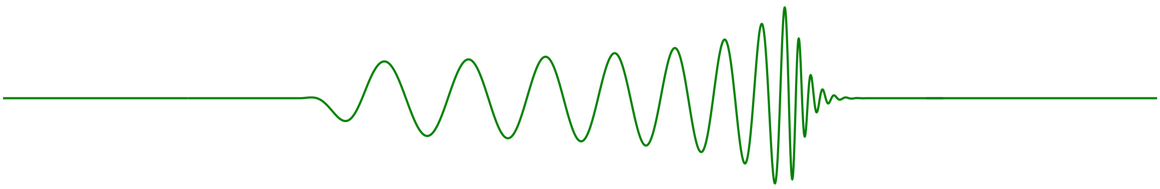
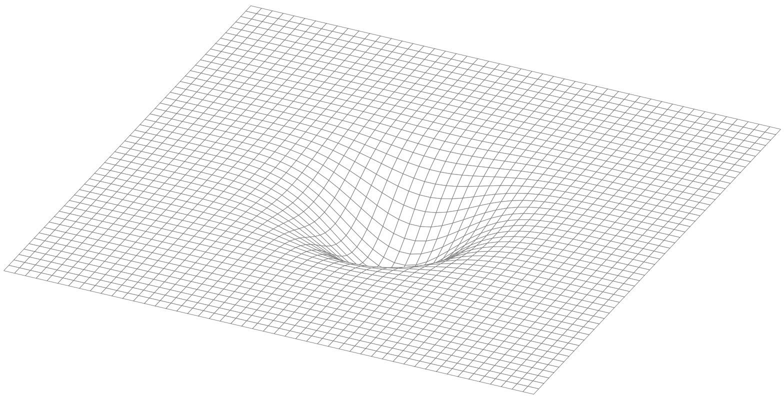


Summary

Relativity



Max Melching

Acknowledgements:

Code for spacetime diagrams is inspired by the one presented in https://de.overleaf.com/latex/templates/minkowski-spacetime-diagram-generator/kqskfzgkjrvq

Contents

List of Figures

Acronyms

SR special relativity

GR general relativity

Mathematical Notation

a thin letters (lower and upper case) are used to denote scalar values and variables with scalar values

$\vec{a}$ thin letters with arrows are used to denote three-vectors

$\underline{a}$ thin, underlined letters are used to denote four-vectors

$\mathbf{A}$ bold, capital letters are used to denote matrices

$a^\mu b_\mu$ the Einstein summation convention is adopted, where repeated indices denote sums:

$$a^\mu b_\mu = \sum_\mu a^\mu b_\mu$$

1 Mathematics

1.1 Gravitational Physics Summary – Math Part

1.1.1 Basics of Manifolds

to describe curved spacetime, we need a coordinate-independent notion of spaces; this is given by manifolds, which are described using charts=coordinates but have an independent, invariant meaning; similarly, they can often be pictured to be embedded in some higher-dimensional Euclidian space, but that need not be the case

therefore, physics happens on manifolds, so events are points on it and more

defining vectors on manifolds is a non-trivial topic, they are now completely distinct notion from points and cannot be visualized as pointing from some origin to this point (problem: thinking of an embedded manifold for now, the vectors would point out of the manifold); instead, we can define vectors locally (infinitesimally) via derivatives of curves (i.e. as *tangent vectors*; the corresponding set of all tangent vectors is called *tangent space* V); for a function $f(x^\alpha)$ on the manifold, we can calculate the derivative along a curve $\gamma = \gamma(\sigma)$ ($\sigma \in \mathbb{R}$ parametrizes γ):

$$\frac{df}{d\sigma} = \frac{df(x^\alpha(\sigma))}{d\sigma} = \frac{\partial f}{\partial x^\alpha}\frac{dx^\alpha}{d\sigma} \tag{1.1}$$

this is a simply application of the chain rule and it yields the tangent vector components

$$t^\alpha = \frac{dx^\alpha}{d\sigma} = \underline{t} \cdot x^\alpha \tag{1.2}$$

because the vector is supposed to act as

$$\underline{t} \cdot f = \left(t^\alpha \frac{\partial}{\partial x^\alpha}\right) \cdot f = \frac{\partial f}{\partial x^\alpha} t^\alpha \tag{1.3}$$

(remember: we identify it with a derivative, which can in turn be expressed using partial derivatives with respect to the coordinates)

formally, we can express this as

$$\underline{t} = t^\alpha \underline{e_\alpha} = t^\alpha \frac{\partial}{\partial x^\alpha} = \frac{d}{d\sigma} \tag{1.4}$$

(note: $\frac{\partial}{\partial x^\alpha}$ is often abbreviated as ∂_α)

tangent vectors are invariant quantities, they do not (and should not!) depend on the coordinates we use to express them; their components, on the other hand, are *not* invariant; they obey the following transformation rule:

$$\underline{t} = t^\alpha \frac{\partial}{\partial x^\alpha} = t^\alpha \frac{\partial x'^\beta}{\partial x^\alpha} \frac{\partial}{\partial x'^\beta} = t'^\beta \frac{\partial}{\partial x'^\beta} \qquad \Leftrightarrow \qquad \boxed{t'^\beta = t^\alpha \frac{\partial x'^\beta}{\partial x^\alpha}} \tag{1.5}$$

once again, this is basically just an application of the chain rule

next natural step: linear maps on tangent space V (= set/space of tangent vectors); these are called *covectors* or *one forms* (elements of the dual space or *cotangent space* V^*) and it turns out that we can identify them with differentials/gradients of functions

$$df = \frac{\partial f}{\partial x^\alpha} dx^\alpha \tag{1.6}$$

where we chose a convenient basis $\left\{\underline{e}^\alpha\right\}_\alpha = \{dx^\alpha\}_\alpha$ of the dual vector space; these satisfy

$$dx^\alpha \left(\frac{\partial}{\partial x^\beta} \right) = \frac{\partial x^\alpha}{\partial x^\beta} = \delta^\alpha_\beta \tag{1.7}$$

more generally, such covectors $w : V \to \mathbb{R}$ obey

$$w\left(\alpha \underline{a} + \beta \underline{b}\right) = \alpha w\left(\underline{a}\right) + \beta w\left(\underline{b}\right), \qquad \forall a, b \in \mathbb{R},\ \underline{a}, \underline{b} \in V \tag{1.8}$$

we can also characterize covectors via tuples of components

$$w_\alpha = w\left(\underline{e_\alpha}\right) = w(\partial_\alpha) = \partial_\alpha w \tag{1.9}$$

in general, we can also write

$$w\left(\underline{a}\right) = w_\alpha \underline{e}^\alpha \left(a^\beta \underline{e_\beta}\right) = w_\alpha a^\alpha \tag{1.10}$$

to see how covector components in different coordinates are related, we look at the following inner product (which is also invariant)

$$w\left(\underline{t}\right) = w_\alpha t^\alpha \stackrel{!}{=} w'_\beta t'^\beta = w'_\beta \frac{\partial x'^\beta}{\partial x^\alpha} t^\alpha \qquad \Leftrightarrow \qquad \boxed{w'_\beta = w_\alpha \frac{\partial x^\alpha}{\partial x'^\beta}} \tag{1.11}$$

1.1.2 Tensors

We have seen how covectors are maps from V to the real numbers. Similarly, one can show that there is a unique identification between vectors from V and maps from the dual space V^* to the real numbers – vectors are also maps. It is possible to generalize this concept to coordinate-independent entities which map multiple vectors, covectors or mixes of them to the real numbers. Linear maps

$$T : V^n \times (V^*)^m = V \times \ldots \times V \times V^* \times \ldots V^* \to \mathbb{R} \tag{1.12}$$

are called *tensors* of rank $m+n$. Due to their invariance under coordinate-transformations, every physical quantity has to be expressed as a tensor.

Just like vectors can be collected in components $t^\alpha = \underline{t} \cdot x^\alpha = \partial_\sigma x^\alpha$ and covectors in components $w_\alpha = w(\partial_\alpha)$, we can characterize a tensor of rank $m+n$ using components

$$\boxed{T^{\alpha_1 \ldots \alpha_m}{}_{\beta_1 \ldots \beta_n} = T\left(\underline{e_1}, \ldots, \underline{e_n}, \underline{e^1}, \ldots, \underline{e^m}\right)}\,. \tag{1.13}$$

Remark: it is no typo that there are m upper and n lower indices. This reflects the fact that a tensor of rank $m+n$ can map m covectors with its m "vectorial" indices and n vectors with its n "covectorial" indices.

These components do change under coordinate transformations. The corresponding behaviour can be derived from the ones for vectors (1.5) and covectors (1.11),

$$\boxed{T'^{\alpha\beta\ldots}{}_{\gamma\delta\ldots} = T^{\mu\nu\ldots}{}_{\lambda\sigma\ldots} \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\beta}{\partial x^\nu} \cdots \frac{\partial x^\lambda}{\partial x'^\gamma} \frac{\partial x^\sigma}{\partial x'^\delta} \cdots}\,. \tag{1.14}$$

This is the important *tensor transformation law*.

The rank of a tensor can be reduced if we insert a fixed object into one of the "slots", i.e. in the example of a rank-4-tensor

$$T(\cdot,\cdot,\cdot,\cdot) \to T'(\cdot,\cdot,\cdot) = T(\underline{t},\cdot,\cdot,\cdot) \qquad T^{\alpha\beta}{}_{\gamma\delta} T'^{\alpha\beta}{}_{\delta} = T^{\alpha\beta}{}_{\gamma\delta} t^\gamma \tag{1.15}$$

or

$$T(\cdot,\cdot,\cdot,\cdot) \to T'(\cdot,\cdot,\cdot) = T(\cdot,\cdot,w,\cdot) \qquad T^{\alpha\beta}{}_{\gamma\delta} \to T'^{\beta}{}_{\gamma\delta} = T^{\alpha\beta}{}_{\gamma\delta} w_\alpha \tag{1.16}$$

Remark: might be inconsistent to write components like this because vectorial indices come first but the first arguments in T are also vectorial (which they connect to a covectorial index).

Example 1.1: Known Tensors

We have already encountered several examples of tensors. Vectors and covectors are rank-1-tensors, which should not be surprising because we used them to derive general tensors. However, scalars are also tensors, namely of rank 0 – they can be thought of as mapping the real numbers to themselves without taking any further arguments.

Another example of a tensor, which plays a great role in geometry on manifolds and thus – as we will see later – also in physics, is the *metric*. The following properties can be used as a definition for this 2-tensor:

(1.) The metric is symmetric.

(2.) The metric is non-degenerate.

Together with the usual properties of a tensor, like linearity, this defines a (pseudo)metric. In case of spacetime, this is characterized by the fact that the metric has three positive and one negative eigenvalue.

Just like any other tensor, the metric can be characterized by their components. These we can also be read off from the line element

$$ds^2 = g_{\mu\nu}\, dx^\mu dx^\nu := g_{\mu\nu}\, dx^\mu \otimes dx^\nu \tag{1.17}$$

which is a frequently used tool in geometry (for example to measure lengths). Often, one thinks of the covectors dx^μ in this expression as infinitesimal changes in the coordinate x^μ and of the corresponding component a in adx^μ as the effect of this change. This is justified by the fact that $adx^\mu(\partial_\nu) = a\delta^\mu_\nu$, the coefficient of dx^μ indeed contains all information about the direction ∂_μ which is present in the whole object.

Metrics can be used to define inner products, which are not natively present on manifolds, in the following manner:

$$\underline{A} \cdot \underline{B} := g\left(\underline{A}, \underline{B}\right) = g_{\mu\nu} A^\mu B^\nu \,. \tag{1.18}$$

Inner products shall be symmetric, i.e. $\underline{A} \cdot \underline{B} = \underline{B} \cdot \underline{A}$, which is why we demanded symmetry of g. That means

$$g\left(\underline{A}, \underline{B}\right) = g\left(\underline{A}, \underline{B}\right) \qquad\qquad g_{\mu\nu} A^\mu B^\nu = g_{\nu\mu} A^\nu B^\mu \,. \tag{1.19}$$

The metric provides us with a natural identification between vectors and covectors because $g\left(\cdot, \underline{A}\right)$ is nothing but a map which takes a vector and maps it to a real number – which is the definition of a covector. Similarly, we can identify covectors w with the unique vector $\underline{A}$ that fulfils $w\left(\underline{B}\right) = g\left(\underline{A}, \underline{B}\right)$. In components, these requirements read

$$\boxed{A_\mu = g_{\mu\nu} A^\nu} \qquad\qquad \boxed{A^\mu = g^{\mu\nu} A_\nu} \tag{1.20}$$

where $g^{\mu\nu}$ denote the components of the inverse metric, which is defined by

$$\boxed{g^{\mu\sigma} g_{\sigma\nu} = \delta^\mu_\nu} \,. \tag{1.21}$$

Apparently, it is almost trivial to change from vectors to covectors and vice versa in this component notation. For this reason, the strict distinction between A^μ and A_μ is often dropped (at least for interpretation purposes).

1.1.3 Covariant Derivative

1.2 Notes & Thoughts

to be able to develop an appropriate/meaningful notion of parallelism, we need a "better" derivative. This will be provided by a connection

1.2.1 Giulini GR lectures May 19 and 26

in Riemann normal coordinates, all Christoffel symbols vanish; but they only exist in neighbourhoods around points p (are Riemann normal coordinates *at* p); they are *very* helpful in calculations because tensor equations only have to be proven in a single coordinate system, which we can choose to be Riemann normal coordinates because we have just shown that they do exist

uhh, Christoffel symbols satisfy an affine transformation law, not linear (indicator of not tensorial) because there is term without $\Gamma^{\sigma}_{\mu\nu}$

formel of Koszul only holds like this for torsion-free and metric connections; and by substituting $X = e_\alpha, Y = e_\beta, Z = e_\gamma$ into it and then contracting with certain component of inverse metric gives rise to formula for connection coefficients (commonly called Christoffel symbols for Levi-Civita connection) and these uniquely determine the connection (which is proof for uniqueness); note that connection coefficients (= covariant derivative with only basis vectors) already determines the connection because connection is $\mathbb{R}$-linear, tensorial (C^∞-linear) and obeys the Leibniz rule

$\nabla_X Y = X^\alpha \left(\nabla_\alpha Y^\beta\right) \frac{\partial}{\partial x^\beta}$, so components of $\nabla_X Y$ are partial derivatives of components plus extra term, i.e. $\nabla_{\frac{\partial}{\partial x^\alpha}} Y =: \left(\nabla_\alpha Y^\beta\right) \frac{\partial}{\partial x^\alpha}$; however, they to read this is *not* covariant derivative of Y^β since this would be the covariant derivative of a function, but instead as the β-component of the covariant derivative $\nabla_\alpha Y$

here it is, reason why covariant derivative of covector (with respect to some vector field X) looks the way it looks:

$$
\begin{aligned}
\nabla_X \omega &= \nabla_{X^\alpha \frac{\partial}{\partial x^\alpha}} \left(\omega_\beta dx^\beta\right) \\
&= X^\alpha \left(dx^\beta \nabla_{\frac{\partial}{\partial x^\alpha}} \omega_\beta + \omega_\beta \nabla_{\frac{\partial}{\partial x^\alpha}} dx^\beta\right) \\
&= X^\alpha \frac{\partial \omega_\beta}{\partial x^\alpha} dx^\beta
\end{aligned}
$$

but $dx^\beta \left(\frac{\partial}{\partial x^\alpha}\right) = \delta^\beta_\alpha$, so by taking the derivative of this equation we see that ... $\left(\nabla_{\frac{\partial}{\partial x^\gamma}} dx^\beta\right)\left(\frac{\partial}{\partial x^\alpha}\right) = \left(\nabla_{\frac{\partial}{\partial x^\gamma}} dx^\beta\right)^\alpha = -\Gamma^\beta_{\gamma\alpha}$

from that we get general formula for tensors of arbitrary rank because of Leibniz rule;

therefore, we get the "master formula"

$$\begin{aligned}\nabla_X T &= T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_m} \frac{\partial}{\partial x^{\alpha_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{\alpha_k}} \otimes dx^{\beta_1} \otimes \cdots \otimes dx^{\beta_m} \\ &= X^\gamma \left(\nabla_\gamma T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_m}\right) \frac{\partial}{\partial x^{\alpha_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{\alpha_k}} \otimes dx^{\beta_1} \otimes \cdots \otimes dx^{\beta_m} \\ \nabla_\gamma T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_m} &= \left(\frac{\partial T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_m}}{\partial x^\gamma} + \sum_{i=1}^k \Gamma^{\alpha_i}_{\gamma\lambda} T^{\alpha_1\dots\alpha_{i-1}\lambda\alpha_{i+1}\alpha_k}_{\beta_1\dots\beta_m} - \sum_{j=1}^m \Gamma^\lambda_{\gamma\beta_j} T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_{j-1}\lambda\beta_{j+1}\dots\beta_m} \right) \end{aligned} \tag{1.22}$$

there is also another notion of derivative on manifolds, the Lie derivative; to define it, we do not need any additional structure (unlike for connection, connection coefficients need metric); the Lie derivative can often be used to express symmetries; components of it are

$$(\mathcal{L}_X T)^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_m} = X^\alpha \frac{\partial}{\partial x^\alpha} T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_m} - \sum_{i=1}^k \frac{\partial X^{\alpha_i}}{\partial x^\lambda} T^{\alpha_1\dots\alpha_{i-1}\lambda\alpha_{i+1}\dots\alpha_k}_{\beta_1\dots\beta_m} + \sum_{j=1}^m \frac{\partial X^\lambda}{\partial x^{\beta_i}} T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_{i-1}\lambda\beta_{i+1}\dots\beta_m} \tag{1.23}$$

we notice: for upper index we now have minus, lower index has plus (reversed compared to connection); interesting property: all partial derivatives could be replaced by covariant derivatives without changing the formula (despite them being defined independently of each other!); if a certain vector field defines a symmetry, i.e. the metric does not change under the flow of that vector field (stays constant along it/integral curves defined by it), then we can express that as the vanishing of the Lie-derivative of this symmetry-generating vector field; these vector fields are called Killing fields; note that $(\mathcal{L}_X g)_{\alpha\beta} = (\nabla_\alpha X^\gamma) g_{\gamma\beta} + (\nabla_\beta X^\gamma) g_{\alpha\gamma} = \nabla_\alpha X_\beta + \nabla_\beta X_\alpha$ (we just use writing in terms of covariant derivative for first equality) -¿ shouldn't that be equal to $\nabla_{[\alpha} X_{\beta]}$; would also explain $\mathcal{L} = \mathrm{Alt}(\nabla)$ statement I heard; ah no, this is the *symmetrized* part... But maybe that supports view, symmetric part is zero for Killing field (but this has vanishing Lie derivative, so antisymmetric part also zero, right?)

