Notes on Differential Geometry, 1

Lars Andersson

This is the first part of my notes for the course "Elementary Differential Geometry".

For more details, see for example

Armstrong, Basic Topology

Halmos, Finite Dimensional Vectorspaces

Rudin, Principles of Mathematical Analysis

Spivak, Calculus on Manifolds

Sternberg, Lectures on Differential Geometry (ch. 1 for multilinear algebra)

1. Some Background Material

1.1. **Linear Algebra.** Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$ and let V be a vector space over $\mathbb{F}$. A finite dimensional vector space over $\mathbb{R}$ ($\mathbb{C}$) is isomorphic to $\mathbb{R}^n$ ($\mathbb{C}^n$) for some n, where $\mathbb{R}^n = \{x = (x^1, \dots, x^n)\}$ for $x^i \in \mathbb{R}$ and $\mathbb{C}^n = \{z = (z^1, \dots, z^n)\}$ for $z^i \in \mathbb{C}^n$.

On $\mathbb{R}^n$, the standard Euclidean inner product is given by $\langle x, y\rangle = \sum x^i y^i$ and the Euclidean norm is given by $||x|| = \langle x, x\rangle^{\frac{1}{2}}$. On $\mathbb{C}^n$ the standard Hermitean structure is given by $\langle z, w\rangle = \sum z^i \bar{w}^i$ and similarly the corresponding norm is given by $||z|| = \langle z, z\rangle^{\frac{1}{2}}$. Note that in the case of the Hermitean structure the inner product is a *sesquilinear form*, i.e. $\langle az, w\rangle = a\langle z, w\rangle$ and $\langle z, aw\rangle = \bar{a}\langle z, w\rangle$ for $a \in \mathbb{C}$.

Let V, W be vector spaces. A mapping $A : V \to W$ is said to be linear if for all $a, b \in \mathbb{F}$ and $v, w \in V$,

$$A(av + bw) = aA(v) + bA(w).$$

Let V be a vector space. The dual space V^* is defined to be the space of linear functionals (i.e. $\mathbb{F}$-valued linear functions) on V. We will sometimes use the notation $v^*(v) = \langle v, v^*\rangle$, i.e. we will use the $\langle\cdot,\cdot\rangle$ to denote the duality pairing between V^* and V.

Note that $(V^*)^* = V$ if V is finite dimensional (reflexivity) and it is natural therefore to consider the elements of V as linear functionals on V^*.

In the case $V = \mathbb{R}^n$ one shows that $V^* = \mathbb{R}^n$ (linear isomorphism).

Let $\{e_1, \dots, e_n\}$ be a basis of V. Then there is a *unique* dual basis $e^1, \dots, e^n$ of V^* such that $e^i(e_j) = \delta^i_j$ where δ^i_j is the Kronecker delta, i.e. $\delta^i_j = 1$ if and only if $i = j$ and $= 0$ otherwise. Then for $v \in V$ we can write

$$v = \sum_i e^i(v) e_i = \sum_i v^i e_i \tag{1}$$

where the $v^i := e^i(v)$ are the (linear) *coordinate functions* on V defined w.r.t. the basis e_i. Similarly, *and this is important* for $v^* \in V^*$, we can write

$$v^* = \sum_i e_i(v^*)e^i = \sum_i v_i^* e^i \tag{2}$$

where the v_i^* are the coordinate functions on V^* defined w.r.t. the basis e_i.
Note: the placing of the indices here (the index up on the dual basis, the index down on the basis elements) is significant. Later we will work with multiindex objects like $R_{abc}{}^d$ and in general the indices down correspond to factors of V^* in a tensor product, while indices up correspond to factors of V in a tensor product, more about this later.
Note: In differential geometry a basis of a tangent space is often called a frame. Note also that the summation in (1) and in (2) is always over an index which is *repeated upstairs and downstairs*! This leads to the *Einstein summation convention* which is to interpret an expression like $T_{ij}S^{jk}$ as $\sum_j T_{ij}S^{jk}$ where the in the summation j runs over its *range*, for example if j is an index associated with a basis on $\mathbb{R}^n$ then its range might be $\{1, \ldots, n\}$. More about this later.

1.1.1. *Metrics.* Let V, W and Z be vector spaces over $\mathbb{R}$. A mapping $A : V \times W \to Z$ is said to be bilinear if it is linear in each argument separately, i.e. if

$$A(a_1v_1 + a_2v_2, w) = a_1A(v_1, w) + a_2A(v_2, w) \text{ for all } a_i \in \mathbb{F} \text{ and } v_i \in V, w \in W$$

and

$$A(v, b_1w_1 + b_2w_2) = b_1A(v, w_1) + b_2A(v, w_2) \text{ for all } b_i \in \mathbb{F} \text{ and } w_i \in W, v \in V$$

Expanding this means that

$$A(a_1v_1 + a_2v_2, b_1w_1 + b_2w_2) = a_1b_1A(v_1, w_1) + a_1b_2A(v_1, w_2) + a_2b_1A(v_2, w_1) + a_2b_2A(v_2, w_2)$$

Exercise 1.1. *Prove that the Euclidean inner product $\langle \cdot, \cdot \rangle$ is a bilinear mapping $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$.* □

A bilinear map which takes values in $\mathbb{R}$ (or more generally, in $\mathbb{F}$) is often called a bilinear form.
A bilinear map $A : V \times V \to \mathbb{R}$ is said to be symmetric if $A(v, w) = A(w, v)$ for all $v, w \in V$. A is said to be nondegenerate if $A(v, w) = 0$ for all $w \in V$ implies $v = 0$.
A symmetric nondegenerate bilinear form is called an *inner product* or *metric*. We will often use the notation $\langle v, w \rangle$ for an inner product or metric.
Note: The terminology "inner product" is standard in linear algebra. However, here we will be interested in inner products defined on the tangent spaces of a manifold and these correspond to "metrics". Therefore we will often use this terminology when discussing pure linear algebra.
A metric $\langle \cdot, \cdot \rangle$ is positive definite if $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.

Let $\{e_i\}$ be a basis of V. Then to a metric $\langle\cdot,\cdot\rangle$ there corresponds a symmetric invertible matrix $g_{ij} = \langle e_i, e_j\rangle$. Then we can write (using the summation convention and expanding v and w in terms of $\{e_i\}$)

$$\langle v, w\rangle = g_{ij} v^i w^j$$

Later we will deal with metrics defined on the tangent spaces of a manifold. Then the g_{ij} will be the metric *tensor field*.
For $v \in V$ there is a naturally defined dual vector given by $v^*(x) = \langle x, v\rangle$. Then we can write (using the summation convention)

$$v^*(x) = v^*_i e^i(x^j e_j) = x^i v^j g_{ij}$$

and thus

$$v^*_i = g_{ij} v^j.$$

Let the Kronecker delta δ^i_j be defined by

$$\delta^i_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Define g^{ij} by $g^{ij} g_{jk} = \delta^i_k$, then g^{ij} is the inverse of g_{ij} and we have $v^j = g^{ij} v^*_i$. Thus we can use g_{ij} to lower indices and g^{ij} to raise indices. This will be used systematically later.

Exercise 1.2. *Given a nondegenerate symmetric bilinear form $\langle\cdot,\cdot\rangle$ on V, prove the existence of an orthonormal basis $e_1, \ldots e_n$ of V which diagonalizes $\langle\cdot,\cdot\rangle$, i.e. if $x = \sum x^i e_i$*

$$\langle x, x\rangle = -\sum_{i=1}^{k} |x^i|^2 + \sum_{i=k+1}^{n} |x^i|^2$$

□

Here k is called the *index* of the metric. We will mainly be interested in the Riemannian case (when $\langle\cdot,\cdot\rangle$ is positive definite, i.e. when $k = 0$) or the Lorenzian case (when $\langle\cdot,\cdot\rangle$ has index $k = 1$).
In the Lorenzian case it is often useful to number the basis elements $e_0, e_1, \ldots, e_{n-1}$ and let e_0 be a negative basis element, i.e. $\langle e_0, e_0\rangle < 0$. In terms of a normed basis we then get

$$\langle x, x\rangle = -|x^0|^2 + \sum_{i=1}^{n-1} |x^i|^2.$$

Given a metric on V and an orthonormal basis $\{e_i\}$ of V there is a natural dual basis of V^* defined by the relations $e^i(v) = \langle e_i, v\rangle$ for all i and $v \in V$.

