↓ **Lecture 8** [03.12.25]

3. Mathematical Tools I: Tensor Calculus

In this chapter we introduce tensor calculus (↑ *Ricci calculus*) for general coordinate transformations φ (which will be useful both in SPECIAL RELATIVITY and GENERAL RELATIVITY). The coordinate transformations φ relevant for SPECIAL RELATIVITY are Lorentz transformations (and therefore linear) which simplifies expressions often significantly (→ *Chapter 4*). However, this special feature of coordinate transformations in SPECIAL RELATIVITY is not crucial for the discussions in this chapter.

Goal: Construct Lorentz covariant (form invariant) equations
(for mechanics, electrodynamics, quantum mechanics)

Question: How to do this *systematically*?

Note that (we suspect that) Maxwell equations *are* Lorentz covariant. Clearly this is not obvious and requires some work to prove; we say that the Lorentz covariance is *not manifest*: it is there, but it is hard to see. Conversely, without additional tools that make Lorentz covariance more obvious, it is borderline impossible to *construct* Lorentz covariant equations from scratch (which we must do for mechanics and quantum mechanics!).

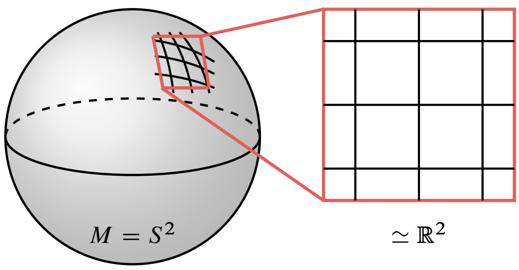
We are therefore looking for a "toolkit" that provides us with elementary "building blocks" and a set of rules that can be used to construct Lorentz covariant equations. This toolbox is known as *tensor calculus* or ↑ *Ricci calculus*; the "building blocks" are tensor fields and the rules for their combination are given by index contractions, covariant derivatives, etc. The rules are such that the expressions (equations) you can build with tensor fields are *guaranteed* to be Lorentz covariant. This implies in particular that if you can rewrite any given set of equations (like the Maxwell equations) in terms of these rules, you automatically show that the equations were Lorentz covariant all along. We then say that the Lorentz covariance is *manifest*: one glance at the equation is enough to check it.

Later, in GENERAL RELATIVITY, our goal will be to construct equations that are invariant under *arbitrary* (differentiable) coordinate transformations (not just global Lorentz transformations). Luckily, the formalism we introduce in this chapter is powerful enough to allow for the construction of such → *general covariant* equations as well. This is why we keep the formalism in this chapter as general as possible, and specialize it to SPECIAL RELATIVITY in the next Chapter 4. The discussion below is therefore already a preparation for GENERAL RELATIVITY; it is based on Schröder [1] and complemented by Carroll [56].

3.1. Manifolds, charts and coordinate transformations

1 | D-dimensional Manifold

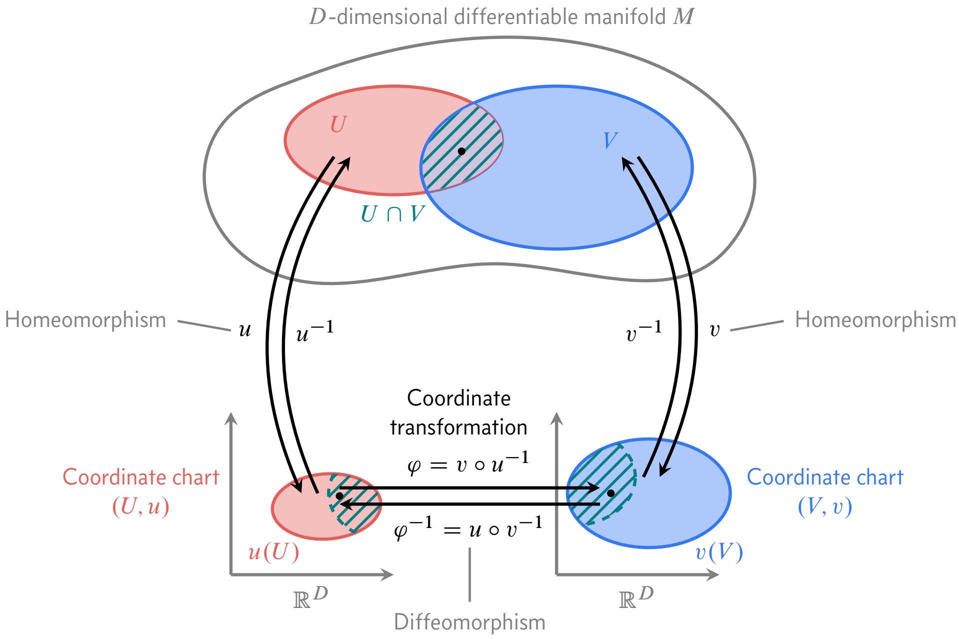
= Topological space that *locally* "looks like" D-dimensional Euclidean space $\mathbb{R}^D$:



- ¡! In RELATIVITY, the manifold of interest is the set of coincidence classes $\mathcal{E}$; it makes up the $D = 4$-dimensional manifold we call *spacetime*.
- A space that "locally looks like $\mathbb{R}^D$" is formalized as a ↑ *topological space* that is locally ↑ *homeomorphic* to Euclidean space $\mathbb{R}^D$. The structure defined in this way is then called a ↑ *topological manifold*.

2 | Differentiable Manifolds:

We want to formalize this idea and introduce additional structure to the manifold so that we can differentiate functions on it:



i | ⁂ *Coordinate system / Chart* (U, u):

$$u : \quad U \subseteq M \quad \to u(U) \subseteq \mathbb{R}^D \tag{3.1a}$$

$$u^{-1} : u(U) \subseteq \mathbb{R}^D \to \quad U \subseteq M \tag{3.1b}$$

$U \subseteq M$: open subset of M; u and u^{-1} are continuous and $u \circ u^{-1} = \mathbb{1}$.

