{1}{2} g^{ml} (\partial_l g_{jl} + \partial_j g_{il} - \partial_l g_{ij}).
\]  \hspace{1cm} (3.5)

Note that \(\Gamma^m_{ij} = \Gamma^m_{ji}\). If \(X = X^i \partial_i\) and \(Y = Y^i \partial_i\) are vectorfields, we obtain
\[
\nabla_X Y = \nabla_X (Y^j \partial_j) = X(Y^j) \partial_j + Y^j \nabla_X \partial_j = X(Y^j) \partial_j + Y^j X^i \nabla_{\partial_i} \partial_j
\]
\[
= X(Y^j) \partial_j + Y^j X^i \Gamma^k_{ij} \partial_k.
\]
Thus
\[
\nabla_X Y = \left[ X(Y^k) + \Gamma^k_{ij} X^i Y^j \right] \partial_k.
\]
When defining parallel transport, we shall use the following consequence of this formula.

**Lemma 49.** Let \((M, g)\) be a semi-Riemannian manifold and \(\nabla\) be the associated Levi-Civita connection. Let \(v \in T_p M\) for some \(p \in M\), and \(Y \in \mathcal{X}(M)\). Let \(X_i \in \mathcal{X}(M), i = 1, 2,\) be such that \(X_i(p) = v\). Then
\[
\langle \nabla_{X_i} Y \rangle_p = \langle \nabla_{X_2} Y \rangle_p.
\]

This lemma justifies defining \(\nabla_v Y\) in the following way.

**Definition 50.** Let \((M, g)\) be a semi-Riemannian manifold and \(\nabla\) be the associated Levi-Civita connection. Let \(v \in T_p M\) for some \(p \in M\), and \(Y \in \mathcal{X}(M)\). Given any vectorfield \(X \in \mathcal{X}(M)\) such that \(X_p = v\), define \(\nabla_v Y\) by
\[
\nabla_v Y = \langle \nabla_X Y \rangle_p.
\]

### 3.2 Parallel translation

At the beginning of the present chapter, we justified the introduction of the notion of a connection by the (vague) statement that it would constitute an “infinitesimal” version of a notion of parallel translation. In the present section, we wish to justify this statement by using the Levi-Civita connection to define parallel translation. To begin with, let us introduce some terminology.

Let \((M, g)\) be a semi-Riemannian manifold, \(I \subseteq \mathbb{R}\) be an open interval and \(\gamma : I \to M\) be a smooth curve. Then a smooth map from \(I\) to the tangent bundle of \(M\), say \(X\), is said to be an element of \(\mathcal{X}(\gamma)\) if the base point of \(X(t)\) is \(\gamma(t)\). If \(X \in \mathcal{X}(M)\), we let \(X_\gamma\) denote the element of \(\mathcal{X}(\gamma)\) which assigns the vector \(X_{\gamma(t)}\) to the number \(t \in I\).

**Proposition 51.** Let \((M, g)\) be a semi-Riemannian manifold, \(I \subseteq \mathbb{R}\) be an open interval and \(\gamma : I \to M\) be a smooth curve. Then there is a unique function taking \(X \in \mathcal{X}(\gamma)\) to
\[
X' = \frac{\nabla X}{dt} \in \mathcal{X}(\gamma),
\]
satisfying the following properties:

\[
\begin{align*}
(a_1X_1 + a_2X_2)' &= a_1X_1' + a_2X_2', \\
(fX)' &= f'X + fX', \\
(Y_\gamma)' &= \nabla_{\gamma'}Y,
\end{align*}
\]

for all \(X, X_i \in \mathfrak{X}(\gamma)\), \(a_i \in \mathbb{R}\), \(f \in C^\infty(I)\) and \(Y \in \mathfrak{X}(M)\), \(i = 1, 2\). Moreover, this map has the property that

\[
\frac{d}{dt}(X_1, X_2) = (X_1', X_2) + (X_1, X_2').
\]

**Remark 52.** How to interpret the expression \(\nabla_{\gamma'}Y\) appearing in (3.8) is explained in Lemma 49 and Definition 50.

**Proof.** Let us begin by proving uniqueness. Let \(X \in \mathfrak{X}(\gamma)\). Then we can write \(X\) as

\[
X(t) = X^i(t)\partial_i|_{\gamma(t)},
\]

where the \(X^i\) are smooth functions on \(X^{-1}(U)\) (where \(U\) is the set on which the coordinates \((x^i)\) are defined). Assume now that we have derivative operator satisfying (3.6)–(3.8). Applying (3.6)–(3.8) to (3.10) yields

\[
X'(t) = \frac{dX^i}{dt}(t)\partial_i|_{\gamma(t)} + X^i(t)(\partial_i|_{\gamma(t)})'(t) = \frac{dX^i}{dt}(t)\partial_i|_{\gamma(t)} + X^i(t)\nabla_{\gamma'(t)}\partial_i.
\]

Since the right hand side only depends on the Levi-Civita connection, we conclude that uniqueness holds.

In order to prove existence, we can define \(X'(t)\) by (3.11) for \(t \in X^{-1}(U)\). It can then be verified that the corresponding derivative operator satisfies the conditions (3.6)–(3.9); we leave this as an exercise. Due to uniqueness, these coordinate representations can be patched together to produce an element \(X' \in \mathfrak{X}(\gamma)\).

**Exercise 53.** Prove that the derivative operator defined by the formula (3.11) has the properties (3.6)–(3.9).

It is of interest to write down a formula for \(X'\) in local coordinates. Let \((x^i)\) be local coordinates, \(\gamma' = x^i \circ \gamma\) and \(X^i\) be defined by (3.10). Then, since

\[
\gamma'(t) = \frac{dx^i}{dt}(t)\partial_i|_{\gamma(t)},
\]

(3.11) implies

\[
X'(t) = \frac{dX^i}{dt}(t)\partial_i|_{\gamma(t)} + X^i(t)\frac{dx^j}{dt}(t)\Gamma^k_{ji}[\gamma(t)]\partial_k|_{\gamma(t)} = \left(\frac{dX^k}{dt}(t) + X^i(t)\frac{dx^j}{dt}(t)\Gamma^k_{ji}[\gamma(t)]\right)\partial_k|_{\gamma(t)}.
\]

Given the derivative operator of Proposition 51, we are now in a position to assign a meaning to the expression parallel translation used in the introduction to the present chapter.

**Definition 54.** Let \((M, g)\) be a semi-Riemannian manifold, \(I \subseteq \mathbb{R}\) be an open interval and \(\gamma : I \to M\) be a smooth curve. Then \(X \in \mathfrak{X}(\gamma)\) is said to be parallel along \(\gamma\) if and only if \(X' = 0\). Note that, in local coordinates, the equation \(X' = 0\) is a linear equation for the components of \(X\); cf. (3.12). For this reason, we have the following proposition (cf. also the arguments used to prove the existence of integral curves of vectorfields).
Proposition 55. Let \((M, g)\) be a semi-Riemannian manifold, \(I \subseteq \mathbb{R}\) be an open interval and \(\gamma : I \to M\) be a smooth curve. If \(t_0 \in I\) and \(\xi \in T_{\gamma(t_0)}M\), then there is a unique \(X \in \mathfrak{X}(\gamma)\) such that \(X' = 0\) and \(X(t_0) = \xi\).

Exercise 56. Prove Proposition 55.

Due to Proposition 55 we are in a position to define parallel translation along a curve. Given assumptions as in the statement of Proposition 55, let \(t_0, t_1 \in I\). Then there is a map

\[ P : T_{\gamma(t_0)} \to T_{\gamma(t_1)} \]

defined as follows. Given \(\xi \in T_{\gamma(t_0)}\), let \(X \in \mathfrak{X}(\gamma)\) be such that \(X' = 0\) and \(X(t_0) = \xi\). Define \(P(\xi) = X(t_1)\). Here \(P\) depends (only) on \(\gamma\), \(t_0\) and \(t_1\). In some situations, it may be useful to indicate this dependence, but if these objects are clear from the context, it is convenient to simply write \(P\). The map \(P\) is called parallel translation along \(\gamma\) from \(\gamma(t_0)\) to \(\gamma(t_1)\). Parallel translation has the following property.

