iii | → ※ *Proper time* accumulated by the spaceship clock along the trajectory $\mathcal{P}$:

$$\Delta\tau[\mathcal{P}] = \lim_{N\to\infty} \sum_{\substack{\text{Segment}\\i=1}}^{N} \Delta\tau_i = \int_{\mathcal{P}} d\tau = \int_{\mathcal{P}} \frac{ds}{c} := \int_{t_D}^{t_A} dt \sqrt{1 - \frac{\dot{\vec{x}}(t)^2}{c^2}} \qquad (2.25)$$

- As constructed, the proper time $\Delta\tau[\mathcal{P}]$ of a time-like trajectory $\mathcal{P}$, parametrized by $\vec{x}(t)$ for $t \in [t_0, t_1]$, is the time elapsed by a clock that follows this trajectory in spacetime.

- ¡! This result is valid for *accelerated* clocks.

  In general, SPECIAL RELATIVITY *can* described the physics of accelerated objects as long as the descpription of the process is given in an inertial coordinate system (as is the case here).

- ¡! The right-most expression in Eq. (2.25) yields the same result *in all inertial systems* K [recall Eq. (2.24)]. This is why $\tau[\mathcal{P}]$ is a function of the *event trajectory* $\mathcal{P}$ and not its coordinate parametrization $\vec{x}(t)$. This is important: It tells us that all inertial observers will agree on the reading of the spaceship clock τ_A at arrival A (although their *parametrization* $\vec{x}(t)$ may look different).

- Note that since $\vec{x}(t)$ is assumed to be time-like, it is $\forall_t : |\dot{\vec{x}}(t)| < c$ such that the radicand is always non-negative.

- $\tau[\bullet]$ is a *functional* of the trajectory $\mathcal{P}$; this is why we use square-brackets.

4 | Which trajectory $\mathcal{P}^*$ between the two events D and A *maximizes* the proper time $\Delta\tau$?
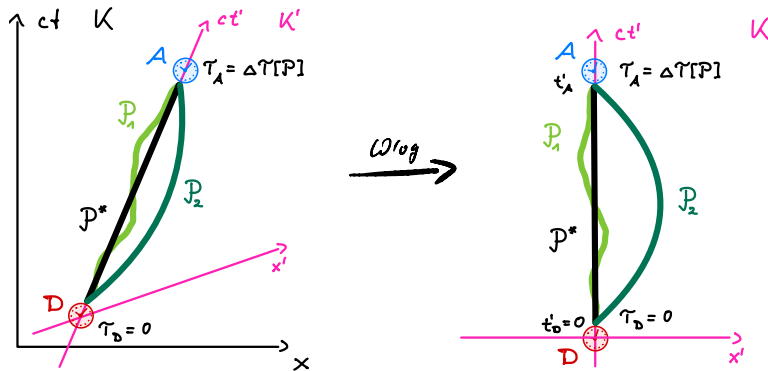
i | D and A are *time-like* separated → ∃ Inertial system $K' = K(D, A)$ with

$$[D]_{K'} = (t_D' = 0, \vec{x}_D' = \vec{0}) \quad \text{and} \quad [A]_{K'} = (t_A', \vec{x}_A' = \vec{0}) \qquad (2.26)$$

That is, without loss of generality, we can Lorentz transform into an inertial system where the two events happen at the *same* location (and by translations we can assume that this location is the origin $\vec{0}$ and that the coordinate time is $t_D' = 0$ at D). We label the time and space coordinate in K' by t' and $\vec{x}'$. Because of the relativity principle SR, K' is as good as any system to describe events.

ii | Time of an arbitrary path $\mathcal{P} \ni D, A$ with $[\mathcal{P}]_{K'} = (t', \vec{x}'(t'))$:

$$\Delta\tau[\mathcal{P}] = \int_{t_D'}^{t_A'} dt' \sqrt{1 - \frac{\dot{\vec{x}}'(t')^2}{c^2}} \overset{\vec{x}'(t')\equiv\vec{0}}{\leq} \int_{t_D'}^{t_A'} dt' = t_A' - t_D' = \Delta\tau[\mathcal{P}^*] \qquad (2.27)$$