$U = M$ is allowed. This is the situation we assumed so far in SPECIAL RELATIVITY: Our inertial coordinate systems cover all of spacetime $M = \mathcal{E}$.

ii | ⊲ Two charts (U, u) and (V, v) and let $U \cap V \neq 0$:

$$\varphi := v \circ u^{-1} : u(U \cap V) \to v(U \cap V) \tag{3.2a}$$

$$\varphi^{-1} := u \circ v^{-1} : v(U \cap V) \to u(U \cap V) \tag{3.2b}$$

φ: ⁂ *Coordinate transformation / Transition map*

$U = M = V$ and $U \cap V = M$ is allowed. This is the situation we assume so far in SPECIAL RELATIVITY where $(U = \mathcal{E}, u)$ and $(V = \mathcal{E}, v)$ correspond to the coordinate systems of two different inertial systems. The coordinate transformation φ would then be a Lorentz transformation (defined on $U \cap V = \mathcal{E}$).

iii | ⁂ *Atlas* := Family of charts $(U_i, u_i)_{i\in I}$ such that $M = \bigcup_{i\in I} U_i$

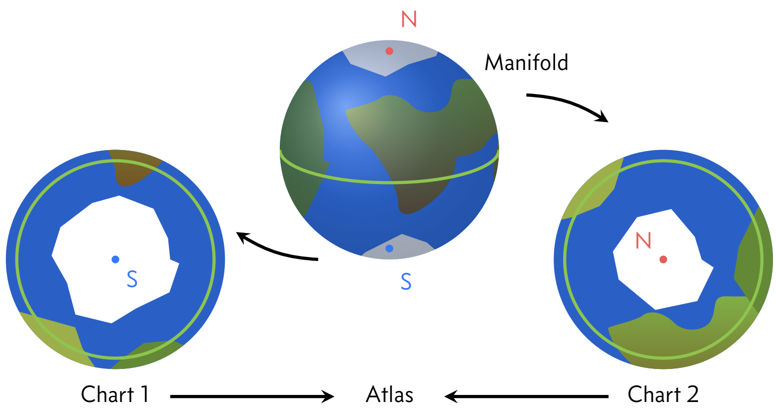
This definition of an atlas formalizes the notion of an atlas in real life (of the book variety): It contains many charts that, taken together, cover the complete manifold (typically earth). The different charts (on different pages of the book) all overlap on their edges such that you can draw any route on earth without gaps.

All φ, φ^{-1} differentiable → M: ⁂ *Differentiable Manifold*

- φ and φ^{-1} are maps from $\mathbb{R}^D$ to itself. It is therefore clear what "differentiable" means.
- In mathematics one is of course more precise about the degree of differentiability of the transition functions, and subsequently assigns this degree to the manifold. For example, if all coordinate transformations are infinitely often differentiable (= smooth), the manifold is called a ↑ *smooth manifold*. We are sloppy in this regard: For us all functions are differentiable as often as we need them to be.

In RELATIVITY we will only be concerned with *differentiable* manifolds.

3 | Example:



→ In general, a manifold *cannot* be covered by a single chart (Earth, mathematically S^2, needs at least two charts). In SPECIAL RELATIVITY this is not a problem: There we assume that spacetime is a flat (pseudo-)Euclidean space $\mathcal{E} \simeq \mathbb{R}^4$ and the coordinates given by our inertial systems cover all of spacetime. Later, in GENERAL RELATIVITY, this will not necessarily be the case.

3.2. Scalars

4 | ⁂ *Scalar (field)* := Function $\phi : M \to \mathbb{R}/\mathbb{C}$

- If ϕ maps to $\mathbb{R}$ ($\mathbb{C}$), we call ϕ a real (complex) scalar field.
- ¡! ϕ is a *geometric object* because it only depends on the manifold itself. It does not rely on charts/coordinates and does not depend on a particular set of charts you might choose to parametrize the manifold. The notion of a mathematical object to be "geometric in nature" or "independent of the choice of coordinates" is *absolutely crucial* for the understanding of GENERAL RELATIVITY. The reason why these "geometric objects" are so important for physics is the following insight that took physicists (including Einstein) a long time to fully

comprehend and implement mathematically:

> Coordinates (charts) do *not* represent physical entities.
> They are (useful) "mathematical auxiliary structures."

- One reason why it is so hard for us to grasp and implement the "physical irrelevance" of coordinates is, so I believe, that the first (and often only) coordinates we encounter in school are *Cartesian coordinates*. They are particularly intuitive because they are simply the *distances* of a point to some coordinate axes. Distances are a geometric property and physically relevant (you can measure them with rods); they are not the invention of mathematicians. This makes students draw the (wrong) conclusion that coordinates in general have intrinsic physical meaning. The problem is that coordinates *are* inventions of mathematicians; they do not share the ontological status of physical quantities like lengths etc. To undo this misconception is key to understand GENERAL RELATIVITY (→ *much later*).
- Since both M and $\mathbb{R}/\mathbb{C}$ are ↑ *topological spaces*, it makes sense to ask whether (or require that) ϕ is *continuous*. It does *not* make sense to ask whether ϕ is *differentiable* (and what is derivative is) because, in general, M does neither come with a notion of "distance" between two points in M nor can you add or subtract points (M does not have to be a ↓ *metric space* and/or a ↓ *linear space*). So an expression like $\partial_p \phi(p)$ does not make sense (→ *below*)!

5 | We just declared that coordinates are "not physical." The problem is that *without* coordinates it is really hard (at least for physicists) to do actual calculations with the geometric objects we are interested in (for example: compute derivatives). In addition, comparing theoretical predictions with experimental observations typically requires some sort of coordinate representation. Our ← *inertial systems*, for example, are elaborate measurement devices that produce a specific coordinate representation of the observed events.

This is why we always assume in the following that we have one (or more) charts that allow us to parametrize a (part of the) manifold, and then express the geometric quantities as functions of these coordinates. This means for the scalar field:

⊲ Two overlapping charts u and v:

$$\Phi(x) := \phi(u^{-1}(x)) \quad x \in u(U \cap V) \tag{3.3a}$$
$$\bar{\Phi}(\bar{x}) := \phi(v^{-1}(\bar{x})) \quad \bar{x} \in v(U \cap V) \tag{3.3b}$$

Φ and $\bar{\Phi}$ are functions on (subsets of) $\mathbb{R}^D$; in contrast to ϕ which is a function on the manifold M. In an abuse of notation, some authors do not make this distinction and write ϕ and $\bar{\phi}$ instead.