Proposition 57. Let \((M, g)\) be a semi-Riemannian manifold, \(I \subseteq \mathbb{R}\) be an open interval and \(\gamma : I \to M\) be a smooth curve. Finally, let \(t_0, t_1 \in I\) and \(p_i = \gamma(t_i), i = 0, 1\). Then parallel translation along \(\gamma\) from \(\gamma(t_0)\) to \(\gamma(t_1)\) is a linear isometry from \(T_{p_0}M\) to \(T_{p_1}M\).

Proof. We leave it to the reader to prove that parallel translation is a vector space isomorphism. In order to prove that it is an isometry, let \(v, w \in T_{p_0}M\) and \(V, W \in \mathfrak{X}(\gamma)\) be such that \(V' = W' = 0\), \(V(t_0) = v\) and \(W(t_0) = w\). Then \(P(v) = V(t_1)\) and \(P(w) = W(t_1)\). Compute

\[
\langle P(v), P(w) \rangle = \langle V(t_1), W(t_1) \rangle = \langle V(t_0), W(t_0) \rangle + \int_{t_0}^{t_1} \frac{d}{dt} \langle V, W \rangle dt \\
= \langle v, w \rangle + \int_{t_0}^{t_1} \left( \langle V', W \rangle + \langle V, W' \rangle \right) dt = \langle v, w \rangle,
\]

where we have used property (3.9) of the derivative operator \(\cdot\), as well as the fact that \(V' = W' = 0\). The proposition follows. \(\square\)

Exercise 58. Prove that parallel translation is a vector space isomorphism.

Let us analyze what parallel translation means in the case of Euclidean space and Minkowski space. Let \(I\) and \(\gamma\) be as in the statement of Proposition 55, where \((M, g)\) is either Euclidean space or Minkowski space. Note that the Christoffel symbols of \(g_E\) and \(g_M\) vanish with respect to standard coordinates on \(\mathbb{R}^n\) and \(\mathbb{R}^{n+1}\) respectively. An element \(X \in \mathfrak{X}(\gamma)\) is therefore parallel if and only if the components of \(X\) with respect to the standard coordinate vectorfields are constant (just as we stated in the introduction). In particular, the result of the parallel translation does not depend on the curve. It is of importance to note that, even though this is true in the case of Euclidean space and Minkowski space, it is not true in general.

3.3 Geodesics

A notion which is extremely important both in Riemannian geometry and in Lorentz geometry is that of a geodesic. In Riemannian geometry, geodesics are locally length minimizing curves. In the case of general relativity (Lorentz geometry), geodesics are related to the trajectories of freely falling test particles, as well as the trajectories of light.

Definition 59. Let \((M, g)\) be a semi-Riemannian manifold, \(I \subseteq \mathbb{R}\) be an open interval and \(\gamma : I \to M\) be a smooth curve. Then \(\gamma\) is said to be a geodesic if \(\gamma'' = 0\); in other words, if \(\gamma' \in \mathfrak{X}(\gamma)\) is parallel.
3.3. GEODESICS

Keeping (3.12) in mind, geodesics are curves which with respect to local coordinates \((x^i)\) satisfy the equation

\[
\ddot{\gamma}^k + \left( \Gamma^k_{ij} \circ \gamma \right) \dot{\gamma}^i \dot{\gamma}^j = 0,
\]

where we use the notation

\[
\gamma^i = x^i \circ \gamma, \quad \dot{\gamma}^i = \frac{d\gamma^i}{dt}, \quad \ddot{\gamma}^i = \frac{d^2\gamma^i}{dt^2}.
\]

It is important to note that, even though (3.13) is an ODE, it is (in contrast to the equation \(X' = 0\) for a fixed curve \(\gamma\)) a non-linear ODE. Due to the fact that (3.13) is an autonomous ODE for \(\gamma\) and the fact that the Christoffel symbols are smooth functions, it is clear that geodesics are smooth curves. Due to local existence and uniqueness results for ODE’s, we have the following proposition.

**Proposition 60.** Let \((M, g)\) be a semi-Riemannian manifold, \(p \in M\) and \(v \in T_p M\). Then there is a unique geodesic \(\gamma : I \to M\) with the properties that

- \(I \subseteq \mathbb{R}\) is an open interval such that \(0 \in I\),
- \(\gamma'(0) = v\),
- \(I\) is maximal in the sense that if \(\alpha : J \to M\) is a geodesic (with \(J\) an open interval, \(0 \in J\) and \(\alpha'(0) = v\)), then \(J \subseteq I\) and \(\alpha = \gamma|_J\).

**Proof.** Since the uniqueness is clear from the definition, let us focus on existence.

**Local existence and uniqueness.** To begin with, note that there is an open interval \(I_0\) containing \(0\) and a unique geodesic \(\beta : I_0 \to M\) such that \(\beta'(0) = v\); this is an immediate consequence of applying standard results concerning ODE’s to the equation (3.13). In other words, local existence and uniqueness holds.

**Global uniqueness.** In order to proceed, we need to prove global uniqueness. In other words, we need to prove that if \(I_i, i = 0, 1,\) are open intervals containing \(0\) and \(\beta_i : I_i \to M\) are geodesics such that \(\beta_i'(0) = v\), then \(\beta_0 = \beta_1\) on \(I_0 \cap I_1\). In order to prove this statement, let \(A\) be the set of \(t \in I_0 \cap I_1\) such that \(\beta_0'(t) = \beta_1'(t)\). Note that if \(t \in A\), then \(\beta_0(t) = \pi \circ \beta_0'(t) = \pi \circ \beta_1'(t) = \beta_1(t)\), where \(\pi : TM \to M\) is the projection taking a tangent vector to its base point. In other words, if we can prove that \(A = I_0 \cap I_1\), it then follows that \(\beta_0 = \beta_1\) on \(I_0 \cap I_1\). Since \(0 \in A\), it is clear that \(A\) is non-empty. Due to local uniqueness, \(A\) is open. In order to prove that \(A\) is closed, let \(t_1 \in I_0 \cap I_1\) belong to the closure of \(A\). Then there is a sequence \(s_j \in A\) such that \(s_j \to t_1\). Since \(\beta_1' : I_1 \to TM\) are smooth maps, it is clear that

\[
\beta_1'(t_1) = \lim_{j \to \infty} \beta_1'(s_j) = \lim_{j \to \infty} \beta_0'(s_j) = \beta_0'(t_1).
\]

Thus \(t_1 \in A\). Summing up, \(A\) is an open, closed and non-empty subset of \(I_0 \cap I_1\). Thus \(A = I_0 \cap I_1\). In other words, global uniqueness holds.

**Existence.** Let \(I_a, a \in A\), be the collection of open intervals \(I_a \subseteq \mathbb{R}\) such that

- \(0 \in I_a\),
- there is a geodesic \(\gamma_a : I_a \to M\) such that \(\gamma_a'(0) = v\).

Due to local existence, we know that this collection of intervals is non-empty. Define

\[
I = \bigcup_{a \in A} I_a.
\]

Then \(I \subseteq \mathbb{R}\) is an open interval containing \(0\). Moreover, due to global uniqueness, we can define a geodesic \(\gamma : I \to M\) such that \(\gamma'(0) = v\); simply let \(\gamma(t) = \gamma_a(t)\) for \(t \in I_a\). Finally, it is clear, by definition, that \(I\) is maximal. \(\square\)
The geodesic constructed in Proposition 60 is called the *maximal geodesic* with initial data given by \( v \in T_p M \).

**Exercise 61.** Prove that the maximal geodesics in Euclidean space and in Minkowski space are the straight lines.

Let \( \gamma \) be a geodesic. Then
\[
\frac{d}{dt} \langle \gamma', \gamma' \rangle = \langle \gamma'', \gamma' \rangle + \langle \gamma', \gamma'' \rangle = 0.
\]
In other words, \( \langle \gamma', \gamma' \rangle \) is constant so that the following definition makes sense.

**Definition 62.** Let \((M, g)\) be a semi-Riemannian manifold and \( \gamma \) be a geodesic on \((M, g)\). Then \( \gamma \) is said to be *spacelike* if \( \langle \gamma', \gamma' \rangle > 0 \) or \( \gamma' = 0 \); \( \gamma \) is said to be *timelike* if \( \langle \gamma', \gamma' \rangle < 0 \); and \( \gamma \) is said to be *lightlike* or *null* if \( \langle \gamma', \gamma' \rangle = 0 \), \( \gamma' \neq 0 \). A geodesic which is either timelike or null is said to be *causal*.

