↓Lecture 19 [30.04.24]

10. Mathematical Tools II: Curvature

Here we continue our discussion of differential geometry in Chapter 3. We study two structures on a differentiable manifold that are particularly important for GENERAL RELATIVITY: → *connections* and the ← *Riemannian metric* (the latter we already know). Since most of our results are not specific to GENERAL RELATIVITY, we mostly consider general D dimensional manifolds, and only specialize to the case of $D = 3 + 1$ spacetime dimensions later.

- The mathematical framework of GENERAL RELATIVITY is ↑ *Riemannian geometry*, i.e., the field of differential geometry that studies differentiable manifolds equipped with a Riemannian (pseudo-)metric. The field was kickstared in 1854 by German mathematician BERNHARD RIEMANN with his inaugural lecture in Göttingen titled *"Über die Hypothesen, welche der Geometrie zu Grunde liegen"* [129]. In the audience was CARL FRIEDRICH GAUß, who had also picked the topic for Riemann's habilitation (Gauß died one year later).

  Fun fact: In his 1854 lecture, Riemann speculated that the material bodies might determine the metric of space; many years before Einstein worked out GENERAL RELATIVITY (see Part III, Paragraph 3 in Ref. [129], highlights are mine):

  > *Die Frage über die Gültigkeit der Voraussetzungen der Geometrie im Unendlichkleinen hängt zusammen mit der Frage nach dem innern Grunde der Maßverhältnisse des Raumes. Bei dieser Frage, welche wohl noch zur Lehre vom Raume gerechnet werden darf, kommt [..] zur Anwendung, daß bei einer diskreten Mannigfaltigkeit das Prinzip der Maßverhaltnisse schon in dem Begriffe dieser Mannigfaltigkeit enthalten ist, bei einer stetigen aber anders woher hinzukommen muß. Es muß also entweder das dem Raume zugrunde liegende Wirkliche eine diskrete Mannigfaltigkeit bilden,* oder der Grund der Maßverhaltnisse außerhalb, in darauf wirkenden bindenden Kräften gesucht werden.

  He continues …

  > *Die Entscheidung dieser Fragen kann nur gefunden werden, indem man von der bisherigen durch die Erfahrung bewährten Auffassung der Erscheinungen, wozu Newton den Grund gelegt, ausgeht und diese durch Tatsachen, die sich aus ihr nicht erklären lassen, getrieben allmählich umarbeitet; [..].*

  …and closes:

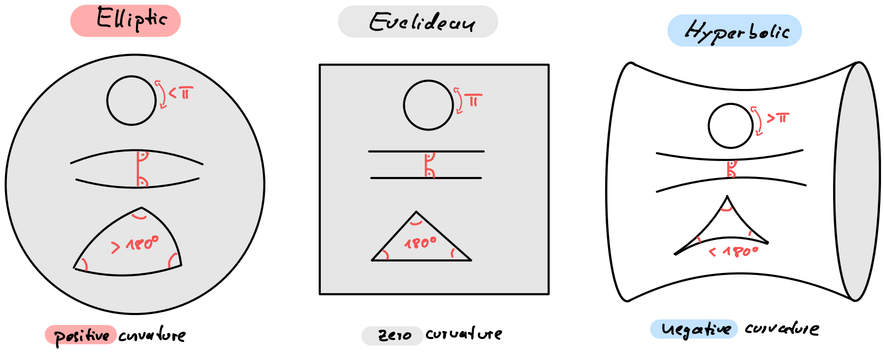
  > *Es führt dies hinüber in das Gebiet einer andern Wissenschaft, in das Gebiet der Physik, welches wohl die Natur der heutigen Veranlassung nicht zu betreten erlaubt.*

  Not only did he sketch the route Einstein would take half a century later, he even seemed intrigued exploring it himself.

- The mathematical field of geometry was conceived in ancient times as a formalization of observable facts about physical space and culminated in the axiomatization of ↓ *Euclidean geometry*. One of the facts/axioms of Euclidean geometry is the ⁂ *parallel postulate*:

  > *If a line segment intersects two straight lines forming two interior angles on the same side that are less than two right angles, then the two lines, if extended indefinitely, meet on that side on which the angles sum to less than two right angles.*

  For two millenia (!) it was suspected that this axiom can be derived from the other four axioms of Euclidean geometry (so that it doesn't deserve the title "axiom" after all). Finally, GAUß (and contemporaries) recognized that the parallel postulate cannot be proven from the other four; it is an independent axiom that can be modified to define consistent geometries that differ from Euclid's! The result is ⁂ *non-Euclidean geometry* which comes in two flavours, ↑ *elliptic geometry* and ↑ *hyperbolic geometry*:



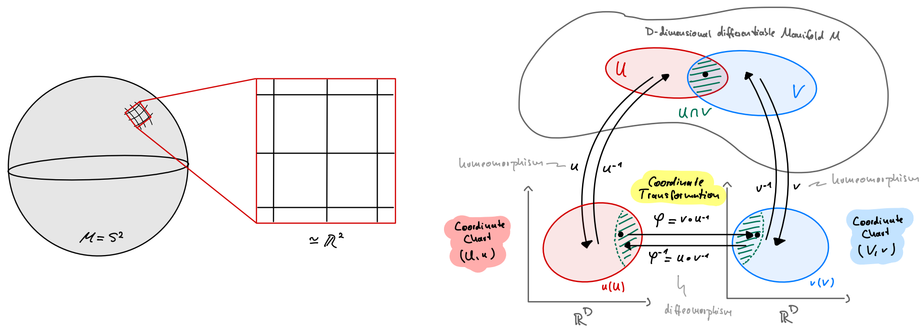
The realization by mathematicians that there are many consistent geometries opened a new question for physics: *Are we sure that the geometry of space really is Euclidean?* The answer of GENERAL RELATIVITY is: No, on large scales space is only *approximately* Euclidean, and it can be very non-Euclidean in regimes of strong gravitational fields.