$\overset{\circ}{\rightarrow}$

$$\bar{\Phi}(\bar{x}) = \Phi(x) \quad \text{for} \quad \bar{x} = \varphi(x) \quad \text{with} \quad \varphi = v \circ u^{-1}\,. \tag{3.4}$$

Note that $\bar{\Phi}(\bar{x}) \overset{\text{def}}{=} \phi(p) \overset{\text{def}}{=} \Phi(x)$ with $u^{-1}(x) = p = v^{-1}(\bar{x})$.

- In RELATIVITY we typically work in a particular chart (coordinate system). Thus we write our fields as functions of coordinates (and not points on the manifold); e.g., when working with scalars, we typically work with Φ (and not ϕ).
- ¡! The special transformation of a field Eq. (3.4) (given as function of coordinates) tells us that it actually encodes a geometric, chart-independent function ϕ (given as function of

points on the manifold). This idea will be prevalent throughout this chapter and is the basis of our modern formulation of RELATIVITY: We work with functions that depend on specific coordinates (and therefore change when we transition to another chart); however, these functions satisfy certain transformation laws [like Eq. (3.4)] that guarantee that they actually encode geometric, chart-independent objects (which is what physics is about).

- As a function of coordinates, scalar fields are those fields the values of which do not change under coordinate transformations. A typical example would be the temperature as a function of position: When you move your coordinate system, the temperature of a particular point in space still is the same (only your coordinates of this particular point have changed!). This is exactly what Eq. (3.4) demands.

  Note that being a scalar (field) does not simply mean "being a number." The z-component of the electric field strength $E_z(x)$, for example, assigns a number to every point x; however, it does *not* transform like Eq. (3.4) under coordinate transformations. (Do you see why? What happens to E_z if you rotate your coordinate system?)

- In the literature, you will find the notation $\bar{\Phi} = \Phi$ to characterize scalars. This does *not* mean $\bar{\Phi}(x) = \Phi(x)$ for all $x \in \mathbb{R}^D$ (which characterizes *form-invariance* or *functional equivalence*), but rather $\bar{\Phi}(\bar{x}) = \Phi(x)$ (which characterizes scalar fields). Note that with $x = \varphi^{-1}(\bar{x})$ it follows $\bar{\Phi}(\bar{x}) = \Phi(\varphi^{-1}(\bar{x}))$ such that the function $\bar{\Phi}$ is typically *not* functionally equivalent to Φ. This ambiguity is the price we have to pay if we want to express geometric objects in terms of coordinates.

- Since $\Phi : \mathbb{R}^D \to \mathbb{R}$, it is well-defined what "differentiability" of Φ means. So expressions like $\frac{\partial \Phi(x)}{\partial x^k}$ make sense now (if Φ is differentiable). One then defines that ϕ is differentiable on M iff Φ is differentiable for all charts of an atlas of M.

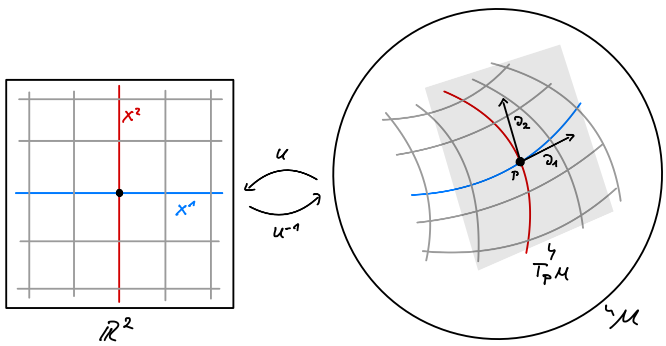
3.3. Covariant and contravariant vector fields

Are scalar fields the only geometric objects that can be defined on a manifold? The answer is *no*, there are many more! And these objects are not just toys for mathematicians: they are necessary to represent physical quantities like the electromagnetic field. Unfortunately, the definition of these quantities is not so straightforward as for scalars. We will not be mathematically precise in our discussion; however, it is important to understand the conceptual ideas:

6 | ⁂ *Tangent space $T_p M$ at $p \in M$*

= *Vector space* of directional derivative operators with evaluation at $p \in M$ (=derivations)

These operators can be applied to differentiable functions on the manifold (i.e., scalar fields).



- The tangent space T_pM is the mathematical formalization of the intuitive concept of the plane $\mathbb{R}^2$ that you can attach tangentially at any point p of a two-dimensional manifold. The problem with this picture is that it only works if you *embed* the manifold M into a higher-dimensional Euclidean space. Mathematically, such an approach is not satisfying because it presupposes additional structure to characterize the manifold (which, as it turns out, is not needed). Physically, the approach is also problematic: The manifold we are interested in is all of spacetime $\mathcal{E}$. But $\mathcal{E}$ is all there is, it is (to the best of our knowledge) not embedded into anything. It is therefore crucial that we can work with manifolds "stand alone", without assuming any embedding into a higher-dimensional space. The price we have to pay is that tangent vectors must be defined, rather abstractly, as directional derivative operators.
- There is a different tangent space T_pM at every point $p \in M$; these vector spaces all have the same dimension D (like the manifold) and are therefore all isomorphic. However, without additional structure, there is no natural connection (isomorphism) between these different vector spaces at different points. The disjoint union of all tangent spaces is called ↑ *tangent bundle* TM.
- Mathematically, the vectors in the tangent space can be defined as equivalence classes of smooth curves through p with the same derivative (with respect to their parametrization) at p. This equivalence class corresponds to a particular directional derivative that one can apply to smooth functions on the manifold at p. We do not need this abstract "bootstrapping procedure" for T_pM in the following.

$\sphericalangle$ Chart (U,u) with coordinates $x = (x^0, x^1, \dots, x^D)$

→ ⁂ *Coordinate basis* $\{\partial_i \equiv \frac{\partial}{\partial x^i}\}$ for T_pM

Recall that partial derivatives are special kinds of directional derivatives (namely in the direction where you keep all but one coordinate fixed). You can therefore think of ∂_i as the tangent vector at $p \in M$ that points into the x^i-direction mapped by u^{-1} onto the manifold.

