↓Lecture 22 [28.05.24]

11. Classical physics on curved spacetime

Our mathematical toolbox is now fully equipped to formulate GENERAL RELATIVITY. In this chapter, we start by assuming a spacetime metric as given, and study how relativistic mechanics and electrodynamics can be formulated on this (curved) spacetime. Where the metric actually comes from will be discussed in the next Chapter 12.

11.1. Spacetime

1 | Setting the stage:

Here are some facts:

- We live in 3 spatial and 1 time dimension.

  For an argument why $3+1$-dimensional spacetimes are special, recall Section 4.4.

- The EEP requires the existence of ← *locally inertial coordinates* (← Section 10.3.1).

  Recall that in such coordinates the metric locally looks like the Minkowski metric.

→ Spacetime is a ← *4D Lorentzian manifold*:

> Spacetime ≡ 4D Lorentzian manifold (M, g)
> with pseudo-Riemannian metric g of signature $(1, 3)$

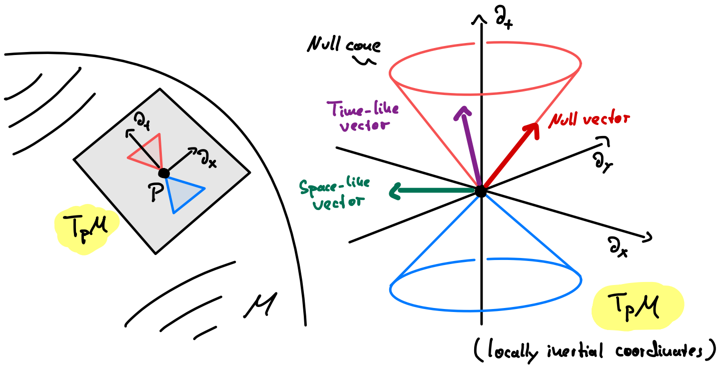
- Henceforth all manifolds are of this type. We indicate this by using Greek indices $\mu, \nu, \ldots = 0, 1, 2, 4$ for tensors; Latin indices $i, j, \ldots = 1, 2, 3$ are now reserved for the spatial components of tensors.
- With the metric g we can measure lengths of curves on the spacetime manifold and norms of and angles between vectors in the tangent bundle. There is also a lot of bonus structure: The metric defines a Levi-Civita connection, which, in turn, defines concepts like parallel transport, covariant derivatives, and curvature.
- Note that the *global topology* of M is not specified by GENERAL RELATIVITY, e.g., whether M is compact in all or some dimensions. For example, the universe could be periodic in one or more spatial dimensions, i.e., it could be a *torus*. While currently there are no observations that indicate a non-trivial topology, such topologies are also not conclusively ruled out and subject to ongoing research [139]. (Note that even assuming a completely flat universe – which is consistent with observations – does not rule out non-trivial topologies; recall the flat torus in Section 10.2.3.)

2 | Geodesics on Lorentzian manifolds:

In Section 10.3.3 we considered generic (pseudo-)Riemannian manifolds. We are now interested in $D = 4$-dimensional Lorentzian manifolds of signature $(1, 3)$ ("Spacetime"). This comes along with a few peculiarities concerning geodesics on this spacetime:

i | Null cones:

⊲ Tangent space $T_p M$ with basis $\{\partial_\mu\}$ induced by ← *locally inertial coordinates*:



→ The ⁂ *null cone* is the subset of tangent vectors $v = v^\mu \partial_\mu \in T_p M$ with

$$\|v\|_p^2 = \mathrm{d}s_p^2(v, v) = \underbrace{\eta_{\mu\nu} v^\mu v^\nu}_{\text{locally inertial coordinates}} \stackrel{!}{=} 0\,. \tag{11.1}$$

- That the null vectors of the Minkowski metric η form a cone was discussed in Section 1.6.
- Recall [← *Eq.* (4.16)] that all other vectors with strictly positive (negative) Minkowski norm are called ← *time-like* (← *space-like*). We adopt this nomenclature for vectors in the tangent spaces of Lorentzian manifolds.
- We call the cone "*null* cone" and not "*light* cone" because the latter term is reserved for a similar but distinct structure on the manifold (→ *below*).

→ A Lorentzian metric induces a "null cone texture" on the manifold (→ *below*).

This means that you can think of a Lorentzian manifold as being covered with little null cones that vary smoothly from point to point (not only their orientation, but also their "opening angle" can vary!). The null cones live in the tangent spaces and indicate which directions on the manifold are time-like, light-like (null), or space-like.

ii | Classification of geodesics:

⊲ Geodesic $\gamma^\mu(t)$ in an arbitrary coordinate system and parametrization

We can use the null cone structure to classify geodesics on a Lorentzian manifold. To this end, consider the sign of the norm (squared) of the "velocity vector" of a geodesic:

⊲ Sign of norm of tangent at geodesics:

$$\operatorname{sign} \|\dot\gamma(t)\|_{\gamma(t)}^2 = \operatorname{sign}\left[g_{\mu\nu}(\gamma(t))\dot\gamma(t)^\mu \dot\gamma(t)^\nu\right] \tag{11.2}$$

$\overset{\circ}{\rightarrow}$ Eq. (11.2) is …

- … independent of the coordinate system.

  The tangent vectors $\dot\gamma(t)^\mu$ contracted with the metric tensor yield a scalar.

- … constant along the geodesic.

  It is easy to check by straightforward calculation that the norm of the tangent vector is constant along a geodesic:

$$\frac{\mathrm{d}\|\dot\gamma\|_\gamma^2}{\mathrm{d}\lambda} = g_{\mu\nu,\sigma}\dot\gamma^\mu\dot\gamma^\nu\dot\gamma^\sigma + 2g_{\mu\nu}\dot\gamma^\mu\ddot\gamma^\nu \tag{11.3a}$$

$$= 2g_{\mu\nu,\sigma}\dot\gamma^\mu\dot\gamma^\nu\dot\gamma^\sigma - g_{\nu\sigma,\mu}\dot\gamma^\mu\dot\gamma^\nu\dot\gamma^\sigma + 2g_{\mu\nu}\dot\gamma^\mu\ddot\gamma^\nu \tag{11.3b}$$

$$= 2\dot\gamma^\mu\left(g_{\mu\nu,\sigma}\dot\gamma^\nu\dot\gamma^\sigma - \frac{1}{2}g_{\nu\sigma,\mu}\dot\gamma^\nu\dot\gamma^\sigma + g_{\mu\nu}\ddot\gamma^\nu\right) \tag{11.3c}$$

$$\overset{10.129}{=} 0 \tag{11.3d}$$

→ $\|\dot\gamma\|_\gamma^2 = \text{const}$ along a geodesic γ.

This of course immediately follows from our observation that geodesics are autoparallel curves, together with the metric-compatibility of the Levi-Civita connection.

- … invariant under reparametrizations.

  The independence of the sign on the parametrization of the curve is easy to show if one remembers that a reparametrization $\tilde\gamma(\tau) = \gamma(t)$ is given by a *strictly monotone* function $\tau = \tau(t)$:

$$\operatorname{sign}\left[\frac{\mathrm{d}\gamma(t)_\mu}{\mathrm{d}t}\frac{\mathrm{d}\gamma(t)^\mu}{\mathrm{d}t}\right] = \operatorname{sign}\Big[\frac{\mathrm{d}\tilde\gamma(\tau)_\mu}{\mathrm{d}\tau}\frac{\mathrm{d}\tilde\gamma(\tau)^\mu}{\mathrm{d}\tau}\underbrace{\left(\frac{\mathrm{d}\tau}{\mathrm{d}t}\right)^2}_{>0}\Big] = \operatorname{sign}\Big[\frac{\mathrm{d}\tilde\gamma(\tau)_\mu}{\mathrm{d}\tau}\frac{\mathrm{d}\tilde\gamma(\tau)^\mu}{\mathrm{d}\tau}\Big]. \tag{11.4}$$

  Note that the norm of the "velocity vector" itself (without the sign) *does* depend on the parametrization! This makes sense if you think of the parameter as time: Changing how you measure time of course changes how you measure velocity.