10.1. Summary: What we know and what comes next

1 | Concepts we already know:

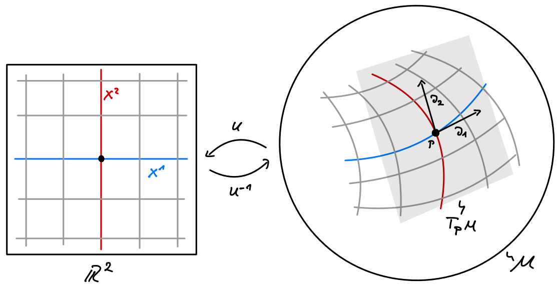
- ← *Differentiable manifolds* (Section 3.1):

  A D-dimensional manifold is locally homeomorphic (continuously isomorphic) to $\mathbb{R}^D$ (it locally "looks like" Euclidean space). A continuous, invertible function that maps a region of the manifold to a subset of $\mathbb{R}^D$ is called a *(coordinate) chart*. A collection of overlapping charts that covers the whole manifold is called an *atlas*. If the transition functions that map between different coordinates in regions where two charts overlap are all differentiable (smooth) on $\mathbb{R}^D$, the manifold is a ← *differentiable (smooth) manifold*. On a differentiable manifold we can talk about the differentiation of functions defined on the manifold. In physics we consider almost exclusively such manifolds:



- ← *Tangent and cotangent spaces* (Section 3.3):

  Given a differentiable manifold (which is not a vector space in general!), there is a canonical way to associate a vector space to every point of the manifold: the ← *tangent space* T_pM. Mathematically, it is the vector space of directional derivative operators that act on smooth functions on that point. Given a coordinate chart, the directional derivatives along the coordinates (evaluated at $p \in M$) induce a basis $\{\partial_i|_p\}$ of the tangent space T_pM (different coordinates lead to different bases). In addition, to every vector space there is an associated *dual space* spanned by the linear forms on the vector space; thus there is a dual tangent space: the ← *cotangent space* $T_p^* M$. It is spanned by the dual basis $\{\mathrm{d}x_p^i\}$ of differential forms:



$$\underbrace{\text{Tangent space } T_pM \text{ at } p \in M}_{\substack{\text{Vector space of directional derivatives}\\ \text{with evaluation at } p\in M.}} \quad = \quad \underbrace{\operatorname{span}\left\{\partial_i|_p \,|\, i=1,\dots,D\right\}}_{\substack{\text{Spanned by coordinate basis}\\ \text{derived from given chart.}}} \tag{10.1}$$

With the dual basis (we often drop the subscript p)

$$\mathrm{d}x_p^i(\partial_j|_p) := \delta^i_j = \left.\frac{\partial x^i}{\partial x^j}\right|_p \tag{10.2}$$

we can define the

$$\text{Cotangent space } T_p^*M \text{ at } p\in M \quad = \quad \operatorname{span}\left\{\mathrm{d}x_p^i \,|\, i=1,\dots,D\right\} \tag{10.3}$$

- ← *Tensor fields* (Sections 3.2 to 3.4):

  Since there are canonical vector and covector spaces associated to every point of the manifold, we can consider (reasonably smooth) functions that map every point of the manifold to a tensor product of p vectors and q covectors; we call such functions ← *tensor fields* of rank (p,q). They are "geometric objects" in that they are independent of coordinate charts; physical quantities (like the electromagnetic field) must be represented by such fields. Once we have chosen a coordinate chart, we can encode these fields in terms of their *components* wrt. the coordinate basis on tangent and cotangent space. The coordinate independence of tensor fields translates then into a particular transformation law for their components:

  $\sphericalangle$ Coordinate transformation $\bar x = \varphi(x) \;\Leftrightarrow\; x = \varphi^{-1}(\bar x)$

  $\rightarrow$ (p,q)-Tensor (field) $T :\Leftrightarrow$

$$\overbrace{\bar T^{i_1\dots i_p}{}_{j_1\dots j_q}(\bar x)}^{=\bar T^I{}_J(\bar x)} = \underbrace{\left[\frac{\partial \bar x^{i_1}}{\partial x^{m_1}}\cdots\frac{\partial \bar x^{i_p}}{\partial x^{m_p}}\right]}_{=:\frac{\partial \bar x^I}{\partial x^M}} \underbrace{\left[\frac{\partial x^{n_1}}{\partial \bar x^{j_1}}\cdots\frac{\partial x^{n_q}}{\partial \bar x^{j_q}}\right]}_{=:\frac{\partial x^N}{\partial \bar x^J}} \underbrace{T^{m_1\dots m_p}{}_{n_1\dots n_q}(x)}_{=T^M{}_N(x)} \tag{10.4}$$

  (Einstein sum convention = Sums over pairs of up- and down indices are implied.)

Examples:

$$(0,0)\text{-Tensor} \equiv \text{Scalar:} \qquad \bar{\Phi}(\bar{x}) = \Phi(x) \tag{10.5a}$$

$$(1,0)\text{-Tensor} \equiv \text{Contravariant vector:} \qquad \bar{A}^i(\bar{x}) = \frac{\partial \bar{x}^i}{\partial x^k} A^k(x) \tag{10.5b}$$

$$(0,1)\text{-Tensor} \equiv \text{Covariant vector:} \qquad \bar{B}_i(\bar{x}) = \frac{\partial x^k}{\partial \bar{x}^i} B_k(x) \tag{10.5c}$$

$$(1,1)\text{-Tensor} \equiv \text{(Mixed) Tensor:} \qquad \bar{T}^i{}_j(\bar{x}) = \frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^l}{\partial \bar{x}^j} T^k{}_l(x) \tag{10.5d}$$

- ← *Riemannian metric* (Section 3.5):

  A Riemannian metric is a $(0,2)$ tensor field with a few additional properties (symmetry and non-degeneracy) so that it defines a (pseudo-)inner product on the tangent space at every point of the manifold. A differentiable manifold equipped with such a metric is called a ← *Riemannian manifold*. On a Riemannian manifold we can measure angles between tangent vectors and lengths of curves:

$$\text{Riemannian (pseudo-)metric } ds^2 := \left\{\begin{matrix}\text{Symmetric}\\ \text{non-degenerate}\\ (0,2)\text{-tensor field}\end{matrix}\right\} \tag{10.6}$$

  More formally:

$$ds^2 : M \ni p \mapsto \underbrace{\left(ds_p^2 : T_pM \times T_pM \to \mathbb{R}\right)}_{\text{Bilinear \& symmetric \& non-degenerate}} \in T_p^*M \otimes T_p^*M \tag{10.7}$$

  with coordinate representation

$$ds_p^2 = \sum_{i,j=1}^{D} g_{ij}(x)\, dx^i \otimes dx^j \equiv g_{ij}(x)\, dx^i dx^j \tag{10.8}$$

  where $g_{ij} = g_{ji}$ (symmetry) and $g = \det(g_{ij}) \neq 0$ (non-degeneracy).