very interesting: Einstein tensor is divergence-free, i.e. $\nabla_\alpha G^{\alpha\beta} = 0$

we know that, in general, $[\nabla_\mu, \nabla_\nu] T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_m} \neq 0$; expanding this quantity for an arbitrary vector field X^α, $[\nabla_\mu, \nabla_\nu] X^\alpha = \cdots = \left(\frac{\partial}{\partial x^\mu}\Gamma^\alpha_{\nu\beta} - \frac{\partial}{\partial x^\nu}\Gamma^\alpha_{\mu\beta}\right) X^\beta$ + terms proportional to $\Gamma = R^\alpha_{\beta\mu\nu} X^\beta$ + terms proportional to Γ; in Riemann-normal coordinates, $\Gamma = 0$ and $[\nabla_\mu, \nabla_\nu] X^\alpha = R^\alpha_{\beta\mu\nu} X^\beta$ and since both sides are tensors, this equation holds in general; curvature is related to (Giulini said "obstruction") commutivity of second derivatives; furthermore, pulling down the index α yields $[\nabla_\mu, \nabla_\nu] X_\alpha = -R^\beta_{\alpha\mu\nu} X_\beta$, which tells us how this quantity acts on a covector (again, has on other sign and acts on other index, like it was for covariant derivative itself); thus, we get the general formula $[\nabla_\mu, \nabla_\nu] T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_m} = \sum_{i=1}^k R^{\alpha_i}_{\lambda\mu\nu} T^{\alpha_1\dots\alpha_{i-1}\lambda\alpha_{i+1}\dots\alpha_k}_{\beta_1\dots\beta_m} - \sum_{j=1}^m R^\lambda_{\beta_j\mu\nu} T^{\alpha_1\dots\alpha_k}_{\beta_1\dots\beta_{i-1}\lambda\beta_{i+1}\dots\beta_m}$

application of that formula: $\nabla_\mu \nabla_\nu X_\beta = R^\alpha_{\mu\nu\beta} X_\alpha$ holds for any Killing vector field X; second derivatives are determined by vector field itself; similarly, $\frac{\partial}{\partial x^\alpha} X_\beta + \frac{\partial}{\partial x^\beta} X_\alpha = 2\Gamma^\gamma_{\alpha\beta} X_\gamma$, the symmetric part of first derivative of Killing vector field is determined by field itself as well; the only free parameters are value of the field itself and anti-symmetric part $\frac{\partial}{\partial x^\alpha} X_\beta - \frac{\partial}{\partial x^\beta} X_\alpha$ of first derivative (all derivatives of higher order are determined by relation to curvature tensor); solutions of linear differential equations (no matter of partial or not) constitute a vector space because we can add them together and multiply with numbers and maximum number of dimensions is given by number of freely specifiable initial conditions; here, these are values $X^\alpha\big|_p$ of Killing field at a specific point, i.e. $n = \dim(M)$, and $\frac{\partial}{\partial x^\alpha} X_\beta - \frac{\partial}{\partial x^\beta} X_\alpha\Big|_p$, i.e. $\frac{1}{2}n(n+1)$; in total, that means there are at most $n + \frac{1}{2}n(n+1)$ independent solutions to the Killing equation; in Minkowski space there are indeed 10, in that sense it is maximally symmetric (these generate symmetries of the space, which are given by Poincare group)

Riemann tensor has 20 components and there are 10 different traces (because of antisymmetry in first two indices, which means trace is always zero); our goal is now decomposing it into trace and traceless parts, both contain information about Riemann tensor; the trace part is nothing but the Ricci tensor $R_{\alpha\beta} = R^\lambda_{\alpha\lambda\beta}$, the "rest" (trace-free part) is the Weyl tensor (which he writes down in terms of some weird product); this product has the same symmetries as the Riemann tensor, so Weyl tensor also has them and it is trace free in addition, i.e. $W^\lambda_{\alpha\lambda\beta} = 0$ (taking these conditions into account, the Weyl tensor has 10 independent components in 4 dimensions; very interesting property is that *only* in 4 dimensions, the amount of information in Weyl, Ricci Tensor is the same; in 3 dimensions, Weyl tensor has no information and in higher ones much more than Ricci); in index-form it is given by $W^\alpha_{\beta\mu\nu}$

oof:

$$\begin{aligned}
R_{\alpha\beta\gamma\delta} &= \frac{1}{2}(g \cdot \mathrm{Ric})_{\alpha\beta\gamma\delta} - \frac{1}{12} R (g \cdot g)_{\alpha\beta\gamma\delta} + W_{\alpha\beta\gamma\delta} \\
&= \frac{1}{2}\big(g_{\alpha\gamma}R_{\beta\delta} + g_{\beta\delta}R_{\alpha\gamma} - g_{\alpha\delta}R_{\beta\gamma} - g_{\beta\gamma}R_{\alpha\delta}\big) \\
&\quad - \frac{1}{12} R \big(g_{\alpha\gamma}g_{\beta\delta} + g_{\beta\delta}g_{\alpha\gamma} - g_{\alpha\delta}g_{\beta\gamma} - g_{\beta\gamma}g_{\alpha\delta}\big) + W_{\alpha\beta\gamma\delta} \\
&= \frac{1}{2}\big(g_{\alpha\gamma}R_{\beta\delta} + g_{\beta\delta}R_{\alpha\gamma} - g_{\alpha\delta}R_{\beta\gamma} - g_{\beta\gamma}R_{\alpha\delta}\big) - \frac{1}{6} R \big(g_{\alpha\gamma}g_{\beta\delta} - g_{\alpha\delta}g_{\beta\gamma}\big) + W_{\alpha\beta\gamma\delta}
\end{aligned}$$

long talk about constant curvature; interesting statement (Schur's theorem): constant curvature $\Leftrightarrow$ Gaussian/sectional curvature of each point does not depend on the choice of the 2-tangent-plane through the point

the Weyl curvature $W^\alpha_{\beta\mu\nu}$ (which is a function of the metric g) has the important property of being conformally invariant, i.e. $W^\alpha_{\beta\mu\nu}(\Omega^2 g) = W^\alpha_{\beta\mu\nu}(g)$ for some function Ω and even the reverse statement is true: if $W^\alpha_{\beta\mu\nu}(g_1) = W^\alpha_{\beta\mu\nu}(g_2)$, then $\exists$ locally a function $\Omega \in C^\infty(M;\mathbb{R})$ without zeros such that $g_1 = \Omega^2 g_2$; for example, a vanishing Weyl tensor means that the space is locally, conformally flat -¿ all of these statements are valid only for $n \geq 3$

1.2.2 Order

basically take order from Penrose?; other way to put it: from summary H_Analysis, but with less math; also Carroll?

first: do manifolds; then go to tangent space (we want vectors); then go to bundles (first tangent bundle, then more general vector bundles); then define tensors and tensor bundle; then go to differential geometry

1.2.3 General Thoughts

Schwarzschild metric contains information on many effects of BHs in its components! coefficient $1 - \frac{2M}{r}$ in front of dt^2 tells us about time dilation close to BH (more t goes by the closer you get) and $\frac{1}{1-\frac{2M}{r}}$ in front of dr^2 tells us about curvature of space (increases as r decreases)

1.2.4 Math Stuff

regarding tensor product: throughout the discussions, linearity of objects was very important (we have used it for the differential, many mappings, etc.); however, a very important notion that is not linear is the underlying spaces we have looked at; take for example $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$: here, we do not have linearity, which would require $(2,1) = 2 \cdot (1,1)$ and clearly, this is not true; however, we might be interested in such a space and this is what is called a tensor product space; there are more tensor products, we also have to make sense of the one of objects in this space, but this is a good motivation

goal of derivatives is approximation to first order, which is expressed in demanding linearity of operators

differentiation is linear and has Leibniz rule, so it already fulfils requirements for tensor derivative (thus it makes sense to demand ∇, d acting like D on functions); multiple generalizations of Df to something like "Ds" for sections s exist, which is fine because from ∇ we can easily get many others e.g. by $d = \mathrm{Alt}(\nabla)$ (not sure if equality is true, but from Carroll eq 1.82 it looks like this), that is by suitable mappings

1.2.5 From Wald

the notion of curvature, intuitively, corresponds to the one of a 2-sphere in 3D space; however, this is extrinsic curvature which is only visible in embeddings, but what we are interested in is something like intrinsic curvature; how can we detect that?

1.2.6 From Penrose

tangent space in point $p \in M$ is immediate/infinitesimal vicinity of M "stretched out"; more formally, a linearisation of the manifold

to do physics, we cannot just work with vector spaces or affine spaces like the Euclidian space (basically $\mathbb{R}^n$, but no need to fix origin), but we need manifolds; however, manifolds do not have enough natural structure to build up the theory that is needed to describe physics, so we need some additional (local) structure (e.g. enabling us to measure infinitesimal distances in case of a metric structure); this structure is often encoded to/using the tangent spaces (which are present naturally for manifolds), which are vector spaces again

problem of abstract notion of "no structure" is for example: no general, meaningful (well-defined) notion of differentiation (does exist for functions, but not for vector fields, 1-forms or other tensors); exterior derivative is something like that, but it does not really give information about varying of the forms (nice is that it maps p-forms to $p + 1$-forms)

some structure can also be provided by connection; although not every structure can reproduced, metrics uniquely determine a connection (Levi-Civita connection)

goal of derivative operators: measure constancy and deviations from it; in case of vectors, this is equivalent to a notion of parallelism; note: we will go reverse route, define derivative and get parallelism from that; this notion will have the unusual feature of path-dependence, where unusual is meant with respect to what we know from Euclidian space; while it is possible to do this (see Wald), but this is mainly by making the "right" guess and thus not really helpful (idea is to say we want something where change of v is proportional to difference Δx and then we say: this works; welp)

which requirements make sense? since tangent space is linearisation of manifold, there should also be linear dependence on direction that we differentiate along; more generally, pointwise linearity means that functions can be dragged across the operator; when acting on tensors however, a product rule has to be specified: $\nabla_X(fs) = (\nabla_X f)s + f\nabla_X s$ makes sense (without argument X, this becomes $\nabla(fs) = (\nabla f) \otimes s + f\nabla s$) (?)

ideas come from the fact that our goal is to generalize action of derivative D; therefore, demanding $\nabla f = df$ also makes a lot of sense

extension to more than one tensor field is possible by demanding additivity $\nabla(s+t) = \nabla s + \nabla t$ and by specifying product/Leibniz rule $\nabla(s \otimes t) = (\nabla s) \otimes t + s \otimes (\nabla t)$; to uniquely determine this generalization, it is also necessary to demand compatibility with trace/contraction (which also helps with defining these things in the first place)

interesting: local connection can be defined uniquely from Gaussian basis vectors

2 Special Relativity

In modern day physics, there are often two competing viewpoints. One is very much based on intuition and the other is based solely on a mathematical description. This shows especially in the theory of special relativity, where one can deal (i) in much detail with groups and transformations or (ii) with a much more pictorial version of the theory, mostly utilizing very basic geometry in so-called spacetime diagrams.

Both approaches can lead to a rich and of course equivalent understanding, but it is often tempting to focus on only one of them. In my personal experience, this is often the mathematical description because students are often more familiar with the required math, so teaching the alternative and rather new intuitive-based approach would actually more complicated. This, however, can often lead to a lack of intuition, which is still fundamental to fully understand relativity as a whole. For this reason, our goal is to learn about both approaches.

2.1 Newtonian Physics

2.1.1 Space & Time

surely state that Newton is conform with our intuition and holds at small distances, speeds; for larger ones, however, inconsistencies show up

mention fictitious forces, e.g. Coriolis

Definition 2.1: Inertial Frame

An *inertial frame of reference* is a frame where $F^k = ma^k = m\ddot{r}^k$ holds.

Newtonian description also admits changing coordinates, via Galilei transform -¿ idea: for $a_1 = \frac{dv_1}{dt} = \frac{dv_2}{dt} = a_2$ as long as $v_1 - v_2 = const.$; therefore they show same physics in $F = ma$

2.1.2 Newton-Cartan Gravity

can also formulate Newtonian description in other mathematical framework, fibre bundles and connections; is more in line with geometrical formulations brought forward by Einstein

2.2 Relativity

2.2.1 What Is Wrong with Newton?

Newtonian theory works beautifully for many applications, even today where the theories of relativity and quantum mechanics are available. However, it does not describe the entirety of physics. This is also what physicists realized in the late 19th/early 20th century. At least to some degree, if not fully, every contradiction that surfaced at this time can be traced back to light.

In the 1860s, James Clerk Maxwell derived the *Maxwell equations* describing electromagnetic fields. They predicted electromagnetic waves, whose existence was confirmed in 1886 in experiments conducted by Heinrich Hertz. These waves propagate with the speed of light,

$$\boxed{c = \frac{1}{\sqrt{\epsilon_0\,\mu_0}} = 299\,792\,458\,\frac{\mathrm{m}}{\mathrm{s}}}, \tag{2.1}$$

so it was (and is) natural to identify them with light.

Several conflicts with Newtonian physics arise from this:

1. c is constant for all observers measuring it. The Galilei transform describing transformations in Newtonian theory, however, predicts different speeds. More specifically, light emitted by a moving observer is measured to have a speed of $c + v$ by a resting observer.

2. Maxwell's equations are not invariant under the Galilei transform. Instead, the corresponding symmetry transformations are modified versions of them (confer section 2.4 on those Lorentz transformations).

On the other hand, Newton's theory had tremendous success itself over many decades, so why would one conclude that this theory is at fault rather than the new Maxwell theory? Because (a) not all of Newtonian physics is to be replaced and (b) because of overwhelming experimental evidence. In addition to Hertz's experiments, the Michelson-Morley experiment in 1887 confirmed that c is constant for all observers to high precision.

For this reason, some new concepts are needed. While one can handle the subject strictly mathematically and e.g. simply derive the "correct" coordinate transformations that leave the Maxwell equations invariant, this tells us only little about the new physics that may arise. As we will learn, quite some rethinking of the concepts of space and time is needed to obtain equivalent results from a more intuitive approach and this work was mainly done by Einstein, e.g. in his famous paper "On the electrodynamics of moving bodies".

2.2.2 Einstein Postulates

Einstein's theory is based on two fundamental ideas, which are formulated in postulates.

The first idea is to make the invariance under chosen reference frame, which has been known a long time beforehand, a building block of the theory rather than a consequence.

> **Postulate 2.2: Principle of Relativity**
>
> All physical observations must hold independently of the inertial frame that is chosen.

This means if you are in a closed room without any windows, you cannot perform any experiment to determine if you are at rest or moving at a constant velocity. In other words, there is no absolute notion of rest; it's all relative (hence the name "relativity principle").

The second postulate concerns the speed of light and is definitely something new.

> **Postulate 2.3: Universality of Speed of Light**
>
> The vacuum speed of light c is constant for observers in all inertial frames.

Note that we do not demand "equivalent" statements like the relativity principle did – c has *the exact same* value for all observers. This postulate was based on the insights mentioned in the previous subsection, i.e. based on theory *and* experiment, and even more experiment evidence has been collected since.

However, it has profound implications. One of them concerns the very notion of space itself. In Newtonian physics, the notion of space is an absolute one. This means there is one reference frame that describes "real" space and while frames moving with respect to this one exhibit the same physical observations, this frame always has a special role. One way to characterize this observer is by measuring the speed of light sent out by him. If it is c and not differing by some amount, then the frame is at rest (otherwise it would be $c + v$).[1] However, since all observers measure c now, there is no way to determine this preferred observer at rest – the notion of rest/motion and therefore the connected notion of space itself becomes completely relative.

At the same time, this constance makes c very special because statements related to it are independent of the observer/inertial system and thus allow to make invariant statements. This property will be used routinely when properties in relativity are derived.

A direct corollary of postulate 2.3 is that c is the maximum speed at which *any* signal, not just light, can be transmitted.

[1] Note that we can assume the value of c to be known because it can be measured e.g. using a mirror by measuring the round-trip time. Strictly speaking, this only yields the two-way speed of light, but details on inferring one-way speeds require theory we are yet to built up.

Property 2.4: Universal Speed Limit

The speed of light c is the maximum velocity for all interactions.

The proof of this is very similar to the argument provided beforehand.[2]

Proof. If there was a speed c' higher than c, then for moving observers the speeds would be dependent on their velocity v and observed to be $c' + v$. Consequently, one could single out an observer sending signals at c' and this observer would be at rest – a violation of the relativity principle. □

This implication is a reason why experimental tests of this speed limit are important, they are confirmations of postulate 2.3. Such tests have been conducted successfully for neutrinos and gravitational waves, both of which do indeed propagate at the speed of light (to high experimental accuracy).[3] Interestingly, both the constancy of c and it being the maximum speed attainable by any signal/information implies that relativistic addition states

$$\boxed{c + c = c} \,. \tag{2.2}$$

As we will see later, this statement is indeed true.

2.2.3 Light Cones & Spacetime Diagrams

Before we start to evaluate implications of these two postulates, it is customary to introduce a technique to visualize the time evolution of physical systems. The most straightforward choice is to map position x on the x-axis and time t on the y-axis, i.e. a space-time diagram (which we will call *spacetime diagram* instead, for reasons that will become clear later). Since we are often interested in velocities close to c, it is customary to rescale the time-axis and display ct-values instead of t-values on it (otherwise slopes would be very small, small time step would correspond to large change in position).

In doing that, we discard two of the three spatial dimensions of Euclidian space. Nonetheless, it is sufficient to visualize important ideas (see figure 2.1 for a simple example).

- A single point in such a diagram gives position and time of an object (e.g. particle, person or rocket). We will refer to these points as *events* and use the symbol E.
- The trajectory of an object as a function of time (essentially a collection of events) is called *world line* Γ. An example for $v = 0.5c$ is shown as a red line in figure 2.1.

[2] The idea was taken from Dragon [**?**], but I tried to provide more details. However, I am not 100% sure about them being correct, so if you find an error to it, you might be right.

[3] This also means confirmations of other theories, which introduce new fields to circumvent the relativity principle and still allow speeds $> c$, cannot be confirmed.

- Light is described by the relation $x = ct$, which means it is always a diagonal at $\pm 45°$ (corresponds to slope ± 1), irrespective of the observer for which the diagram is drawn. This is why we often depict light sent out from and received by the origin (in both spatial directions), see the yellow lines in figure 2.1.

For the first time, we encounter what is often called *light cone structure* of relativity, which is essentially a corollary of c being a universal speed limit. This fact shapes the causal structure in relativity and for spacetime diagrams, it tells us that only events connected by straight world lines with slope $|v| \leq c$ can causally influence each other. This defines two cones above and below the event, the possible future and past of this event, each defined by events that the event can have interacted with or can potentially interact with. Together, these cones are called *light cone* (see figure 2.1 for an example). In terms of light cones, the structure of relativity can be states as follows.

Definition 2.5: Timelike, Lightlike, Spacelike

A straight world line Γ connecting events E_1, E_2 is called

- *timelike*, if it lies inside of E_1's and E_2's light cone ($|v| < c$)
- *null/lightlike*, if it lies on the edge of E_1's and E_2's light cone ($|v| = c$)
- *spacelike*, if it lies outside of E_1's and E_2's light cone ($|v| > c$)

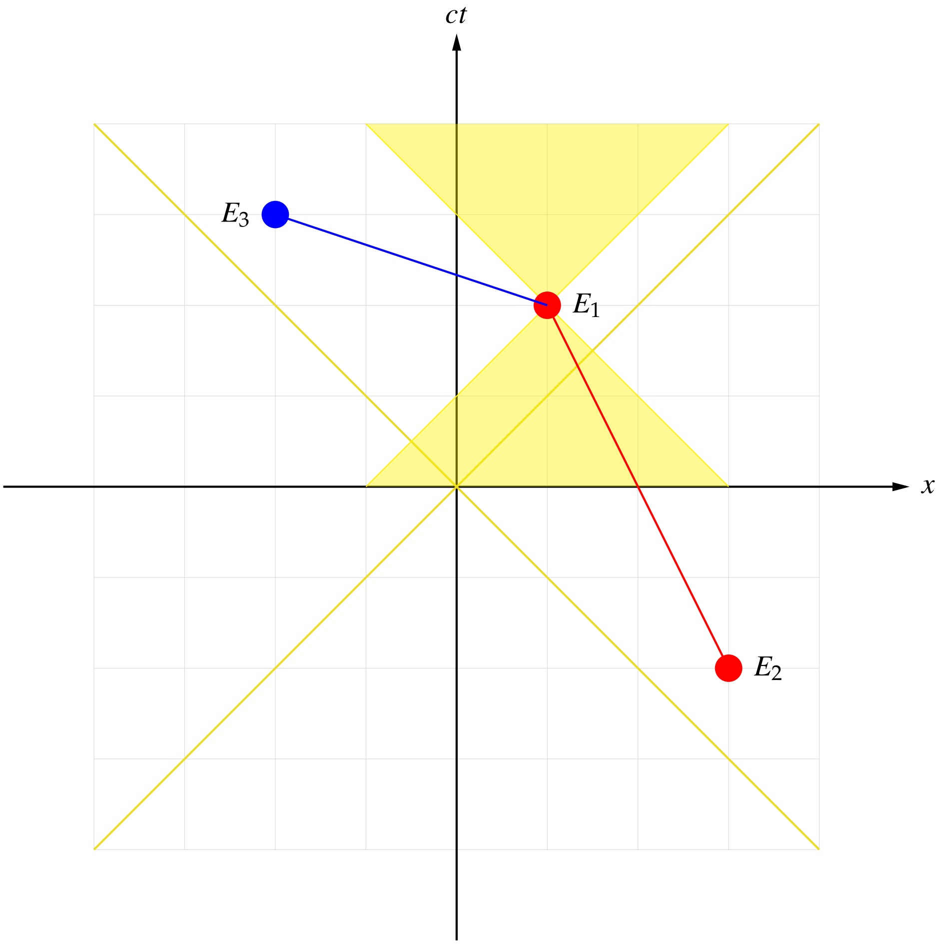


Figure 2.1: A simple spacetime diagram.
Red dots visualize three events E_1, E_2, E_3 at $(x, ct) = (1,2), (3,-2), (-2,3)$. For E_1, the light cone is drawn as well. Additionally, the world lines connecting E_1, E_2 (red; timelike) and E_1, E_3 (blue; spacelike) are shown.