→ $\operatorname{sign}\|\dot\gamma\|_\gamma^2$ characterizes geodesics:

$$\left.\begin{matrix}\gamma \text{ time-like}\\ \gamma \text{ light-like (or } \textit{null}\text{)}\\ \gamma \text{ space-like}\end{matrix}\right\} \quad :\Leftrightarrow \quad \operatorname{sign}\|\dot\gamma\|_\gamma^2 = \begin{cases}+1\\ 0\\ -1\end{cases} \tag{11.5}$$

Hence there are three types of geodesics on a Lorentzian manifold.

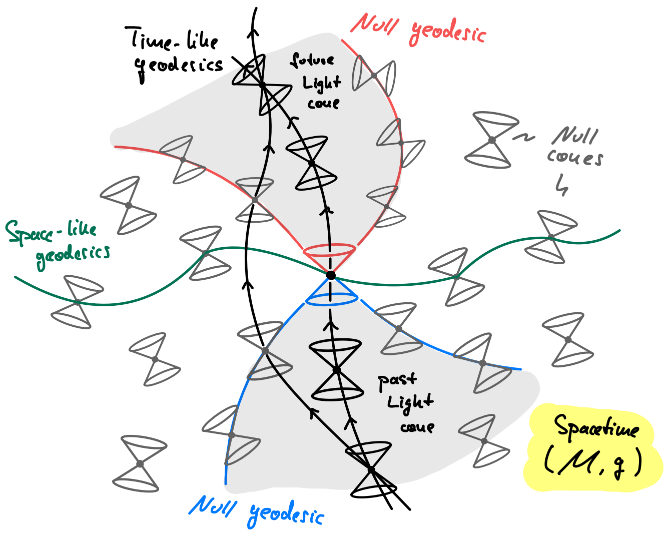
- We adopt the same nomenclature also for spacetime curves that are *not* geodesics. In this case, the claim that the sign is constant along the curve is not (necessarily) the consequence of some dynamical law, but simply a feature of a particular curve.
- On the $D = 4$-dimensional spacetime of GENERAL RELATIVITY, the time-like geodesics correspond to possible trajectories of free-falling bodies (also: possible time axes). The light-like geodesics are the trajectories of, well, light rays. Space-like geodesics are the analog of "straight lines" in space.
- There is a subtlety regarding light-like/null geodesics: Since their "velocity" vanishes (by definition), their length Eq. (10.123) vanishes as well. As a consequence, we cannot use their length s as an affine parameter λ. To see what goes wrong, note that for $\chi(y) = \sqrt{y}$ setting $y = \|\dot\gamma\|_\gamma^2 = 0$ in Eq. (10.128) is undefined (division by zero).

  Luckily, this is only a technical inconvenience. Recall that in our setting, the equations for autoparallel curves Eq. (10.60) and geodesics Eq. (10.131) are identical. While the *norm* $\|\dot\gamma\|_\gamma^2$ of a null vector vanishes, the *vector* itself $\dot\gamma^i$ is a perfectly normal vector in the tangent space (courtesy of $g_{\mu\nu}$ being a *pseudo*-metric). We then can simply fall back to the autoparallel equation Eq. (10.60) to describe null geodesics. The only difference is then that the affine parameter of light-like solutions of Eq. (10.60) (or, equivalently, Eq. (10.131)) cannot be interpreted as the length along the geodesic anymore.

iii | Light cones:

$\triangleleft$ Point/Event $E \in M$; Draw all null geodesics emanating from E

→ ⁂ *Light cone* of E:



Notes:

- Null geodesics remain null everywhere, i.e., their tangent vectors at every point lie on the null cone of the corresponding tangent space. Since the metric is Lorentzian (but otherwise arbitrary) the null cones can point in "different directions" at different points, so that the light cone can be warped and deformed.

  In summary: The *null* cones live in the tangent spaces attached to the manifold, the *light* cone lives on the manifold itself and warps according to the local null cones (and thereby the metric).

- Note how all null cones on the future light cone point "inward", whereas all null cones on the past light cone point "outward". They act like unidirectional "pores" in a membrane that allow time-like trajectories (not necessarily geodesics) to *leave* the past light cone and *enter* the future light cone (but not to other way around).

- All time-like geodesics through E stay within its past- and future light cone. Conversely, all space-like geodesics remain outside of this light cone.

  Note that because of curvature in the metric [← Eq. (10.140)], geodesics can "attract" each other; in particular, two time-like geodesics emanating from a common event might cross again at another event! (Example: Imagine two satellites orbiting earth on the same orbit in opposite directions. Both are falling freely and – according to GENERAL RELATIVITY– follow geodesics in spacetime. But they periodically meet each other, i.e., their geodesics cross in spacetime repeatedly.)

- Not every time-/light-/space-like *curve* is a time-/light-/space-like *geodesic*! Here is an example of a completely light-like curve on Minkowski space (in inertial coordinates):

$$\gamma^\mu(\lambda) = (\lambda, \cos(\lambda), \sin(\lambda), 0)\ . \tag{11.6}$$

Indeed:

$$\|\dot{\gamma}\|_\gamma^2 = \eta_{\mu\nu}\dot{\gamma}^\mu\dot{\gamma}^\nu = 1 - \left[\sin^2(\lambda) + \cos^2(\lambda)\right] = 0\,. \tag{11.7}$$

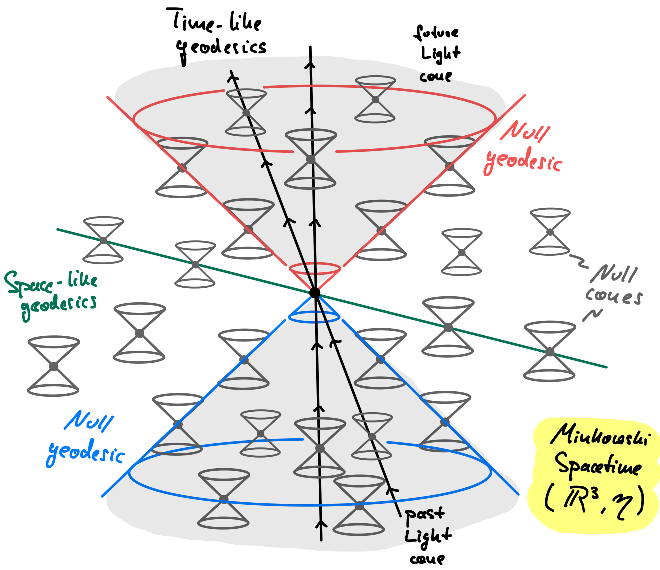
On Minkowski space, all geodesics are straight lines in inertial coordinates (because the Christoffel symbols vanish in them); the helical curve above clearly isn't a straight line, i.e., it is no geodesic but still *null* everywhere.

- The null cone texture (also called a ↑ *cone field*) induces a ← *partial order of events*, which encodes a ← *causality structure* on the spacetime manifold (recall Section 1.6 for the case of Minkowski space). Up to a local (conformal) deformation of time- and length scales, this structure is essentially *equivalent* to the Lorentzian metric [140]! This suggests the intriguing possibility that the null cone texture (equivalently: the causal structure of events) might be the truly fundamental field of GENERAL RELATIVITY, and the Lorentzian metric is just a convenient tool to encode it.

  (Note that a local "stretching" of the metric by a strictly positive scalar field, $\Omega(x)g_{\mu\nu}(x)$, does neither alter the null cone texture nor angles between tangent vectors, thus it is a ↑ *conformal transformation*. This is why one says that the null cone texture determines the ↑ *conformal class* of the Lorentzian metric.)

  For more details on Lorentzian manifolds, null cones, light cones, and the causal structure of spacetime, see the monograph [141].

- By comparison, in flat Minkowski space all geodesics are straight lines and never cross twice:



  Note how the *null cone* (which lives in the tangent space) of the reference event coincides with its *light cone* (which lives on the manifold). Mathematically, Minkowski space $\mathbb{R}^{1,3}$ is not just a Riemannian manifold (with Minkowski metric η) but also an ↓ *affine space*; this allows for a natural embedding of its tangent spaces into the manifold itself. Minkowski space is therefore a rather "degenerate" case of a generic spacetime and is not well suited to carve out the essential features of GENERAL RELATIVITY.