Here $\mathcal{P}^*$ is the trajectory between D and A that is parametrized by the constant function $\vec{x}'(t') \equiv \vec{0}$ in K'. In other inertial systems, this trajectory will not be constant; however, it is inertial, i.e., $\mathcal{P}^*$ is described by a trajectory between D and A with uniform velocity.

Check this by applying a Lorentz transformation to the coordinates $(t', \vec{0})_{K'}$!

→ Clocks that travel along the *inertial trajectory $\mathcal{P}^*$* between D and A collect the largest proper time $\tau^* = \Delta\tau[\mathcal{P}^*]$.

Collecting the "largest time" means that the these clocks run the *fastest*.

5 | It is important to let this result sink in:

Let K' be the rest frame of earth (which is located in the origin $\vec{0}$) and consider two twins of age τ_D:

- **Twin S** departs with a **S**paceship at D, flies away from earth, turns around and returns to earth at A. **Twin S** therefore follows a trajectory similar to $\mathcal{P}_2$ in the sketches above.

- **Twin E** stays on **E**arth. He follows the inertial trajectory $\mathcal{P}^*$ in the sketches above.

We just proved above:

$$\langle\text{Age of } \textbf{Twin S} \text{ at } A\rangle = \Delta\tau[\mathcal{P}_2] + \tau_D < \Delta\tau[\mathcal{P}^*] + \tau_D = \langle\text{Age of } \textbf{Twin E} \text{ at } A\rangle$$

This is the famous ❄ *Twin "paradox"*: **Twin S** aged <u>less</u> than **Twin E**.

6 | Why there is *no* paradox:

- If you don't see why the above result should be paradoxical:

  Good! Move along. Nothing to see here! ☺

- Why one *could* conclude that the above result is paradoxical (= logically inconsistent):

  – From the view of **Twin E**, **Twin S** speeds around quickly, thus time-dilation tells him that **Twin S** should age slower. And indeed, when **Twin S** returns, he actually didn't age as much.

  – Now, you conclude, due to the relativity principle `SR`, we could also take the perspective of **Twin S** (i.e., our system of reference is now attached to the spaceship). Then **Twin S** would conclude that time-dilation makes **Twin E** (who now, together with earth, speeds around quickly) age more slowly. But this does not match up with the above result that, when both twins meet again at A, **Twin S** is the younger one! *Paradox!*

  The resultion is quite straightforwad:

  The invocation of the relativity principle `SR` in the last point is not admissible! Remember that `SR` only makes claims about the equivalence of *inertial systems*. Now have a look at the trajectory $\mathcal{P}_2$ of the spaceship again: it is clearly accelerated and cannot be inertial. And that there *is* at least a period where the spaceship (and **Twin S**) is accelerating is a *neccessity* for **Twin S** to *return* to **Twin E** (at least in flat spacetimes, but not so in curved ones [50])! This implies that the reunion of both twins at A requires at least one of them to *not* stay in an inertial system. This breaks the symmetry between the two twins and explains why the result can be (and is) asymmetric.

- ¡! For historical (and anthropocentric) reasons, the "twin paradox" is called a "paradox." We stick to this term because we have to – and not because it is appropriate name. The term "paradox" suggests an intrinsic inconsistency of RELATIVITY. As we explained above: *This is not the case.* All "paradoxes" in RELATIVITY are a consequence of unjustified, seemingly "intuitive" reasoning. The root cause is almost always an inappropriate, vague notion of "absolute simultaneity" that cannot be operationalized.

- An overview on different geometric approaches to rationalize the phenomenon can be found in Ref. [51].

