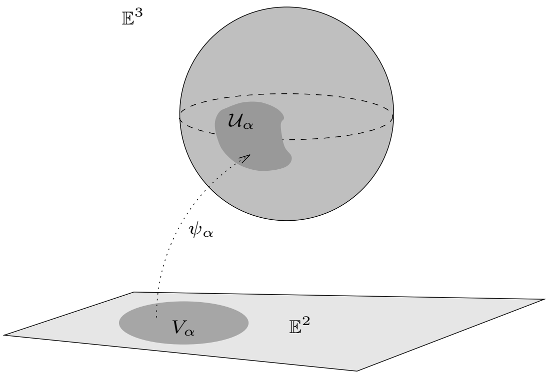
# Differential Geometry Part III Notes

## 1 Introduction

### 1.1 From smooth surfaces to smooth manifolds

The simplest way that the objects of the form we call smooth surfaces $\mathcal{S} \subset \mathbb{E}^3$ arise are as level sets of a smooth function, say $f(x, y, z) = c$, at a non-critical[1] value $c$. Example: $\mathbb{S}^2$ as $x^2 + y^2 + z^2 = 1$. It is the implicit function theorem[2] that says that these objects are, in some sense, two dimensional, i.e. that $\mathcal{S}$ can be expressed as the union of the images of a collection of maps $\psi_\alpha : V_\alpha \to \mathbb{E}^3$, $V_\alpha \subset \mathbb{E}^2$, such that $\psi_\alpha$ is smooth, $D\psi_\alpha$ is one-to-one, and denoting $\psi_\alpha(V_\alpha)$ as $\mathcal{U}_\alpha$, $\psi_\alpha$ is a homeomorphism[3] $\psi_\alpha : V_\alpha \to \mathcal{U}_\alpha$.




Let us denote the inverse of the $\psi_\alpha$'s by $\phi_\alpha : \mathcal{U}_\alpha \to V_\alpha$. The collection $\{(\mathcal{U}_\alpha, \phi_\alpha)\}$ is known as an *atlas* of $\mathcal{S}$. Each $\mathcal{U}_\alpha, \phi_\alpha$ is called a *chart*, or alternatively, a *system of local coordinates*[4].

The word "differential" in the title of this course indicates that we should be able to do calculus. The point about local coordinates is that it allows us to do calculus on the surface.

The first issue: When can we even say that a function on the surface is differentiable?

Answer:

**Definition 1.1.** *We say that $f : \mathcal{S} \to \mathbb{R}$ is $C^\infty$ at a point $p$ if $f \circ \phi_\alpha^{-1} : V_\alpha \to \mathbb{R}$ is $C^\infty$ for some $\alpha$.*

[1] i.e. a value $c \in \mathbb{R}$ such that $df(p)$ is surjective for all $p \in f^{-1}(c)$

[2] The reader is assumed familiar with standard results in multivariable analysis.

[3] Here we are taking $\mathcal{S}$ to have the induced topology from $\mathbb{E}^3$. We assume that the reader is familiar with basic notions of point set topology.

[4] Actually, more correctly, one says that the system of local coordinates are the projections $x^i \circ \phi_\alpha$ to the standard coordinates on $\mathbb{R}^2$.

For this to be a good definition, it should not depend on the chart. Let $\phi_\alpha$, $\phi_\beta$ be different charts containing $p$. We have

$$f \circ \phi_\alpha^{-1} = f \circ \phi_\beta^{-1} \circ (\phi_\beta \circ \phi_\alpha^{-1})$$

where this is defined.

**Proposition 1.1.** *$\phi_\beta \circ \phi_\alpha^{-1}$ is $C^\infty$ on the set where it is defined.*

*Proof.* Exercise □

Thus, the definition holds for any compatible chart.

We can think of our surface as the topological space $\mathcal{S}$, together with the homeomorphisms $\phi_\alpha : \mathcal{U}_\alpha \to V_\alpha$. Question: What did we need to define consistently the notion that $f$ is smooth? Answer: Proposition 1.1. The notion of an *abstract smooth surface* distills this property from that of a surface.

**Definition 1.2.** *An* abstract smooth surface *is a topological space $\mathcal{S}$ together with an open cover $\mathcal{U}_\alpha$ and homeomorphisms $\phi_\alpha : \mathcal{U}_\alpha \to V_\alpha$, with $V_\alpha$ open subsets of $\mathbb{R}^2$, such that $\phi_\beta \circ \phi_\alpha^{-1}$, where defined, are $C^\infty$.*

The notion of a smooth $n$-dimensional manifold $\mathcal{M}$ is defined now precisely as above, where $\mathbb{R}^2$ is replaced by $\mathbb{R}^n$.

**Definition 1.3.** *A map $f : \mathcal{M} \to \tilde{\mathcal{M}}$ is smooth if $\tilde{\phi}_\beta \circ f \circ \phi_\alpha^{-1}$ is smooth.*

Check consistency of the definition.

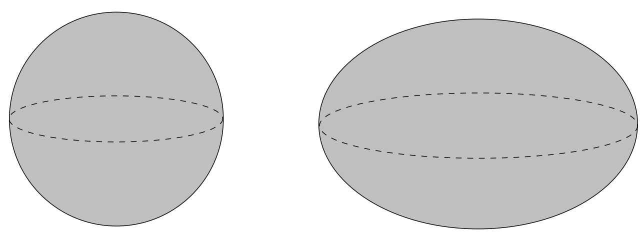
**Definition 1.4.** *$\mathcal{M}$ and $\tilde{\mathcal{M}}$ are said to be* diffeomorphic *if there exists an $f : \mathcal{M} \to \tilde{\mathcal{M}}$ such that $f$ and $f^{-1}$ are both smooth.*

Exercise: The dimension $n$ of a manifold is uniquely defined and a diffeomorphism invariant.

Examples: $\mathbb{E}^n$, $\mathbb{S}^n$, products, quotients, twisted products (fiber bundles, etc.), connected sums. The point is that manifolds are a very flexible category and there is the usual economy provided by a good definition. We will discuss all this soon enough in the course.

## 1.2 What defines geometry?

The study of smooth manifolds is what is known as *differential topology.* From the point of view of the smooth structure, the sphere $\mathbb{S}^n$ and the set $\frac{x_1^2}{a_1^2} + \cdots \frac{x_{n+1}^2}{a_{n+1}^2} = 1$ are diffeomorphic as manifolds.



To speak about geometry, we must define additional structure. To speak about "differential" geometry, this structure should be defined via the calculus. Without a doubt, the most important such structure is that of a *Riemannian* (or more generally semi-Riemannian) metric.

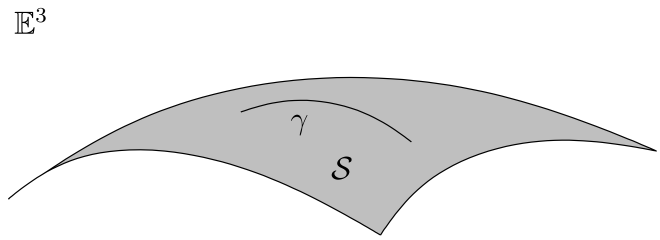
This concept again arises from distilling from the theory of surfaces in $\mathbb{E}^3$ a piece of structure: A surface $\mathcal{S} \subset \mathbb{E}^3$ comes with a notion of how to measure the lengths of curves. This notion can be characterized at the differential level. Formally, we may write

$$dx^2 + dy^2 + dz^2 = E(u,v)du^2 + 2F(u,v)dudv + G(u,v)dv^2, \tag{1}$$

where

$$\begin{aligned} E &= \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2 \\ F &= \frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v} + \frac{\partial z}{\partial u}\frac{\partial z}{\partial v} \\ G &= \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2. \end{aligned}$$

This is motivated by the chain-rule à la Leibniz. The expression on the right hand side of (1) is called the *first fundamental form*. What does this actually mean? Say that a smooth curve $\gamma : I \to \mathcal{S}$ is given by $(x(t), y(t), z(t)) = (u(t), v(t))$.




Then we can compute its length $L$ in the standard way:

$$L = \int \sqrt{x'^2 + y'^2 + z'^2}dt,$$

and, by the chain rule, we obtain

$$L = \int \sqrt{Eu'^2 + 2Fu'v' + Gv'^2}dt \tag{2}$$

in our local coordinates on $\mathcal{S}$. It turns out that if $(\tilde{u}, \tilde{v})$ is another coordinate system, then writing $dx^2 + dy^2 + dz^2 = \tilde{E}d\tilde{u}^2 + 2\tilde{F}d\tilde{u}d\tilde{v} + \tilde{G}d\tilde{v}^2$, we can compute the relation between $E$ and $\tilde{E}$:

$$\tilde{E} = E\frac{\partial u}{\partial \tilde{u}}\frac{\partial u}{\partial \tilde{u}} + 2F\frac{\partial u}{\partial \tilde{u}}\frac{\partial v}{\partial \tilde{u}} + G\frac{\partial v}{\partial \tilde{u}}\frac{\partial v}{\partial \tilde{u}}. \tag{3}$$

Now we ask, let us again forget about $\mathbb{E}^3$. Question: What was it about $\mathcal{S}$ that allowed us to unambiguously define lengths of curves? Answer: A set of functions $E, F, G$ defined for each chart, transforming via (3). We distill from the above the following:

**Definition 1.5.** *A* Riemannian metric *on an abstract* 2*-surface is a collection of smooth functions* $\{E_i\}$*,* $\{F_i\}$*,* $\{G_i\}$ *on an atlas* $\{\mathcal{U}_i\}$*, transforming like in* (3)*, satisfying in addition*

$$E_i G_i - F_i^2 > 0. \tag{4}$$

In particular, the formula (2) now allows one to define consistently the notion of the length of a smooth curve $\phi : I \to \mathcal{S}$.[5] The condition (4) ensures that our notion of length is positive.[6]

The expression on the right hand side can be generalized to $n$ dimensions, and this defines the notion of a Riemannian metric on a smooth manifold.[7] We leave this until we have more sophisticated language to describe these issues.

The reader familiar with the geometry of surfaces has no doubt encountered the so-called *Theorema Egregium* of Gauss. This says that the curvature, originally, defined using the so-called second fundamental form[8], in fact can be expressed as a complicated expression in local coordinates involving up to second derivatives of the compontents $E$, $F$, $G$ of the first fundamental form. That is to say, it could have been defined in the first place as said expression. In particular, the notion of curvature can thus be defined for abstract surfaces. One main difficulty in Riemannian geometry in higher dimensions is the algebraic complexity of the analogue of this curvature curvature, which is no-longer a scalar, but a so-called "tensor".

## 1.3 Geometry, curvature, topology

The following remarks are meant to give a taste of the kinds of results one wants to prove in geometry. Some familiarity with curvature of surfaces will be useful for getting a sense of what these statements mean.

The common thread in these examples is that they relate completeness, curvature and global behaviour (e.g. topology):

**Theorem 1.1.** *Let* $(\mathcal{M}, g)$ *be a simply-connected*[9] $n$*-dimensional complete Riemannian manifold with nonpositive sectional curvature. Then* $(\mathcal{M}, g)$ *is diffeomorphic to* $\mathbb{R}^n$*.*

**Theorem 1.2.** *Let* $(\mathcal{M}, g)$ *be complete, orientable, even-dimensional and of positive sectional curvature. Then* $(\mathcal{M}, g)$ *is simply connected.*

[5] How to define a smooth curve?

[6] The semi-Riemannian case replaces (4) with the assumption that this determinant is non-zero. See Section 1.4.

[7] As we have tentatively defined them, not all manifolds admit Riemannian metrics. But Hausdorff paracompact ones do...

[8] For those who know about the geometry of curves and surfaces...

[9] We will often make reference to basic notions of algebraic topology.

### 1.3.1 Aside: Hyperbolic space and non-euclidean geometry

The set $\mathbb{H}^2$ can be covered by one chart $\{(u, v) : v > 0\}$, and the Riemannian metric is given by

$$\frac{1}{v^2}(du^2 + dv^2). \tag{5}$$

Later on we will recognize $\mathbb{H}^2$ as a complete space form with the topology of $\mathbb{R}^2$ and with constant curvature $-1$. This defines a so-called non-Euclidean geometry, a geometry satisfying all the axioms of Euclid with the exception of the so-called fifth postulate. In particular, the existence of the Riemannian geometry (5) shows the necessity of the Euclidean fifth postulate to determine Euclidean geometry.

The enigma of why it took so long for this to be understood is in part explained by the following global theorem:

**Theorem 1.3.** *Let $(\mathcal{S}, g)$ be an abstract surface with Riemannian metric. If $\mathcal{S}$ is complete with constant negative curvature, then $\mathcal{S}$ cannot arise as a subset of $\mathbb{E}^3$ so that $g$ is induced as in* (1) *(in fact, not even as an immersed surface.)*

Compare this with the case of the sphere.

## 1.4 General relativity

A subject with great formal similarity, but a somewhat diverging epistomological basis, with Riemannian geometry is "general relativity". The basis for this theory is a four dimensional manifold: $\mathcal{M}$, called *spacetime*, together with a so-called *Lorentzian* metric, i.e. a smooth quadratic form $\sum g_{ij}dx^i dx^j$ such that the signature of $g$ is $(-, +, +, +)$. (In two dimensions, Lorentzian vs. Riemannian would just mean that the sign of (4) is flipped.) Pure Lorentzian geometry in full generality is more complicated and less studied than pure Riemannian geometry. What sets general relativity apart from pure geometry, is that in this theory, the Lorentzian metric must satisfy a set of partial differential equations, the so-called Einstein equations. These equations constitute a relation between a geometric quantity, the Einstein tensor[10], and the energy-momentum content of matter. The central questions in general relativity are questions of the dynamics of this system. It is thus a much more rigid subject.

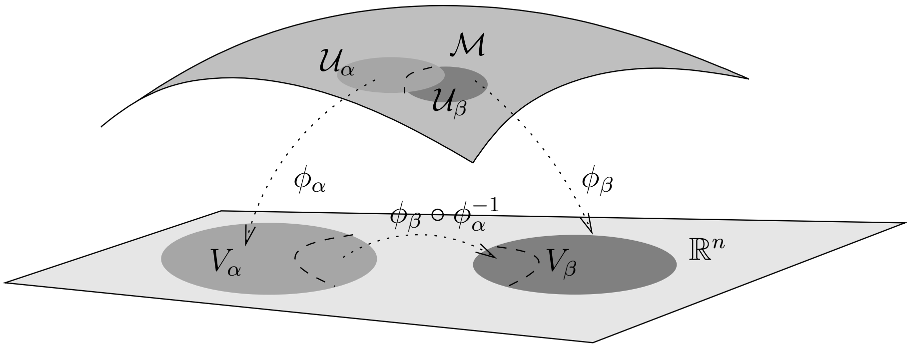
The above comments notwithstanding, there are (quite surprisingly!) some spectacular theorems in general relativity which can be proven via pure geometry. The reason: When the dynamics of matter is not specified, the Einstein equations still yield *inequalities* for this curvature tensor, analogous in many ways to the inequalities in the statement of the previous theorems. The statements of the theorems obtained–unfortunately often referred to as "singularity theorems"–are not easy to state without some terminology. Time permitting, we will discuss these later...

[10] This is an expression derived from the Ricci and scalar curvatures.

# 2 Manifolds

## 2.1 Basic definitions

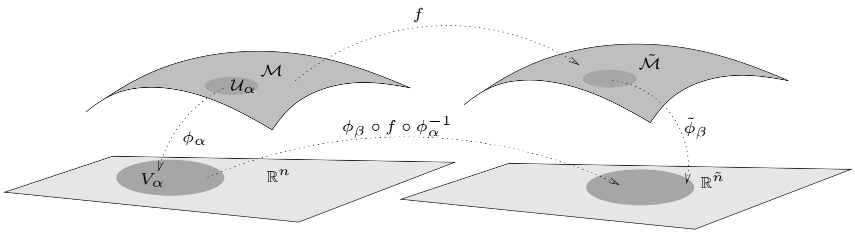
**Definition 2.1.** *A $C^\infty$ manifold of dimension $n$ is a topological space $\mathcal{M}$, together with an open cover $\mathcal{U}_\alpha$, and a collection of homeomorphisms $\phi_\alpha : \mathcal{U}_\alpha \to V_\alpha$, where $V_\alpha \subset \mathbb{R}^n$, such that $\phi_\alpha \circ \phi_\beta^{-1}$ is $C^\infty$ where defined:*




*We call the collection $\{(\mathcal{U}_\alpha, \phi_\alpha)\}$ an* atlas, *and each $(\mathcal{U}_\alpha, \phi_\alpha)$, a* chart. *We call $\pi^i \circ \phi_\alpha$ a system of* local coordinates, *where $\pi^i$ denote the projections to standard coordinates on $\mathbb{R}^n$.*

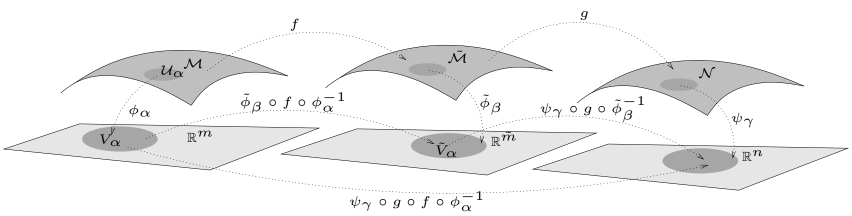
Notation: Very often in notation we completely suppress $\phi_\alpha$ and talk about local coordinates $(x^1, \dots, x^n)$. It is understood that $x^1 = \pi^1 \circ \phi_\alpha$ for a $\phi_\alpha$.

