

Thus an inner product is a rather versatile structure and nice to have!

Problem: We cannot define an inner product on the manifold directly because M is not a linear space.

However: We can introduce an inner product on each of its tangent spaces $T_p M!$ →

26 | **** Riemannian (Pseudo-)Metric** $ds^2 :=$ Symmetric, non-degenerate (0, 2)-tensor field:

$$ds^2 : M \ni p \mapsto \underbrace{(ds_p^2 : T_p M \times T_p M \rightarrow \mathbb{R})}_{\text{Bilinear \& symmetric \& non-degenerate}} \quad (3.46a)$$

$$\begin{aligned} ds_p^2 \text{ bilinear} &\Rightarrow ds_p^2 \in T_p^* M \otimes T_p^* M \\ &\Rightarrow ds_p^2 = \sum_{i,j=1}^D g_{ij}(x) dx^i \otimes dx^j \equiv g_{ij}(x) dx^i dx^j \end{aligned} \quad (3.46b)$$

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

- The tensor product is *non-commutative*: $dx^i \otimes dx^j \neq dx^j \otimes dx^i$. However, you can always decompose a tensor product as

$$dx^i \otimes dx^j = \underbrace{\frac{1}{2}(dx^i \otimes dx^j + dx^j \otimes dx^i)}_{=: dx^i \vee dx^j} + \underbrace{\frac{1}{2}(dx^i \otimes dx^j - dx^j \otimes dx^i)}_{=: dx^i \wedge dx^j} \quad (3.47)$$

with the symmetrized tensor product $dx^i \vee dx^j$ and the anti-symmetrized tensor product $dx^i \wedge dx^j$ (↑ *wedge product*).

Since g_{ij} is assumed to be symmetric, only the symmetric component survives:

$$g_{ij}(x) dx^i \otimes dx^j = g_{ij}(x) dx^i \vee dx^j \equiv g_{ij}(x) dx^i dx^j \quad (3.48)$$

This means that when writing $dx^i dx^j$ in the above formula, you can be sloppy and either mean $dx^i \otimes dx^j$ or, equivalently, $dx^i \vee dx^j$. You will find both conventions in the literature. I will use $dx^i dx^j \equiv dx^i \vee dx^j$ so that $dx^i dx^j = dx^j dx^i$.

- It would be more appropriate to write $g = g_{ij} dx^i dx^j$ for the metric (0, 2)-tensor; it is conventional, however, to reserve g for the determinant $\det(g_{ij})$ so that we are stuck with ds^2 for the metric. Note that the d in ds^2 does *not* refer to an ↑ *exterior derivative*, it is purely symbolical.
- To define a proper ↓ *inner product* on $T_p M$, we should demand ↓ *positive-definiteness* instead of *non-degeneracy*. This, however, is often (for example in RELATIVITY) too restrictive; as it turns out, non-degeneracy is all we need for an isomorphism between $T_p M$ and $T_p^* M$ (“pulling indices up and down”, → *below*). This is why *negative* eigenvalues of g_{ij} are fine for many purposes, and motivates the concept of a → *signature*:

27 | Signature:

Since $g_{ij}(x) = g_{ji}(x)$ and $\det(g_{ij}(x)) \neq 0$

→ $g_{ij}(x)$ has r *positive* and s *negative* real eigenvalues for all $p \in M$

Since $\det(g_{ij}(x)) \neq 0$, these numbers must be the *same* for all $p \in M$.

→ (r, s) : **** Signature** of the metric ds^2

This classification does not depend on the coordinate basis (↑ *Sylvester’s law of inertia*).

- $(r > 0, s = 0)$

→ ds^2 : Riemannian metric → (M, ds^2) : $**$ Riemannian manifold

I.e., g_{ij} has only positive eigenvalues for all $p \in M$ and is therefore \downarrow positive-definite. This produces a true, positive-definite inner product on $T_p M$.

- $(r > 0, s > 0)$

→ ds^2 : pseudo-Riemannian metric → (M, ds^2) : $**$ pseudo-Riemannian manifold

I.e., g_{ij} has both positive and negative eigenvalues and is therefore \downarrow indefinite.

- $(r > 0, s = 1)$ or $(r = 1, s > 0)$:

→ ds^2 : Lorentzian metric → (M, ds^2) : $**$ Lorentzian manifold

In RELATIVITY we are only interested in metric tensors with one positive and three negative eigenvalues (equivalently: three positive and one negative eigenvalue). Mathematically speaking, spacetime is then a four-dimensional Lorentzian manifold and a special case of a pseudo-Riemannian manifold.