7 | Since T_pM is a vector space for each point p of the manifold M, we can define *fields* on M that assign to each point p a tangent vector:

⁂ *Vector field*: $A(p) = \sum_{i=1}^{D} A^i(x)\partial_i$ with $x = u(p)$

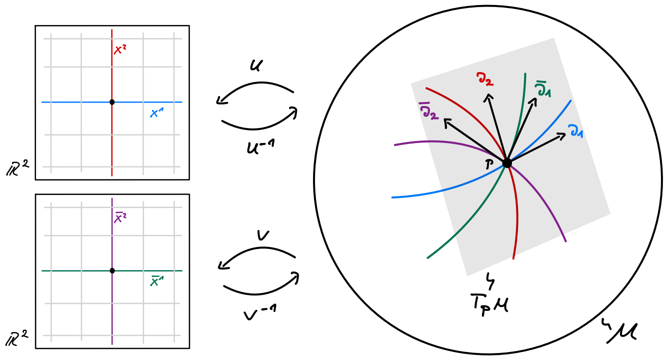
At every point $p \in M$ the vector field yields a tangent vector $A(p) = \sum_i A^i(u(p))\partial_i \in T_pM$.

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

→ Chain rule:

$$\underbrace{\frac{\partial}{\partial \bar{x}^i}}_{\bar{\partial}_i} = \sum_{k=1}^{D} \frac{\partial x^k}{\partial \bar{x}^i} \underbrace{\frac{\partial}{\partial x^k}}_{\partial_k} \tag{3.5}$$

→ For $x = u(p)$ and $\bar{x} = v(p)$ this is a *basis change* on the tangent space T_pM from one coordinate basis $\{\partial_i\}$ to another coordinate basis $\{\bar{\partial}_i\}$ via the (invertible) matrix $\frac{\partial x^k}{\partial \bar{x}^i}$:



9 | ∢ Vector field A and expand it in different coordinate bases:

$$\sum_i A^i(x)\partial_i = A(p) = \sum_i \bar{A}^i(\bar{x})\bar{\partial}_i \tag{3.6}$$

with $x = u(p)$ and $\bar{x} = v(p)$.

- ¡! The vector field A is a geometric object, just as the scalar field ϕ was. That it does *not* depend on the chosen chart is the statement of this equation.
- You learned this (with different notation and without the x/p-dependency) in your first course on linear algebra: Given a vector space V, a vector $\vec{v} \in V$, and a basis $\{\vec{e}_i\}$ with $V = \text{span}\{e_i\}$, you can encode the vector in a basis-dependent set of numbers v_i called *components* via linear combination: $\vec{v} = \sum_i v_i \vec{e}_i$. The *same* vector can be encoded by *different* components v'_i in a *different* basis $\{\vec{e}'_i\}$: $\vec{v} = \sum_i \vec{v}'_i \vec{e}'_i$. In our terminology, the vector $\vec{v}$ is a "geometric object" that does not depend on your choice of basis; only its components do. In this context, the gist of the story is that $\vec{v}$ represents something physical (like the velocity of a particle). The *components* v_i do so only indirectly because they depend on your choice of the basis $\{\vec{e}_i\}$ – and this choice does not bear any physical meaning.

Eq. (3.6) →

$$A = \sum_i A^i(x)\partial_i \stackrel{!}{=} \sum_i \bar{A}^i(\bar{x})\bar{\partial}_i \stackrel{\text{Eq. (3.5)}}{=} \sum_k \underbrace{\left[\sum_i \frac{\partial x^k}{\partial \bar{x}^i}\bar{A}^i(\bar{x})\right]}_{\stackrel{!}{=} A^k(x)} \partial_k \tag{3.7}$$

This motivates the following definition (we replace $x \leftrightarrow \bar{x}$ and the indices $i \leftrightarrow k$):

10 | ∢ D-tuple $\{A^i(x)\}$ of fields (in some chart with coordinates x):

$$⁂\ \textit{Contravariant vector field } \{A^i(x)\} \quad :\Leftrightarrow \quad \bar{A}^i(\bar{x}) = \sum_{k=1}^{D} \frac{\partial \bar{x}^i}{\partial x^k} A^k(x) \tag{3.8}$$

*Contra*variant vector (field) → *Super*script indices!

This is a *convention* which relates syntax and semantics and is at the heart of *tensor calculus*. The idea is that whenever you are given a collection of fields $A^i(x)$, you immediately know that they transform like Eq. (3.8) under coordinate transformations. (Unfortunately, there are exceptions to this rule, e.g., the → *Christoffel symbols*.)

- ¡! Not every D-tuple of fields transforms as Eq. (3.8). To deserve the name "contraviarant vector (field)," (and superscript indices) one has to check this transformation law explicitly!
- The rationale of Eq. (3.8) is the same as that of Eq. (3.4): Whenever we find a family of fields that transform under coordinate transformations as Eq. (3.8), we immediately know that together they encode a geometric, chart-independent object on the manifold that can be used to describe a physical quantity.

11 | (Counter)Examples:

- $\triangleleft$ Only *linear* coordinate transformations: $\bar{x} = \varphi(x) = \Lambda x$

  $\triangleleft$ Coordinate functions $X^i(x) := x^i$ as fields:

$$\underbrace{\bar{X}^i(\bar{x})}_{\bar{x}^i} = \sum_{k=1}^{D} \Lambda^i_k \underbrace{X^k(x)}_{x^k} = \sum_{k=1}^{D} \underbrace{\frac{\partial \bar{x}^i}{\partial x^k}}_{\Lambda^i_k} X^k(x) \tag{3.9}$$

  → Coordinate functions are contravariant vectors for linear transition maps.

  This is useful in SPECIAL RELATIVITY because there we only consider global Lorentz transformations (which are linear).

- $\triangleleft$ D *scalar* fields $\Phi^i(x)$ $(i = 1, \dots, D)$:

$$\text{For general } \bar{x} = \varphi(x): \quad \bar{\Phi}^i(\bar{x}) = \Phi^i(x) \neq \sum_{k=1}^{D} \underbrace{\frac{\partial \bar{x}^i}{\partial x^k}}_{\neq \delta^i_k} \Phi^k(x) \tag{3.10}$$

  → $\{\Phi^i(x)\}$ are *not* components of a contravariant vector field.