  A Riemannian metric allows us to define the following geometric concepts:

  - Angle between two vectors $A = A^i\partial_i$, $B = B^i\partial_i \in T_pM$:

$$\langle A, B\rangle \equiv ds_p^2(A,B) = g_{ij}(p)A^iB^j \equiv \|A\|_p\|B\|_p\cos\theta \tag{10.9}$$

    with the norm on T_pM

$$\|A\|_p := \sqrt{ds_p^2(A,A)} = \sqrt{g_{ij}(p)A^iA^j}\,. \tag{10.10}$$

  - Length of curve $\gamma : [a,b] \to M$:

$$L[\gamma] := \int_a^b \sqrt{g_{ij}(\gamma(t))\frac{d\gamma^i(t)}{dt}\frac{d\gamma^j(t)}{dt}}\, dt = \int_a^b \underbrace{\|\dot{\gamma}(t)\|_{\gamma(t)}}_{\text{“Velocity”}} dt\,. \tag{10.11}$$

- ← *Pulling indices up and down* (Section 3.5):

  A symmetric, non-degenerate bilinear form defines a canonical isomorphism between a vector space and its dual. A special case is a Riemannian metric which provides us with a isomorphism between tangent and cotangent spaces at every point of the manifold. In tensor calculus, this isomorphism is applied by "pulling indices up and down" with the metric:

$$\text{Pulling } \textit{down}: \quad T^{i_1\dots\square\dots i_p\,\square\dots\square}{}_{\square\dots\, i\,\dots\square\, j_1\dots j_q} := g_{ik}\, T^{i_1\dots\, k\,\dots i_p\,\square\dots\square}{}_{\square\dots\square\dots\square\, j_1\dots j_q} \tag{10.12a}$$

$$\text{Pulling } \textit{up}: \quad T^{i_1\dots i_p\,\square\dots\, j\,\dots\square}{}_{\square\dots\square\, j_1\dots\square\dots j_q} := g^{jk}\, T^{i_1\dots i_p\,\square\dots\square\dots\square}{}_{\square\dots\square\, j_1\dots\, k\,\dots j_q} \tag{10.12b}$$

  where g^{ij} is the *inverse metric* defined via $g^{ik}g_{kj} \overset{!}{=} \delta^i_j$.

- ← *Christoffel symbols and covariant derivatives* (Section 3.6):

  We realized that the partial derivatives of tensor fields are *not* tensor fields themselves (this only works for scalars). This motivated the introduction of a "patched up derivative," the so called ← *covariant derivative* that transforms again like a tensor. To define the covariant derivative, we needed a set of (non-tensorial) functions called ← *Christoffel symbols* that were defined by a given Riemannian metric:

$$\downarrow \text{Covariant derivative wrt. } k$$

$$\text{Scalar:} \quad \Phi_{;k} := \Phi_{,k} \tag{10.13a}$$

$$\text{Contravariant vector:} \quad A^i{}_{;k} := A^i{}_{,k} + \Gamma^i{}_{kl}\, A^l \tag{10.13b}$$

$$\text{Covariant vector:} \quad B_{i;k} := B_{i,k} - \Gamma^l{}_{ik}\, B_l \tag{10.13c}$$

  with $\Phi_{,k} \equiv \partial_k \Phi$ etc. and the ← *Christoffel symbols (of the second kind)*

$$\Gamma^i{}_{kl} := \frac{1}{2} g^{im}\left(g_{mk,l} + g_{ml,k} - g_{kl,m}\right). \tag{10.14}$$

  ¡! In the following two sections we will revisit, motivate and study the concept of a covariant derivative in more detail. We will also see where the Christoffel symbols come from and which role they play geometrically on the manifold.

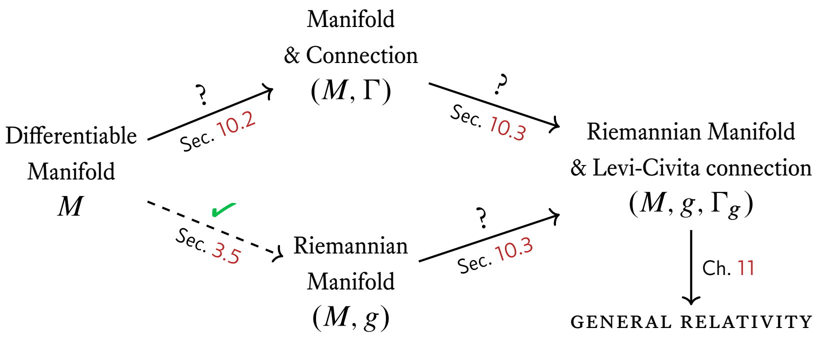
  So if you were not satisfied with the way the covariant derivative and the Christoffel symbols appeared out of thin air in Section 3.6: Now comes the proper introduction!

- ← *Manifest covariance* (Section 3.6):

  The whole point of our endeavor was to find a mathematical toolbox that allows us to write down equations that are guaranteed to be form-invariant under arbitrary coordinate transformations. These equations describe relations between geometric objects on a manifold, such that their content is independent of the chosen coordinate chart. This toolbox is called ← *tensor calculus* and consists of rules how to combine/construct tensors (e.g. via multiplication, contraction of indices, covariant derivatives, ...) to form generally covariant equations. The general covariance of tensorial equations is *manifest* because their mere structure guarantees general covariance:

$$T^I{}_J(x) = 0 \quad \xleftrightarrow[\bar{T}^I{}_J(\bar{x}) = \frac{\partial \bar{x}^I}{\partial x^M}\frac{\partial x^N}{\partial \bar{x}^J} T^M{}_N(x)]{\text{Coordinate trafo: } \bar{x}=\varphi(x)} \quad \bar{T}^I{}_J(\bar{x}) = 0 \tag{10.15}$$

2 | Plan:



- Section 10.2:
  Introduce and study → *connections* and the concept of parallel transport and curvature.
- Section 10.3:
  Use a Riemannian metric to derive a special connection: the → *Levi-Civita connection*.
  Study properties of this special connection: Riemannian curvature and geodesic curves.