- Remember that there are locally inertial coordinates for every *point* of the manifold where (1) the Christoffel symbols vanish and (2) the metric has the Minkowski form (Section 10.3.1). This concept can be generalized:

  For every *geodesic*, there is a coordinate system (defined in a "tube" around the geodesic) such that on the geodesic, the metric takes the Minkowski form and the Christoffel symbols vanish (and so do the first derivatives of the metric). Such coordinates are called ↑ *Fermi normal coordinates* [142] and are useful for a freely falling observer to describe physics along (and close to) its time-like geodesics (which is then the time-axis of these coordinates).

iv | Extremal properties of geodesics:

We defined geodesics by a variational principle Eq. (10.125). Hence they *extremize* their Riemannian length *locally*. Since null geodesics have vanishing length, we focus here on time-like and space-like geodesics.

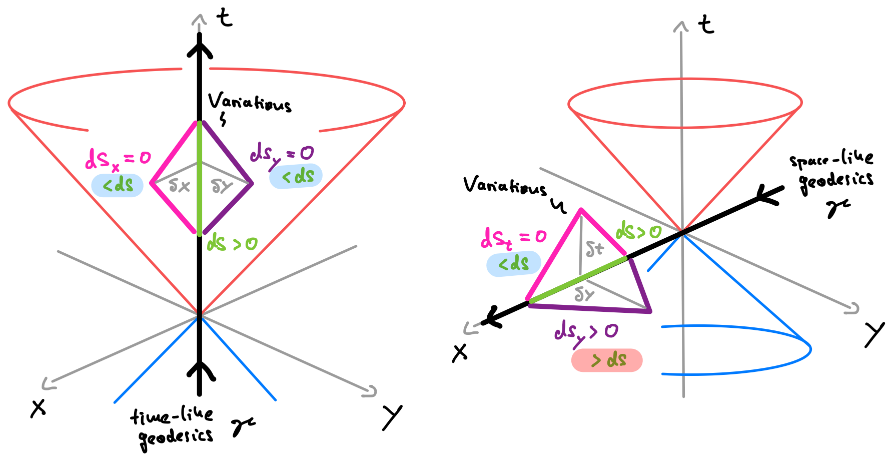
⊲ Length of time/space-like geodesics γ:

$$\text{Proper time:} \quad L_{\text{Time}}[\gamma] = \int_\gamma \underbrace{\sqrt{+g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu}}_{\sim c^2\mathrm{d}t^2-(\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2)>0} \tag{11.8a}$$

$$\text{Proper distance:} \quad L_{\text{Space}}[\gamma] = \int_\gamma \underbrace{\sqrt{-g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu}}_{\sim(\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2)-c^2\mathrm{d}t^2>0} \tag{11.8b}$$

Recall that the metric has signature $(1,3) = (+,-,-,-)$. The expressions below the integrals are valid approximately in locally inertial coordinates.

⊲ Local variations (in locally inertial coordinates; geodesic *w.l.o.g.* along coordinate axis):



→

> *Time-like* geodesics are local *maxima* of *proper time*.
> *Space-like* geodesics are local *saddle points* of *proper distance*.

3 | Proper time:

Which quantitiy corresponds to the time interval $\Delta\tau$ (*proper time*) measured by an ideal clock in GENERAL RELATIVITY?

Requirements:

- Correspondence principle:

  GENERAL RELATIVITY must reduce to SPECIAL RELATIVITY if the spacetime manifold is flat Minkowski space: $(M, g) = (\mathbb{R}^4, \eta) = \mathbb{R}^{1,3}$.

Remember that we established in Section 2.4 that the time measured by a clock moving along an arbitrary *time-like* trajectory $\gamma^\mu : [\lambda_a, \lambda_b] \to \mathbb{R}^{1,3}$ in Minkowski space is given by [Eqs. (2.25) and (4.14)]

$$\Delta\tau[\gamma] = \frac{1}{c}\int_{\lambda_a}^{\lambda_b} \mathrm{d}\lambda \sqrt{\eta_{\mu\nu}\dot{\gamma}^\mu\dot{\gamma}^\nu} \,. \tag{11.9}$$

This expression is valid in global inertial coordinate systems.

- General covariance:

  Following GRP, the expression for $\Delta\tau$ must be a geometric property of the trajectory γ that depends on the metric of the spacetime manifold, but not on the chosen coordinates and/or parametrization of the curve.

These conditions suggest the following definition of the propert time in GENERAL RELATIVITY:

⊲ Clock following *time-like* trajectory $\gamma : [\lambda_a, \lambda_b] \to M$ on arbitrary spacetime (M, g):

¡! γ is not required to be a geodesic.

→ ⁂ *Proper time* measured by this clock:

$$\Delta\tau[\gamma] := \frac{1}{c}\int_{\lambda_a}^{\lambda_b} \mathrm{d}\lambda \sqrt{g_{\mu\nu}(\gamma(\lambda))\dot{\gamma}^\mu\dot{\gamma}^\nu} \equiv \frac{1}{c}\int_\gamma \sqrt{g_{\mu\nu}(x)\mathrm{d}x^\mu\mathrm{d}x^\nu} \tag{11.10}$$

- Note that, because γ is a *time-like* curve by assumption, the expression under the squareroot is always strictly positive.
- That Eq. (11.10) is the correct expression for the reading of ideal clocks following arbitrary time-like trajectories on arbitrary spacetimes is reasonable, but it is not a "mathematical necessity" – it is a *prediction* of GENERAL RELATIVITY that can be experimentally assessed by its physical implications (→ *later*).

  This is actually a rather subtle point: What *is* an ideal clock? The only sensible thing to do is to *declare* any dynamic physical process that counts time according to Eq. (11.10) as an ideal clock. That ideal clocks measure Eq. (11.10) becomes then a tautology and physically vacuos. That physical systems *exist* that (up to some limiting acceleration) measure Eq. (11.10) as predicted by GENERAL RELATIVITY is not, however. Atomic clocks, for instance, turn out to be rather good and robust approximations of ideal clocks, whereas pendulum clocks are very sensitive to accelerations and quickly deviate from Eq. (11.10). This deviation, however, is not a feature of time itself, but a consequence of the particular dynamical law governing the motion of a pendulum under accelerated motion. Conversely, to verify that atomic clocks do not suffer from such effects, and therefore are good proxies for measuring proper time, one can check whether their readings match the predictions of GENERAL RELATIVITY for $\Delta\tau$ in various situations, e.g., in the presence of gravitational fields (→ *later*). There are also more direct, operational procedures (using light rays and freely falling test particles) to assess how closely a physical process resembles an ideal clock [143] (→ *below*).
- Here is an analogy to demystify clocks: The voltage U_{ab} between two points $\vec{r}_a$ and $\vec{r}_b$ is given by the line integral of the electric field $\vec{E}(\vec{r})$:

$$U_{ab} = -\int_{\vec{r}_a}^{\vec{r}_b} \vec{E} \cdot \mathrm{d}\vec{r} \,. \tag{11.11}$$

A ↓ *voltmeter* is a measurement device that exploits electrodynamic processes to measure U_{ab}, and thereby a particular property of the electromagnetic field. Voltmeters are no magical devices that, by decree, always measure the quantity Eq. (11.11) (this would be an *ideal* voltmeter, which, unfortunately, you cannot buy). A "good" voltmeter is a device that exploits a physical process such that its output correlates with Eq. (11.11) for a wide range of voltages; however, if you exceed its voltage ratings, this is no longer true and the readings are no longer reliable.

Similarly, clocks are measurement devices that exploit some physical process to produce outputs that correlate with the quantity Eq. (11.10), *and thereby measure a property of the metric field* $g_{\mu\nu}(x)$. An ideal clock does so for all curves γ in all conceivable metric fields $g_{\mu\nu}$; a "good" clock (like an atomic clock) does so *approximately* under a wide variety of circumstances, while a "bad" clock (like a pendulum clock) has only a very narrow range of applicability (e.g., unaccelerated trajectories).