1.2. **Topology, continuity etc.** A *topological space* is a set X together with a collection of subsets called *open sets* such that any union of open sets is open, finite intersections of open sets is open and X and the empty set $\emptyset$ are open. The complement of an open set is called closed. The elements of a topological space are often called points.
A collection of open sets $\mathcal{O} = \{O_\alpha\}_{\alpha \in I}$ which cover X (i.e. any element of X is an element of a O_α for some α) is called an *open cover* of X. An open cover of X by a subset of $\mathcal{O}$ is called a refinement.
An open cover $\mathcal{O}$ is called *locally finite* if any point $x \in X$ is a member of at most finitely many of the O_α.
X is called *paracompact* if any open covering has a locally finite refinement.
Given two topological spaces X and Y we say that a mapping $f : X \to Y$ is *continuous* if $f^{-1}(O)$ is open in X for any O open in Y. A map $f : X \to Y$ is said to be a *homeomorhism* if f is continuous, 1-1 and if f^{-1} is continuous.
Let $f : X \to Y$, $g : Y \to Z$ be mappings. Then $g \circ f(x) = g(f(x)$ is the composition mapping $g \circ f : X \to Z$.
Let $f : X \to \mathbb{R}$ be continuous. Then we define the *support* of f, denoted $\mathrm{supp}(f)$ to be the set $\{x : f(x) \neq 0\}$. The analogous definition holds for f with values in a vector space. A function f is said to have *compact support* if the closure $\overline{\mathrm{supp}(f)}$ is compact.
A mapping $f : X \to Y$ is said to be *proper* if $f^{-1}(K)$ is compact for any compact $K \subset Y$. In particular, if X is compact and $f : X \to Y$ proper, then $f^{-1}(p)$ is finite for any $p \in Y$.
A collection of subsets $\mathcal{B}$ of a topological space X is called a *base* for the topology if every open set is a union of members of B.
If X is a topological space with a countable base, then X is said to be *second countable*. A topological space is called Haussdorff if it has the property that two distinct points can be surrounded by disjoint open sets.
A metric on X is a function $d : X \times X \to \mathbb{R}$ on the cartesian product of X with itself such that for any $x, y, z \in X$,

- D1 $d(x, y) \geq 0$, $d(x, y) = 0 \Leftrightarrow x = y$ (nonnegativity)
- D2 $d(x, y) = d(y, x)$ (symmetry)
- D3 $d(x, y) \leq d(x, z) + d(z, y)$ (triangle inequality)

A metric space (X, d) is in a natural sense a Haussdorff topological space if we take as a base for the topology the neighborhoods $\{y \in X \mid d(x, y) < \epsilon\}$ for all $x \in X$ and $\epsilon > 0$.
The typical example of a metric space is $\mathbb{R}^n$ with the metric $d(x, y) = ||x - y||$, where $|| \cdot ||$ is some norm. In the case of metric spaces the concept of continuity corresponds to the ordinary ϵ, δ definition.

1.3. **Differentiability, integration etc.** Consider a map (function) $f : U \to V$ where $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$. We say that f is (Frechet) differentiable at x if there is a linear map $Df(x) : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$||f(y) - f(x) - Df(x)(y-x)|| = o(||y-x||).$$

$Df(x)$ is then called the derivative of f at x. We say that $f : U \to V$ is continuously differentiable if the map $x \mapsto Df(x)$ is continuous as a mapping from $\mathbb{R}^n \to L(\mathbb{R}^n, \mathbb{R}^m)$. Here $L(\mathbb{R}^n, \mathbb{R}^m)$ denotes the space of linear mappings with norm

$$||A||_{L(\mathbb{R}^n,\mathbb{R}^m)} = \sup_{||x||=1} ||Ax||.$$

We will say that $f \in C^k$ as a map from U to V if f has continuous (Frechet) derivatives up to order k and we will say that f is C^∞ if f has continuous derivatives up to *any* order.

In the following we will unless otherwise stated mean C^∞ when we talk about differentable or smooth maps.

The *inverse function theorem* states that if $f : U \to V$ with U, V as above is C^∞ at x_0 then if $Df(x_0)$ is invertible ($\Rightarrow n = m$) then f is a C^∞ homeomorphism (i.e. f has a C^∞ inverse) from a neighborhood of x_0 to a neighborhood of $f(x_0)$.

Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ be C^∞ and write elements of $\mathbb{R}^n \times \mathbb{R}^m$ as (x, y). Then we can write the derivative of f w.r.t. y as $D_y f$. The *implicit function theorem* states that if $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is such that $D_x f(x_0, y_0)$ is invertible at (x_0, y_0), then there is a C^∞ mapping $g : \mathbb{R}^m \to \mathbb{R}^n$ defined in a neighborhood of y_0 (the implicit function) such that $g(y_0) = x_0$ and $f(g(y), y) = f(x_0, y_0)$ for y in a neighborhood of y_0.

1.3.1. *Integration on* $\mathbb{R}^n$. Let $\Omega \subset \mathbb{R}^n$ be an open set, let $\Phi : \Omega \to \mathbb{R}^n$ be a proper C^1 mapping and let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function with compact support in $\Phi(\Omega)$. Let the Jacobean of Φ be the function

$$J_\Phi(x) = \det(D\Phi(x)).$$

Then

$$\int_{\Phi(\Omega)} f(x)dx = \int_\Omega f(\Phi(x))|J_\Phi(x)|dx$$

1.3.2. *Gradients and vector fields on* $\mathbb{R}^n$. Let $V = \mathbb{R}^n$ and let $f : V \to \mathbb{R}$ be a differentiable function. Then we denote the gradient of f by df. It is natural to think of $df(x)$ as an element of the dual space V^*, since if $v \in V$, we have

$$df(v)|_x = \langle df, v\rangle|_x = \frac{d}{dt} f(x+tv)\Big|_{t=0}.$$

From this it is natural to associate to the vector v a 1:st order differential operator v and write $vf = df(v)$. This notation will show up when we deal with functions and vector fields on manifolds.

Note that the operation vf satisfies the rules

- $v(f+g) = vf + vg$ (linearity)
- $v(fg) = (vf)g + f(vg)$ (Leibniz rule)

Now note that if we denote the point in V by x and the coordinate functions defined w.r.t. the basis $\{e_i\}$ by x^i then we have the natural correspondence

$$\partial_{x^i} \leftrightarrow e_i$$

and

$$dx^i \leftrightarrow e^i$$

Thus the partial derivatives in the coordinate directions correspond to the basis elements e_i considered as *differential operators* and the gradients of the coordinate functions x^i correspond to the dual basis elements e^i.
Now we see that the natural expression for df considered as a geometric object is

$$df = \frac{\partial f}{\partial x^i} dx^i$$

This is an example of a differential form of order 1, i.e. a function with values in V^*.
A vector field X on V is simply a function on V with values in V. If $\{e_i\}$ is a basis on V then we can write $X(x) = \sum_i X^i(x)e_i$ or using the Einstein summation convention, suppressing the reference to x, and using the ∂_{x^i} instead of e_i we can write $X = X^i \partial_{x^i}$.
Let X be a vector field on V. Then we can solve the system of ODE's

$$\dot{x} = X(x)$$

(explicitely, $\frac{d}{dt} x^i = X^i(x), i = 1, \ldots, n$) locally. This defines for some sufficiently small $\epsilon > 0$ a differentiable flow $\Phi_{X,t}$, $-\epsilon < t < \epsilon$ such that $t \mapsto \Phi_{X,t}(x)$ is the solution curve for the above ODE.
Then we can define the derivative of a function in the direction of X as

$$Xf(x) = \frac{d}{dt} f(\Phi_{X,t}(x)) = \frac{d}{dt} \Phi_{X,t}^* f(x)$$

where $\Phi^* f = f \circ \Phi$ denotes the pullback of f by Φ.

Exercise 1.3. *Check that* $Xf = \sum_i X^i \frac{\partial f}{\partial x^i}$ *and* $Xf = \langle X, df \rangle$ *where* $\langle \cdot, \cdot \rangle$ *denotes the pairing between* V *and* V^*. □

Let X, Y be vector fields. Then the *Lie bracket* $[X, Y]$ of X with Y is a vector field defined by its action on functions by its

$$[X, Y]f = X(Yf) - Y(Xf)$$

A computation shows that with $X = X^i \partial_{x^i}$ and $Y = Y^i \partial_{x^i}$ (using the summation convention) we have

$$[X, Y] = \sum_i \sum_j \left(\frac{\partial Y^i}{\partial x^j} X^j - \frac{\partial X^i}{\partial x^j} Y^j \right) \partial_{x^i}$$

The Lie bracket $[X, Y]$ is an example of the Lie derivative, $\mathcal{L}_X Y = [X, Y]$. The Lie derivative can be defined on any tensor field, as we shall see later. It can be considered to be a derivative along the flow of X.