  - You see: not every collection of D fields is a vector!
  - ¡! δ^i_k is the Kronecker symbol: $\delta^i_k = 1$ for $i = k$ and $\delta^i_k = 0$ for $i \neq k$. The notation δ_{ik} is *not* used in tensor calculus (→ *later*).

12 | Reminder: ↓ *Dual spaces*

i | Remember: Linear algebra

Consider the vector space $V = \mathbb{R}^D$ and a column vector $\vec{v} = (v_1, \dots, v_D)^T \in V$ (a $1 \times D$-matrix). Let $\vec{w}^T = (w_1, \dots, w_D)$ be a row vector (a $D \times 1$-matrix). We can then perform a matrix multiplication between the vectors and interpret it as a linear map $\vec{w}^T$ acting on the vector $\vec{v}$ and producing a number:

$$\vec{w}^T : \vec{v} \in V \mapsto \vec{w}^T \cdot \vec{v} = (w_1 \quad \dots \quad w_D) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_D \end{pmatrix} = \sum_i w_i v_i \in \mathbb{R}\,. \tag{3.11}$$

In mathematical parlance $\vec{w}^T$ is a *linear functional* on the vector space V. All linear functionals of this form make up another vector space V^* called the ↓ *dual space* of V. You can think of V^* as the vector space of all D-dimensional *row* vectors and V as the vector space of all D-dimensional *column* vectors. The elements of the dual space are referred to as a ↓ *covectors*.

ii | Remember: Quantum mechanics

In quantum mechanics, the state of a physical system is described by ↓ *state vectors* in some Hilbert space $\mathcal{H}$ (which is a special kind of vector space). Vectors in this space are written as ↓ *kets*: $|\Psi\rangle \in \mathcal{H}$. You can produce a ↓ *bra* $\langle\Psi| = |\Psi\rangle^\dagger$ by applying the complex transpose operator. As in the example above, the bra $\langle\Psi|$ is a covector from the dual space $\mathcal{H}^*$; indeed, it acts as a linear functional on state vectors via the inner product of the Hilbert space:

$$\langle\Psi|(|\Phi\rangle) := \langle\Psi|\Phi\rangle \in \mathbb{C}\,. \tag{3.12}$$

This is the gist of the famous ↓ *Dirac bra-ket notation*.

iii | Hopefully these examples convinced you that the dual space is just as important and useful as the vector space itself.

→ Dual space of the tangent space T_pM?

Given a coordinate basis $\{\partial_i\} \in T_pM$ of a vector space, there is a standard way to define a basis of of the dual space T_p^*M:

↓ *Dual basis* $\{\mathrm{d}x^i\}$ with

$$\mathrm{d}x^i(\partial_j) := \delta^i_j = \frac{\partial x^i}{\partial x^j} \tag{3.13}$$

→ $\{\mathrm{d}x^i\}$ is a basis of the ⁂ *Cotangent space* T_p^*M

T_p^*M is the dual space of T_pM; it is common to write T_p^*M and not $(T_pM)^*$.

13 | Since T_p^*M is just another vector space for each point p of the manifold M, we can again define *fields* on M that map into this space:

⁂ *Covector field*: $B(p) = \sum_{i=1}^D B_i(x)\,\mathrm{d}x^i$ with $x = u(p)$

14 | Just like the coordinate basis, the dual coordinate basis depends on the chart and changes under coordinate transformations:

⊲ Coordinate transformation $\bar{x} = \varphi(x)$:

$$\mathrm{d}\bar{x}^i = \sum_{k=1}^D \frac{\partial \bar{x}^i}{\partial x^k}\,\mathrm{d}x^k \tag{3.14}$$

- Check that this is the correct transformation for the dual coordinate basis:

$$\begin{aligned}\mathrm{d}\bar{x}^i(\bar\partial_j) &= \left[\sum_k \frac{\partial \bar{x}^i}{\partial x^k}\,\mathrm{d}x^k\right]\left(\sum_l \frac{\partial x^l}{\partial \bar{x}^j}\partial_l\right)\\ &= \sum_{k,l}\frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^l}{\partial \bar{x}^j}\underbrace{\mathrm{d}x^k(\partial_l)}_{\delta^k_l} = \underbrace{\sum_k \frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^k}{\partial \bar{x}^j}}_{\frac{\partial \bar{x}^i}{\partial \bar{x}^j}} = \delta^i_j \quad ☺\end{aligned} \tag{3.15}$$

- You might recognize Eq. (3.14): This is simply the rule to compute the ↓ *total differential* of the function $\bar{x} = \varphi(x)$. This is no coincidence and explains why we use the differential notation $\mathrm{d}x^i$ for the dual vectors: The objects $\mathrm{d}x^i$ that we physicists like to illustrate as "infinitesimal shifts" in x^i are actually *linear functionals* (↑ *1-forms*).

15 | Now we can play the same game on T_p^*M as before on T_pM:

$\triangleleft$ Covector field B and expand it in different dual coordinate bases:

$$\sum_i B_i(x)\mathrm{d}x^i = B(p) = \sum_i \bar{B}_i(\bar{x})\mathrm{d}\bar{x}^i \tag{3.16}$$

with $x = u(p)$ and $\bar{x} = v(p)$.

¡! The covector field B is another geometric object, just as the vector field A was. That it does *not* depend on the chosen chart is the statement of this equation.

Eq. (3.16) →

$$B = \sum_i B_i(x)\,\mathrm{d}x^i \stackrel{!}{=} \sum_i \bar{B}_i(\bar{x})\,\mathrm{d}\bar{x}^i \stackrel{\text{Eq. (3.14)}}{=} \sum_k \underbrace{\left[\sum_i \frac{\partial \bar{x}^i}{\partial x^k}\bar{B}_i(\bar{x})\right]}_{\stackrel{!}{=} B_k(x)} \mathrm{d}x^k \tag{3.17}$$

This motivates the following definition (we replace $x \leftrightarrow \bar{x}$ and the indices $i \leftrightarrow k$):

16 | $\triangleleft$ D-tuple $\{B_i(x)\}$ of fields (in some chart with coordinates x):

$$\text{⁂ } \textit{Covariant vector field } \{B_i(x)\} \quad :\Leftrightarrow \quad \bar{B}_i(\bar{x}) = \sum_{k=1}^{D} \frac{\partial x^k}{\partial \bar{x}^i} B_k(x) \tag{3.18}$$

*Co*variant vector (field) → *Sub*script indices!