**Remark 63.** In a spacetime, we can also speak of future oriented timelike, null and causal curves.

In general relativity, timelike geodesics are interpreted as the trajectories of freely falling test particles and null geodesics are interpreted as the trajectories of light. In particular, in Lorentz geometry, we can think of the timelike geodesics as freely falling observers. Moreover, if \( \gamma : I \to M \) is a future oriented timelike geodesic in a spacetime and \( t_0 < t_1 \) are elements of \( I \), then
\[
\int_{t_0}^{t_1} \left( -\langle \gamma'(t), \gamma'(t) \rangle \right)^{1/2} \, dt
\]
is the proper time between \( t_0 \) and \( t_1 \) as measured by the observer \( \gamma \). Since the integrand is constant, it is clear that if \( I = (t_-, t_+) \) and \( t_+ < \infty \), then the amount of proper time the observer can measure to the future is finite (there is an analogous statement concerning the past if \( t_- > -\infty \)). This can be thought of as saying that the observer leaves the spacetime in finite proper time. One way to interpret this is that there is a singularity in the spacetime. It is therefore of interest to analyze under what circumstances \( I \neq \mathbb{R} \). To begin with, let us introduce the following terminology.

**Definition 64.** Let \((M, g)\) be a semi-Riemannian manifold and \( \gamma : I \to M \) be a maximal geodesic in \((M, g)\). Then \( \gamma \) is said to be a *complete geodesic* if \( I = \mathbb{R} \). A semi-Riemannian manifold, all of whose maximal geodesics are complete is said to be *complete*.

Euclidean space and Minkowski space are both examples of complete semi-Riemannian manifolds. On the other hand, removing one single point from Euclidean space or Minkowski space yields an incomplete semi-Riemannian manifold. In other words, the notion of completeness is very sensitive. Moreover, it is clear that in order to interpret the presence of an incomplete causal geodesic as the existence of a singularity (as is sometimes done), it is necessary to ensure that the spacetime under consideration is maximal in some natural sense. Nevertheless, trying to sort out conditions ensuring that a spacetime (which is maximal in some natural sense) is causally geodesically incomplete is a fundamental problem. Due to the work of Hawking and Penrose, spacetimes are causally geodesically incomplete under quite general circumstances. The relevant results, which are known under the name of “the singularity theorems”, are discussed, e.g., in [1, Chapter 14].

### 3.4 Variational characterization of geodesics

Another perspective on geodesics is obtained by considering the variation of the length of curves that are close to a fixed curve. To be more precise, let \((M, g)\) be a semi-Riemannian manifold, \( t_0 < t_1 \) and \( \epsilon > 0 \) be real numbers, and
\[
\nu : [t_0, t_1] \times (-\epsilon, \epsilon) \to M.
\]
The function $\nu$ should be thought of as a variation of the curve $\gamma(t) = \nu(t, 0)$. Let

$$L(s) = \int_{t_0}^{t_1} |(\partial_t \nu(t, s), \partial_t \nu(t, s))|^{1/2} dt.$$ 

A natural question to ask is: what are the curves $\gamma$ such that for every variation $\nu$ (as above, with an appropriate degree of regularity and fixing the endpoints $t_0$ and $t_1$), $L'(0) = 0$? Roughly speaking, it turns out to be possible to characterize geodesics as the curves for which $L'(0)$ for all such variations. In particular, geodesics in Riemannian geometry are the locally length minimizing curves.

Here we shall not pursue this perspective further, but rather refer the interested reader to, e.g., [1, Chapter 10].
Chapter 4

Curvature

The notion of curvature arose over a long period of time. Some of the history can be found in [3]. Here we proceed in a more formal way. As indicated at the beginning of the previous chapter, one way to define curvature is through parallel translation along a closed curve. Here we define the curvature tensor via an “infinitesimal” version of this idea.

4.1 The curvature tensor

**Proposition 65.** Let \((M,g)\) be a semi-Riemannian manifold and \(\nabla\) denote the associated Levi-Civita connection. Then the function \(R : \mathfrak{X}(M)^3 \to \mathfrak{X}\) defined by

\[
R_{XY}Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]}Z
\] (4.1)

is linear over \(C^\infty(M)\). In particular, it can thus be interpreted as a tensor field, and it is referred to as the Riemannian curvature tensor of \((M,g)\).

**Remark 66.** Even though the notation \(R(X,Y,Z)\) may seem more reasonable, the convention (4.1) is the one commonly used. Since \(R\) does not take its values in \(C^\infty(M)\), the statement that it is a tensor field requires some justification. The reason for the terminology is that we can easily consider \(R\) to be a map from \(\mathfrak{X}(M)^3 \times \mathfrak{X}^*(M) \to C^\infty(M)\) according to

\[(X, Y, Z, \eta) \mapsto \eta(R_{XY}Z).\]

That the corresponding map is linear over the functions in \(\eta\) is obvious. If it is linear over the functions in the other arguments, the tensor characterization lemma [2, Lemma 12.24, p. 318] thus yields the conclusion that we can think of \(R\) as of a tensor field.

**Proof.** That \(\mathbb{R}\) is linear over the real numbers is clear. The only thing we need to prove is thus that

\[
R_{(fX)Y}Z = fR_{XY}Z,
\] (4.2)

\[
R_{X(fY)}Z = fR_{XY}Z,
\] (4.3)

\[
R_{XY}(fZ) = fR_{XY}Z.
\] (4.4)

We prove one of these equalities and leave the other two as exercises. Compute

\[
R_{XY}(fZ) = \nabla_X[f \nabla_Y Z + Y(f)Z] - \nabla_Y[X(f)Z + f \nabla_X Z] - [X,Y](f)Z - f \nabla_{[X,Y]}Z
\]

\[
= X(f) \nabla_Y Z + XY(f)Z + Y(f) \nabla_X Z + f \nabla_X \nabla_Y Z - YX(f)Z - X(f) \nabla_Y Z
\]

\[
- Y(f) \nabla_X Z - f \nabla_Y \nabla_X Z - [X,Y](f)Z - f \nabla_{[X,Y]}Z
\]

\[
= fR_{XY}Z.
\]
Due to Exercise 67, the proposition follows.

Exercise 67. Prove (4.2) and (4.3).

By an argument similar to the proof of the tensor characterization lemma, cf. [2, pp. 318–319], Proposition 65 implies that it is possible to make sense of $R_{xyz}$ for $x, y, z \in T_pM$. In fact, choosing any vector fields $X, Y, Z \in \mathfrak{X}(M)$ such that $X_p = x$, $Y_p = y$, and $Z_p = z$, we can define $R_{xyz}$ by

$$R_{xyz} = (R_{XZy})(p);$$

the right hand side is independent of the choice of vectorfields $X, Y, Z$ satisfying $X_p = x$, $Y_p = y$, and $Z_p = z$. Moreover, we can think of $R_{xyz}$ as a linear map from $T_pM$ to $T_pM$. The curvature tensor has several symmetries.

Proposition 68. Let $(M, g)$ be a semi-Riemannian manifold and $x, y, z, v, w \in T_pM$, where $p \in M$. Then

$$R_{xy} = -R_{yx},$$

$$(R_{xy}v, w) = -\langle v, R_{xy}w \rangle,$$  \hspace{1cm} (4.5)

$$R_{xy} + R_{yx}x + R_{zx}y = 0,$$ \hspace{1cm} (4.6)

$$\langle R_{xy}v, w \rangle = \langle R_{vw}x, y \rangle.$$ \hspace{1cm} (4.7)

Proof. Choose vector fields $X, Y, Z, V, W$ so that $X_p = x$ etc. Without loss of generality, we may assume the Lie brackets of any pairs of vector fields in $\{X, Y, Z, V, W\}$ to vanish; simply choose these vector fields to have constant coefficients relative to coordinate vector fields (it is sufficient to carry out the computation locally).