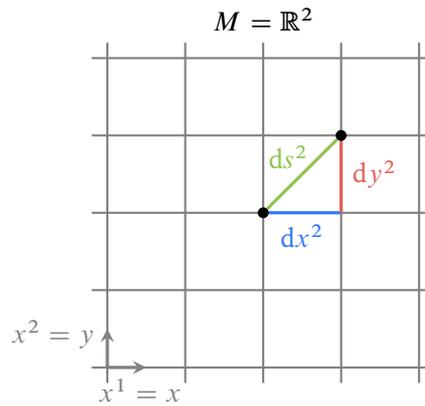
28 | Example: (Details: ➔ Problemset 4)

- i | $\sphericalangle D = 2$ Euclidean space $E_2 \equiv (\mathbb{R}^2, ds_E^2)$

The Euclidean metric in Cartesian coordinates $x^1 = x$ and $x^2 = y$ reads:

$$ds_E^2 := dx^2 + dy^2 = g_{ij}(x) dx^i dx^j \quad \text{with} \quad (g_{ij}) = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}}_{\substack{\text{Signature} \\ (2,0)}}. \quad (3.49)$$

This is consistent with the notion of dx and dy as infinitesimal shifts in coordinates and ds^2 as the infinitesimal distance (squared) that corresponds to this shift:



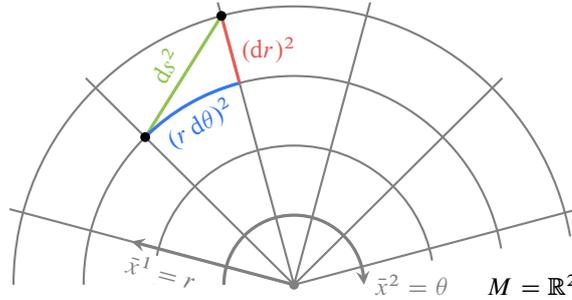
- ii | We can now transition to a new chart, namely polar coordinates $\bar{x}^1 = r$ and $\bar{x}^2 = \theta$. The induced basis change on the cotangent space is given by the total differential of the coordinate functions Eq. (3.14):

$$\varphi^{-1} : \begin{cases} x = r \cos(\theta) \\ y = r \sin(\theta) \end{cases} \xrightarrow{\text{Eq. (3.14)}} \begin{cases} dx = \cos(\theta) dr - r \sin(\theta) d\theta \\ dy = \sin(\theta) dr + r \cos(\theta) d\theta \end{cases} \quad (3.50)$$

iii | We find the components of the metric tensor field in the new basis $\{d\bar{x}^1 = dr, d\bar{x}^2 = d\theta\}$:

$$ds_{E_2}^2 \doteq dr^2 + r^2 d\theta^2 = \bar{g}_{ij}(\bar{x}) d\bar{x}^i d\bar{x}^j \quad \text{with} \quad (\bar{g}_{ij}) = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}}_{\substack{\text{Signature} \\ (2,0)}}. \quad (3.51)$$

This expression is again compatible with infinitesimal shifts in the (new) coordinates r and θ :



- The Euclidean plane E_2 is therefore an example for a Riemannian manifold with metric signature $(2, 0)$; its distinctive feature is that it is *flat*.
- Note that here we compute *the same* infinitesimal length in different coordinates (with the same result)! We did not change the *metric*, only the *coordinates* and thereby the coordinate basis in which we express the metric tensor. This is *flat* Euclidean space in \uparrow *curvilinear coordinates*. By contrast, later in GENERAL RELATIVITY we will study curved (non-flat, non-Euclidean) metric tensors, i.e., we will modify the geometry of space(time) itself.

29 | Since the metric ds^2 is a $(0, 2)$ -tensor field:

$$\bar{g}_{ij}(\bar{x}) d\bar{x}^i d\bar{x}^j = ds^2 = g_{ij}(x) dx^i dx^j \quad (3.52)$$

Eq. (3.14) $\overset{\circ}{\rightarrow}$

$$\bar{g}_{ij}(\bar{x}) = \frac{\partial x^l}{\partial \bar{x}^i} \frac{\partial x^m}{\partial \bar{x}^j} g_{lm}(x) \quad (3.53)$$

The metric (components) transforms as any other $(0, 2)$ tensor. Nothing special!