2.3 Clocks

Until now, we have not really touched the notion of time. Partly, this is because we natively have a very clear, intuitive understanding of time: we look at clocks to measure it and this notion can be employed anywhere in space – just take a look at equal clocks in different points and compare their readings. This definition is employed in Newtonian physics, without much more attention being needed.

However, as it is the story for much of relativity theory, this notion essentially breaks down once we go to more extreme situations like distances on cosmic scales or clocks moving with respect to each other with high velocities. In both of those cases, timing the reading of a clock and hence comparing if they show the same time is difficult. For large distances, this is rather easy to see because information is transmitted at a finite speed $\leq c$, so when receiving information about the measurement result t of a far-away clock, we have to take into account the time it travelled to us in order to find out which event happened simultaneously to t.

This is problematic since many notions implicitly rely on the fact that we can measure quantities at the same time, i.e. on a notion of simultaneity. A very important example are lengths, which are defined as the separation of points – at a fixed time. Therefore, a well-defined notion of "at a fixed time" is required for us to be able to measure lengths and until now, we have no such notion. In everyday life, it is easy to avoid such difficulties: after all, we can look at clocks side-by-side, make sure they show the same time and then move one of them away to the desired position. This procedure ensures the clocks are synchronized, so we can simply take the desired measurements and compare the times later on. However, this is not really feasible to do that for measurements between planets or galaxies and clearly, an alternative, perhaps more general, way of communicating time measurements is needed.

All of that motivates the need for a synchronization procedure of clocks. We will here present the one proposed by Einstein, starting with its definition for resting observers and then look at it for the case of moving observers. Throughout this section, we will adopt visualizations from [**?**], while many of the definitions follow [**?**] more closely.

2.3.1 Synchrony of Clocks

Resting Observers Our setting is identical copies of an ideal clock being attached event/ point in space. For a consistent, well-defined notion of "time", however, we now have to make sure these clocks show equivalent times. To do that, we will synchronize them by adopting the following definition, originally proposed by Einstein.

Definition 2.6: Einstein Synchronization

Two clocks C, C' with times t, t' attached to observers O, O' at rest are *synchronized*, i.e. $t = t'$, if light signals sent out from them meet exactly in the midpoint of $\overline{OO'}$.

In principle, there is no unique procedure to synchronize clocks. However, in accordance with the relativity principle, it would be desirable to for the procedure to work independently of the chosen inertial frame. A very straightforward idea is to exploit the constancy of c and use light to communicate times between different observers and their respective clocks. That is what lead Einstein to this definition of synchrony.[4] It gives rise to the following notion of simultaneity.

Definition 2.7: Simultaneity

Two events E at t and E' at t' are called *simultaneous* if the locally simultaneous clock readings of synchronized clocks at these events are identical.

Essentially by definition, the following properties hold.

Property 2.8: Simultaneity as an Equivalence Relation

Simultaneity defines an equivalence relation on the set of all clocks in an inertial frame, i.e. the following properties hold:

1. Every event is simultaneous to itself.
2. If E is simultaneous to E', then E' is simultaneous to E.
3. If E is simultaneous to E' and E' is simultaneous to E'', then E is simultaneous to E''.

Moreover, a notion of simultaneity attached to some clock C partitions the set of all events $\{E\}$ into several, mutually disjoint subsets (equivalence classes), each containing events which are simultaneous to each other. A representative of the former is a family of synchronized clocks which show equivalent times.

Our main takeaway is the following: events that simultaneous for one observer O are simultaneous for every other observer O' at rest with respect to O.

This also justifies the following interpretation: two events E at t and E' at t' are simultaneous if light signals sent out by them reach the *referee* $\mathcal{R}$ at the same time. Here, the referee is a third observer, which is defined by the property that he always has an equal distance to the two other observers O, O'. This is meant in the sense that light signals sent out from $\mathcal{R}$ take the same time to go to O and come back as they do to go to O' and come back. Therefore, we reduce simultaneity for the spatially separated observers O, O' to simultaneity for a single observer $\mathcal{R}$, which is precisely the requirement stated in 2.7. More precisely, events E at t (measured by clock C) and E' at t' (measured by clock C') are simultaneous, i.e. $t = t'$, if they show the same time to the referee.

Besides governing which events happen at the time, the synchronization procedure further paves way to possibilities to determine times t' shown on a clock C' from times measured by another clock C. As figure 2.2 (a) shows, the travel time of a light signal from O to O' is

[4] As Giulini elaborates on in 2.1 of [**?**], this freedom in defining synchronization indeed exists.

the same as the travel time from O' to O. Using t' to denote the time light from O sees on C' when intersecting O' and using that C' is synchronized with C (which implies $t' = t$), equality of light travel times implies:

$$t - t_- = t_+ - t \qquad \Leftrightarrow \qquad \boxed{t = \frac{t_+ + t_-}{2}}. \tag{2.3}$$

Here, t_+ is the time on C when the first signal is sent out and t_+ the time on C when the second signal arrives. In just the same manner,

$$\boxed{t' = \frac{t'_+ + t'_-}{2}} \tag{2.4}$$

which should also be clear since $t'_- = t_-, t'_+ = t_+$.

In the definition of synchrony, we have assumed to be the observers O, O' to be at rest, which is also what 2.2 (a) represents. However, one requirement was that synchronization should also work in other inertial systems. This situation where O, O' are at rest with respect to each other, but move uniformly with respect to another observer, is shown in 2.2 (b).[5] Indeed, by constructing the referee $\mathcal{R}'$ and drawing the light signals sent to and from each observer (which form a rectangle, sometimes named *lightangle*), one obtains the times $t''_-, t''_+, t'''_-, t'''_+$ by prolonging the edges of the lightangle. This is equivalent to what has been done for resting observers in 2.2 (a) and likewise,

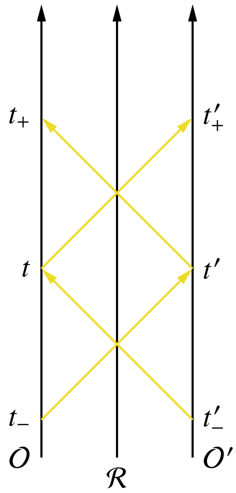
$$t'' = \frac{t''_+ + t''_2}{2} \qquad\qquad t''' = \frac{t'''_+ + t'''_-}{2}, \tag{2.5}$$

just as before, so the observers agree on $t'' = t'''$ (to verify that, we can also look at the distance of t_-, t, t_+ on the vertical axis we use to depict time). But nonetheless, something seems off. Figure 2.2 (c), where for simplicity we assumed that $t = t''$ in the event where O, O'' intersect, shows this more clearly. While the synchronization process still works, its induced notion of simultaneity for C'', C''' is *not* the same as the one for C, C'. Geometrically speaking, the "lines of simultaneity" (diagonal from left to right in lightangle) change from being horizontal in figure 2.2 (a) to being tilted in (b).

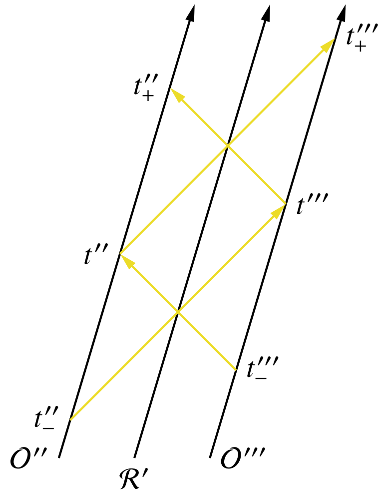
The origin of this difference is that the time of flight for light on $\overline{OO'}$ is different from the one on $\overline{O'O}$ (equivalent: time on $\overline{O\mathcal{R}}$ differs from the one on $\overline{\mathcal{R}O'}$) because the observers are moving uniformly in the same direction. Does that point to an inconsistency and thus error in the synchronization procedure? The answer is no because the roundtrip time is equal for $\overline{\mathcal{R}O\mathcal{R}}$ and $\overline{\mathcal{R}O'\mathcal{R}}$ – and this is the only distance/time that can be quantified using a single clock like the one at $\mathcal{R}$.[6] The events at t'' and t''' are said to be simultaneous because they do have the same roundtrip time of light signals, as one can verify in figure 2.2 (b).

[5] For a better distinction between the situations, the observers are named O'', O''' instead of O, O'. This does not change any interpretations that were mentioned.

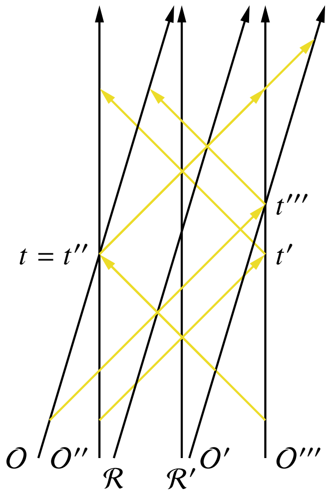
[6] To do that, synchronized clocks would be required. However, we wish to accomplish synchronization using the referee and the roundtrip time for him, so reasoning in this way does not work.



(**a**) Observers at rest



(**b**) Moving observers a rest with respect to each other



(**c**) Comparison of all observers from (a) and (b). For simplicity, we assume that the clock C and C'' show the same time $t = t''$ on intersection. Despite this fact, they cannot agree on notion of simultaneity.

Figure 2.2: Synchronization procedure in spacetime diagrams for different sets of observers. We draw the observers O, O' whose clocks we wish to synchronize, the referee $\mathcal{R}$ for them and the light pulses they exchange in order to achieve the synchronization.
Clocks carried by O and O' measure the events at times $t = \frac{t_+ + t_-}{2}$ and $t' = \frac{t'_+ + t'_-}{2}$ to be simultaneous, i.e. $t = t'$. This is because light (represented as yellow lines) sent out simultaneously by the referee $\mathcal{R}$ to O and O' "sees" times t and t' on the clocks before returning simultaneously to $\mathcal{R}$, where the results are "reported".

Moving Observers We have already seen how uniform movement changes the notion of simultaneity, even if the corresponding observers remain at rest with respect to each other (which means it is a relative notion, dependent on the motional state of the observer). Now, we will deal with the case where O and O' move with respect to each other with a relative velocity v (see figure 2.3 for examples of that).

This relative velocity results in a change of the distance $\overline{OO'}$ over time, making comparisons of clock readings much harder than before. One can prove the following theorem.

Theorem 2.9: Minkowski's Theorem

For two observers O, O' that move relative to each other with velocity v and an event E' occurring on the world line of O' at time τ',

$$\tau' = \sqrt{t_+ t_-} = \sqrt{1 - v^2/c^2}\, t\,. \tag{2.6}$$

Here, t_-, t_+ are the times measured by a synchronized clock on O where light signals to E' have been sent out and received back.

We assume that the synchronization when world lines of O, O' intersected and that clocks have been set to $t = 0 = t'$ there (although it is not necessary that this happened at $t = 0$, just at any time smaller than τ'). Just like before, we can visualize the synchronization process in terms of a referee $\mathcal{R}$ (figure 2.3). For him, the time passing between emitting light and receiving it back is equal for O and O', i.e. $\tau = \tau'$ for synchronized clocks and the corresponding events are simultaneous.

However, we have already seen that observers in other inertial frames can perceive simultaneity differently. Indeed, the travel time seen by O is $t_+ - t_-$. By the synchronization procedure we have employed, he assumes an equal light travel time and hence that the time τ light signals sent out by him "saw" is the same as the time t defined by

$$t - t_- = t_+ - t \qquad \Leftrightarrow \qquad t = \frac{t_+ + t_-}{2}\,,$$

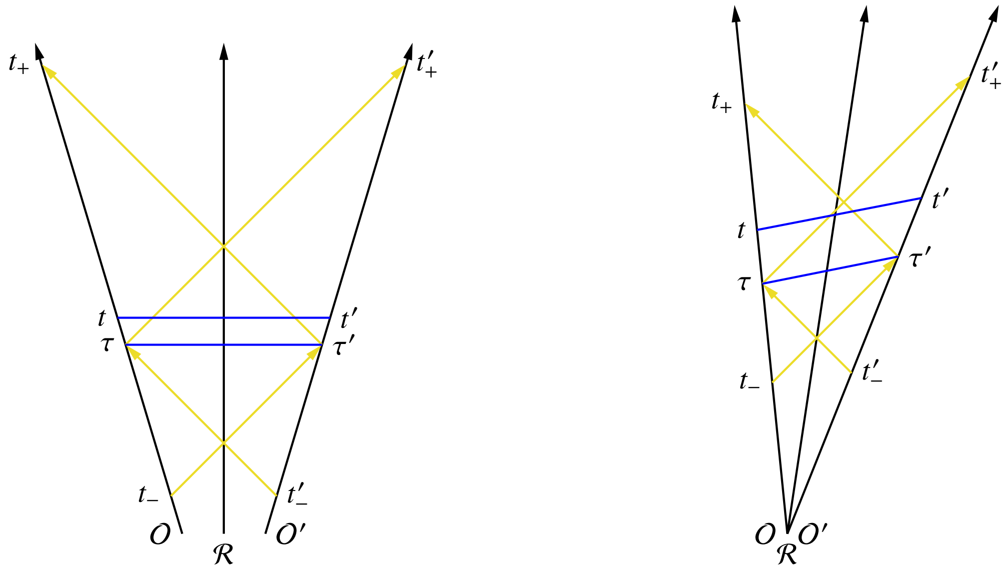
just like (2.3). However, theorem 2.9 shows that

$$\boxed{\tau' = \sqrt{t_+ t_-} = \sqrt{1 - v^2/c^2}\, t \leq t = \frac{t_+ + t_-}{2}}\,. \tag{2.7}$$

For the observer O, a time t has passed since the synchronization at $t = 0 = t'$, while for O' the *smaller* time τ' has passed – moving clocks tick slower.

This seems very puzzling. How does the situation look from the rest frame of O'? For a clock attached to O' the time $t' = \frac{t'_+ + t'_-}{2}$ has passed, while for one attached to O it is $\tau =$

[7] Beware that the diagrams are not perfect, e.g. the referee should be exactly where the light beams cross again. This is most likely due to an error in my code, which I was unable to locate. Nonetheless, the most important ideas should still be conveyed, which is why I decided to keep the graphics. -¿ ahhh, forgot that addition of velocities works differently in relativity; bug fix still not there, though



(**a**) Observers moving with relative velocity $v = 0.6c$

(**b**) Observers moving with relative velocity $v = 0.5c$

Figure 2.3: Synchronization procedure in spacetime diagrams for moving observers.[7]

We also note that it is really the relative velocity that matters because we could always go in the rest frame of one of the observers and examine the situation from there. Also, it does not matter which rest frame because the effect is mutual, i.e. it does not change if we replace $v \to -v$.

$\sqrt{t'_+ t'_-} = \sqrt{1 - v^2/c^2}\, t' \leq t'$ – the effect is mutual. That makes things even more puzzling and in fact seems paradoxical: how can both claim that less time went by for the other observer? The answer lies in the fact that in (special) relativity, there is no absolute, universal time anymore. Of course, if the time standard becomes relative, so do notions like simultaneity. As a consequence, both observers have the right to claim less time went by for the other and both are right in doing so. This is a key fact to understand in relativity and no different from the uniform motion of observers relative to each other. Both can claim to be at rest, while the other one moves because it does not change the physics exhibited by the situation.

The same observation can explain why the synchronization procedure for the referee, who does indeed see E' at τ' simultaneous to E at τ (which are the events and times light signals see which are sent out by him), does not yield the same notions of simultaneity for O, O': both of them are not at rest to $\mathcal{R}$ anymore and this differing motional state leads to the effects observed now. If this was the case, the relative velocity v would be $v = 0$ and the geometric mean $\sqrt{t_+ t_-}$ would reduce back to the arithmetic mean $\frac{t_+ + t_-}{2}$, yielding the familiar result

$$t = \tau = \tau' = t' \,. \tag{2.8}$$

Apparently, Minkowski's theorem is a more general version of the results (2.3), (2.4) and it reduces back to them in case of $v = 0$.

2.3.2 Time Dilation

The behaviour of time for moving observers seems puzzling at first, one may ask: is this a bug or a feature? In other words, is this effect that less time seems to pass on moving clocks a real physical effect or is it caused by our choice of synchronization? It is especially the mutuality of effect that may add to these doubts. However, as it was also argued in the last subsection, this phenomenon is fundamental *feature* of time. Resting observers agree on times elapsed along clocks carried by them, moving observers do not. Abandoning this idea of universal time may seem very unintuitive compared to what we experience in everyday life, but it turns out to hold true. Just like the answer to "where does event E happen?" depends on the coordinates we choose, the answer to "at which time did event E happen?" depends on the (world line of the) clock we use in our measurements. The slowing of time for moving observers is called *time dilation*.

There are multiple ways to see that this is not a consequence of our notion of simultaneity being flawed, causing the inconsistent result. A very striking one, at least in my personal opinion, is that what we have shown here as a synchronization procedure is equivalent to setting the same time on clocks and then moving them to different points in space (in the slow transport limit, such that no time dilation occurs). This is what we would do intuitively to synchronize clocks on Earth, which shows us that our unfamiliarity with these effects comes from our lack of experience with velocities $v \approx c$ or cosmic distances rather than pointing to unphysical effects.

Moreover, time dilation not only occurs for spatially separated events, but also for a scenario where the clocks are read off next to each other. It is commonly stated in the following form.

Example 2.10: Twin Paradox

It deals with two observers (commonly taken to be twins, i.e. of the same age), one of them at rest and the other one moving uniformly to some destination, for example Mars M. Both start at the same time and at the same point in space and meet again in this point at some later time. The interesting question is how much time has passed for both of them, after all we have just learned that uniform motion affects clocks. While in reality, this experiment would involve some kind of acceleration (turning around), but as Dragon points out in [**?**] this would not change the results because one can accelerate the resting observer in this time as well. Hence, any result we infer from instantaneous accelerations (and uniform movement in between) already contain the relevant ingredients and can be attributed to uniform motion. It should also be noted that certainly no general relativity is needed to explain the effects (neglecting any effect of gravity), contrary to what is claimed in many popular-science explanations.

From now on, O_R will be used to denote the resting observer and O_M to denote the moving observer. If O_R resting in Earth's orbit measures O_M to reach M at time t, then $\tau = \sqrt{1 - v^2/c^2}\, t$ has elapsed on a clock carried by O_M (v denotes the velocity of O_M). Immediately turning around at Mars and travelling back with velocity v' (could be $v' = -v$, but no need for that), then times t' and $\tau' = \sqrt{1 - v'^2/c^2}\, t'$ elapse for O_R and O_M, respectively. All in all, for the roundtrip times we have

$$t + t' = \frac{\tau}{\sqrt{1 - v^2/c^2}} + \frac{\tau'}{\sqrt{1 - v'^2/c^2}} \geq \tau + \tau' \,.$$

Furthermore, there is no ambiguity here in the perception of O_M because they end up witnessing the same event, $E = E'$, so there can be no ambiguity in simultaneity – the twin paradox is not paradoxical after all. This is demonstrated explicitly in the second treatment of this example using spacetime diagrams (2.11).

Therefore, it is indeed true that less time goes by on moving clocks. This effect has also been verified experimentally. It should be noted, though, that this effect only becomes relevant for velocities $v \approx c$, otherwise $\sqrt{1 - v^2/c^2} \approx 1$ and $\tau \approx t$. Once again, we emphasize that this is the reason why the Newtonian way of describing time as an absolute, universal notion has worked for so long and is perhaps also more intuitive.