- According to our discussion above, time-like geodesics correspond to trajectories of clocks along which they run *fastest*. This generalizes our discussion of the twin "paradox" in Section 2.4, where we concluded that the twin staying home (in an inertial system, we ignore the gravitational effects of Earth) ages quicker than the one following an accelerated trajectory with his rocket. In our new reading, the earth-bound twin follows a *geodesic* in Minkowksi space; by contrast, the rocket-twin follows a *non-geodesic* time-like curve in Minkowski space.
- In specific coordinate systems, the integral Eq. (11.10) can look simpler.

  For example, one can always choose a coordinate system $\hat{x}^\mu$ with the clock fixed in the origin $\vec{0}$ (recall the discussion in Section 9.2 about the role of coordinates in GENERAL RELATIVITY). In such coordinates, the proper time integral simplifies to

$$\Delta\tau[\gamma] = \frac{1}{c}\int_{x_a^0}^{x_b^0} \mathrm{d}\hat{x}^0\,\sqrt{\hat{g}_{00}(\hat{x}^0,\vec{0})} \equiv \int_{\tau_a}^{\tau_b}\mathrm{d}\tau\,, \tag{11.12}$$

  so that the proper time interval is given by

$$\mathrm{d}\tau = \frac{1}{c}\sqrt{\hat{g}_{00}}\,\mathrm{d}\hat{x}^0 \quad \text{with} \quad \hat{g}_{00} > 0\,. \tag{11.13}$$

  But why stop there? Nothing prevents you from locally "stretching" and "squeezing" the time coordinate $\mathrm{d}\tilde{x}^0 := \sqrt{\hat{g}_{00}}\mathrm{d}\hat{x}^0$ to absorb the time-dependence of the metric so that

$$\mathrm{d}\tau = \frac{1}{c}\mathrm{d}\tilde{x}^0\,, \tag{11.14}$$

  and thereby

$$\Delta\tau[\gamma] = \frac{1}{c}\int_{\tilde{x}_a^0}^{\tilde{x}_b^0}\mathrm{d}\tilde{x}^0\,. \tag{11.15}$$

  Such coordinates can be systematically constructed (↑ *proper reference frames*) for clocks on arbitrary time-like trajectories ("observers"); see MISNER *et al.* [2] (§13.6, pp. 327–332) for details.

  Note that the evaluation of $\Delta\tau$ is not simplified by Eq. (11.15) in general because you must know the integral boundries $\tilde{x}^0_{a,b}$ in these coordinates (which is tantamount to knowing $\Delta\tau$).

4 | <u>Radar coordinates:</u>

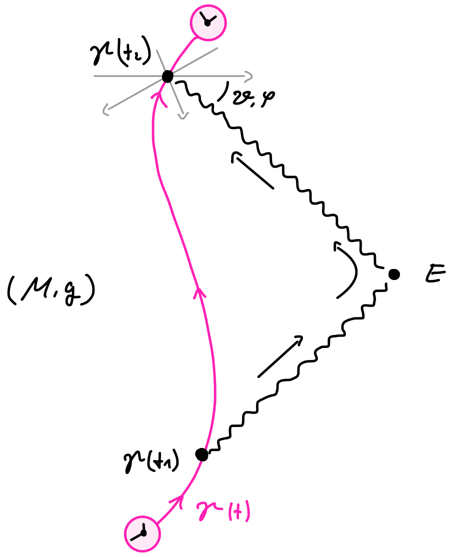
In GENERAL RELATIVITY, coordinates are mathematical artifices that are used to catalog events, while preserving their local causal relations ("continuity"). In contrast to the inertial coordinate systems of SPECIAL RELATIVITY, there is *no operational meaning* associated to most coordinates! To find special coordinate charts that have a physical interpretation (at least in some region of spacetime) one can proceed the other way around: Define an operational procedure that assigns four numbers to events in spacetime; this procedure then defines a particular kind of coordinate system that can be identified with measurable quantities by construction. A particularly simple example of such coordinates are *radar coordinates*:

This discussion is based on Ref. [144].

i | ⊲ Observer ≡ Clock following a time-like trajectory $\gamma(t) : [a, b] \to M$

¡! The trajectory does not need to be a geodesic. The clock displays t along γ – but the parameter t is not required to be an affine parameter (in particular: proper time). If t does equal proper time Eq. (11.10) along γ (up to some offset), we call the clock an ← *ideal clock*.

ii | ⊲ Event $E \in M$
⊲ Light signals emitted at $\gamma(t_1)$ & reflected at E & received at $\gamma(t_2)$:



→ ⁂ *Radar coordinates* (T, R, θ, φ):

$$T := \tfrac{1}{2}(t_2 + t_1) \qquad \text{⁂ Radar time} \tag{11.16a}$$

$$R := \tfrac{c}{2}(t_2 - t_1) \qquad \text{⁂ Radar distance} \tag{11.16b}$$

$$\theta := \langle \downarrow \textit{Altitude} \text{ of reflection at } \gamma(t_2)\rangle \tag{11.16c}$$

$$\varphi := \langle \downarrow \textit{Azimuth} \text{ of reflection at } \gamma(t_2)\rangle \tag{11.16d}$$

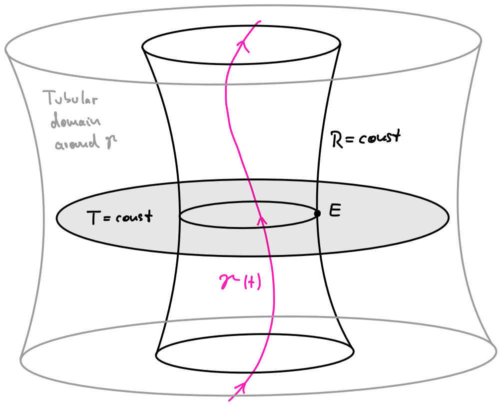
To define the altitude and azimuth, one must fix a smooth orthonormal ↑ *tetrad* along γ.

Note that one can really measure (T, R, θ, φ): Think of E as a point on the trajectory of a space probe flying away from Earth. You can periodically send directional radar pulses – that are reflected by the space probe – and use your Earth-bound clock to measure t_1 and t_2 (together with the angles θ and φ of the received reflection).

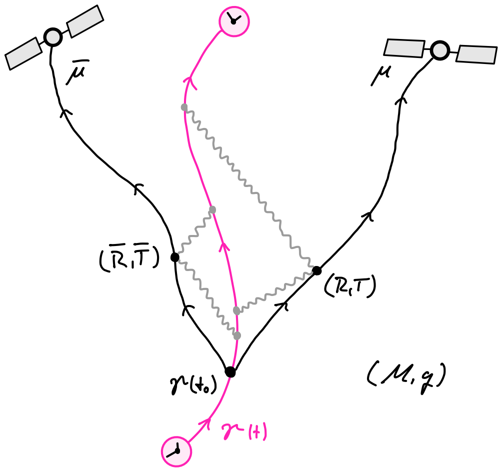
iii | Radar coordinates *cannot* cover all of spacetime in general!

The method can fail to assign coordinates to events E that are "shadowed" by other objects, or because γ and E are separated by an event horizon (e.g., a ↑ *Rindler horizon*). It can also happen that the assignment is not unique if there are different null geodesics from $\gamma(t_1)$ to E and/or from E to $\gamma(t_2)$; this can happen due to spacetime curvature (→ *gravitational* *lensing*). However, one can show that there is always a finite "tube" around γ where these peculiarities can be excluded.