10.2. Affine connections

- An *affine connection* is an additional structure on a differentiable manifold (no metric needed!) that allows for the definition of the following concepts:
  - Parallel transport
  - Covariant derivatives
  - Autoparallel curves
  - Curvature

  "Additional" means that it is not intrinsic or canonical to a manifold; *you* can add a connection to obtain more structure. It also implies that typically there are many connections to choose from.
- Terminology:

  In modern differential geometry, the term "connection" has a rather broad meaning. Generally speaking, a connection is a structure that allows one to "parallel transport" objects along curves on a manifold. The most straightforward objects to move around are vectors taken from the tangent spaces of the manifold; this type of connection is called an → *affine connection*, and it is this variety we use in GENERAL RELATIVITY.

  However, you can also (artificially) attach other spaces to every point of a manifold (e.g., Lie groups like U(1)). Then you can ask how objects of these spaces are parallel transported around the manifold. This gives rise to other types of connections that are particularly important in modern formulations of gauge theories (↑ *gauge connections*). The gauge field A^μ of electrodynamics is an example of a U(1) gauge connection; it transports U(1) phases around (not tangent vectors) and is therefore not an affine connection.

  In the following we will often drop the "affine" and simply talk about "connections." Keep in mind, however, that we only consider affine connections in this chapter (and this course).

1 | ⊲ Differentiable D-dimensional manifold M; vector field $A = A^i \partial_i$; scalar field Φ:

$$\partial_k \Phi \quad \rightarrow \quad \checkmark \text{ covariant vector field } (\leftarrow \textit{Eq.}\ (3.19)) \tag{10.16a}$$

$$\partial_k A^i \quad \rightarrow \quad \text{✗ } \textit{no} \text{ tensor field } (\leftarrow \textit{Eq.}\ (3.73)) \tag{10.16b}$$

This is a problem because we often need derivatives of tensors to formulate physical models; and since these equations must be generally covariant (GRP!), we need them to transform as tensors!

2 | Problem:

i | $\sphericalangle$ Directional derivative of $A = A^i\partial_i$ along a curve $\gamma(\lambda)$ with $\gamma(0) = p \in M$:

$$\left.\frac{\mathrm{d}A(\gamma(\lambda))}{\mathrm{d}\lambda}\right|_{\lambda=0} \stackrel{?}{=} \lim_{\delta\lambda\to 0}\frac{A(\gamma(\delta\lambda)) - A(\gamma(0))}{\delta\lambda} \equiv \lim_{\delta\lambda\to 0}\frac{\overbrace{A(q) - A(p)}^{\text{↯ Undefined!}}}{\delta\lambda} \tag{10.17}$$

Note that $A(q) \in T_qM$ and $A(p) \in T_pM$, i.e., these values of the vector field belong to *different* vector spaces. Hence their difference is completely undefined!

ii | We can of course try to work with the *components* of the vector field wrt. a given chart instead. Since $A^i \in \mathbb{R}$, the following expression is at least well-defined:

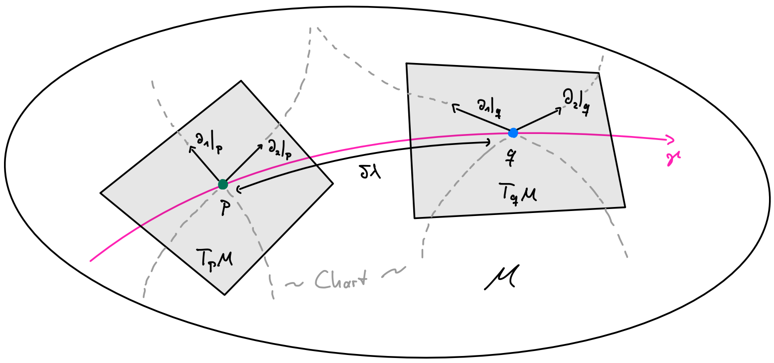
$$\left.\frac{\mathrm{d}A^i(\gamma(\lambda))}{\mathrm{d}\lambda}\right|_{\lambda=0} = \lim_{\delta\lambda\to 0}\frac{A^i(q) - A^i(p)}{\delta\lambda} \tag{10.18}$$

Unfortunately this does not solve the problem, because these components are given wrt. to *different*, *coodinate-dependent* bases on T_qM and T_pM, respectively:

$$A(q) = A^i(q)\partial_i|_q \quad \text{with} \quad \mathrm{span}\left\{\partial_i|_q\right\} = T_qM \tag{10.19}$$

$$A(p) = A^i(p)\partial_i|_p \quad \text{with} \quad \mathrm{span}\left\{\partial_i|_p\right\} = T_pM \tag{10.20}$$

iii | To understand why this is a problem, imagine you fix the basis $\{\partial_i|_p\}$ of T_pM; this does not fix the basis $\{\partial_i|_q\}$ of T_qM because choosing different (curvilinear) coordinates can be used to modify the induced basis $\{\partial_i|_q\}$ without changing $\{\partial_i|_p\}$:



As a consequence, the components $A^i(q)$ can be modified arbitrarily without changing the vector field A itself. Thus the difference $A^i(q) - A^i(p)$, and thereby the directional derivative above, do not encode a geometric, coordinate independent object! Mathematically, this is reflected in the *non-tensorial* transformation of the difference under arbitrary coordinate transformations

$$\bar{A}^i(q) - \bar{A}^i(p) = \frac{\partial\bar{x}^i(q)}{\partial x^k}A^k(q) - \frac{\partial\bar{x}^i(p)}{\partial x^k}A^k(p) \neq \frac{\partial\bar{x}^i(p)}{\partial x^k}\left[A^k(q) - A^k(p)\right] \tag{10.21}$$

This explains why partial derivatives of the form $\partial_k A^i$ (which are simply directional derivatives along coordinate axes) fail to transform as tensors.

3 | Idea:

The problem is conceptually most transparent in Eq. (10.17) which is mathematically undefined. However, if we could make it well-defined, we would immediately obtain a geometric, coordinate-independent object. To make the difference between the two vectors well-defined, they must live in the same tangent space, though.