  Below are two widely used spacetime diagrams of an idealized version where **Twin S** changes inertial systems only once from S_D to S_A halfway through the journey at R. You can think of this as an instantenous acceleration at the kink. Note, however, that the acceleration itself is dynamically irrelevant for the arguments; it is only important that the inertial frames in which **Twin S** departs and returns are not the same:
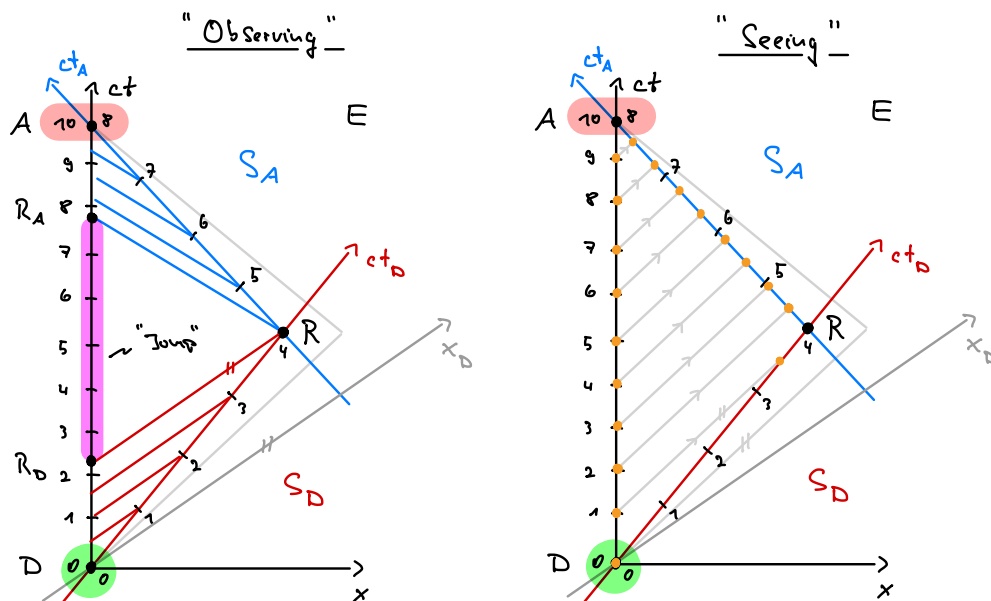
  – In the left diagram the *slices of simultaneity* in the two systems S_D and S_A are drawn. As predicted by time-dilation (and mandated by SR), **Twin S** observes the clocks of **Twin E** to run *slower* during his "inertial periods", i.e., while he stays in a single inertial system. However, the moment **Twin S** "jumps" from S_D to S_A at R, his notion of simultaneity changes instantaneously: In S_D, R and R_D are simultaneous; in S_A, however, R and R_A are simultaneous. Due to this jump, the record of **Twin S** contains now a temporal gap for events on earth (highlighted interval). It is this "missing" time interval that overcompensates the slower running clocks on earth (as observed from S_D and S_A) and makes **Twin S** conclude that **Twin E** ages faster (in agreement with the actual outcome of the experiment).

    If you wonder what happened to the (missing) observations of events in the triangle $R_A R R_D$: there is a nice explanation in Schutz [4]. (The bottom line is that **Twin S** constructs a bad coordinate system by stopping the recording of events in system S_D when he reaches R.)