The rationale of Eq. (3.18) is the same as that of Eq. (3.8): Whenever we find a family of fields that transform under coordinate transformations as Eq. (3.18), we immediately know that together they encode a geometric, chart-independent object on the manifold that can be used to describe a physical quantity. To indicate that this object is a *covariant* vector field, we use *sub*script indices.

17 | Example:

First, let us introduce an even shorter notation for partial derivatives: $\Phi_{,i} \equiv \partial_i \Phi$

Following our index convention, the lower index in these expressions is only warranted *if* the field transforms as a covariant vector field according to Eq. (3.18). Let us check this:

$$\bar{\Phi}_{,i}(\bar{x}) = \bar{\partial}_i \bar{\Phi}(\bar{x}) \stackrel{\substack{\text{Eq. (3.4)}\\ \text{Eq. (3.5)}}}{=} \sum_{k=1}^{D} \frac{\partial x^k}{\partial \bar{x}^i} \frac{\partial \Phi(x)}{\partial x^k} = \sum_{k=1}^{D} \frac{\partial x^k}{\partial \bar{x}^i} \Phi_{,k}(x) \tag{3.19}$$

→ The gradient of a scalar is a covariant vector field.

18 | What happens if we apply a covector field on a vector field at each point $p \in M$?

$$\phi(p) := B(p)A(p) = \sum_{i,j} B_i(x)A^j(x)\underbrace{\mathrm{d}x^i(\partial_j)}_{\delta^i_j} = \sum_i A^i(x)B_i(x) =: \Phi(x) \tag{3.20}$$

→ $\Phi(x)$ must be a scalar!

This is a good point to introduce a new (and very convenient) notation:

⁂ *Einstein sum convention:*

$$\sum_{i=1}^{D} A^i(x)B_i(x) \equiv \underbrace{A^i(x)B_i(x)}_{\text{⁂ Einstein summation, ⁂ Contraction}} = A^l(x)B_l(x) \tag{3.21}$$

The *Einstein sum convention* or *Einstein summation* is a syntactic convention according to which a sum is automatically implied (but not written) whenever two indices show up twice in an expression and one is up (contravariant) and one down (covariant). Note that such indices are "dummy indices" in the sense that you can rename them to whatever you want (as long as you do not use the same letter for other indices already!). The sum over one co- and one contravariant index is called a *contraction*.

With this new notation it is straightforward to check that Φ transforms according to Eq. (3.4) by using the transformations Eq. (3.8) and Eq. (3.18):

$$\bar{\Phi}(\bar{x}) = \bar{A}^i(\bar{x})\bar{B}_i(\bar{x}) = \left[\frac{\partial \bar{x}^i}{\partial x^k}A^k(x)\right]\left[\frac{\partial x^l}{\partial \bar{x}^i}B_l(x)\right] \tag{3.22a}$$

$$= \underbrace{\frac{\partial \bar{x}^i}{\partial x^k}\frac{\partial x^l}{\partial \bar{x}^i}}_{\text{Chain rule} \to \delta^l_k} A^k(x)B_l(x) = A^l(x)B_l(x) = \Phi(x) \tag{3.22b}$$

The intermediate expression contains *three* sums over the colored indices (which we don't write)!

→ The contraction of a contra- and a covariant vector field yields a scalar field.

19 | Note on nomenclature:

- If you compare Eq. (3.18) with Eq. (3.5) you find that the *components* B_i of a covector field transform like the *basis vectors* ∂_i of the tangent space. We say the components *co*vary ("vary together") with the basis. This is why they are called *covariant*.
- A comparison of Eq. (3.8) and Eq. (3.14) shows that the components A^i of a vector field transform like the basis dx^i of the cotangent space – which is the *inverse* ("opposite") transformation as for the basis of the tangent space ∂_i. Thus we say the components A^i *contra*vary ("vary opposite to") the basis ∂_i. This is why they are called *contravariant*.

3.4. Higher-rank tensors

You learned in your linear algebra course that two vector spaces V and W can be used to construct a new vector space $V \otimes W$ called the ↓ *tensor product*. This allows us to generalize the notion of contra- and covariant *vector* fields to *tensor* fields, all of which are geometric, chart-independent objects defined on the manifold that are needed to describe physical quantities:

20 | An ⁂ *(absolute)* (p,q)*-tensor (field)* T of rank $r = p + q$

$$T^{i_1 i_2 \dots i_p}{}_{j_1 j_2 \dots j_q} \equiv T^{i_1 i_2 \dots i_p}{}_{j_1 j_2 \dots j_q}(x) \quad \text{or} \quad T^I{}_J \equiv T^I{}_J(x)\,, \tag{3.23}$$

with ↓ *multi-indices* $I = (i_1 \dots i_p)$ and $J = (j_1 \dots j_q)$, transforms like the tensor product of p contravariant and q covariant vector fields:

$$\underbrace{\bar{T}^{i_1\dots i_p}{}_{j_1\dots j_q}(\bar{x})}^{=\bar{T}^I{}_J(\bar{x})} = \underbrace{\left[\frac{\partial \bar{x}^{i_1}}{\partial x^{m_1}}\cdots\frac{\partial \bar{x}^{i_p}}{\partial x^{m_p}}\right]}_{=:\frac{\partial \bar{x}^I}{\partial x^M}}\underbrace{\left[\frac{\partial x^{n_1}}{\partial \bar{x}^{j_1}}\cdots\frac{\partial x^{n_q}}{\partial \bar{x}^{j_q}}\right]}_{=:\frac{\partial x^N}{\partial \bar{x}^J}}\underbrace{T^{m_1\dots m_p}{}_{n_1\dots n_q}(x)}_{=T^M{}_N(x)} \tag{3.24}$$

There are $r = p + q$ sums in this transformation rule (Einstein summation!).