An interesting consequence of time dilation arises for observers moving with the speed of light, like e.g. photons. No other light pulse sent from some distance can ever reach him, so how can this observer agree on time measurements with others? Well, they cannot agree because the speed of light is a universal limit, no signal can reach the moving observer, from which we conclude that no time passes for this observer. This result is in accordance with

$$\tau = \sqrt{1 - c^2/c^2}\, t = 0 \,. \tag{2.9}$$

2.3.3 Length Contraction

there is an induced effect on lengths, which are defined as distances measured *at the same time*; here we see how observers may measure different lengths, they have potentially different notions of simultaneity

-¿ moving rulers are shorter (again, mutual effect)

2.3.4 Doppler Shift

do it? Or maybe do in time dilation section? But I do think it is not 100% necessary, so should not be priority

ah, while $\kappa(-v) = 1/\kappa(v)$, this is not true for times: $\tau(-v) = \tau(v)$; thus it is not needed to explain time shifts, but instead it is useful for frequencies

Doppler factors could be helpful because frequencies are inverse time intervals, i.e. through arriving of light signals (thus blue-, redshift), one can argue for time intervals that the observers see as well! Due to longer travel time of light, light signals reach at different times, leading to different perceptions of simultaneity (Wikipedia on twin paradox is nice for that)

also helpful because from them, we can deduce how addition of velocities works; so maybe might be worthwhile to treat them after all... -¿ however, doing that in Lorentz transformation section might also be fine...

2.4 Lorentz Transformation

The effects of time dilation and length contraction point to new effects that occur upon changing inertial frames to look at problems. These effects are not predicted by the Galilei transform, which maps spatial coordinates between two inertial frames (one unprimed, one primed, moving relative to each other with velocity v) according to

$$x' = x - vt \qquad \Leftrightarrow \qquad x = x' + vt$$

and does not affect time at all, i.e.

$$t' = t\,.$$

Over the course of the last sections, however, we have seen that time is a notion which depends on the observer measuring it as well. Therefore, it is natural that it has to be transformed as well. Additionally, observers measure different speeds of light for this Galilei transform, a violation of the relativity principle since observers have to agree on c by postulate 2.3. The goal of this section is to find a new, corrected transformation between the unprimed and primed coordinates. Note that we will restrict ourselves to one spatial dimension x instead of three for now (a generalized treatment will be done in subsection 2.5.2).

Since we now that the Galilei transformation does work for velocities $v \ll c$, there is no need for a completely new expression. Our ansatz can simply be

$$x' = \gamma(x - vt) \qquad \text{and} \qquad x = \gamma(x' + vt') \tag{2.10}$$

for some factor $\gamma = \gamma(v)$, which fulfils $\gamma \approx 1$ for $v \ll c$. Note that we have used $\gamma(v) = \gamma(-v)$ here, which is the reason why $x \to x'$ has the same structure as $x' \to x$. This is a necessary condition on the transformation (and thus γ) since otherwise, there would be a preferred direction, which is forbidden by the relativity principle.

To obtain the way $t \to t'$ transforms, we use postulate 2.3 again, which tells us that $x = ct$ and $x' = ct'$ for light. We can rearrange this to read

$$ct' = x' = \gamma(x - vt) = \gamma(ct - \frac{v}{c}x) \qquad \text{and} \qquad ct = x = \gamma(x' + vt') = \gamma(ct' + \frac{v}{c}x')\,. \tag{2.11}$$

We are free to switch the expressions here since we are looking at light and it is necessary because $t' \overset{!}{=} t$ for $v = 0$.

The last step is determining the factor γ, which can be done by inserting (2.11) into (2.10)

$$\begin{aligned} x' &= \gamma(x - vt) = \gamma(\gamma(x' + vt') - v\gamma/c(ct' + \frac{v}{c}x')) \\ &= \gamma^2(x' + vt' - vt' - \frac{v^2}{c^2}x') = \gamma^2\left(1 - \frac{v^2}{c^2}\right)x' \end{aligned}$$

from which we obtain the *Lorentz-factor*

$$\boxed{\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}} \tag{2.12}$$

Inserting this result into (2.11) and (2.10) yields the *Lorentz transformation*

$$\boxed{x' = \gamma(x - vt) = \frac{x - vt}{\sqrt{1-\frac{v^2}{c^2}}}} \qquad \boxed{t' = \gamma(t - \frac{v}{c^2}x) = \frac{t - \frac{v}{c^2}x}{\sqrt{1-\frac{v^2}{c^2}}}} \tag{2.13a}$$

$$\boxed{x = \gamma(x' + vt') = \frac{x' + vt'}{\sqrt{1-\frac{v^2}{c^2}}}} \qquad \boxed{t = \gamma(t' + \frac{v}{c^2}x') = \frac{t' + \frac{v}{c^2}x'}{\sqrt{1-\frac{v^2}{c^2}}}}\,. \tag{2.13b}$$

Often, one abbreviates $\beta = \frac{v}{c}$, which is the velocity given in multiples of c (equal to v in units where $c = 1$). This set of equations shows how different inertial frames are related, respecting the the constancy of c.[8]

From its definition and figure 2.4, it is clear that γ diverges as v approaches c. The Lorentz transformation, on the other hand, does not diverge because of $x = ct$, which implies

$$x' = \left.\frac{c - v}{\frac{1}{c^2}\sqrt{c^2 - v^2}}\right|_{v=c} t = 0 \qquad\qquad t' = \left.\frac{c - v}{\frac{1}{c^2}\sqrt{c^2 - v^2}}\right|_{v=c} \frac{t}{c} = 0\,.$$

Hence, $x' = ct'$ still holds, as required. Them being equal to zero, as measured from the unprimed coordinates, simply comes from the fact that exchanging signals with light is impossible, so one cannot receive any information regarding times t' or positions x'.

The other limit we have to check is $v \ll c$, where we should re-obtain the Galilei transform. As figure 2.4 shows, γ exhibits the desired behaviour of $\gamma(v) \approx 1$, $v \ll c$. More quantitatively, a Taylor expansion in v/c shows

$$\begin{aligned}\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} &\simeq \left.\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\right|_{v=0} + \left.\frac{\frac{1}{2}\cdot\frac{-2v}{c}}{(1-\frac{v^2}{c^2})^{3/2}}\right|_{v=0} \frac{v}{c} + \ldots \\ &= 1 + \mathcal{O}(v^2/c^2)\,.\end{aligned}$$

Combining that with $v \ll c \ll c^2$, we obtain the desired result to zero-th order in v:

$$x' \simeq x - vt \qquad\qquad t' \simeq t\,.$$

[8]The inertial part is crucial here. It implies that the transformation have to be linear in x, ct since otherwise, $\frac{d^2x'}{dt^2} \neq \frac{d^2x}{dt^2}$ and the frames would not be inertial anymore.

One final note: the derivation presented here is perhaps not be the most rigorous one. Alternatively, one can argue using a diagram like 2.5, which is the spacetime diagram of a moving rod. Using a bit of geometry, one can show that the angle α between x, x' and the one between ct, ct' are equal. Moreover ct' is the line of equilocality of $x' = 0$, while in unprimed coordinates it is the line $x = vt$. [**?**] argues in detail how $\tan(\alpha) = \frac{v}{c}$ and more relations, which also leads to (2.13). Another approach is mentioned in subsection 2.5.2.

2.4.1 Time Dilation & Length Contraction from Lorentz Transformation

In principle, it is not surprising that a new transformation has to be used. After all, we have demanded some new postulates to be true and there is no reason for the "old" transformation to fulfil it. On the other hand, we have already examined quite a few predictions of these postulates in the clock section 2.3 and likewise, it is not guaranteed that our new Lorentz transformation reproduces the results which have been found there.

Most prominently, time dilation and length contraction showed up as new effects. We will now check if they can be reproduced by coordinates related via the Lorentz transformation. For that, we will look at a rod with ends at positions x_1, x_2. Their spatial distance and thus the length of the rod in unprimed coordinates is $L = x_2 - x_1$. For a ruler in the primed coordinates, however

$$L' = x_2' - x_1' = \gamma(x_2 - vt_2) - \gamma(x_1 - vt_1) = \gamma(x_2 - x_1) = \gamma L = \frac{L}{\sqrt{1 - \frac{v^2}{x^2}}} \,. \tag{2.14}$$

Here we assume $t_1 = t_2$ because by definition, distances are measured at the same time (simultaneously). This result is not what we expect from the previous discussions. Is the Lorentz transformation at fault? Luckily, no. We have emphasized how simultaneity is important for the length to be meaningful. However,

$$t_1' = \gamma(t_1 - \frac{v}{c^2}x_1) \neq \gamma(t_2 - \frac{v}{c^2}x_2) = t_2' \quad \Leftrightarrow \quad \Delta t' = t_2' - t_1' = \gamma\frac{v}{c^2}(x_1 - x_2) \,, \tag{2.15}$$

the measurements in primed coordinates are not simultaneous, so this L' is *not* what a ruler in primed coordinates measures. In order to simulate this result, we have to tune the times t_1, t_2 to achieve $t_1' = t_2'$. Choosing

$$\Delta t = t_2 - t_1 = -\Delta t' = -\gamma\frac{v}{c^2}(x_1 - x_2) = \frac{v}{c^2}\gamma L \tag{2.16}$$

leads to a cancellation of the time difference in (2.15). The "true" length measured in

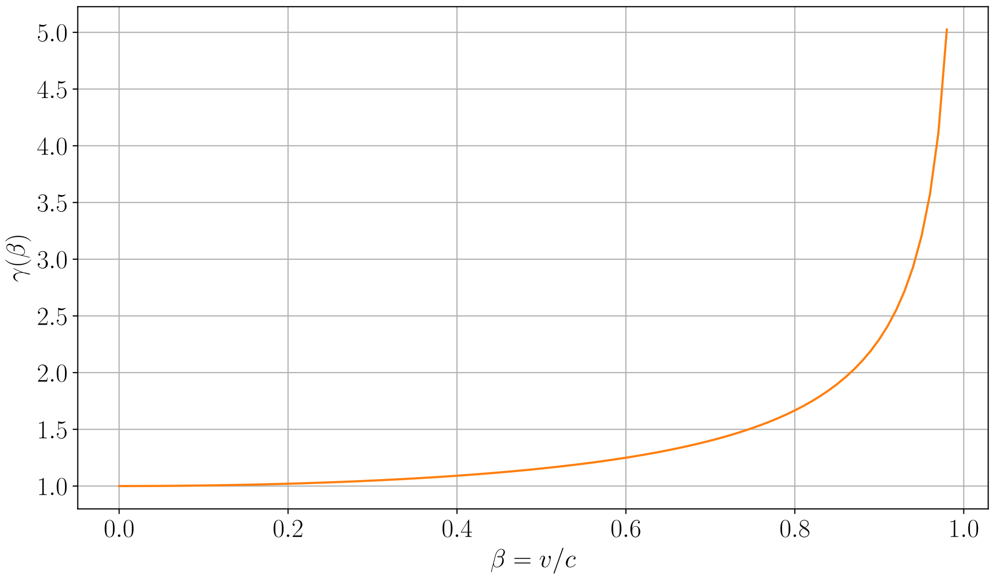


Figure 2.4: Plot of the Lorentz-factor $\gamma = \frac{1}{\sqrt{1-v^2/c^2}}$. Clearly, its effect only becomes significant for speeds which are significant fractions of c.

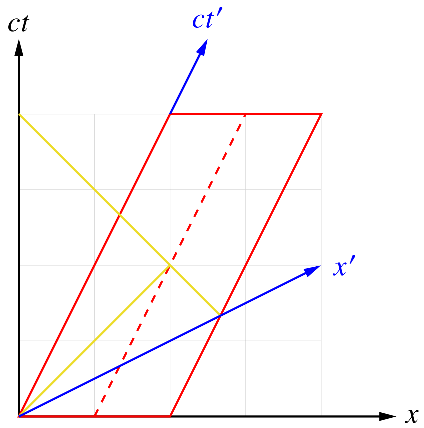


Figure 2.5: Spacetime diagram for a rod moving with velocity v, whose ends are depicted as red lines (midpoint using dashed line).
By definition, the axis ct' is the left end of the rod at $x' = 0$. In the (ct, x)-frame, it is given by the line $x = vt$. For the x'-axis, one can construct the point on the world line of the right end of the rod that is simultaneous to $t' = 0$ as measured by the left end of the rod. In accordance with the Einstein synchronization, this can be done by looking at light sent out from the origin and its intersection with the midpoint (dashed line). Tracing back this intersection point using light that is sent out by the right end of the rod in negative x-direction, one finds the point simultaneous to $t = 0 = t'$ in the origin (i.e. the line of simultaneity, which is nothing but x').

unprimed coordinates then becomes

$$\begin{aligned}\boxed{L'} &= x_2' - x_1' = \gamma(x_2 - vt_2) - \gamma(x_1 - vt_1) = \gamma L - v\Delta t \\ &= \gamma L - \frac{v^2}{c^2}\gamma L = \frac{1 - \frac{v^2}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} L = \sqrt{1 - \frac{v^2}{c^2}} L = \boxed{\frac{L}{\gamma}}.\end{aligned} \tag{2.17}$$

Moving rulers measure smaller distances, physics is still intact. Effectively, this corresponds to waiting a longer time until the second measurement in unprimed coordinates, which in turn causes the transformed times t_1', t_2' to be equal (note that the length L does not change because it is at rest in unprimed coordinates, i.e. x_1, x_2 are constant, independent of time). An intuitive explanation for this is that the moving observer claims the system at rest is moving, i.e. the end of the rod moves away from him, shortening it. In order to compensate for that, we have to wait a bit longer until taking the second measurement at t_2' and correspondingly, increase t_2.

Looking at temporal distances $T = t_2 - t_1$ measured by a clock resting in unprimed coordinates (i.e. $x_1 = x_2$), we obtain

$$\boxed{T'} = t_2' - t_1' = \gamma(t_2 - \frac{v}{c^2}x_2) - \gamma(t_1 - \frac{v}{c^2}x_1) = \gamma(t_2 - t_1) = \gamma T \boxed{= \frac{T}{\sqrt{1 - \frac{v^2}{c^2}}}}. \tag{2.18}$$

The primed observer measures a time T' between two events on his clock, while on the clock resting in unprimed coordinates, $T = \sqrt{1 - \frac{v^2}{c^2}} T'$ go by. Since the primed observer sees the unprimed one moving, that confirms what we already knew: moving clocks tick slower. This effect can be attributed to the synchronization of clocks. Despite $x_1 = x_2$, $x_1' \neq x_2'$, so the measurements in primed coordinates are taken by two different (but of course properly synchronized) clocks.

In the end, we can can confirm that both length contraction and time dilation are reproduced by the Lorentz transformation (the mutuality can be confirmed in analogous calculations). Furthermore, the calculations are much more straightforward compared to what had to be done previously (setting up synchronization, drawing complicated diagrams etc.). For this reason, it is customary to introduce the Lorentz transformation without covering any details on clocks. However, in my personal experience, that often results in a lack of intuition on these topics – one is bound to the understanding in terms of Lorentz transformations, although they are *not* necessary to understand what is going on. Since special relativity is a confusing topic in itself due to several frames and new concepts, intuition is an important part and helps tremendously with interpreting calculations.

2.4.2 Spacetime Diagrams 2

Even more help with understanding relativity is provided by visualizations like spacetime diagrams, which are especially helpful because one can visualize multiple observers in a single diagram (see e.g. figure 2.5 or figure 2.6, where grid lines are shown as well). Essentially, this is a geometric way of visualizing what Lorentz transformations do. By marking the coordinates (x, ct) of an event in a single, resting frame we can immediately read off the coordinates in other frames as well by showing the axes as it is done in 2.6. In principle, a Lorentz transformation shifts the points $(x, ct) = (0, ct)$ by $-vt$ to the left. This means the ct'-axis is rotated to the left compared to the ct-axis, so events would have different positions in diagrams. For this reason, the ct'-axis is rotated to the right, such that the position of events stays the same (an analogous argument can be made for the x'-axis). Reading off coordinates then works by looking at the transformed lines of simultaneity and equilocality (parallel to x'-axis and ct'-axis, respectively).

We have already discussed the causal structure of relativity, timelike trajectories are always in the light cone and spacelike ones are outside of it. For this reason, the events marked by red dots in figure 2.7 are clearly spacelike. Now we get a a visual explanation of why spacelike events are problematic: while the unprimed, resting observer sees the left event E_1 happening before the right one E_2, the primed observer sees E_2 happening before E_1. Thus, if c was not the speed limit and spacelike events were able to communicate with each other, observers could disagree on cause and effect. Luckily, no evidence for such a transmission with $v > c$ has been found (yet), so we do not have to rebuild our understanding of causality.

The new addition of being able to depict multiple observers to spacetime diagrams makes them a powerful tool suitable to explain many effects of relativity. For example, the fairly complicated twin paradox can be explained and – perhaps, even more important – visualized conveniently.

Example 2.11: Twin Paradox 2

As promised, here comes the detailed demonstration of the twin paradox, which has been started in example 2.10. We will discuss the setup shown in figure 2.8, i.e. treat one observer at rest (will be commonly referred to as "unprimed" one) and two observers moving with velocities $v = \pm 0.5c$ relative to the unprimed observer (these will be called "primed" and "double-primed", in accordance with their axis labels in 2.8).

Our approach will be to compute the roundtrip time needed to go from S to E (i.e. the time passing on the world line on ct axis) and the time needed to go from S to T to E (i.e. the time passing on the other world line shown in 2.8). Each of these quantities will be computed from clocks resting in all three of the inertial frames shown in figure 2.8 (reminder: three are involved due to turning around, rest frame of moving observer changes there).

During the process, we have to distinguish between four times: (i) the time t_{ST} passing on the resting clock between S and T, (ii) the time t_{TE} passing on the resting clock

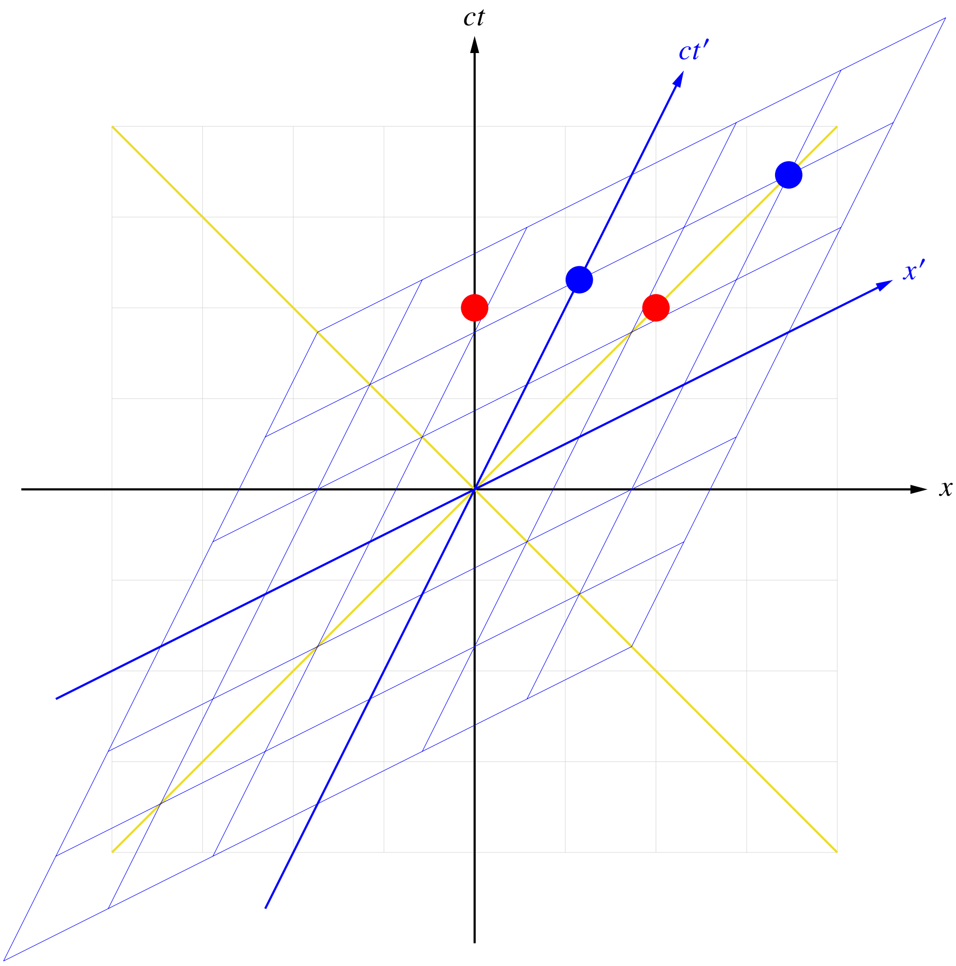


Figure 2.6: Events in frames moving with $v = 0.5c$ relative to each other. Red dots show two events at spacetime points $(x, ct) = (0, 2), (2, 2)$ (the corresponding primed coordinates are $(x', ct') = (-1.15, 2.3), (1.15, 1.15)$). Blue dots show the same coordinates in the (x', ct')-frame (corresponding unprimed coordinates: $(x, ct) = (1.15, 2.3), (3.46, 3.46)$).
We can see very nicely how each observer perceives time differently. Events happening simultaneously to both red dots (i.e. which lie on the line between them at $t = 2$), do *not* happen at $t' = t$, but at $t' = \tau = \sqrt{1 - v^2}\, t$ (which is evident from the fact that the blue dots are at $t' = t$). The same can be said for the moving observer in blue, which sees events at $t = \sqrt{1 - v^2}\, t'$ simultaneous to the blue dots at $t' = 2$. Analogous arguments can be made for positions and equilocality.