Our only way out is to assume that we are given some function $\Gamma_{p\to q} : T_pM \to T_qM$ that establishes a correspondence between the two nearby tangent spaces by "parallel transporting" vectors between them. We then could "parallel transport" $A(p)$ from T_pM to T_qM like so: $\Gamma_{p\to q}(A(p)) \in T_qM$. With this new vector, the difference is mathematically well-defined:

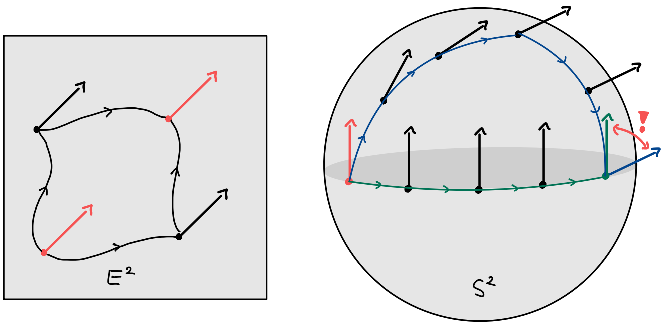
$$\frac{\mathrm{D}A}{\mathrm{D}\lambda} := \lim_{\delta\lambda\to 0} \frac{\overbrace{A(q) - \Gamma_{p\to q}(A(p))}^{\in T_qM}}{\delta\lambda} \quad \text{or} \quad \frac{\mathrm{D}A^i}{\mathrm{D}\lambda} := \lim_{\delta\lambda\to 0} \frac{A^i(q) - A^i(p \xrightarrow{\Gamma} q)}{\delta\lambda} \tag{10.22}$$

We use the capital letter D to indicate that the difference in the numerator of the difference quotient has been modified by (and depends on) Γ.

→ $\Gamma_{p\to q}$ is an ⁂ *affine connection*

(This is not yet very rigorous, we will specify our idea more formally below.)

- As already mentioned, the interpretation of an affine connection Γ is that it formalizes the notion of "parallel translating" or "parallel transporting" tangent vectors along curves on the manifold from one tangent space to another. It is important to realize that the notion of "parallel transport" is mathematically subtle and not trivial. It must be carefully defined and can lead to quite surprising results when considering curved manifolds:



  Note that the (intuitive) parallel transport on the Euclidean plane (left) is independent of the path along which the vector is transported. By contrast, intuitively transporting vectors on a sphere (right) yields different results depending on the chosen path. The fact that there is no unique "parallel vector" to a given vector, but that the notion of parallelelism depends on the path taken, is the hallmark of → *curvature*.

- To be clear: the failure to produce a tensorial object from the directional derivative of a vector field is a fundamental and not a technical issue. We were neither too naïve when performing the derivative in Eq. (10.17), nor will our "solution" Eq. (10.22) render it magically tensorial. It is *impossible* to define a tensorial derivative on a manifold without specifying an additional structure (namely an affine connection Γ).

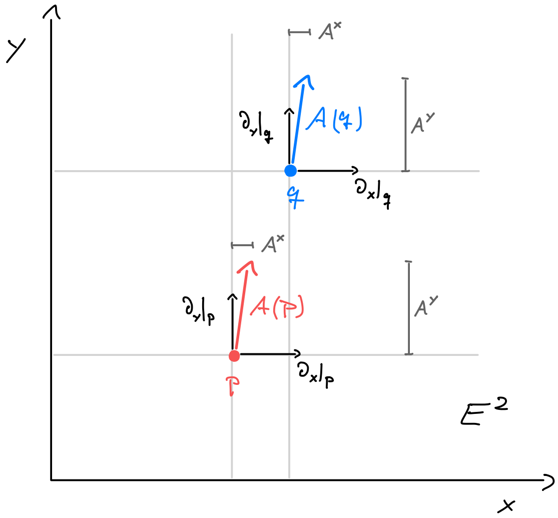
4 | Motivation:

To understand the properties of parallel transport (and thereby an affine connection Γ) better, we consider the simple example of parallel transporting a vector in the affine space $M = \boldsymbol{E}^2 = \mathbb{R}^2$ (the Euclidean plane), described in curvilinear polar coordinates:

i | $\triangleleft$ $M = \boldsymbol{E}^2$ with Cartesian coordinates $\vec{x} = (x, y)$ and Cartesian basis $\{\partial_x, \partial_y\}$:

$\triangleleft$ "Constant" vector field

$$A = A^x \partial_x + A^y \partial_y \quad \text{with } A^x = \text{const},\ A^y = \text{const:} \tag{10.23}$$



If you are given a differentiable manifold without additional structure, it does not make sense to ask whether a vector field "is constant." For example, if we consider $M = \boldsymbol{E}^2$ as a manifold (forgetting about its Euclidean metric and affine structure), it does not make sense to call the vector field A "constant"; its *components* are constant wrt. to the basis induced by a specific coordinate system. However, a coordinate-independent statement like $A(p) = A(q)$ for all $p, q \in M$ is nonsensical because $A(p) \in T_pM$ and $A(q) \in T_qM$, and there is no canonical isomorphism connecting T_pM and T_qM; without an affine connection Γ, these are completely unrelated vector spaces and we do not know how to compare vectors at different points on the manifold (there is no concept of "parallel" vectors).

ii | $\triangleleft$ Coordinate transformation $(x, y) = \varphi^{-1}(r, \theta)$ to polar coordinates:

$$x = r \cos\theta \tag{10.24a}$$
$$y = r \sin\theta \tag{10.24b}$$

$\overset{\circ}{\rightarrow}$ Induced basis change on tangent spaces (← *Eq.* (3.5)):

$$\partial_r = \cos\theta\, \partial_x + \sin\theta\, \partial y \tag{10.25a}$$
$$\partial_\theta = -r \sin\theta\, \partial_x + r \cos\theta\, \partial_y \tag{10.25b}$$

$\overset{\circ}{\rightarrow}$ Components of vector field:

$$A = A^x \partial_x + A^y \partial_y = A^r \partial_r + A^\theta \partial_\theta \tag{10.26}$$

with (no longer constant!)

$$A^r(r, \theta) = A^x \cos\theta + A^y \sin\theta \tag{10.27a}$$
$$A^\theta(r, \theta) = \tfrac{1}{r}(A^y \cos\theta - A^x \sin\theta)\,. \tag{10.27b}$$

