Summary

Theory of Relativity and Cosmology

Lecture at the RWTH-Aachen University

Read by
Prof. Julien Lesgourges

Summary by
Jonas El Gammal

in
WS 2018/19

Last change:
October 12, 2020

Contents

Die hier angegebene Literatur ist nur eine Auswahl an Büchern, die ich selbst hilfreich für das Verstehen der Vorlesung fand. Sie ist keinesfalls komplett.
Außerdem baut diese Vorlesung maßgeblich auf den Theorievorlesungen des Bachelors auf. Für allgemeine Informationen zu diesen Themen möchte ich auf die Zusammenfassung von Jannis Zeller [1] (sehr kompakt), sowie auf die Lehrbuchreihe von T. Fließbach [2],[3],[4],[5] (ausführlichere Erklärungen, sowie Übungsaufgaben) verweisen.
Zusätzlich wird Spezielle Relativitätstheorie zum Verständnis der Vorlesung unbedingt benötigt. Obwohl Am Anfang dieser Zusammenfassung eine kleine Einführung gegeben wird empfehle ich für mehr Informationen [6] und [7].

Diese Zusammenfassung wurde nach bestem Wissen und Gewissen geschrieben. Trotzdem kann ich kann nicht für die Richtigkeit der Angaben garantieren. Falls grobe Fehler auftauchen sollten oder für sonstige Anmerkungen bitte ich um eine E-Mail an *jonas.el.gammal@rwth-aachen.de*.

Bibliography

[1] Jannis Zeller. *Zusammenfassung Theoretische Physik I-IV*. WS 2017/18. URL: https://www.studydrive.net/kurse/rwth-aachen/vernetzungsmodul-theoretische-physik/zusammenfassungen/theo-bachelor-zusammenfassung/vorschau/501368.

[2] T. Fließbach. *Mechanik: Lehrbuch zur Theoretischen Physik I*. Springer Berlin Heidelberg, 2014. ISBN: 9783642554322.

[3] T. Fließbach. *Elektrodynamik: Lehrbuch zur Theoretischen Physik II*. Spektrum Akademischer Verlag, 2012. ISBN: 9783827430366.

[4] T. Fließbach. *Quantenmechanik: Lehrbuch zur Theoretischen Physik III*. Springer Berlin Heidelberg, 2018. ISBN: 9783662580318.

[5] T. Fließbach. *Statistische Physik: Lehrbuch zur Theoretischen Physik IV*. Spektrum Akademischer Verlag, 2010. ISBN: 9783827425287.

[6] Malgorzata Worek. *Vorlesungsskript zur speziellen Relativitätstheorie*. WS 2017/18.

[7] T. Fließbach. *Allgemeine Relativitätstheorie*. Springer Berlin Heidelberg, 2016. ISBN: 9783662531068.

[8] Helmut Fischer and Helmut Kaul. *Mathematik für Physiker - Band 1: Grundkurs*. Berlin Heidelberg New York: Springer-Verlag, 2011. ISBN: 978-3-834-89863-0.

[9] Bernard Schutz. *A First Course in General Relativity*. Cambridge: Cambridge University Press, 2009. ISBN: 978-0-521-88705-2.

[10] Helmut Fischer and Helmut Kaul. *Mathematik für Physiker - Band 3: Variationsrechnung, Differentialgeometrie, Mathematische Gundlagen der allgemeinen Relativitätstheorie*. Berlin Heidelberg New York: Springer-Verlag, 2011. ISBN: 978-3-658-00474-3.

1. Special Relativity

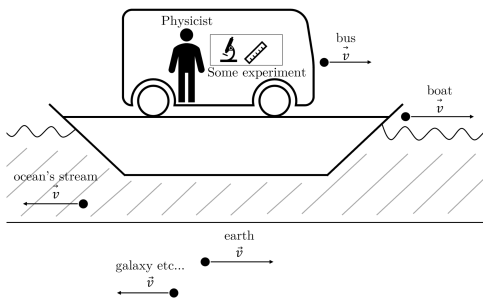
1.1. Principles of special relativity

We start with the Special Theory of Relativity (SRT) which we need as foundation to build the structure of General Relativity (GR). This is just a small introduction, for a more detailed course see [6].
First, we need three postulates, from which we can build up the mathematical framework of the SRT.

Principle A ("‘old relativity principle"’)

All *inertial observers* are equivalent and experience the same laws of physics.

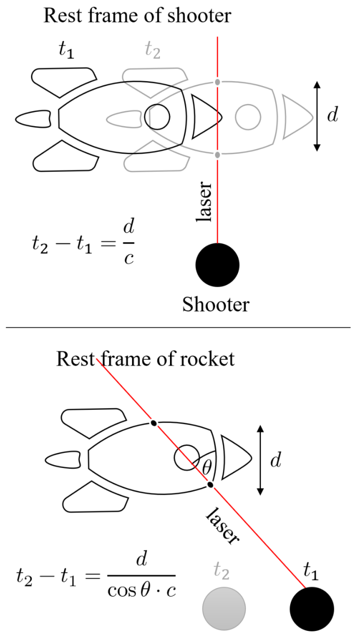
An inertial system is a system, which experiences no acceleration. In particular, in two systems moving at constant velocity relative to each other, the laws of physics are the same. This in turn means that physical laws explored on Earth are generalizable, despite the fact that not all velocities of the Earth relative to other objects are known to us.



Principle B

The *speed of light* is the same in all inertial frames.

This postulate, which was first implied by the experiment performed by Michelson & Morley, showed that the Galileo transformation could not be sufficient to describe describe the transformation from on inertial frame to another. This can be made clear by the following thought experiment. If one imagines a rocket through which a laser beam is shot and requires a constant speed of light in the frame of the laser shooter and the rocket, one notices that the time spent by the laser beam in the rocket is not the same in both inertial systems (laser, rocket). The paradox was solved when Lorentz proposed an alternative transformation that requires a different understanding of space and time for each observer.



Principle C

The *dynamics* in special relativity are given by the equation

$$\boxed{\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = m\frac{\mathrm{d}\vec{U}}{\mathrm{d}t} = \vec{F}} \tag{1.1.0.1}$$

Here $\vec{p}$ describes the 4-momentum, $\vec{U}$ the 4-velocity ($\frac{\mathrm{d}\vec{U}}{\mathrm{d}t}$ then is the 4-acceleration) and $\vec{F}$ the 4-force. These will be defined in 1.6.

1.2. Lorentz-Boosts

We will now work in *4-dimensional space-time*. We call a point in this space *Event* to clarify the concept of space-time. Such a point is therefore a vector

$$\vec{v} = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t \\ \boldsymbol{x} \end{pmatrix}$$

A trajectory now becomes 4-Dimensional as well and we call it *World line*.

We want to represent the connection between two different frames by a transformation. We therefore consider a simple example in which there is one observer in the O and on in the $\bar{O}$ system, whereas $\bar{O}$ moves with the speed

$$\boldsymbol{v} = \begin{pmatrix} v \\ 0 \\ 0 \end{pmatrix}$$

relative to O. This simplifies the general transformation

$$\begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \leftrightarrow \begin{pmatrix} \bar{t} \\ \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix}$$

to

$$\begin{pmatrix} t \\ x \\ 0 \\ 0 \end{pmatrix} \leftrightarrow \begin{pmatrix} \bar{t} \\ \bar{x} \\ 0 \\ 0 \end{pmatrix} .$$

Without sacrificing the generality, we can set our origin to $\vec{x} = 0$.
Let's start with some graphical considerations. First we draw into O a x and a t-line, which we will multiply by c to get the same units. The choice of c is due to the fact that this is the only thing which *must* remain constant in O and $\bar{O}$. It is clear that a light pulse then travels along a 45° line.
We cab easily construct the $c\bar{t}$ line if we look at an event stays at rest in $\bar{0}$. We then obtain $\bar{x} = 0$. If we look at the object from O its position is given by

$$x = v \cdot t \quad \text{which leads to} \quad ct = \frac{c}{v} x \ .$$

The angle θ between the ct and $c\bar{t}$-axis is then given as

$$\tan(\theta) = \frac{v}{c}$$

We can graphically construct the $\bar{x}$-axis by plotting two light pulses at the same time on the $\bar{x} = 0$ line, moving at 45° in both inertial systems. Thus, we immediately find that the $\bar{x}$-axis is also tilted by the angle θ. A diagram of this type is called **Minkowski diagram**.