**Definition 2.2.** *A map $f : \mathcal{M} \to \tilde{\mathcal{M}}$ is $C^\infty$ if $\tilde{\phi}_\beta \circ f \circ \phi_\alpha^{-1}$ is $C^\infty$.*




**Proposition 2.1.** *If $f : \mathcal{M} \to \tilde{\mathcal{M}}$ is smooth, and $g : \tilde{\mathcal{M}} \to \mathcal{N}$ is smooth, then $g \circ f$ is smooth.*

*Proof.*




$\square$

**Definition 2.3.** *We say that two atlases $\{(\mathcal{U}_\alpha, \phi_\alpha)\}$ and $\{(\tilde{\mathcal{U}}_\beta, \tilde{\phi}_\beta)\}$ are compatible, if their union is also an atlas.*

**Remark 2.1.** *Let $\mathcal{M}$ be a manifold with respect to an atlas $\{\mathcal{U}_\alpha, \phi_\alpha\}$, and also, with respect to $\{\tilde{\mathcal{U}}_\alpha, \tilde{\phi}_\beta\}$. Then the notion of smooth maps is the same.*

For this reason, it is often the case that one adds the condition of completeness for the atlas in the definition of manifold. Any atlas as defined previously is contained in a complete atlas. To check properties, however, one always picks a sub-atlas compatible with this maximal atlas.

**Definition 2.4.** *$\mathcal{M}$ and $\tilde{\mathcal{M}}$ are said to be* diffeomorphic *if there exists an $f : \mathcal{M} \to \tilde{\mathcal{M}}$ such that $f$ and $f^{-1}$ are both smooth.*

Exercise: The above defines an equivalence relation.

**Example 2.1.** *The set $\mathbb{R}^n$ is an $n$-dimensional manifold defined by an atlas consisting of a single chart, the identity map.*

**Example 2.2.** *$\mathbb{S}^n$, with topology given as the subset*

$$(x^1)^2 + \cdots (x^{n+1})^2 = 1$$

*of $\mathbb{R}^{n+1}$, can be given the structure of an $n$-dimensional smooth manifold with coordinate charts the projections to the coordinate hyperplanes.*

To see this, note the transition functions are of the form:

$$(x^1, \ldots, x^{k-1}, x^{k+1}, \ldots x^{n+1}) \mapsto (x^1, \ldots, x^{k-1}, \sqrt{1 - \sum_{i \neq k} (x^i)^2}, x^{k+1}, \ldots, x^{h-1}, x^{h+1}, \ldots, x^{n+1})$$

Note. In various dimensions, for instance, 4, 7, there are differentiable structures inequivalent to the above[11] which live on the same topology. These are called *exotic spheres.*

**Example 2.3.** *Denote by $\mathbb{R}P^n$ the set of all lines through the origin in $n+1$-dimensional space. This space can be endowed with the structure of an $n$-dimensional manifold, and is then called* real projective space. *With this structure the map $\pi : \mathbb{S}^n \to \mathbb{R}P^n$ is smooth.*

This is an example of the quotient by a discrete group action. For an extension of this kind of construction, see the first example sheet.

**Example 2.4.** *Let $\mathcal{M}$ and $\mathcal{N}$ be manifolds. Then one can define a natural manifold structure on $\mathcal{M} \times \mathcal{N}$.*

Take $\{(\mathcal{U}_\alpha \times \tilde{\mathcal{U}}_\beta, \phi_\alpha \times \tilde{\phi}_\beta)\}$. Complete the details...

[11] i.e. such that the resulting manifold is not diffeomeorphic to the above

## 2.2 Tangent vectors

Let $\mathcal{M}$ be a smooth manifold, let $p \in \mathcal{M}$. Let $X$ denote the algebra of locally $C^\infty$ functions at $p$.[12] Note that if $f \in X$ and $g \in X$ then $fg \in X$, where $fg$ is a locally defined function.

**Definition 2.5.** *A* derivation *$D$ at $p$ is a mapping $D : X \to \mathbb{R}$ satisfying $D(\lambda f + \mu g) = \lambda Df + \mu Dg$, for $\lambda$, $\mu$ scalars, and, in addition, $D(fg) = (Df)g + f(Dg)$.*

**Proposition 2.2.** *The set of derivations at $p$ define a vector space of dimension $n$, denoted $T_p\mathcal{M}$.*

*Proof.* The fact that $T_p\mathcal{M}$ is a vector space is clear. Let $x^i$ be a system of local coordinates centred at $p$. Define a map $\frac{\partial}{\partial x^i}$ by $\frac{\partial}{\partial x^i} f = \partial_{x^i} f$. (Note $\frac{\partial}{\partial x^i} x^j = \delta^i_j$. This in particular implies that the $\frac{\partial}{\partial x^i}$ are linearly independent.) Clearly $\frac{\partial}{\partial x^i}$ is a derivation, by the well-known properties of derivatives. We want to show that the $\left\{\frac{\partial}{\partial x^i}\right\}$ span $T_p\mathcal{M}$. It suffices to show that if $Dx^i = 0$ for all $x^i$, then $D = 0$. So let $D$ be such a $D$, and let $f$ be arbitrary. Locally, $f = \alpha_i x^i + g_i x^i$ where $\alpha_i \in \mathbb{R}$ and where $g_i$ are $C^\infty$. Thus, $Df(p) = \alpha_i Dx^i + x_i Dg^i = 0$, since $p$ is the origin of coordinates. □

Note. We have used above the Einstein summation convention, i.e. the convention that whenever we the same index "up" and "down", as in the expression $\alpha_i x^i$, we are to understand $\sum_{i=1}^n \alpha_i x^i$. Here $n = dim\mathcal{M}$.

**Definition 2.6.** *We will call $T_p\mathcal{M}$ the* tangent space *of $\mathcal{M}$ at $p$, and we will call its elements* tangent vectors.

**Proposition 2.3.** *Let $x^i$ and $\tilde{x}^i$ denote two coordinate systems. Then $\frac{\partial}{\partial \tilde{x}^i} = \frac{\partial x^j}{\partial \tilde{x}^i} \frac{\partial}{\partial x^j}$, in the region where both these are defined.*

*Proof.* If we apply an arbitrary $f$ to both sides, then by the chain rule, the left and right hand side coincide. Thus, the two expressions correspond to one and the same derivation. □

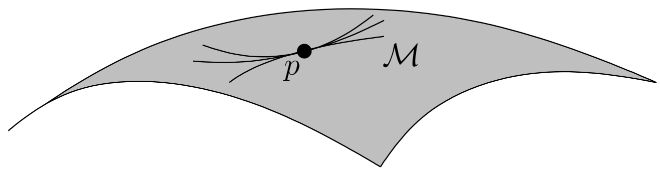
Notation. In writing $\frac{\partial x^j}{\partial \tilde{x}^i}$ one is to understand $\partial_i(\pi^j \circ \phi \circ \tilde{\phi}^{-1})$, where $\tilde{\phi}$ and $\phi$ are the two charts corresponding to the local coordinates.[13]

The geometric interpretation of derivations: Let $\gamma$ be a smooth curve through $p$, i.e. a smooth map $\gamma : (-\epsilon, \epsilon) \to \mathcal{M}$ such that $\gamma(0) = p$. Given $f$, define a derivation $D_\gamma$ by $D_\gamma f = (f \circ \gamma)'(0)$. All derivations in fact arise in this way. For given $\alpha^i \frac{\partial}{\partial x^i}$, then one can consider the curve $t \mapsto (\alpha^1 t, \ldots, \alpha^n t)$, and it is clear from the definition of partial differentiation in local coordinates that the action of $D_\gamma$ coincides with that of $\alpha^i \frac{\partial}{\partial x^i}$. We will often denote this tangent vector as $\gamma'$ or $\dot{\gamma}$.

[12] Exercise: define these formally, as equivalence classes of functions coinciding on a neighborhood.

[13] It is assumed that you know what this means because $\pi^j \circ \phi \circ \tilde{\phi}^{-1} : \mathbb{R}^n \to \mathbb{R}$, so this is partial differentiation from calculus of many variables.

The curves depicted below, suitably parametrized, all correspond to the same derivation at $p$.




We thus often visualize tangent vectors as arrows of a given length (related to the above mentioned parametrization) through $p$ in the direction distuingished by these curves. Exercise: Draw on top of this picture such a vector!

## 2.3 The tangent bundle

From multidimensional calculus, one knows the importance of considering smooth vector fields. We would like a geometric way of describing these in the case of manifolds. It turns out that there is sufficient "economy" in the definition of manifold so as to apply it also to the natural space where these tangent vectors "live". This will allow us to discuss smoothness.

Let $\mathcal{M}$ be an $n$-dimensional smooth manifold. Define $T\mathcal{M}$ to be the set of tangent vectors in $\mathcal{M}$, i.e.

$$T\mathcal{M} = \bigcup_{p\in\mathcal{M}} T_p\mathcal{M}.$$

Note the natural map

$$\pi : T\mathcal{M} \to \mathcal{M},$$

taking a vector in $T_p\mathcal{M}$ to $p$. Define an atlas $\{\tilde{\mathcal{U}}_\alpha, \tilde{\phi}_\alpha\}$ as follows:

$$\tilde{\mathcal{U}}_\alpha = \pi^{-1}(\mathcal{U}_\alpha)$$

$$\tilde{\phi}_\alpha : \left\{ \alpha^i \frac{\partial}{\partial x^i}\bigg|_p \right\} \mapsto \phi_\alpha(p) \times (\alpha^1, \ldots, \alpha^n).$$

**Proposition 2.4.** *The above choice of atlas makes $T\mathcal{M}$ into a smooth manifold such that $\pi$ is smooth.*

Note that, for fixed $x \in \mathcal{M}$, $\tilde{\phi}$ restricted to $T_p\mathcal{M}$ is a linear map.

**Definition 2.7.** *A* vector field *is just a smooth map $V : \mathcal{M} \to T\mathcal{M}$ such that $\pi \circ V = id$, where $id$ denotes the identity map.*

# 3 Vector bundles

The tangent bundle is a special case of the following:

**Definition 3.1.** *A* smooth vector bundle *is a map of manifolds $\pi : \mathcal{E} \to \mathcal{M}$ such that there exists an atlas $\mathcal{U}_\alpha$ for $\mathcal{M}$ such that*

$$\tilde{\phi}_\alpha : \pi^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times \mathbb{R}^m$$

*yields an atlas for $\mathcal{E}$, and where the restrictions of the transition functions to the fibres*

$$\tilde{\phi}_\beta \circ \tilde{\phi}_\alpha^{-1}|_{\tilde{\phi}_\alpha(\pi^{-1}(p))} : \{\phi_\alpha(p)\} \times \mathbb{R}^m \to \{\phi_\beta(p)\} \times \mathbb{R}^m \tag{6}$$

*are linear maps.*

In the above $m$ is a positive integer, not necessarily the dimension $n$ of the manifold! By general nonsense, it follows from (6) that the fibres $\mathcal{E}_p = \pi^{-1}(p)$ acquire the structure of a vector space. For defining

$$\lambda v_p + \mu w_p = \tilde{\phi}_\alpha^{-1}(\lambda \tilde{\phi}_\alpha v_p + \mu \tilde{\phi}_\alpha w_p)$$

we have by (6) that

$$\tilde{\phi}_\alpha^{-1}(\lambda \tilde{\phi}_\alpha v_p + \mu \tilde{\phi}_\alpha w_p) = \tilde{\phi}_\beta(\lambda \tilde{\phi}_\beta v_p + \mu \tilde{\phi}_\beta w_p).$$

**Definition 3.2.** *A* smooth section *of a vector bundle $\mathcal{E}$ is a map $\sigma : \mathcal{M} \to \mathcal{E}$ such that $\pi \circ \sigma = id$.*

# 4 Dual bundles and the cotangent bundle

First a little linear algebra. Given a finite dimensional vector space $V$ (over $\mathbb{R}$), we can associate the dual space $V^*$ consisting of all linear functionals $f : V \to \mathbb{R}$. This is a vector space of the same dimension as $V$.

Given now a map $\phi : V \to W$, then there is a natural map $\phi^* : W^* \to V^*$ defined by $\phi^*(g)(v) = g(\phi(v))$. Thus, given an *isomorphism* $\phi : V \to W$, there exists a map $\psi : V^* \to W^*$ defined by $\psi = (\phi^*)^{-1}$.

Now, given a vector bundle $\pi : \mathcal{E} \to \mathcal{M}$, we can define a vector bundle $\mathcal{E}^*$ called the dual bundle, where

$$\mathcal{E}^* = \bigcup_{p \in \mathcal{M}} (\mathcal{E}_p)^*$$

and where the charts $\chi_\alpha : \pi^{-1}(\mathcal{U}_\alpha) \to V_\alpha \times \mathbb{R}^m$ of $\mathcal{E}^*$, when restricted to the fibres,

$$\chi_\alpha|_{\mathcal{E}^*} = \theta \circ \tilde{\psi}_\alpha|_{\mathcal{E}^*}$$

where $\tilde{\psi}_\alpha|_{\pi^{-1}(p)}$ denotes the map from $\mathcal{E}_p^* \to \mathbb{R}^{m*}$ induced from $\tilde{\phi}_\alpha|_{\mathcal{E}}$, where $\tilde{\phi}$ denote the coordinate charts of $\mathcal{E}$, and $\theta$ denotes some fixed[14] linear isomorphism $\theta : \mathbb{R}^{m^*} \to \mathbb{R}^m$.

[14] i.e. not depending on $\alpha$

**Definition 4.1.** *The dual bundle of the tangent bundle is denoted $T^*\mathcal{M}$ and is called the* cotangent bundle. *Elements of $T^*\mathcal{M}$ are called* covectors, *and sections of $T^*\mathcal{M}$ are called* 1*-forms.*

Let $dx^i$ denote the dual basis[15] to $\frac{\partial}{\partial x^i}$.

**Proposition 4.1.** *Change of basis:* $d\tilde{x}^j = \frac{\partial \tilde{x}^j}{\partial x^i} dx^i$.

Note if $f \in C^\infty(\mathcal{M}, \mathbb{R})$, then there exists a one one form, which we will denote $df$, defined by

$$df(X) = X(f).$$

This is called the differential of $f$. Clearly, in local coordinates,

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

We can think of $d$ as a linear operator

$$d : C^\infty(\mathcal{M}, \mathbb{R}) \to \Gamma(T^*M).$$

Much more about this point of view later.

# 5 The pull-back and the push forward

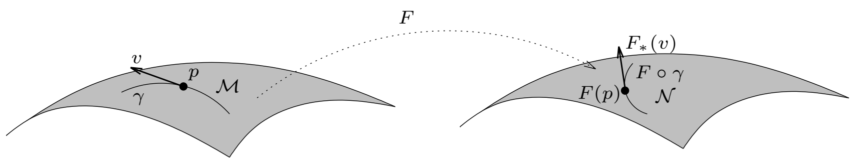
Let $F : \mathcal{M} \to \mathcal{N}$ be a smooth map.

**Definition 5.1.** *For each $p$, the differential of $F$ is a map $(F_*)_p : T_p\mathcal{M} \to T_{F(p)}\mathcal{N}$ which takes $D$ to $\tilde{D}$ with $\tilde{D}g = D(g \circ F)$.*

We can also describe the map $F_*$ in terms of the equivalent characterization of tangent vectors as explained at the end of Section 2.2. Let $v$ be a tangent vector and let $\gamma$ be a curve such that $v = \gamma'$. Then

$$F_*(v) = (F \circ \gamma)'.$$

See below:




We have now

**Definition 5.2.** *We can define a map $F^* : \Gamma(\mathcal{T}^*\mathcal{N}) \to \Gamma(\mathcal{T}^*\mathcal{M})$ by $F^*(\omega)(X) = \omega(F_*(X))$.*

[15] Recall this notion from linear algebra!

**Definition 5.3.** *Let $F : \mathcal{M} \to \mathcal{N}$ be smooth. We say that $F$ is an* immersion *if $(F_*)_p : T_p\mathcal{M} \to T_{F(p)}\mathcal{N}$ is injective for all $p$. We say that $f$ is an embedding if it is an* immersion *and $F$ itself is* 1-1*. In the latter case, if $\mathcal{M} \subset \mathcal{N}$ and $F$ is the identity, we call $F$ a* submanifold.[16]

**Example 5.1.** *Let $\mathcal{M}$ be a manifold and $\mathcal{U} \subset \mathcal{M}$ an open set. Then $\mathcal{U}$ is a submanifold with the induced maps as charts.*

More interesting:

**Proposition 5.1.** *Let $\mathcal{M}$ be a smooth manifold, and let $f_1, \ldots f_d$ be smooth functions. Let $\mathcal{N}$ denote the common zero set of $f_i$ and assume $df_1, \ldots df_m$ span a subset of dimension $d'$ in $T_p^*\mathcal{M}$, for all $p$, where $d'$ is constant. Then $\mathcal{N}$ can be endowed with the structure of a closed submanifold of $\mathcal{M}$.*

See the example sheet!

# 6 Multilinear algebra

The tangent and cotangent bundles are the simplest examples of *tensor bundles*. These are where the objects of interest to us in geometry "live". To understand them, we will need a short diversion into multilinear algebra.