→ Radar coordinates cover a "tube" around γ:



iv | Now that we can construct radar coordinates in the vicinity of a clock γ, we can perform the following experiment to check whether this clock is an ← *ideal clock* [i.e., the parameter t measures proper time Eq. (11.10)] [143]:



a | Eject two free falling space probes along trajectories μ and $\bar{\mu}$ at t_0.

b | Track their trajectories with radar pulses →

$$\mu = (R(T), T) \quad \text{and} \quad \bar{\mu} = (\bar{R}(\bar{T}), \bar{T}) \tag{11.17}$$

(We omit the polar coordinates.)

c | γ is an ideal clock at t_0 iff $\|\dot{\gamma}(t_0)\|_{\gamma(t_0)} = c$ since [← Eq. (11.10)]

$$d\tau \equiv \frac{1}{c}ds = \frac{1}{c}\sqrt{g_{\mu\nu}dx^\mu dx^\nu} = \frac{1}{c}\sqrt{g_{\mu\nu}\dot{\gamma}^\mu\dot{\gamma}^\nu}dt = \frac{1}{c}\underbrace{\|\dot{\gamma}\|_\gamma}_{c} dt = dt\,. \tag{11.18}$$

One can show [143, 144] that this is the case if and only if

$$\left[\frac{R''_T}{1-(R'_T)^2}\right]_{T=t_0} = -\left[\frac{\bar{R}''_{\bar{T}}}{1-(\bar{R}'_{\bar{T}})^2}\right]_{\bar{T}=t_0} \tag{11.19}$$

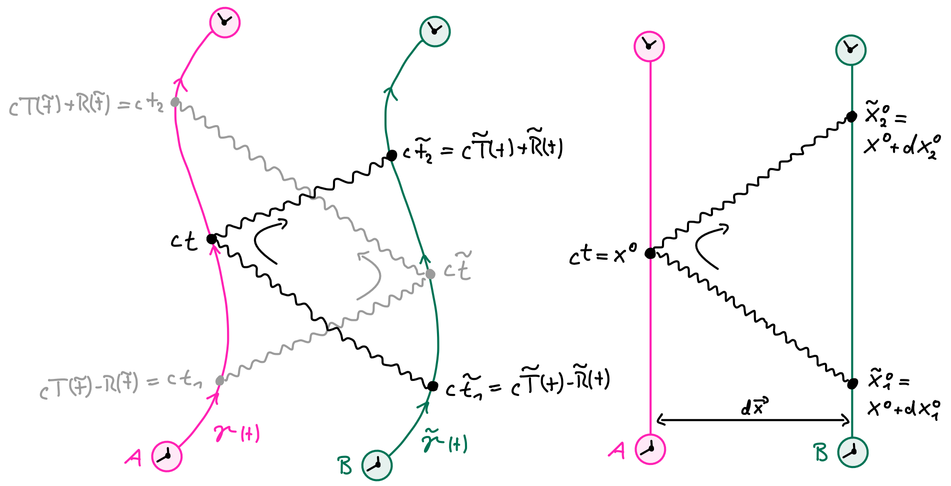
with the shorthand notation $R'_X \equiv \frac{dR}{dX}$.

This provides an operational procedure to check (in principle) that atomic clocks indeed measure the proper time Eq. (11.10) and are therefore ideal clocks.

5 | Simultaneity:

i | ⊲ Two nearby clocks γ and $\tilde{\gamma}$:

We assume that they are within each others "tube" where radar coordinates can be defined.



We say:

$$B \text{ is ⁂ } \textit{Einstein synchronous} \text{ to } A \quad :\Leftrightarrow \quad T(\tilde{t}) = \tfrac{1}{2}(t_2 + t_1) = \tilde{t} \tag{11.20a}$$

$$A \text{ is ⁂ } \textit{Einstein synchronous} \text{ to } B \quad :\Leftrightarrow \quad \tilde{T}(t) = \tfrac{1}{2}(\tilde{t}_2 + \tilde{t}_1) = t \tag{11.20b}$$

- Note that $\tilde{t}$ is measured by clock B while $T(\tilde{t})$ is computed from t_2 and t_1 which are measured by clock A.
- This is simply ← *Einstein synchronization* in SPECIAL RELATIVITY (← Section 1.1 and ➔ Problemset 1 last semester) generalized to arbitrary spacetimes. The synchronization constraint Eq. (11.20) is also known as ⁂ *Radar synchronization*.
- Recall that Einstein synchronization in SPECIAL RELATIVITY (i.e., on Minkowski space) could be proven to be *symmetric* and *transitive* for clocks at rest in the same inertial frame (➔ Problemset 1 last semester). In the more general situation considered here, Einstein synchronization is neither symmetric nor transitive (note that, in general, there is no inertial frame that encompasses both clocks), see [144].
- One can show [144] that if the synchronization is *symmetric* (as any good synchronization should be), then the radar distances $R(\tilde{t})$ from A to B and $\tilde{R}(t)$ from B to A are necessarily constant and equal: $R(\tilde{t}) \equiv \tilde{R}(t) \equiv \text{const}$.

ii | We now want to study possible obstructions to synchronizing clocks in curved spacetimes. For simplicity, we consider two *infinitesimally* separated clocks (right sketch).

This calculation follows LANDAU & LIFSHITZ [145] (§84, pp. 233–236).

⊲ Two infinitesimally close clocks A and B separated by $\mathrm{d}\vec{x} = \{\mathrm{d}x^n\}$:

$$A \text{ synchronous to } B \quad \overset{\text{Eq. (11.20)}}{\Longleftrightarrow} \quad x^0 + \underbrace{\delta x^0}_{\text{Offset}} = \tfrac{1}{2}(\tilde{x}_2^0 + \tilde{x}_1^0) \tag{11.21}$$

- We assume that the position of the clocks is labeled by x^m and their reading corresponds to the coordinate x^0, i.e., these are *coordinate* clocks. They are *not* required to be ideal clocks ticking off proper time; the following argument therefore applies to arbitrary coordinate systems with time-like coordinate x^0 and space-like coordinates $x^{1,2,3}$.

We say that a coordinate x^0 is time-like (at a given point $p \in M$) if the curve on M defined by varying x^0 and keeping x^m constant is a time-like curve in p (in a similar way we define x^m to be space-like). Mathematically, this means that $\partial_0 \equiv \frac{\partial}{\partial x^0} \in T_p M$ is a time-like vector:

$$\|\partial_0\|_p^2 = g_p(\partial_0, \partial_0) = g_{\mu\nu}(p) \underbrace{\mathrm{d}x^\mu(\partial_0)}_{\delta_0^\mu} \underbrace{\mathrm{d}x^\nu(\partial_0)}_{\delta_0^\nu} = g_{00} > 0\,. \tag{11.22}$$

Not every coordinate system has this property, but because of the Lorentzian signature of g, there are always (many) such coordinate systems. These coordinate systems are useful because their time-axis corresponds (at least locally) to possible trajectories of physical bodies (not necessarily free-falling ones). Thus one can think of the coordinate x^0 as *the time* (not necessarily proper time) measured by some (not necessarily free-falling) clock tracing out the time-axis through spacetime. Note that a coordinate can be time-like in one region of spacetime, become null at some point, and then space-like in another region. So the "type" of a coordinate is not fixed like that of a geodesic. Note also that not every coordinate system is guaranteed to have a time-like coordinate at all (this is possible for non-orthogonal coordinates which are rarely used).

- Note that the offset δx^0 could be absorbed into one of the clocks by shifting its reading (corresponding to a coordinate transformation). But we can also simply agree that two events at A and B are simultaneous iff their local clocks differ by δx^0 (→ *below*). This is a generalization of the synchronization condition Eq. (11.20) with no downsides, at least for the comparison of two clocks.

Let *w.l.o.g.* $\tilde{x}_i^0 \equiv x^0 + \mathrm{d}x_i^0 \rightarrow$

$$x^0 + \delta x^0 = x^0 + \underbrace{\tfrac{1}{2}(\mathrm{d}x_2^0 + \mathrm{d}x_1^0)}_{\delta x^0} \tag{11.23}$$

iii | For the light signals used to synchronize the clocks we have:

$$\mathrm{d}s^2 = g_{00}\left(\mathrm{d}x^0\right)^2 + 2g_{0m}\mathrm{d}x^0\mathrm{d}x^m + g_{mn}\mathrm{d}x^m\mathrm{d}x^n \overset{!}{=} 0 \tag{11.24}$$

Here we separated the temporal from the spatial components $n, m = 1, 2, 3$.