Exercise 1.4. *Prove that*

$$[X,Y] = \frac{d}{dt}\Phi^*_{X,t}Y\Big|_{t=0}$$

where $\Phi^*Y = (D\Phi)^{-1} \circ Y \circ \Phi$ *is the pullback on vector fields.* □

Exercise 1.5. *Prove the Jacobi identity*

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$$

□

On V it is clear that we can define partial differentiation on vector fields by operating on the components. We can write this as

$$D_X Y = \sum_i \sum_j X^j (\partial_{x^j} Y^i) \partial_{x^i}$$

Then it is easy to check the following rules: Let $f \in C^\infty$ and let X, Y be vector fields on V, then

- $D_X Y - D_Y X = [X,Y]$ (torsion free)
- $D_X(fY) = (Xf)Y + f D_X Y$ (Leibniz rule)
- $D_{fX} Y = f D_X Y$

Later we will define the notion of torsion free *covariant derivative* or connection on manifolds. This is a 1:st order differential operator which satisfies the above rules. The important difference is that for the *partial derivative* on $\mathbb{R}^n$ it holds that

$$D_X D_Y Z - D_Y D_X Z - D_{[X,Y]} Z = 0 \tag{3}$$

while for a general connection, this is no longer true. This leads us to introduce the so called curvature tensor. Let ∇ denote a connection and let X, Y, Z be vector fields. Then define the vector field $R(X,Y)Z$ by

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

$R(X,Y)Z$ is the (3,1)-form of the so called Riemann curvature tensor.

2. Differentiable Manifolds

A manifold of dimension n is a space which is locally like the Euclidean space $\mathbb{R}^n$. Formally,

Definition 2.1. *A topological manifold modelled on* $\mathbb{R}^n$ *is a second countable paracompact Haussdorff space* M *together with an open cover* $\{O_\alpha\}$ *and a system of maps* $\psi_\alpha : O_\alpha \to U_\alpha = \psi_\alpha(O_\alpha) \subset \mathbb{R}^n$ *such that each* ψ_α *is a homeomorphism from* O_α *onto* U_α.
The cover $\{O_\alpha\}$ *together with the maps* ψ_α *is called an* atlas *and the maps* ψ_α *are called* charts.

A differentiable manifold is a topological manifold M together with a *differentable structure*, i.e. a way of defining differentiable functions on M. The natural way of doing this is to use the charts ψ_α to transfer the definition of differentiable functions from $\mathbb{R}^n$ to M.
A function $f : M \to \mathbb{R}$ defined in O_α can be considered also as a function on $\mathbb{R}^n$ by looking at $f_\alpha = f \circ \psi_\alpha^{-1} \,|\, \psi_\alpha(O_\alpha) \to \mathbb{R}$ and it is natural to say that f is differentiable on M if the f_α are differentiable for all α. In order to make this definition consistent however, we need to assume that the mappings

$$\psi_{\alpha'} \circ \psi_\alpha^{-1} : \psi_\alpha(O_\alpha \cap O_{\alpha'}) \to \psi_{\alpha'}(O_\alpha \cap O_{\alpha'})$$

are differentiable whenever $O_{\alpha'}$ and O_α have nonempty intersection. This leads to the following definition:

Definition 2.2. *A topological manifold M with atlas $\{(O_\alpha, \psi_\alpha)\}$ is a* differentiable manifold *if*

$$\psi_{\alpha'} \circ \psi_\alpha^{-1} : \psi_\alpha(O_\alpha \cap O_{\alpha'}) \to \psi_{\alpha'}(O_\alpha \cap O_{\alpha'})$$

is a C^∞ mapping whenever O_α and $O_{\alpha'}$ have nonempty intersection.

Let M be a differentiable manifold. Then it makes sense to define differentiable functions on M as above.

Examples:

- The standard sphere $S^n = \{x \in \mathbb{R}^{n+1} \,|\, ||x|| = 1\}$ where $||x||^2 = \sum x_i^2$ is the Euclidean norm. S^n is given a topology by considering it as a subspace of $\mathbb{R}^{n+1}$.

  A C^∞ atlas on S^n can be defined by taking the open cover defined by $O_1 = \{x \in S^n | x_1 > -\epsilon\}$ and $O_2 = \{x \in S^n | x_1 < \epsilon\}$ for some small $\epsilon > 0$. Then O_1 and O_2 are neoghborhoods of the upper and lower hemisphere, respectively. These subsets of S^n are in a natural way C^∞ homeomorphic to open balls in $\mathbb{R}^n$ and S^n is a C^∞ manifold of dimension n.
- Let M and N be C^∞ manifolds. Then the cartesian product $M \times N$ is a C^∞ manifold.
- The torus $T^n = S^1 \times S^1 \times \cdots \times S^1$ (the cartesian product of n copies).

Exercise 2.1.

(a) *Construct an explicit atlas for T^2.*
(b) *Let M and N be differentiable manifolds. Construct an atlas for $M \times N$.*

□

Let M and N be differentiable manifolds of dimension m and n, repsectively and let $\phi : M \to N$ be a mapping. Then we say that ϕ is differentiable at $x \in M$ if $\psi_{N,\alpha} \circ \phi \circ \psi_{M,\beta}^{-1}$ is differentiable for any charts $(O_{N,\alpha}, \psi_{N,\alpha})$ of N and $(O_{M,\beta}, \psi_{M,\beta})$ of M, covering x and $\phi(x)$.
Let $\mathcal{F}_M$ and $\mathcal{F}_N$ denote the space of differentiable functions defined on open sets on M and N respectively. Then we can define the *pullback* on functions under ϕ

by $\phi^* f(x) = f(\phi(x))$ and $\phi^*\mathcal{F}_N = \{\phi^* f | f \in \mathcal{F}_N\}$. Then ϕ is differentiable if and only if $\phi^*\mathcal{F}_N \subset \mathcal{F}_M$.

The *rank* of a differentiable mapping ϕ at $x \in M$ is defined as the rank of the differentiable mapping $\psi_{N,\alpha} \circ \phi \circ \psi_{M,\beta}^{-1}$ for charts covering x and $\phi(x)$.

A differentiable mapping $\phi : M \to N$ which has rank m is called an *immersion*.

A differentiable mapping $\phi : M \to N$ which is a homeomorphism and which has a differentiable inverse is called a *diffeomorphism*. This implies in particular that $\dim M = \dim N$.

A group G which is also a C^∞ manifold and such that the group operation $(g_1, g_2) \mapsto g_1 g_2$ is a C^∞ mapping $G \times G \to G$ is called a *Lie group*. The following are examples of Lie groups.

- Let $Gl(n)$ be the group of invertible (real) matrices of order n. Then $Gl(n)$ is a manifold of dimension n^2. To see this intuitively recall that invertibility for a matrix is stable under small perturbations. This means that any $A \in Gl(n)$ has a neighborhood homeomorphic to an open subset in the space $\mathcal{M}(n)$ of *all* matrices of order n. But $\mathcal{M}(n)$ is isomorphic to the space $\mathbb{R}^{n^2}$.
- Let $O(n)$ be the group of orthogonal matrices with unit determinant. Then $O(n)$ is a manifold of dimension $n(n-1)/2$. It is not connected in general (why?).
- Consider $\mathbb{R}^n$ and let $< \cdot, \cdot >$ be an inner product of index k, i.e. a nondegenerate symmetric bilinear form with k negative and k positive eigenvalues. Then we can define $O(n,k)$ to be the group of linear isometries w.r.t. $< \cdot, \cdot >$. In particular, the group $O(3,1)$ is called the *Lorenz group*. This is a Lie group of dimension 6.
- $Sl(n, \mathbb{C})$ is the group of complex matrices of order n with unit determinant. The real dimension is $2(n^2 - 1)$. In particular $Sl(2, \mathbb{C})$ has real dimension 6, the same as the Lorenz group $O(3,1)$. We will see later that $Sl(2, \mathbb{C})$ is a double covering of $O(3,1)$.

2.1. **Coordinates.** Let $x \in M$ and let ψ_α be a chart covering x. Then if x^i are coordinate functions on $\mathbb{R}^n$ we can transfer these from U_α to O_α by letting

$$x^i(x) = x^i(\psi_\alpha(x))$$

Such coordinates are called local coordinates defined by the chart ψ_α.

This allows us to express functions on M in terms of coordinates: $f(x) = f(x^1, \ldots, x^n)$.