That (4.5) holds is an immediate consequence of the definition. Note that (4.6) is equivalent to

$$\langle R_{xy}z, z \rangle = 0 \hspace{1cm} (4.9)$$

for all $x, y, z \in T_pM$. In order to prove (4.9), compute (using (3.2) and the fact that $[X, Y] = 0$)

$$\langle R_{XY}Z, Z \rangle = \langle \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z, Z \rangle = X \langle \nabla_Y Z, Z \rangle - \langle \nabla_Y Z, \nabla_X Z \rangle - Y \langle \nabla_X Z, Z \rangle + \langle \nabla_X Z, \nabla_Y Z \rangle = \frac{1}{2} \langle X, Y \rangle \langle Z, Z \rangle = 0.$$

Thus (4.6) holds. In order to prove (4.7), compute

$$R_{XY}Z + R_{YZ}X + R_{ZX}Y = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z + \nabla_Y \nabla_Z X - \nabla_Z \nabla_Y X + \nabla_Z \nabla_X Y - \nabla_X \nabla_Z Y = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z = 0.$$

In order to prove (4.8), note that

$$\langle R_{xy}v + R_{yu}x + R_{ux}y, w \rangle = 0$$

due to (4.7). Adding up the four cyclic permutations of this equation and using (4.5) and (4.6) yields (4.8). We leave the details to the reader.

Exercise 69. Prove (4.8).
4.2 Calculating the curvature tensor

It is of interest to derive a formula for the curvature in terms of a frame. Let \( \{e_i\} \) be a local frame. Then we define the associated connection coefficients \( \Gamma^i_{jk} \) by

\[
\nabla e_j e_k = \Gamma^i_{jk} e_i.
\]

In case \( e_i = \partial_i \), the connection coefficients are the Christoffel symbols given by (3.5). However, for a general frame, the relation \( \Gamma^k_{ij} = \Gamma^k_{ji} \) does typically not hold. This is due to the fact that the Lie bracket \([e_i, e_j]\) typically does not vanish. Note that the information concerning the Lie brackets is contained in the functions \( \gamma^k_{ij} \) defined by

\[
[e_i, e_j] = \gamma^k_{ij} e_k.
\]

Let us compute

\[
R_{e_i e_j e_k} e_l = \nabla_{e_i} \nabla_{e_j} e_k - \nabla_{e_j} \nabla_{e_i} e_k - \nabla_{[e_i, e_j]} e_k
\]

\[
= \nabla_{e_i} (\Gamma^l_{jk} e_l) - \nabla_{e_j} (\Gamma^l_{ik} e_l) - \gamma^l_{ij} \nabla_{e_i} e_k
\]

\[
= e_i (\Gamma^l_{jk}) e_l + \Gamma^l_{jk} \nabla_{e_k} e_i - e_j (\Gamma^l_{ik}) e_l - \Gamma^l_{ik} \nabla_{e_i} e_l - \gamma^l_{ij} \Gamma^m_{ik} e_m
\]

\[
= e_i (\Gamma^l_{jk}) e_l + \Gamma^l_{jk} \Gamma^m_{il} e_m - e_j (\Gamma^m_{ik}) e_m - \Gamma^l_{ik} \Gamma^m_{jl} e_m - \gamma^l_{ij} \Gamma^m_{ik} e_m.
\]

This equality can be written

\[
R_{e_i e_j e_k} e_l = -R^m_{ij k} e_m,
\]

where

\[
R^m_{ij k} = e_j (\Gamma^m_{ik}) - e_i (\Gamma^m_{jk}) + \Gamma^m_{ik} \Gamma^l_{jl} - \Gamma^l_{ik} \Gamma^m_{jl} + \gamma^l_{ij} \Gamma^m_{lk}.
\]

The motivation for including a minus sign in (4.11) is perhaps not so clear. There are several different conventions, but we have included the minus sign to obtain consistency with some of the standard references. The symbol \( R^m_{ij k} \) should be thought of as the components of the curvature tensor (which is a \((1,3)\) tensor field) with respect to the frame \( \{e_i\} \). In case the frame is given by \( e_i = \partial_i \), the \( \gamma^k_{ij} \)'s vanish, and we obtain the formula

\[
R^m_{ij k} = \partial_j \Gamma^m_{ik} - \partial_i \Gamma^m_{jk} + \Gamma^l_{ik} \Gamma^m_{jl} - \Gamma^l_{ik} \Gamma^m_{jl}.
\]

Moreover, in this case, the \( \Gamma^k_{ij} \)'s are given by (3.5). With respect to the standard coordinates, the Christoffel symbols of the Euclidean metric and the Minkowski metric vanish. In particular, the associated curvature tensors thus vanish. Moreover, this property (essentially) characterizes Euclidean space and Minkowski space. To prove this statement is, however, non-trivial.

The general strategy for computing the components of the curvature tensor is the following. First, choose a suitable local frame. Which frame is most appropriate depends on the context. Sometimes it is convenient to use a coordinate frame, but sometimes it is easier to carry out the computations with respect to an orthonormal frame. Once a choice of frame has been made, one first calculates the functions \( \gamma^k_{ij} \) determined by the Lie bracket. Then, one calculates the coefficients \( \Gamma^k_{ij} \) using the Koszul formula, (3.4). After this has been done, the components of the curvature can be calculated using (4.12). Needless to say, this is a cumbersome process in most cases.

4.3 The Ricci tensor and scalar curvature

Since the curvature tensor is a \((1,3)\) tensor field, we can contract two of the indices in order to obtain a covariant 2-tensor field. In fact, we define the Ricci tensor to be the covariant 2-tensor field whose components are given by

\[
R_{ik} = R^j_{ijk}.
\]
In terms of local coordinates, the components of the Ricci tensor are given by
\[ R_{ik} = \partial_j \Gamma^j_{ik} - \partial_i \Gamma^j_{jk} + \Gamma^l_{ik} \Gamma^j_{jl} - \Gamma^l_{jk} \Gamma^j_{il}. \]

Again, the Ricci tensor of Euclidean space and Minkowski space vanish. In what follows, we denote the tensor field whose components are given by (4.13) by Ric. In other words, if \( R_{ijk} \) are the components of the curvature tensor relative to a frame \( \{e_i\} \), then
\[ \text{Ric}(e_i, e_k) = R_{ijk}^j. \]

The Ricci tensor is an extremely important object in semi-Riemannian geometry in general, and in general relativity in particular. It is of interest to derive alternate formulae for the Ricci tensor.

**Lemma 70.** Let \((M, g)\) be a semi-Riemannian manifold and \( \{e_i\} \) be an orthonormal frame such that \( \langle e_i, e_i \rangle = \epsilon_i \) (no summation on \( i \)). Then
\[ \text{Ric}(X, Y) = \sum_j \epsilon_j \langle R_{e_j X Y}, e_j \rangle. \quad (4.14) \]

**Proof.** Note that with respect to a local frame \( \{e_i\} \),
\[ \langle R_{e_i e_j e_k}, e_l \rangle = -R_{ijk}^m \langle e_m, e_l \rangle = -R_{ijk}^m g_{ml}, \quad (4.15) \]
where all the components are calculated with respect to the frame \( \{e_i\} \). Assume now that the frame is orthonormal so that \( \langle e_i, e_j \rangle = 0 \) if \( i \neq j \) and \( \langle e_i, e_i \rangle = \epsilon_i \) (no summation on \( i \)), where \( \epsilon_i = \pm 1 \). Letting \( l = j \) in (4.15) then yields
\[ -\epsilon_j R_{ijk}^j = \langle R_{e_i e_j e_k}, e_j \rangle \]
(no summation on \( j \)). Thus
\[ R_{ijk}^j = -\epsilon_j \langle R_{e_i e_j e_k}, e_j \rangle = \epsilon_j \langle R_{e_j e_i e_k}, e_j \rangle \]
(no summation on \( j \)), where we have appealed to (4.5). Summing over \( j \) now yields
\[ \text{Ric}(e_i, e_k) = \sum_j \epsilon_j \langle R_{e_j e_i e_k}, e_j \rangle. \]

If \( X = X^i e_i \) and \( Y = Y^i e_i \) are elements of \( \mathfrak{X}(M) \), we then obtain
\[ \text{Ric}(X, Y) = X^i Y^k \text{Ric}(e_i, e_k) = X^i Y^k \sum_j \epsilon_j \langle R_{e_j e_i e_k}, e_j \rangle = \sum_j \epsilon_j \langle R_{e_j X Y}, e_j \rangle. \]

The lemma follows.

It is of interest to note that, as a consequence of (4.14), (4.5), (4.6) and (4.8),
\[ \text{Ric}(X, Y) = \sum_j \epsilon_j \langle R_{e_j X Y}, e_j \rangle = \sum_j \epsilon_j \langle R_{Y e_j} e_j, X \rangle = \sum_j \epsilon_j \langle R_{e_j Y} X, e_j \rangle = \text{Ric}(Y, X). \]

In other words, the Ricci tensor is a symmetric covariant 2-tensor field.