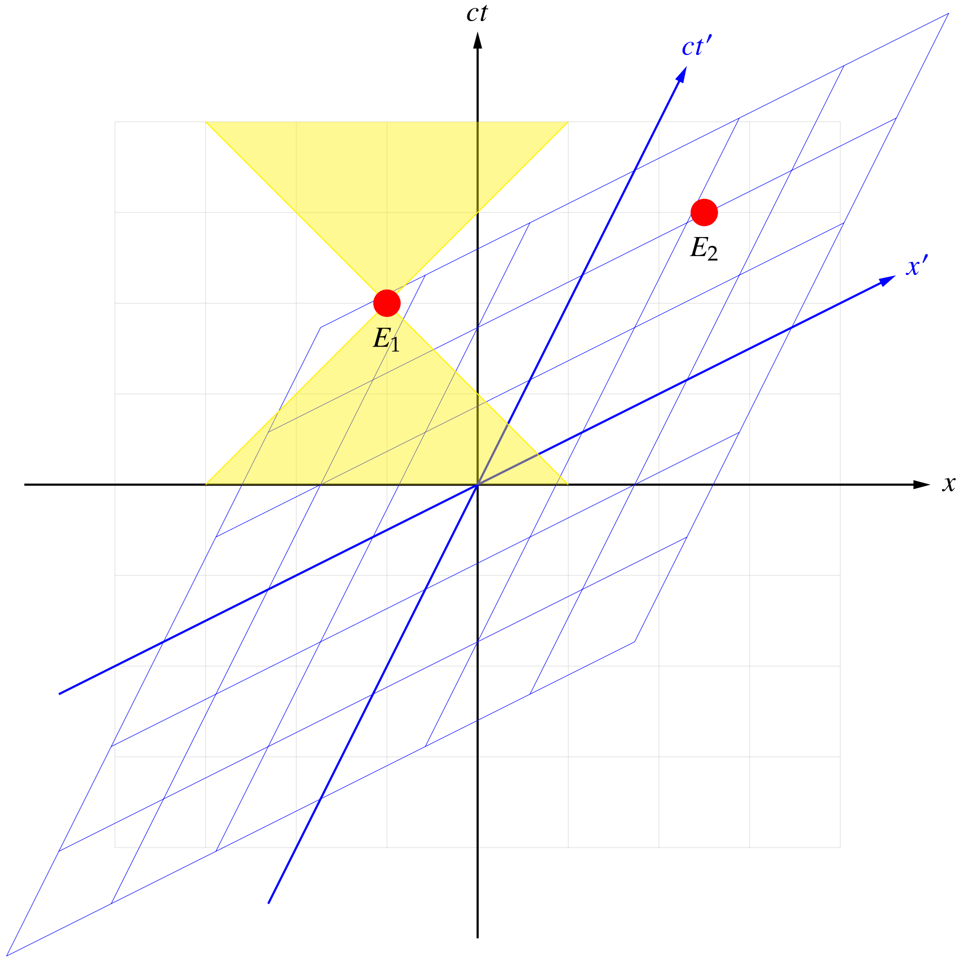


Figure 2.7: Spacelike events.
It is very obvious that E_2 does not lie in the light cone of E_1, which has been visualized to make that clear (this implies E_1 does not lie in the light cone of E_2). The unprimed, black observer sees E_1 happening before E_2, while the blue observer moving at $v = 0.5c$ sees it the other way around.

between T and E, (iii) the time τ_{ST} passing on the moving clock between S and T and (iv) the time τ_{TE} passing on the moving clock between T and E. From that we get the total times for resting and moving clock,

$$t_{SE} = t_{ST} + t_{TE} \qquad\qquad \tau_{SE} = \tau_{ST} + \tau_{TE} = t'_{ST} + t''_{TE} \,.$$

One final note concerns the velocities involved: the world line is drawn for $v = 0.5c$ on the way from S to T and $v = -0.5c$ on the way from T to E (same velocity, different direction), where v is the velocity of the respective moving frame compared to the unprimed, resting one. This implies a relative velocity of $v_2 = \frac{0.5c+0.5c}{1+0.5^2} = 0.8c$ from double-primed to primed frame (note that addition of velocities works differently in relativity).

- **Measuring from unprimed coordinates**

  Clearly,
  $$t_{ST} = 2 = t_{TE} \qquad \Leftrightarrow \qquad t_{SE} = t_{ST} + t_{TE} = 4$$
  in arbitrary time units (where one time unit goes by between two grid lines). From that, Minkowski's theorem tells us
  $$t'_{SE} = \sqrt{1-v^2}\, t_{SE} = 3.464 \,.$$
  Simultaneously, by looking at how much time elapses between the intersection of gray grid lines and the blue ct'-axis, one can see that as a rough estimate $t'_{ST} \lesssim 2$. For the exact result, we apply Minkowski's theorem:
  $$t'_{ST} = \tau_{ST} = \sqrt{1 - v^2/c^2}\, t_{ST} = \sqrt{1-0.5^2} \cdot 2 = 1.732 \,.$$
  Applying the same procedure to the double-primed coordinates yields
  $$t''_{TE} = \tau_{TE} = \sqrt{1 - v^2/c^2}\, t_{TE} = \sqrt{1-(-0.5)^2} \cdot 2 = 1.732 = \tau_{ST} \,.$$
  This is because time dilation does not depend on the direction, only on the absolute velocity. Therefore, we can confirm that
  $$t_{SE} = t_{ST} + t_{TE} = 4 > \tau_{SE} = \tau_{ST} + \tau_{TE} = 3.464 \,,$$
  less time goes by on the moving clock.

- **Measuring from primed coordinates**

  While from these coordinates one still sees four time units going by on the roundtrip for the resting observer, i.e. $t_{SE} = 4$, a rough estimate for the roundtrip time measured by a clock resting in the primed coordinates ct' is $t'_{SE} \gtrsim 4$. More precisely, Minkowski's theorem yields

$$t_{SE} = \sqrt{1 - v^2/c^2}\, t'_{SE} \qquad \Leftrightarrow \qquad t'_{SE} = \frac{t_{SE}}{\sqrt{1 - v^2/c^2}} = \frac{4}{\sqrt{1 - (-0.5)^2}} = 4.619\,.$$

This is a consequence from the mutuality of time dilation, an observer resting in primed coordinates sees the unprimed observer moving at $v = -0.5c$ and therefore measures more time passing in his own frame.

However, t'_{SE} is not what a clock in the primed coordinates sees. Instead,

$$\tau_{SE} = \tau_{ST} + \tau_{TE} = t'_{ST} + t''_{TE}$$

as stated before. For t'_{ST}, however, we cannot simply use Minkowski's theorem and thus the time

$$\frac{t_{ST}}{\sqrt{1 - v^2/c^2}} = 2.309\,,$$

which cannot be quite correct since from the diagram we get the estimate $t'_{ST} \lesssim 3$. This is because the world line is moving with respect to the rest frame of ct, so we would only get statements about what is the time as seen from this frame. However, we wish to measure the primed time from the primed coordinates. This requires a Lorentz transformation of the events S, T, E:

$$t'_{ST} = t'_T - t'_S = \frac{t_T - \frac{v}{c} x_T}{\sqrt{1 - v^2/c^2}} - \frac{t_S - \frac{v}{c} x_S}{\sqrt{1 - v^2/c^2}} = \frac{t_{ST} - \frac{v}{c} x_T}{\sqrt{1 - v^2/c^2}} = \frac{2 - 0.5 \cdot 1}{\sqrt{1 - 0.5^2}} = 1.732\,.$$

In the same manner, we obtain

$$t'_{TE} = t'_E - t'_T = \frac{t_E - \frac{v}{c} x_E}{\sqrt{1 - v^2/c^2}} - \frac{t_T - \frac{v}{c} x_T}{\sqrt{1 - v^2/c^2}} = \frac{t_{TE} + \frac{v}{c} x_T}{\sqrt{1 - v^2/c^2}} = \frac{2 + 0.5 \cdot 1}{\sqrt{1 - (-0.5)^2}} = 2.887\,.$$

To get t''_{TE}, we have to prolong the axes of simultaneity of the primed coordinates and count the number of time units which go by between the intersections of them with ct'' (i.e. the number of intersections with green lines on the way). This yields roughly $t''_{TE} \gtrsim 1.5$ again. For the time passing simultaneously on a clock in primed coordinates, we can read off roughly $t'_{TE} \approx 4$. The correct numbers can be obtained from Minkowski's theorem again:

$$t''_{TE} = \sqrt{1 - (-v_2)^2/c^2}\, t'_{TE} = \sqrt{1 - (-0.8)^2}\, 2.887 = 1.732\,.$$

All together, the primed observer measures a roundtrip time

$$t'_{SE} = t'_{ST} + t'_{TE} = 4.619 > \tau_{SE} = \tau_{ST} + \tau_{TE} = t'_{ST} + t''_{TE} = 3.464\,.$$

We find the same result that less time has passed on the moving clock. Furthermore, we see that the absolute value for this $\tau_{SE} = t'_{SE}$ and the one computed from

Minkowski's theorem in the first calculation agree (because corresponding world line is parallel to ct, i.e. in rest frame). On the other hand, it should not be surprising that the absolute values measured for τ_{ST} and τ_{TE} do not agree. After all, they are still measured from different inertial frames and they measure with respect to frames moving with different relative velocities (± 0.5 for primed; $0.5, 0.8$ for unprimed).

- **Measuring from double-primed coordinates**

  Just as before, $t_{SE} = 4$ is measured, while roughly $t''_{SE} \gtrsim 4$ and precisely

$$t''_{SE} = \frac{t_{SE}}{\sqrt{1 - v^2/c^2}} = \frac{4}{\sqrt{1 - (-0.5)^2}} = 4.619$$

  are measured by a clock resting in the double-primed coordinates. As one can confirm by looking at the result above, this is the same time a clock resting in primed coordinates measures. We expect more of these equal results since the double-primed coordinate system is moves with the same relative velocity as the primed one, just in the other direction (sign of velocity is different).

  A rough estimate for t''_{TE} is $t''_{TE} \gtrsim 2$ and the exact result is

$$t''_{TE} = t''_E - t''_T = \frac{t_E - \frac{v}{c} x_E}{\sqrt{1 - v^2/c^2}} - \frac{t_T - \frac{v}{c} x_T}{\sqrt{1 - v^2/c^2}} = \frac{t_{TE} + \frac{v}{c} x_T}{\sqrt{1 - v^2/c^2}} = \frac{2 - 0.5 \cdot 1}{\sqrt{1 - (-0.5)^2}} = 1.732\,.$$

  An analogous calculation yields

$$t''_{ST} = t''_T - t''_S = \frac{t_T - \frac{v}{c} x_T}{\sqrt{1 - v^2/c^2}} - \frac{t_S - \frac{v}{c} x_S}{\sqrt{1 - v^2/c^2}} = \frac{t_{ST} - \frac{v}{c} x_T}{\sqrt{1 - v^2/c^2}} = \frac{2 + 0.5 \cdot 1}{\sqrt{1 - 0.5^2}} = 2.887\,.$$

  As promised, we get more results that we have already seen when making computations in the primed coordinates, but now they are switched (due to the transition $v \longrightarrow -v$).

  Now we wish to compute t'_{ST} and expect the rough estimate $t'_{ST} \gtrsim 1.5$. Indeed, we obtain

$$t'_{ST} = \sqrt{1 - v_2^2/c^2}\, t'_{TE} = \sqrt{1 - 0.8^2}\, 2.887 = 1.732\,.$$

  All in all, the double-primed observer measures a roundtrip time

$$t''_{SE} = t''_{ST} + t''_{TE} = 4.619 > \tau_{SE} = \tau_{ST} + \tau_{TE} = t'_{ST} + t''_{TE} = 3.464\,.$$

  Once again, these results look very familiar.

A conclusion of these extensive calculations is that physics is not broken, despite relativity sometimes being unintuitive at first glance. Although all inertial frames play equal roles, which shows in the mutual slowing of moving clocks, all agree on the effect of time dilation: moving clocks tick slower than resting ones. In frames where both clocks seem to be moving, we can further confirm that the effect increases with the velocity v.

It should also be noted that the agreement of all three observers regarding τ_{SE} is really a coincidence due to the symmetric setup we have chosen – for other scenarios, e.g. with unequal velocities on the first and second part of the journey or perhaps even if we assume that ct is not at rest after all (but keeping the relative velocities of primed and unprimed system, i.e. rotating the whole setup), this will not be the case anymore. In the same manner, $t'_{ST} \neq t''_{TE}$ in general. However, if one of the clocks moves on a world line parallel to the ct-axis of one of the observers, *all* of them will agree on the time elapsed along this clock (so t_{SE} being equal for all observers is really not a coincidence; the same goes for t'_{ST} and t''_{TE} in this setup). This is despite observers measuring different times on their own clocks and is due to Minkowski's theorem, which tells us how much time has passed simultaneously on a clock in another frame.

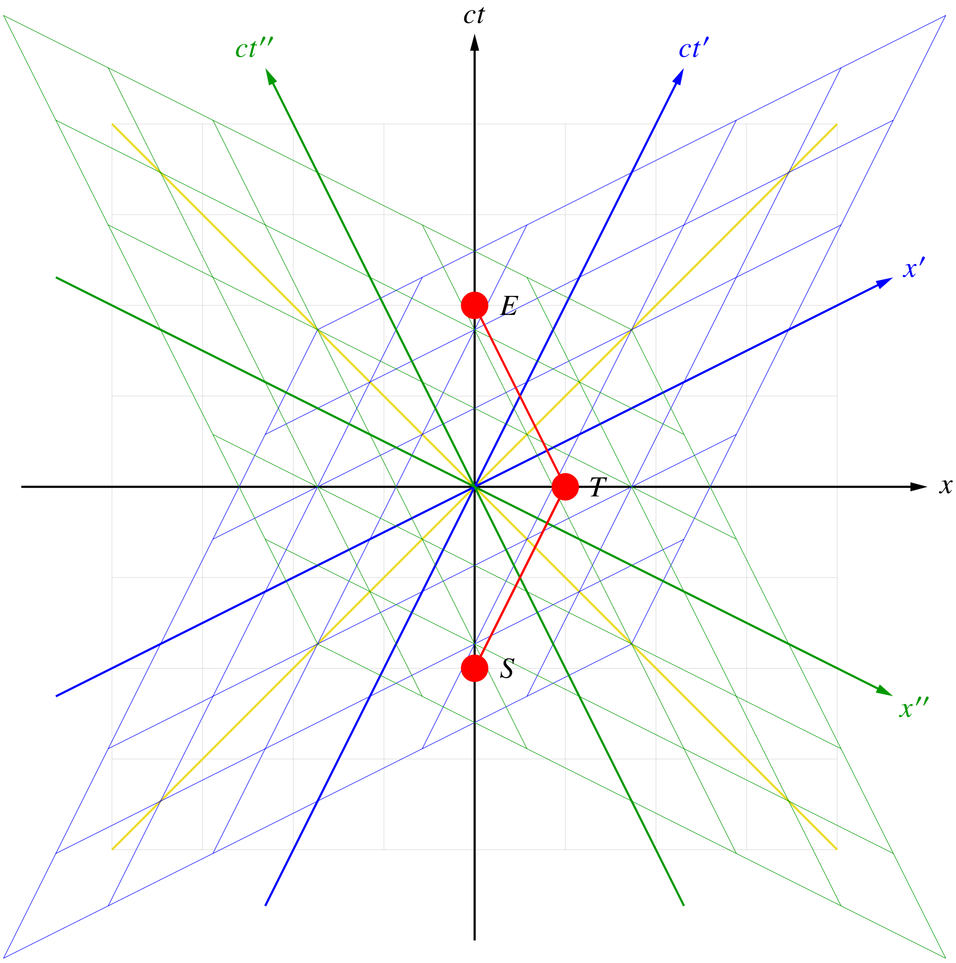


Figure 2.8: Visual explanation of twin paradox. Event S is the starting point, T turning point and E end point. The primed observer has a relative velocity of $v = 0.5v$ with respect to the unprimed one, while the double-primed observer has $v = -0.5c$. We can measure times passing between events in a certain frame O by prolonging the corresponding lines of simultaneity (parallel to spatial axis of O) from event to the time axis of O and then count the number of time units passing on the time axis between the intersection points. Similarly, we could also prolong them until they intersect the time axis of any other frame O'. In this case, counting the number of time steps passing between the intersection points on the time axis of O' would yield the time passing for O', measured from O.

2.5 Minkowski Space

Throughout the last sections, we discovered more and more how space and time work in relativity and how they are related. Important contributions to that picture were made by the insights of Einstein regarding synchronization of clocks and Lorentz, who developed a corrected version of the Galilei transform.

It might be clear to the reader that this implies space and time are not independent anymore, but instead have to be treated on the same footing. Historically speaking, however, this final step was not made until Minkowski proposed his viewpoint that physics should take place in a four-dimensional *spacetime*. This unification is an essential part of how the theory of relativity is described in modern literature, in particular because it allows a description in terms of a well-developed mathematical theory – the theory of manifolds. We will also adopt the usage from now on.

First, it is necessary to state what spacetime is in a formal, mathematical way.

Definition 2.12: Minkowski Space

In special relativity, *Spacetime* is described as a 4-manifold M with one time and three spatial coordinates. The corresponding tangent spaces $T_p M$ are called *Minkowski space*.

4-manifold, points called events; coordinates usually denoted as $\xi = \begin{pmatrix} ct \\ \vec{x} \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$

events can then be described using 4-vectors $\xi = \begin{pmatrix} ct \\ \vec{x} \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$ (called *event* from now on)

-¿ but that is point, right? Not vector; we choose to scale time such that all components have the same units; factors of c are kind of nasty in relativity, which is why very often c is set to 1 (this way, first coordinate is simply time but scaling and computations are still convenient)

we choose convenient scaling again for time coordinate -¿ important: here, we have Cartesian coordinates (but we can easily change; just for definition we use Cartesian) -¿ adopt name inertial Cartesian? I kind of like it...

Lorentz- or pseudo-Riemannian manigfold

physical way to say things: we want observer-independent statements, for example on spatial distances and time; mathematical: coordinate-independent, invariant notions like distance and metric -¿ this is why language of manifolds is so convenient: its underlying, mathematical ideas are conceptually equivalent to key physical ideas of relativity

events are now spacetime points, world lines are collections of events and thus curves (can usually be parametrized; parameter with specific normalization is called proper time) introduce four-velocity as tangent vector; then also state that as of now, we are not able to take norm of it -¿ motivates metric section

just like coordinate planes with $x-, y-$axis visualize Euclidian space, spacetime diagrams visualize spacetime; although often time-axis versus only one spatial axis; for this reason, they are also commonly called *Minkowski diagrams*

2.5.1 Metric & Inner Product

We have now seen how physics can be conveniently described using a 4D manifold, which we called spacetime. Points on this manifold are events and we can change coordinates or inertial frames using Lorentz-transformations. Moreover, there are several quantities that can be defined naturally on manifolds, for example curves, vectors, and covectors (maps that take vectors as input). While manifolds do have an additional natural structure, this is given by topology. In physics, however, we are also interested in statements concerned with distances between events and to measure them we need additional structure. More specifically, we have to specify a metric that will allow measuring distances, as well as norms of vectors via the induced inner product.