  – In the right diagram, we draw *light signals* ("pings") of an earth-bound clock next to **Twin E** sent to **Twin S**. **Twin S** receives these signals and measures their period. This idealizes how **Twin S** *sees* (not *observes*!) the clocks ticking on earth (and, by proxy, how fast **Twin E** ages). It is important to understand the difference between this "seeing" and our operational definition of *observing* (using the contraption called an ← *inertial system*, as used in the left diagram). As demonstrated by the diagram, **Twin S** first sees the clock on earth ticking *slower*; but when he turns around at R, the clocks on earth (apparantly) speed up significantly. In the end, this speedup overcompensates for the slowdown during the first part of the journey so that **Twin S** again arrives at the (correct) conclusion that **Twin E** ages faster. Note that the speedup of the earth-bound clock *seen* by **Twin S** during the second half of his journey does not contradict time-dilation

because *seeing* is not *observing*. This is similar to the ↑ *Penrose-Terrell effect* in that a genuine relativistic effect (here: time-dilation) is distorted by an additional "imaging effect" due to the finite speed of light.

- In our careful derivation above, we not only showed that **Twin S** ages less than **Twin E**; we also showed that this conclusion is *independent* of the inertial observer! Thus we know that there will be *no dispute* about the different ages between different inertial observers.

- The Hafele-Keating experiment [46, 47] and the muon decay experiments [44], mentioned previously in the context of time-dilation, are experimental confirmations of the twin "paradox." So our theoretical prediction above (that **Twin S** ages less than **Twin E**) is experimentally confirmed. End of discussion.

- Our derivation of the accumulated proper time along trajectories in spacetime is both mathematically sound and experimentally confirmed. This qualifies SPECIAL RELATIVITY as a successfull theory of physics. *Operationally* there is nothing to complain about: the theory does its job to produce quantiative predictions of real phenomena. So why do so many people (physicists included) – despite the various efforts to visualize the phenomenon – have this nagging feeling of dissatisfaction that they cannot get rid of? The reason, so I would argue, is the human brain and its proclivity to inject concepts of absolute simulateneity into its model building. This qualifies the historical overemphasis of the twin "paradox" as a *meta problem*: The question to study is not how to "solve" the twin "paradox" (as we showed above, there is nothing to solve); the question to study is why so many peoply thought (and still think) that there is a problem in the first place. This *meta problem* is an actual problem to study; but it falls into the domain of cognitive science, and not physics!

7 | Two lessons to be learned from this:

> You can outlive your inertial-system-dwelling peers
> by changing inertial systems (= accelerating) at least once.

- ¡! You "live longer" when speeding around than your twin on earth, i.e., when you return, your twin might be 80 and have reached the end of his lifespan while you are still in your fourties. This is a real, observable effect, not an illusion of sorts. However, "living longer" does *not* mean that you somehow have "more time to spend" than your twin because all physical phenomena in your spaceship experience the same effect. It is not your metaboslism that slows down wrt. other physical phenomena around you, it is *time* itself. Put differently: If you and your twin both try to read as many books as possible during your lifetimes (say one per month), both of you will have read roughly the same amount of books when either of you dies (say at the age of 80).

- The mere fact that our universe *really* allows for this (at least in theory) makes it much more interesting than its boring alternative: a Galilean universe.

and

> Phenomena like length contraction and the twin "paradox" are physically *real*.
> Their "paradoxical" flavor is a phenomenon of human cognition, not physics.

This is why we put "paradox" always in quotes in the context of RELATIVITY.

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 [52].

3.1. Manifolds, charts and coordinate transformations

1 │ *D*-dimensional Manifold

= Topological space that *locally* "looks like" *D*-dimensional Euclidean space $\mathbb{R}^D$:
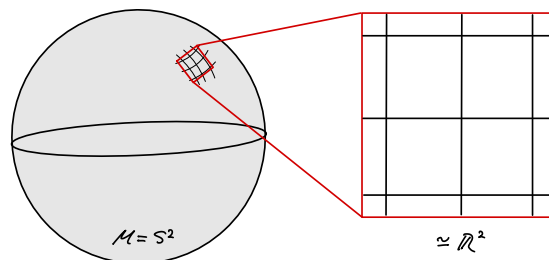
- ¡! In RELATIVITY, the manifold of interest is the set of coincidence classes $\mathcal{E}$; it makes up the $D = 4$-dimensional manifold we call *spacetime*.