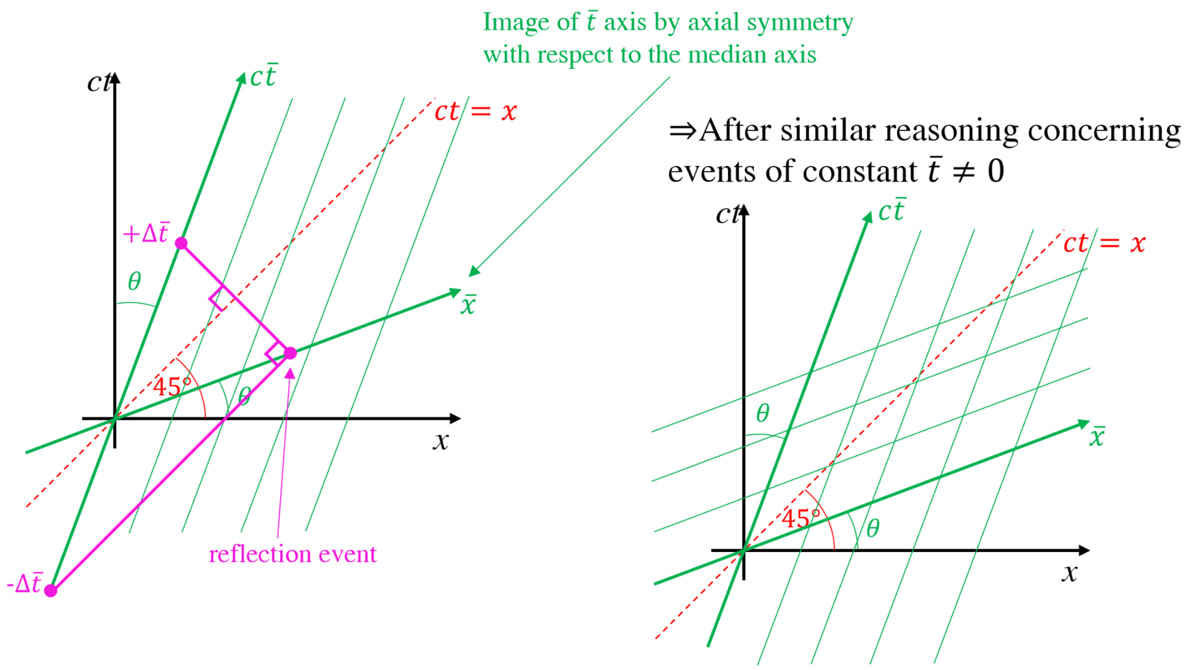


Figure 1.1.: Construction of the O and $\bar{O}$ systems in the Minkowski diagram

We will now write down the most general transformation that connects these inertial frames. For this we impose the important condition that the transformation is **linear**. This is justified by the fact that in our diagram, straight lines are mapped to straight lines and areas are preserved. We simply write this transformation down, for a derivation see [6].

$$\begin{pmatrix} c\bar{t} \\ \bar{x} \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\frac{v}{c} \\ -\frac{v}{c} & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}$$

γ is a normalization factor that is unknown so far, but we know that

- γ can only be a function of $\boldsymbol{v}$ or c or both
- γ can only depend on $|\boldsymbol{v}|$ (if we rotate the coordinate system the norm should be preserved)
- If $\bar{\bar{O}}$ moves with $-\boldsymbol{v}$ relative to $\bar{O}$ the transformation should give us back O.

We can determine γ by doing a double-transformation ($O \to \bar{O}$,$\bar{O} \to \bar{\bar{O}}$):

$$\begin{pmatrix} ct \\ x \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\frac{v}{c} \\ -\frac{v}{c} & 1 \end{pmatrix} \cdot \gamma \begin{pmatrix} 1 & \frac{v}{c} \\ \frac{v}{c} & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}$$
$$\Rightarrow \gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Alternatively we can calculate γ by using $\Lambda \cdot \Lambda^{-1} = 1$ and

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\det \Lambda} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} .$$

Note: In the following, we will often use Λ for a general Lorentz Transformation (LT), and $\beta := v/c$ will be used as an abbreviation (later we set $c = 1$, so $\beta = v$). This factor is called Lorentz factor.
In addition, we will use the common signature in cosmology $(-,+,+,+)$ (versus $(+,-,-,-)$ in high energy physics) for the metric tensor. This changes a sign in some places (especially in some norms, which follow from the preservation of the length element), but leaves the structure of the equations invariant.

to avoid confusion, I will try to consistently use the notation $\boldsymbol{v}$ (bold) for 3 vectors and $\vec{v}$ (vector arrow) for 4 vectors.

1.3. Consequences of Lorentz-Boosts

1.3.1. Lorentz-Contraction

We imagine a staff at rest in $\bar{O}$ with the length $\bar{l}$ hat. What is its length in O? We will go back to graphic:

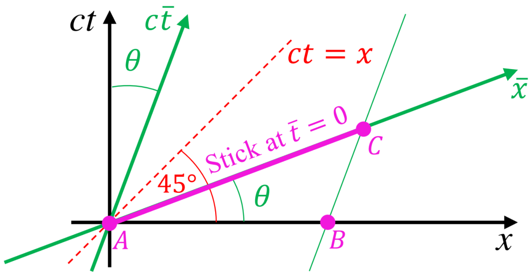
$$x^B - x^A = x^B = x^C - vt^C$$

$$\bar{l} = \bar{x}^C - \bar{x}^A = \bar{x}^C$$

because we chose our systems to fulfil $x^A, \bar{x}^A = 0$. We now compute the speed:

$$v = \frac{x^C - x^B}{t^C - t^B} \Rightarrow x^C - x^B = vt^C$$

$$\begin{pmatrix} c\bar{t}^C \\ \bar{x}^C \end{pmatrix} = \gamma \begin{pmatrix} 1 & -\frac{v}{c} \\ -\frac{v}{c} & 1 \end{pmatrix} \begin{pmatrix} ct^C \\ x^C \end{pmatrix}$$



Using the above equations, we obtain:

$$\begin{cases} c\bar{t}^C = \gamma(ct^C - \beta x^C) = 0 \\ \bar{x}^C = \gamma(x^C - vt^C) = \bar{l} \end{cases}$$

$$\boxed{\bar{l} = \gamma l}$$

So the staff is the **longest in its rest frame!**

1.3.2. Time dilation

We place a clock that rests in $\bar{O}$ and ticks every $\Delta\bar{t}$. $\bar{O}$ moves again at v in the x-direction relative to O. We are looking for Δt. We examine two events:

$$\begin{aligned} \text{First tick:} \quad & (\bar{x}, \bar{t}_1) \leftrightarrow (x_1, t_1) \\ \text{Second tick:} \quad & (\bar{x}, \bar{t}_2) \leftrightarrow (x_2, t_2) \end{aligned}$$

We carry out the LT and receive:

$$\begin{aligned} c\Delta\bar{t} &= \gamma c\Delta t - \gamma\frac{v}{c}\Delta x \\ 0 = \Delta\bar{x} &= -\gamma\frac{v}{c}\cdot c\Delta t + \gamma\Delta x \end{aligned}$$

We put Δx from the second equation into the first one and get:

$$\begin{aligned} c\Delta\bar{t} &= \gamma\cdot c\underbrace{\left(1-\beta^2\right)}_{1/\gamma^2}\Delta t \\ &= \frac{c}{\gamma}\Delta t \end{aligned}$$

$$\Rightarrow \boxed{\Delta t = \gamma\Delta\bar{t} > \Delta t}$$

Times are thus the **shortest in their rest frame!**

1.3.3. Composition of Velocities

Unlike the Galilei transformation, the Lorentz transformation **does not** simply add up velocities:

$$\Lambda(\boldsymbol{v}_1)\Lambda(\boldsymbol{v}_2) \neq \Lambda(\boldsymbol{v}_2)\Lambda(\boldsymbol{v}_1) \neq \lambda(\boldsymbol{v}_1 + \boldsymbol{v}_2)$$

We envision three systems, where $\bar{O}$ moves at velocity v relative to O and $\bar{\bar{O}}$ moves at w relative to $\bar{O}$. The added velocity is then (proof exercise 1, task 1)

$$\Lambda(w)\Lambda(v) = \Lambda\left(\frac{v+w}{1+\frac{vw}{c^2}}\right)$$

1.4. Intervals and Lorentz-Transformations

In "classical" mechanics (Galilei transformation, hereafter called *Galilean mechanics*) we know that the norm φ^{CS} is invariant under

- 3D-rotations
- 3D-translations

We measure the norm by $d = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$.
As we have already seen lengths and times are not preserved in SR, but it is true that φ^{SR} is invariant under