Solve for $\mathrm{d}x^0 \equiv \mathrm{d}x_i^0 \overset{\circ}{\rightarrow}$

$$\mathrm{d}x_i^0 = \frac{1}{g_{00}}\left[-g_{0m}\mathrm{d}x^m \mp \sqrt{(g_{0m}g_{0n} - g_{mn}g_{00})\mathrm{d}x^m\mathrm{d}x^n}\right] \tag{11.25}$$

The minus sign corresponds to $\mathrm{d}x_1^0 < 0$.

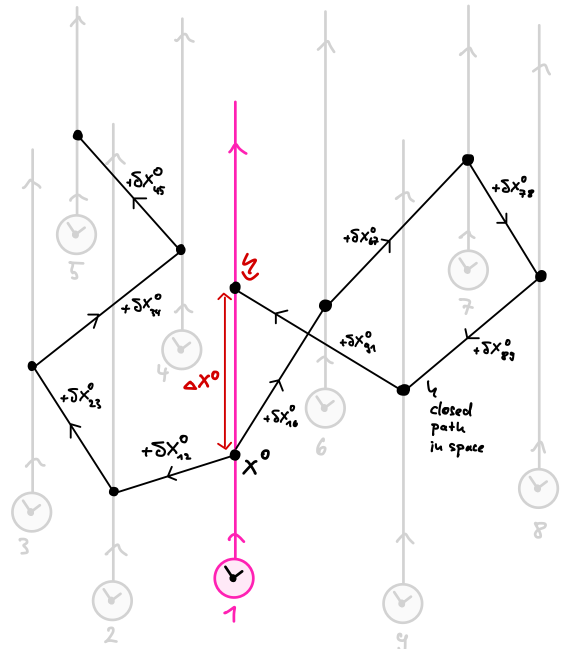
iv | Eq. (11.23) $\xrightarrow{\text{Eq. (11.25)}}$

$$\boxed{\delta x^0 = -\frac{g_{0m}}{g_{00}}\mathrm{d}x^m \quad \Leftrightarrow \quad g_{00}\delta x^0 + g_{0m}\mathrm{d}x^m = 0} \tag{11.26}$$

- *Interpretation:* δx^0 is the difference of the reading of two infinitesimally nearby clocks A at $\vec{x}$ and B at $\vec{x} + \mathrm{d}\vec{x}$ that indicates the time of two events happening *simultaneously* according to Einstein synchronization.

- You can think of δx^0 as a ← *connection* relating nearby clocks (*cf* Eq. (10.34)): If A displays the time x^0, we consider the time $x^0 + \delta x^0 = \frac{1}{2}(\tilde{x}_2^0 + \tilde{x}_1^0)$ displayed by B as "equal" (in the terminology of connections: "parallel", here better: *simultaneous*). If $\delta x^0 \neq 0$, the change in reading of nearby clocks is considered "fake"; this is not a problem in principle: If you have two clocks where the reading t of one always coincides with the reading $t + \Delta t$ of another, you don't loose anything and can consider them as being synchronized (as long as you know what Δt is). However, there *is* a problem coming from $\delta x^0 \neq 0$ if you consider different paths in space to synchronize your clocks (→ *next*).

v | ⊲ Closed path in space:

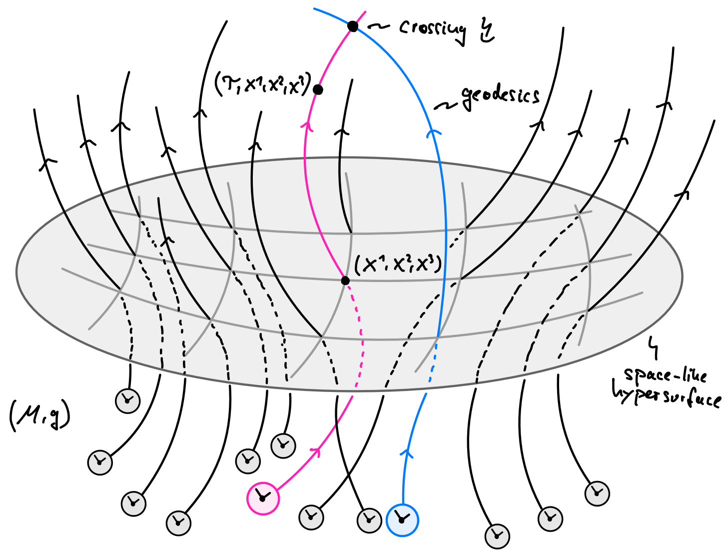


$$\to g_{0m} \neq 0 \;\Rightarrow\; \delta x^0_{e=(ij)} \neq 0 \;\Rightarrow\; \Delta x^0 \equiv \textstyle\sum_{e \in \text{Loop}} \delta x^0_e \neq 0 \text{ ☹}$$

Only in coordinates with $g_{0m} = 0 \;\Leftrightarrow\; \delta x^0 = 0$ the synchronization of clocks is *path-independent*. (Example: Minkowski space with clocks corresponding to inertial coordinates where $\eta_{\mu\nu} = \text{diag}\,(1, -1, -1, -1)$.) If not, synchronizing clocks along a closed path can lead to the identification of different times x^0 and $x^0 + \Delta x^0$ *of the same clock* as simultaneous!

- In the vicinity of every space-like slice ("hypersurface") of an arbitrary spacetime it is possible to construct a coordinate system in which $g_{0m} = 0$. Because of Eq. (11.26), on such slices the synchronization of clocks is consistently possible (= path-independent). Furthermore, it is possible to tweak the coordinates such that $g_{00} = 1$ so that stationary coordinate clocks ($\mathrm{d}\vec{x} = 0$) measure proper time (= are ideal clocks): $\mathrm{d}\tau^2 = \frac{1}{c^2}\mathrm{d}s^2 = g_{00}\mathrm{d}t^2 + 0$; such coordinates are called ↑ *synchronous*, see Misner *et al.* [2] (§27.4, p. 717).
- The above argument shows that, in general, it is possible to fill space with ideal clocks and synchronize all of them. The question is whether they *stay* synchronized for all times (i.e., whether is it possible to synchronize clocks throughout *spacetime*).

  The answer turns out to be negative because the synchronous coordinates, while being defined throughout *space* around a particular space-like hypersurface, cannot be extended to encompass all of *spacetime* (except for special cases like flat Minkowski space); they necessarily form "time singularities" [146]. This conclusion is reasonable if one thinks of synchronous coordinates as being constructed by ejecting free-falling ideal clocks from the hypersurface (with arbitrary spatial coordinates (x^1, x^2, x^3)):



These clocks (the "time axes" of the synchronous coordinate system) follow geodesics. But we already know that, in generic spacetimes with curvature, geodesics tend to attract/repel each other and, eventually, *cross*. At this point the coordinate system becomes singular because the map between events (= points on the spacetime manifold) and coordinates is no longer unique.

6 | Spatial distances:

i | ◁ *Space-like* curve $\gamma : [\lambda_a, \lambda_b] \to M$ on arbitrary spacetime (M, g):

¡! γ is not required to be a geodesic.

→ ⁂ *Proper distance*:

$$L[\gamma] := \int_{\lambda_a}^{\lambda_b} \mathrm{d}\lambda \sqrt{-g_{\mu\nu}(\gamma(\lambda))\dot{\gamma}^\mu \dot{\gamma}^\nu} \equiv \int_\gamma \sqrt{-g_{\mu\nu}(x)\mathrm{d}x^\mu \mathrm{d}x^\nu} \tag{11.27}$$

The minus is necessary because $\|\dot{\gamma}\|^2 < 0$ for a space-like curve.

¡! While mathematically the proper distance is defined completely analogous to the proper time Eq. (11.10), its operational/physical role is very different: Whereas proper time can be immediately identified as the time displayed by an ideal clock that follows a time-like trajectory, there is nothing that "follows" a space-like trajectory; hence there is no immediate physical interpretation associated to the proper distance defined above.

ii | To obtain an operationally meaningful concept of distance, it is reasonable to use the ← *radar distance* R defined in Eq. (11.16) as a distance measure of spatially separated points.