Let $U$, $V$ be vector spaces. We can define a vector space $U \otimes V$ as the free vector space generated by the symbols $u \otimes v$ as $u \in U$, $v \in V$, modulo the subspace generated by $u \otimes (\alpha v + \beta\tilde{v}) - \alpha u \otimes v + \beta u \otimes \tilde{v}$ and $(\alpha u + \beta\tilde{u}) \otimes v - \alpha u \otimes v + \beta\tilde{u} \otimes v$. This space is indeed a vector space. In fact, if $U$ has dimension $n$, with basis $e_1, \ldots, e_n$, and $V$ has dimension $m$, with dimension $f_1, \ldots, f_n$, then $U \otimes V$ has dimension $nm$, with basis $\{e_i \otimes f_j\}$.

**Proposition 6.1.** *We collect some facts about $U \otimes V$.*

1. *$U \otimes V$ has the following universal mapping property. If $B : U \times V \to W$ is bilinear then it factors uniquely as $\tilde{B} \circ h$ where $h : U \times V \to U \otimes V$ is defined by $h : (u, v) \mapsto u \otimes v$, and where $\tilde{B}$ is linear.*

2. *$(U \otimes V) \otimes W = U \otimes (V \otimes W)$. So we can write without fear $U \otimes V \otimes W$.*

3. *$U \otimes V \cong V \otimes U$,*

4. *$Hom(U, V) \cong U^* \otimes V$.*

5. *$(U \otimes V)^* \cong U^* \times V^*$*

The proof of this proposition is left to the reader. Let us here give only the definition of isomorphism number 4 above, more precisely the map $\leftarrow$ as follows:

[16]Note other conventions where $F$ an embedding is required to be a homeomorphism onto its image.

If $\sum c_{ij} u_i^* \otimes v_j$ is an element of $U^* \otimes V$ we send it to the element of $Hom(U, V)$ defined by

$$u \mapsto \sum_{ij} c_{ij} u_i^*(u) v_j.$$

(You also have to check that this is well defined. . . )

**Definition 6.1.** *Let $f : U \to \tilde{U}$, $g : V \to \tilde{V}$ be linear. Then we can define a map $f \otimes g : U \otimes V \to \tilde{U} \otimes \tilde{V}$ taking $\sum u_\alpha \otimes v_\alpha \to \sum \tilde{u}_\alpha \otimes \tilde{v}_\beta$, where $\tilde{u}_\alpha = f(u_\alpha)$, $\tilde{v}_\alpha = g(v_\alpha)$.*

**Definition 6.2.** *Define the map $C : U^* \otimes U \to \mathbb{R}$ by*

$$C\left(\sum a_\alpha u_\alpha^* \otimes u_\alpha\right) = a_\alpha \sum u_\alpha^*(u_\alpha)$$

Finally, we note that if we compose the map $C$ with the isomorphism from 4 of Proposition 6.1 (with $U = V$), we obtain a map $Hom(U, U) \to \mathbb{R}$. This map is called the *trace*.

Exercise: Show this map indeed coincides with the trace of an endomorphism as you may have seen it in linear algebra.

# 7 Tensor bundles

Now let $\mathcal{E}$, $\mathcal{E}'$ be vector bundles. We can define $\mathcal{E} \otimes \mathcal{E}'$, etc., in view of Definition 6.1. (This tells us how to make transition functions.) The bundles of the form

$$T\mathcal{M} \otimes \cdots T\mathcal{M} \otimes T^*\mathcal{M} \otimes \cdots T^*\mathcal{M}$$

are known as *tensor bundles*. If there are say $d$ copies of $T\mathcal{M}$, and $d'$ of $T^*\mathcal{M}$, we notate the bundle by $T_d^{d'}\mathcal{M}$, and say the bundle of $d$-contravariant and $d'$-covariant tensors. A basis for the fibres, in local coordinates is given by

$$\frac{\partial}{\partial x^{i_1}} \otimes \cdots \frac{\partial}{\partial x^{i_d}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_{d'}}.$$

The transformation law (omitted for now). Note let $F : \mathcal{M} \to \mathcal{N}$. Then can define

$$F^* : \Gamma\left(\bigotimes_{i=1}^{d'} T^*\mathcal{N}\right) \to \Gamma\left(\bigotimes_{i=1}^{d'} T^*\mathcal{M}\right)$$

How? Exercise.

Get used to the following notation: "Let $A^{i_1 \ldots i_d}_{j_1 \ldots j_{d'}}$ be a tensor" meaning: Let $A : \mathcal{M} \to T_d^{d'}\mathcal{M}$ be a smooth section given in local coordinates by

$$A = A^{i_1 \ldots i_d}_{j_1 \ldots j_{d'}} \frac{\partial}{\partial x^{i_1}} \otimes \cdots \frac{\partial}{\partial x^{i_d}} \otimes dx^{j_1} \otimes \cdots \otimes dx^{j_{d'}}.$$

With the results of Section 6, we can play all sorts of games in the spirit of the above. We can construct the bundle $Hom(\mathcal{E}, \mathcal{E}')$. We can construct a natural isomorphism of bundles $Hom(\mathcal{E}, \mathcal{E}') \cong \mathcal{E}^* \otimes \mathcal{E}'$.

# 8 Riemannian manifolds

A Riemannian metric is to be an inner product on all the fibres, varying smoothly. The point is, in view of the previous section, we can now define what "varying smoothly" means: Since an inner product is a bilinear map $T_p\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R}$, which is also symmetric and positive definite, then it can be considered an element of $(T_p\mathcal{M} \otimes T_p\mathcal{M})^*$, and thus, $T_p^*\mathcal{M} \otimes T_p^*\mathcal{M}$.[17] We will thus define

**Definition 8.1.** *A* Riemannian metric *$g$ on a smooth manifold $\mathcal{M}$ is an element $g \in \Gamma(T^*\mathcal{M} \otimes T^*\mathcal{M})$ such that for all $V, W \in T_p\mathcal{M}$, $g(V,W) = g(W,V)$ and $g(V,V) \geq 0$, with $g(V,V) = 0$ iff $V = 0$. A pair $(\mathcal{M}, g)$, where $\mathcal{M}$ is a smooth manifold and $g$ a Riemannian metric on $\mathcal{M}$, is called a* Riemannian manifold.

In local coordinates we have

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

The symmetry condition $g(V,W) = g(W,V)$ gives in local coordinates $g_{ij} = g_{ji}$.

Note: Comparison with the classical notation. In differential geometry of surfaces, one writes classically expressions like the right hand side of (1). In interpreting this notation, you are supposed to remember that this is a symmetric 2-tensor, and thus you are to replace $dudv$ in our present notation by $\frac{1}{2}(du \otimes dv + dv \otimes du)$. On the other hand, one also encounters expressions like $dudv$ in a completely different context, namely in double integrals. Here, one is supposed to interpret $dudv$ as an antisymmetric 2-tensor, a so-called 2-form, and replace it, in more modern notation, by $du \wedge dv$. To avoid confusion, we will never again see in these notes expressions like $dudv$...

**Definition 8.2.** *Let $\gamma : I \to \mathcal{M}$ be a curve. We define the length of $\gamma$ as*

$$\int_I \sqrt{g(\gamma', \gamma')} dt.$$

*Let $\gamma$ and $\tilde{\gamma}$ be curves in $\mathcal{M}$ going through $p$, such that $\gamma' \neq 0$, $\tilde{\gamma}' \neq 0$. We define the angle between $\gamma$ and $\tilde{\gamma}$ to be*

$$\cos^{-1}(g(\gamma', \tilde{\gamma}')(g(\gamma', \gamma')g(\tilde{\gamma}', \tilde{\gamma}'))^{-1/2}).$$

Note. Invariance under reparametrizations.

**Example 8.1.** *Let $\mathbb{R}^n$ with $g = \sum_i dx^i \otimes dx^i$, and let $\mathcal{M}$ be a submanifold of any dimension, $i : \mathcal{M} \to \mathbb{R}^n$. Then $i^*(g)$ is a Riemannian metric on $\mathcal{M}$.*

**Let us assume from now on that all manifolds are *Hausdorff* and *paracompact*.**

[17]Exercise: make this identification formal in the language of isomorphisms of bundles described in the end of the previous section.

**Definition 8.3.** *A topological space is said to be Hausdorff if for all $x \neq y$ there exist open sets $x \in U$, $y \in V$ such that $U \cap V = \emptyset$.*

**Definition 8.4.** *A topological space is said to be paracompact if every open cover $\{V_\beta\}$ admits a locally finite, refinement, i.e. a collection of open sets $\{U_\alpha\}$ such that for every $p$, there exists only finitely many $U_\alpha$ such that $p \in U_\alpha$.*

Note. We can construct a Riemannian metric by applying the following:

**Theorem 8.1.** *(Whitney) Let $\mathcal{M}^n$ be second countable*[18]*. Then there exists an embedding (homeomorphic to its image with subspace topology) $F : \mathcal{M} \to \mathbb{R}^{2n+1}$.*

Actually, it turns out that all Riemannian metrics arise in this way:

**Theorem 8.2.** *(Nash) Let $(\mathcal{M}, g)$ be a Riemannian manifold. Then there exists an embedding $F : \mathcal{M} \to \mathbb{R}^{(n+2)(n+3)/2}$ such that $g = F^*(e)$ where $e$ denotes the euclidean metric on $\mathbb{R}^{(n+2)(n+3)/2}$.*[19]

Embeddings of the above form are known as *isometric embeddings*.

Although by the above Riemannian geometry is nothing other than the study of submanifolds of $\mathbb{R}^N$ with the induced metric from Euclidean space, the point of view of the above theorem is rarely helpful. We shall not refer to it again in this course.

We can construct on any manifold a Riemannian metric in a much more straightforward fashion using a so-called *partition of unity*.

**Proposition 8.1.** *Let $\mathcal{M}$ be paracompact. Let $\{U_\alpha\}$ be a locally finite atlas such that $\bar{U}_\alpha$ is compact. Then there exists a collection $\chi_\alpha$ of smooth functions $\chi_\alpha : \mathcal{M} \to \mathbb{R}$, compactly supported in $U_\alpha$, such that $1 \geq \chi_\alpha \geq 0$, $\sum \chi_\alpha = 1$.*[20]

We call the collection $\chi_\alpha$ a *partition of unity* subordinate to $\mathcal{U}_\alpha$.

Using this, we can construct a Riemannian metric on any paracompact manifold: We just take $\sum \chi_\alpha e_\alpha$, where $e_\alpha$ denotes the Euclidean metric in local coordinates.

Here we should point out the difference between *geometric topologists* and *Riemannian geometers*. Geometric topologists study smooth manifolds. In the study of such a manifold, it may be useful for them to define a Riemannian metric on it, and to use this metric to assist them in defining more structures, etc. At the end of the day, however, they are interested in aspects that don't depend on which Riemannian metric they happened to construct. An example of topological invariants constructed with the help of a Riemannian metric are the so-called Donaldson invariants and the Seiberg-Witten invariants.

For Riemannian geometers, on the other hand, the objects of study are Riemannian manifolds. You don't get to choose the metric. The metric is given to you, and your task is to understand its properties. The Riemannian geometry

[18] Note that a connected component of a paracompact manifold is second countable.

[19] Note that the original theorem of Nash needed a higher exponent.

[20] Evaluated at any point $p$, this sum is to be interpreted as a finite sum, over the (by assumption!) finitely many indices $\alpha$ where $\chi_\alpha(p) \neq 0$.

of $(\mathcal{M}, g)$ is interesting even if the topology of $\mathcal{M}$ is diffeomorphic to $\mathbb{R}^n$. In fact, for the first half-century of its existence, higher dimensional Riemannian geometry concerned precisely this case.

# 9 Isometries

**Definition 9.1.** *A diffeomorphism $F : (\mathcal{M}, g) \to (\mathcal{N}, \tilde{g})$ is called an* isometry *if $F^*(\tilde{g}) = g$. The manifolds $\mathcal{M}$ and $\mathcal{N}$ are said to be isometric.*

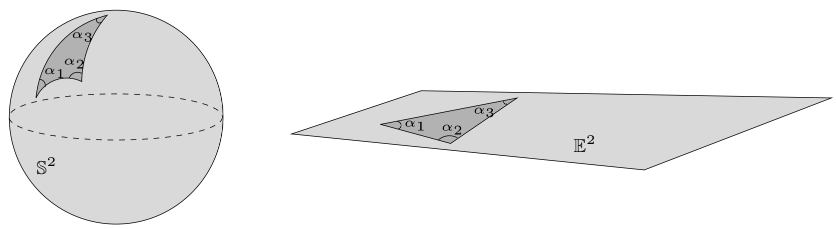
We can also define a local isometry.

**Definition 9.2.** *$(\mathcal{M}, g)$ and $(\mathcal{N}, \tilde{g})$ are* locally isometric *at $p$, $q$, if there exist neighborhoods $\mathcal{U}$ and $\tilde{\mathcal{U}}$ of $p$, $q$, and an isometry $F : \mathcal{U} \to \tilde{\mathcal{U}}$. $F$ is then called a* local isometry*.*

How can we tell if two metrics are really different, i.e. not locally isometric?

**Definition 9.3.** *A curve $\gamma$ is said to be a geodesic if it locally minimizes arc length.*[21]

**Example 9.1.** *The 2-sphere and the plane. The geodesics are great circles, and lines, respectively. In the 2-sphere, the sum of the interior angles of a geodesic triangles is given by $\alpha_1 + \alpha_2 + \alpha_3 = \pi + 2\pi Area$. In the plane $\alpha_1 + \alpha_2 + \alpha_3 = \pi$. Since lengths, areas*[22] *and angles are preserved by local isometries, these spaces can thus not be locally isometric.*




Now we would like in general in Riemannian geometry to figure out which curves are geodesics in the sense described above, and moreover, we want to relate this notion to that of parallelism. The fact that this is possible is one of the fundamental insights of Riemannian geometry, and is due to Levi-Civita.

But first we have to return to vector fields.

# 10 Vector fields and O.D.E.'s

## 10.1 Existence of integral curves

Back to $\mathbb{R}^n$. I will assume the following fact from the theory of ode's[23].

[21] Later, we will find it convenient to define these curves otherwise and infer this property...
[22] as yet undefined...
[23] ode's=ordinary differential equations

**Theorem 10.1.** *Consider the initial value problem*

$$(x^i)' = f^i(x_1, \dots x_n), \tag{7}$$

$$x^i(0) = x^i_0, \tag{8}$$

*where $f$ is a Lipschitz function in $\mathcal{U} \subset \mathbb{R}^n$. Then there exists a unique maximal $(T_-, T_+)$, with $-\infty \leq T_- < 0 < T_+$, and a unique continuously differentiable solution*

$$x^i : (T_-, T_+) \to \mathbb{R}^n$$

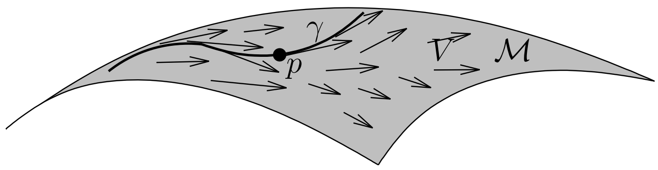
*satisfying (7), (8). If $f$ is smooth then $x$ is smooth. Moreover, if $T_+ < \infty$, then given any compact subset $K \subset U$, there exists a $t_K$ such that $x(t_K, T_+) \cap K = \emptyset$.*

The geometric interpretation of this theorem is:

**Theorem 10.2.** *Let $V$ be a $C^\infty$ vector field on an open subset $\mathcal{U} \subset \mathbb{R}^n$. Then through any point $p$ in $\mathcal{U}$, there exists a maximal parametrized integral curve $\gamma$ of $V$, i.e. a curve $\gamma : (T_-, T_+)$ such that $\gamma' = V$, $\gamma'(0) = p$. Moreover, if $T_+ < \infty$, then given any compact subset $K \subset U$, there exists a $t_K$ such that $x(t_k, T_+) \cap K = \emptyset$.*

In the example sheet you shall show that this can be extended to manifolds as follows:

**Theorem 10.3.** *Let $\mathcal{M}$ be a $C^\infty$ manifold, and let $V$ be a $C^\infty$ vector field on $\mathcal{M}$, i.e. $V \in \Gamma(T\mathcal{M})$. Then through any point $p$ in $\mathcal{U}$, there exists a maximal*[24] *parametrized integral curve $\gamma$ of $V$, i.e. a curve $\gamma : (T_-, T_+)$ such that $\gamma' = V$, $\gamma'(0) = p$.*




*Moreover, if $T_+ < \infty$, then given any compact subset $K \subset U$, there exists a $t_K$ such that $x(t_K, T_+) \cap K = \emptyset$. In particular, if $\mathcal{M}$ itself is compact, then $\gamma$ "exists for all $t$", i.e. $T_\pm = \pm\infty$.*

**Definition 10.1.** *If for all $x$, $T_\pm = \pm\infty$, then we call $X$* complete.

In this language, on a compact manifold $\mathcal{M}$, all vector fields $X$ are complete.

Exercise: Write down a manifold and an incomplete vector field. Write down a non-compact manifold and a complete vector field. Does every non-compact manifold admit an incomplete vector field?

[24] i.e. a curve not a subset of a larger such curve

## 10.2 Smooth dependence on initial data; 1-parameter groups of transformations

Classical O.D.E. theory tells us more than Theorem 10.1. It tells us that solutions depend continuously (in the Lipschitz case) and smoothly (in the smooth case) on initial conditions.

To formulate this in the language of vector fields, let $V \in \Gamma(T\mathcal{M})$.