Side note:

Let $g := \det(g_{ij})$ and $\bar{g} := \det(\bar{g}_{ij}) \xrightarrow{\text{Eq. (3.53)}}$

$$\sqrt{|\bar{g}|} = \left| \det \left(\frac{\partial x}{\partial \bar{x}} \right) \right| \sqrt{|g|} \quad (3.54)$$

$\rightarrow \sqrt{|g|}$ is a *pseudo* scalar tensor density of weight $w = +1$. The “pseudo” indicates that the absolute value of the Jacobian determinant shows up, cf. Eq. (3.37).

$\llcorner g < 0 \xrightarrow{\text{Eq. (3.39)}} d^D x \sqrt{-g}$ is a scalar (\rightarrow later)!

30 | Length of curves on M :

One immediate benefit of having a Riemannian manifold is that we can now compute the length of curves $\gamma(t)$ on M (parametrized by $t \in [a, b]$ and given in some chart):

$$L[\gamma] \equiv \int_{\gamma} ds := \int_a^b \sqrt{g_{ij}(\gamma(t)) \frac{d\gamma^i(t)}{dt} \frac{d\gamma^j(t)}{dt}} dt \tag{3.55}$$

$$\equiv \int_a^b \|\dot{\gamma}(t)\|_{\gamma(t)} dt \tag{3.56}$$

! If ds^2 is a true *pseudo* metric (i.e., g_{ij} has at least one negative eigenvalue), one must make sure that the chosen curve γ does not produce negative values under the square root. In RELATIVITY these will be \uparrow *time-like* curves.

Example:

Let γ be the circle with radius R in the Euclidean plane E_2 . A possible parametrization in Cartesian coordinates (with origin in the center of the circle) is $\vec{\gamma}_{xy}(t) = (x_t, y_t) = (R \cos(t), R \sin(t))$ with $0 \leq t < 2\pi$ so that one finds for the circumference:

$$L = \int_{\gamma} \sqrt{dx^2 + dy^2} = \int_0^{2\pi} \sqrt{\dot{x}_t^2 + \dot{y}_t^2} dt \doteq 2\pi R \tag{3.57}$$

The same length can of course be calculated with the parametrization $\vec{\gamma}_{r\theta}(t) = (r_t, \theta_t) = (R, t)$ and $0 \leq t < 2\pi$ in polar coordinates:

$$L = \int_{\gamma} \sqrt{dr^2 + r^2 d\theta^2} = \int_0^{2\pi} \sqrt{\dot{r}_t^2 + r_t^2 \dot{\theta}_t^2} dt \doteq 2\pi R \tag{3.58}$$

Details: ➔ Problemset 4

- 31 | Besides computing lengths of curves (and other geometric quantities, → later), there is another benefit of having a metric tensor:

Pulling indices down:

$$\tilde{T} \begin{matrix} i_1 \dots \square \dots i_p \square \dots \square \\ \square \dots i \dots \square j_1 \dots j_q \end{matrix} := g_{ik} T \begin{matrix} i_1 \dots k \dots i_p \square \dots \square \\ \square \dots \square \dots \square j_1 \dots j_q \end{matrix} \tag{3.59}$$

→ \tilde{T} is a tensor of type $(p - 1, q + 1)$

- In Eq. (3.59) we indicate “empty” slots for indices by \square to emphasize that in each index “column” an index can either be *up* (contravariant) or *down* (covariant). It is conventional to omit the \square -markers. Note that this explains why you never should write two indices directly above each other (except for special cases, → below).

Furthermore, since g is fixed, it makes sense to label \tilde{T} again by T (note that the difference between the original tensor and the new one is manifest in the different index patterns!):

$$\tilde{T} \begin{matrix} i_1 \dots \square \dots i_p \square \dots \square \\ \square \dots i \dots \square j_1 \dots j_q \end{matrix} \mapsto T \begin{matrix} i_1 \dots \dots i_p \\ i \dots j_1 \dots j_q \end{matrix} \tag{3.60}$$

Example:

$$A^i \begin{matrix} k \\ j \end{matrix} \begin{matrix} l \end{matrix} := g_{jm} A^{imk} \tag{3.61}$$

- This convention matches perfectly with the computation of an inner product (which is determined by the metric tensor g) of two contravariant vectors:

$$\langle A, B \rangle \stackrel{\text{def}}{=} g_{ij} A^i B^j \stackrel{\text{def}}{=} \underbrace{A^i B_i}_{\text{Scalar}} \quad (3.62)$$

32 | Pulling indices up:

We would like to have a $(2, 0)$ -tensor g^{ij} with the property

$$\delta_j^k T^j = T^k \stackrel{!}{=} g^{ki} T_i \stackrel{\text{def}}{=} g^{ki} g_{ij} T^j. \quad (3.63)$$

g^{ij} allows us to revert the pulling-down of indices defined by the metric g_{ij} . Note that g^{ij} is a *different* tensor than g_{ij} , we could call it \tilde{g}^{ij} ; however, it is conventional to denote it with the same label due to the following close relationship with g :

$$g^{ki} g_{ij} \stackrel{!}{=} \delta_j^k \quad (3.64)$$

This is an implicit equation for g^{ki} !