To this end, consider again the two infinitesimally close clocks A and B from above:

→ *Coordinate* time needed by radar pulse from B to A back to B:

$$\mathrm{d}x^0_{\rightleftarrows} := \tilde{x}^0_2 - \tilde{x}^0_1 = \mathrm{d}x^0_2 - \mathrm{d}x^0_1 \stackrel{11.25}{=} \frac{2}{g_{00}}\sqrt{(g_{0m}g_{0n} - g_{mn}g_{00})\mathrm{d}x^m \mathrm{d}x^n} \tag{11.28}$$

→ *Proper time* elapsed at position B during round trip:

$$\mathrm{d}\tau_{\rightleftarrows}^2 = \frac{1}{c^2}\mathrm{d}s_{\rightleftarrows}^2 = \frac{1}{c^2}g_{00}\left(\mathrm{d}x_{\rightleftarrows}^0\right)^2 = \frac{4}{c^2}\left(\frac{g_{0m}g_{0n}}{g_{00}} - g_{mn}\right)\mathrm{d}x^m\mathrm{d}x^n \tag{11.29}$$

Note that B is stationary in the considered coordinates so that $\mathrm{d}x_{\rightleftarrows}^m = 0$ for $m = 1, 2, 3$.

→ Infinitesimal *distance* between A and B:

$$\mathrm{d}l^2 := \left(\frac{c\,\mathrm{d}\tau_{\rightleftarrows}}{2}\right)^2 = \underbrace{\left(-g_{mn} + \frac{g_{0m}g_{0n}}{g_{00}}\right)}_{=:\tilde{g}_{mn}}\mathrm{d}x^m\mathrm{d}x^n \tag{11.30a}$$

$$\equiv \tilde{g}_{mn}(x)\mathrm{d}x^m\mathrm{d}x^n \tag{11.30b}$$

$\tilde{g}_{mn}$: metric of *three-dimensional space* (in vicinity of B)

iii | Notes:

- It is straightforward to show that

$$-g^{lm}\tilde{g}_{mn} \overset{\circ}{=} \delta^l_n \tag{11.31}$$

so that the inverse spatial metric is the *negative* of the spatial part of the inverse metric:

$$\tilde{g}^{lm} = -g^{lm} \quad \text{with} \quad \tilde{g}^{lm}\tilde{g}_{mn} = \delta^l_n\,. \tag{11.32}$$

- In general, it is operationally meaningless to integrate $\mathrm{d}l$ over a spatial curve η: $\int_\eta \mathrm{d}l$. While every infinitesimal distance element $\mathrm{d}l$ does make sense for some observer (because we constructed it as the radar distance), this does not mean that *adding* different such elements along a curve η with constant coordinate time x^0 makes sense. This is only reasonable if one can establish an unambiguous notion of simultaneity along the spacetime curve defined by η and $x^0 = \text{const}$ – which is not always possible (as discussed above). Thus, in GENERAL RELATIVITY, there is no general concept of a "distance" between bodies that has objective and operational meaning.

  (Note that we are not claiming that the length of a curve in space somehow depends on "how fast it is traversed": The spacetime curves we are considering are *space-like*, one cannot "traverse" them in any meaningful way! Only for the special cases where the metric $g_{\mu\nu}(x)$ is independent of time x^0, the length of a spatial curve η can be defined by $\int_\eta \mathrm{d}l$ and has a meaning that is independent of coordinates.)

7 | Speed of light:

Let us briefly comment on the role played by the speed of light in GENERAL RELATIVITY.

- Recall (Section 1.5):

  In the global *inertial systems* of SPECIAL RELATIVITY, light always propagates with the same velocity $v_{\max} = c$, and no signal can move faster.

- Problem:

  "Velocity" is an observer-/coordinate-dependent quantity that depends on the choice of time and space coordinates. Objective statements about velocity therefore require the choice of a distinguished class of coordinate systems.

Note that this was also the case in SPECIAL RELATIVITY: The coordinate velocity is only constant c in *inertial coordinates* (where the *coordinate* velocity corresponds to a *physical* velocity because inertial coordinates are, by definition, Cartesian). By contrast, in the Rindler coordinates of an accelerated observer, the coordinate speed of light is only c *locally*, but can be less or more than c away from the observer.

→ ∢ Local inertial coordinates:

The EEP suggests: →

> The speed of light is constant (c) in locally inertial coordinates
> at any point of spacetime (= as measured in a freely falling laboratory).

- ¡! Remember that in SPECIAL RELATIVITY we went to great lengths to link the abstract notion of an inertial coordinate system to an operationally defined contraption of synchronized clocks and rods forming a Cartesian lattice. *Observing* (or *measuring*) events was then defined via this information-gathering contraption, and, as stressed previously, is different from *seeing* events (i.e., waiting for light signals to reach the camera of someone sitting in the origin of the rod lattice; recall the ← *Penrose-Terrell effect* mentioned in Section 2.1 and ➔ Problemset 3 of last semester). Since, in SPECIAL RELATIVITY, inertial systems were assumed to be *global*, covering all of spacetime (i.e., the rod lattice was assumed to cover all of space and the clocks remained synchronized for all of times) we could *measure* events (times, distances, speeds) everywhere in spacetime, in particular: far away. Thus a statement like "the velocity of a light signal at Alpha Centauri was measured to be c" makes sense because we have a (magical) grid of synchronized clocks that reaches from Earth to Alpha Centauri (note that synchronized is short for "synchronized for all times", i.e., in particular, the clocks tick with the same rate).

  In Section 8.2 we argued that global inertial systems do not survive the presence of gravity: they shrink to small, local patches on spacetime, namely free falling laboratories that are small enough to be not affected by tidal forces. But this means that we also *cannot* construct a universe-encompassing latticework of synchronized clocks, and, as a consequence, there is no longer a well-defined concept of *observing/measuring* distant events! In particular, there is no well-defined way for an observer located on Earth to *measure* the speed of a light signal at Alpha Centauri, we can only point a telescope into the direction and *watch* (= *see*). This is what astronomers do (it is all they *can* do) and they call it *observing* (they do it even in *observatories*); but keep in mind that this is not what we refereed to as *observing* in the context of SPECIAL RELATIVITY! We will adopt this new terminology henceforth.

  Thus, in GENERAL RELATIVITY, we can only *measure* the speed of light *locally* (if one manages to setup a pair of synchronized clocks). The speed of *distant* light signals can only be *observed*, not *measured*. The constancy of the speed of light above refers only to *local measurements*, not to remote observations; the observed speed of light can be both smaller an larger than c!

  (Whether the lab in which a local measurement of the speed of light is performed is inertial or accelerated actually doesn't matter: one always measures c. This is so because accelerated observers can describe physics locally by Rindler coordinates (➔ Problemset 3), and in these the *local* (coordinate) speed of light is also c.)

- ¡! If you *calculate* the speed of light in *non-inertial* coordinates, the result is not necessarily c. For example, in the → *Schwarzschild metric* of a spherically symmetric mass, the (coordinate) speed of light in Schwarzschild *coordinates* decreases when approaching the → *Schwarzschild radius*. This corresponds to the well-known phenomenon that an observer far away from a black hole *sees* light freeze when approaching the event horizon.

Similarly, in Rindler coordinates (a useful coordinate system for observers with constant proper acceleration, ➔ Problemset 3), the (coordinate) velocity of distant light signals can be less or more than c, depending on their position.

- Recall our discussion ← *above* on the cone field on Lorentzian manifolds that encodes the local causal structure of spacetime. The light cones are locally generated by null cones which, by definition, are spanned by the tangent vectors of light rays. This makes the statement that "the speed of light is constant in local inertial systems" a tautology in GENERAL RELATIVITY: There is no fixed background of distinguished coordinates with respect to which you could measure that c is constant. It is the *finiteness* of the speed of information propagation (e.g., by light) that guarantees a *local* causality structure on spacetime; this causality structure can be encoded by the cone field. One can then, without loss of generality, choose units of time and length such that, locally, the signals that span the null cones propagate with constant velocity c.

  We already touched this topic in Section 1.4 where we derived the Lorentz transformation. There we realized that it is not so much the *constancy* of the speed of light that is important but its *finiteness* (the constancy follows from the finiteness, recall Eq. (1.73)). It is the finiteness of the speed of information propagation that induces a *local* causal structure of events.