**Proposition 10.1.** *For every $p \in \mathcal{M}$ there exists an open set $\mathcal{U}$, a nonempty open interval $I$ and a collection of local transformations*[25]

$$\phi_t : \mathcal{U} \to \mathcal{M},$$

*such that $\phi_t(q)$ is the integral curve of $V$ through $q$ given by Proposition 10.1. Moreover $\phi : \mathcal{U} \times I \to \mathcal{M}$ is a smooth map, and*

$$\phi_t \circ \phi_s = \phi_{t+s}, \tag{9}$$

*whenever $t, s, t+s \in I$.*

A family of local transformations satisfying (9) is called a 1-parameter local group of transformations. If $I = \mathbb{R}$ then $\phi_t$ are in fact "global" and (9) defines a group structure on $\{\phi_t\}$.

Note that in particular, the above theorem says that $|T_\pm|$ can be uniformly bounded below in a neighborhood of any point.

It is easy to see using (9) that there is a one to one correspondence between local 1-parameter groups of transformations and vector fields. Check the following: Given such a family $\phi_t$, define $X(p)$ to be the tangent vector of $\phi_t(p)$ at $t = 0$. The local 1-parameter group of transformations associated to $X$ is again $\phi_t$.

## 10.3 The Lie bracket

Let $\mathcal{M}$ be a smooth manifold, let $X$ and $Y$ be smooth vector fields: $X, Y \in \Gamma(T\mathcal{M})$.

**Definition 10.2.** $[X, Y]$ *is the vector field defined by the derivation given by*

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

**Claim 10.1.** $[X, Y]$ *is indeed a derivation.*

*Proof.* Check the properties of a derivation! □

**Proposition 10.2.** *The following hold*

1. $[X, Y] = -[X, Y]$

[25] i.e. a smooth map $\mathcal{U} \to \mathcal{M}$, where $\mathcal{U} \subset \mathcal{M}$, such that the map is a diffeomorphism to its image

*2.* $[X_1 + X_2, Y] = [X_1, Y] + [X_2, Y]$

*3.* $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y]$ *(Jacobi identity)*

*4.* $[fX, gY] = fg[X, Y] + f(Xg)Y - g(Yf)X$

The proof of this proposition is a straightforward application of the definition, left to the reader.

So we can say that $\Gamma(T\mathcal{M})$ is a (non-associative) algebra with the bracket operation as multiplication. In general, algebras whose multiplication satisfies 3 above are known as *Lie algebras*. Note finally, that if $\phi$ is a diffeomorphism then

$$\phi_*[X, Y] = [\phi_* X, \phi_* Y]. \tag{10}$$

We say that $\phi_*$ is a *Lie algebra isomorphism*.

The above Proposition allows us to easily obtain a formula for $[X, Y]$ in terms of local coordinates. If $x^i$ is a system of local coordinates, first note that

$$\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] = 0.$$

Now using in particular identity 4 of Proposition 10.2, setting $X = X^i \frac{\partial}{\partial x^i}$, $Y = Y^i \frac{\partial}{\partial x^i}$, we have

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

Geometric interpretation of $[X, Y]$. Let $\phi_t$ denote the one-parameter group of transformations corresponding to $X$

**Proposition 10.3.**

$$[X, Y]|_p = \lim_{t \to 0} t^{-1}(Y|_p - ((\phi_t)_* Y)_p). \tag{11}$$

*Proof.* Let $f_t$ denote $f \circ \phi_t$. Claim: $f_t = f + t(Xf) + t^2 h_t$ where $h_t$ is smooth.

Now, we clearly have

$$(\phi_t)_* Y)_p f = Y_{\phi^{-1}(p)}(f \circ \phi_t).$$

Thus the right hand side of (11) applied to $f$ is

$$\begin{aligned}
\lim t^{-1}(Y|_p f - Y_{\phi^{-1}(p)}(f \circ \phi_t)) &= \lim t^{-1}(Y_p f - Y_{\phi^{-1}(p)}(f + t(Xf) + t^2 h_t)) \\
&= \lim t^{-1}(Y_p f - T_{\phi^{-1}(p)} f) - Y_p(Xf) \\
&= X_p(Yf) - Y_p(Xf) \\
&= [X, Y]_p
\end{aligned}$$

as desired. $\square$

**Proposition 10.4.** *Let $\phi$ be a diffeomorphism. If $\phi_t$ generates $X$, then $\phi^{-1} \circ \phi_t \circ \phi$ generates $\phi_* X$. In particular, for $\phi$ and $\phi_t$ to commute, we must have $\phi_* X = X$.*

**Proposition 10.5.** $[X, Y] = 0$ *if and only if the* 1*-parameter local groups of transformation commute.*

*Proof.* Apply (11) to $(\phi_s)_* Y$, use (10) and the relationship $(\phi_s)_* \circ (\phi_t)_* = (\phi_{s+t})_*$. $\square$

# 11 Geodesics and parallelism in $\mathbb{R}^n$

With our toolbox from the theory of ode's full, let us now return to the study of geometry.

We will begin with a discussion of the relevant concepts in the special case of Euclidean space.

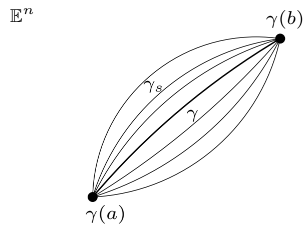
Let us call *geodesics* in $\mathbb{R}^n$ curves

$$\gamma : I \to \mathbb{R}^n$$

which extremize arc length in the following sense: Let $I = [a, b]$, $V = (-\epsilon, \epsilon)$, and consider a smooth[26] map

$$\tilde{\gamma} : I \times V \to \mathbb{R}^n$$

such that $\gamma_s = \tilde{\gamma}|_{I \times \{s\}}$ is a smooth curve in $\mathbb{R}^n$ with $\gamma_0(t) = \gamma(t)$ for all $t \in I$, $\gamma_s(a) = \gamma(a)$, $\gamma_s(b) = \gamma(b)$ for all $s \in J$. We shall call $\tilde{\gamma}$ a smooth variation of $\gamma$.




Define $L(s)$ to be the length of the curve $\gamma_s$. We would like to derive conditions for $s = 0$ to be a critical point of $L$ for all smooth variations $\tilde{\gamma}$.

Let us for convenience assume $\gamma_0$ is parametrized by arc length, i.e. $|\partial_t \gamma_0(t)| =$

[26] Exercise: define this in view of the fact that $[a, b]$ is closed.

1. We compute

$$\begin{aligned}
L'(s)|_{s=0} &= \frac{d}{ds}\int_a^b \sqrt{\partial_t\gamma_s\cdot\partial_t\gamma_s}dt|_{s=0} \\
&= \int_a^b \partial_s\sqrt{\partial_t\gamma_s\cdot\partial_t\gamma_s}dt|_{s=0} \\
&= \int_a^b \partial_t\gamma_s\cdot\partial_s\partial_t\gamma_s dt|_{s=0} \\
&= \int_a^b \partial_t\gamma_s\cdot\partial_t\partial_s\gamma_s dt|_{s=0} \\
&= \int_a^b \partial_t(\partial_t\gamma_s\cdot\partial_s\gamma_s)-\partial_t\partial_t\gamma_s\cdot\partial_s\gamma_s dt|_{s=0} \\
&= \partial_t\gamma_0\cdot\partial_s\gamma_s|_{s=0}(b)-\partial_t\gamma_0\cdot\partial_s\gamma_s|_{s=0}(a)-\int_a^b \partial_t\partial_t\gamma_s\cdot\partial_s\gamma_s)dt|_{s=0} \\
&= \int_a^b \frac{d^2}{dt^2}\gamma_0\cdot\partial_s\gamma_s|_{s=0}dt.
\end{aligned}$$

Now (exercise) it is easy to see that one can construct a variation $\tilde\gamma$ of $\gamma$ such that $\partial_s\gamma_s(t)|_{s=0}$ for $t\in(a,b)$, is an arbitrary smooth vector field along[27] $\gamma$, vanishing at the endpoints. Thus, for the identity $L'(0)=0$ to hold for all variations $\tilde\gamma$,[28] we must have

$$\frac{d^2}{dt^2}\gamma_0 = 0. \tag{12}$$

This is the geodesic equation in $\mathbb{R}^n$. It is a second order ode. The general solution is

$$\gamma(t) = (x_0^1+a^1t,\ldots,x_0^n+a^nt),$$

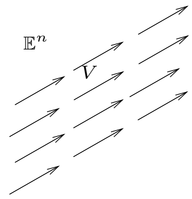
i.e. straight lines.

We didn't need any of the so-called qualitative theory of Section 10 to say that there exist solutions to (12), for we could just write them down explicitly! This is related to the following fact: In the case of $\mathbb{R}^n$, it turns out that straight lines are distinguished not only in the variational sense just discussed above but also from the group action point of view. For on $\mathbb{R}^n$ we have the well known translations, which act by isometry. Given a vector $V_p$ at a point $p$, we can construct a vector field $V:\mathbb{R}^n\to T\mathbb{R}^n$ such that $V(q)=(T_q^p)_*(V_p)$, where $T_q^p$ denotes the translation map $\mathbb{R}^n\to\mathbb{R}^n$ which sends $p$ to $q$. Geodesics through $p$

[27] For a formal definition of this, see Section 14.

[28] Remember, $L$ depends on the variation $\tilde\gamma$, i.e. we should really write $L^{\tilde\gamma}$.

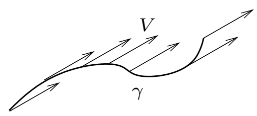
tangent to $V_p$ are then integral curves of the vector field $V$.




Vector fields $V$ constructed as above are known as *parallel*. Geodesics are thus curves whose tangent vector is parallel.

It is somewhat of a miracle that in Euclidean geometry we can identify certain vector fields as parallel so as for this notion to relate to geodesics (defined as length extremizers) in the above sense.

In Riemannian geometry, things are not as simple. An *absolute parallelism* in the sense above does not exist. Nonetheless, one may still define the notion of a vector field being parallel *along a curve*:




More generally, we may still define the notion of the directional derivative of a vector field $X$ in the direction of a vector $\xi$, to be denoted $\nabla_\xi X$. The vector field $X$ will be called parallel along a curve $\gamma$ if $\nabla_\xi X = 0$. In particular, we shall be able to define this so that the equation for length extremizing curves is again $\nabla_{\gamma'}\gamma' = 0$, i.e. so that length extremizing curves can again be characterized as those whose tangent is parallel along itself.

At this point one should stop and point out that it is truly remarkable that one can again relate length-extremization and a suitable notion of parallelism, albeit, more restricted than that of an absolute parallelism. The realization that this concept is useful is essentially the contribution of Levi-Civita to the subject of Riemannian geometry.

The task of defining this operation $\nabla$ belongs to the next section.

# 12 Connections

Our goal in the next section is to relate to a Riemannian manifold $(\mathcal{M}, g)$, an operation $\nabla$ which will allow us to call certain vector fields parallel along curves, and more generally, will allow us to differentiate vector fields along curves, the ones with vanishing derivative called parallel. It turns out, however, that a $\nabla$ operation is a useful concept in more general contexts, independent of Riemannian geometry. Let us start thus in more generality.

**Definition 12.1.** *Let $\mathcal{M}$ be a smooth manifold, and let $\pi : \mathcal{E} \to \mathcal{M}$ be a vector bundle. A* connection $\nabla$ *on $\mathcal{E}$ is a mapping*

$$\nabla : T\mathcal{M} \times \Gamma(\mathcal{E}) \to \mathcal{E}$$

*(we will write $\nabla(\xi, X)$ as $\nabla_\xi X$!) with the following properties:*

1. *If $\xi \in T_p\mathcal{M}$ then $\nabla_\xi X \in \mathcal{E}_p$*
2. $\nabla_{(a\xi+b\tilde{\xi})}X = a\nabla_\xi X + b\nabla_{\tilde{\xi}}X$
3. $\nabla_\xi(X+Y) = \nabla_\xi X + \nabla_\xi Y$
4. $\nabla_\xi fX = (\xi f)X + f\nabla_\xi X$[29]
5. *If $Y \in \Gamma(T\mathcal{M})$, then $p \mapsto \nabla_{Y(p)}X$ is an element of $\Gamma(\mathcal{E})$, i.e. is smooth.*

In this class we will be interested in connections on the tangent bundle and related tensor bundles.

**Example 12.1.** *The flat connection on the tangent bundle of $\mathbb{R}^n$. Let $\frac{\partial}{\partial x^i}$ denote standard coordinates on $\mathbb{R}^n$. (We often call such coordinates Euclidean coordinates.) Let us define $(\nabla_\xi X)^j = \xi^i \partial_i X^j$. Check that this is indeed a connection, and that $\nabla_\xi X = 0$ iff it is parallel in the sense described previously.*

Let $\nabla$ be a connection in the tangent bundle $T\mathcal{M}$, and let $x^i$ be a system of local coordinates. Let us introduce the symbols $\Gamma^i_{jk}$ by

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma^k_{ij} \frac{\partial}{\partial x^k}.$$

Note as always the Einstein summation convention. By the defining properties of connections, the $\Gamma^i_{jk}$ determine the connection completely by the following formulas: Let $\xi = \xi^i \frac{\partial}{\partial x^i}$, and let $X^i = X^i \frac{\partial}{\partial x^i}$,

$$\begin{aligned} \nabla_\xi X &= \xi^i \frac{\partial X^j}{\partial x^i} \frac{\partial}{\partial x^j} + \Gamma^k_{ij} \xi^i X^j \frac{\partial}{\partial x^k} \\ &= \left.\frac{dX^j \circ \gamma(t)}{dt}\right|_{t=0} \frac{\partial}{\partial x^j} + \Gamma^k_{ij} \xi^i X^j \frac{\partial}{\partial x^k}, \end{aligned} \tag{13}$$

where $\gamma(t)$ is any curve in $\mathcal{M}$ with $\gamma'(0) = \xi$.

Clearly, connections can be constructed by prescribing arbitrarily the functions $\Gamma^k_{ij}$ and patching together with partitions of unity.

# 13 The Levi-Civita connection

Let us now return to Riemannian manifolds $(\mathcal{M}, g)$. It turns out that there exists a distinguished connection $\nabla$ that one can relate to $g$:

**Proposition 13.1.** *Let $(\mathcal{M}, g)$ be Riemannian. There exists a unique connection $\nabla$ in $T\mathcal{M}$ characterized by the following two properties*

[29] Here we are dropping evaluation at a point $p$ from notation. To check the syntax of the formulas, always remember that vectors act on functions and that $\Gamma(\mathcal{E})$ is a module over $C^\infty(\mathcal{M})$.

1. *If $X$, $Y$ are vector fields then,* $\nabla_{X(p)}Y - \nabla_{Y(p)}X = [X,Y](p)$.

2. *If $X$, $Y$ are vector fields and $\xi$ is a vector then* $\nabla_\xi g(X,Y) = g(\nabla_\xi X, Y) + g(X, \nabla_\xi Y)$.

*Proof.* Compute $g(\nabla_X Y, Z)$ explicitly using the rules above, and show that this gives a valid connection. ☐

The Levi-Civita connection in local coordinates. First some notation. We will define the inverse metric $g^{ij}$ as the components of the bundle transformation $T^*\mathcal{M} \to T\mathcal{M}$ inverting the isomorphism $T\mathcal{M} \to T^*\mathcal{M}$ enduced by the Riemannian metric $g$. More pedestrianly, it is the inverse matrix of $g_{ij}$, i.e. we have $g^{ij}g_{jk} = \delta^i_k$ where $\delta^i_k = 1$ if $i = k$ and 0 otherwise. Check that

$$\Gamma^k_{ij} = \frac{1}{2} g^{kl}(\partial_j g_{il} + \partial_i g_{jl} - \partial_l g_{ij}).$$

(Check also that the first condition in the definition of the connection is equivalent to the statement $\Gamma^k_{ij} = \Gamma^k_{ji}$.)

## 13.1 Raising and lowering indices with the metric

Since we have just used the so-called inverse metric, we might as well discuss this topic now in more detail. This is the essense of the power of index notation.

The point is that given any tensor, i.e. a section of

$$T^*\mathcal{M} \otimes \cdots \otimes T^*\mathcal{M} \otimes T\mathcal{M} \otimes \cdots \otimes T\mathcal{M}$$

we can apply the bundle isomorphism defined by the inverse metric on any of the $T^*\mathcal{M}$ factors so as to convert it into a $T\mathcal{M}$. And similarly, we can apply the isomorphism defined by the metric itself to turn any of the factors $T\mathcal{M}$ to convert it to a $T^*\mathcal{M}$.

It is traditional in local coordinates to use the same letter for all the tensors one obtains by applying these isomorphisms to a given tensor. I.e., if $S^{j_1,\ldots j_m}_{i_1,\ldots i_n}$ is a tensor, then the tensor produced by applying the above isomorphism say to the factor corresponding to the index $i_k$ is given in local coordinates by

$$S^{j_1\ldots j_n h}_{i_1\ldots \hat{i_k}\ldots i_n} = g^{i_k h} R^{j_1\ldots j_m}_{i_1\ldots i_n}.$$

The hat above denotes that the index is omitted.

This process is known as raising and lowering indices.

Thus, using the metric, we can convert a tensor to one with all indices up or down, or however we like, and we think of these, as in some sense being the "same" tensor.

Finally, this process can be combined with the contraction map. For if say $S_{ijkl}$ is a tensor, then we can raise the index $i$ to obtain $S^i_{jkl}$, and now apply the contraction map of Definition 6.2 on the factors corresponding to the indices $i$ and $j$ to obtain a tensor $S_{kl}$. Again, one often uses the same letter to denote this new tensor, as we have just done here, although there is a potential for confusion, as one can define several different contractions, depending on the indices selected.