Mathematically, metrics are objects called tensors and they have the convenient property that they are invariant under coordinate changes. Therefore, distances are physically meaningful statements because they do not depend on the inertial frame we compute them on. In the tradition of invariant quantities that have been encountered so far, we may guess that the metric will be related to light in some way. From the universality of the speed if light c, distances s are equivalent to times t for light, $s = ct$. Because of that, a natural measure for distances is the time elapsed on a clock, i.e. the geometric structure of Minkowski space is determined by Minkowski's theorem 2.9. Instead of denoting time with the usual variable t, we will now switch to the *proper time* τ since the time elapsed a clock between events $(0, 0, 0, 0)$ and (ct, x, y, z) is

$$\boxed{\tau = \sqrt{1 - \frac{v^2}{c^2}}\, t = \sqrt{1 - \left(\frac{x}{ct}\right)^2 - \left(\frac{y}{ct}\right)^2 - \left(\frac{z}{ct}\right)^2}\, t = \sqrt{t^2 - (x^2 + y^2 + z^2)/c^2}} \tag{2.19}$$

This distance notion depends on the trajectory taken by the clock/corresponding observer (more specifically, on the uniform velocity v), but will in general not be equal to t, which is the time measured simultaneously by a clock resting in the corresponding frame. This does *not* mean τ depends on the coordinates we use to compute it, i.e. the specific inertial frame chosen, the striking factor is the path taken by the clock we measure time for.

Example 2.13: Proper Time vs. Coordinate Time

We will use an example to elaborate a bit more on the meaning of all the symbols in (2.19). Say we are in an inertial frame with coordinates (ct, x, y, z).

The time elapsed between two events A clock resting in this frame will measure the proper time between events $E_1 = (ct_1, 0, 0, 0), E_2 = (ct_2, 0, 0, 0)$, i.e. it will show

$$\tau_{E_2,E_1} = t_2 - t_1$$

to be elapsed between them. If we look at events $E_3 = (ct_1, x_1, y_1, z_1), E_4 = (ct_2, x_2, y_2, z_2)$, however, it will still measure $t_2 - t_1$. This is not equal to the proper time

$$\tau_{E_3,E_4} = \sqrt{(t_2 - t_1)^2 - ((x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2)/c}$$

between these events, which would be measured by a clock moving on the straight world line that connects them.

Going to a frame with coordinates (ct', x', y', z'), which moves uniformly between E_3 and E_4, the situation is different. A clock resting in this frame moves on a trajectory between E_3 and E_4, which means the time t' measured by it now coincides with the proper time between these two events,

$$\tau_{E_3,E_4} = t'_2 - t'_1 \,.$$

This is because the spatial coordinates of E_3 and E_4 are equal in the primed frame, the Lorentz transformation automatically incorporates all spatial movement happening in unprimed coordinates into t'. However,

$$t'_2 - t'_1 = \tau_{E_1,E_2} = \sqrt{(t'_2 - t'_1)^2 - ((x'_2 - x'_1)^2 + (y'_2 - y'_1)^2 + (z'_2 - z'_1)^2)/c}$$

because the trajectory connecting E_1 and E_2 is not parallel to ct'. This is the mutuality of time dilation. Nonetheless, it shows a general way how different observers can agree on times: by using the proper time between events. Not only do the numbers agree in this case, conceptually it also makes a lot of sense to look at times which are measured by clocks which actually "see" both events, i.e. move on a world line connecting them, instead of using clocks far away from the event.

This whole behaviour might seem familiar, we have already encountered it in example 2.11, where the twin paradox has been discussed in detail. Here it already showed how different observers do not necessarily agree on times t their own clocks measure between two events, but they do agree on times they infer to be measured by a clock moving on a trajectory connecting these two events – the proper time. Moreover, since all observers agree on the proper time, one can immediately infer effects like time dilation, which becomes as easy as

$$t'_2 - t'_1 = \tau'_{E_3,E_4} = \tau_{E_3,E_4} < t_2 - t_1 \,,$$

the primed, moving observers measures smaller times than the unprimed, resting one.

As we have seen in this example, in some coordinates it is very easy to compute proper times because the object or particle we look at is at rest in this frame, which implies that the coordinate time t is already the proper time, $\tau = t$. These coordinates are often given a special name.

Definition 2.14: Instantaneous Rest Frame

A frame (t, x, y, z) where

$$\boxed{dx = dy = dz = 0 \qquad \Rightarrow \qquad d\tau = dt} \tag{2.20}$$

along the world line of a particle is called *instantaneous rest frame* or *comoving frame*.

This definition can also be extended to non-uniform velocities $v = v(t)$ because instantaneous rest frames taken at different times are related by Lorentz transforms. We will now see how the general definition of the proper time may be extended to this case.

Generalized Proper Time What was denoted with τ in equation (2.19), in reality is a difference $\Delta\tau$ of proper times (just measured with respect to $\tau = 0$), and the same is true for the coordinates (ct, x, y, z). Making these differences infinitesimally small, i.e. $\Delta \to d$, we obtain the infinitesimal distance or *proper time element*[9]

$$\boxed{d\tau^2 = dt^2 - \frac{dx^2 + dy^2 + dz^2}{c^2} = (1 - \frac{v^2}{c^2})\, dt^2 = \gamma^2\, dt^2}\,. \tag{2.21}$$

This implies the following *(proper) line element*

$$\boxed{ds^2 = c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2} \tag{2.22}$$

and corresponding *proper distance* $s = c\tau$ between events (note that this implicitly uses the constancy of c again). Writing this line element out in its general form $ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu$ immediately allows to read off the components of the *Minkowski metric* η, which can be conveniently arranged in a matrix

$$\boxed{\eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}}\,. \tag{2.23}$$

At this point, we shall make a remark regarding the signs: in principle, we could have chosen η to have only one minus (in the time component) since only the proper time is constrained by physical observations, while the proper distance could equally well be defined as $ds^2 = -c^2 d\tau^2$. There are good arguments for both conventions, but I personally find it more natural to keep the signs when going from times to distances, which is why the $(+, -, -, -)$-convention is adopted here.

[9] We adopt the common notation $dx^2 := (dx)^2$.

Perhaps a bit contrary to popular belief (at least that is what many authors claim), the scope special relativity is not restricted to observers with uniform velocities. Using some of the quantities and tools we have just derived, it is possible to extend the description of certain dynamics to observers moving with non-uniform velocities, i.e. accelerating ones. For distances, this is possible in the scope of integration on manifolds, where the length of a curve $\Gamma(t)$ defined on an interval $I \subset \mathbb{R}$ is

$$L(\Gamma) = \int_I \sqrt{g_{\mu\nu} v^\mu v^\nu}\, dt = \int_I \sqrt{g(v,v)}\, dt =: \int_I d\Gamma \, .$$

$v = \frac{d\Gamma(t)}{dt}$ is the tangent vector (field) along Γ. Because we are equipped with a metric $g_{\mu\nu} = \eta_{\mu\nu}$ and corresponding line element $d\tau$, we can compute proper times for arbitrary kinds of movements.

Postulate 2.15: Clock Postulate

Given a world line Γ parametrized by $\sigma \in I = [a,b]$, i.e.

$$\Gamma(\sigma) : \sigma \mapsto (t,x,y,z) = (t(\sigma), x(\sigma), y(\sigma), z(\sigma)) \, ,$$

the proper time elapsed along Γ is

$$\boxed{\begin{aligned} \tau &= \int_\Gamma d\tau = \int_\Gamma \gamma(t)\, dt = \int_\Gamma \sqrt{1 - \frac{v(t)^2}{c^2}}\, dt \\ &= \int_a^b \frac{d\tau}{d\sigma}\, d\sigma = \int_a^b \sqrt{\left(\frac{dt}{d\sigma}\right)^2 - \frac{1}{c^2}\frac{dx^\alpha}{d\sigma}\frac{dx^\alpha}{d\sigma}}\, d\sigma \end{aligned}} \, . \tag{2.24}$$

Note that while this "derivation" we provided here does make a lot sense, there is no guarantee for it to be correct – accelerating clocks could be vastly different from uniformly moving and resting ones. For this reason, we call (2.24) a postulate rather than a property. Once again, experiments have tested this postulate to very high accelerations of $\approx 10^{16}$ times the acceleration on earth, which showed no dependence on it (only on the speed v).

Intuitively, we can see that (2.24) works because it utilizes infinitesimal steps $d\tau$ where $v(t)$ does not change, so we can apply what we know at this point and integrate up the results from all points (idea is similar to the rectification of curves). It includes the case of uniform movement v = const., whence the integrand γ is constant and evaluation of the integral simply yields $\tau = \gamma(t_b - t_a)$. In case $v = v(t)$, the most convenient formula to use from (2.24) really depends on what is given – it may be the parametrization that is explicitly known or the velocity as a function of time.

Instead of σ, one could also choose arbitrary linear combinations $\sigma' = e\sigma + f$ $(e, f \in \mathbb{R})$ to parametrize Γ and then use $d\sigma' = e\, d\sigma$, which means that the integral in (2.24) is invariant under changes of this *affine parameter*. A very common choice is $\sigma = \tau$ and we will later see how this specific parametrization can be characterized.

Another remark on this definition is related to the metric. In coordinates, which are not inertial Cartesian ones, the metric is very likely to have different components $\eta'_{\mu\nu} \neq \eta_{\mu\nu}$. In this case, the last equality would not hold true anymore, although analogous formulas can be obtained from (again, assuming a world line Γ with tangent vector $\underline{v}$)

$$\boxed{\tau = \int_\Gamma d\tau = \frac{1}{c} \int_I \sqrt{\eta'_{\mu\nu} v^\mu v^\nu} \, d\sigma}\,, \tag{2.25}$$

which still holds due to the transformation rule for integrals. This rule is also what implies the invariance of τ under the coordinates/frame it is computed in, which is a mathematical way of stating that all observers agree on the proper time.

It is this invariance that allows us to infer statements about the geometry of Minkowski space (where we now assume $v = \text{const.}$ again and thus use (2.19)). In Euclidian space, points of constant distance s lie on a circle around the origin, which is determined by the equation $s^2 = x^2 + y^2 + z^2$. In Minkowski space, events of equal distance lie on a hyperboloid of constant proper times τ and are determined by

$$\boxed{c^2\tau^2 = c^2t^2 - x^2 - y^2 - z^2}\,. \tag{2.26}$$

Setting $c\tau = 1$ yields a hyperbola, whose intersection with time axes in spacetime diagrams determines the "length of one time unit", a very convenient way to see how moving clocks are perceived to be slower from a resting observers point of view (figure 2.9).

One final note concerns an alternative derivation of the metric. It is also based c being constant, but instead of constructing clocks etc. explicitly, it uses that light propagates as a spherical wave with velocity c. Writing this equation in multiple inertial frames yields

$$c^2t^2 = x^2+y^2+z^2 \;\text{ and }\; c^2t'^2 = x'^2+y'^2+z'^2 \qquad \Leftrightarrow \qquad c^2t^2-x^2-y^2-z^2 \overset{!}{=} c^2t'^2-x'^2-y'^2-z'^2\,. \tag{2.27}$$

by the relativity principle. This points to the invariant proper distance we have called s (which vanishes for light). At the same time, it explains the ambiguity in overall sign because we can get the same physics no matter which sign is used in ds^2 (only the sign in $d\tau^2$ is fixed, but this can be adjusted by defining $ds^2 = \pm c^2 d\tau^2$, respectively).

Inner Product The notion of a metric allows for the construction of a rich theory. We have seen how it can be used to define integrals and now we will deal with another important structure on manifolds, which has already been used in (2.25).

Definition 2.16: Minkowski Inner Product

The *Minkowski inner product* of two vectors $\underline{v}, \underline{w}$ is

$$\boxed{\underline{v} \cdot \underline{w} := \eta(\underline{v}, \underline{w}) = \eta_{\mu\nu} dx^\mu(\underline{v}) dx^\nu(\underline{w}) = \eta_{\mu\nu} v^\mu w^\nu}\,. \tag{2.28}$$

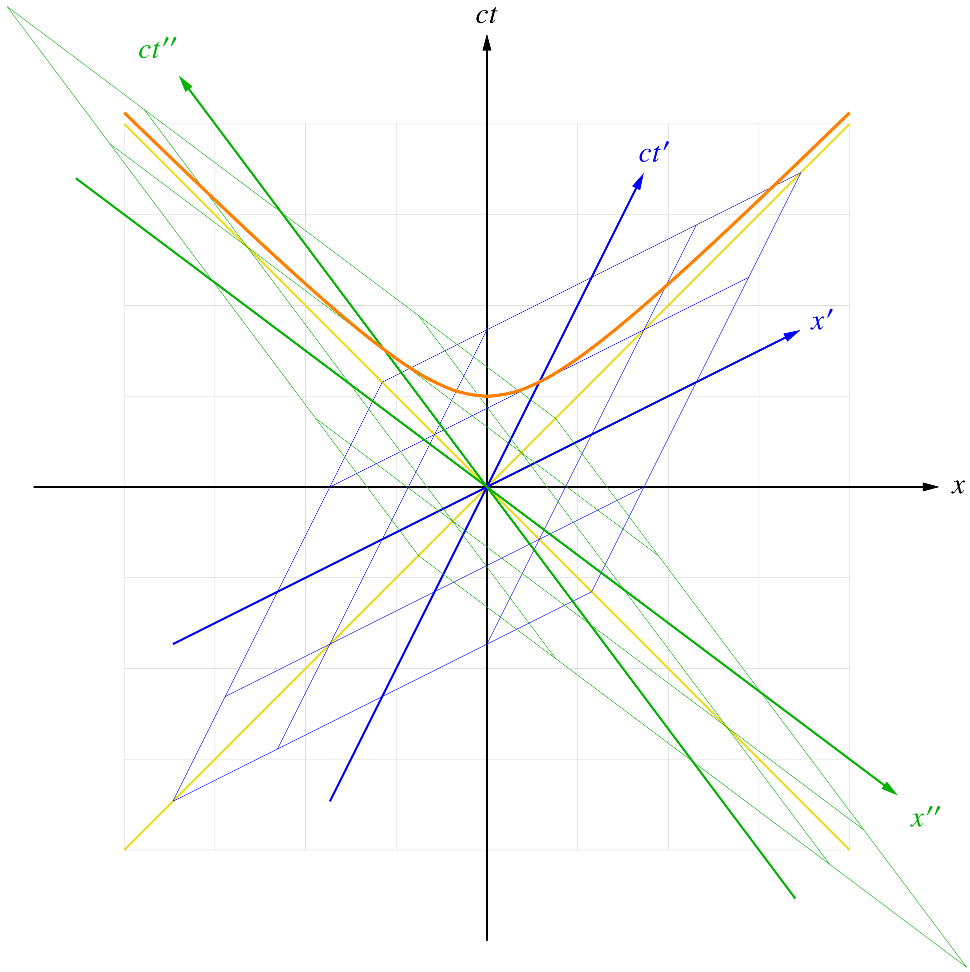


Figure 2.9: Plot of the curve fulfilling $1 = c^2t^2 - x^2$ (orange), along with three inertial frames (rest frame in black and frames moving with $v = 0.5c, -0.75$ relative to resting one in blue, green, respectively).
As we can see, this yields same time steps as Lorentz transform, the hyperbola intersects time-axes of all observers perfectly at time unit on their respective axis (as it should).

Strictly mathematically speaking, this does not define an inner product because it is not positive-definite, "only" non-degenerate. For this reason, η is also called a pseudo-Riemannian metric. For simplicity (in typical physicist-manner) it is commonly called Minkowski inner product despite that.

The induced norm of a vector is

$$\boxed{\|\underline{v}\|^2 = \eta(\underline{v}, \underline{v}) = \eta_{\mu\nu} v^\mu v^\nu} \, . \tag{2.29}$$

This is also the line element of the curve that $\underline{t}$ is tangent to.

Due to the non-degeneracy of η, vectors can have negative or even vanishing norm.

Based on (2.28), we naturally get an equivalent, more formal way of quantifying causality.

Definition 2.17: Timelike, Lightlike, Spacelike

A world line Γ with tangent vector $\underline{t}$ is called

- *timelike*, if $|v| < c \;\Leftrightarrow\; ds^2 = \eta(\underline{t}, \underline{t}) > 0$ along Γ
- *null/lightlike*, if $|v| = c \;\Leftrightarrow\; ds^2 = \eta(\underline{t}, \underline{t}) = 0$ along Γ
- *spacelike*, if $|v| > c \;\Leftrightarrow\; ds^2 = \eta(\underline{t}, \underline{t}) < 0$ along Γ

This very compact characterization adds to the intuitive definition from before.

-¿ Penrose continues to make interesting point on that on page 407: light cones more fundamental than metric

2.5.2 Lorentz Transformations 2

Lorentz transformations have been introduced in section 2.4, from which we know how they look like, that they correctly reproduce effects of relativity and how they can be visualized in spacetime diagrams. Nonetheless, it is worthwhile to revisit them now because of the knowledge we have gained since then, in particular regarding the mathematical interpretation of spacetime as a 4-manifold. We have already seen how this formulation comes with a natural structure, like vectors. A crucial part of the theory of manifolds is yet to be discussed. After all, one of the key motivations to use manifolds was that while they are described in terms of coordinates, but their properties exist independently of them, i.e. they do not depend on the specific choice of coordinates. Consequently, changing coordinates is an important part of the whole theory and the coordinate transformations of spacetime are those between different inertial frames – Lorentz transformations. This interpretation is the first time we encounter the more general role they play since there is a rich mathematical structure related to them.

-¿ after metric is nice, then we can present mathematical viewpoint; do not change norm of four-vector (this is well-known property of rotations, but there is also a second class of Lorentz transformations, which are called boosts; they represent what we have derived in previous section, change to other inertial frame that moves uniformly with respect to first one); also note that they admit group structure, i.e. note properties here; maybe even introduce rapidity and note connection of Lie group and Lie algebra? But not elaborate on this

maybe just note that Lorentz transformation = change of coordinates/charts (which is corresponding term in language of manifolds); one condition to obtain them is by demanding norm of four-vector does not change (ahh, might be confusing because I think Nolting refers to this kind of in Euclidian space; what he calls norm is really proper time passing on clock which moves from origin to point, isn't it? Would make sense to demand this better stay the same after transformation after we have put so much effort into invariant definition) -¿ I like this, but does not fit before introduction of norm; so maybe make subsection on transformations after metric?

Lorentz-scalar or Lorentz-invariant is scalar quantity, which does not change under Lorentz-transformation; example is proper time, but also mass etc. are of this nature

when interpreting Minkowski space as a manifold and working with coordinates/charts ξ, we know the results should be independent of ξ; in particular, that means they hold in other charts as well and changing coordinates is an important part; the basis changes even have a distinct name, *Lorentz transformation*; this is basically group theory due to the symmetries that Minkowski space possesses (known from logic and experiments) ?right?

2.6 Relativistic Dynamics

we have dependence on world line of our distance measure; this is nothing unusual in mathematical theory of metric spaces, but it raises an important physical question: what is the preferred trajectory of particles, i.e. what is the time that usually elapses for them? Turns out that it is extremal proper time (minimal for us, depends on sign convention of metric, right?), which yields straight lines.

idea: new transformation law means we have new dynamics; these can be formulated conveniently using points and vectors in Minkowski space; talk about four-momentum and forces etc.