- 3D-rotations
- 4D-translations
- Lorentz-Boosts

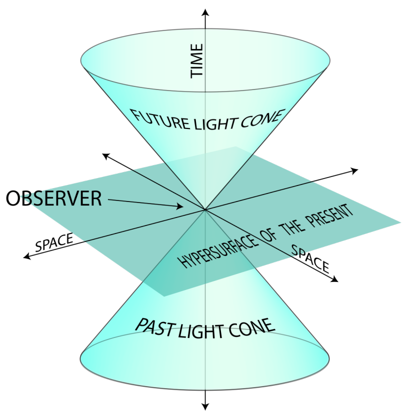
So we need a new invariant under Lorentz transformations. This is (proof in exercise 1, task 2)

$$\boxed{\Delta s^2 = -c^2 \Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2}$$

The square should be viewed with caution, as it is **not positive definite**. In particular:

- $\Delta s^2 = 0$ Events are separated by a light ray, we call them *Light-like interval*
- $\Delta s^2 < 0$ Events are separated by something slower than light. We call them *Time-like interval*
- $\Delta s^2 > 0$ Events can not be connected by any world line, because the speed of light would have to be exceeded. We call them *space-like interval*

Since Δs^2 is invariant under LT, the type of event **is the same in all inertial frames**. In particular, however, this also **removes** the concept of simultaneity (see ex. 1, task 3). This makes it clear that only certain events (time-like, light-like) can be perceived by us at all. We now want to write down the general Lorentz transformation, which consists of a Lorentz boost and a 3D rotation. To simplify the formulas we will consider the case of a x-direction boost and a rotation around the z-axis:



$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

1.5. Lorentz-Algebra

1.5.1. Absolute geometrical objects

Since we want to transform from one coordinate system into another, but despite frame dependency, things need to exist independently of coordinate systems[1], we try to find well-defined objects, which we call *absolute geometrical objects* and which we can use to construct other physical variables. These are for example:

- *Event*: An event is well-defined and has a place and a time, even though the individual components $\{x_\alpha^E\}$ are frame-dependent.
- *Vector between two Events:* Since events are well-defined, the vector $\overrightarrow{E_1E_2}$ is also well-defined. Again, although $\vec{v} = \{v_\alpha\}$ is absolute, its components change when transforming into another base.
- *Scalar fields:* A scalar function $\Phi(E)$ is absolute, it too can be represented in a base, with $\Phi(x_\alpha)$ being frame-dependent though. If we now transform an event into another coordinate system, we also need to transform the scalar field to preserve totality.

$$\begin{aligned} E : x_\alpha \mapsto \bar{x}_\alpha \\ \Rightarrow \Phi(x_\alpha) = \bar{\Phi}(\bar{x}_\alpha) \end{aligned}$$

- *Scalar products* $\vec{v}_A,\ \vec{v}_B \quad \Rightarrow \quad s = \vec{v}_A \cdot \vec{v}_B$. The scalar product is completely **independent of basis** in the sense, that

$$\vec{v}_A(x^\alpha) \cdot \vec{v}_B(x^\beta) = \vec{v}_{\bar{A}}(\bar{x}^\alpha) \cdot \vec{v}_{\bar{B}}(\bar{x}^\beta)$$

  It is thus a **Lorentz-invariant**.

- *Direct product* $\vec{v}_A \otimes \vec{v}_B$ with

$$(\vec{v}_A \otimes \vec{v}_B)_{\alpha\beta} = v_{A,\alpha} v_{B,\beta}$$

1.5.2. Lorentz-Algebra with standard vectors and matrices

For an event we will use the notation:

$$E : \quad x_\alpha, \quad \bar{x}_{\bar{\alpha}} \quad (\text{or } \beta \text{ or} \ldots)$$

and for a vector:

$$\vec{AB} : \quad v_\alpha, \quad \bar{v}_{\bar{\alpha}}$$

[1] We cannot just make objects disappear by switching frames

For a LT we need a sum, which we will usually leave out in the following (Einstein notation):

$$\bar{x}_{\bar{\alpha}} = \sum_{\alpha} \Lambda_{\bar{\alpha}\alpha} x_{\alpha}$$

The summation over α is often called *contraction* (since we essentially contract one dimension). The indices are switched for the reverse-transformation:

$$x_{\alpha} = (\Lambda^{-1})_{\alpha\bar{\alpha}} \bar{x}_{\bar{\alpha}}$$

Vectors transform like coordinates:

$$\bar{v}_{\bar{\alpha}} = \sum_{\alpha} \Lambda_{\bar{\alpha}\alpha} v_{\alpha}$$

When we look at the gradient of a scalar field it quickly becomes clear why we need co- and contravariant vectors. Here we will use the names *vector* and *covector*. If we consider a small variation $\mathrm{d}\Phi$ we obtain:

$$\mathrm{d}\Phi = \frac{\partial \Phi}{\partial x_{\alpha}}\ \mathrm{d}x_{\alpha}$$

By LT, we realize that we need an object that transforms with Λ^{-1} to leave the scalar field invariant:

$$\mathrm{d}\bar{\Phi} = \Lambda^{-1}_{\bar{\alpha}\alpha} \frac{\partial \Phi}{\partial x_{\alpha}} \Lambda_{\bar{\alpha}\beta}\ \mathrm{d}x_{\beta}$$

We call the objects that transform withΛ^{-1} covectors. In particular, we see that scalars that are invariant under LT always have to consist of a composition of vectors and covectors.

1.5.3. Lorentz-Algebra with Covariant notation

In order to be able to unambiguously classify vectors (contravariant) or covectors (covariant) in complicated calculations, we now introduce the following notation:

- **Index up: Contravariant index** v^{α}
- **Index down: Covariant index** v_{α}
- **Latin Letters:** $i = 1, 2, 3$ (3D-space)
- **Greek Letters:** $\alpha = 0, 1, 2, 3$ (4D-spacetime)

Lorentz-tranformation

We implement our new notation:

$$\Lambda_{\bar{\alpha}\alpha} \quad \text{becomes} \quad \Lambda^{\bar{\alpha}}{}_{\alpha}$$

Here we have left the order of the indices equal to underline the tensor nature.
Thus, a vector transforms like

$$\bar{v}^{\bar{\alpha}} = \sum_{\alpha} \Lambda^{\bar{\alpha}}{}_{\alpha} v^{\alpha}$$

It can already be seen that two indices, one appearing at the top and one at the bottom are contracted (summed up). This is no coincidence, the order of tensors must be conserved, and the symmetry of the LT, as discussed in the last section, requires contracting over a co- and a contravariant object. This brings us to the next convention.

Einstein sum convention

Since we contract **always** on the same indices, one of which is a co- and one a contravariant, we will omit the sum-sign in the following. This more compact notation is called Einstein sum convention.

Tensors of rank $\substack{i \\ j}$

If we want to Lorentz-transform any tensor (e.g. $R^{\alpha}{}_{\beta\gamma}$) we need to apply the corresponding LT to every index:

$$\bar{R}^{\bar{\alpha}}{}_{\bar{\beta}\bar{\gamma}} = \Lambda^{\bar{\alpha}}{}_{\alpha} (\Lambda^{-1})^{\beta}{}_{\bar{\beta}} (\Lambda^{-1})^{\gamma}{}_{\bar{\gamma}} \; R^{\alpha}{}_{\beta\gamma}$$

Scalar products between vectors

We have already constructed the Lorentz transformation to leave Δs^2 invariant. By rewriting Δs^2 we get:

$$\begin{aligned}
\Delta s^2 &= -\Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2 \\
&= \Delta x^{\mu} \eta_{\mu\nu} \Delta x^{\nu} \\
&= \Delta \bar{x}^{\bar{\mu}} \eta_{\bar{\mu}\bar{\nu}} \Delta \bar{x}^{\bar{\nu}} \\
&= \Delta x^{\mu} \Lambda^{\bar{\mu}}{}_{\mu} \eta_{\bar{\mu}\bar{\nu}} \Lambda^{\bar{\nu}}{}_{\nu} \Delta x^{\nu}
\end{aligned}$$

which gives us a transformation for $\eta_{\mu\nu}$:

$$\boxed{\eta_{\mu\nu} = \Lambda^{\bar{\mu}}{}_{\mu} \eta_{\bar{\mu}\bar{\nu}} \Lambda^{\bar{\nu}}{}_{\nu}}$$

We call $\eta_{\mu\nu}$ the *Minkowski-* or *metric tensor*. It is of rank

$$\text{Rang} : \begin{pmatrix} 0 \to \text{ “Contravariant” rank} \\ 2 \to \text{ “Covariant” rank} \end{pmatrix}$$

Furthermore we can explicitly write out $\eta_{\mu\nu}$:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Note: This metric only holds in SR, in GR it becomes considerably more complicated.