2.2. **Tangent space.** Recall the relation between 1:st order derivatives on $\mathbb{R}^n$ and vectors in $\mathbb{R}^n$. Let $x \in M$. Then we define the *tangent space* of M at x as the space of 1:st order derivatives at x. This means the space of mappings which are linear and satisfies the Leibniz rule

- $v(fg)(x) = (vf)(x)g(x) + f(x)vg(x)$

It is not hard to show that if M is n-dimensional then this is a space isomorphic to $\mathbb{R}^n$ and we denote this space by T_xM.
An element of T_xM is called a tangent vector at x.
Note: For some reason Spivak uses the notation M_x for the tangent space.
Note: If we consider $M = \mathbb{R}^n$ as a manifold, then at $x \in M$, we have $T_xM = M$ by the correspondence between tangent vectors and 1:st order differental operators discussed earlier. Thus when the underlying space is $\mathbb{R}^n$, there is a *global* identification of the tangent space with $\mathbb{R}^n$. It is important to realize that this does not hold in general.
Now let $x \in M$ and consider a chart ψ_α with corresponding local coordinates $(x^1_\alpha, \ldots, x^n_\alpha)$. Then we can write $f(x) = f(x^1_\alpha, \ldots, x^n_\alpha)$ and from the relation between 1:st order differential operators and tangent vectors we find that any element $X \in T_xM$ can be written $X = X^i_\alpha \partial_{x^i_\alpha}$.
Here we make the dependence on the chart explicit by keeping the α.
In terms of another chart ψ_β at x with corresponding coordinate system $(x^1_\beta, \ldots, x^n_\beta)$ we similarly get $X = X^i_\beta \partial_{x^i_\beta}$.
Thus there basis (frame) of T_xM given by ∂_{x^i}, which is determined by a choice of local coordinates (chart). The frame depends in a simple way on the chart: From the chain rule we have the following relation

$$\partial_{x^i_\alpha} = \sum_j \frac{\partial x^j_\beta}{\partial x^i_\alpha} \partial_{x^j_\beta}$$

which gives

$$X = X^i_\alpha \partial_{x^i_\alpha} = X^j_\beta \partial_{x^j_\beta}$$

with

$$X^j_\beta = X^i_\alpha \frac{\partial_{x^j_\beta}}{\partial x^i_\alpha}$$

Notes on Differential Geometry 2

Lars Andersson

Now we describe the remaining aspects of the coordinate dependence of vectors and covectors, which motivates the notion of tensors on manifolds. Then, after introducing lots of index expressions (the "debauch of indices"), we will start to introduce notation which allow us to describe geometric quantities in invariant form, both with and without indices.

In this part of the notes, I will discuss a number of concepts without giving full details, to give a "map" of the field we are discussing. Some of this we will come back to in fuller detail, but some of these can be discussed properly only in a course devoted exclusively to Differential Geometry.

2.3. **Transition function.** Recalling the relation between charts and coordinates, $x^i(p) = \psi^i(p)$, we find the correspondence $\psi_\beta \circ \psi_\alpha^{-1} \leftrightarrow x^i_\beta(x^j_\alpha)$ so we can understand that

$$D(\psi_\beta \circ \psi_\alpha^{-1}) \leftrightarrow \frac{\partial x^i_\beta}{\partial x^j_\alpha}$$

Thus to the atlas $\{\psi_\alpha\}$ there corresponds a system of matrices $T_{\beta\alpha} = D(\psi_\beta \circ \psi_\alpha^{-1}) : U_\alpha \cap U_\beta \to Gl(n)$. These satisfy the relations $T_{\alpha\beta} = T^{-1}_{\beta\alpha}$ and $T_{\alpha\beta}T_{\beta\gamma} = T_{\alpha\gamma}$ (the *cocycle conditions*).

2.4. **Cotangent Space.** The tangent space $T_x M$ at $x \in M$ is naturally defined. The dual space $T^*_x M$ is called the cotangent space at x. As in the case of $\mathbb{R}^n$, there is a natural dual basis to $\{\partial_{x^i}\}$ given by the gradients of the coordinate functions: $\{dx^i\}$.Let $\xi \in T^*_x M$, i.e. $\xi : T_x M \to \mathbb{R}$ is a linear functional. Then $\xi = \sum_i \xi_i dx^i$ or if we make the dependence on the chart ψ_α explicit:

$$\xi = \sum_i \xi_{\alpha,i} dx^i_\alpha.$$

Since dx^i_α is a dual basis to $\partial_{x^i_\alpha}$, we have

$$\delta^i_j = dx^i_\alpha \partial_{x^j_\alpha} \tag{4}$$

This relation can be used to determine the transformation rules for the dx^i:

$$\begin{aligned}
\delta^i_j &= dx^i_\alpha \partial_{x^j_\alpha} \\
&= dx^i_\alpha \frac{\partial x^k_\beta}{\partial x^j_\alpha} \partial_{x^k_\beta} \\
&= dx^\ell_\beta B^i_\ell \frac{\partial x^k_\beta}{\partial x^j_\alpha} \partial_{x^k_\beta} \\
&= \delta^\ell_k B^i_\ell \frac{\partial x^k_\beta}{\partial x^j_\alpha}
\end{aligned}$$

where we used that the relation (4) holds also in the chart ψ_β. This determines the matrix

$$B^i_\ell = \left(\frac{\partial x^\ell_\beta}{\partial x^i_\alpha}\right)^{-1} = \frac{\partial x^i_\alpha}{\partial x^\ell_\beta}$$

which means that

$$dx^i_\alpha = \frac{\partial x^i_\alpha}{\partial x^\ell_\beta} dx^\ell_\beta$$

This means that

$$\xi = \xi_{\alpha,i} dx^i_\alpha = \xi_{\alpha,i} dx^\ell_\beta \frac{\partial x^i_\alpha}{\partial x^\ell_\beta} = \xi_{\beta,\ell} dx^\ell_\beta$$

which gives the transformation rule

$$\xi_{\beta,\ell} = \xi_{\alpha,i} \frac{\partial x^i_\alpha}{\partial x^\ell_\beta}$$

If we let $X \in T_x M$, then $\xi(X) \in \mathbb{R}$ should be independent of the choice of chart ψ_α (or equivalently of the coordinate system $x^1_\alpha, \dots, x^n_\alpha$). We this fact to check the above calculations:

$$\begin{aligned} \xi(X) &= \xi_{\alpha,i} X^i_\alpha = \xi_{\beta,\ell} \frac{\partial x^\ell_\beta}{\partial x^i_\alpha} X^k_\beta \frac{\partial x^i_\alpha}{\partial x^k_\beta} \\ &= \xi_{\beta,\ell} X^k_\beta \delta^\ell_k = \xi(X) \end{aligned}$$

Here we note an *important fact:* the quantity $\xi(X)$ is *independent of the coordinate system*, in other words, it transforms as a function. This is the fundamental fact which underlies the definition of a tensor on a manifold.

2.5. **Tangent and Cotangent bundles.** Let

$$TM = \cup_{x \in M} T_x M$$

and

$$T^*M = \cup_{x \in M} T^*_x M$$

where the unions are disjoint. We will call TM the *tangent bundle* of M and T^*M the *cotangent bundle* of M. It is not hard to show that these are C^∞ manifolds of dimension $2n$. For example coordinates on TM can be given by $(x^1, \dots, x^n, X^1, \dots X^n)$.

From the definition of TM as disjoint union, it is natural to associate to a point $z \in TM$ a point $x \in M$. This can be described as follows: there is a C^∞ map $\pi : TM \to M$. Further, one may show that

$$\pi^{-1}(O_\alpha) \cong O_\alpha \times \mathbb{R}^n$$

(TM is locally trivial).

This together with the cocycle conditions described above, show that TM is a vector bundle with *fiber* $\mathbb{R}^n$ and *structure group* $Gl(n)$.

Similar facts hold for T^*M.
In general, a *vector bundle* is a manifold B called the *bundle manifold*, a manifold M called the *base manifold* and a map $\pi : B \to M$ such that $\pi^{-1}(x) = \mathbb{R}^m$ for some m (plus some more conditions like local triviality and cocycle conditions etc. which we won't discuss now).
A *section* of a bundle B is a C^∞ map $\xi : M \to B$, so that $\pi \circ \xi(x) = x$.
More about this later.

2.6. **Vector fields and 1-forms.** Now consider TM and T^*M as bundles. Then we see that a function which assigns $x \to X_x \in T_xM$ is a *section* of TM. This is called a vector field. Similarly, a function $x \to \xi_x \in T^*_xM$ is a section of T^*M, called a 1-form.
These are the most elementary examples of tensor fields, which we will discuss next.

2.7. **Frames.** A collection of vectorfields $X_1, \dots, X_n$ defined on an open set $O \subset M$ is called a *frame* if at each $x \in O$, the collection $X_1, \dots, X_n$ is a basis of T_xM. Considering the situation on $\mathbb{R}^n$ one sees that the space of frames at x is isomorphic to $Gl(n)$.