Finally, we define the **scalar curvature** \( S \) of a semi-Riemannian manifold by the formula
\[ S = g^{ij} R_{ij}. \]
4.4 The divergence, the gradient and the Laplacian

The divergence of a vector field. Let \((M,g)\) be a semi-Riemannian manifold with associated Levi-Civita connection \(\nabla\). If \(X \in \mathfrak{X}(M)\), we can think of \(\nabla X\) as \((1,1)\)-tensor field according to \((Y,\eta) \mapsto \eta(\nabla_Y X)\); note that this map is bilinear over the smooth functions and thus defines a \((1,1)\)-tensor field due to the tensor characterization lemma. The components of this tensor field with respect to coordinates would in physics notation be written \(\nabla_i X^j\). They are given by

\[
\nabla_i X^j = dx^j(\nabla_i X) = dx^j[\nabla_{\partial_i} (X^k \partial_k)] = dx^j[(\partial_i X^k) \partial_k + X^k \nabla_{\partial_i} \partial_k]
\]

Contracting the components of this tensor field yields a smooth function. We define the divergence of \(X \in \mathfrak{X}(M)\), written \(\text{div} X\), to be the function which in local coordinates is given by

\[
\text{div} X = \nabla_i X^i = \partial_i X^i + \Gamma^i_{ik} X^k.
\]

In Euclidean space, this gives the familiar formula, since \(\Gamma^k_{ij} = 0\).

The gradient of a function. If \(f \in C^\infty(M)\), then \(df \in \mathfrak{X}^*(M)\). Applying the isomorphism \(\sharp\) to \(df\), we thus obtain a vectorfield referred to as the gradient of \(f\):

\[
\text{grad} f = (df)\wedge.
\]

In local coordinates,

\[
\text{grad} f = g^{ij}(\partial_i f)\partial_j.
\]

The Laplacian of a function. Finally, taking the divergence of the gradient yields the Laplacian

\[
\Delta f = \text{div}(\text{grad} f).
\]

In the case of Euclidean space, this definition yields the ordinary Laplacian. However, in the case of Minkowski space, it yields the wave operator.

4.5 Computing the covariant derivative of tensor fields

So far, we have only applied the Levi-Civita connection to vectorfields. However, it is also possible to apply it to tensor fields. To begin with, let us apply it to a one-form. To this end, let \(\eta \in \mathfrak{X}^*(M)\) and \(X, Y \in \mathfrak{X}(M)\). Then we define

\[
(\nabla_X \eta)(Y) = X[\eta(Y)] - \eta(\nabla_X Y).
\]

Exercise 71. Prove that \((\nabla_X \eta)(Y)\) defined by (4.16) is linear over the smooth functions in the argument \(Y\) (so that \(\nabla_X \eta\) is a one-form due to the tensor characterization lemma). Prove, moreover, that

\[
\nabla_X \eta = (\nabla_X \eta)\wedge.
\]

Note that \(\nabla \eta\) can be thought of as a covariant 2-tensor field according to

\[
(X, Y) \mapsto (\nabla_X \eta)(Y).
\]

The components of this tensor field with respect to local coordinates is given by

\[
(\nabla_{\partial_i} \eta)(\partial_j) = \partial_i[\eta(\partial_j)] - \eta(\nabla_{\partial_i} \partial_j) = \partial_i \eta_j - \eta(\Gamma^k_{ij} \partial_k) = \partial_i \eta_j - \Gamma^k_{ij} \eta_k.
\]
where $\Gamma^k_{ij}$ are the Christoffel symbols associated with the coordinates $(x^i)$. Physicists would write this equation as
\[ \nabla_i \eta_j = \partial_i \eta_j - \eta_k \Gamma^k_{ij}. \]

In order to generalize this to tensorfields, let $T$ be a tensorfield of type $(k, l)$. We can then think of $T$ as a map from $\mathfrak{X}(M)^* \times \cdots \times \mathfrak{X}^*(M) \times \mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)$ ($k$ copies of $\mathfrak{X}(M)^*$ and $l$ copies of $\mathfrak{X}(M)$) to $C^\infty(M)$ which is multilinear over the smooth functions. If $\eta_1, \ldots, \eta_k \in \mathfrak{X}^*(M)$ and $Y, X_1, \ldots, X_l \in \mathfrak{X}(M)$, then $\nabla_X T$ is defined by the relation
\[
(\nabla_X T)(\eta_1, \ldots, \eta_k, X_1, \ldots, X_l) = X[T(\eta_1, \ldots, \eta_k, X_1, \ldots, X_l)] \\
- T(\nabla_X \eta_1, \eta_2, \ldots, \eta_k, X_1, \ldots, X_l) - \cdots - T(\eta_1, \ldots, \eta_{k-1}, \nabla_X \eta_k, X_1, \ldots, X_l) \\
- T(\eta_1, \ldots, \eta_k, \nabla_X X_1, X_2, \ldots, X_l) - \cdots - T(\eta_1, \ldots, \eta_k, X_1, \ldots, X_{l-1}, \nabla_X X_l).
\]

**Exercise 72.** Prove that $(\nabla_X T)(\eta_1, \ldots, \eta_k, X_1, \ldots, X_l)$ defined by the formula (4.17) is linear over the smooth functions in $\eta_1, \ldots, \eta_k, X_1, \ldots, X_l$. Due to the tensor characterization lemma, this implies that $\nabla_X T$ is a tensorfield of type $(k, l)$.

It is of interest to calculate $\nabla g$, where $g$ is the metric.

**Exercise 73.** Let $(M, g)$ be a semi-Riemannian manifold and let $\nabla$ be the associated Levi-Civita connection. Prove that $\nabla g = 0$.

### 4.5.1 Divergence of a covariant 2-tensor field

In the context of Einstein’s equations, it is of interest to calculate the divergence of symmetric covariant 2-tensor fields. For that reason, we here wish to define the divergence and to derive a convenient formula for calculating it.

Let $T$ be a symmetric covariant 2-tensorfield. Then we can think if $\nabla T$ a covariant 3-tensor field according to
\[
(X, Y, Z) \mapsto (\nabla_X T)(Y, Z).
\]

The components of this tensorfield with respect to local coordinates $(x^i)$ are given by
\[
(\nabla_{\partial_i} T)(\partial_j, \partial_k) = \partial_i T_{jk} - T(\nabla_{\partial_i} \partial_j, \partial_k) - T(\partial_j, \nabla_{\partial_i} \partial_k) = \partial_i T_{jk} - \Gamma^l_{ij} T_{lk} - \Gamma^l_{ik} T_{jl}.
\]

Again, in physics notation, this equation would be written
\[
\nabla_i T_{jk} = \partial_i T_{jk} - \Gamma^l_{ij} T_{lk} - \Gamma^l_{ik} T_{jl}.
\]

We here follow this convention by denoting the components of the tensorfield defined by (4.18) by $\nabla_i T_{jk}$. We use this notation also in the case that the components are calculated with respect to a frame as opposed to only coordinate frames. However, which frame we use should be clear from the context. Note that $\nabla_i T_{jk} = \nabla_i T_{kj}$ (this is a consequence of the fact that $T$ is symmetric). We define $\text{div} T$ to be the one-form whose components are given by
\[
(\text{div} T)_k = g^{ij} \nabla_i T_{jk}.
\]

In physics notation, this equation would be written
\[
(\text{div} T)_k = \nabla^i T_{ik};
\]

first you raise the first index and then you contract with the second index.
4.6. AN EXAMPLE OF A CURVATURE CALCULATION

Say that \( \{e_i\} \) is an orthonormal frame with \( \epsilon_i = \langle e_i, e_i \rangle \). Then
\[
(\text{div} T)_k = g^{ij} \nabla_i T_{jk} = g^{ij} \nabla_i T_{jk} = \sum_i \epsilon_i \nabla_i T_{ik} = \sum_i \epsilon_i \nabla_i(e_i, e_k),
\]
where it is taken for granted that all the indices are calculated with respect to the frame \( \{e_i\} \).