2.6.1 Four-velocity and Four-momentum

Trajectories or world lines are nothing but curves on manifolds, as we have already seen. A natural question, however, was left unanswered until now: what is the velocity of such a trajectory? Since its role is to quantify how fast and in which direction the trajectory is changing, velocity now becomes a tangent vector. In analogy to the previous description, we may try

$$U^\alpha = \frac{dx^\alpha}{dt} = (c, \vec{v})\,. \tag{2.30}$$

Thinking back to the previous sections on clocks and proper time, though, we can immediately see how this definition is flawed: coordinate times t are not invariant, i.e. the velocity of a world line would change depending on the observer. At the same time, we can immediately come up with a solution: replacing t with the proper time τ a clock would measure along the world line. This leads to

$$\boxed{U^\alpha = \frac{dx^\alpha}{d\tau} = (c\frac{dt}{d\tau}, \frac{d\vec{x}}{d\tau}) = \frac{dt}{d\tau}\,(c, \frac{d\vec{x}}{dt}) = \gamma(c, \vec{v})}\,, \tag{2.31}$$

the *four-velocity*. $\vec{v}$ is the "regular" three-velocity, which finds its way into the equation by a simple application of the chain rule.

-¿ has only three independent velocities!

-¿ only defined for timelike world lines (because it vanishes trivially for light and is complex, i.e. not defined, for spacelike)

Information we get from $\vec{v}$ is the direction of an object (through the unit vector $\vec{v}/\|v\|$) and how fast (through $v = \|v\|$). As a tangent vector along $x^\alpha(\tau)$, $\underline{U}$ naturally contains information about the direction, so what about $\|\underline{U}\|$?

$$\boxed{\|\underline{U}\|^2 = \eta_{\mu\nu} U^\mu U^\nu = \eta_{\mu\nu} \frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = \frac{\eta_{\mu\nu} dx^\mu dx^\nu}{d\tau^2} = c^2 \frac{d\tau^2}{d\tau^2} = c^2}\,. \tag{2.32}$$

That four-velocity possesses a "built-in" normalization, it has a constant magnitude, the speed of light – no matter at which actual speed v the particle moves (and also, independent of the time τ we evaluate it at, i.e. independent of the point in manifold)! Not only does that imply all observers trivially agree on its value, it also points to the special role of τ as affine parameter σ of the world line. Note that the calculation here only works because the Minkowski metric has constant components. Equivalently, we could calculate

$$\left\|\underline{U}\right\|^2 = \eta_{\mu\nu} U^\mu U^\nu = \gamma^2 (c^2 - v^2) = c^2 \frac{1 - v^2/c^2}{1 - v^2/c^2} = c^2 .$$

four-velocity: using chain rule, we can write $U^\alpha = \frac{dx^\alpha}{d\tau} = \frac{dt}{d\tau}\frac{dx^\alpha}{dt} =: \gamma(c, \vec{v})$ in general; this $\gamma = \frac{dt}{d\tau} = \frac{\partial t}{\partial \tau}$ then depends on the metric, it is the square root of the (negative of, depends convention that is chosen regarding metric) component g_{tt}; for the flat Minkowski metric in SR (where only uniform velocities occur -¿ not true; should also hold for $v = v(\tau)$, but not in GR where gravity alters metric), we can further simplify this by using $d\tau^2 = ds^2/c^2 = (c^2 dt^2 - dx^2 - dy^2 - dz^2)/c^2 = dt^2(1 - \frac{dx^\alpha}{dt}\frac{dx^\alpha}{dt}/c^2) =: dt^2/\gamma^2$ -¿ of course different!!! Norm is always c^2 because

4-momentum is related to velocity, that is tangent vector; therefore, we can also compute inner product for it; note that this is tangent vector with four independent velocities (mentioned on Wikipedia, due to multiplication with Lorentz scalar)

redshift leads to higher energy, but frequency also changes such that all in all speed of light remains constant (might be wrong this way, but heard something like that)

2.6.2 Acceleration

using the generalization we just made, we can also deal with accelerations, i.e. non-uniform movement; this is easier (or at least more straightforward) to do than intuitive version

-¿ people claiming general relativity is needed here often comes from differences in nomenclature, what the "special" and "general" part refer to; Einstein himself saw special for uniform movement and then included acceleration into the general part (along with gravity, which is very similar to acceleration as he soon realized); but in principle, there is nothing to prevent us from dealing with acceleration already in special relativity, using inertial frames etc., which is what is done in modern descriptions quite commonly (feature setting apart general is then just gravity, i.e. a curved spacetime instead of flat, Minkowski)

we have seen how to compute proper times in special relativity, the formula could be broken down to

$$\tau = \int d\tau = \frac{1}{c} \int \sqrt{\eta_{\mu\nu} v^\mu v^\nu}\, dt$$

in fact, this formula also holds for $v = v(t)$, i.e. when a time-dependent velocity and thus acceleration is present (we can compute dynamics); this is because it does not involve absolute differences like $x - x'$, but infinitesimal ones dx along the whole path, so changes in v are incorporated automatically; however, we have to use other metrics in this case and general relativity presents a general way to compute the metric

interesting (from https://math.ucr.edu/home/baez/physics/Relativity/SR/clock.html): clock postulate cannot be proven because we do not know speed of light in accelerating frame; however, using clock postulate we can show that this is the case (locally); but note that postulate is required!!!

-¿ however, being able to compute proper times does not mean everything can be done by simply replacing v with $v(t)$; for example, Lorentz transformations fundamentally rely on $v =$ const., so accelerated frames are not straightforward to obtain from what we have done so far (does not mean it is impossible, though); we may still examine the behaviour of accelerating world lines, though (from an inertial frame)

good start I think: https://de.wikipedia.org/wiki/Zeitdilatation#Bewegung_mit_konstanter_Beschleunigung

for non-uniform velocity, connection procedure does not work between macroscopic events anymore (because getting distance from velocity is not multiplication, now integration), but only infinitesimally (equivalent: just like we have to integrate velocity for distance, we have to integrate for travel time of light; since time = distance for light -¿ nice visualization: rectification of curve, works perfect for straight line, but if changing derivative comes into play, not perfect anymore, have to go to infinitesimal distances for that); motivation for $d\tau$ -¿ better: we have infinitesimal line element, allows natural transition to non-uniform, i.e. accelerated movement

3 General Relativity

SR dealt with uniformly moving frames, now we want to use the insights gained there to generalize discussions to accelerated frames – this is what general relativity does (as it turns out, acceleration is very closely related to gravity, so GR is a theory of gravity as well)

-¿ wrong, SR can handle acceleration (contrary to popular belief I feel)! GR is really about incorporating gravity

3.1 Generalizing Relativity

3.1.1 What is wrong with Newton (and SR)?

gravitational redshift and instantaneous effect of gravity

special relativity came with the abandonment of absolute space and time – so radical changes are to be expected if we want to incorporate gravity now... indeed, it will turn out that gravity is *not* a force, but a fundamental geometrical property/feature of spacetime

3.1.2 Einstein Postulates

do postulates by Einstein again as start, but now the ones for GR; weak equivalence principle + Einstein equivalence principle

-¿ what about Mach principle? Ah, indeed needed (see https://de.wikipedia.org/wiki/Relativit%C3%A4tsprinzip)

3.1.3 Notes

Penrose has incredibly well written section 17.9 on intuition about metric and light cone structure in GR

we have seen how to compute proper times in special relativity, the formula could be broken down to

$$\tau = \int d\tau = \frac{1}{c} \int ds = \frac{1}{c} \int \sqrt{g_{\mu\nu} v^\mu v^\nu} \, dt \tag{3.1}$$

in fact, this formula also holds for $v = v(t)$, i.e. when a time-dependent velocity and thus acceleration is present (we can compute dynamics); this is because it does not involve absolute differences like $x - x'$, but infinitesimal ones dx along the whole path, so changes in v are incorporated automatically; however, we have to use other metrics in this case and general relativity presents a general way to compute the metric

3.2 Giulini Lectures

from 10 on (until 16) he deals with GWs, noice

3.3 Gravitational Physics Summary – Physics Part

Remark: we use units of $c = G = 1$ (*geometric units*).

Remark: we adopt the Einstein summation convention where repeated combinations of upper and lower indices are summed over, that is we abbreviate $\sum_\mu x^\mu y_\mu = x^\mu y_\mu$.

3.3.1 Newtonian Gravity

Newtonian gravity can be captured by his famous formula

$$\boxed{F_g = -\frac{m_1 m_2}{r^2}} \tag{3.2}$$

which describes the gravitational force that an object with mass m_1 exerts onto another object with mass m_2. From Newton's second law, we know that the same force is exerted from the second object onto the first.

This force can also be brought into the form

$$F_g = m_2 \frac{d}{dr}\left(\frac{m_1}{r}\right) = -m_2 \frac{d\Phi_g}{dr} \tag{3.3}$$

which tells us that gravitation is a conservative force with potential

$$\boxed{\Phi_g = -\frac{m_1}{r}} \tag{3.4}$$

produced by some object with mass m_1. We get the conservative property from equation (3.3) alone because gravitational force only has a radial and no angular component (thinking in polar/spherical coordinates) such that any derivative with respect to angular coordinates vanishes. Thus, (3.3) is equivalent to the more general condition for conservative forces,

$$\boxed{\vec{F} = -\vec{\nabla}\Phi} \qquad\qquad F^k = -\delta^{kl}\frac{\partial \Phi}{\partial x^l}\,. \tag{3.5}$$

As a consequence, knowing the potential is sufficient to know how gravity acts. Thus, we are interested in how to determine Φ and this can be done using the Poisson equation. For a point particle, it takes the form

$$\boxed{\Delta\Phi = \nabla^2\Phi = 0} \tag{3.6}$$

and for a continuous mass distribution $\rho(\vec{x})$

$$\boxed{\Delta\Phi(\vec{x}) = 4\pi\rho(\vec{x})} \qquad\qquad \delta^{ij}\frac{\partial^2\Phi(\vec{x})}{\partial x^i \partial x^j} = 4\phi\rho(\vec{x})\,. \tag{3.7}$$

Another perspective is not to look at forces $\vec{F}$, but at associated accelerations which comes from Newton's second law

$$\boxed{\vec{F} = m\vec{a} = m\frac{d^2\vec{r}}{dt^2}} \qquad F^k = ma^k = m\ddot{r}^k \tag{3.8}$$

or at momenta $\vec{p}$ which are defined by

$$\boxed{\vec{F} = \frac{d\vec{p}}{dt} \quad \Leftrightarrow \quad \vec{p} = m\vec{v}} \qquad p^k = mv^k\,. \tag{3.9}$$

Example 3.1: Gravity on Earth

The gravity exerted by Earth on objects with mass m (assuming they stand on Earth's surface for now) is

$$F_g = -m\frac{m_e}{r_e^2} = -mg \tag{3.10}$$

Comparing that with Newton's second formula, $F = ma$, we see that such an object experiences an acceleration

$$a = -g = -9.81\frac{\mathrm{m}}{\mathrm{s}^2} = -1.1 \cdot 10^{-16}\frac{1}{\mathrm{m}}\,. \tag{3.11}$$

Remark: note that we implicitly assume that gravitational mass and inertial mass are equal here. This is a non-trivial statement, which has been experimentally verified with high accuracy.

To see how much potential energy is needed to lift objects of mass m to a height $h \ll r_e$ above Earth's surface, we can do a Taylor expansion around $h = 0$:

$$\begin{aligned}\Phi_g = -\frac{m_e}{r_e + h} &\simeq -\left.\frac{m_e}{r_e + h}\right|_{h=0} + h\left.\frac{d}{dh}\left(-\frac{m_e}{r_e + h}\right)\right|_{h=0} + O(h^2)\\ &= -\frac{m_e}{r_e} + h\left.\frac{m_e}{(r_e + h)^2}\right|_{h=0} + O(h^2)\\ &= -\frac{m_e}{r_e} + h\frac{m_e}{r_e^2} + O(h^2) = -\frac{m_e}{r_e} + hg + O(h^2)\end{aligned}$$

However, the first contribution is nothing but the energy at Earth's surface. The potential energy that at h and thus the energy which is needed to lift an object of mass m to this height h (which is what one is interested in most of the time) is given to first order by the difference

$$\Phi_g = -\frac{m_e}{r_e} + gh - -\frac{m_e}{r_e} = gh\,. \tag{3.12}$$

This corresponds to gauging our measurements such that Earth's surface is the value with zero potential energy.

We see that gravity is related to a potential and thus to potential energy. Hence, we expect an objects energy to change if it moves in a gravitational field (in radial direction). This has interesting consequences, for example because light will also be affected by this.

Example 3.2: Gravitational Redshift

somehow we could build perpetual motion machine is the argument, don't get it

I rather think about it like this (should be equivalent): SR tells us that photons have a certain mass $m = \frac{E}{c^2}$; therefore, it is also affected by a gravitational potential and to move against gravity, some of its energy has to be converted; that corresponds to a change in frequency, (since $f = \frac{E}{h}$):

$$\frac{f_{\text{top}}}{f_{\text{bottom}}} = \frac{E_{\text{top}}}{E_{\text{bottom}}} = \frac{m - mgh}{m} = 1 - gh \tag{3.13}$$

Remark: in script, this is only true to first order, so derivation might be wrong... Result there reads $\frac{1}{m+mgh}$. Ahhh, because there things are defined differently: photon starts from top, thus it has more energy at ground

A natural consequence because time is inversely proportional to frequency (time differences are) is that clocks tick faster at higher altitude, i.e. for a stronger gravitational potential. It is also possible to derive it in reverse order, that is by showing that clocks tick slower in a stronger gravitational field. This causes a change in frequency and thus also a redshift.

3.3.2 Special Relativity

The theory of relativity is about how physical laws depend on the observer. We will begin with the theory of *special relativity* (SR), which generalizes Newtonian dynamics.

Definition 3.3: Inertial Frame

An *intertial frame of reference* is a coordinate frame where $\vec{F} = m\vec{a}$ holds. In particular, that means objects move with constant speed when no force is acting on them.

From the definition we can immediately see that there is no unique inertial frame because we can always get other inertial frames from existing ones by looking at frames which move with constant speed with respect to them.

We also see that all laws of physics hold equally in all inertial frames because

$$\vec{F} = m\vec{a} = m\frac{d\vec{v}(t)}{dt} = m\frac{d\vec{v}'(t)}{dt} \tag{3.14}$$

as long as $\vec{v} - \vec{v}' = \text{const}$. That also means only laws and quantities which are invariant under transformations between inertial frames have physical meaning. In particular, that means coordinates of events have no physical meanings. As an alternative, we can look at distances between events which turn out to be invariant because they are related to a metric:

$$\begin{aligned}(\Delta s)^2 &= -(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 \\ &= \underline{\Delta x} \cdot \eta \cdot \underline{\Delta x} = \begin{pmatrix}\Delta t & \Delta x & \Delta y & \Delta z\end{pmatrix} \cdot \begin{pmatrix}-1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1\end{pmatrix} \cdot \begin{pmatrix}\Delta t \\ \Delta x \\ \Delta y \\ \Delta z\end{pmatrix} \\ &= \eta_{\mu\nu}(\Delta x)^\mu(\Delta x)^\nu = (\Delta x)^\mu(\Delta x)_\mu \,.\end{aligned}$$

Making these differences Δ infinitesimally small gives the line element of the metric η

$$ds^2 := (ds)^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 \,. \tag{3.15}$$

Let us now think about gravity in special relativity. In principle, the Newtonian description is kept, but some effects can be examined in different manner now, e.g. due to the new notion/tool of different inertial frames. However, a frame where gravity acts is *not* inertial (because we have external force in gravity, right?)! Thus, to do physics on Earth, we have to find a reference frame in which the effect of gravity is cancelled out. Obviously, earths surface is not sufficient and neither is a uniformly moving one. In free fall, however, we experience no gravity, that is a freely falling frame cancels out the effect of gravity. This can be stated more formally:

Property 3.4: Weak Equivalence Principle

The effects of a gravitational field are indistinguishable from an accelerated frame of reference.

Basically, that means only a freely falling frame can serve as an inertial frame on Earth. That raises the question what happens to the laws of physics in such a freely falling frame.

Property 3.5: Einstein Equivalence Principle

The laws of physics in a freely falling frame are locally described by SR without gravity. For this reason, such a frame is also called *local inertial frame (LIF)*.

Speaking strictly mathematically, "locally" means in an infinitesimally small neighbourhood of points. The degree to which this can be extended in practice depends on the physical effects of interest.

Since gravity acts radially, its direction changes on different places around Earth. That implies there is no uniform direction of acceleration, so there can be no global freely falling frame/LIF. Other properties of gravity which are known from experience are the following:

(a) All bodies which start with the same initial velocity move through a gravitational field along the same curve

(b) Bodies which move initially parallel to each other in a freely falling frame do not necessarily move parallel at all times if an external gravitational field is present (this effect is due to *tidal forces* acting on them)

Property (b) can be further examined and quantified. To do that, we note that for a particle with world line $x^k(\tau)$ we have

$$\frac{d^2 x^k}{d\tau^2} = -\delta^{kl}\frac{\partial \Phi}{\partial x^l} = -\delta^{kl}\left.\frac{\partial \Phi}{\partial x^l}\right|_{\vec{x}}$$

because of (3.5) and Newton's second law $F^k = m\frac{d^2 x^k}{d\tau^2}$. Similarly, for another particle starting close to the first one (i.e. with world line $x^k + \xi^k$, where $\left|\xi^k \xi_k\right| \ll 1$)

$$\begin{aligned}\frac{d^2\left(x^k + \xi^k\right)}{d\tau^2} &= -\delta^{kl}\left.\frac{\partial \Phi}{\partial x^l}\right|_{\vec{x}+\vec{\xi}} \\ &\simeq -\delta^{kl}\left.\frac{\partial \Phi}{\partial x^l}\right|_{\vec{x}} - \delta^{kl}\xi^m \frac{\partial}{\partial x^m}\left.\frac{\partial \Phi}{\partial x^l}\right|_{\vec{x}}\end{aligned}$$

where we used a Taylor approximation to first order. The linearity of derivatives yields

$$\boxed{\frac{d^2 \xi}{d\tau^2} = \frac{d^2\left(x^k + \xi^k\right)}{d\tau^2} - \frac{d^2 x^k}{d\tau^2} = -\delta^{kl}\frac{\partial^2 \Phi}{\partial x^m \partial x^l}\xi^m} \tag{3.16}$$

Remark: note that the evaluation is still at the point $\vec{x}$, not at something related to ξ^k!

This is the *Newtonian deviation equation*. We see that tidal forces are governed by the tidal acceleration tensor $\frac{\partial^2 \Phi}{\partial x^m \partial x^l}$. Tidal forces are a way to detect gravity as opposed to constant acceleration (which would affect the world lines x^k and $x^k + \xi^k$ equally)

3.3.3 Curved Spacetime

One problem in SR is that the Newtonian description of gravity is still taken to be valid. That, however, is a problem because there are many inconsistencies between them, for example the instantaneous effect of gravity (gravitational redshift is also puzzling). However, we can come up with generalized description: for anybody familiar with differential/Riemannian geometry, the effects (a), (b) of gravity stated above sound very much like the ones associated with a curved space. This motivates the (mathematical) description of gravity as a geometrical effect in Minkowskian spacetime (which will become a curved space in this process). Many relations known from SR will remain, but with different quantities and most prominently, a different metric other than η. The basic goal of *general relativity* (GR) will be to find ways to derive the metric which contains information about spacetime curvature and thus gravity.

The approach in this subsection will always be to look how generalizations can be made using the metric and other tools of geometry, while recovering SR in a LIF. Such a check, however, has not been made for the metric itself yet! Thus, we will now look at how the mathematical term "locally" is to be thought of. Taking an arbitrary metric with components $g_{\mu\nu}$ in some basis, we can always transform to other coordinates using the tensor transformation law. This

say something about how LIF can be characterized using metric; is important we always want to recover results in there, for example statements on timelike etc. (around 2.36)

The most basic thing we need to know is how test particles move in curved spaces. We will start with the case where no force is present, i.e. free movement. The Newtonian theory/SR gives us $\ddot{\vec{a}} = 0$ and thus a movement on straight lines. This has to be reproduced locally (that is in a LIF), but originating from a more general concept. Finding this generalization is based on the observation that tangent vectors remain constant along straight lines. That leads to the *geodesic equation*

$$\boxed{\nabla_{\underline{t}} t^\beta = t^\alpha \nabla_\alpha t^\beta = 0}, \tag{3.17}$$

which just expresses that t^β remains constant as long as we take the derivative along the curve that it is tangent to (which explains the $t^\alpha \nabla_\alpha$ part). Consequently, the world lines test particles are *geodesics*.