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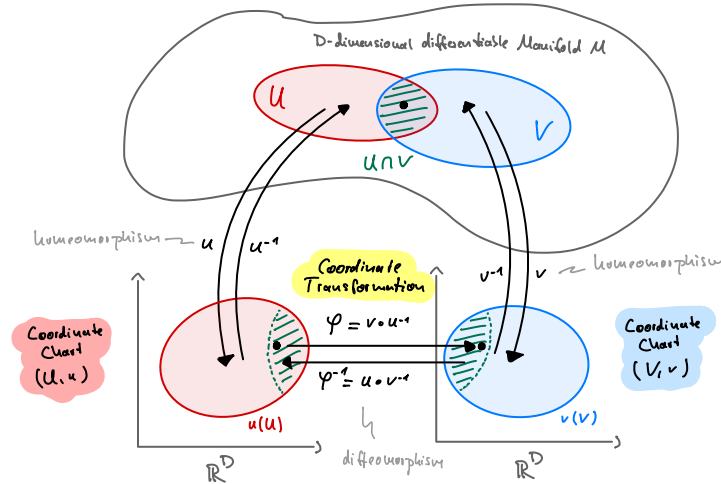
i | ⁂ *Coordinate system / Chart* (U, u):

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

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

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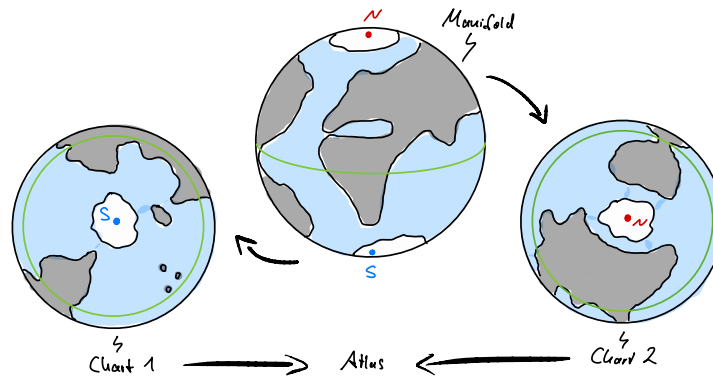
→ In general, a manifold *cannot* be covered by a single chart (Earth, mathematically S^2, needs at least two charts). In SPECIAL RELATIVITY this is not a problem: There we assume that spacetime is a flat (pseudo-)Euclidean space $\mathcal{E} \simeq \mathbb{R}^4$ and the coordinates given by our inertial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

3.2. Scalars

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

- 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.

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

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

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).
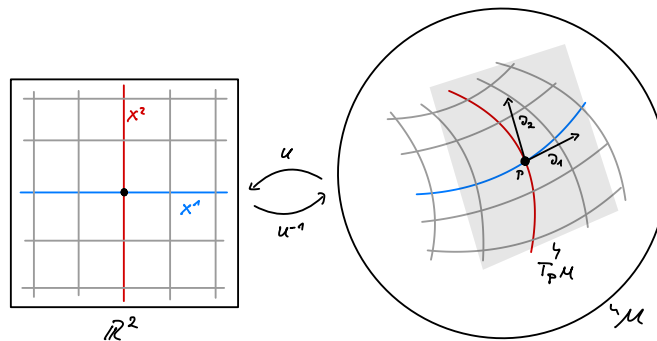
- The tangent space $T_p M$ is the mathematical formalization of the intuitive concept of the plane $\mathbb{R}^2$ that you can attach tangentially at any point p of a two-dimensional manifold. The problem with this picture is that it only works if you *embed* the manifold M into a higher-dimensional 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 $T_p M$ 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 $T_p M$ in the following.

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

→ ❊ *Coordinate basis* $\{\partial_i \equiv \frac{\partial}{\partial x^i}\}$ for $T_p M$

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.