### 13.2 Levi-Civita and local isometry

Try and prove the following

**Proposition 13.2.** *Let $(\mathcal{M}, g)$, $(\tilde{\mathcal{M}}, \tilde{g})$ be Riemannian and suppose that $p \in \mathcal{U} \subset \mathcal{M}$, $q \in \tilde{\mathcal{U}} \subset \tilde{\mathcal{M}}$, and $\phi : \mathcal{U} \to \tilde{\mathcal{V}}$ is an isometry with $\phi(p) = q$. Let $\gamma$ be a curve in $\mathcal{U}$, and let $V$ be a vector field along $\gamma$ and let $T$ be a vector tangent to $\gamma$. Let $\nabla$ and $\tilde{\nabla}$ denote the Levi-Civita connections of $\mathcal{M}$, $\tilde{\mathcal{M}}$, respectively. Then*

$$\phi_* \nabla_T V = \tilde{\nabla}_{\phi_* T} \phi_* V.$$

*In particular, if $V$ is parallel along $\gamma$ (i.e., $\nabla_T V = 0$), then $\phi_* V$ is parallel along $\phi \circ \gamma$.*

## 14 Geodesics and the first variation

We may now make the definition

**Definition 14.1.** *Let $(\mathcal{M}, g)$ be a Riemannian manifold with Levi-Civita connection $\nabla$. A curve $\gamma : I \to \mathcal{M}$ is said to be a geodesic if*

$$\nabla_{\gamma'} \gamma' = 0. \tag{14}$$

Strictly speaking, equation (14) does not make sense, since $\gamma'$ is a vector field along $\gamma$, i.e. it can be though of as a section of the bundle $\gamma^*(T\mathcal{N}) \to I$. Nonetheless, we can use formula (13) to *define* the left hand side of (14). In general, when $V$ is a vector field along a curve $\gamma$, and $W$ is a vector tangent to $\gamma$, we will use unapologetically the notation $\nabla_W V$. Similarly for vector fields "along" higher dimensional submanifolds, defined in the obvious sense. (Exercise: What is this obvious sense?)

Note that in view of Proposition 13.2, local isometries map geodesics to geodesics. (Show it!)

It turns out that the above notion of geodesic coincides with that of curves locally minimizing arc length:

**Theorem 14.1.** *Let $\gamma$ be a geodesic. Then for all $p \in \gamma$, there exists a neighborhood $\mathcal{U}$ of $p$, so that for all $q, r \in \gamma \cap \mathcal{U}$, denoting by $\gamma_{q,r}$ the piece of $\gamma$ connecting $q$ and $r$, we have $d(q, r) = L(\gamma_{q,r})$, where $L$ here denotes length, and moreover, if $\tilde{\gamma}$ is any other piecewise smooth curve in $\mathcal{M}$ connecting $q$ and $r$, then $L(\tilde{\gamma}) > d(q, r)$.*

The proof of Theorem 14.1 is not immediate, and in fact, reveals various key ideas in the calculus of variations. We will complete the proof several sections later in these notes. For now let us give the following:

**Proposition 14.1.** *A $C^2$ curve $\gamma : [a, b] \to \mathcal{M}$ is a geodesic parametrized by a multiple of arc length iff for all $C^2$ variations $\tilde{\gamma} : [a, b] \times [-\epsilon, \epsilon]$ of $\gamma$ with $\tilde{\gamma}(a, s) = \gamma(a)$, $\tilde{\gamma}(b, s) = \gamma(b)$, we have*

$$\frac{d}{ds} L\left(\tilde{\gamma}(\cdot, s)\right)\big|_{s=0} = 0.$$

It is this Proposition that relates parellelism with length extremization, i.e. that allows us to recover the analogue in Riemannian geometry of the picture of Section 11.

*Proof.* Let $\gamma$ be a $C^2$ curve, and let $\tilde{\gamma}$ be an arbitrary variation of $\gamma$.

Let us introduce the notation $N = \tilde{\gamma}_* \frac{\partial}{\partial s}$, $T = \tilde{\gamma}_* \frac{\partial}{\partial t}$, where $s$ is a coordinate in $(-\epsilon, \epsilon)$ and $t$ is a coordinate in $[0, L]$. And define $L(s)$ to be the length of the curve $\tilde{\gamma}(\cdot, s)$.

We have

$$L(s) = \int_a^b \sqrt{g(T_{\tilde{\gamma}(s,t)}, T_{\tilde{\gamma}(s,t)})}dt.$$

**We now have the technology to mimick the calculation in Section 11.** Differentiating $L$ in $s$, we obtain

$$\begin{aligned}
L'(s) &= \frac{d}{ds}\int_a^b \sqrt{g(T,T)}dt \\
&= \int_a^b N\sqrt{g(T,T)}dt \\
&= \int_a^b (g(T,T))^{-1/2} g(\nabla_N T, T)dt \\
&= \int_a^b (g(T,T))^{-1/2} g(\nabla_T N, T)dt \\
&= \int_a^b T((g(T,T))^{-1/2} g(N,T)) \\
&\quad - T(g(T,T))^{-1/2})g(N,T) - (g(T,T))^{-1/2} g(N, \nabla_T T)dt \qquad (15) \\
&= g(T,T)^{1/2} g(N,T)]_a^b \\
&\quad - \int_a^b T(g(T,T))^{-1/2} g(N,T) - (g(T,T))^{-1/2} g(N, \nabla_T T)dt \qquad (16) \\
&= \int_a^b -T(g(T,T))^{-1/2} g(N,T) - (g(T,T))^{-1/2} g(N, \nabla_T T)dt \\
&= \int_a^b g(T,T)^{-3/2} g(\nabla_T T, T) g(N,T) - (g(T,T))^{-1/2} g(N, \nabla_T T)dt \qquad (17)
\end{aligned}$$

Here we have used $[N,T] = 0$, $\nabla_N T - \nabla_T N = 0$, and the fact that $N(a,s) = 0$, $N(b,s) = 0$. Exercise: Why are these statements true?

Now suppose that $\gamma$ is a geodesic in the sense of Definition 14.1. Since $\nabla_T T|_{\tilde{\gamma}(0,t)} = 0$, the whole expression on the right hand side of (15) vanishes when evaluated at $s = 0$. Since $\tilde{\gamma}$ is arbitrary, this proves one direction of the equivalence.

To prove the other direction, first we note that given any vector field $N$ along $\gamma$, there exists some variation $\tilde{\gamma}$ such that $N = \tilde{\gamma}_* \frac{\partial}{\partial s}$. (A nice way to construct such a vector field is via the exponential map discussed in later sections. But this is not necessary.) Thus it suffices to show that if $T$ does not satisfy $\nabla_T T = 0$, then there exists an $N$ such that the expression on the right hand side of (15) is non-zero.

Suppose then that $\nabla_T T(t_0, 0) \neq 0$. There exists a neighborhood $(t_1, t_2)$ of $t_0$ such that $\nabla_T T(t, 0) \neq 0$ for $t \in (t_1, t_2)$. Let $N$ be the vector field along $\gamma$ such that $N(t) = \nabla_T T(t, 0)$. We have

$$
\begin{aligned}
&\int_a^b g(T,T)^{-3/2} g(\nabla_T T, T) g(N, T) - (g(T,T))^{-1/2} g(N, \nabla_T T) dt \\
&= \int_a^b g(T,T)^{-3/2} g(\nabla_T T, T) g(\nabla_T T, T) - g(T,T))^{-1/2} g(\nabla_T T, \nabla_T T) dt \\
&\leq \int_{t_1}^{t_2} g(T,T)^{-3/2} g(\nabla_T T, T) g(\nabla_T T, T) - g(T,T))^{-1/2} g(\nabla_T T, \nabla_T T) dt \\
&< 0.
\end{aligned}
$$

The last inequality follows by noting that the above expression is $-g(T,T)^{-1/2}$ times the norm squared of the projection of $\nabla_T T$ to the orthogonal complement of $T$, and the latter is certainly nonnegative.

Thus, we must have $\nabla_T T = 0$. ☐

# 15 Parallel transport

Our goal is to prove the existence of geodesics by reducing to the theory of ordinary differential equations. The geodesic equation in local coordinates is of course a second order equation for $\gamma$. It is a first order equation for the tangent vector.

Let us first consider the following simpler situation. Let $\gamma : I \to \mathcal{M}$ be a fixed smooth curve, $I = [a, b]$, denote $\gamma(a)$ as $p$, $\gamma(b)$ as $q$, and let $T$ denote the tangent vector $\gamma'$. Suppose $V$ is an arbitrary vector at $p$.

**Proposition 15.1.** *There exists a unique smooth vector field $\tilde{V}$ along $\gamma$ such that $\tilde{V}(a) = V$ and*

$$\nabla_T \tilde{V} = 0, \tag{18}$$

*i.e. so that $\tilde{V}$ is parallel along $\gamma$.*

*Proof.* Writing (18) for the components $\tilde{V}^x$ of $\tilde{V}$ with respect to a local coordinate system, we obtain

$$\frac{d}{dt} \tilde{V}^\alpha = -\Gamma^\alpha_{\beta\gamma} T^\beta(\gamma(t)) \tilde{V}^\gamma. \tag{19}$$

Let us consider the manifold $(-\infty, \infty) \times \mathbb{R}^n$, and consider the vector field

$$(s, x) \mapsto (s, -\Gamma^\alpha_{\beta\gamma} T^\beta(\gamma(t)) x^\beta),$$

where we set $\Gamma^\alpha_{\beta\gamma} T^\beta(\gamma(s)) = \Gamma^\alpha_{\beta\gamma} T^\beta(\gamma(a))$ for $s < a$, and $\Gamma^\alpha_{\beta\gamma} T^\beta(\gamma(s)) = \Gamma^\alpha_{\beta\gamma} T^\beta(\gamma(b))$ for $s > b$. Then integral curves $(s(t), \tilde{V}(t))$ of this vector field are precisely solutions of (19) for $t$ values in $[a, b]$, after dropping the $s(t)$ which clearly must satisfy $s(t) = t$.[30] Thus, by Theorem 10.3, we have that for the initial value problem at $t = a$, there exists a maximum future time $T$ of existence, and a solution $\tilde{V}(t)$ on $[a, T)$.

On the other hand, since the equation (19) is linear, we know a priori that a solution $\tilde{V}$ is bounded by

$$\sum_\delta |\tilde{V}^\delta(t-a)| \leq \sum_\delta |V^\delta| \exp\left(\left(\sup_{t,\alpha,\gamma} \Gamma^\alpha_{\beta\gamma} T^\beta\right) |t-a|\right)$$

Thus, $(s(t), \tilde{V}(t))$ cannot leave every compact subset of $(-\infty, \infty) \times \mathbb{R}^n$ in finite time, and thus $T = \infty$. In particular, $\tilde{V}$ is defined in all of $[a, b]$ $\square$

We call the vector $W = \tilde{V}(q)$ the *parallel transport* of $V$ to $q$ along $\gamma$. One easily sees that parallel transport defines an isometry $T_\gamma : T_p\mathcal{M} \to T_q\mathcal{M}$ of tangent spaces.

# 16 Existence of geodesics, the exponential map

Now for the existence of geodesics:

Since the geodesic equation is second order, to look for a first order equation, we must go to the tangent bundle.[31] Let $x^\alpha$ be a system of local coordinates on $\mathcal{M}$. Extend this to a system of local coordinates $(x^1, \ldots x^n, p^1, \ldots p^n)$ on $T\mathcal{M}$, where the $p^\alpha$ are defined by

$$V = \sum p^\alpha(V) \frac{\partial}{\partial x^\alpha}$$

for any vector $V$.

The geodesic equation (14), which in local coordinates can be written in second order form

$$\frac{d^2 x^\alpha}{dt^2} = -\Gamma^\alpha_{\beta\gamma} \frac{dx^\beta}{dt} \frac{dx^\gamma}{dt},$$

can now be written as

$$\frac{dx^\alpha}{dt} = p^\alpha \tag{20}$$

$$\frac{dp^\alpha}{dt} = -\Gamma^\alpha_{\beta\gamma} p^\beta p^\gamma. \tag{21}$$

Solutions of the system (20)–(21) are just integral curves on $T\mathcal{M}$ of the vector field

$$p^\alpha \frac{\partial}{\partial x^\alpha} - \Gamma^\alpha_{\beta\gamma} p^\beta p^\gamma \frac{\partial}{\partial p^\alpha}$$

[30] We have just here performed a well known standard trick from ode's for turning a so-called non-autonomous system to an autonomous system.

[31] Again, this is just a sophisticated version of the well known trick from ode's of making a second order equation first order.

on $\mathcal{TM}$. Remember, the latter is an element of $\Gamma(T(\mathcal{TM}))$. Don't be too confused by this...

We now apply Theorem 10.3, and Proposition 10.1. We call the one-parameter local group of transformations $\phi_t : T\mathcal{M} \to T\mathcal{M}$ generated by this vector field geodesic flow.

Projections $\pi \circ \phi_t$ to $\mathcal{M}$ are then the geodesics we have been wanting to construct. We have thus shown in particular the following:

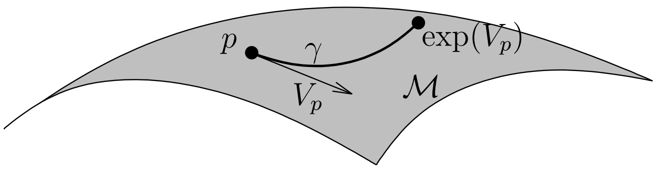
**Proposition 16.1.** *Let $V_p \in T\mathcal{M}$. Then there exists a unique maximal arc-length-parametrized geodesic $\gamma : (T_-, T_+) \to \mathcal{M}$ such that $\gamma'(0) = V_0$.*

Thus, we have shown the existence of geodesics.

Back to $\phi_t$. It is easy to see that the domain $\mathcal{U}$ of $\phi_t$ is star-shaped in the sense that if $V_p \in \mathcal{U}$, for some vector $V_p$ at a point $p$, then $\lambda V_p \in \mathcal{U}$ for all $0 \le \lambda \le 1$. Moreover, (exercise) $\phi_t(V_p) = \phi_{\lambda^{-1}t}(\lambda V_p)$. This implies, that $\phi_1$ is defined in a non-empty star-shaped open set.

**Definition 16.1.** *The map* $\exp : \mathcal{U} \to \mathcal{M}$ *defined by $\pi \circ \phi_1$, where $\pi$ denotes the standard projection $\pi : T\mathcal{M} \to \mathcal{M}$ and $\mathcal{U}$ denotes the domain of $\phi_1$, is called the* exponential map.

The map is depicted below:




The curve $\gamma(t)$ is a geodesic tangent to $V_p$, parametrized by arc length and $\exp(V_p) = \gamma(|V_p|t)$. Here $|V_p| = \sqrt{g(V_p, V_p)}$.

As a composition of smooth maps, the exponential map is clearly a smooth map of manifolds. In the next section, we shall compute its differential.

We end this section with a definition:

**Definition 16.2.** *Let $(\mathcal{M}, g)$ be Riemannian. We say that $(\mathcal{M}, g)$ is* geodesically complete *if the domain $\mathcal{U}$ of the exponential map is $T\mathcal{M}$.*

Equivalently, $(\mathcal{M}, g)$ is geodesically complete if all geodesics can be continued to arbitrary positive and negative values of an arc length parameter.

# 17 Geodesic polar coordinates; Gauss lemma

## 17.1 The differential of exp

First a computation promised at the end of the last section. What is $\exp_*$?

First, what seems like a slightly simpler situation: for any $p \in \mathcal{M}$ let us denote by $\exp_p$ the restriction of the map exp to $\mathcal{U}_p = \mathcal{U} \cap T_p\mathcal{M}$. This is also clearly a smooth map.

We will compute $\left(\exp_{p*}\right)_{0_p}$. It turns out that this is basically a tautology. The only difficulty is in the notation. Remember

$$\left(\exp_{p*}\right)_{0_p} : T_{0_p}(T_p\mathcal{M}) \to T_p\mathcal{M}$$

On the other hand, in view of the obvious[32]

$$T_{0_p}(T_p\mathcal{M}) \cong T_p\mathcal{M}, \tag{22}$$

we can consider the map as a map:

$$\left(\exp_{p*}\right)_{0_p} : T_p\mathcal{M} \to T_p\mathcal{M}.$$

**Proposition 17.1.** *We have*

$$\left(\exp_{p*}\right)_{0_p} = id \tag{23}$$

*Proof.* Let $v \in T_p\mathcal{M}$, and consider the curve $t \mapsto tv$. Denote this curve in $T_p\mathcal{M}$ by $\kappa(t)$. This curve is tangent to $v$. The curve $\exp_p(\kappa(t))$ is found by noting

$$\exp_p(\kappa(t)) = \exp_p(tv) = \phi_t(v) = \gamma(|v|t) \doteq \tilde{\gamma}(t)$$

where $\phi_t$ denotes geodesic flow, and $\gamma$ denotes the arc-length geodesic through $p$ tangent to $v$. By definition of the differential map, we have that

$$(\exp_*)_{0_p} = \tilde{\gamma}'(t) = v,$$

thus, we have obtained (23). ☐

A little more work (equally tautological as the above) show that

$$(\pi \times \exp_*)_{0_p} : T_p\mathcal{M} \times T_p\mathcal{M} \to T_p\mathcal{M} \times T_p\mathcal{M} = id \tag{24}$$

Here it is understood that

$$\pi \times \exp : T\mathcal{M} \to \mathcal{M} \times \mathcal{M}$$

$\pi \times \exp(v_p) = (p, \exp(v_p))$. Again, the domain and range of the differential map have been identified with $T_p\mathcal{M} \times T_p\mathcal{M}$ by obvious identifications analogous to (22) that the reader is here meant to fill in.