One can obtain the same statement from a completely different approach: by demanding that world lines are the curves in Minkowski space which extremize the proper time/distance

$$\tau_{AB} = \int_A^B d\tau = \int_A^B \sqrt{-g_{\alpha\beta} dx^\alpha dx^\beta} = \int_0^1 \sqrt{-g_{\alpha\beta} \frac{dx^\alpha}{d\sigma} \frac{dx^\beta}{d\sigma}} \, d\sigma \tag{3.18}$$

between two events A, B.

In both cases, we obtain the following coordinate version of the geodesic equation:

$$\boxed{\frac{d^2 x^\beta}{d\sigma^2} + \Gamma^\beta_{\alpha\delta} \frac{dx^\alpha}{d\sigma} \frac{dx^\delta}{d\sigma}} \tag{3.19}$$

Proof. In the first approach, we use

$$\begin{aligned}
t^\alpha \nabla_\alpha t^\beta &= \frac{dx^\alpha}{d\sigma} \nabla_\alpha \frac{dx^\beta}{d\sigma} \\
&= \frac{dx^\alpha}{d\sigma} \left(\frac{\partial}{\partial x^\alpha} \frac{dx^\beta}{d\sigma} + \Gamma^\beta_{\alpha\delta} \frac{dx^\delta}{d\sigma} \right) \\
&= \frac{dx^\alpha}{d\sigma} \frac{\partial}{\partial x^\alpha} \frac{dx^\beta}{d\sigma} + \Gamma^\beta_{\alpha\delta} \frac{dx^\alpha}{d\sigma} \frac{dx^\delta}{d\sigma} \\
&\underset{\text{chain rule}}{=} \frac{d}{d\sigma} \frac{dx^\beta}{d\sigma} + \Gamma^\beta_{\alpha\delta} \frac{dx^\alpha}{d\sigma} \frac{dx^\delta}{d\sigma} \\
&= \frac{d^2 x^\beta}{d\sigma^2} + \Gamma^\beta_{\alpha\delta} \frac{dx^\alpha}{d\sigma} \frac{dx^\delta}{d\sigma} \,.
\end{aligned}$$

In the second approach, we derive the Euler-Lagrange equations by varying the proper time integral:

$$\boxed{\frac{\partial L}{\partial x^\alpha} = \frac{d}{d\sigma} \frac{\partial L}{\partial dx^\alpha / d\sigma}} \tag{3.20}$$

where we introduced the Lagrangian

$$\boxed{L = L\left(x^\alpha, \frac{dx^\alpha}{d\sigma}\right) = \sqrt{-g_{\alpha\beta} \frac{dx^\alpha}{d\sigma} \frac{dx^\beta}{d\sigma}}} \,. \tag{3.21}$$

Notable and useful properties in this context are that L is constant along the geodesic because

$$\begin{aligned}
\frac{d}{d\sigma}\left(g_{\alpha\beta} \frac{dx^\alpha}{d\sigma} \frac{dx^\beta}{d\sigma} \right) &= t^\gamma \partial_\gamma \left(g_{\alpha\beta} t^\alpha t^\beta \right) = t^\gamma \nabla_\gamma \left(g_{\alpha\beta} t^\alpha t^\beta \right) \\
&= t^\alpha t^\beta \nabla_\gamma g_{\alpha\beta} + g_{\alpha\beta} t^\beta t^\gamma \nabla_\gamma t^\alpha + g_{\alpha\beta} t^\alpha t^\gamma \nabla_\gamma t^\beta = 0 \,.
\end{aligned}$$

Note that this does *not* imply $\partial_\alpha L = 0$ (so the Euler-Lagrange equations still make sense). It does, however, mean that we are free to change the parametrization from σ to any *affine parameter* $\sigma' = a\sigma + b,\ a, b \in \mathbb{R}$ while only picking up a factor $\frac{1}{a}$. The proper time τ is defined as the parameter σ' with

$$L = 1 \qquad \Leftrightarrow \qquad g_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} = -1 \tag{3.22}$$

which comes from the known normalization of the four-velocity $U^\alpha = \frac{dx^\alpha}{d\tau}$. Therefore, we can always replace $d\sigma \to d\tau = L d\sigma$ or vice versa. □

The advantage of having two approaches is that, in some cases, calculating the Christoffel symbols might be easier or might have already been done, while in others the Euler-Lagrange equations reveal very useful properties of the problem/system (for example in case one coordinate does not appear explicitly in L; then, we have immediately found a quantity which is conserved by the system, i.e. it does not change as time evolves/as we vary σ when going along the geodesic, this quantity being $\frac{\partial L}{\partial dx^\alpha / d\sigma}$).

A physical consequence from $L = \text{const}$ is that geodesics/world lines which are timelike/null/spacelike somewhere have this property everywhere! Test particles (with mass $m > 0$) always move along timelike geodesics while massless particles like photons move along null geodesics (spacelike ones violate causality).

3.3.4 Weakly Curved Spacetimes

basic idea here: re-derive first subsection from GR (confirm that it reproduces Newtonian results), that is using the spacetime metric

$$ds^2 = -(1 + 2\Phi)dt^2 + (1 - 2\Phi)\left(dx^2 + dy^2 + dz^2\right) \tag{3.23}$$

which is the first order approximation in case the Newtonian potential $\Phi = -\frac{m}{r}$ fulfils $|\Phi| \ll 1$; we will see where it comes from later on, will be used as motivation for now

we can formulate the geodesic equation for test particles in many ways, for example

$$U^\alpha \nabla_\alpha U^\beta = 0 \qquad\qquad p^\alpha \nabla_\alpha p^\beta \tag{3.24}$$

where $U^\alpha = \frac{dx^\alpha}{d\tau}$ is the four-velocity of a particle with world line x^α and $p^\alpha = mU^\alpha$ is the four-momentum (sometimes preferred quantity because it is well-defined also for photons)

for the first/time component $p^0 = E \approx m$ ($c^2 = 1$, this is an approximation to lowest order) we obtain

$$0 = m\frac{dp^0}{d\tau} + \Gamma^0_{\alpha\beta} p^\alpha p^\beta \approx m\frac{dp^0}{d\tau} + \Gamma^0_{00}\left(p^0\right)^2 \qquad \Leftrightarrow \qquad \frac{dp^0}{d\tau} \approx -m\frac{\partial\Phi}{\partial t} \tag{3.25}$$

which matches the Newtonian result that if the gravitational field does not change over time, then the energy p^0 will be conserved over time

similarly, the equations for the spatial components

$$0 = m\frac{dp^k}{d\tau} + \Gamma^k_{\alpha\beta} p^\alpha p^\beta \approx m\frac{dp^k}{d\tau} + \Gamma^k_{00} m^2 \qquad \Leftrightarrow \qquad \frac{dp^k}{d\tau} \approx -m\delta^{kl}\frac{\partial\Phi}{\partial x^l} \tag{3.26}$$

match the Newtonian result (3.5) that gravity acts as a conservative force

In the Newtonian case, tidal forces (3.16) could be used to detect gravity as opposed to constant acceleration; something similar should be possible for curvature (which is how we described gravity now) and indeed, we obtain the *geodesic deviation equation*

$$\boxed{\nabla_{\underline{U}} \nabla_{\underline{U}} \xi^\beta = U^\sigma \nabla_\sigma U^\alpha \nabla_\alpha \xi^\beta = -R^\beta{}_{\gamma\delta\epsilon} U^\gamma \xi^\delta U^\epsilon}\,. \tag{3.27}$$

Therefore, gravity/curvature is determined and measured by the *Riemann curvature tensor*

$$\boxed{R^{\beta}{}_{\gamma\delta\epsilon} = \frac{\partial \Gamma^{\beta}{}_{\gamma\epsilon}}{\partial x^{\delta}} - \frac{\partial \Gamma^{\beta}{}_{\gamma\delta}}{\partial x^{\epsilon}} + \Gamma^{\beta}{}_{\delta\mu}\Gamma^{\mu}{}_{\gamma\epsilon} - \Gamma^{\beta}{}_{\epsilon\mu}\Gamma^{\mu}{}_{\gamma\delta}}\,. \tag{3.28}$$

A helpful way to think about the Riemann tensor is as the "difference"/commutator of subsequent derivatives, which is motivated by the following equation:

$$\boxed{\left[\nabla_{\alpha}, \nabla_{\beta}\right] V^{\mu} = R^{\mu}{}_{\gamma\alpha\beta} V^{\gamma}}\,. \tag{3.29}$$

We can simplify this equation by choosing a LIF with $\underline{e}'_0 = \underline{e}_{\tau}$ (one basis vector along proper time). Since this coordinate system is parallel transported along the geodesic,

$$U^{\gamma}\nabla_{\gamma}\frac{dx'^{\beta}}{dx^{\alpha}} = U^{\gamma}\nabla_{\gamma}\left(\underline{e}'^{\beta}\right)_{\alpha} = 0$$

and we obtain

$$\frac{d^2\xi^{\beta}}{d\tau^2} = \frac{\partial x'^{\beta}}{\partial x^{\alpha}} U^{\sigma}\nabla_{\sigma}U^{\delta}\nabla_{\delta}\xi^{\alpha} = -\frac{\partial x'^{\beta}}{\partial x^{\alpha}} R^{\alpha}{}_{\gamma\delta\epsilon}\frac{\partial x^{\gamma}}{\partial \tau}\xi^{\delta}\frac{\partial x^{\epsilon}}{\partial \tau} = -R'^{\beta}{}_{\tau\delta\tau}\xi'^{\delta} \tag{3.30}$$

Remark: note that the evaluation is still at the point $\vec{x}$, not at something related to ξ^k! due to the transformation law for vectors and tensors (?). This equation looks very much like the Newtonian expression (3.16). Moreover, we see that the Riemann tensor might play a role which is similar to the tidal acceleration tensor $\frac{\partial^2\Phi}{\partial x^m \partial x^l}$. Hence, it might also play a similar role in causing gravity...

3.3.5 Einstein Equation

We have seen how Newtonian theory can be generalized, for example by going from a potential Φ for the gravity field to the metric $g_{\mu\nu}$. Many effects could be derived from that, e.g. equations of motion and the effect of gravity in tidal forces – but we have no way of determining the metric yet! Thus, we will now look for a field equation like (3.7). As it turns out, there is no way of truly deriving the result. Instead, we can only motivate it sufficiently.

First of all, we need an analogue to the mass distribution $\rho(\vec{x})$ to describe the effect of matter. This should be a tensorial quantity in accordance with previous generalizations.

This definition also allows to summarize several conservation formulas into one, coordinate-independent equation:

$$\nabla_{\alpha}T^{\alpha\beta} = 0\,. \tag{3.31}$$

This gives us

$$\partial_t\rho + \partial_i\pi^i = 0 \qquad\qquad \partial_t\pi^i + \partial_j T^{ij} = 0 \tag{3.32}$$

which corresponds to energy and momentum conservation (which is the continuum version of $F^k = ma^k$ since spatial components of the stress-energy tensor describe forces).

Now that we have something which causes gravity, we need to equate it with something that describes gravity – a natural quantity would be the Riemann tensor $R^\alpha{}_{\beta\gamma\delta}$. However, the stress-energy tensor only has two indices, while the Riemann tensor has four. Thus, we have to get rid of two indices, which can be done by contraction. Since $T = T^\alpha{}_\alpha \neq 0$, we are only interested a 2-tensor trace component, which turns out to be the *Ricci tensor*

$$\boxed{R_{\alpha\beta} = R^\mu{}_{\alpha\mu\beta}} = R_{\beta\alpha}\,. \tag{3.33}$$

However, the Ricci tensor is not divergent free, so a field equation of the kind $R^{\alpha\beta} = bT^{\alpha\beta}$ cannot hold. Luckily, we can construct divergent-free tensor from it by rearranging

$$\nabla_\alpha R^{\alpha\beta} = \frac{1}{2} g^{\alpha\beta} \nabla_\alpha R$$

which can be obtained using the Bianchi identity. Since $\nabla_\alpha g^{\alpha\beta} = 0$,

$$\nabla_\alpha \left(R^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} R \right) = 0 \tag{3.34}$$

and we see that the *Einstein tensor*

$$\boxed{G^{\alpha\beta} = R^{\alpha\beta} - \frac{1}{2} g^{\alpha\beta} R = G^{\beta\alpha}} \tag{3.35}$$

fulfils the required conservation law. Here, R is the *Ricci scalar* (or *scalar curvature*)

$$\boxed{R = R^\alpha_\alpha}\,, \tag{3.36}$$

part of another trace component of the Riemann tensor. Thus, the field equation of interest (also called *Einstein equation*) is

$$\boxed{G^{\alpha\beta} = 8\pi T^{\alpha\beta}} \tag{3.37}$$

where the constant of proportionality 8π is determined by requiring that they reduce to the Newtonian weak-field limit. Now, remembering that $\nabla_\gamma g^{\alpha\beta} = 0$ tells us we are free to add a term of the kind $ag^{\alpha\beta}$ to the field equation. Therefore,

$$\boxed{G^{\alpha\beta} + \Lambda g^{\alpha\beta} = 8\pi T^{\alpha\beta}} \tag{3.38}$$

is an equally valid version of the Einstein equation. Λ is the (in)famous *cosmological constant* and its contribution can be thought of as "stress-energy of empty space". However, this can be absorbed in $T^{\alpha\beta}$, so we will mostly stick to using (3.37). The Einstein equation can be interpreted following this famous quote by John Archibald Wheeler:

"Spacetime tells matter how to move; matter tells spacetime how to curve."

To simplify things one can look empty space first, where $T^{\alpha\beta} = T = 0$. For the Einstein equation, that means

$$G^{\alpha\beta} = 0\,.$$

But since trace of the Einstein equation yields $R = -8\pi T$ ($= 0$ in vacuum), we obtain the following *vacuum Einstein equation*:

$$\boxed{R^{\alpha\beta} = 0}\,. \tag{3.39}$$

Now we will come to few examples of solutions of the Einstein equation.

Example 3.6: Solutions of the Einstein equation

(a) Minkowski space with $g_{\alpha\beta} = \eta_{\alpha\beta}$ and

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \tag{3.40}$$

(b) Schwarzschild black hole: unique (non-trivial) solution for a static, spherically symmetric spacetime containing a mass M (a *black hole*)

$$ds^2 = -\left(1 - \frac{2M}{r}\right)dt^2 + \frac{1}{1 - \frac{2M}{r}}dr^2 + r^2\left(d\theta^2 + \sin(\theta)^2 d\phi^2\right) \tag{3.41}$$

(c) weak (Newtonian) gravitational potential: Schwarzschild solution far away from the source (such that it can be approximated as point mass)

$$ds^2 = -(1 + 2\Phi)dt^2 + (1 - 2\Phi)\left(dR^2 + R^2 d\theta^2 + R^2\sin(\theta)^2 d\phi^2\right) \tag{3.42}$$

$$= -(1 + 2\Phi)dt^2 + (1 - 2\Phi)\left(dx^2 + dy^2 + dz^2\right) \tag{3.43}$$

Is obtained from defining the potential $\Phi = -\frac{r}{r-M}$, Taylor-expanding to first order in Φ and then making the gauge transformation $r \to R = r - M$.

(d) Kerr solution of a rotating black hole

$$\begin{aligned} ds^2 = &-\frac{\Delta - a^2\sin(\theta)^2}{\rho^2}dt^2 - 2a\frac{2Mr\sin(\theta)^2}{\rho^2}dt\,d\phi + \frac{\rho^2}{\Delta}dr^2 \\ &+ \frac{\left(r^2 + a^2\right)^2 - a^2\Delta\sin(\theta)^2}{\rho^2}\sin(\theta)^2 d\phi^2 + \rho^2 d\theta^2 \end{aligned} \tag{3.44}$$

where $\Delta = r^2 - 2Mr + a^2, \rho = r^2 + a^2\cos(\theta)^2$ and a parametrizes the rotation speed

(e) Friedmann-Lemaitre-Robertson-Walker (FLWR) metric of an isotropic, homogenous, expanding universe

$$ds^2 = -dt^2 - a(t)^2 \left[\frac{1}{1 - kr^2} dr^2 + r^2 \Big(d\theta^2 + \sin(\theta)^2 d\phi^2 \Big) \right] \tag{3.45}$$

As we can see, there are not too many analytical solutions. This is because the Einstein equation forms a coupled system of 10 non-linear, second-order partial differential equations, which makes them very hard to solve in general (even numerically).

3.3.6 Existence of Gravitational Waves

gravitational wave = GW

3.3.7 Effect of Gravitational Waves

Gravitational waves are small perturbations of spacetime, so they should have a measurable effect. To find this effect, we will now look at a particles in the TT gauge. We assume it to be at some position x^α and at rest at $t = 0$, that is

$$U^\alpha \Big|_{t=0} = \frac{dx^\alpha}{d\tau} \Big|_{t=0} = (1, 0, 0, 0) \,.$$

Evaluating the geodesic equation at $t = 0$ yields

$$\begin{aligned} \frac{dU^\alpha}{d\tau} = \frac{d^2 x^\alpha}{d\tau^2} &= -\Gamma^\alpha{}_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = -\Gamma^\alpha{}_{tt} \\ &= -\frac{1}{2} g^{\alpha\mu} \left(2\partial_t g_{t\mu} - \partial_\mu g_{tt} \right) = 0 \,. \end{aligned} \tag{3.46}$$

The four-velocity remains constant, so the coordinates of the particle do not change! That, however, does not mean GWs have no effect – after all, coordinates have no invariant meaning anyway. Instead, we have to look at physically meaningful quantities like the proper distance between particles which is related to the metric and thus may be affected by a GW.

To simplify calculations, we will now assume a GW propagating in z-direction, with linear polarization and $h_+ = h_+(t - z), h_\times = 0$. The proper distance between two particles which are initially separated by L in x-direction then becomes

$$L_x(t) = \int_0^L \sqrt{g_{xx} dx dx} = \int_0^L \sqrt{1 + h_+} \, dx \approx \int_0^L \left(1 + \frac{h_+}{2} \right) dx = L \left(1 + \frac{h_+(t)}{2} \right) \tag{3.47}$$

assuming that the wavelength is much longer than L (whence the amplitude h_+ does not change much during the propagation from $x = 0$ to $x = L$). Therefore, a GW causes a (time-dependent) relative length change

$$\frac{\Delta L_x}{L} = \frac{L_x - L}{L} \approx \frac{h_+(t)}{2} \tag{3.48}$$

which is often called *strain*.

Another way to quantify the effect of a GW is to look at geodesic deviation. In a LIF, we recall that the corresponding equation (3.27) reads

$$\frac{d^2\xi^\alpha}{d\tau^2} = -R^\alpha{}_{\tau\mu\tau}\xi^\mu \,.$$

It is possible to calculate the relevant components of the Riemann tensor in TT gauge:

$$R^x{}_{\tau x\tau} = -\frac{1}{2}\frac{\partial^2 h_+}{\partial\tau^2} \qquad R^y{}_{\tau x\tau} = -\frac{1}{2}\frac{\partial^2 h_\times}{\partial\tau^2} \qquad R^y{}_{\tau y\tau} = \frac{1}{2}\frac{\partial^2 h_+}{\partial\tau^2} \,. \tag{3.49}$$

Assuming an initial separation in x-direction again, i.e. $\underline{\xi} = (0,\xi,0,0)$, we obtain

$$\frac{d^2\xi^x}{d\tau^2} = \frac{\xi}{2}\frac{\partial^2 h_+}{\partial\tau^2} \qquad \frac{d^2\xi^y}{d\tau^2} = \frac{\xi}{2}\frac{\partial^2 h_\times}{\partial\tau^2} \,. \tag{3.50}$$

For an initial separation in y-direction, $\underline{\xi} = (0,0,\xi,0)$, the roles are reversed:

$$\frac{d^2\xi^x}{d\tau^2} = \frac{\xi}{2}\frac{\partial^2 h_\times}{\partial\tau^2} \qquad \frac{d^2\xi^y}{d\tau^2} = -\frac{\xi}{2}\frac{\partial^2 h_+}{\partial\tau^2} \,. \tag{3.51}$$

Consequently, GWs exert a force which changes the proper distance between particles, i.e. they stretch and squeeze spacetime between particles (not only in a LIF). Moreover, we see that the two polarizations have an equivalent effect, they are just rotated against each other.