3. Tensors

3.1. **Multilinear Algebra.** The first nontrivial concept that we need to introduce is that of *tensor* and tensor product. One might say that the concept of tensor is a systematic way of dealing with multilinear mappings. Here we will discuss how this arises in a purely linear algebra context. Later we will introduce tensors on manifolds. This is easy once the linear algebra background is understood properly.
The following is slightly more sophisticated than we'll need at the moment. This section can be skipped on the first reading
Let $V_1, V_2, \dots, V_k$ and W be vector spaces. A mapping $A : V_1 \times V_2 \times \cdots \times V_k \to W$ is said to be *multilinear* if it is linear in each argument separately.
NOTE: Usually we associate the Cartesian product space with a vector space structure in the following way: For example it seems natural to think of the Cartesian product $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ as the vector space $\{x = (x^1, \dots, x^{n_1}, x^{n_1+1}, \dots, x^{n_1+n_2})\}$. However, since a multilinear mapping is NOT LINEAR in general *Exercise: Why??* it is not correct to think of a multilinear mapping A as defined on the *vector space* given by the Cartesian product $V_1 \times \cdots \times V_k$. However, it is natural to associate to A a linear mapping $\bar{A}$ on the free vector space $\bar{U}$, generated by the elements of $V_1 \times \cdots \times V_k$, by defining $\bar{A}(v_1, \dots, v_k) = A(v_1, \dots, v_k)$ and extending linearly.
Consider a multilinear mapping $A : V_1 \times V_2 \times \cdots \times V_k \to W$. Then it is easy to see that the associated mapping $\bar{A}$ defined on $\bar{U}$ vanishes on the subspace $\bar{R}$ of

$\bar{U}$ spanned by elements of the form

$$a(v_1, v_2, \dots, v_k) - (v_1, v_2, \dots, av_j, \dots, v_k), 1 \leq j \leq k \qquad \text{for } a \in \mathbb{F}$$

and

$$(v_1, \dots, v_i + w_i, \dots, v_k) - (v_1, \dots, v_i, \dots, v_k) - (v_1, \dots, w_i, \dots, v_k)$$

Thus $\bar{A}$ can be considered as defined on the quotient space $\bar{U}/\bar{R}$. This space is the *tensor product space* $V_1 \otimes \cdots V_k$.
Thus we can relate to the *multilinear* mapping $A : V_1 \times \cdots \otimes V_k \to W$ a *linear* mapping (which we also denote by A,

$$A : V_1 \otimes \cdots \otimes V_k \to Z.$$

Let now $(v_1, \dots, v_k) \in V_1 \times \cdots \times V_k$ (i.e. $v_i \in V_i$). Then we can consider $(v_1, \dots, v_k)$ as an element in $\bar{U}$ and we will denote the coset of $(v_1, \dots, v_k)$ in $\bar{U}/\bar{R}$ by $v_1 \otimes \cdots \otimes v_k$. Thus the elements of $V_1 \otimes \cdots \otimes V_k$ are of the form $v_1 \otimes \cdots \otimes v_k$.
Given a basis $\{e_{i,j}\}$ for each V_j, a basis for the tensor product space $V_1 \otimes \cdots \otimes V_k$ is given by elements of the form

$$e_{i_1,1} \otimes \cdots \otimes e_{i_k,k}$$

and thus we find that the dimension of $V_1 \otimes \cdots \otimes V_k$ is given by $\prod_{j=1}^k \dim V_j$ (*Note:* $\dim V_1 \times \cdots \times V_k = \sum_{j=1}^k \dim V_j$).
Note: the definition of tensor product depends on the field $\mathbb{F}$. Thus for example $\mathbb{C} \otimes \mathbb{C} = \mathbb{C}$ when the tensor product is defined over $\mathbb{C}$ but the *real* dimension of $\mathbb{C} \otimes \mathbb{C}$ is 4 when the tensor product is taken w.r.t. $\mathbb{R}$. In general we have that $\mathbb{F} \otimes V = V$ for a vector space V defined over $\mathbb{F}$.
Example: $L(V, W)$, the space of linear mappings $V \to W$ is isomorphic to $W \otimes V^*$. In particular, $L(\mathbb{R}^n, \mathbb{R}^m)$, the space of linear mappings $\mathbb{R}^n \to \mathbb{R}^m$ is isomorphic to $\mathbb{R}^m \otimes (\mathbb{R}^n)^* = \mathbb{R}^m \otimes \mathbb{R}^n = \mathbb{R}^{mn}$. This is the representation of $L(\mathbb{R}^n, \mathbb{R}^m)$ as $m \times n$ matrices.
If V is a vector space then V^* is the space of linear functionals on V. This means that if we have a multilinear mapping $A : V_1 \times \cdots \times V_k \to \mathbb{F}$, then A is a linear mapping $A : V_1 \otimes \cdots \otimes V_k \to \mathbb{F}$ and thus A can be considered as an element of $(V_1 \otimes \cdots V_k)^* = V_1^* \otimes \cdots \otimes V_k^*$.
Example: Let $V = \mathbb{R}^n$ and let $\{e_i\}$ and $\{e^i\}$ be a basis on V and its dual basis. Then a basis for $T = V \otimes V^* \otimes V \otimes V$ is given by $\{e_{i_1} \otimes e^{i_2} \otimes e_{i_3} \otimes e_{i_4}\}$ and an element $t \in T$ can be expanded in the form

$$t = t^{i_1}{}_{i_2}{}^{i_3 i_4} e_{i_1} \otimes e^{i_2} \otimes e_{i_3} \otimes e_{i_4}$$

Thus we can represent the *tensor* t w.r.t. the basis $\{e_i\}$ and its dual basis as the indexed object $t^{i_1}{}_{i_2}{}^{i_3 i_4}$, a kind of generalized matrix.
A tensor field on a manifold can be said to give, when evaluated at a point, an object of precisely the type that I have just described.

3.2. **Tensor fields.** Let M be a C^∞ manifold. A tensor field t of type (k,l) on M can be considered as a function which assigns to each $x \in M$, a multilinear map

$$t : T_x^*M \times \cdots \times T_x^*M \times T_xM \times \cdots \times T_xM$$

(k copies of T_x^*M, l copies of T_xM, the order in which the T_x^*M and T_xM occur is irrelevant at the moment) such that if we take k co-vectorfields $\xi_1, \ldots, \xi_k$ and l vector fields $X_1, \ldots, X_l$, the expression

$$x \to t(\xi_1, \ldots, \xi_k, X_1, \ldots, X_l)$$

is a *function*. Recalling the transformation rules for the components of vectors and co-vectors and the expression for a tensor in terms of a basis we easily find the transformation rules for a tensor of type (k,l):

$$(t_\beta)^{i_1,\ldots,i_k}{}_{j_1,\ldots,j_l} = (t_\alpha)^{i'_1,\ldots,i'_k}{}_{j'_1,\ldots,j'_l} \cdot \frac{\partial x_\beta^{i_1}}{\partial x_\alpha^{i'_1}} \cdot \cdots \cdot \frac{\partial x_\beta^{i_k}}{\partial x_\alpha^{i'_k}} \cdot \frac{\partial x_\alpha^{j'_1}}{\partial x_\alpha^{j_1}} \cdot \cdots \cdot \frac{\partial x_\alpha^{j'_l}}{\partial x_\alpha^{j_l}}$$

A more sophisticated point of view is to consider tensor fields as multi-linear maps over the ring C^∞. See Penrose-Rindler for this point of view.
A *contravariant tensor* is a tensor field of type $(k,0)$ and a covariant tensor field is a tensor field of type $(0,l)$.
A symmetric tensor (e.g. covariant) is such that $t(X_1, \ldots, X_l) = t(X_{\pi(1)}, \ldots, X_{\pi(l)})$ for any permutation π of $1, \ldots l$.
We will return later to the natural operations on tensors.

3.3. **Curves and vectorfields.** A smooth map $c : [a,b] \to M$ is called a curve. Let $f \in C^\infty(M)$. Then the action $f \to \frac{d}{dt} f(c(t))$ defines a vector $X = \dot{c}$ at $c(t)$. We have

$$Xf = \frac{d}{dt} f \circ c = \sum \partial_{x^i} f \frac{dc^i}{dt}$$

3.4. **Flows and vectorfields.** Locally, to each vector field there exists a flow (i.e. one-parameter group of smooth maps, actually diffeomorphisms in this case) solving the equation $\frac{d}{dt}\phi_t = X \circ \phi_t$. In particular, $\phi_0 = id$, $\phi_{s+t} = \phi_s \circ \phi_t$.