What is the point of all this? We can now apply the following inverse function theorem

**Theorem 17.1.** *Let $F : \mathcal{M} \to \mathcal{N}$ be a smooth map such that $(F_*)_p : T_p\mathcal{M} \to T_q\mathcal{N}$ is invertible. Then there exists a neighborhood $\mathcal{U}$ of $p$, such that $F|_\mathcal{U}$ is a diffeomorphism onto its image.*

*Proof.* Prove this from the inverse function theorem on $\mathbb{R}^n$. ☐

[32] define it!

Applied to the map $\pi \times \exp$, in view of this gives the following:

**Proposition 17.2.** *Let $p \in (\mathcal{M}, g)$. There exists a neighbhorhood $\tilde{\mathcal{U}} = \mathcal{U} \times \mathcal{U}$ of $(p,p)$ in $\mathcal{M} \times \mathcal{M}$ such that, denoting by $\mathcal{W} = (\pi \times \exp)^{-1}(\tilde{\mathcal{W}})$, we have that$\pi \times \exp|_{\mathcal{W}} : \mathcal{W} \to \mathcal{U} \times \mathcal{U}$ is invertible.*

That is to say, for any points $q_1, q_2 \in \mathcal{U}$, there exists a $v_{q_1} \in T_{q_1}\mathcal{M}$ such that $\exp(v_{q_1}) = q_2$. (Exercise in tautology: why does this statement follow from the proposition?)

A moment's thought tells us that we can slightly refine the above Proposition. Let us choose $\epsilon > 0$ so that

$$B_{0_q}(\epsilon) \cap \mathcal{W} = \cup q \in \mathcal{U} B_{0_q}(\epsilon) \doteq \tilde{\mathcal{W}}.$$

Again, $\pi \times \exp|_{\mathcal{W}}$ is a diffeomorphism, and its projection to the first component is $\mathcal{U}$. Let $\mathcal{V} \subset \mathcal{U}$ such that $\mathcal{V} \times \mathcal{V}$ is in the image of this. Let $q_1, q_2 \in \mathcal{V}$. Then there exists a $v_{q_1} \in T_{q_1}\mathcal{M}$ such that $\exp(v_{q_1}) = q_2$, and moreover, such that $|v_{q_1}| < \epsilon$.

Note that the curve $t \to \exp(tv_{q_1}), 0 \le t \le 1$ is contained completely in $\mathcal{W}$.

We have thus produced a neighborhood $\mathcal{V}$ with the property that there exists an $\epsilon > 0$ such that any two points $q_1$ and $q_2$ of $\mathcal{V}$ can be joined by a geodesic $\gamma$ of length $< \epsilon$. Moreover, any other *geodesic* joining $q_1$ and $q_2$ must have length $\ge \epsilon$. Why?

We can in fact refine this further: We shall prove that $\gamma$ has length $<$ *any* curve joining $q_1$ and $q_2$. Moreover, we shall show that $\mathcal{V}$ can be chosen so that $\gamma$ is completely contained in $\mathcal{V}$. Such a neighbhorhood is called a *geodesically convex neighborhood*.

## 17.2 The Gauss lemma

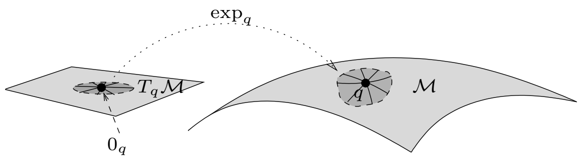
For this, we need a computation originally done in the setting of surfaces in $\mathbb{R}^3$ by Gauss.

First, note the following: By specializing our discussion from before, we have that for every point $q$, there exists a $B_{0_q}(\epsilon) \subset T_q\mathcal{M}$ such that $\exp_q$ is a diffeomorphism

$$\exp_q : B_{0_q}(\epsilon) \to \mathcal{U}$$

for some $\mathcal{U} \in \mathcal{M}$. In particular, given any coordinate system on $B_{0_q}(\epsilon)$, we obtain a coordinate system on $\mathcal{U}$ by pulling them back via $\exp_q^{-1}$.

For instance, one could choose a system of Cartesian coordinates. The associated coordinates on $\mathcal{M}$ are known as *normal coordinates*. Alternatively, one can choose polar coordinates $r, \theta^1, \theta^{n-1}$, where $\theta^\alpha$ are local coordinates on $\mathbb{S}^{n-1}$. (Note of course that these coordinates defined only on a subset of $B_{0_q}(\epsilon)$.) The associated coordinates on $\mathcal{M}$, $r \circ \exp_q^{-1}$, etc., are known as *geodesic polar* *coordinates.*




The so-called Gauss' lemma is the following proposition:

**Proposition 17.3.** *On $\mathcal{M}$,*

$$g\left(\frac{\partial}{\partial r}, \frac{\partial}{\partial r}\right) = 1, g\left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta^\alpha}\right) = 0,$$

*i.e., in geodesic polar coordinates, the metric can be written as*

$$dr^2 + g_{\alpha\beta} d\theta^\alpha \otimes d\theta^\beta.$$

*Proof.* Let us prove the first identity first. We can interpret $\frac{\partial}{\partial r}$ as a vector field either on $B_{0_q}$ and on $\mathcal{M}$. On the former, its integral curves are lines through the origin $0_q$, parametrized by arc length.[33] By the definition of the map $\exp_q$, it follows that the integral curves of $\frac{\partial}{\partial r}$, interpreted now as a vector field on $\mathcal{M}$, are geodesics parametrized by arc length. Thus, the first identity follows.

For the second, let $\alpha$ be given, let $p \in \mathcal{U}$ and fix a line through the origin $0_q$ connecting it with the pre-image of $p$ under $\exp_q$. Let us consider the vector fields $\frac{\partial}{\partial r}$ and $\frac{\partial}{\partial \theta^\alpha}$ along the image of this line in $\mathcal{M}$, which is a geodesic $\gamma$ connecting $q$ and $p$.[34] Let us denote $T = \frac{\partial}{\partial r}$, $N = \frac{\partial}{\partial \theta^\alpha}$.

We are interested in the quantity $g(N, T)$. Since this is differential geometry, let's differentiate and see if we are lucky. We compute

$$Tg(N, T) = g(\nabla_T N, T) + g(N, \nabla_T T).$$

The second term above vanishes in view of the geodesic equation, thus

$$Tg(N, T) = g(\nabla_T N, T).$$

On the other hand, since $[N, T] = 0$, we have

$$Tg(N, T) = g(\nabla_T N, T) = g(\nabla_N T, T) = \frac{1}{2} N g(T, T) = 0$$

since $g(T, T) = 1$ identically. Thus $g(N, T)$ is constant along $\gamma$. Since $N = 0$ at $q$, then $g(N, T) = 0$ identically. □

The above lemma leads immediately to the following

[33] The vector field is of course not defined at the origin. Exercies: deal with this issue.

[34] Again, address for yourself the issue of the fact that *a priori*, these may not be defined everywhere.

**Proposition 17.4.** *Let $q$, $\epsilon$, $\tilde{\mathcal{U}}$ be as in the above lemma, and let $p \in \tilde{\mathcal{U}}$. Then the radial geodesic $\gamma$ joining $q$ and $p$, of length $r(p)$, is length minimizing, i.e. if $\tilde{\gamma}$ is any other piecewise regular curve joining $q$ and $p$, then $L(\tilde{\gamma}) > L(\gamma) = r(p)$.*

The statement is in fact true where piecewise regular is replaced by any rectifiable curve. These are the most general curves for which one can define the notion of length.

*Proof.* For a piecewise regular curve, we can write the length as

$$L = \int \sqrt{g(\tilde{\gamma}', \tilde{\gamma}')} dt.$$

Let us consider separately the case where $\gamma \subset \mathcal{U}$, and when it is not. In the former case, since the curve is contained in our geodesic polar coordinate chart[35], we can write

$$\begin{aligned} L &= \int \sqrt{\left(\frac{dr}{dt}\right)^2 + g_{\alpha\beta}\frac{d\theta^\alpha}{dt}\frac{d\theta^\beta}{dt}} dt \qquad (25) \\ &\leq \int \sqrt{\left(\frac{dr}{dt}\right)^2} dt \\ &\leq \int \frac{dr}{dt} dt \\ &= r(p), \end{aligned}$$

with equality iff[36] $g_{\alpha\beta}\frac{d\theta^\alpha}{dt}\frac{d\theta^\beta}{dt} = 0$, and thus $\theta^\alpha = c$, and $\frac{dr}{dt} \neq 0$. Thus, after reparametrization, $\tilde{\gamma}$ is the radial geodesic.

In the other case, there is a first time $t_0$ when $\tilde{\gamma}$ crosses $r = r(p)$. Redo the above with $\tilde{\gamma}$ replaced by $\tilde{\gamma}|_{[0,t_0]}$. □

The above argument actually illustrates a general technique in the calculus of variations for showing that the solution of a variational problem is actually a minimizer. The technique is called: embedding in a field of variations.

## 17.3 Geodesically convex neighborhoods

We can now turn to finishing up a task left undone, namely showing the existence of geodesically convex neighborhoods in the sense described previously.

First a remark. There is something we can say about no-radial geodesics completely contained in a geodesic polar coordinate chart. If $\kappa$ is such a geodesic then $r \circ \kappa$ cannot have a strict maximum. For suppose $t_{max}$ were such a point.

[35] Again, with the usual caveat.

[36] Exercise: Why is this expression positive definite?

Note that a $t_{max}$, $\frac{d^2r}{dt^2} \leq 0$, $\frac{dr}{dt} = 0$. But,

$$\begin{aligned} \frac{d^2r}{dt^2}(t_{max}) &= -\Gamma^r_{\beta\gamma}\frac{d\theta^\beta}{dt}\frac{d\theta^\gamma}{dt} \\ &= -\frac{1}{2}g^{rr}(-g_{\beta\gamma})\frac{d\theta^\beta}{dt}\frac{d\theta^\gamma}{dt} \\ &= \frac{1}{2}g_{\beta\gamma}\frac{d\theta^\beta}{dt}\frac{d\theta^\gamma}{dt} \\ &> 0 \end{aligned}$$

Exercise: why the last strict inequality? This is a contradiction.

We can finally complete our construction of a geodesically convex neighborhood. Let $\mathcal{W}$ be as in Proposition 17.2. We may choose $\epsilon$ as in the discussion after that Proposition, and and choose a $\mathcal{V}$ as before, but with $\epsilon/2$ in place of $\epsilon$, and so that addition $\mathcal{V}$ is of the form $\exp_p(B_{0_p}(\epsilon/4)$. We have that $\mathcal{V} \subset \exp_p(B_{0_p}(\epsilon)) \subset \mathcal{W}$. Moreover, we have that any two points $q_1$, $q_2$ in $\mathcal{V}$ can be joined by a geodesic in $\mathcal{W}$ of length $< \frac{1}{2}\epsilon$.

Repeating a previous computation (namely, (25), it follows that such a geodesic necessarily must remain in $\exp_p(B_{0_p}(\epsilon))$. Thus, by our result on the absense of maxima of $r$, it follows that $r$ cannot have a maximum. Since $r(q_1) < \epsilon/4$, $r(q_2) < \epsilon_4$, it follows that $r < \epsilon/4$ throughout the geodesic. That is to say, the geodesic is contained in $\mathcal{V}$. So $\mathcal{V}$ is geodesically convex in the sense claimed.

## 17.4 Application: length minimizing curves are geodesics

In later sections, we shall give conditions on a Riemannian manifold implying the existence of length minimizing geodesics joining any two given points. The construction of these will involve global considerations.

On the other hand, given a curve which is *not* a geodesic, one can show that it can *not* be length minimizing by completely local considerations. This is in fact yet another application of what we have just done.

**Proposition 17.5.** *Let $(\mathcal{M}, g)$ be Riemannian, with $p, q \in \mathcal{M}$, and let $\gamma$ be a piecewise regular curve from $p$ to $q$. If $\gamma$ is not a geodesic, then there exists a piecewise regular curve $\tilde{\gamma}$ connecting $p$ to $q$, such that $L(\tilde{\gamma}) < L(\gamma)$.*

*Proof.* Suppose $\gamma$ is not a geodesic. Then there exists a point $\tilde{p} = \gamma(T)$ where $\gamma$ does not satisfy the geodesic equation. Consider a geodesically convex neighborhood $\mathcal{V}$ centered at $\tilde{p}$. Consider two points $\tilde{q}_1$, $\tilde{q}_2$ on $\gamma$, in $\mathcal{V}$, such that say $\gamma(t_1) = \tilde{q}_1$, $\gamma(t_2) = \tilde{q}_2$, with $t_1 < T < t_2$. We know that there exists a geodesic connecting $\tilde{q}_1$ and $\tilde{q}_2$ which is length minimizing. If this geodesic coincides with $\gamma|_{[t_1,t_2]}$, then there is nothing to show. If not, then the curve $\gamma|_{[t_1,t_2]}$ has strictly greater length than this curve, in which case we can replace $\gamma$ by a shorter curve joining $p$ and $q$. □

# 18 The metric space structure

An application of the previous is to show that a Riemannian manifold inherits the structure of a metric space related to the Riemannian metric.

**Definition 18.1.** *Let $(\mathcal{M}, g)$ be a Riemannian manifold. For $x, y \in \mathcal{M}$, let $G_{x,y}$ denote the set of all piecewise smooth curves joining $x$ and $y$, and for $\gamma \in G_{x,y}$, let $L(\gamma)$ denote the length of the curve $\gamma$¿ Define*

$$d(x, y) = \inf_{\gamma \in G_{x,y}} L(\gamma)$$

We have

**Proposition 18.1.** *The function $d$ defines a distance function on $\mathcal{M}$, i.e. $(\mathcal{M}, d)$ is a metric space.*

*Proof.* That $d(x, y) = d(y, x)$ is obvious. The triangle inequality is similarly immediate. To spell it out: If $\gamma : [a, b] \to \mathcal{M}$ is a piecewise smooth curve joining $x$ and $y$, and $\tilde{\gamma} : [\tilde{a}, \tilde{b}]$ is a piecewise curve joining $y$ and $z$, then $\hat{\gamma} : [a, b+\tilde{b}-\tilde{a}] \to \mathcal{M}$ defined by $\hat{\gamma}(t) = \gamma(t)$ for $t \in [a, b]$, $\hat{\gamma}(t) = \tilde{\gamma}(t)$ for $t \in (b, b + \tilde{b} - \tilde{a}]$ is a piecewise smooth curve joining $x$ and $x$ with $L(\hat{\gamma}) = L(\tilde{\gamma}) + L(\gamma)$. Taking infimums over $G_{x,y}$, $G_{y,z}$, one obtains the triangle inequality $d(x, z) \leq d(x, y) + d(y, z)$.

That $d(x, x) \geq 0$ is obvious. If $x \neq y$, then let $\mathcal{U}$ be a geodesically convex neighborhood of $x$ not containing $y$. In particular, $\mathcal{U}$ contains a geodesic sphere of radius $\epsilon$. Any curve $\gamma$ joining $x$ and $y$ must cross this geodesic sphere at some point $p$. Since the radial geodesic from $x$ to $p$ minimizes the length of all curves from $x$ to $p$, and the length of this curve is $\epsilon$, it follows that $L(\gamma) > \epsilon$. Thus $d(x, y) \geq \epsilon$. □

# 19 Geodesic completeness and the Hopf-Rinow theorem

Recall from Section 16 the definition of geodesic completeness. In view of the metric space structure, we now have a "competing" notion of completeness, namely, metric completeness. In this section we shall show that this notion is actually equivalent to the notion of geodesic completeness defined earlier. In the process, we shall show that in a geodesically complete Riemannian manifold, it follows that any two points can be connected by a (**not** necessarily unique!) length-minimizing geodesic. This result is essential for global arguments in Riemannian geometry. We shall get a taste of this in the final section.

**Theorem 19.1.** *Let $(\mathcal{M}, g)$ be a Riemannian manifold, let $x, y \in \mathcal{M}$, and suppose $\exp_x$ is defined on all of $T_x\mathcal{M}$. Then there exists a geodesic $\gamma : [0, L] \to \mathcal{M}$ such that*

$$L(\gamma) = d(x, y). \tag{26}$$

Clearly, the assumptions of the above theorem are satisfied if $(\mathcal{M}, g)$ is geodesically complete. Note that in view of the fact that $L(\gamma) \geq d(x, y)$, (26) is equivalent to the statement that for any other curve $\tilde{\gamma}$ joining $x$ and $y$, then $L(\tilde{\gamma}) \geq L(\gamma)$.

*Proof.* Let $\mathcal{U}_x$ be a geodesically convex neighborhood of $x$, and let $\mathcal{S}_\epsilon$ be a geodesic sphere around $x$ of radius $\epsilon > 0$, contained in $\mathcal{U}_x$, such that $\epsilon < d(x, y) = d$. Since $d(\cdot, y)$ is a continuous function on $\mathcal{S}_\epsilon$, and $\mathcal{S}_\epsilon$ is compact, it follows that there exists a $p \in \mathcal{S}_\epsilon$ such that $d(p, y) = \inf_{q \in \mathcal{S}_\epsilon} d(q, y)$. Consider now the radial geodesic $\gamma$ from $x$ through $p$, parametrized by arc length. By assumption, $\gamma(t)$ is defined for all values of $t$. We will show that $\gamma(d) = y$. This will give the result of the theorem.