Say that \( X = X^i e_i \) is a smooth vector field. Then
\[
(\text{div} T)(X) = X^k (\text{div} T)(e_k) = X^k (\text{div} T)_k = X^k \sum_i \epsilon_i \nabla_i T(e_i, e_k) = \sum_i \epsilon_i \nabla_i T(e_i, X).
\]

To conclude
\[
(\text{div} T)(X) = \sum_i \epsilon_i \nabla_i T(e_i, X).
\]

4.6 An example of a curvature calculation

In the present section, we calculate the Ricci tensor of one specific metric; cf. (4.19) below. Our motivation for doing so is that the calculations themselves illustrate the theory. However, the particular metric we have chosen is such that \( n \)-dimensional hyperbolic space and the 2-sphere are two special cases. Moreover, the Lorentz manifolds used by physicists nowadays usually have a metric of the form (4.19).

**The metric.** Define the metric \( g \) by the formula
\[
g = \epsilon dt \otimes dt + f^2(t) \sum_{i=1}^{n} dx^i \otimes dx^i \tag{4.19}
\]
on \( I \times U \), where \( I \) is an open interval and \( U \) is an open subset of \( \mathbb{R}^n \). Moreover, \( t \) is the coordinate on the interval \( I \) and \( x^i \) are the standard coordinates on \( \mathbb{R}^n \). Finally, \( f \) is a strictly positive smooth function on \( I \) and \( \epsilon = \pm 1 \). If \( \epsilon = 1 \), the metric is Riemannian, and if \( \epsilon = -1 \), \( g \) is a Lorentz metric.

**The orthonormal frame.** The curvature calculations can be carried out in many different ways. Here we shall use an orthonormal frame, denoted \( \{e_\alpha\}, \alpha = 0, \ldots, n \), and defined by
\[
e_0 = \partial_t, \quad e_i = \frac{1}{f} \partial_i,
\]
where \( i = 1, \ldots, n \); we shall here use the convention that Greek indices range from 0 to \( n \) and that Latin indices range from 1 to \( n \).

**The coefficients of the Lie bracket.** Our goal here is to compute the curvature of the metric (4.19). The strategy is to first compute the \( \gamma^\lambda_{\alpha\beta} \), defined by the relation
\[
[e_\alpha, e_\beta] = \gamma^\lambda_{\alpha\beta} e_\lambda.
\]
Then the idea is to use the Koszul formula (3.4) to calculate the connection coefficients, defined by
\[
\nabla_{e_\alpha} e_\beta = \Gamma^\lambda_{\alpha\beta} e_\lambda.
\]
Since \( \gamma_{\alpha\beta} = -\gamma^\lambda_{\beta\alpha} \), it is sufficient to compute \( \gamma^\lambda_{\alpha\beta} \) for \( \alpha < \beta \). Compute
\[
[e_0, e_i] = -\frac{\dot{f}}{f^2} \partial_i = H e_i,
\]
where \( H \) is the function defined by
\[
H = -\frac{\dot{f}}{f}.\]
Thus $\gamma^i_0 = 0$ unless $\alpha = i$ and $\gamma^i_i = H$.

Since $[e_i, e_j] = 0$ for all $i, j = 1, \ldots, n$, we have

$$\gamma^i_0 = 0$$

for all $\alpha = 0, \ldots, n$. To conclude, the only $\gamma^i_0$'s that do not vanish are

$$\gamma^i_0 = H, \gamma^i_i = -H,$$

(4.23)

where $i = 1, \ldots, n$ and we do not sum over $i$.

### 4.6.1 Computing the connection coefficients

The next step is to compute the connection coefficients. Let us first derive a general formula for the connection coefficients of an orthonormal frame.

**Lemma 74.** Let $(M, g)$ be a semi-Riemannian manifold and let $\{e_\alpha\}$, $\alpha = 0, \ldots, n$, be an orthonormal frame on an open subset $U$ of $M$. Define the connection coefficients $\Gamma^{\lambda}_{\alpha\beta}$ by the formula (4.22) and $\gamma^\alpha_{\beta\alpha}$ by (4.21). Then

$$\Gamma^{\lambda}_{\alpha\beta} = \frac{1}{2} \left(-\epsilon_\alpha \epsilon_\beta \gamma^\alpha_{\beta\lambda} + \epsilon_\lambda \epsilon_\beta \gamma^\beta_{\alpha\lambda} + \gamma^\lambda_{\alpha\beta}\right)$$

(no summation on any index), where $\epsilon_\alpha = g(e_\alpha, e_\alpha)$.

**Proof.** Due to the Koszul formula, (3.4),

$$2 \langle \nabla e_\alpha e_\beta, e_\mu \rangle = \epsilon_\alpha \epsilon_\beta \gamma^\alpha_{\beta\mu} + \epsilon_\lambda \epsilon_\beta \gamma^\beta_{\alpha\mu} + \gamma^\lambda_{\alpha\beta}$$

where we used the fact that the frame is orthonormal in the second step and $g_{\alpha\beta} = (e_\alpha, e_\beta)$. On the other hand,

$$2 \langle \nabla e_\alpha e_\beta, e_\mu \rangle = 2 \langle \Gamma^{\nu}_{\alpha\beta} e_\nu, e_\mu \rangle = 2 \Gamma^{\nu}_{\alpha\beta} g_{\mu\nu}.$$ Combining these two equations yields

$$\Gamma^{\nu}_{\alpha\beta} g_{\mu\nu} = \frac{1}{2} \left(-\epsilon_\nu \epsilon_\beta \gamma^\nu_{\beta\mu} + \epsilon_\nu \epsilon_\alpha \gamma^\nu_{\alpha\mu} + \epsilon_\nu \gamma^\nu_{\alpha\beta}\right)$$

Multiplying this equation with $g^{\lambda\mu}$ and summing over $\nu$ yields

$$\Gamma^{\lambda}_{\alpha\beta} = \frac{1}{2} \left(-\gamma^\nu_{\beta\mu} g^{\lambda\mu} + \gamma^\nu_{\alpha\mu} g^{\lambda\nu} + \gamma^\nu_{\alpha\beta} g^{\nu\mu}\right)$$

(no summation on any index), where we have used the fact that $g_{\alpha\beta} = \epsilon_\alpha \delta_{\alpha\beta}$. The lemma follows. \hfill \square
4.6. AN EXAMPLE OF A CURVATURE CALCULATION

Calculating the connection coefficients in the case of the metric \((4.19)\). Let us now return to the metric \((4.19)\). Considering the formula \((4.24)\), it is of interest to note the following. In all the terms on the right hand side, the indices in the \(\gamma^\mu_\alpha_\beta\)’s are simply permutations of the indices in \(\Gamma^\mu_\alpha_\beta\). In our particular setting, the only combination of indices in \(\gamma^\mu_\alpha_\beta\) that (may) give a non-zero result is if one of the indices is 0 and the other two are equal and belong to \(\{1, \ldots, n\}\).

Let us compute, using \((4.24)\),

\[
\Gamma^i_0_0 = \frac{1}{2} (-\epsilon^{\gamma_0_0} + \gamma^i_0_0) = 0,
\]

\[
\Gamma^i_0_\alpha = \frac{1}{2} (-\gamma^0_0 + \epsilon^{\gamma_0_0} + \gamma^i_0) = -\gamma^0_i = -H,
\]

\[
\Gamma^0_\alpha_\beta = \frac{1}{2} (-\epsilon^{\gamma_0_\alpha_\beta} + \epsilon^{\gamma_0_\alpha} + \gamma^0_0) = \epsilon^i_0 = \epsilon H.
\]

To conclude, the only connection coefficients which are non-zero are

\[
\Gamma^i_0 = -H, \quad \Gamma^0_\alpha = \epsilon H \tag{4.25}
\]

(no summation on \(i\)).

4.6.2 Calculating the components of the Ricci tensor

Let us now turn to the problem of calculating the Ricci tensor. Again, it is useful to derive a general expression for the the components of the Ricci tensor with respect to an orthonormal frame.

Lemma 75. Let \((M, g)\) be a semi-Riemannian manifold and let \(\{e_\alpha\}, \alpha = 0, \ldots, n\), be an orthonormal frame on an open subset \(U\) of \(M\). Define the connection coefficients \(\Gamma^\lambda_\alpha_\beta\) by the formula \((4.22)\) and \(\gamma^\lambda_\alpha_\beta\) by \((4.21)\). Then

\[
\text{Ric}(e_\mu, e_\nu) = e_\alpha (\Gamma^\alpha_\mu_\nu) + \Gamma^\lambda_\mu_\nu \gamma^\alpha_\lambda - \epsilon_\mu (\Gamma^\alpha_\mu_\nu) - \lambda_\alpha_\nu \Gamma^\alpha_\mu_\nu - \gamma^\lambda_\alpha_\nu \Gamma^\alpha_\nu,
\]

where Einstein’s summation convention applies.