3.5. **Pushforward and pullback.** Let M and N be differentiable manifolds and let $F : M \to N$ be a differentiable map. Given a vector field X on M we define the pushforward F_*X to be the vector field on N given by

$$F_*X\Big|_{F(x)} = DF(x).X = \frac{d}{dt} F(\phi_t(x))\Big|_{t=0}$$

where ϕ_t denotes the flow on M defined by X.
Thus F_* is a generalization of the derivative to manifold-maps. At x, F_* defines a linear map $F_*\Big|_x : T_xM \to T_{F(x)}N$.

Now we can define the pullback of a covariant tensor t on N. This is defined as the tensor field F^*t on M which at $x \in M$ is given by

$$(F^*t)(X_1, \dots, X_l)\Big|_x = t(F_*X_1, \dots, F_*X_l)\Big|_{F(x)}$$

It is a useful exercise to check that this corresponds exactly to the transformation rules for covariant tensors described above.

3.6. **Metrics.** Now we will introduce the fundamental notion in Riemannian geometry, which is also the *physical field* in relativity, namely the metric.

Definition 3.1. *A metric on a manifold M is a covariant tensor field g of order 2 which is symmetric and non degenerate, in the sense that at each $x \in M$,*

$$g(X, Y) = 0, \forall Y \in T_xM \qquad \Leftrightarrow \qquad X = 0$$

at x. Thus g induces on T_xM, a metric in the sense discussed above in the context of $\mathbb{R}^n$. In particular, each metric has a well defined index k. The important cases are:

- *Riemannian, $k = 0$*
- *Lorenzian, $k = 1$.*

In general, an indefinite metric ($k \neq 0$) is called semi-Riemannian.

Using the above notions we are now ready to describe the components of metrics in terms of local coordinates, which is sometimes necessary for calculations. In terms of a coordinate system $x^1, \dots, x^n$ we have

$$g = g_{ij} dx^i dx^j$$

where g_{ij} is a symmetric invertible matrix at each $x \in M$.
It is convenient to denote by g^{ij} the components of the inverse of g, i.e. $g^{ij} g_{jk} = \delta^i_k$.
It is often convenient to use the notation $\langle \cdot, \cdot \rangle$ for the metric on M, thus we can write for example $\langle X, Y \rangle\Big|_x$ for the inner product of X_x with Y_x in terms of the metric defined on T_xM.
In some (especially older) literature, it is common to think of the metric as given by the *line element*

$$ds^2 = g_{ij} dx^i dx^j.$$

Examples: Note in the following that we often use *names* for the coordinates in explicit calculations, such as x, y, z instead of x^1, x^2, x^3. It is important to remember that this is just a matter of notation.

- Euclidean metric on $\mathbb{R}^3$ with Cartesian Coordinates x, y, z: $g_{ij} = \delta_{ij}$, $g = dx^2 + dy^2 + dz^2$.
- Euclidean metric on $\mathbb{R}^3$ in spherical coordinates r, θ, ϕ: $g = dr^2 + r^2(d\theta^2 + \sin^2(\theta) d\phi^2)$.
- Minkowski metric on $\mathbb{R}^4$ with Cartesian coordinates t, x, y, z, $g_{ab} = \mathrm{diag}(-1, 1, 1, 1)$ ($\mathbb{R}^4$ endowed with this metric is called *Minkowski space* and is often written $\mathbb{R}^{3,1}$), $g = -dt^2 + dx^2 + dy^2 + dz^2$.

- Minkowski metric on $\mathbb{R}^4$, after the change of coordinates $(t,x,y,z) \to (t,r,\theta,\phi)$: $g = -dt^2 + dr^2 + r^2(d\theta^2 + \sin^2(\theta)d\phi^2)$.

Exercises. Numbers refer to Wald.

(1) 2.8

(2) Let g be the metric on Minkowski space in the coordinate system t,r,θ,ϕ.

   (a) Compute the form of the metric in the coordinate system t,x,θ,ϕ where $x = 1/r$.

   (b) Introduce coordinates T,ψ by the relations

$$\begin{aligned} t &= T\cosh(\psi) \\ r &= T\sinh(\psi) \end{aligned}$$

   and compute the form of the metric in the coordinate system t,ψ,θ,ϕ.

   (c) Find the expression for the hypersurfaces $T =$ constant and compute the normal vectorfield.

   (d) What is the induced Riemannian metric on the hypersurfaces $\{T = \text{constant}\}$?

   (e) Consider the t,x,θ,ϕ coordinate system. Let $T^2 = t^2 - x^{-2}$ and compute the form of the metric in the T,x,θ,ϕ coordinate system. Compute the induced metric on the $\{T = \text{constant}\}$ hypersurfaces in this coordinate system.

Notes on Differential Geometry, 3

Lars Andersson

Metrics, continued:
On a C^∞ manifold M with atlas $\{(O_\alpha, \psi_\alpha)\}$ there is always a C^∞ *partition of unity* $\{f_\alpha\}$ subordinate to the cover $\{O_\alpha\}$, i.e. $f_\alpha \in C^{\infty(M)}, \forall\alpha$ and for all α we have

(1) $0 \le f_\alpha \le 1$,
(2) $\sum_\alpha f_\alpha \equiv 1$,
(3) $\mathrm{supp} f_\alpha \subset O_\alpha$.

cf. Wald, appendix A.
Let $\{f_\alpha\}$ be a partition of unity on M and let $g_\alpha = \psi_\alpha^* e$ be a metric on O_α where e denotes the standard Euclidean metric on $\mathbb{R}^n$. Then it is easy to prove that $g = \sum f_\alpha g_\alpha$ is a Riemannian metric on M. Thus any C^∞ manifold has a C^∞ Riemannian metric.
Any noncompact manifold has a Lorenz metric. However, a compact manifold has a Lorenz metric if and only if the Euler characteristic $\chi(M) = \sum_{p=0}^n (-1)^p \beta_p$ vanishes, i.e. $\chi(M) = 0$ (see Steenrod, The topology of fibre bundles).
This shows that for example the 3-dimensional sphere S^3 has a Lorenz structure but the 4-dimensional sphere S^4 does not.

3.7. **Tensor product.** Let s and t be tensor fields of type (k, l) and (k', l') respectively. Then we define the tensor product $s \otimes t$ to be the tensor field of type $(k + k', l + l')$ such that

$$\begin{aligned}(s \otimes t)&(\xi_1, \dots, \xi_k, X_1, \dots, X_l, \xi_{k+1}, \dots, \xi_{k+k'}, X_{l+1}, \dots, X_{l+l'}) \\ &= s(\xi_1, \dots, \xi_k, X_1, \dots, X_l) t(\xi_{k+1}, \dots, \xi_{k+k'}, X_{l+1}, \dots, X_{l+l'})\end{aligned}$$

for 1-forms ξ_j and vector-fields X_j.
The tensor product is associative and distributive.
Let us denote by $\mathcal{T}_l^k(M)$ the space of (smooth) tensorfields of type (k, l). Elements of $\mathcal{T}_l^k(M)$ can be viewed as sections of the bundle $T_l^k(M)$ with fiber at x given by $\bigotimes^k T_x M \bigotimes^l T_x^* M$. Thus there is a tensor product operation on bundles, which underlies the above tensor product defined on tensor fields.

3.7.1. *Contraction.* Let t be a tensor field of type (k, l). Then a contraction of t is a tensor field of type $(k - 1, l - 1)$ defined by

$$(\mathcal{C}t)(\xi_1, \dots, \xi_{k-1}, X_1, \dots, X_{k-1}) = \sum_{i=1}^n t(\xi_1, \dots, \xi_r, e^i, \xi_{r+1}, \dots, \xi_{k-1}, X_1, \dots, X_q, e_i, X_{q+1}, \dots, X_{l-1})$$

where the e^i, e_j are elements of a dual basis.
Note that the contraction involves a choice of r and q.
Example: let t be a covariant tensor. Then

$$t(X_1, \dots, X_k) = \mathcal{C} \cdots \mathcal{C} t \otimes X_1 \otimes \cdots \otimes X_k$$

i.e. the tensor t evaluated on the vectorfields $X_1, \dots, X_k$ can be viewed as the result of a sequence of contractions on the tensor product $t \otimes X_1 \otimes \cdots \otimes X_k$.