The metric tensor enables us to rewrite the scalar product as:

$$\boxed{\vec{v} \cdot \vec{w} = v^\alpha \cdot \eta_{\alpha\beta} \cdot w^\beta = v_\alpha w^\alpha}$$

Explicit LT in particular shows that the scalar product is a **Lorentz-scalar**. (It is also visible the structure, as it is a contraction of a co- and a contravariant vector)

We furthermore see from

$$v^\mu \, \eta_{\mu\nu} = v_\nu$$

that we can identify η as a *index-lowering* tensor. Assuming the existence of an *index-raising* tensor $N^{\mu\nu}$, we immediately find that this tensor should be the inverse of $\eta_{\mu\nu}$. With $\eta_{\mu\nu}$ being its own inverse this gives:

$$\boxed{N^{\mu\nu} = (\eta_{\mu\nu})^{-1} = \eta_{\mu\nu} = \eta^{\mu\nu}}$$

1.6. Lorentz-vectors for physical variables

In classical mechanics we defined our variables as 3-variables ($\boldsymbol{v}$,$\boldsymbol{a}$,$\boldsymbol{F}$,...) and found some laws ($\dot{E} = 0, m\boldsymbol{a} = \boldsymbol{F}$). However, these no longer necessarily apply in SRT. In particular, we want to express all quantities by 4-vectors, so that we get back the structure of our equations after Lorentz transformation on both sides. For this, we have to generalize our old laws to new ones. This first requires some definitions:

1.6.1. Proper time

If we again use the concept of world lines, we can draw a small distance separating two objects into our diagram

$$\mathrm{d}\vec{x} = (\mathrm{d}x^0, \mathrm{d}x^1, \mathrm{d}x^2, \mathrm{d}x^3)$$

We then have two possibilities for constructing the proper time:

- Constant velocity: There is a rest frame fulfilling $\mathrm{d}\vec{x} = (\mathrm{d}\bar{x}^0, 0, 0, 0)$. In this frame the proper time is given by $t = \bar{x}^0$.

- Acceleration: At any given moment we can find a rest frame with the same characteristics as above (*MCRF, Momentarily Comoving Reference Frame*). t is then given as the integral of $\mathrm{d}\bar{x}^0$ over all MCRFs.
  The proper time is the time, that is **measured by an observer in the MCRF**, while the coordinate time is obviously dependent on the choice of coordinates and **does not always have a direct physical interpretation**.

  Note: In literature the proper time is often written as τ (in contrast to the *coordinate time* t). For now t will be written as x^0 though, which is also a consistent notation, later in GR we will use τ and t when we switch to spherical coordinates.

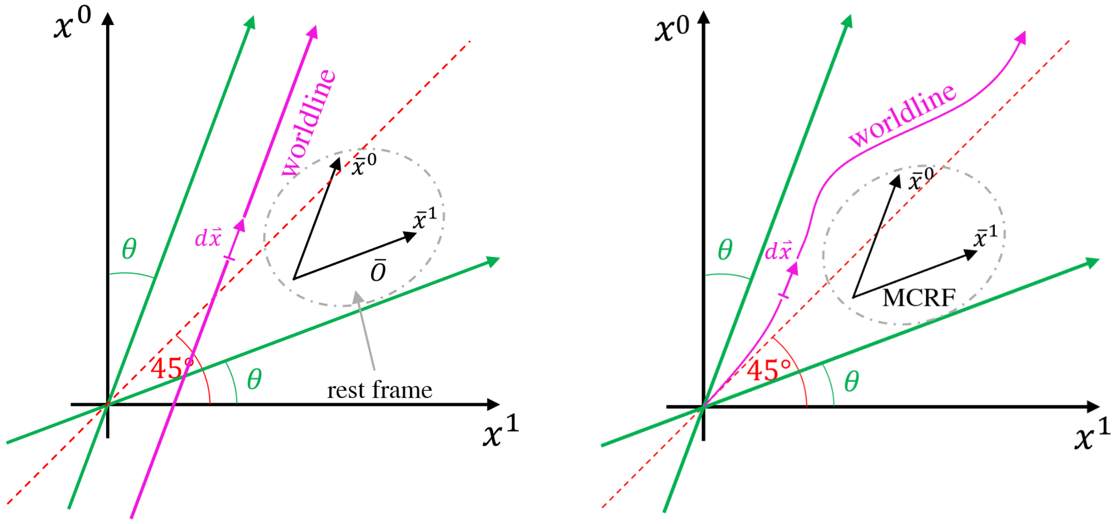


Figure 1.2.: World lines for a inertial frame (left) and an accelerated frame (right). It immediately becomes clear why the condition we imposed on direction and norm uniquely defines the proper time.

1.6.2. 4-velocity

We motivate the definition of 4-velocity from classical mechanics where the velocity is given as $\boldsymbol{v} = \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}$. We generalize this to SR by taking the 4-vector and the proper time. This is a Lorentz vector since the length transforms with Λ and t is a Lorentz scalar:

$$\boxed{\vec{U} = \frac{\mathrm{d}\vec{x}}{\mathrm{d}t}} \tag{1.6.2.1}$$

$\vec{U}$ is **tangent to its world line**. In particular we find for the norm of $\vec{U}$:

$$\begin{aligned}\vec{U}\cdot\vec{U} &= \frac{\mathrm{d}x^\alpha}{\mathrm{d}t}\eta_{\alpha\beta}\frac{\mathrm{d}x^\beta}{\mathrm{d}t}\\ &= \frac{\mathrm{d}x^\alpha\eta_{\alpha\beta}\mathrm{d}x^\beta}{\mathrm{d}t^2}\\ &= \frac{\mathrm{d}s^2}{\mathrm{d}t^2} = -\frac{\mathrm{d}t^2}{\mathrm{d}t^2} = -1\end{aligned}$$

So we can alternatively define $\vec{U}$ as the Lorentz-vector which is tangent to the world line with norm -1.
In practice we sometimes see the case where only x^0 is known and $t(x^0)$ has a complicated form. Then we use the chain rule and get

$$\vec{U} = \frac{\partial\vec{x}}{\partial x^0}\frac{\mathrm{d}x^0}{\mathrm{d}t} \quad\Rightarrow\quad \left(\frac{\partial\vec{x}}{\partial x^0}\right)^2\left(\frac{\mathrm{d}x^0}{\mathrm{d}t}\right)^2 = -1$$

Which gives us a handy way of calculating $\frac{\mathrm{d}x^0}{\mathrm{d}t}$:

$$\dot{x}^0 := \frac{\mathrm{d}x^0}{\mathrm{d}t} = \sqrt{-\frac{1}{\left(\frac{\partial\vec{x}}{\partial x^0}\right)^2}}$$

(We will use this relation in ex. 3, task 3.3)

The explicit form of the 4-velocity is given by:

$$\vec{U}_{MCRF} = (1,0,0,0) \quad\Rightarrow\quad \vec{U} = (\gamma,\gamma v_1,\gamma v_2,\gamma v_3) \tag{1.6.2.2}$$

1.6.3. 4-acceleration

The definition of the 4-acceleration follows directly from the definition of the 4-velocity:

$$\vec{a} = \frac{\mathrm{d}\vec{U}}{\mathrm{d}t} = \frac{\mathrm{d}^2\vec{x}}{\mathrm{d}t^2}$$

In particular the 4-acceleration of an arbitrary MCRF is always **perpendicular** to its 4-velocity. Geometrically the four-acceleration is the *curvature* of a world line.