→ g^{ij} is the *inverse matrix* of g_{ij}

(Which always exists because ds^2 is non-degenerate: $\det(g_{ij}) \neq 0$.)

→ In general:

$$\tilde{T}^{\square i_1 \dots i_p \square \dots j \dots \square}_{\square \dots \square j_1 \dots \square \dots j_q} := g^{jk} T^{\square i_1 \dots i_p \square \dots \square}_{\square \dots \square j_1 \dots k \dots j_q} \quad (3.65)$$

→ \tilde{T} is a tensor of type $(p + 1, q - 1)$

- Again we relabel \tilde{T} to T and omit the \square -markers:

$$\tilde{T}^{\square i_1 \dots i_p \square \dots j \dots \square}_{\square \dots \square j_1 \dots \square \dots j_q} \mapsto T^{i_1 \dots i_p j}_{j_1 \dots \dots j_q} \quad (3.66)$$

- Example:

$$A^{ijkl} := g^{lm} A^{ijk}_m \quad (3.67)$$

- With these new definitions, we can now raise and lower contractions:

$$A^i B_i = A^i \delta_i^j B_j = A^i g_{ik} g^{kj} B_j = A^i g_{ik} B^k = A_k B^k = A_i B^i \quad (3.68)$$

- What happens if you pull the indices of the Kronecker symbol down or up?

$$\delta_{ij} := g_{ik} \delta^k_j = g_{ij} \quad \text{and} \quad \delta^{ij} := g^{jk} \delta^i_k = g^{ij} \quad (3.69)$$

! $\delta_{ij} \equiv g_{ij}$ and $\delta^{ij} \equiv g^{ij}$ denote the metric and its inverse!

→ We never use the notation δ^{ij} and δ_{ij} to prevent confusion!

- Note that in general

$$g^{jk} T_k^i = T^{ij} \neq T^{ji} = g^{jk} T_k^j. \tag{3.70}$$

This means that the “column” in which the index is located is *important*, and notations like T_k^i are ill defined (if you pull k up by g^{jk} , do you get T^{ij} or T^{ji} ?). However, if the tensor is *symmetric*, $T^{ij} = T^{ji}$, this does not matter and you can get away with the sloppy notation T_k^i . This explains why writing δ_k^i for the Kronecker symbol is fine: $g^{ji} = g^{jk} \delta_k^i$ is symmetric.

33 | Mathematical side note:

“Pulling indices up and down” is mathematically the application of an \downarrow *isomorphism* between $T_p M$ and $T_p^* M$:

$$g(\bullet, \bullet) : T_p M \ni A \mapsto g(A, \bullet) \in T_p^* M \tag{3.71}$$

This has nothing to do with differential geometry or manifolds in particular; it is a general feature of non-degenerate bilinear forms on vector spaces. In differential geometry, this canonical isomorphism between the tangent bundle TM and the cotangent bundle T^*M is known as \uparrow *musical isomorphism*.

For example, you are using the same kind of isomorphism all the time in quantum mechanics, namely whenever you “dagger” a *ket* $|\Psi\rangle$ to obtain a *bra* $\langle\Psi|$:

$$(\bullet)^\dagger : \mathcal{H} \ni |\Psi\rangle \mapsto \langle\Psi| \equiv |\Psi\rangle^\dagger \in \mathcal{H}^* \quad \text{with} \quad \langle\Psi||\Phi\rangle \stackrel{!}{=} \langle\Psi|\Phi\rangle_{\mathcal{H}} \quad \text{for all } |\Phi\rangle \in \mathcal{H}. \tag{3.72}$$

Note how the *bra* $\langle\Psi|$ associated to the *ket* $|\Psi\rangle$ is *defined* via the inner product $\langle\bullet|\bullet\rangle_{\mathcal{H}}$ (and therefore metric) of the Hilbert space (\uparrow *Riesz representation theorem*!).