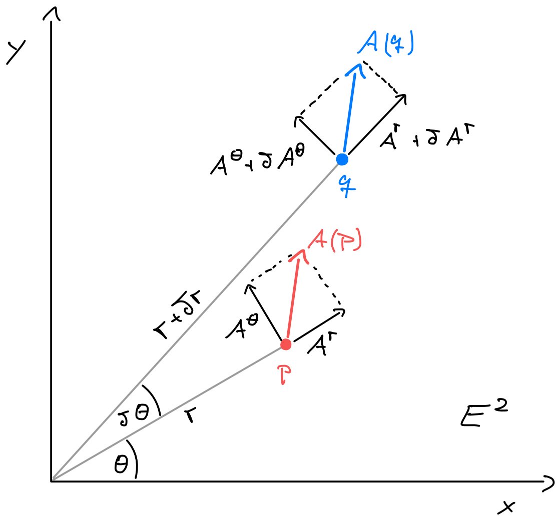
iii | $\triangleleft$ Two infinitesimally separated points $p, q \in \boldsymbol{E}^2$ with coordinates

$$u(p) = (r, \theta) \quad \text{and} \quad u(q) = (r + \delta r, \theta + \delta\theta) \tag{10.28}$$

and associated vectors ($A^r = A^r(p)$ and $A^\theta = A^\theta(p)$)

$$A(p) = A^r \partial_r + A^\theta \partial_\theta \,, \tag{10.29a}$$

$$A(q) = [A^r + \delta A^r]\partial_r + [A^\theta + \delta A^\theta]\partial_\theta \,. \tag{10.29b}$$



Eq. (10.27) $\overset{\circ}{\to}$ (via Taylor expansion)

$$\delta A^r = r A^\theta \delta\theta \tag{10.30a}$$

$$\delta A^\theta = -\tfrac{1}{r}(A^\theta \delta r + A^r \delta\theta) \tag{10.30b}$$

If we now *declare* the vector field A to be *constant*, the variations Eq. (10.30) must be "fake" in the sense that they are caused by our choice of curvilinear coordinates rather an "intrinsic" variation of the vector field itself.

→ This choice specifies an → *affine connection*.

iv | Now that we *specified* which changes of the components of vector fields (in our coordinate system) are considered to be "fake", i.e., artifacts of the coordinates, we can define the "real" changes of arbitrary vector fields (which then can be non-constant wrt. our specific notion of parallel vectors) as their "complete" variation corrected by the "fake" variation δA^i:

$\sphericalangle$ Arbitrary ("non-constant") vector field with $B^i(p) = B^i(r,\theta)$

$\xrightarrow{\text{Eq. (10.30)}}$ "True change" due to "intrinsic" variation of the vector field:

$$[B^r(q) - B^r(p)] - \delta B^r = \frac{\partial B^r}{\partial r}\delta r + \left(\frac{\partial B^r}{\partial \theta} - r B^\theta\right)\delta\theta \tag{10.31a}$$

$$[B^\theta(q) - B^\theta(p)] - \delta B^\theta = \left(\frac{\partial B^\theta}{\partial r} + \frac{1}{r}B^\theta\right)\delta r + \left(\frac{\partial B^\theta}{\partial \theta} + \frac{1}{r}B^r\right)\delta\theta \tag{10.31b}$$

The idea is to use such "corrected" differences in the numerator of a difference quotient like Eq. (10.22) to define a derivative of the vector field that transforms like a tensor.

That is, we define

$$A^i(p \xrightarrow{\Gamma} q) = A^i(p) + \delta A^i(p) \,. \tag{10.32}$$

5 | Generalization:

Drawing from the example and the form of the particular connection Eq. (10.30), we can select reasonable properties that an general affine connection should satisfy (in terms of components):

(i) $A^i(p \xrightarrow{\Gamma} q)$ is *linear* in $A^i(p)$.

(ii) The variation $\delta A^i(p)$ is *linear* in the first-order variation δx^i of coordinates.

We can satisfy both conditions if the variation has the general form (the minus is convention)

$$\boxed{\delta A^i(p) = -\Gamma^i{}_{kl}(p) A^k(p)\, \delta x^l} \tag{10.33}$$

$\rightarrow$

$$A^i(p \xrightarrow{\Gamma} q) = A^i(p) + \delta A^i(p) = \left[\delta^i_k - \Gamma^i{}_{kl}(p)\,\delta x^l\right] A^k(p) \tag{10.34}$$

with some undetermined set of coefficients $\Gamma^i{}_{kl}$ that completely specify the affine connection (in the particular coordinates chosen):

$$\boxed{\Gamma^i{}_{kl}(x) \,:\, ⁂\ \textit{(coefficients of the) affine connection}\ \Gamma\ (\text{in } x)}$$

Example:

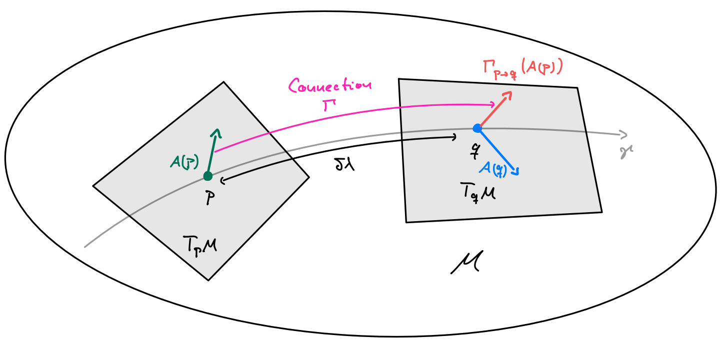
From Eq. (10.30) and Eq. (10.33) it follows for the coefficients of the affine connection of the Euclidean plane, expressed in polar coordinates (➔ Problemset 2):

$$\Gamma^r{}_{kl}(r,\theta) = \begin{pmatrix} 0 & 0 \\ 0 & -r \end{pmatrix}_{kl} \quad \text{and} \quad \Gamma^\theta{}_{kl}(r,\theta) = \begin{pmatrix} 0 & \frac{1}{r} \\ \frac{1}{r} & 0 \end{pmatrix}_{kl} . \tag{10.35}$$

6 | Interpretation:

The affine connection establishes a connection (hence the name) between tangent spaces at different points on the manifold by establishing a notion of *parallelism*:

$$\begin{aligned} \overbrace{A(p)}^{\in T_pM} \xrightarrow[\text{parallel transport}]{\text{Infinitesimal}} \overbrace{\Gamma_{p\to q}(A(p))}^{\in T_qM} &= A^i(p \xrightarrow{\Gamma} q)\, \partial_i|_q \\ &= [A^i(p) + \delta A^i(p)]\, \partial_i|_q \\ &= [\delta^i_k - \Gamma^i{}_{kl}\delta x^l]\, A^k(p)\partial_i|_q \end{aligned} \tag{10.36}$$



We say: $\Gamma_{p\to q}(A(p))$ is the vector in q that is *parallel* to $A(p)$ in p.