Proof. Due to \((4.14)\),

\[
\text{Ric}(e_\mu, e_\nu) = \sum_\alpha e_\alpha \langle R_{e_\alpha e_\mu} e_\nu, e_\alpha \rangle. \tag{4.27}
\]

On the other hand, \((4.10)\) yields

\[
\langle R_{e_\alpha e_\mu} e_\nu, e_\alpha \rangle = e_\alpha (\Gamma^\alpha_\mu_\nu) e_\beta + \lambda_\mu_\nu \Gamma^\alpha_\lambda e_\beta - \epsilon_\mu (\Gamma^\alpha_\mu_\nu) e_\beta - \lambda_\alpha_\nu \Gamma^\alpha_\mu_\nu e_\beta - \gamma^\lambda_\alpha_\nu \Gamma^\alpha_\nu e_\beta, e_\alpha
\]

(no summation on \(\alpha\)), where we have used the fact that \(\langle e_\alpha, e_\beta \rangle = \epsilon_\alpha \delta_\alpha_\beta\). Combining this observation with \((4.27)\) yields \((4.26)\).

\[
\boxed{\text{The components of the Ricci tensor of the metric (4.19).}}
\]

Let us now compute the components of the Ricci tensor of the metric \((4.19)\). Since the Ricci tensor is symmetric, it is sufficient to compute \(\text{Ric}(e_\mu, e_\nu)\) for \(\mu \leq \nu\). Before computing the individual components, let us make the following observations. Since the \(\Gamma^\mu_\alpha_\beta\)’s only depend on \(t\), \(e_\lambda (\Gamma^\mu_\alpha_\beta) = 0\) unless \(\lambda = 0\). Keeping in mind that the only non-zero connection coefficients are given by \((4.25)\), we conclude that

\[
e_\alpha (\Gamma^\alpha_\mu_\nu) = e_0 (\Gamma^0_\mu_\nu) = 0
\]

unless \(\mu = \nu = i\). Moreover,

\[
e_\alpha (\Gamma^\alpha_i) = e_0 (\Gamma^0_i) = \epsilon H.
\]
The 00-component of the Ricci tensor. Compute, using the above observations as well as the fact that the only non-zero connection coefficients are given by (4.25),
\[
\text{Ric}(e_0, e_0) = e_0(\Gamma_{0i}^0) + \Gamma_{00}^\lambda \Gamma_{\alpha\lambda}^\alpha - e_0(\Gamma_{\alpha0}^0) - \Gamma^\lambda_{\alpha0} \Gamma^\alpha_{0\lambda} - \gamma^\lambda_{\alpha0} \Gamma^\alpha_{\lambda0}
\]
\[
= - \sum_i e_0(\Gamma_{0i}^i) + \sum_i \gamma^i_{0i} \Gamma^i_{0i} = n\dot{H} - nH^2,
\]
where we have used the fact that the only non-zero \(\gamma^\alpha_{\mu\nu}\)'s are given by (4.23).

The 0i-components of the Ricci tensor. Compute
\[
\text{Ric}(e_0, e_i) = e_0(\Gamma_{0i}^0) + \Gamma_{0i}^\lambda \Gamma_{\alpha\lambda}^\alpha - e_0(\Gamma_{\alpha0}^i) - \Gamma^\lambda_{\alpha0} \Gamma^\alpha_{i\lambda} - \gamma^\lambda_{\alpha0} \Gamma^\alpha_{i\lambda} = 0;
\]
since \(\Gamma^\lambda_{0j} = 0\) regardless of what \(\lambda\) and \(\beta\) are, the first, second and fourth terms on the right hand side vanish; since \(\Gamma^\lambda_{0i} = 0\) and \(\Gamma^j_{ij} = 0\) regardless of the values of \(i\) and \(j\), it is clear that \(\Gamma^\alpha_{ai} = 0\) (so that the third term on the right hand side vanishes); in order for the first factor in the fifth term to be non-vanishing, we have to have \(\lambda = \alpha = j\) for some \(j = 1, \ldots, n\), cf. (4.23), but if \(\lambda = \alpha = j\), then the second factor in the fifth term vanishes.

The \(ij\)-components of the Ricci tensor, \(i \neq j\). If \(i \neq j\), then
\[
\text{Ric}(e_i, e_j) = e_i(\Gamma_{ij}^0) + \Gamma_{ij}^\lambda \Gamma_{\alpha\lambda}^\alpha - e_j(\Gamma_{\alpha j}^0) - \Gamma^\lambda_{\alpha j} \Gamma^\alpha_{i\lambda} - \gamma^\lambda_{\alpha j} \Gamma^\alpha_{i\lambda} = 0.
\]
(4.28)
To justify this calculation, note that in order for \(\Gamma_{\mu\nu}^\alpha\) or \(\gamma_{\mu\nu}^\alpha\) to be non-zero, two of the indices have to be non-zero and equal, and the remaining index has to be zero. For this reason, the first three terms on the right hand side of (4.28) vanish. Turning to the last two terms, there are indices such that one of the factors appearing in these terms is non-vanishing. However, then the other factor has to vanish.

The \(ij\)-components of the Ricci tensor, \(i = j\). By arguments similar to ones given above,
\[
\text{Ric}(e_i, e_i) = e_i(\Gamma_{ii}^0) + \Gamma_{ii}^\lambda \Gamma_{\alpha\lambda}^\alpha - e_i(\Gamma_{\alpha i}^0) - \Gamma^\lambda_{\alpha i} \Gamma^\alpha_{i\lambda} - \gamma^\lambda_{\alpha i} \Gamma^\alpha_{i\lambda} = e_0(\Gamma_{ii}^i) + \Gamma_{ii}^0 \Gamma_j^j - \Gamma_{ii}^0 \Gamma_i^0 - \gamma^i_{0i} \Gamma^i_{0i} = e\dot{H} - e\dot{H}^2 + e\dot{H} - e\dot{H}^2 = e\dot{H} - e\dot{H}^2
\]
(no summation on \(i\)).

Summing up, we obtain the following lemma.

**Lemma 76.** Let \(I\) be an open interval, \(U\) be an open subset of \(\mathbb{R}^n\) and \(g\) be defined by (4.19), where \(t\) is the coordinate on the interval \(I\) and \(x^i\) are the standard coordinates on \(\mathbb{R}^n\). Moreover, \(f\) is a strictly positive smooth function on \(I\) and \(\epsilon = \pm 1\). Let \(\{e_\alpha\}, \alpha = 0, \ldots, n\), be the orthonormal frame defined by (4.20). Then
\[
\text{Ric}(e_\alpha, e_\beta) = 0
\]
if \(\alpha \neq \beta\). Moreover,
\[
\text{Ric}(e_0, e_0) = n\dot{H} - nH^2,
\]
\[
\text{Ric}(e_i, e_i) = e\dot{H} - e\dot{H}^2
\]
(no summation on \(i\)), where \(i = 1, \ldots, n\).

### 4.7 Calculating the Ricci curvature of the 2-sphere and \(n\)-dimensional hyperbolic space

It is of interest to apply the calculations of the previous section to two special cases, namely that of the 2-sphere and that of the \(n\)-dimensional hyperbolic space. Let us begin with the 2-sphere.
4.7. THE 2-SPHERE AND HYPERBOLIC SPACE

4.7.1 The Ricci curvature of the 2-sphere

Recall that the $n$-sphere and the metric on the $n$-sphere were defined in Example 27. Here we calculate the Ricci curvature of this metric in case $n = 2$.

**Proposition 77.** Let $S^2$ denote the 2-sphere and $g_{S^2}$ denote the metric on the 2-sphere. If $\text{Ric}[g_{S^2}]$ denotes the Ricci curvature of $g_{S^2}$, then

$$\text{Ric}[g_{S^2}] = g_{S^2}.$$  

**Remark 78.** Note that the Ricci curvature of the 2-sphere is positive definite. There is a general connection between the positive definiteness of the Ricci tensor and the compactness of the manifold. In fact, the so-called Myers theorem implies the following: If $(M, g)$ is a geodesically complete and connected Riemannian manifold such that

$$\text{Ric}[g](v, v) \geq c_0 g(v, v)$$  

for some constant $c_0 > 0$, all $v \in T_p M$ and all $p \in M$, then $M$ is compact (and $\pi_1(M)$ is finite). We shall not prove this theorem here but refer the interested reader to [1, Theorem 24, p. 279] for a proof.