Note that $\gamma(\epsilon) = p$. Consider the subset $X \subset [\epsilon, d]$ defined by

$$X = \{s \in [\epsilon, d] : d(y, \gamma(s')) = d - s', \forall d \leq s' \leq s\}.$$

Clearly, our result follows if we show $d \in X$. Moreover, $X$ is non-empty, because, we have $\epsilon \in X$. To see this note first that every curve $\kappa$ joining $x$ and $y$ must cross $\mathcal{S}_\epsilon$, points in the connected component of $\mathcal{M} \setminus \mathcal{S}_\epsilon$ containing $p$ have distance $< \epsilon$ away from $x$. Thus the length of $\kappa$ is greater than $\epsilon$ plus the length of a curve from $\mathcal{S}_\epsilon$ to $y$. Thus $d(x, y) \geq d(p, y) + \epsilon$. On the other hand the reverse inequality is immediate from the triangle inequality. So indeed $\epsilon \in X$.

Thus, since $[\epsilon, d]$ is connected, it suffices to show that $X$ is open and closed in the topology of $[\epsilon, d]$.

The closedness of $X$ is immediate in view of the continuity of the function $d(y, \cdot)$. To show that $X$ is open, it suffices to show that if $s \in X$ with $s < d$, then $s + \delta \in X$ for $\delta$ sufficiently small.

So let $s \in X$, $s < d$. There exists a geodesically convex neighborhood $\tilde{\mathcal{U}}$ of $\gamma(s)$. Let $\delta$ be such that the geodesic sphere around $\gamma(s)$ of radius $\delta$, denoted $\tilde{\mathcal{S}}_\delta$, is contained in $\tilde{\mathcal{U}}$, and so that, moreover, $\delta < s < s + \delta < d$. (This will be true for all $\delta \leq \delta_0$ for some $\delta_0 > 0$.) As before, let $\tilde{p}$ minimize the distance from $\tilde{\mathcal{S}}_\delta$ to $y$. Defining $\tilde{\gamma}$ to be the radial geodesic from $\gamma(s)$ to $\tilde{p}$, it follows as before that $d(\tilde{p}, y) = d(\gamma(s), y) + \delta$.

Consider now the distance between $\gamma(s - \delta)$ and $\tilde{p}$. On the one hand, we have, since $s - \delta \in X$, that $d(y, \gamma(s - \delta)) = d - s + \delta$. On the other hand, we know that

$$d(y, \gamma(s - \delta)) \leq d(\gamma(s - \delta), \tilde{p}) + d(\tilde{p}, y) = d(\gamma(s - \delta), \tilde{p}) + d - s - \delta.$$

Thus

$$d(\gamma(s - \delta), \tilde{p}) \geq 2\delta.$$

On the other hand, by the triangle inequality, we have

$$d(\gamma(s - \delta), \tilde{p}) \leq 2\delta.$$

Thus we have $d(\gamma(s - \delta), \tilde{p}) = 2\delta$. But since $\gamma$ followed by $\tilde{\gamma}$ is a curve joining $\gamma(s - \delta)$ and $\tilde{p}$ of precisely length $2\delta$, it follows by the properties of geodesically convex neighborhoods that this curve is a geodesic, i.e. $\tilde{\gamma}$ must coincide with $\gamma$. But now the claim follows, since $\tilde{p} = \gamma(s + \delta)$. $\square$

**Theorem 19.2.** *Let $(\mathcal{M}, g)$ be a Riemannian manifold, and suppose there exists a point $x$ satisfying the assumptions of the previous Theorem. Then $(\mathcal{M}, d)$ is complete as a metric space.*

In particular, the theorem applies in the case $(\mathcal{M}, g)$ is assumed geodesically complete.

*Proof.* A metric space is complete if every Cauchy sequence converges. Compact spaces are complete, and Cauchy sequences are clearly bounded. Thus it suffices to show that all bounded subsets of $\mathcal{S}$ are contained in compact sets.

Let $\mathcal{B} \subset \mathcal{M}$ be bounded. This means that there exists an $\infty > M > 0$ such that $d(x, y) \leq M$ for all $y \in \mathcal{M}$. By Theorem 19.1, this implies that $\mathcal{B} \subset \exp_x(B_M(0))$ wheere $B_M(0)$ denotes the closed ball of radius $M$ in the $T_x\mathcal{M}$. But $\exp_x : T_x\mathcal{M} \to \mathcal{M}$ is a continuous function and $B_M(0)$ is compact. Thus $\exp_x(B_M(0))$ is compact and our theorem follows.

□

**Theorem 19.3.** *Let $(\mathcal{M}, g)$ be a Riemannian manifold, and suppose $(\mathcal{M}, d)$ is metrically complete. Then $(\mathcal{M}, g)$ is geodesically complete.*

*Proof.* Let $\gamma : [0, T) \to \mathcal{M}$ be a geodesic parametrized by arc length, with $T < \infty$. We shall show that $\gamma$ can be extended to a geodesic $\gamma : [0, T+\delta) \to \mathcal{M}$.

Let $t_i \to T$. Since $d(\gamma(t_i), \gamma(t_j)) \leq |t_i - t_j|$, it follows that $\gamma(t_i)$ is Cauchy. By assumption then, $\gamma(t_i)$ converges to a point $y$. Let $\mathcal{U}$ denote a geodesically convex neighborhood of $y$, and let $\epsilon$ be such that for every $x_1, x_2 \in \mathcal{U}$, there exists a unique geodesic $\tilde{\gamma}$ of length $\epsilon$ in $\mathcal{M}$ starting from $x_1$, with the property that $x_2$ is on this geodesic within length $\epsilon$. Choose $\gamma(t_i), \gamma(t_{i+1}) \in \mathcal{U}$, with $T - t_i < \epsilon$, and consider the geodesic $\tilde{\gamma}$ joining $\gamma(t_i)$ and $(\gamma(t_{i+1})$ described above. Since $\gamma$ is also a geodesic joining $\gamma(t_i)$ and $\gamma(t_{i+1})$ of length less than $\epsilon$, it follows that $\tilde{\gamma}(t - t_i)$ must coincide with $\gamma(t)$ for $t \in [t_i, T)$. Define now $\gamma(t)$ for $t \in [T, t_i + \epsilon)$ by $\gamma(t) = \tilde{\gamma}(t - t_i)$. This is then the required extension. □

# 20 The second variation

The existence of length minimizing geodesics rests on a *global* argument. On the other hand, a length minimizing geodesic certainly must be length minimizing to *local* deformations. The condition that a geodesic be length minimizing to local deformations is characterized by a differential condition, analogous to the characterization $f'' \leq 0$ for local minima of a function of one variable. It is in this characterization that the so-called curvature tensor first makes its appearance.

Let $\tilde{\gamma} : [0, L] \times (-\epsilon, \epsilon) \to \mathcal{M}$ denote a variation of a smooth curve $\gamma : [0, L] \to \mathcal{M}$, i.e., let $\tilde{\gamma}$ be a smooth map such that $\tilde{\gamma}(\cdot, 0) = \gamma$, and such that $\tilde{\gamma}(\cdot, s)$ is a curve for all $s \in (-\epsilon, \epsilon)$. We have already computed a formula for the first variation in Section 14. We now go further and compute a formula for the second variation.

We shall compute this formula in the special case that the original $\gamma$ is a geodesic parametrized by arc length (so then $L$ is its length), and such that

$$\tilde{\gamma}(0,s) = \gamma(0,s), \tilde{\gamma}(L,s) = \tilde{\gamma}(L,s). \tag{27}$$

It is a good exercise for the reader to write down the general case!

Let us recall the notation $N = \tilde{\gamma}_* \frac{\partial}{\partial s}$, $T = \tilde{\gamma}_* \frac{\partial}{\partial t}$, where $s$ is a coordinate in $(-\epsilon, \epsilon)$ and $t$ is a coordinate in $[0, L]$. Also recall the notation $L(s)$ got the length of the curve $\tilde{\gamma}(\cdot, s)$.

We have from before that for all $s$.

$$L'(s) = \int_0^L -g(T,T)^{-3/2} g(\nabla_T T, T) g(N,T) - (g(T,T))^{-1/2} g(N, \nabla_T T) dt$$

Differentiating in $s$, and evaluating at 0, we obtain (using that $g(T_{\tilde{\gamma}(t,0)}, T_{\tilde{\gamma}(t,0)}) = 1$, $\nabla_T T|_{\tilde{\gamma}(t,0)} = 0$)

$$\begin{aligned}
L''(0) &= -\frac{d}{ds} \int_0^L -g(T,T)^{-3/2} g(\nabla_T T, T) g(N,T) - (g(T,T))^{-1/2} g(N, \nabla_T T) dt|_{s=0} \\
&= -\int_0^L N \left( -g(T,T)^{-3/2} g(\nabla_T T, T) g(N,T) - (g(T,T))^{-1/2} g(N, \nabla_T T) \right) dt|_{s=0} \\
&= -2 \int_0^L g((\nabla_N \nabla_T T)_{\gamma(t)}, N_{\gamma(t)}) dt
\end{aligned}$$

At this point, let us momentarily specialize to the case of $\mathbb{R}^n$ with its Euclidean metric. **In this case, covariant derivatives commute.** (Why?) I.e., we may write

$$\begin{aligned}
L''(0) &= -2 \int_0^L g(\nabla_T \nabla_N T, N) dt \\
&= -2 \int_0^L T g(\nabla_N T, N) + g(\nabla_N T, \nabla_T N) \\
&= -2 g(\nabla_N T, N)]_0^T + 2 \int_0^L g(\nabla_T N, \nabla_T N) dt \\
&= 2 \int_0^L g(\nabla_T N, \nabla_T N) dt
\end{aligned}$$

where we have used here both the boundary condition (27) implying $N(\gamma(0)) = N(\gamma(L)) = 0$, and the relation $[N,T] = 0$. (Exercise: Why does the latter hold?) In particular, we see that in this case $L''(0) \geq 0$, and thus geodesics minimize arc length with respect to near-by variations.

Of course, in the case of Euclidean space, we already knew much more, namely that geodesics globally minimize arc length, i.e. that a geodesic from $p$ to $q$ has the property that its length is strictly less than the length of any other curve from $p$ to $q$. (In particular, geodesics are unique.) This follows from what we have done already, in view of the existence (for $\mathbb{R}^n$) of **global** geodesic normal coordinates, and Gauss' lemma.

But no matter. It is merely this computation for $L''$ that we wish to generalize to Riemannian manifolds. Now, however, covariant derivatives no-longer commute. The analogue of the previous is given as below:

$$\begin{aligned} L''(0) &= \int_0^L g(\nabla_T\nabla_N T, N) + g(\nabla_N\nabla_T T - \nabla_T\nabla_N T, N)dt \\ &= \int_0^L g(\nabla_T N, \nabla_T N) + g(R(T,N)T, N)dt \qquad (28) \end{aligned}$$

where

$$R(T,N)T = \nabla_N\nabla_T T - \nabla_T\nabla_N T.$$

The expression $R(T,N)T$ is called the *curvature*. A priori it seems that its value at a point $p$ should depend on the behaviour of the vector fields $T$ and $N$ up to second order. (If this were the case, then the formula (28) would not be particularly useful.) But it fact, $R(T,N)T$ depends only on $T(p)$, $N(p)$, i.e. $R$ is a *a tensor*. (See the next section!) As we shall see, $R$ is an invariant of the metric $g$; it is the above expression for $L''$ which gives it perhaps its most natural geometric interpretation in terms of the lengths of nearby curves.

We commence in the following section a systematic discussion of the curvature tensor. We will then return to the second variation formula, and use it to prove theorems relating local assumptions on the behaviour of the curvature tensor, and global geometric properties of the manifold.

# 21 The curvature tensor

Having got a taste of curvature we now make the general definitions.

**Definition 21.1.** *Let $(\mathcal{M}, g)$ be a Riemannian manifold, $p \in \mathcal{M}$, and $X$, $Y$, and $Z$ be vetor fields in a neighborhood of $p$. We define*

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

**Proposition 21.1.** *$R(X,Y)Z$ only depends on $X(p)$, $Y(p)$, and $Z(p)$, i.e. $R$ can be thought of as a section of $T\mathcal{M} \otimes T^*\mathcal{M} \otimes T^*\mathcal{M} \otimes T^*\mathcal{M}$.*

*Proof.* As usual, an uninspired calculation. □

We call $R$ the Riemann curvature tensor. The curvature tensor satisfies the following symmetries

**Proposition 21.2.**

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

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

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

$$R(X,Y)Z + R(Z,X)Y + R(Y,Z)X = 0$$

*Proof.* Again, this is just a computation. The last identiy is known as the first *Bianchi identity.* $\square$

## 21.1 Ricci and scalar curvature

The algebraic complexity of curvature, and the more subtle relations of this complexity and geometry, is something which we will not in fact really come to terms with in this class. Here, we shall consider algebraically more simple objects that can be obtained from the Riemann curvature.

We can define the so called *Ricci curvature* by applying the contraction $C$ of Definition 6.2. Specifically, considering $R$ as a section of $T\mathcal{M}\otimes T^*\mathcal{M}\otimes T^*\mathcal{M}\otimes T\mathcal{M}$, we may apply the map

$$C : T\mathcal{M}\otimes T^*\mathcal{M}\otimes T^*\mathcal{M}\otimes T^*\mathcal{M} \to T^*\mathcal{M}\otimes T^*\mathcal{M}$$

acting as in Definition 6.2 but only in $T\mathcal{M}$ and the second $T^*\mathcal{M}$ factor, i.e the one "corresponding" to $Y$ in the notation $R(X,Y)Z$. The composition of $C$ and $R$ produces a tensor Ric which is a section of $T^*\mathcal{M}\otimes T^*\mathcal{M}$. This is the so-called Ricci curvature. Alternatively, in view of the relation between contraction and trace of a homomorphism, one easily sees that

$$\mathrm{Ric}(X,Z) = \mathrm{trace}Y \mapsto R(X,Y)Z.$$

From the symmetries of curvature tensor one sees that this is the only non-trivial contraction. Moreover, one sees that the Ricci tensor is symmetric, i.e. that $\mathrm{Ric}(X,Y) = \mathrm{Ric}(Y,X)$.

It is hard to give a sense at this stage for why the Ricci curvature is so important an object in geometry. Written in local coordinates (see next section) it has the same number of components as the metric, and this very pedestrian reason may be why it was first written down. But its true geometric character is much more subtle. In this class, we will only scratch the surface, but we shall see it used in Theorem 24.1 of Section 24.

In general relativity, the importance of Ricci curvature is clear from the beginning. The vacuum Einstein equations are the statement that

$$\mathrm{Ric} = 0$$

A further contraction of Ric (after raising one of the indices!) yields the so-called *scalar curvature*. This is often denoted by $R$, in the spirit of remarks given in Section 13.1. For a formula in terms of indices, see Section 21.3.

## 21.2 Sectional curvature

The reader familiar with the classical differential geometry of surfaces will have some intuition for the Gauss curvature of a surface, in particular, what it means geometrically for the Gauss curvature to be positive, negative, zero.

It turns out that the curvature tensor defines, for each 2-plane, a notion of curvature that can be thought to correspond to the Gauss curvature. This is called the sectional curvature. It is defined as follows. Given a plane

$$\Pi \subset T_p\mathcal{M}$$

choose vectors $X$, $Y$ spanning $T_p$. Define

$$K(\Pi) = \frac{g(R(X,Y)X,Y)}{g(X,X)g(Y,Y) - g(X,Y)^2}.$$

Note that if $X$ and $Y$ are orthonormal, then

$$K(\Pi) = g(R(X,Y)X,Y)$$

Note that if $\mathcal{M}$ is two dimensional, then $K$ coincides with the usual Gaussian curvature.

It turns out that the condition $K(\Pi) > 0$ for *all* $\Pi$ is most like the condition of positive Gaussian curvature for surfaces. Similarly, $K(\Pi) < 0$. The subtlety of higher dimensional geometry lies in that $K(\Pi)$ can have different signs for different planes. The condition that Ric is positive (resp. negative) definite can be thought of as a weaker analogue of $K(\Pi) > 0$ for all $\Pi$, (resp. $K(\Pi) < 0$ for all $\Pi$). We shall get at least some taste of all this in Section 24.

Finally, as an exercise, test your multilinear algebra skills by showing that the collection of all sectional curvatures determines the Riemann curvature tensor.

## 21.3 Curvature in local coordinates

Since as we saw in the previous section, $R$ can be though of as a section of $T\mathcal{M} \otimes T^*\mathcal{M} \otimes T^*\mathcal{M} \otimes T^*\mathcal{M}$, we can write it as

$$R = R^i_{jkl}\frac{\partial}{\partial x^i} \otimes dx^j \otimes dx^k \otimes dx^l$$

where

$$R^i_{jkl}\frac{\partial}{\partial x^i} = R\left(\frac{\partial}{\partial x^k}, \frac{\partial}{\partial x^l}\right)\frac{\partial}{\partial x^j}$$

Note: There is never universal agreement in the literature as to the order of the indices, or even the sign convention in the definition of the Riemann curvature tensor. Sectional and Ricci curvature are always defined so as to be positive on, say $\mathbb{S}^n$.

One can derive from the definition of curvature the formulas

$$R^i_{jkl} = \frac{\partial\Gamma^i_{jk}}{\partial x^l} - \frac{\partial\Gamma^i_{jl}}{\partial x^k} + \Gamma^m_{jk}\Gamma^i_{ml} - \Gamma^m_{jl}\Gamma^i_{mk}$$

$$\mathrm{Ric}_{ij} = R^k_{ijk}$$

$$R = g^{ij}R_{ij}.$$

Thus, if one has explicit formulas for the metric, one can compute the curvature tensor by brute force.