7 | ⁂ *Absolute derivative:*

We can now express the absolute derivative using the connection:

$$\boxed{\frac{\mathrm{D}A^i}{\mathrm{D}\lambda} \overset{10.22}{\underset{10.32}{:=}} \lim_{\delta\lambda\to 0} \frac{\overbrace{A^i(\gamma(\lambda+\delta\lambda)) - A^i(\gamma(\lambda))}^{dA^i} - \delta A^i}{\delta\lambda} \overset{10.33}{=} \frac{\mathrm{d}A^i}{\mathrm{d}\lambda} + \Gamma^i{}_{kl}A^k \frac{\mathrm{d}x^l}{\mathrm{d}\lambda}} \tag{10.37}$$

We want the absolute derivative to transform as a contravariant vector:

$$\frac{\mathrm{D}\bar{A}^i}{\mathrm{D}\lambda} \overset{!}{=} \frac{\partial\bar{x}^i}{\partial x^k}\frac{\mathrm{D}A^k}{\mathrm{D}\lambda} \tag{10.38}$$

A straightforward but cumbersome calculation shows [recall Section 3.6] that this is the case if and only if the connection coefficients transform as follows:

$$\boxed{\bar{\Gamma}^i{}_{kl} \overset{\circ}{=} \underbrace{\frac{\partial\bar{x}^i}{\partial x^m}\frac{\partial x^n}{\partial\bar{x}^k}\frac{\partial x^o}{\partial\bar{x}^l}\Gamma^m{}_{no}}_{\text{Tensor ✓}} + \underbrace{\frac{\partial\bar{x}^i}{\partial x^p}\frac{\partial^2 x^p}{\partial\bar{x}^k\partial\bar{x}^l}}_{\text{No tensor ✗}}} \tag{10.39}$$

→ $\Gamma^i{}_{kl}$ does *not* transform as a tensor!

- ¡! For a given manifold M, there are *infinitely many* choices for an affine connection Γ.
- ¡! The definition Eq. (10.37) makes sense for any contravariant vector A^i that is defined (and differentiable) along the curve $\gamma(\lambda)$ [for example, a particle trajectory $x^\mu(\lambda)$]. Although we considered a vector field A^i in our discussion, it is *not* necessary for A^i to be defined *in the neighborhood* of the trajectory $\gamma(\lambda)$; i.e., partial derivatives $\partial_j A^i$ do not need to be defined for the definition of the absolute derivative Eq. (10.37). This is why we distinguish between the *absolute derivative* and the → *covariant derivative*.
- The additional term that makes the transformation of the connection coefficients non-tensorial is needed to compensate for a corresponding non-tensorial term from the total (non-covariant) derivative $\frac{\mathrm{d}A^i}{\mathrm{d}\lambda}$.
- Every set of fields $\Gamma^i{}_{kl}$ that transforms according to Eq. (10.39) can be used to define a connection (and therefore a notion of what "parallel" means on a manifold). This definition allows for more solutions than the specific type of connection that we used for our motivation, namely connections derived from declaring a given vector field as "constant." Interestingly, not all connections can be constructed in this way (the ones that can are actually quite boring because they do not have → *curvature*), and in Section 10.3 we will find a recipe to construct a special connection from every Riemannian metric.

8 | Torsion:

In general, the connection coefficients are *not* symmetric in their lower two indices. →

$$\Gamma^i{}_{kl} = \underbrace{\frac{1}{2}\left(\Gamma^i{}_{kl} + \Gamma^i{}_{lk}\right)}_{\Gamma^i{}_{(kl)}} + \underbrace{\overbrace{\frac{1}{2}\left(\Gamma^i{}_{kl} - \Gamma^i{}_{lk}\right)}^{S^i{}_{kl}}}_{\Gamma^i{}_{[kl]}} \tag{10.40}$$

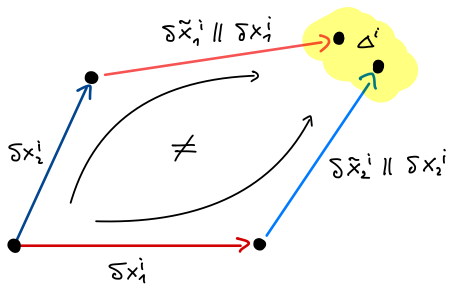
$\xrightarrow{\text{Eq. (10.39)}}$ (Note that the non-tensorial part in Eq. (10.39) is symmetric in k and l!)

$$\bar{S}^i{}_{kl} = \frac{\partial \bar{x}^i}{\partial x^m}\frac{\partial x^n}{\partial \bar{x}^k}\frac{\partial x^o}{\partial \bar{x}^l} S^m{}_{no} \tag{10.41}$$

→ *Antisymmetric* part $S^i{}_{kl}$ of connection is a *tensor*: ⁂ *Torsion tensor*

- ¡! This is not true for the *symmetric* part.
- GENERAL RELATIVITY is based on the assumption that the affine connection of spacetime is *torsion-free*. Hence it is sufficient to focus on symmetric, torsion-free connections to formulate the theory.
- Interpretation:

  On a manifold with torsion, infinitesimal parallelograms do not close:



  To see this, consider two infinitesimal vectors δx_1^i and δx_2^i at some point $p \in M$. Then parallel transport δx_1^i along δx_2^i to produce $\delta\tilde{x}_1^i$ and vice versa:

$$\delta\tilde{x}_1^i = \delta x_1^i - \Gamma^i{}_{kl}(\delta x_1^k)(\delta x_2^l)\,, \tag{10.42a}$$
$$\delta\tilde{x}_2^i = \delta x_2^i - \Gamma^i{}_{kl}(\delta x_2^k)(\delta x_1^l)\,. \tag{10.42b}$$