8 | Implementing Einstein's Equivalence Principle EEP:

To implement the EEP into the physical models that are defined on the spacetime of GENERAL RELATIVITY, one can employ the following procedure:

§ Principle 3: Minimal-Coupling Principle MCP (*"Comma-goes-to-Semicolon Rule"*)

i | Take a physical model (equation) in manifestly Lorentz covariant form.

The model is of course assumed to be valid in SPECIAL RELATIVITY (i.e., describe the laws of nature correctly in the globally inertial coordinates of flat Minkowski space).

ii | Convert it into a generally covariant form by the following substitutions:

$$\partial_\mu \mapsto \nabla_\mu \quad (\text{or }, \mapsto ;) \qquad \text{and} \quad \eta_{\mu\nu} \mapsto g_{\mu\nu}(x) \tag{11.33}$$

If the physical model is given by a Lagrangian density (i.e., in an integral form), one must also ensure that the integrand transforms as a scalar by the substitution $\mathrm{d}^4x \mapsto \mathrm{d}^4x\sqrt{g}$ as discussed in Section 10.3.1.

iii | Assert the validity of this model in the curved spacetimes of GENERAL RELATIVITY.

Examples: → Sections 11.2 and 11.3.

- The MCP has a similar status as ↓ *canonical quantization* in quantum mechanics: It provides a (mathematically supported) *guiding principle* to "update" an "old" physical model to a "new" form that adapts the model to a more fundamental theory, while respecting some sort of correspondence principle (which is necessary because the outdated model works well in some domain).

  Furthermore, the MCP ensures (at least in the cases relevant to us, → *below*) that the constructed models respect the EEP, in that it asserts the *absence* of explicit (non-minimal) couplings to the curvature tensor (which, if present, would allow local experiments to detect the presence of a gravitational field, → *below*).

- Just as canonical quantization works for most situations encountered in physics – but fails in certain edge cases (↑ *Weyl quantization*, ↑ *Groenewold's theorem*) –, the MCP works for most relativistic theories (in particular, all problems that we encounter in this course), but has some subtle ambiguities that prevent a unique outcome. The problem is that higher-order covariant derivatives do not commute, whereas higher-order partial derivatives do:

$$\dots\partial_\mu\partial_\nu\dots = \dots\partial_\nu\partial_\mu\dots \tag{11.34a}$$

$$?\downarrow \qquad\qquad \downarrow ?$$

$$\dots\nabla_\mu\nabla_\nu\dots \overset{10.71}{\neq} \dots\nabla_\nu\nabla_\mu\dots \tag{11.34b}$$

  Depending on which order of covariant derivatives you pick, you end up with theories that are equivalent in SPECIAL RELATIVITY (i.e., on flat Minkowski space), but differ by a curvature-dependent term in GENERAL RELATIVITY. Thus the MCP is only a unique recipe for first-order differential equations [147]. (Luckily, we are physicist and can experiments let decide which generally covariant model describes the laws of nature correctly.) For more details on this ordering ambiguity, see MISNER *et al.* [2] (pp. 388–389, §16.3 and pp. 390–391, Box 16.1).

- Here is an example to illustrate the MCP and contrast it to *non-minimal* coupling. The example also shows that non-minimal coupling typically leads to a violation of the EEP (put differently, the EEP lends credence to the MCP):

  The real Klein-Gordon field is given by the Lorentz covariant action (in ← Section 7.1 we discussed the complex Klein-Gordon field)

$$S[\phi] = \int d^4x\left[\eta^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\phi) - m^2\phi^2\right] = \int d^4x\left[\phi^{,\mu}\phi_{,\mu} - m^2\phi^2\right], \tag{11.35}$$

  the Euler-Lagrange equations of which are the Klein-Gordon equation

$$\left(\partial^2 + m^2\right)\phi(x) = 0\,. \tag{11.36}$$

  If we want to study the Klein-Gordon field in the curved spacetime of GENERAL RELATIVITY, the MCP tells us to construct the generally covariant action (➔ Problemset 4)

$$S_g[\phi] = \int d^4x\sqrt{g(x)}\left[g^{\mu\nu}(x)(\nabla_\mu\phi)(\nabla_\nu\phi) - m^2\phi^2\right] \tag{11.37a}$$

$$= \int d^4x\sqrt{g}\left[\phi^{;\mu}\phi_{;\mu} - m^2\phi^2\right], \tag{11.37b}$$

  which reduces to Eq. (11.35) on flat Minkowski space ($g = \eta$). The corresponding equation of motion is the generally covariant Klein-Gordon equation

$$\left(\Delta + m^2\right)\phi(x) = 0\,, \tag{11.38}$$

  where Δ is the ← *Laplace-Beltrami operator* Eq. (10.97). In locally inertial coordinates, Eq. (11.38) reduces to Eq. (11.36), realizing the EEP (check this!).

  Now let us couple the Klein-Gordon field in a *non-minimal* way to gravity by adding a scalar interaction with the Ricci scalar $R(x)$ [← Eq. (10.117)] to the action (green):

$$\tilde{S}_g[\phi] = \int d^4x\sqrt{g}\left[\phi^{;\mu}\phi_{;\mu} - m^2\phi^2 - \xi R(x)\phi^2\right], \tag{11.39}$$

  where $\xi \in \mathbb{R}$ is a coupling constant. The generally covariant equation of motion is clearly

$$\left[\Delta + m^2 + \xi R(x)\right]\phi(x) = 0\,, \tag{11.40}$$

where $R(x)$ depends on the metric $g_{\mu\nu}(x)$.

But Eq. (11.40) does *not* reduce to the Klein-Gordon Eq. (11.36) of SPECIAL RELATIVITY in local inertial frames! This is so because the term $\xi R(x)$ is a *scalar* that does not vanish on a curved spacetime in *any* coordinate system (in particular, locally inertial coordinates). Thus Eq. (11.40) explicitly violates the **EEP** because, using local measurements of the evolution of the Klein-Gordon field ϕ, a local observer can detect the presence of curvature (and thereby gravity).

- Be careful when making statements about *higher-order* covariant derivatives! For example, on a *flat* Minkowski space in *globally inertial coordinates*, all covariant derivatives in a generally covariant equation become partial derivatives:

$$T^{\cdots}{}_{;\alpha;\beta} \xrightarrow{g_{\mu\nu}\to\eta_{\mu\nu}} T^{\cdots}{}_{,\alpha,\beta}\,. \tag{11.41}$$

This is true because the Christoffel symbols are identically zero everywhere, so that all their derivatives vanish as well.

By contrast, on a *curved* spacetime in *locally inertial coordinates* [with metric as in Eq. (10.89)], this is *not* true:

$$T^{\cdots}{}_{;\alpha;\beta} \xrightarrow{g_{\mu\nu}\to\bar{g}_{\mu\nu}} \cancel{T^{\cdots}{}_{,\alpha,\beta}}\,. \tag{11.42}$$

To see this, note that the left-hand side contains *derivatives* of the Christoffel symbols – which do not necessarily vanish in locally inertial coordinates (because coordinate transformations cannot make curvature go away).

This also follows from the Ricci identity Eq. (10.71):

$$T_{k;l;m} - T_{k;m;l} = R^{i}{}_{klm} T_i\,. \tag{11.43}$$

On a curved space, the right-hand side does not vanish *in any coordinate system*, so that covariant derivatives do not commute; in particular, they cannot become partial derivatives.

This line of reasoning leads to a peculiar conclusion: Applying the **MCP** to higher-order differential equations can lead to generally covariant equations that contain *curvature terms*, and thereby violate the **EEP** (in its strictest form)! (Note that they do obey a correspondence princple in the sense that they reproduce the physics of SPECIAL RELATIVITY on flat Minkowski space.) This phenomenon is of course rooted in the fact that in locally inertial coordinates the metric is only Minkowskian *to first order*. For more details on this predicament (that most textbooks seem to be silent about) see Ref. [148]. CARROLL argues that curvature terms (from non-minimal or higher-order minimal coupling) may actually be present, but should be supressed by the Planck scale (↑ Ref. [102], §4.7, pp. 179–181).