3.8. **Abstract index notation (Wald, §2.4).** It is often useful to do calculations by expanding tensors in terms of a particular frame, for example $g = g_{\mu\nu} dx^\mu \otimes dx^\nu$. However, the components of g depend on the choice of coordinates x^μ. It is possible to avoid this problem and still keep most of the advandages for explicit calculation by using the *abstract index notation*. This consists of associating upper indices to contravariant and lower indices to covariant components of the tensor: Thus the expression $t^{abc}{}_{de}$ simply corresponds to a tensor of order $(3, 2)$. The tensor product can be written for example $(s \otimes t)^{abc}{}_{def}{}^{gh} = s^{abc}{}_{de} t_f{}^{gh}$.
Wald uses the convention that greek indices correspond to coordinate frames but lower case latin indices are abstract indices.
In terms of the abstract index notation we denote contraction by

$$(\mathcal{C}t)^{a_1 \cdots a_{k-1}}{}_{b_1 \cdots b_{l-1}} = t^{a_1 \cdots e \cdots a_{k-1}}{}_{b_1 \cdots e \cdots b_{l-1}}$$

The inverse metric tensor is denoted by g^{ab} and satisfies the rules $g^{ab} g_{bc} = \delta^a{}_c$ where $\delta^a{}_c$ is the tensor which is given at $x \in M$ by the identity mapping on $T_x M$. g^{ab} and g_{ab} can be used to raise and lower indices: $g^{ab} T_{bc} = T^a{}_c$ and $g_{ab} T^{bc} = T_a{}^c$.

3.8.1. *Symmetrization and antisymmetrization.* Indices enclosed with round brackets denote symmetrization and square brackets denote antisymmetrization. For example

$$T_{(ab)} = \frac{1}{2}(T_{ab} + T_{ba})$$

and

$$T_{[ab]} = \frac{1}{2}(T_{ab} - T_{ba})$$

More generally,

$$t_{(a_1 \cdots a_k)} = \frac{1}{k!} \sum_\pi t_{a_{\pi(1)} \cdots a_{\pi(k)}}$$

and

$$t_{[a_1 \cdots a_k]} = \frac{1}{k!} \sum_\pi \text{sign}(\pi) t_{a_{\pi(1)} \cdots a_{\pi(k)}}$$

with the obvious generalization to mized tensors.
Note: symmetrization and antisymmetrization is only defined w.r.t. indices of the same type.

3.9. **Differential forms, exterior derivative, integration, Wald, appendix B.** An antisymmetric covariant tensor field of order p is called a differential form of order p or p-form.
Let ξ and ω be differential form of order p and p‘ respectively. The antisymmetrized tensor product (wedge product) $\xi \wedge \omega$ is a differential form of order $p + p'$ defined by

$$(\xi \wedge \omega)_{a_1 \cdots a_{p+p'}} = \frac{(p+p')!}{p!p'!} \xi_{[a_1 \cdots a_p} \omega_{a_{p+1} \cdots a_{p+p'}]}$$

The space of p-forms is denoted by $\Lambda^p(M)$ and is the space of sections of a bundle with fiber at x given by the antisymmetrized tensor product of $T_x^* M$ of order p. This space has dimension $\binom{n}{p}$, in particular forms of order $p > n$ always vanish. These facts can be proved from the identities for the totally antisymmetric tensor ε, see Wald, Appendix B.2.
In particular the space of n-forms are sections of a line-bundle, i.e. with fiber $\mathbb{R}$.
A differential form can be expanded in terms of coordinates as $\omega = \omega_{\mu_1 \cdots \mu_p} dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_p}$.
The exterior derivative d is defined by $(d\omega)_{ea_1 \cdots a_p} = \nabla_{[e} \omega_{a_1 \cdots a_p]}$ and is independent of the chioce of the derivative operator ∇. In particular, in terms of coordinates,

$$d\omega = \frac{\partial \omega_{\mu_1 \cdots \mu_p}}{\partial x^\gamma} dx^\gamma \wedge dx^{\mu_1} \wedge \cdots \wedge dx^{\mu_p}.$$

Thus d is well defined without introducing a metric on M.
An important fact is that $d^2 = 0$. This follows from the definition and the fact that the second partial derivative on functions is symmetric.
The Leibniz rule for the exterior derivative takes the form

$$d(\xi \wedge \omega) = d\xi \wedge \omega + (-1)^p \xi \wedge d\omega$$

where p is the order of ξ.
A manifold with an everywhere nonzero section of $\Lambda^n(M)$ is called orientable.
On an orientable manifold, there is a well defined notion of integration of n-forms.

4. Covariant Derivative and Curvature, see Wald §3.1-3.3

Here I discuss these concepts in a slightly different notation than is used in Wald.
Let (M, g) be a manifold with a semi-Riemannian structure. Let $c : [0, 1] \to M$ be a curve. Then if $\dot{c}(t) = dc(t)/dt \in T_{c(t)} M$ satisfies $g(\dot{c}, \dot{c}) > 0$ (in which case c is said to be space like, we define the length of c to be

$$\ell[c] = \int g(\dot{c}(t), \dot{c}(t))^{\frac{1}{2}} dt$$

while if $g(\dot{c}, \dot{c}) < 0$ (the timelike case) we define the *proper time* along c to be

$$\tau[c] = \int (-g(\dot{c}(t), \dot{c}(t)))^{\frac{1}{2}} dt$$

One proves easily that $\ell[c]$ resp. $\tau[c]$ are independent of the parametrization.

4.1. **Geodesics.** A geodesic is defined to be a curve which is stationary w.r.t. ℓ or τ in the sense that

$$\frac{d}{ds}\ell[c_s]\Big|_{s=0}$$

resp.

$$\frac{d}{ds}\tau[c_s]\Big|_{s=0}$$

where c_s is a one-parameter family of curves such that $c_0 = c$.
By a direct calculation shows that stationarity is equivalent to the equation

$$\ddot{c}^\mu(t) + \Gamma^\mu_{\gamma\nu}\dot{c}^\gamma(t)c^\nu(t) = 0$$

where $\Gamma^\mu_{\gamma\nu}$ denote the Christoffel symbols.
This leads one to introduce the notation $\nabla_X Y$, in coordinates

$$\nabla_X Y = X^\mu(\partial_{x^\mu}Y^\nu + \Gamma^\nu_{\gamma\mu}Y^\gamma)\partial_{x^\nu}$$

Parallel translation along c is defined by solutions to the ODE

$$\nabla_{\dot{c}}Y = 0$$

It can be shown that parallel translating from $c(0)$ to $c(t)$ defines an isometry $T_{c(0)}M \to T_{c(t)}M$.
The geodesic differential equation says simply that the velocity field is parallel, i.e. the motion of the point $c(t)$ is not accelerated.
In Riemannian geometry, it makes sense to define

$$d(x, y) = \inf \ell[c]$$

where the inf is taken over all curves connecting x and y. It is easily shown that this defines a metric. The Hopf-Rhinow theorem states that this metric is complete if and only if all geodesics can be continued indefinitely and in that case there is a minimizing geodesic connecting each pair of points in M.

4.2. **Exponential Map.** For $x \in M$, there is defined a map $Exp_x : T_xM \to M$ by putting $Exp_x(v) = c_v(1)$ where c_v is the unique geodesic defined by solving the geodesic differential equation with initial data $c(0) = x, \dot{c}(0) = v$ up to time $t = 1$.
Using the exponental map one introduces geodesic normal coordinates by pulling back to T_xM by Exp_x and then using spherical coordinates on T_xM.

4.3. **The covariant derivative.** ∇ defined as above is a torsion free covariant derivative, i.e. it satisfies the following rules: Let $f \in C^\infty$ and let X, Y be vectorfields on M, then

- $\nabla_X Y - \nabla_Y X = [X, Y]$ (torsion free)
- $\nabla_X(fY) = (df)Y + f\nabla_X Y$ (Leibniz rule)
- $\nabla_{fX}Y = f\nabla_X Y$

From the above it is clear that we can view ∇_X as a map $\nabla_X : \mathcal{T}_0^1(M) \to \mathcal{T}_0^1(M)$, or in fact $\nabla : \mathcal{T}_0^1(M) \to \mathcal{T}_1^1(M)$. It is not hard, using the properties of the tensor product and ∇ to show that there is a generalization of $\nabla_X : \mathcal{T}_l^k(M) \to \mathcal{T}_l^k(M)$ which acts on general tensors fields. This satisfies the Leibniz rule in the form

$$\nabla_X(s \otimes t) = (\nabla_X s) \otimes t + s \otimes (\nabla_X t)$$

and commutes with contractions, i.e. $\nabla_X \mathcal{C} t = \mathcal{C} \nabla_X t$ for any contraction $\mathcal{C}$. From this it follows directly that for a covariant tensor t, we have

$$Xt(Y_1, \dots, Y_k) = (\nabla_X t)(Y_1, \dots, Y_k) + \sum_{j=1}^{k} t(Y_1, \dots, \nabla_X Y_j, \dots, Y_k)$$

The covariant derivative defined as above is metric, i.e.