1.6.4. 4-momentum

We define the 4-momentum in analogy to the classical momentum as

$$\boxed{\vec{p} = m\cdot\vec{U}}$$

We immediately see, that $\vec{p}$ is a Lorentz-vector (m is a frame-independent scalar). We can then perform a LT into another frame and see that

$$\vec{p}_{MCRF} = (m, 0, 0, 0) \quad \Rightarrow \quad \vec{p} = (m\gamma, m\gamma v_1, m\gamma v_2 m\gamma v_3) = (E, p_1, p_2, p_3)$$

whereas E is the energy in our frame and $\boldsymbol{p} = p_i$ is the 3-momentum. Why is the 0-component E the energy? This question will be clarified below:
First we do a Lorentz-boost in the x-direction out of our rest frame.

$$\bar{p} = (m, 0) \Rightarrow p = \gamma \begin{pmatrix} 1 & v \\ v & 1 \end{pmatrix} \bar{p} = (\gamma m, \gamma m v)$$

In order to find a connection with the energy we first look at the classical limit:

$$\begin{aligned} E = \gamma m = \frac{m}{\sqrt{1 - v^2}} &\overset{v \ll 1}{\Rightarrow} \left(1 + \frac{v^2}{2} + \mathcal{O}(v^4)\right) \\ &\Rightarrow E \approx \underbrace{m}_{\text{rest energy}} + \underbrace{m\frac{v^2}{2}}_{\text{kinetic energy}} \end{aligned}$$

We see that the mass is independent of the observer (at least according to our definition of the mass, often in the literature E is called "mass". While this is a valid way of defining mass, we then have to call m the rest mass which would not correspond to our intuitive perception of mass). We calculate the square of $\vec{p}$:

$$\begin{aligned} \vec{p} \cdot \vec{p} &= m\vec{U} \cdot m\vec{U} = -m^2 \qquad \text{as well as} \\ p^\alpha \eta_{\alpha\beta} p^\beta &= -E^2 + \sum_i (p^i)^2 \end{aligned}$$

From this we get the (very famous and practical) equation

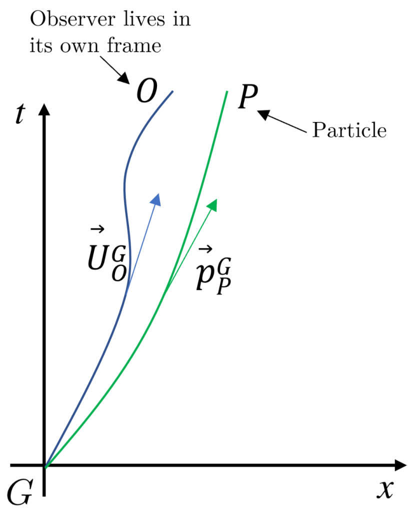
$$\boxed{E^2 = m^2 + \sum_i (p^i)^2} \; .$$

In summary, we find the following properties of the the 4-momentum:

- $\vec{p} \parallel$ Worldline
- $\vec{p} \cdot \vec{p} = -m^2$
- $p^0 = E$

1.6.5. Energy in arbitrary frames

We imagine a particle and an observer moving at different speeds. We call the particle's rest frame P and the observer's O. We look at both systems from a third frame G which moves relative to O and P.
We know $\vec{p}_{\text{particle}}$ and $\vec{U}_{\text{observer}}$ as seen from G. If we now want to compute the particle's energy as seen by the observer.
Since each of the three systems sees a different energy of the particle, we need to define a Lorentz-invariant which we can compute in any frame which will give us (E_{Obs}). We define E_{Obs} as



$$
\begin{aligned}
E_{\text{Obs}} &= -\,\vec{p}_p^{\,G} \cdot \vec{U}_O^G \\
&\overset{\text{Lorentz-scalar}}{=} -\,\vec{p}_p^{\,\text{Obs}} \cdot \vec{U}_O^{\text{Obs}} \\
&= \vec{p}_p^{\,0,\text{Obs}} = E_{\text{Obs}}
\end{aligned}
$$

Since this is a scalar product of two Lorentz-vectors it is invariant under LT and thus gives us E_{Obs} regardless of the frame we compute it in:

$$E_{\text{Obs}} = p_p^0 \cdot U_O^0 - \sum_i p_p^i U_0^i$$

By explicitly calculating the scalar product we see that E_{Obs} in fact corresponds to

$$E_{\text{Obs}} = m \cdot \gamma(\boldsymbol{v}_{\text{Obs}}, \boldsymbol{v}_{\text{particle}}) \ .$$

Now that we have our energy *defined*, we need to verify that it satisfies definition of energy in a physical sense. In particular we need to check the conservation of energy. To do this, we define the 4-force and examine the conservation laws.

1.6.6. 4-force

The 4-force is again defined in analogy to the classical force (derived in [6]) and gives us postulate C (see 1.1):

$$\boxed{\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = m\frac{\mathrm{d}\vec{U}}{\mathrm{d}t} = m\vec{a} = \vec{F}}$$

We see no (obvious) connection between 4-force and 3-force from this definition. This will be discussed later.

Using this definition we see that if no force is applied to an object conservation of energy and momentum still hold:

$$\vec{F} = 0 \Rightarrow \frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = 0 \Rightarrow \begin{cases} \frac{\mathrm{d}E}{\mathrm{d}t} = 0 & \text{conservation of energy} \\ \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = 0 & \text{conservation of momentum} \end{cases}$$

1.6.7. Massless particles

For a massless particle, we have the problem that we cannot find a rest frame since in this frame $c = 1$ would be violated. This also prevents us from finding a proper time and therefore we have no valid definition for $\vec{u} = \frac{\mathrm{d}\vec{x}}{\mathrm{d}t}$. Taking the limit $m \to 0$ would also not solve the problem since $\vec{u}$ would diverge while while $ds^2 = -1$. This forces us to chose a different approach by just *defining*

$$\boxed{\vec{p} = (E, p_1, p_2, p_3)}$$

with

$$\boxed{p_\alpha p^\alpha} = 0$$

We define E as the energy, which we measure in our reference frame, which allows us to construct the momentum from the direction of the particle and the condition above. We check if these definitions are compatible with our previous ones:

- $\vec{p}$ is (obviously) a 4-vector
- $E^2 = m^2 + \boldsymbol{p}^2$ is consistent since $m = 0$ gives us the correct energy
- The energy which an observer measures is $E = h \cdot \nu$

With this we can now calculate the relativistic Doppler effect.

1.6.8. Doppler effect

We look at two reference frames G and $\bar{G}$, with

$$\begin{aligned} \vec{p} &= (E, -E, 0, 0) \qquad & E &= h\nu \\ \vec{\bar{p}} &= (\bar{E}, -\bar{E}, 0, 0) \qquad & \bar{E} &= h\bar{\nu} \end{aligned}$$

(Each of the frames sees the light ray moving away from it).

$$\begin{aligned} \bar{p}^\alpha &= \Lambda^\alpha{}_\beta p^\beta \\ \bar{p}^0 &= \Lambda^0{}_\beta p^\beta = \gamma E + \gamma v E \\ \Rightarrow h\bar{\nu} &= \gamma h\nu(1+v) = h\nu\sqrt{\frac{1+v}{1-v}} \end{aligned}$$

We thus get a shift in frequency of:

$$\boxed{\frac{\bar{\nu}}{\nu} = \sqrt{\frac{1+v}{1-v}}}$$

In the classical limit we get the same proportionality as with the normal (acoustic) Doppler effect:

$$\frac{\Delta\nu}{\nu} = \sqrt{\frac{1+v}{1-v}} - 1 = v + \mathcal{O}(v^2)$$

1.7. Relativistic hydrodynamics

This part will focus on relativistic hydrodynamics. The goal here is to transfer $n(\boldsymbol{x})$, $\rho(\boldsymbol{x})$, $P(\boldsymbol{x})$, $T(\boldsymbol{x})$, that is, thermodynamic quantities to the theory of relativity. We will look at three different scenarios:

- **Dust:** Collision less particles, particles keep their velocities
- **Perfect fluid:** Strong interaction between the particles $\Rightarrow$ (isotropic) pressure, thermodyn. equilibrium
- **Imperfect fluid:** Weak interactions $\Rightarrow$ viscosity, no thermodyn. equilibrium.

We will look at the first two cases as limits of the (only realistic) third case and introduce the concepts of the *number-flux-vector* and the *stress-energy tensor* by using the exemplary cases above.