- ¡! It is important that we do *not* write contra- and covariant indices above each other like so: T^i_j (at least not with additional knowledge about the tensor). This will become important below.
- Henceforth we always encode tensor fields by their chart-dependent *components*. The actual tensor field is of course chart-independent and maps each point $p \in M$ to an element of the tensor product

$$\underbrace{T_pM \otimes\cdots\otimes T_pM}_{p\text{ factors}} \otimes \underbrace{T_p^*M \otimes\cdots\otimes T_p^*M}_{q\text{ factors}}\,. \tag{3.25}$$

  like so

$$T(p) = \sum_{I,J} T^{i_1\dots i_p}{}_{j_1\dots j_q}(x)\,\partial_{i_1}\otimes\cdots\otimes\partial_{i_p}\otimes \mathrm{d}x^{j_1}\otimes\cdots\otimes\mathrm{d}x^{j_q}\,. \tag{3.26}$$

- Note that while tensors (more precisely: tensor components) are indicated by upper and lower indices (corresponding to their rank), not every object that is conventionally written with upper and lower indices does encode a tensor. For example, the transformation matrices $\frac{\partial \bar{x}^i}{\partial x^m}$, which describe a basis change on T_p^*M, do not encode a tensor field.

21 | Examples:

Scalar $\Phi(x)$	$\to$	$(0,0)$-tensor
Contravariant vector $A^i(x)$	$\to$	$(1,0)$-tensor
Covariant vector $B_i(x)$	$\to$	$(0,1)$-tensor
Tensor product $T^i{}_j(x) := A^i(x)B_j(x)$	$\to$	$(1,1)$-tensor (Check this!)

22 | Properties:

- *Equality:*

$$A = B \quad :\Leftrightarrow\quad \forall_{i_1\dots i_p}\forall_{j_1\dots j_q}:\ A^{i_1\dots i_p}{}_{j_1\dots j_q} = B^{i_1\dots i_p}{}_{j_1\dots j_q} \tag{3.27}$$

- *Symmetry:*

$$T\text{ (anti-)symmetric in } k \text{ and } l \quad :\Leftrightarrow\quad T^{\dots k\dots l\dots}{}_{\dots} = (-)\,T^{\dots l\dots k\dots}{}_{\dots} \tag{3.28}$$

  Every contra- or covariant rank-2 tensor can be decomposed into a sum of symmetric and antisymmetric tensors:

$$T_{ij} = \underbrace{\frac{1}{2}(T_{ij}+T_{ji})}_{=:T_{(ij)}} + \underbrace{\frac{1}{2}(T_{ij}-T_{ji})}_{=:T_{[ij]}} = T_{(ij)} + T_{[ij]}\,. \tag{3.29}$$

Institute for Theoretical Physics

23 | Constructing tensors:

New tensors can be constructed from known tensors as follows (Proofs: ➔ Problemset 4):

- *Sum of (p,q)-tensors A and B yields (p,q)-tensor C:*

$$C^{i_1\dots i_p}{}_{j_1\dots j_q} := A^{i_1\dots i_p}{}_{j_1\dots j_q} + B^{i_1\dots i_p}{}_{j_1\dots j_q} \tag{3.30a}$$

$$\text{or} \quad C^I{}_J := A^I{}_J + B^I{}_J \tag{3.30b}$$

- *Product of (p,q)-tensor A and scalar Φ yields (p,q)-tensor C:*

$$C^I{}_J := \Phi\, A^I{}_J \tag{3.31}$$

- *Tensor product of (p,q)-tensor A and (r,s)-tensor B yields $(p+r,q+s)$-tensor C:*

$$C^{IK}{}_{JL} := A^I{}_J \cdot B^K{}_L \tag{3.32}$$

- *Contractions:*

  Summing over a pair of contra- and covariant indices yields a tensor of rank $(p-1,q-1)$:

$$\tilde{A}^{i_1\dots\bullet\dots i_p}{}_{j_1\dots\bullet\dots j_q} := A^{i_1\dots k\dots i_p}{}_{j_1\dots k\dots j_q} \tag{3.33}$$

  The • indicates that the index summed over on the right side is missing in the list.

  Proof: ➔ Problemset 4

  A special case of a contraction (in combination with a tensor product) is the scalar obtained from a contra- and a covariant vector field above:

$$\Phi = C^i{}_i = A^i B_i\,. \tag{3.34}$$

- *Quotient theorem:*

$$\overset{\sqcap}{AB} = C \text{ tensor for all tensors } B \quad\Rightarrow\quad A \text{ is tensor} \tag{3.35}$$

  Here, $\overset{\sqcap}{AB}$ denotes (potentially multiple) contractions between indices of A and B (but not within A and B).

  - As an example, rewrite an arbitrary contravariant vector A^i as $A^i = \delta^i_j A^j$ with Kronecker symbol δ^i_j. The above theorem then implies that δ^i_j transforms as a $(1,1)$-tensor (verify this using the definition!). Hence we actually should write $\delta^i{}_j$ instead of δ^i_j. However, because the Kronecker symbol is symmetric in its indices, this simplified notation is allowed (→ *later*).
  - Special case:

$$A_{ik}B^k = C_i \text{ covector for all vectors } B^k \quad\Rightarrow\quad A_{ik} \text{ is } (0,2)\text{-tensor} \tag{3.36}$$

    Proof: ➔ Problemset 4

24 | Relative tensors:

i | Relative tensor are a generalization of the (absolute) tensors defined above. This generalization is useful because most of the rules for computing with tensors discussed so far carry over to relative tensors.

A ⁂ *relative tensor of weight* $w \in \mathbb{Z}$ picks up an additional power w of the ↓ *Jacobian determinant* under coordinate transformations:

$$\bar{R}^I{}_J(\bar{x}) = \det\left(\frac{\partial x}{\partial \bar{x}}\right)^w \frac{\partial \bar{x}^I}{\partial x^M}\frac{\partial x^N}{\partial \bar{x}^J} R^M{}_N(x) \quad \text{with weight } w \in \mathbb{Z} \tag{3.37}$$

and Jacobian determinant

$$\det\left(\frac{\partial x}{\partial \bar{x}}\right) := \sum_{\sigma \in S_D} (-1)^\sigma \prod_{i=1}^{D} \frac{\partial x^i}{\partial \bar{x}^{\sigma_i}}\,. \tag{3.38}$$

Here S_D is the group of permutations σ on D elements.