This leads to a nice dictionary between concepts in tensor calculus (and therefore RELATIVITY) and the bra-ket formalism of quantum mechanics:

	Relativity (fixed $p \in M$)	Quantum mechanics
Inner product space	$T_p M$	\mathcal{H}
Basis	$\{\partial_i\}$	$\{ i\rangle\}$
Vector	$A = A^i \partial_i$	$ \Psi\rangle = \Psi_i i\rangle$
Dual space	$T_p^* M$	\mathcal{H}^*
Dual basis	$\{dx^i\}$	$\{\langle i \}$
...	$dx^i(\partial_j) = \delta_j^i$	$\langle i j\rangle = \delta_{ij}$
Covector	$B = B_i dx^i$	$\langle\Psi = \Psi_i^* \langle i $
Inner product	$g(A_1, A_2) = g_{ij} A_1^i A_2^j$	$\langle\Psi \Phi\rangle$
Tensor	$A = A^{ij} \partial_i \otimes \partial_j$	$ \Psi\rangle \otimes \Phi\rangle \equiv \Psi\rangle \Phi\rangle$
...	$B = B_{ij} dx^i \otimes dx^j$	$\langle\Psi \otimes \langle\Phi \equiv \langle\Psi \langle\Phi $
Operator	$T = T^i_j \partial_i \otimes dx^j$	$ \Phi\rangle \otimes \langle\Psi \equiv \Phi\rangle\langle\Psi $
Trace	T^i_i	$\text{Tr}[\Phi\rangle\langle\Psi]$
Scalar	$BA = B_i A^i = g_{ij} B^i A^j$	$\langle\Psi \Phi\rangle = \langle\Psi \Phi\rangle$
Pulling indices down	$A_i = g_{ij} A^j$	$\langle\Psi = \Psi\rangle^\dagger$
Pulling indices up	$A^i = g^{ij} A_j$	$ \Psi\rangle = \langle\Psi ^\dagger$

3.6. Differentiation of tensor fields

- 34** | Remember: $\partial_i \Phi$ is covariant vector if Φ is scalar. However:

◁ Contravariant vector A^i :

$$\bar{A}^i{}_{,k} \equiv \frac{\partial \bar{A}^i}{\partial \bar{x}^k} = \frac{\partial x^m}{\partial \bar{x}^k} \frac{\partial}{\partial x^m} \left[\frac{\partial \bar{x}^i}{\partial x^l} A^l \right] = \underbrace{\frac{\partial^2 \bar{x}^i}{\partial x^m \partial x^l} \frac{\partial x^m}{\partial \bar{x}^k} A^l}_{\neq 0 \text{ (in general) } \ominus} + \underbrace{\frac{\partial x^m}{\partial \bar{x}^k} \frac{\partial \bar{x}^i}{\partial x^l} \frac{\partial A^l}{\partial x^m}}_{(1, 1)\text{-tensor } \ominus} \quad (3.73)$$

Here we used the transformation of \bar{A}^i [Eq. (3.8)] and $\bar{\partial}_k$ [Eq. (3.5)] and the product rule.

→ In general: $\frac{\partial \bar{A}^i}{\partial \bar{x}^k}$ is not a tensor!

35 | How to define a derivative of tensor fields that again transforms as a tensor?

To solve this problem, we first need a new field:

→ ** Christoffel symbols (of the second kind):

$$\Gamma^i{}_{kl} := \frac{1}{2} g^{im} (g_{mk,l} + g_{ml,k} - g_{kl,m}) \quad (3.74)$$

- The Christoffel symbols are symmetric in the lower two indices: $\Gamma^i{}_{kl} = \Gamma^i{}_{lk}$
- ¡! Despite the index notation, the Christoffel symbols are *not* tensors:

$$\bar{\Gamma}^i{}_{kl} \stackrel{\circ}{=} \frac{\partial \bar{x}^i}{\partial x^m} \frac{\partial x^n}{\partial \bar{x}^k} \frac{\partial x^p}{\partial \bar{x}^l} \Gamma^m{}_{np} - \underbrace{\frac{\partial x^n}{\partial \bar{x}^k} \frac{\partial x^p}{\partial \bar{x}^l} \frac{\partial^2 \bar{x}^i}{\partial x^n \partial x^p}}_{\text{No tensor!}} \quad (3.75)$$

This is why they are called “symbols” and not “tensors”!