1.7.1. Number flux vector

As we will see below, the number density $n = \frac{N}{V}$, which in thermodynamics is a conserved, intensive variable (at least in thermodynamic equilibrium), is not conserved in relativity due to the volume being observer-dependent. Therefore we need a new type of 4-vector, which describes the number density and transforms like a Lorentz-vector. To construct this we will look at two cases:

Dust with uniform velocity

In the case of dust, where all particles have the same velocity $\boldsymbol{v}$ there is an MCRF in which all particles rest. We can then define the *number density* $n(x^\alpha) = \frac{N}{V}$ which is **not** Lorentz-invariant, which we can see if we perform a Lorentz-Boost into another frame:

$$O \quad \to \quad \bar{O} \quad \text{with velocity } v \text{ in } \boldsymbol{e}_1\text{-direction}$$

$$V = \Delta x^1 \cdot \Delta x^2 \cdot \Delta x^3 \to \bar{V} = \frac{1}{\gamma} \Delta x^1 \cdot \Delta x^2 \cdot \Delta x^3$$

We thus get

$$\bar{n} = \frac{\gamma N}{V} = \gamma n$$

One can easily see, that n transforms like the 0-component of a Lorentz-vector.
Hence we are searching for a 4-vector with the components

$$(\vec{\dots}) = (n, f^1, f^2, f^3)$$

We identify the 3-vector $\boldsymbol{f}$ as the *particle number flux*. With this in mind we construct f^1:

$$\begin{aligned} f^1 &= \frac{\#\text{particles which cross the } \perp \boldsymbol{e}_1\text{-plane}}{\text{surface} \cdot \text{time}} \\ &= \frac{n \cdot S \cdot v^1 \Delta t}{S \cdot \Delta t} = n v^1 \end{aligned}$$

Now we want to check if we indeed constructed a Lorentz vector. If we take the MCRF of the dust O and another frame $\bar{O}$ which is moving in $\boldsymbol{e}_1$-direction relative to O we get

$$\begin{aligned}\bar{f}^1 &= f^1 \cdot (\text{Velocity of } O \text{ in } \bar{O}) \\ &= \bar{n} \cdot (-v) = -\gamma n v\end{aligned}$$

We thus get the transformation from $(n, 0, 0, 0)$ in O to $(\gamma n, -\gamma n v, 0, 0)$ in $\bar{O}$ which is the desired property for our 4-vector. We call this Lorentz-vector *number flux* $\vec{N}$. In general we can write it as

$$\begin{aligned}\vec{N} &= n_{\text{MCRF}} \cdot \vec{U} \\ \vec{N} \cdot \vec{N} &= n^2_{\text{MCRF}} \cdot \vec{U} \cdot \vec{U} = -n^2_{\text{MCRF}}\end{aligned}$$

From the norm of $\vec{U}$ we obtain a very easy way to get n_{MCRF}:

$$\boxed{n_{\text{MCRF}} = \sqrt{-\vec{N} \cdot \vec{N}} = \sqrt{-N_\mu N^\mu}}$$

Particles with arbitrary velocities

If we have many particles with arbitrary velocities, it is clear, that we can only get a Lorentz-vector if we define $\vec{N}$ as

$$\vec{N} = \sum_p \frac{1}{V_{\text{MCRF},p}} \vec{U}_p \ .$$

The reason here is that $\vec{N}$ can only be a Lorentz-vector if the factor $\frac{1}{V}$ from above is Lorentz-invariant. We therefore see, that we need to calculate V in every particle's rest frame and cannot just simply factor out $\frac{1}{V}$ from the sum. Fortunately, however, we have a relation for scaling the volume $\bar{V} = \gamma V$. This gives us

$$\begin{aligned}\vec{N} &= \sum_p \frac{1}{V \gamma_p} \vec{U}_p \\ &= \frac{1}{V} \sum_p \gamma_p^{-1} \vec{U}_p \\ &= \left(\frac{N}{V}, \frac{\sum_p v_p^1}{V}, \frac{\sum_p v_p^2}{V}, \frac{\sum_p v_p^3}{V} \right)\end{aligned}$$

For this, the definitions of $\vec{U}_p$ from (1.6.2.2) was used. But since $\vec{N}$ is a 4-vector, we can find a frame with $\boldsymbol{N} = 0$ and have thus found a $N^0 = n_{\text{MCRF}}$. The "MCRF" is to be viewed with caution, since this is no *true* rest frame (the thermal velocity of the particles still exists), but only moves at the *bulk velocity* of the particles.

1.7.2. Stress-energy-tensor

Let us now turn to the energy density $\rho(x^\alpha)$. It shall be of the form

$$\frac{\sum_p E_p}{V} = \frac{N\langle E_p\rangle}{V} \ .$$

Again we look at different cases:

Dust

If we naively transform from the MCRF into any arbitrary frame we get

$$\begin{aligned} E: m &\quad\mapsto\quad \gamma m \\ V: v &\quad\mapsto\quad \gamma^{-1} V \end{aligned}$$

This means, that ρ transforms like

$$\rho = \frac{Nm}{V} \quad\mapsto\quad \bar{\rho} = \frac{N\gamma m}{\gamma^{-1}V} = \gamma^2 \rho$$

So we have to consider an object of a higher rank than a vector. We define this tensor so

$$T^{\alpha\beta} \to T^{00} = \rho$$

Which gives us the right behaviour under LT. We define the whole tensor as the tensor product (derived in [6]):

$$T = \vec{p} \otimes \vec{N} = m \cdot n_{\text{MCRF}} \cdot \vec{U} \otimes \vec{U}, \qquad T^{\alpha\beta} = T^{\beta\alpha} = m \cdot n_{\text{MCRF}} \cdot U^\alpha U^\beta$$

Since two identical vectors occur in the tensor product, the stress energy tensor (SET) is symmetric and of rank $\binom{2}{0}$. Since furthermore $\vec{v} = (1,0,0,0)$ in the MCRF, this means that in this frame T is 0 in all components except $T^{00} = \rho_{\text{MCRF}}$.

Let's explicitly calculate an LT of a SET. Imagine again a frame $\bar{O}$, which moves relative to the MCRF with $-\boldsymbol{v}$. We are boosting in a general direction. Then $\Lambda(-\boldsymbol{v})$ takes the following form:

$$\Lambda(-\boldsymbol{v}) = \begin{pmatrix} \gamma & \gamma v^1 & \gamma v^2 & \gamma v^3 \\ \gamma v^1 & & & \\ \gamma v^2 & & & \\ \gamma v^3 & & & \end{pmatrix}$$

The unspecified part contains more complicated terms, luckily we do not need those terms. With this we get

$$\begin{aligned} \bar{T}^{\bar{\alpha}\bar{\beta}} &= \Lambda(-\boldsymbol{v})^{\bar{\alpha}}{}_{\alpha} \cdot \Lambda(-\boldsymbol{v})^{\bar{\beta}}{}_{\beta} \cdot T^{\alpha\beta} \\ &= \Lambda(-\boldsymbol{v})^{\bar{\alpha}}{}_{0} \cdot \Lambda(-\boldsymbol{v})^{\bar{\beta}}{}_{0} \cdot T^{00} \end{aligned}$$

We identify the components of $\bar{T}^{\alpha\beta}$ as

$$\begin{aligned}
&\bullet\bar{T}^{00} = \gamma^2 T^{00} = \gamma^2\rho = \bar{\rho} && = \text{energy density}\\
&\bullet\bar{T}^{0i} = \bar{T}^{i0} = \gamma^2 v^i T^{00} = \bar{\rho} v^i && = \text{energy flux in } \boldsymbol{e}_i\text{-direction}\\
&\quad = (m\gamma^i v^i)(\gamma n_{\text{MCRF}}) = p^i \bar{n} && = \text{momentum density in } \boldsymbol{e}_i\text{-direction}\\
&\bullet\bar{T}^{ij} = \gamma^2 v^i v^j T^{00} = \gamma^2 v^i v^j m \cdot n_{\text{MCRF}} &&\\
&\quad = (\gamma m v^i)(\gamma n_{\text{MCRF}}) v^j = p^i \bar{n} v^j = p^j \bar{n} v^i && = j\text{-momentum flux in } \boldsymbol{e}_i\text{-direction}\\
& && = i\text{-momentum flux in } \boldsymbol{e}_j\text{-direction}
\end{aligned}$$

Gas

In the general case of gas, we always get a $T^{\alpha\beta}$, but it takes very different shapes, depending on whether it is a gas with (without) internal (external) interactions or fields (scalar or vector fields). We will fist look at the case of a gas of **collisionless** particles, the case of collisions is discussed below in 1.7.2 We already know

$$\begin{aligned}
T &= \sum_p m \frac{1}{V_{\text{MCRF}}} \vec{U}_p \otimes \vec{U}_p \\
&= \frac{1}{V} \sum_p m\gamma_p^{-1} \vec{U}_p \otimes \vec{U}_p = \frac{1}{V} \sum_p \gamma_p^{-1} \vec{p}_p \otimes \vec{U}_p
\end{aligned}$$