  The amount by which this infinitesimal parallelogram does not close is:

$$\begin{aligned}\Delta^i &:= (\delta x_1^i + \delta\tilde{x}_2^i) - (\delta x_2^i + \delta\tilde{x}_1^i) = (\delta x_1^i - \delta\tilde{x}_1^i) - (\delta x_2^i - \delta\tilde{x}_2^i)\\ &\overset{10.42}{=} (\Gamma^i{}_{kl} - \Gamma^i{}_{lk})(\delta x_1^k)(\delta x_2^l) \overset{\text{def}}{=} S^i{}_{kl}(\delta x_1^k)(\delta x_2^l)\,.\end{aligned} \tag{10.43}$$

  Non-vanishing torsion therefore implies:

$$\Delta^i = S^i{}_{kl}\,(\delta x_1^k)\,(\delta x_2^l) \neq 0 \quad\Leftrightarrow\quad S^i{}_{kl}\,(\delta x_1^k)\,(\delta x_2^l) \neq S^i{}_{kl}\,(\delta x_2^k)\,(\delta x_1^l) \tag{10.44}$$

  → The direction of paths matters: First going along δx_1^k and then parallel to δx_2^l leads to a different point than doing the opposite. (Similar to the motion of a screw, which is different for clockwise and counterclockwise rotation.)

- It is possible to extend GENERAL RELATIVITY by allowing the torsion of spacetime to be non-zero (and dynamic as well) [130, 131]. In such theories, the ↓ *spin* of particles becomes the source of torsion, just as their *mass* is the source of → *curvature*. Such theories can predict additional forces between spinful particles, see Ref. [132] for a review.
- Since torsion is "just another tensor field" (which is not true for the symmetric part of the connection), it is reasonable to keep a geometric theory of gravity slim and assume torsion to vanish. If the theory matches observations, we didn't produce unnecessary clutter by dragging torsion along (↓ *Occam's razor*); however, if there happen to be phenomena that cannot be explained, we can still "patch" the theory by adding new (tensor) fields (that might play the role of torsion). In any case, there is no experimental evidence to date that makes a torsion field necessary.

→ Henceforth we consider only torsion-free connections:

$$\boxed{\Gamma^i{}_{kl} = \Gamma^i{}_{lk}}$$

9 | Locally geodesic coordinate systems:

Since we know how the coefficients of a connection transform, we can ask whether there are special coordinate systems in which the connection looks particularly simple:

Details: ➲ Problemset 2

i | Goal:

Show that for every point $p \in M$ there is a coordinate system in which the connection coefficients *in this point* vanish:

$$\forall p \in M \; \exists \text{ Chart } u \text{ with } u(p) = x_0 : \; \Gamma^i{}_{kl}(x_0) = 0 \quad \forall_{ijk} \tag{10.45}$$

u: ⁂ *Locally geodesic coordinate system*

ii | First, show the alternative form of the transformation: (recall Eq. (3.75))

$$\bar{\Gamma}^i{}_{kl} \stackrel{\circ}{=} \frac{\partial \bar{x}^i}{\partial x^m}\frac{\partial x^n}{\partial \bar{x}^k}\frac{\partial x^o}{\partial \bar{x}^l}\Gamma^m{}_{no} - \frac{\partial x^m}{\partial \bar{x}^l}\frac{\partial x^p}{\partial \bar{x}^k}\left(\frac{\partial^2 \bar{x}^i}{\partial x^p \partial x^m}\right) \tag{10.46}$$

This follows from Eq. (10.39) by differentiating $\frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^k}{\partial \bar{x}^j} = \delta^i_j$.

iii | $\sphericalangle$ Coordinates v with $v(p) = 0 \in \mathbb{R}^D$ (in general it is $\Gamma^i{}_{kl}(0) \neq 0$ in this chart)

→ Coordinate transformation $\bar{x} = \varphi(x) = u \circ v^{-1}(x)$ in vicinity of $p \in M$:

$$\bar{x}^i = x^i + \tfrac{1}{2}C^i{}_{kl}(0)\, x^k x^l + \dots \tag{10.47}$$

with (*w.l.o.g.*) *symmetric* coefficients $C^i{}_{kl} = C^i{}_{lk}$.

iv | → Partial derivatives at $u(p) = 0 = v(p)$:

$$\left.\frac{\partial \bar{x}^i}{\partial x^m}\right|_{x=0} = \delta^i_m \quad \text{and} \quad \left.\frac{\partial^2 \bar{x}^i}{\partial x^p \partial x^m}\right|_{x=0} = C^i{}_{pm}(0) \tag{10.48}$$

$$\xrightarrow{\text{Eq. (10.46)}} \bar{\Gamma}^i{}_{kl} \stackrel{\circ}{=} \Gamma^i{}_{kl} - C^i{}_{kl}$$

v | $\bar{\Gamma}^i{}_{kl}(0) \stackrel{!}{=} 0$ and $\bar{\Gamma}^i{}_{kl} = \bar{\Gamma}^i{}_{lk}$ (torsion-free!) → $C^i{}_{kl}(0) := \Gamma^i{}_{kl}(0)$ ■

Notes:

- ¡! Note that we only showed that the connection coefficients can be made zero *in a single point*; in general one cannot find a coordinate system where the coefficients vanish everywhere. This also implies that in general the *derivatives* $\partial_m \Gamma^i{}_{kl}(0)$ do *not* vanish in p.
- In locally geodesic coordinates, the absolute derivative Eq. (10.37) is simply the "normal" total derivative. As a consequence, in the context of Riemannian manifolds, the coordinate lines are local geodesics ("shortest paths", → *later*) – hence the name.
- The above argument fails for connections with non-vanishing torsion $S^i{}_{kl} \neq 0$ since the latter transforms as a tensor and cannot be zeroed by a coordinate transformation (unless it vanishes in all coordinates).

- The fact that locally geodesic coordinates exist at every point will be the foundation for the implementation of Einstein's equivalence principle EEP in the mathematical framework of GENERAL RELATIVITY. Physically, these coordinates will be identified with the free falling, local inertial frames.