Since $\bar{x} = \varphi(x)$ is invertible, $x = \varphi^{-1}(\bar{x})$, it is $\frac{\partial \bar{x}}{\partial x} = \left(\frac{\partial x}{\partial \bar{x}}\right)^{-1}$ and therefore $\det\left(\frac{\partial \bar{x}}{\partial x}\right) = \det\left(\frac{\partial x}{\partial \bar{x}}\right)^{-1}$.

ii | Examples:

- *(Absolute) tensors* $\equiv$ Relative tensors of weight $w = 0$
- *Volume form*: Relative tensor of weight $w = -1$:

$$\mathrm{d}^D\bar{x} = \mathrm{d}^D x \det\left(\frac{\partial \bar{x}}{\partial x}\right) = \mathrm{d}^D x \det\left(\frac{\partial x}{\partial \bar{x}}\right)^{-1} \tag{3.39}$$

Remember the rule for integration by substitution with multiple variables!

- ⁂ *Tensor density* $\mathcal{L}(x) :=$ Relative tensor of weight $w = +1 \rightarrow$

$$S = \int \underbrace{\mathrm{d}^D x\,\mathcal{L}(x)}_{\text{Absolute tensor}} = \int \mathrm{d}^D\bar{x}\,\bar{\mathcal{L}}(\bar{x}) \tag{3.40}$$

In this example, we assume that $\mathcal{L}(x)$ is a *scalar* tensor density such that its integral is a (absolute) *scalar* quantity.

In ↑ *relativistic field theories* (like electrodynamics), the *Lagrangian density* $\mathcal{L}(x)$ is a scalar tensor density such that the ↓ *action* S becomes a scalar.

- Let $i_1, i_2, \dots, i_D \in \{1, 2, \dots, D\}$ and define the ⁂ *Levi-Civita symbol* as

$$\varepsilon^I \equiv \varepsilon^{i_1 i_2 \dots i_D} := \begin{cases} +1 & I \textit{ even} \text{ permutation of } 1, 2, \dots, D \\ -1 & I \textit{ odd} \text{ permutation of } 1, 2, \dots, D \\ 0 & \text{(at least) two indices equal} \end{cases} \tag{3.41}$$

An even (odd) permutation of $1, 2, \dots, D$ is constructed by an even (odd) number of *transpositions* (= exchanges of only two indices).

$\overset{\circ}{\rightarrow}$

$$\bar{\varepsilon}^I = \varepsilon^I \mathrel{\hat{=}} \det\left(\frac{\partial x}{\partial \bar{x}}\right)^{+1} \frac{\partial \bar{x}^I}{\partial x^J}\,\varepsilon^J \tag{3.42}$$

$\rightarrow \varepsilon^I = \varepsilon^{i_1 i_2 \dots i_D}$ is a $(D,0)$-tensor density

- ¡! $\bar{\varepsilon}^I = \varepsilon^I$ is true by definition: ε is a *symbol* defined by Eq. (3.41); this definition is independent of the coordinate system. In Eq. (3.42) we compare this trivial transformation with that of a (relative) tensor and conclude that it is equivalent to the statement that ε^I transforms as a $(D,0)$-tensor density with weight $w = +1$. This knowledge is helpful in tensor calculus to construct covariant expressions that contain Levi-Civita symbols (→ *below*).
- To show this, note that the Levi-Civita symbol can be used to compute determinants:

$$\det\left(\frac{\partial \bar{x}}{\partial x}\right) = \sum_{\sigma \in S_D} (-1)^\sigma \prod_{i=1}^{D} \frac{\partial \bar{x}^i}{\partial x^{\sigma_j}} = \frac{\partial \bar{x}^1}{\partial x^{j_1}} \cdots \frac{\partial \bar{x}^D}{\partial x^{j_D}}\,\varepsilon^{j_1 \dots j_D}\,. \tag{3.43}$$

Details: ➔ Problemset 4

↓ **Lecture 9** [10.12.25]

3.5. The metric tensor

A differentiable manifold M does not automatically allow us to measure the length of curves, the angles of intersecting lines, or the area/volume of subsets of the manifold; to do so, we need a *metric* on M (which is an additional piece of information). While the continuity structure (an atlas) that comes with M determines its *topology*, the metric determines its *geometry* (= shape). The same manifold M can be equipped with *different* metrics; this corresponds to different geometries of the same topology (a potato and an egg both have the topology of a sphere, nonetheless they are geometrically distinct).

A differentiable manifold together with a (pseudo-)metric is called ↑ *(pseudo-)Riemannian manifold*. In SPECIAL RELATIVITY and GENERAL RELATIVITY, spacetime is modeled by such (pseudo-)Riemannian manifolds where the metric is used to represent spatial and temporal distances between events.

25 | Motivation:

On linear spaces V, it is convenient to define an ↓ *inner product* (like in quantum mechanics where you consider Hilbert spaces and use their inner product to compute probabilities and transition amplitudes).

Recall the definition of a (real) inner product:

$$\langle \bullet | \bullet \rangle :\; V \times V \to \mathbb{R} \quad \text{with} \dots \tag{3.44a}$$

$$\text{Symmetry:} \quad \langle x|y\rangle = \langle y|x\rangle \tag{3.44b}$$

$$\text{(Bi)linearity:} \quad \langle ax + by|z\rangle = a\langle x|z\rangle + b\langle y|z\rangle \tag{3.44c}$$

$$\text{Positive-definiteness:} \quad x \neq 0 \;\Rightarrow\; \langle x|x\rangle > 0 \tag{3.44d}$$

Once you have an inner product, you get a norm, and subsequently a metric for free:

$$\underbrace{\langle x|y\rangle}_{\text{Inner product}} \;\Rightarrow\; \underbrace{\|x\| := \sqrt{\langle x|x\rangle}}_{\text{Norm}} \;\Rightarrow\; \underbrace{d(x,y) := \|x - y\|}_{\text{Metric}} \tag{3.45}$$