If we again look component by component:

$$\begin{aligned}
T^{00} &= \frac{1}{V} \sum_p m\gamma_p && = \text{energy density}\\
T^{0i} &= \frac{1}{V} \sum_p m\gamma_p^{-1}\gamma_p\gamma_p v_p^i = \frac{1}{V}\sum_p (\gamma_p m v_p^i) = \frac{1}{V}\sum_p p_p^i && = \text{momentum density in } \boldsymbol{e}_i\text{-direction}\\
= T^{i0} &= \frac{1}{V} \sum_p E_p v_p^i && = \text{energy flux in } \boldsymbol{e}_i\text{-direction}\\
T^{ij} &= \frac{1}{V} \sum_p p_p^i v_p^j && = i\text{-momentum flux in } \boldsymbol{e}_{\boldsymbol{j}}\text{-direction}\\
&= \frac{1}{V} \sum_p p_p^j v_p^i && = j\text{-momentum flux in } \boldsymbol{e}_{\boldsymbol{i}}\text{-direction}
\end{aligned}$$

In component notation this gives us:

$$\begin{aligned}
T^{\alpha\beta} &= \frac{1}{V} \sum_p \gamma_p^{-1} p_p^\alpha u_p^\beta \\
&= \frac{1}{V} \sum_p \gamma_p^{-1} m^{-1} p_p^\alpha p_p^\beta \\
&= \frac{1}{V} \sum_p (p_p^0)^{-1} p_p^\alpha p_p^\beta
\end{aligned}$$

We can no proceed to go to the continuous limit by using the density in phase-space

$$f = \frac{\#\text{particles}}{(\text{space volume})(\text{momentum volume})} = \frac{\mathrm{d}N}{\mathrm{d}^3x\ \mathrm{d}^3p}$$

The stress-energy-tensor is then given as

$$\boxed{T^{\mu\nu} = \int \mathrm{d}^3p\ f \cdot \frac{p^\mu p^\nu}{p^0}}\ . \tag{1.7.2.1}$$

This is the very famous formula for the *stress-energy-tensor of a collisionless gas* in SRT. (In fact, we will see that almost the same is true in GR).
Note, however, that the stress-energy tensor is a **local** quantity. Even if we defined it in terms of a volume, we only look at the case, where $V \to 0$, so at an infinitesimally small region around a given point.

General case

Even though, in the most general case, the interpretation of $T^{\alpha\beta}$ stays the same as above we cannot find a rest frame, in which $T^{\alpha\beta}$ becomes trivial. To prove this we look at a gas of colliding particles:

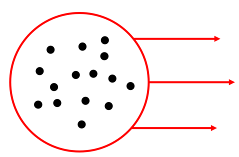
1. $$T^{00} = \frac{\sum_p \sqrt{m^2 \sum_i (p_p^i)^2}}{V}$$

   We see, that even if we define the MCRF to fulfil $\sum_i p_p^i = 0$, the kinetic energy does not vanish in this frame. However it is minimal at this point.

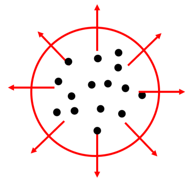
2. T^{0i} vanishes in the MCRF for collision-free particles, but not in the general case. We see in particular that T^{0i} vanishes in thermodynamic equilibrium but not if the system is not in thermodyn. equilibrium (we define thermodynamic equilibrium as a state, in which the particles interact on very short timescales).

3. $T^{ij} \neq 0$ in MCRF, even in the collision-less case. In particular, in MCRF we can identify T^{ij} as 3×3 *stress-tensor*, which describes the **internal forces**, which act on the fluid. As an analogy, we consider a spherical balloon in which the particles are located. With the force $\boldsymbol{F} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t}$ the balloon can

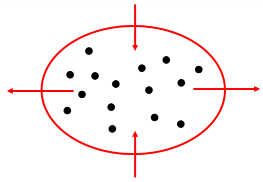
   a) move due to bulk momentum

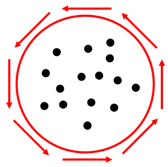
   b) expand due to isotropic pressure



c) deform due to anisotropic pressure



d) rotate due to viscosity



These possibilities are encoded in the 6 degrees of freedom of the (symmetric) T^{ij} tensor.
In particular, we also see that only case b) is compatible with isotropy (we will see that isotropy is often a condition we can assume in GR). In this case T^{ij} is **diagonal** in the MCRF (shown in ex. 03, task 2) and we get

$$\boxed{T^{ij} = \mathfrak{p}\delta^{ij}}$$

Note: In the following I will use $\mathfrak{p}$ as symbol for the **pressure** to avoid confusion with the momentum.

Perfect fluid

In the case of a perfect fluid, there are strong interactions between the particles. This leads to the following effects:

1. Thermodynamic equilibrium at any time $\Rightarrow$ no heat conduction

$$\Rightarrow T^{0i} = T^{i0} = 0$$

   in local MCRF

2. No viscosity $\Rightarrow$ only isotropic pressure

$$\Rightarrow T^{ij} = \mathfrak{p}\delta^{ij}$$

   (see above, shown in ex. 02, task 2).

So the stress energy tensor is represented by

$$T^{\alpha\beta} = \begin{pmatrix} \rho_{\mathrm{MCRF}} & 0 & 0 & 0 \\ 0 & \mathfrak{p}_{\mathrm{MCRF}} & 0 & 0 \\ 0 & 0 & \mathfrak{p}_{\mathrm{MCRF}} & 0 \\ 0 & 0 & 0 & \mathfrak{p}_{\mathrm{MCRF}} \end{pmatrix}$$

in MCRF.
Since the property *perfect fluid* should be frame-independent we need to construct $T^{\alpha\beta}$ in Terms of Lorentz-Tensors as we did with the T-Tensor in case of dust. The 4-vectors we have at our deposal are

- $u^\alpha u^\beta$ (bulk-velocity or center of mass velocity)
- $\eta^{\alpha\beta}$ (metric tensor

In the MCRF $U^\alpha U^\beta$ gives us just a 1 in the 00-component so our S-E-Tensor needs to contain a term $\rho_{\mathrm{MCRF}} U^\alpha U^\beta$. In contrary the metric tensor gives us a -1 in the 00-component and 1 for the ii-components. Since $\mathfrak{p}_{\mathrm{MCRF}}$ should cancel out in the 00-component we see that $T^{\alpha\beta}$ needs to be

$$\boxed{T^{\alpha\beta} = (\rho_{\mathrm{MCRF}} + \mathfrak{p}_{\mathrm{MCRF}}) U^\alpha U^\beta + \mathfrak{p}_{\mathrm{MCRF}} \eta^{\alpha\beta}} \tag{1.7.2.2}$$

Note: The same as we derived above for the perfect fluid also applies for a perfect solid[2] and a non-ideal gas with isotropic, internal interaction.

1.7.3. Conservation of the Stress-energy-Tensor

We think of a small cube from which energy can enter and leave. We then have 6 faces from which the energy can enter or leave. We express the energy gain of each face by taking T^{0i} (x^0 time, x^i =center of each face) and taking the derivative (not derived here)

$$\frac{\partial T^{00}}{\partial x^0} = -\frac{\partial T^{01}}{\partial x^1} - \frac{\partial T^{02}}{\partial x^2} - \frac{\partial T^{03}}{\partial x^3}$$

The same can also be done with $\mathfrak{p}_i$ instead of ρ which finally gives

$$\frac{\partial T^{\alpha 0}}{\partial x^0} + \frac{\partial T^{\alpha 1}}{\partial x^1} + \frac{\partial T^{\alpha 2}}{\partial x^2} + \frac{\partial T^{\alpha 3}}{\partial x^3} = 0$$

We now introduce the convenient notation

$$\boxed{(\dots)_{,\beta} = \frac{\partial(\dots)}{\partial x^\beta}}$$

[2]Perfect solid in this context means isotropy and instantaneous thermodynamic equilibrium.

With Einstein sum convention this gives

$$\boxed{T^{\alpha\beta}{}_{,\beta} = 0}$$

which gives a very clear insight in the structure of the equation. Sometimes this is called the *covariant derivative* and is sometimes written as ∇_β.

The expression derived above only holds in the case that no external forces are applied on the system. In case of external interactions the stress energy tensor is not conserved but we can summarize the external forces into F^α (actually this describes the conservation of the conservation of the 4-Momentum flux):

$$\boxed{T^{\alpha\beta}{}_{,\beta} = F^\alpha}$$

2. General Relativity

2.1. Introduction to curvature

General relativity fundamentally relies on the concepts of differential geometry. In this field of mathematics, the curvature of manyfolds plays an important role. To explain this concept we will look at to new concepts:

1. **Manyfold:** A "smooth" space, which at every point has a flat,tangent space. In principle this simply means that our space has no "spikes", where the gradient of the space is not defined.