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

or

$$\nabla_X g = 0$$

for any vectorfield X. In terms of the abstract index notation this becomes $\nabla_a g_{bc} = 0$.
For a given metric, there is a *unique* torsion free metric covariant derivative (Wald, Thm. 3.1.1).
As mentioned in §1.3.2 of the notes the partial derivative satisfies $(D_X D_Y - D_Y D_X - D_{[X,Y]})Z = 0$ for any vectorfields X, Y, Z on $\mathbb{R}^n$. Now for a covariant derivative we define the *Riemann Curvature tensor* by

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

Here the sign is chosen to coincide with Wald. In indices,

$$R_{abc}{}^d Z^c = (\nabla_b \nabla_a - \nabla_a \nabla_b) Z^d$$

or acting on a 1-form,

$$R_{abc}{}^d \xi_d = (\nabla_a \nabla_b - \nabla_b \nabla_a) \xi_c$$

One checks that $R(X, Y)(fZ) = fR(X, Y)Z$ and thus $R(X, Y)Z$ is a *tensor*!
It is sometimes useful to consider R as defining a 2-form with values in $End(TM)$, i.e. (1,1)-tensors, this is defined by $X, Y \to R(X, Y)$, where the map $Z \to R(X, Y)Z$ can be viewed as a (1,1)-tensor.
It is correct to view the expression $(\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]})Z$ as the antisymmetrized second covariant derivative of Z.
The tensor R can be said to measure the change in a vector after parallel translation around an infinitesimal loop.
Above, the $(1, 3)$-form of R is given. The totally covariant form of R is given by $R(X, Y, Z, W) = g(R(X, Y)Z, W)$ or $R_{abcd} = R_{abc}{}^f g_{fd}$.

4.4. **Ricci identities.** Considering the fact that $R(X,Y) = \nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]} \in End(TM)$ we get using the extension of the action of $End(TM)$ to tensor products the identities

$$\begin{aligned}&(\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]})t(\xi_1,\dots,\xi_k,X_1,\dots,X_l)\\ &= -\textstyle\sum_{j=1}^k t(\xi_1,\dots,R(X,Y)^*\xi_j,\dots,\xi_k,X_1,\dots,X_l)\\ &\quad -\textstyle\sum_{j=1}^l t(\xi_1,\dots,\xi_k,X_1,\dots,R(X,Y)X_j,\dots,X_l)\end{aligned}$$

cf. Wald, eq. (3.2.12). Here we use the skew adjointness of $R(X,Y)$, see point 3) below.

4.5. **Symmetries of the Riemann Tensor. Wald §3.2.**

(1) $R(X,Y,Z,W) = -R(Y,X,Z,W) = -R(X,Y,W,Z)$. The first identity follows from the definition. The second by using the Ricci identities and the fact that ∇ is metric,

$$0 = \left((\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]})g\right)(Z,W) = g(R(X,Y)Z,W)+g(Z,R(X,Y)W)$$

(2) ${R_{[abc]}}^d = 0$. This follows using the fact that ${R_{[abc]}}^d\omega_d$ is equal to a multiple of $\nabla_{[a}\nabla_b\omega_{c]}$ which in turn is proportional to $d^2\omega = 0$ since $d^2 = 0$. Here d is the exterior derivative.

Given 1) this is equivalent (by an easy manipulation) to the cyclic identity

$$R_{abcd} + R_{bcad} + R_{cabd} = 0$$

which alternatively can be proved by noting that

$$R(X,Y)Z+R(Y,Z)X+R(Z,X)Y = -([[X,Y],Z]+[[Y,Z],X]+[[Z,X],Y]) = 0$$

due to the Jacobi identity.

(3) $R(X,Y,Z,W) = R(Z,W,X,Y)$ (see Wald, problem 3.3).

(4) $\nabla_{[a}{R_{bc]d}}^e = 0$. (this is the 2:nd Bianchi identity)

From 1) and 3) it follows that we may view the Riemann tensor as defining a symmetric bilinear form on Λ^2TM.

One may think of the covariant derivative ∇ as defined w.r.t. a connection on the principal bundle $O(M)$ of orthonormal frames on M. In terms of ON frames, ∇ can be written

$$\nabla Y = d^\omega Y = dY + \omega\wedge Y$$

where ω can be considered as a $Lie(O(n))$ valued 1-form, a connection.

Then the curvature tensor is $(d^\omega)^2 = d\omega + \frac{1}{2}[\omega,\omega]$. To see this, one calulates

$$\begin{aligned}d^\omega d^\omega Y &= d^\omega(dY + \omega Y)\\ &= d^2Y + d(\omega Y) + \omega\wedge dY + \omega\wedge\omega Y\\ &= d\omega Y - \omega\wedge dY + \omega\wedge dY + \omega\wedge\omega Y\\ &= (d\omega + \frac{1}{2}[\omega,\omega])Y = \Omega^\omega Y\end{aligned}$$

Here Ω^ω is the curvature of the connection ω. The identity $\omega \wedge \omega = \frac{1}{2}[\omega, \omega]$ was used which holds for Lie-algebra valued 1-forms.
From this we see that (*this calculation should be taken with a grain of salt!!*)

$$\begin{aligned} d^\omega \Omega^\omega &= d\Omega^\omega + [\omega, \Omega^\omega] \\ &= d^2\omega + \frac{1}{2} d[\omega, \omega] + [\omega, \Omega^\omega] \\ &= \frac{1}{2}([d\omega, \omega] - [\omega, d\omega]) + [\omega, \Omega^\omega] \\ &= [d\omega, \omega] + [\omega, \Omega^\omega] \\ &= [d\omega, \omega] + [\frac{1}{2}[\omega, \omega], \omega] + [\omega, \Omega^\omega] \\ &= [\Omega^\omega, \omega] + [\omega, \Omega^\omega] = 0 \end{aligned}$$

which is the same as 4).
Here we used among other identities which hold for Lie-algebra valued forms, the Jacobi identity which says that $[[\omega, \omega], \omega] = 0$, see Bleecker, Gauge Theories and Variational Principles. This is the formalism which underlies the (classical) Yang-Mills theories.
The above point of view is closely related to what one is lead to by doing computations in terms of ON frames, see Wald, §3.4b.

4.6. **Sectional Curvature.** The sectional curvature w.r.t. the 2-plane spanned by X, Y is defined to be

$$K(X, Y) = R(X, Y, X, Y)/(g(X, X)g(Y, Y) - g(X, Y)^2)$$

A manifold which has constant sectional curvature, i.e. K(X,Y) is independent of the choice of (linearly independent) X, Y is called a “space form”. These can be classified, cf. the book by Wolf.
Clearly, a manifold with constant sectional curvature has a Riemann tensor of the form

$$R(X, Y, Z, W) = g(X, Z)g(Y, W) - g(X, Y)g(Z, W)$$

(use polarization).
An important fact is that in the indefinite metric case, a metric which has sectional curvature bounded from above and below actually has constant sectional curvature.
It follows from the Bianchi identities, that if $R(X, Y, Z, W) = fg(X, Z)g(Y, W) - g(X, Y)g(Z, W)$ for some function f then f is a constant function.

4.7. **The Ricci tensor and the Scalar curvature.** Contracting the Riemann tensor gives new tensors:

- $R_{ab} = R_{acb}{}^c$ (the Ricci tensor). This satisfies $R_{ab} = R_{ba}$, i.e. it is a symmetric covariant tensor of order 2.

- Further contraction gives a function: $R = g^{ab}R_{ab}$. This is called the scalar curvature (or the Ricci scalar).

By contracting the Bianchi identities one proves the following formula:

$$\nabla^a R_{ab} = \frac{1}{2}\nabla_b R$$

This shows that the tensor $R_{ab} - \frac{1}{2}Rg_{ab}$ satisfies $\nabla^a(R_{ab} - \frac{1}{2}Rg_{ab}) = 0$.
The operation $\nabla^a T_{ab}$ on a symmetric tensor is called divergence.
The tensor $E_{ab} = R_{ab} - \frac{1}{2}Rg_{ab}$ is called the Einsten tensor.
We are now ready for the following definition:

Definition 4.1. *A 4-dimensional Lorenz manifold satisfying $E_{ab} = 0$ is a vacuum solution to Einstein's equations, or a vacuum spacetime.*

We shall see later the significance of the divergence free property of E_{ab} for the general covariance (coordinate independence) of the theory. It is a necessary consequence of the existence of a generally covariant Lagrangian formulation of the Einstein equations.