**Remark 79.** In the case of 3-dimensions, there is an even deeper connection between the positive definiteness of the Ricci tensor and the topology of the manifold. In fact, if $(M, g)$ is a connected, simply connected and geodesically complete Riemannian manifold of dimension 3 with positive definite Ricci curvature (in the sense that (4.29) holds), then $M$ is diffeomorphic to the 3-sphere. This result is due to Richard Hamilton and we shall not prove it here.

**Proof.** Let

$$\psi : (0, \pi) \times (0, 2\pi) \to S^2$$

be defined by

$$\psi(\theta, \phi) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta).$$  

Then $\psi$ is a diffeomorphism onto its image, the image of $\psi$ is dense in $S^2$ and

$$\psi^* g_{S^2} = d\theta \otimes d\theta + \sin^2 \theta \, d\phi \otimes d\phi.$$  

We leave the verification of these statements as an exercise. Due to these facts (and the smoothness of the Ricci tensor and the metric), it is sufficient to verify that

$$g = d\theta \otimes d\theta + \sin^2 \theta \, d\phi \otimes d\phi$$

satisfies $\text{Ric} = g$ for $0 < \theta < \pi$ and $0 < \phi < 2\pi$.

The metric $g$ is such that we are in the situation considered in Section 4.6; replace $t$ with $\theta$; $x^1$ with $\phi$; $n$ with 1; $f(t)$ with $\sin \theta$; and $\epsilon$ with 1. As in Section 4.6, we also introduce the frame

$$e_0 = \partial_{\theta}, \quad e_1 = \frac{1}{\sin \theta} \partial_{\phi}.$$  

Note that

$$H = -\frac{1}{\sin \theta} \partial_{\theta} \sin \theta = -\cot \theta.$$  

Moreover,

$$\dot{H} = 1 + \frac{\cos^2 \theta}{\sin^2 \theta} = \frac{1}{\sin^2 \theta}.$$  

Appealing to Lemma 76 then yields

$$\text{Ric}(\partial_{\theta}, \partial_{\theta}) = \text{Ric}(e_0, e_0) = \frac{1}{\sin^2 \theta} - \frac{\cos^2 \theta}{\sin^2 \theta} = 1.$$  

Similarly, $\text{Ric}(e_1, e_1) = 1$ and $\text{Ric}(e_0, e_1) = 0$. The proposition follows.

**Exercise 80.** Let $\psi$ be defined by (4.30). Prove that $\psi$ is a diffeomorphism onto its image, that the image of $\psi$ is dense in $S^2$ and that (4.31) holds.
4.7.2 The curvature of the upper half space model of hyperbolic space

Let us define $U^n$ by

$$U^n = \{ x = (x^1, \ldots, x^n) \in \mathbb{R}^n : x^n > 0 \}.$$ 

Moreover, define

$$g_{U^n} = \frac{1}{(x^n)^2} \sum_{i=1}^n dx^i \otimes dx^i.$$ 

Then $(U^n, g_{U^n})$ is called the upper half space model of $n$-dimensional hyperbolic space. Here, we do not sort out the relation between this model and the metric defined in Example 27, but we calculate the Ricci tensor of $g_{U^n}$.

**Lemma 81.** Let $1 \leq n \in \mathbb{Z}$ and let $U^{n+1}$ and $g_{U^{n+1}}$ be defined as above. Then

$$\text{Ric}[g_{U^{n+1}}] = -ng_{U^{n+1}},$$

where $\text{Ric}[g_{U^{n+1}}]$ denotes the Ricci curvature of $g_{U^{n+1}}$.

**Proof.** Let $\psi : \mathbb{R}^{n+1} \to U^{n+1}$ be defined by

$$\psi(x^0, \ldots, x^n) = [x^1, \ldots, x^n, \exp(x^0)].$$

Then $\psi$ is a diffeomorphism from $U^{n+1}$ to $\mathbb{R}^{n+1}$ and

$$g = \psi^* g_{U^{n+1}} = dx^0 \otimes dx^0 + e^{-2x^0} \sum_{i=1}^n dx^i \otimes dx^i.$$ 

Denoting $x^0$ by $t$, introducing $f$ by $f(t) = e^{-t}$, and letting $\epsilon = 1$, we are exactly in the situation considered in Section 4.6. Compute

$$H = -\frac{\dot{f}}{f} = 1.$$ 

Lemma 76 then yields $\text{Ric} = -ng$. The lemma follows. 

\qed
Bibliography

Index

Basis
orthonormal, 2

Causal
curve, 11
geodesic, 22
tangent vector, 11
vectorfield, 11
Causal character
vector, 4
Christoffel symbols, 18
Complete
desic, 22
semi-Riemannian manifold, 22

Connection, 15
Levi-Civita, 16
metric, 16
torsion free, 16

Connection coefficients, 27
Contraction, 10
Covariant derivative, 15

Curve
causal, 11
lightlike, 10
spacelike, 10
timelike, 10

Divergence, 29

Euclidean
metric, 8
scalar product, 1

Euclidean metric, 8
scalar product space, 1

Future oriented
timelike vectors, 6

Geodesic, 20
causal, 22
complete, 22
lightlike, 22
maximal, 22
null, 22
spacelike, 22
timelike, 22

Gradient, 29

Group of isometries, 9

Hyperbolic metric on $H^n$, 8
Hyperbolic space
upper half space model, 36

Index
metric, 7
scalar product, 2

Induced
metric, 8

Inertial frames, 13

Koszul formula, 16

Laplacian, 29
Levi-Civita connection, 16

Lightlike
curve, 10
desic, 22
subspace, 5
tangent vector, 10
vector, 4
vector field, 10

Linear isometry, 3

Lorentz
manifold, 7
metric, 7
scalar product, 3

Lorentz manifold
time orientable, 10
time oriented, 10

Manifold
Lorentz, 7
Riemannian, 7
semi-Riemannian, 7
Time orientation, 10

Maximal geodesic, 22

Metric, 7
cconnection, 16
Euclidean, 8
index, 7
induced, 8
Lorentz, 7
Minkowski, 8
INDEX

Riemannian, 7
    semi-Riemannian, 7
    time orientable, 10
Minkowski
    metric, 8
    scalar product, 1
    scalar product space, 1
Musical isomorphisms, 10

Non-degenerate
    subspace of a scalar product space, 4
Null
    geodesic, 22
    vector, 4
Opposite timecone, 5
Orthonormal
    basis, 2
Parallel
    translation along a curve, 20
    vector field along a curve, 19
Preserving scalar products, 3
Ricci tensor, 27
Riemann curvature tensor, 25
Riemannian
    manifold, 7
    metric, 7
    scalar product, 3
    scalar product space, 3
Round metric on $\mathbb{S}^n$, 8
Scalar curvature, 28
Scalar product, 1
    Euclidean, 1
    index, 2
    Lorentz, 3
    Minkowski, 1
    non-degenerate, 1
    Riemannian, 3
    symmetric, 1
Scalar product space, 1
    Euclidean, 1
    Minkowski, 1
    Riemannian, 3
    time orientation, 6
Semi-Riemannian
    manifold, 7
    metric, 7
    submanifold, 8
Semi-Riemannian manifold
    complete, 22
Spacelike
    curve, 10
    geodesic, 22
    submanifold, 11
    subspace, 5
    tangent vector, 10
    vector, 4
    vector field, 10
Spacetime, 10
Submanifold
    semi-Riemannian, 8
    spacelike, 11
Subspace
    lightlike, 5
    spacelike, 5
    timelike, 5
Tangent vector
    causal, 11
    lightlike, 10
    spacelike, 10
    timelike, 10
Time orientable
    Lorentz manifold, 10
    metric, 10
Time orientation
    manifold, 10
    scalar product space, 6
Time oriented
    Lorentz manifold, 10
Timecone, 5
    opposite, 5
Timelike
    curve, 10
    geodesic, 22
    subspace, 5
    tangent vector, 10
    vector, 4
    vector field, 10
Timelike vectors
    future oriented, 6
Torsion free
    connection, 16
Upper half space model, 36
Vector
    causal character, 4
    lightlike, 4
    null, 4
    spacelike, 4
    timelike, 4
Vector field
    lightlike, 10
    spacelike, 10
    timelike, 10
Vectorfield
   causal, 11

Warped product, 11