- There are also Christoffel symbols of the *first* kind:

$$\Gamma_{ikl} := g_{ij} \Gamma^j{}_{kl} = \frac{1}{2} (g_{ik,l} + g_{il,k} - g_{kl,i}) \quad (3.76)$$

- Mathematically, the Christoffel symbols are the coefficients (in some basis) of the ↑ *Levi-Civita connection* which is determined by the metric tensor g^{ij} (→ later).

36 | ◁ Contravariant vector \bar{A}^i and contract it with $\bar{\Gamma}^i{}_{kl}$:

$$\bar{\Gamma}^i{}_{kl} \bar{A}^l = \underbrace{\frac{\partial \bar{x}^i}{\partial x^m} \frac{\partial x^n}{\partial \bar{x}^k} \Gamma^m{}_{np}}_{(1, 1)\text{-tensor } \ominus} \underbrace{\left[\frac{\partial x^p}{\partial \bar{x}^l} \bar{A}^l \right]}_{A^p} - \underbrace{\frac{\partial x^n}{\partial \bar{x}^k} \frac{\partial^2 \bar{x}^i}{\partial x^n \partial x^p} \left[\frac{\partial x^p}{\partial \bar{x}^l} \bar{A}^l \right]}_{\text{Problematic term in Eq. (3.73)}} \quad (3.77)$$

Idea: Add Eq. (3.73) and Eq. (3.77) to cancel the problematic term:

$$\bar{A}^i{}_{,k} + \bar{\Gamma}^i{}_{kp} \bar{A}^p = \underbrace{\frac{\partial x^m}{\partial \bar{x}^k} \frac{\partial \bar{x}^i}{\partial x^l} \left[A^l{}_{,m} + \Gamma^l{}_{mp} A^p \right]}_{(1, 1)\text{-tensor } \ominus \ominus} \quad (3.78)$$

37 | This motivates the definition of the ** Covariant derivative:

$$\text{Scalar: } \Phi_{;k} := \Phi_{,k} \quad (3.79a)$$

$$\text{Contravariant vector: } A^i{}_{;k} := A^i{}_{,k} + \Gamma^i{}_{kl} A^l \quad (3.79b)$$

$$\text{Covariant vector: } B_{i;k} := B_{i,k} - \Gamma^l{}_{ik} B_l \quad (3.79c)$$

- With this definition, $A^i_{;k}$ is a (1, 1)-tensor and $B_{i;k}$ is a (0, 2)-tensor!
- With this definition, the product rule is valid for the covariant derivative:

$$(A^i B_i)_{;k} = (A^i B_i)_{,k} \stackrel{\circ}{=} A^i_{;k} B_i + A^i B_{i;k} \quad (3.80)$$

- The construction of higher-rank tensors by tensoring contra- and covariant vectors Eq. (3.32) and the definitions of the covariant derivative above Eq. (3.79) can be used to construct covariant derivatives of arbitrary tensor fields. For example:

$$T^i_{k;l} := T^i_{k,l} + \Gamma^i_{ml} T^m_k - \Gamma^m_{kl} T^i_m \quad (3.81)$$

- With this generalization, we can apply the covariant derivative multiple times. For example:

$$A^i_{;k;l} \equiv (A^i_{;k})_{;l} \quad (3.82)$$

- The covariant derivative is *not commutative* in general:

$$A^i_{;k;l} - A^i_{;l;k} \neq 0 \quad (3.83)$$

→ Riemann curvature tensor → GENERAL RELATIVITY (→ later)

(This is not the case for the “normal” derivative: $A^i_{,k,l} = A^i_{,l,k}$.)

38 | Conclusion:

If you can formulate an equation that describes a physical theory in terms of tensors, it can always be brought into the form

$$T^I_J(x) = 0. \quad (3.84)$$

(This equation is meant to hold for all values of indices I and J and all coordinate values x .)

Here is an example:

The (inhomogeneous) Maxwell equations on an arbitrary (potentially curved) spacetime read:

$$\underbrace{F^{\mu\nu}_{;\nu} + \frac{4\pi}{c} J^\mu}_{=: T^\mu(x)} = 0 \quad (3.85)$$

with current density J^μ and field strength tensor $F^{\mu\nu} = g^{\mu\rho} g^{\nu\pi} (A_{\pi;\rho} - A_{\rho;\pi})$.

How does Eq. (3.84) look like in any other coordinate system $\bar{x} = \varphi(x)$?