## 21.4 Curvature as a local isometry invariant

In view of Proposition 13.2, we have immediately that

**Proposition 21.3.** *Let $(\mathcal{M}, g)$, $(\tilde{\mathcal{M}}, \tilde{g})$ be Riemannian and suppose that $p \in \mathcal{U} \subset \mathcal{M}$, $q \in \tilde{\mathcal{U}} \subset \tilde{\mathcal{M}}$, and $\phi : \mathcal{U} \to \tilde{\mathcal{V}}$ is an isometry with $\phi(p) = q$. Let $R$, $\tilde{R}$ denote the Riemann curvature tensors of $\mathcal{M}$, $\tilde{\mathcal{M}}$ respectively. Then*

$$\tilde{R}(\phi_* X, \phi_* Y)\phi_* Z = \phi_*(R(X, Y)Z).$$

This is the statement that curvature is a local isometry invariant.

An alternative (and quite cumbersome) way of proving Proposition 21.3 is via the interpretation of curvature in terms of the second variation of arc length. Try it for yourself if you're of that persuasion.

## 21.5 Spaces of constant curvature

### 21.5.1 $\mathbb{R}^n$

It should be clear from the definition that in this case we have

$$R^l_{ijk} = 0.$$

In fact, the following statement is true. Let $\mathcal{U}$ be homoeomorphic to the $n$-disc, and let $g$ be a metric on $\mathcal{U}$ such that $R^l_{ijk}$. Then there exists a smooth map $\phi : \mathcal{U} \to \mathbb{R}^n$ which is an isometry onto its image. In other words, if $(\mathcal{M}, g)$ has identically vanishing Riemann curvature, then $\mathcal{M}$ is locally isometric to Euclidean space.

The above statement was in fact proven by Riemann in his original lecture which initiated Riemannian geometry. In view also of Propostion 21.3, we see that $R^l_{ijk} = 0$ are necessary and sufficient conditions for a Riemannian manifold to be locally isometric to Euclidean space. This was the original motivation for identifying this expression.

### 21.5.2 $\mathbb{S}^n$

Write down the metric of $\mathbb{S}^n$ explicitly, and compute the curvature. In particular, show that

$$K(\Pi) = 1$$

for all planes $\Pi \subset T\mathbb{S}^n$. We say that $\mathbb{S}^n$ is a space of constant curvature.

Exercise: How can you cheat in this computation?

Note that, in particular, the curvature tensor is not identically 0. In view of Proposition 21.3, this gives a purely infinitessimal proof that the sphere and Euclidean space are not locally isometric. (Compare with Example 9.3, where the "macroscopic" formula for the area of spherical triangles is used. Actually, this formula can be derived by integrating the curvature in the spherical triangle. This is the celebrated Gauss-Bonnet formula.)

### 21.5.3 $\mathbb{H}^n$

This space, called hyperbolic $n$-space, is the higher dimensional analogue of the space $\mathbb{H}^2$ discussed in the introduction.

A useful way of representing this space is as the standard open disc in $\mathbb{R}^n$ with metric
$$g = \frac{4}{(1-(x^1)^2 - \cdots - (x^n)^2)^2} e$$
where $e$ denotes the standard Euclidean metric.

Compute for yourself the curvature. In particular, show
$$K(\Pi) = -1$$
for all $\Pi$.

It turns out that Riemannian manifolds of constant curvature are locally isometric to $\mathbb{H}^n$, $\mathbb{S}^n$ or $\mathbb{R}^n$. This is not so difficult to show, but we will not show it here...

There is no end to the study of $\mathbb{S}^n$, $\mathbb{H}^n$, and their various quotients. A good geometer should be able to navigate these with ease.

# 22 Differential forms and Cartan's method

# 23 Jacobi fields

We commenced in Sections 14 and 20 the study of the set of all piecewise curves from points $p$ to $q$, via the techniques of the variation calculus. A major difficulty arises from the fact that this space is infinite dimensional. It turns out, however, that the behaviour of critical points of the length functional $L$, i.e. geodesics, can be completely analyzed by passing to a finite dimensional subset, so-called geodesic variations. The vector fields $N$ defined by such variations are known as *Jacobi fields*.

**Definition 23.1.** *Let $\gamma : [0, L] \to \mathcal{M}$ be a geodesic. A* Jacobi field *along $\gamma$ is a vector field $J$ satisfying*
$$\nabla_T \nabla_T J + R(T, J)T = 0. \tag{29}$$

Equation (29) is known as the *Jacobi equation.* For reasons that shall soon become clear, we could restrict to fields $J$ such that $g(T, J) = 0$, for fields in the direction of $T$ are in some sense trivial.

**Proposition 23.1.** *Let $\gamma : [0, L] \to \mathcal{M}$ be a geodesic parametrized by arc length, and let $T = \gamma_* \frac{\partial}{\partial t}$. Then $J$ is a Jacobi field if and only if there exists a geodesic variation $\tilde{\gamma}$ of $\gamma$ such that $J = \tilde{\gamma}_* \frac{\partial}{\partial s}|_{s=0}$.*

*Proof.* Suppose $\tilde{\gamma}$ is a geodesic variation, i.e. suppose that $\nabla_T T = 0$ throughout the surface spanned by the variation. Let us denote by $J$ the vector field $\tilde{\gamma}_* \frac{\partial}{\partial s}$. Applying $J$ to $\nabla_T T = 0$, we obtain

$$\nabla_J(\nabla_T T) = 0.$$

But by the definition of curvature and the fact that $[J, T] = 0$, we have

$$\begin{aligned} 0 = \nabla_J \nabla_T T &= \nabla_T \nabla_J T + R(T, J)T \\ &= \nabla_T \nabla_T J + R(T, J)T. \end{aligned}$$

This gives the desired result.

For the other direction take just take $\tilde{\gamma}(s,t) = \exp_{\exp_{\gamma(0)} J(0)s}(T + J's)$. □

In view of our local existence theorem for ode's, and the linearity of Jacobi equation, we easily obtain the following

**Proposition 23.2.** *Let $\gamma$ be a geodesic as before, with $p \in \gamma$ and let $J_0, J_0'$ be vectors at $p = \gamma(t_0)$. Then there exists a Jacobi field along $\gamma$ such that $J(t_0) = J_0$, $\nabla_{\frac{\partial}{\partial t}} J = J_0'$.*

*Proof.* Compare with the Proposition 15.1. □

**Corollary 23.1.** *The space of Jacobi fields along $\gamma$ constitutes an $2n$-dimensional linear subspace of the (infinite dimensional) set of all piecewise regular vector field along $\gamma$.*

## 23.1 The index form $I(V, W)$

The second variation formula motivates the following definition. Given piecewise regular vector fields $V$, $W$ along a geodesic $\gamma$, with $T = \gamma'$, define

$$I(V, W) = \int (g(\nabla_T V, \nabla_T W) - g(R(T, V)T, W))dt. \tag{30}$$

This is a symmetric bilinear form on the space of piecewise regular vector fields. In view of our computation, if $V$ is a variation vector field for a variation $\tilde{\gamma}$ of $\gamma$, with fixed endpoints, then

$$L''(0) = I(V, V).$$

We will be interested in knowing when is $L''(0) \geq 0$ for all variations $\tilde{\gamma}$ of a given curve $\gamma$. In view of the above, we have reduced this to the question, when is $I(V, V) \geq 0$ for all piecewise regular vector fields along $\gamma$, vanishing at the endpoints.

This is progress, for the space of piecewise regular vector fields vanishing at the endpoints is a linear space, and the mapping $I$ is a bilinear form. This is a lot better than understanding the nonlinear mapping $L$ on a huge space of curves.

But in fact, we can do better, and this is the whole point of considering Jacobi fields.

For it turns out that the question of whether $I(V,V) \geq 0$ for all $V$ can be completely resolved by restricting to the finite dimensional subspace of Jacobi fields.

## 23.2 Conjugate points and the index form

To state the relation between the index form and Jacobi fields, we will make the following definition.

**Definition 23.2.** *Let $\gamma : [0,L] \to \mathcal{M}$ be a geodesic parametrized by arc length. Points $p = \gamma(t_0)$, $q = \gamma(t_1)$, $t_0 < t_1$ are said to be conjugate along $\gamma$ if there exists a nontrivial*[37] *Jacobi field $J$ along $\gamma$ such that $J(t_0) = 0$, $J(t_1) = 0$.*

The main result of this section is contained in:

**Proposition 23.3.** *Let $\gamma : [0,L] \to \mathcal{M}$ be a geodesic parametrized by arc length, with $p = \gamma(0)$. Suppose $V$ is a piecewise regular vector field along $\gamma$ vanishing at the endpoints, such that $I(V,V) \leq 0$. Then there exists a point $q = \gamma(t_1)$ for some $t_1 \in (0,L]$ , conjugate to $p$ along $\gamma$. Moreover, if $I(V,V) < 0$, then there exists such a point for $t_1 \in (0,L)$. Conversely, given a conjugate point $q = t_1 \in (0,L)$, there exits a $V$ vanishing at the endpoints such that $I(V,V) < 0$.*

*Proof.* The idea of the proof is to write the vector field $V$ in terms of a frame of Jacobi fields vanishing at $p$, together with $T$, and then apply the definition of the index form and just compute!

Fix an arbitrary orthonormal frame $T$, $E_1$,$\ldots E_{n-1}$ at $p$. For $E_1, \ldots E_{n-1}$, we can define Jacobi field $J_1, \ldots J_n$ such that $J'_\alpha(0) = E_\alpha$.

Note that these Jacobi fields are linearly independent in the space of vector fields. If their span (together with $T$) is not $n$-dimensional at some point $q$, then there exists a linear combination of the vector fields that vanishes at $q$. This linear combination yields our desired Jacobi field.

Otherwise, we may assume that the span of $T$, $J_1, \ldots J_n$ is $n$-dimensional for $t > 0$. Given our arbitrary $V$ we may write $V$ as

$$V = V^1 T + V^2 J_2 + \ldots + V^n J_n.$$

[37] i.e. not identically vanishing!

Applying (30) we obtain

$$\begin{aligned}
I(V,V) &= \int (g(\nabla_T V, \nabla_T V) - g(R(T,V)T,V))dt \\
&= -\int g(V, \nabla_T \nabla_T V) + g(R(T,V)T,V)dt \\
&= -\int V^\alpha V^\beta g(J_\alpha, \nabla_T \nabla_T J_\beta) + 2(V^\alpha)(V^\beta)' g(J_\alpha, \nabla_T J_\beta) \\
&\quad + (V^\alpha)(V^\beta)'' g(J_\alpha, J_\beta) + V^\alpha V^\beta g(R(T,J_\alpha)T, J_\beta)dt \\
&= -\int 2(V^\alpha)(V^\beta)' g(J_\alpha, \nabla_T J_\beta) + (V^\alpha)(V^\beta)'' g(J_\alpha, J_\beta)dt \\
&= \int (V^\alpha)'(V^\beta)' g(J_\alpha, J_\beta)dt \geq 0.
\end{aligned}$$

Actually, in this computation, we have cheated, or better, we haven't justified all the steps. For we have several times applied integration by parts and thrown away the boundary terms. For this, we need that the $V^\alpha$ are regular at 0. Recall that the $J_\alpha$ and $V$ both vanish at 0. Since $J_\alpha$ is piecewise regular, it must vanish at least linearly in $t$. On the other hand, one easily sees by the Jacobi equation that $J_\alpha$ vanishes at most linearly. Thus our integration by parts is justified.[38]

The full strength of the proposition now follows from carefully considering the various cases of equality in the above. This is again left as an exercise for the reader. □

# 24 Simple comparison theorems

The results of the previous section open the door to the study of the local length minimizing properties of geodesics by means of the theory of ordinary differential equations applied to the Jacobi equation. And this equation in turn depends on curvature. A classical corpus of results in the theory of ode's allows one to compare solutions of

$$f'' + Kf = 0$$

and

$$f'' + \bar{K}f = 0$$

when, say $K \geq \bar{K}$. Applied to the Jacobi equation, this theory allows us to infer the vanishing of Jacobi fields on a manifold $\mathcal{M}$ whose curvature satisfies a certain inequality, by "comparing" the solutions of the Jacobi equation, with the solutions on a constant-curvature manifold, which are known explicitly. For this reason, such results are known as *comparison theorems*.

This being said, the results of this section do not in fact depend on the characterization of the index form given by Propostion 23.3. They are in a sense more elementary. For examples where Jacobi fields are central, the reader can consult the literature.

[38] The reader is encouraged to do this carefully.

**Theorem 24.1.** *(Bonnet-Myers) Let $(\mathcal{M}, g)$ be a complete Riemannian manifold with*

$$Ric(X, X) \geq (n-1)kg(X, X)$$

*for all $X \in T\mathcal{M}$, for some $k > 0$. Then the diameter of $\mathcal{M}$ is less than or equal to $\pi/\sqrt{k}$ . In particular, $\mathcal{M}$ is compact.*

This is a protypical result of its kind in Riemannian geometry. A weak global assumption (completeness), together with a curvature assumption, give a strong geometric statement (diameter bounded) and more information on the topology (compactness).[39]

Note that the diameter is defined as $\sup_{x,y \in \mathcal{M}} d(x, y)$.

*Proof.* Let $x, y \in \mathcal{M}$. By Theorem 19.1, there exists a geodesic $\gamma : [0, d] \to \mathcal{M}$, parametrized by arc length, with $\gamma(0) = x$, $\gamma(d) = y$, and $L(\gamma) = d = d(x, y)$. It suffices to show that $d \leq \pi/\sqrt{k}$.

Since $\gamma$ is length minimizing, then if $\tilde{\gamma}$ is any variation with fixed endpoints, we must have, for the corresponding function $L(s)$, $L''(s) \geq 0$. This means that we must have $L(V, V) \geq 0$ for all piecewise smooth vector fields vanishing at the endpoints.

By Proposition 23.3, it clearly suffices then to show the following. For any curve $\gamma$ of length $> \pi/\sqrt{k}$, there exists a $V$ vanishing at the endpoints with $I(V, V) < 0$.

To find our $V$, we will make use of an important technique in differential geometry, namely, the choice of a well adapted frame, i.e. a well adapted set of orthogonal vectors spanning the tangent at each point along $\gamma$. Actually, we have already seen this technique in action, namely in the proof of Propostion 23.3. There, the frame in question consisted of a basis of Jacobi fields.

In our case, a convenient such frame can be constructed as follows. Let $E_1, \ldots E_{n-1}$, denote vectors at $\gamma(0)$ such that the collection $T, E_1, \ldots E_{n-1}$ comprises a set of orthonormal vectors. We may now parallel transport these vectors along $\gamma$ to obtain a set of vector fields $T, E_1, \ldots E_{n-1}$. These vector fields remain orthonormal and linearly independent. (Why?)

We will look for a $V$ in the direction of one of the $E_i$. Note that for

$$V_i = f E_i$$

[39] After studying the proof, explore the situation when the completeness assumption is dropped.

we can compute:

$$\begin{aligned}
I(V_i, V_i) &= \int_0^d g(\nabla_T V_i, \nabla_T V_i) - g(R(T, V_i)T, V_i))dt \\
&= \int_0^d g(\nabla_T (fE_i), \nabla_T (fE_i)) - g(R(T, fE_i)T, fE_j)dt \\
&= \int_0^d (f')^2 g(E_i, E_i) + f^2 g(\nabla_T E_i, \nabla_T E_i) \\
&\qquad +2ff' g(\nabla_T (E_i), E_i) - f^2 g(R(T, E_i)T, E_i)dt \\
&= \int_0^d (f')^2 - f^2 g(R(T, E_i)T, E_i)).
\end{aligned}$$

The reader should remark why the choice of $E_i$ is so convenient.

Since we have no assumptions about $g(R(T, E_i)T, E_i)$ itself, we cannot ensure directly that any particular $I(V_i, V_i) < 0$ from the above. But on summation over $i$, $g(R(T, E_i)T, E_i)$ yields $Ric(T, T)$, and for this, by assumption, we have an inequality. This is what we use to complete the proof.

Explicitly: Let $\tilde{k} < k$ be such that $d\tilde{k} = \pi$.

$$V^i(t) = \sin(t/\sqrt{\tilde{k}}).$$

Note that this vanishes at 0 and $d$.

Summing the above identity in $i$, we obtain

$$\begin{aligned}
\sum_{i=1}^{n-1} I(V_i, V_i) &= \int_0^d (n-1)k\cos^2(t\sqrt{k}) - \sin^2(t\sqrt{k}) \sum_{i=1}^{n-1} g(R(T, E_i)T, E_i)dt \\
&= \int_0^{\pi/\tilde{k}} (n-1)k\cos^2(t\sqrt{k}) - \sin^2(t\sqrt{k})Ric(T, T) \\
&< \int_0^{\pi/\tilde{k}} (n-1)k\cos^2(t\sqrt{k}) - \sin^2(t\sqrt{k})(n-1)k \\
&= (n-1)k \int_0^{\pi/\tilde{k}} \cos^2(t\sqrt{k}) - \sin^2(t\sqrt{k}) = 0
\end{aligned}$$

This clearly implies that $I(V_i, V_i) < 0$ for *some* $i$, and thus, by our previous remarks, this completes the proof of the Theorem. □

# References


[1] I. Chavel *Riemannian Geometry: A modern introduction* Cambridge University Press, 1993

[2] S. Kobayashi and K. Nomizu *Foundations of Differential Geometry: Volume 1* Wiley, 1963

[3] M. Berger *A panoramic view of Riemannian Geometry* Springer, 2003

[4] S. W. Hawking and G. F. R. Ellis *The large scale structure of space-time* Cambridge University Press, 1973