Easy:

$$\bar{T}^I_J(\bar{x}) = \frac{\partial \bar{x}^I}{\partial x^M} \frac{\partial x^N}{\partial \bar{x}^J} \underbrace{T^M_N(x)}_{=0} = 0 \quad \Leftrightarrow \quad \bar{T}^I_J(\bar{x}) = 0. \quad (3.86)$$

This means:

Tensor equations are automatically form-invariant under *arbitrary* coordinate transformations; we say they exhibit $\star\star$ (*manifest*) *general covariance*.

The “manifest” means that checking general covariance is just a matter of checking whether the equation “looks right”, i.e., whether it is built from tensors following the rules discussed in this chapter. If a property of an equation is manifest, you don’t have to do calculations to verify it!

In the next chapter, we take a step back and specialize the allowed coordinate transformations to the Lorentz transformations of **SPECIAL RELATIVITY**. We can then use the form-invariance of equations built from “Lorentz tensors” to construct Lorentz covariant equations from scratch – which was our original goal!

4. Formulation on Minkowski Space

In this section we briefly reformulate what we already learned about **SPECIAL RELATIVITY** in terms of tensor calculus. We use this notation in subsequent chapters to make classical and quantum mechanics relativistic, and reformulate electrodynamics in a form where its Lorentz covariance is manifest. It also allows a smooth transition into **GENERAL RELATIVITY**.

The formulation of **SPECIAL RELATIVITY** on a unified, four-dimensional spacetime manifold goes back to Hermann Minkowski, Albert Einstein's former professor of mathematics at ETH. Minkowski writes in the notes of his lecture "Raum und Zeit" delivered 1908 in Cologne [57]:

Die Anschauungen über Raum und Zeit, die ich Ihnen entwickeln möchte, sind auf experimentell-physikalischem Boden erwachsen. Darin liegt ihre Stärke. Ihre Tendenz ist eine radikale. Von Stund' an sollen Raum für sich und Zeit für sich völlig zu Schatten herabsinken und nur noch eine Art Union der beiden soll Selbständigkeit bewahren.

Einstein, a physicist all through, didn't appreciate this mathematical reformulation of his theory at first. According to Sommerfeld, he (Einstein) commented:

Seit die Mathematiker über die Relativitätstheorie hergefallen sind, verstehe ich sie selbst nicht mehr.

Einstein later changed his views and acknowledged that without Minkowski's introduction of spacetime as a four-dimensional manifold, the development of **GENERAL RELATIVITY** would have been impossible.

For a historical account on the role of Minkowski, and his relationship (or absence thereof) to Einstein, see Ref. [58].

4.1. Minkowski space

1 | Manifold:

$$M = \langle \text{Spacetime of events / coincidence classes } \mathcal{E} \rangle \simeq \mathbb{R}^4 \quad (4.1)$$

It is a well-founded, but nonetheless empirical assumption that the spacetime manifold of events has the *topology* of \mathbb{R}^4 . Note that at this point we do not impose restrictions on the *geometry* of spacetime, e.g., whether it is flat or curved; this follows below when we settle on a metric.

2 | Charts:

In **SPECIAL RELATIVITY**, we restrict the coordinate systems to the ones that correspond to inertial observers / inertial coordinate systems:

$$(\mathcal{E}, K) \leftrightarrow \text{Inertial (coordinate) systems } K \in \mathcal{I} \quad (4.2)$$

Here, (\mathcal{E}, K) corresponds to a chart (U, u) as introduced in Section 3.1 where $U = M$ covers the entire manifold.

The coordinates are the ones obtained by an \uparrow *inertial observer*:

$$K : \mathcal{E} \ni E \mapsto K(E) := [E]_K = x \tag{4.3a}$$

$$\text{with } x^\mu = (x^0, x^1, x^2, x^3)^T = (ct, x, y, z)^T = (ct, \vec{x})^T \tag{4.3b}$$

- $\uparrow!$ Henceforth, *Greek* indices μ, ν, \dots run over 0, 1, 2, 3 where $\mu = 0$ denotes the time component and $\mu = 1, 2, 3$ denote the spatial components. *Roman* indices i, j, \dots run only over the spatial components 1, 2, 3.
- $\uparrow!$ We multiply the time t with the speed of light to measure times and distances in the same units.
- Since we assumed that our inertial systems cover all of spacetime, the domains on which the coordinate functions are defined are the complete manifold.
- The notation above is very suggestive: You can think of our inertial systems, namely the calibrated latticework of clocks and rods, as physical manifestations of the coordinate map of the corresponding chart. That is, an inertial system is a measurement device, or function, which assigns to every event $E \in \mathcal{E}$ the coordinate tuple $x = K(E) = (ct, \vec{x})_K \in E$.

3 | Transition maps:

- i | We worked hard in Section 1.4 to derive and select the correct coordinate transformations between different inertial systems. The most general ones have the form of ...

$$\left. \begin{array}{l} \text{Inhomogeneous Lorentz transformations} \\ \text{Poincaré transformations} \end{array} \right\} : \bar{x} = \varphi(x) = \Lambda x + a \tag{4.4}$$

with $a \in \mathbb{R}^4$ arbitrary and $\Lambda \in \mathbb{R}^{4 \times 4}$ a \uparrow *Lorentz transformation*.

For the special case $a = 0 \in \mathbb{R}^4$ we found:

$$\text{Homogeneous Lorentz transformations: } \bar{x} = \varphi(x) = \Lambda x \tag{4.5}$$

- ii | Since these transformations are affine, we find immediately:

$$\frac{\partial \bar{x}^\mu}{\partial x^\nu} = \Lambda^\mu_\nu \quad \text{and} \quad \frac{\partial x^\mu}{\partial \bar{x}^\nu} = (\Lambda^{-1})^\mu_\nu \equiv \Lambda_\nu^\mu \tag{4.6}$$

- Recall that the derivative of a linear (affine) map is simply the matrix which defines the map.
- $\uparrow!$ We use the tensor-inspired notation Λ^μ_ν for the matrix elements of Λ to allow for well-defined contractions with the metric (\rightarrow later). In Λ^μ_ν , the upper index μ denotes the *rows*, the lower index ν the *columns* of the matrix. The notation Λ_ν^μ for the components of the inverse transformation matrix Λ^{-1} is purely conventional at this point; it will turn out to be consistent with pulling indices up and down with the Minkowski metric (\rightarrow below).

- ¡! The Jacobi matrices of coordinate transformations are *not* tensor fields. (Remember that tensor fields are geometric objects that can be defined without referring to coordinates.) In particular, the components Λ^μ_ν of Lorentz transformations are *not* tensors, despite the index notation! [← *Recall* that Christoffel symbols are also written in index notation without being tensors, Eq. (3.75).]

→ This allows us to rewrite the coordinate transformation Eq. (4.4) in tensor notation:

$$\bar{x}^\mu = \Lambda^\mu_\nu x^\nu + a^\mu \tag{4.7}$$

¡! The matrix-vector product Λx is now given by the Einstein summation (index contraction) highlighted blue. We will stick to this notation whenever possible. Since we are now in the world of tensor calculus, it is strongly discouraged to think of and write rank-2 tensors as “matrices” and contractions as matrix-vector products Λx (even though Λ does not represent the components of a tensor). It is less error-prone (and simpler) to perform computations using the index notation introduced in Chapter 3.

- iii | Writing down the most general homogeneous Lorentz transformation is very complicated (and unnecessary). Here we provide the two special Lorentz transformations (boosts) discussed earlier in the new matrix notation, and an example for a spatial rotation about the z -axis:

- Lorentz boost in x -direction $K \xrightarrow{v_x} \bar{K}$ ($\beta_x = v_x/c$):

$$\text{Eq. (1.77)} \rightarrow \Lambda^\mu_\nu = [\Lambda_{v_x}]^\mu_\nu = \begin{pmatrix} \gamma & -\beta_x \gamma & 0 & 0 \\ -\beta_x \gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}_{\mu\nu}$$

(4.8)

- Lorentz boost in \hat{v} -direction $K \xrightarrow{\vec{v}} \bar{K}$ ($v = |\vec{v}|$ and $\tilde{\gamma} := \gamma - 1$):

$$\text{Eq. (1.75)} \rightarrow \Lambda^\mu_\nu = [\Lambda_{\vec{v}}]^\mu_\nu = \begin{pmatrix} \gamma & -\beta_x \gamma & -\beta_y \gamma & -\beta_z \gamma \\ -\beta_x \gamma & 1 + \tilde{\gamma} v_x^2 / v^2 & \tilde{\gamma} v_x v_y / v^2 & \tilde{\gamma} v_x v_z / v^2 \\ -\beta_y \gamma & \tilde{\gamma} v_x v_y / v^2 & 1 + \tilde{\gamma} v_y^2 / v^2 & \tilde{\gamma} v_y v_z / v^2 \\ -\beta_z \gamma & \tilde{\gamma} v_x v_z / v^2 & \tilde{\gamma} v_y v_z / v^2 & 1 + \tilde{\gamma} v_z^2 / v^2 \end{pmatrix}_{\mu\nu}$$

(4.9)