2. **Curvature:** The curvature of space essentially means the same as what we normally perceive as as curvature (an analogy would be the curvature of the 2D surface of a 3D ball). But how do we determine curvature? There are two ways:

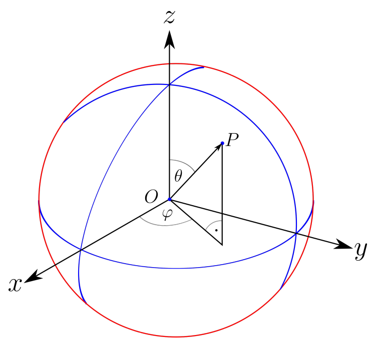
   a) With extra dimensions:
   If we introduce an extra dimension (in this case the 3rd Dimension to our 2D ball-surface) we can simply write down the equation

   $$\Delta x^2 + \Delta y^2 + \Delta z^2 = R$$

   and get the curvature. Here it is important to note that although we used our three-dimensional intuition, the law itself is two dimensional (R is fixed). This leads to the second method.

   b) Without extra dimensions:

   Here we need the definition of an (arbitrary but regular[1]) basis $\{x^i\}$ and a law of infinitesimal distances. For example on our sphere we can introduce the coordinates (θ, φ)



$$(\theta, \varphi) \to \mathrm{d}R^2 = R^2(\mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\varphi^2)$$
$$= (\mathrm{d}\theta, \mathrm{d}\varphi) \begin{pmatrix} R^2 & 0 \\ 0 & R^2 \sin^2\theta \end{pmatrix} \begin{pmatrix} \mathrm{d}\theta \\ \mathrm{d}\varphi \end{pmatrix}$$

[1]Regular means, that no different points have the same coordinates or in mathematical notation: $\bar{p} \neq p \Leftrightarrow x_{\bar{p}} \neq x_p$

We then define the infinitesimal distance $\mathrm{d}l^2$ on a flat Cartesian space (which exists for every manyfold , since we postulated a flat tangent space around any point):

$$\mathrm{d}l^2 = (\mathrm{d}x, \mathrm{d}y) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \end{pmatrix}$$

We thus get for the infinitesimal distance on the sphere

$$\begin{aligned} \mathrm{d}l^2 &= \mathrm{d}r^2 + r^2\mathrm{d}\theta^2 \\ &= (\mathrm{d}r, \mathrm{d}\theta) \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \begin{pmatrix} \mathrm{d}r \\ \mathrm{d}\theta \end{pmatrix} \end{aligned}$$

which leads to the *general law of infinitesimal distances*:

$$\boxed{\mathrm{d}l^2 = \mathrm{d}x^i g_{ij} \mathrm{d}x^j} \tag{2.1.0.1}$$

with the *symmetric metric tensor* g_{ij}

What ways do we have to measure curvature? We will briefly discuss four methods:

- Draw parallel lines (which are on an infinitesimal scale flat). On a flat space these will stay parallel indefinitely, on a curved space they may cross
- Sum the angles of a triangle. Only on a flat space the sum of the angles of a triangle is 180°.
- Try to draw a sphere with angles of $\theta/2$ between every line. In case of a flat space the lines will meet, in case of curvature not.
- Parallel Transport a vector along a closed path. In case of a flat space the vector should still point in the same direction after the transport, in case of curvature it will be rotated.

2.2. Fundamental principles of General Relativity

The fundamental principle of GR is **Galileo's equivalence principle**:

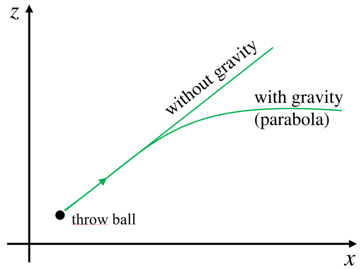
> All gravitating-only[2] bodies passing through a point x^i with a velocity v^i follows a unique trajectory (which is independent of mass, charge, etc.)

This lead Einstein to the assumption that gravity is in fact not a "force" in the sense of all other forces but rather a fundamental curvature of space(-time) and actually trajectories are just straights on a curved space(-time) (*geodesics*). Why is this a curvature of spacetime and not only of 3D space? Let us investigate both cases:

1. The curvature is the **curvature of 3D-space**:

[2] Gravitation is the only force acting on this body. This does not hold for electromagnetic, weak or strong interactions, which depend on charge, colour,. . .

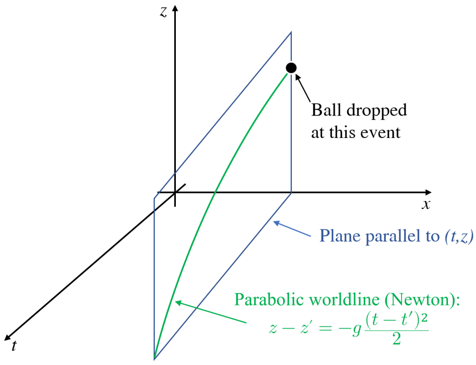
If we look at the difference of throwing a ball in space and on earth we would see that with the gravity of earth the curvature of the trajectory is much larger than without. This would imply that the curvature of space would be around the order of cm or m. We would observe this by the means which we discussed earlier.



2. The curvature is the **curvature of 4D-spacetime**:

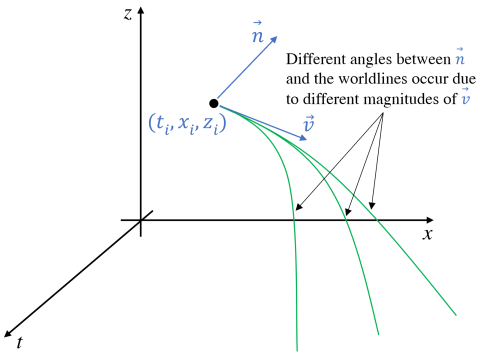
Then even if the object falls along a straight line in 3D-space (e.g. vertical fall on earth) we get a parabolic trajectory in spacetime. The curvature of this spacetime though is more realistic to what we can measure. This becomes clear if we consider a 1 m flight and use natural units:

$$t = 1\,\mathrm{s} = 3 \cdot 10^8 \quad \text{meters of time}$$



which is tiny (the radius of curvature of the order of the distance between the Earth and the moon) and corresponds more to our perception of the curvature of space, which we perceive to be flat.

We furthermore check if one initial coordinate and initial velocity results in a unique trajectory. For that we do another plot, which assumes that we throw an object at a *given event* in a *given direction* but with different initial velocities. We then see that **different v_i give different initial tangent vectors to worldlines**, like we can see in the drawing.



This means that ansatz 2. seems to fulfil the conditions we imposed. This finding leads to a major problem:
Special Relativity is only valid in inertial frames and especially assumes the existence of inertial frames. But if there is gravity at (almost) every point in space we cannot apply SR. Luckily we can simply look at the system from the point of view of an observer at rest in a *free-falling* frame in the gravitational field. This observer will feel **no acceleration**. In the (local) free falling frame, the observer has no means of knowing, that he is in a gravitational field. This leads to the **Einstein equivalence principle**:

> Physics is the same in any **free falling frame with gravity** as in any **inertial frame without gravity**.

We can now write down the **three basic principles of GR**:

(α) Special relativity is locally valid in a free falling frame (This naturally incorporates the postulates of SR)

(β) Spacetime is curved and free falling objects follow geodesics on this curved space

(γ) There is a relation between matter and curvature of space, which is called the *Einstein equation*. (In fact this will describe the dynamics of spacetime and gravity itself)

We will see that we can already solve problems of GR by applying only principle α. To look at effects of GR we now have two possibilities:

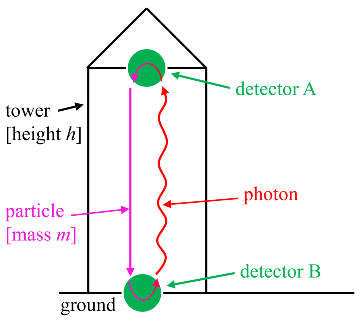
1. We can **investigate the motion of free falling frames**:
   Here we would need β and γ, which we have not sufficiently investigated. But if the gravitational field is *weak*, we can take the Newtonian limit[3]

2. We can **do an experiment in a free falling frame**:
   Here we do not need to consider the effect of curvature of space, but SR holds. If the particle is *slow* in this frame we can take the limit towards classical mechanics.

1.	strong grav. field curvature,geodesics	weak grav. field newtonian limit $\boldsymbol{a} = \boldsymbol{g}$
2.	particles fast/c w.r.t free-falling frame SR	particles slow w.r.t. free falling frame Classical mechanics

This means, that we can look at the case of weak gravitational effects and fast particles just by knowing the laws of SRT. In the following we will look at an example.

2.2.1. Gedankenexperiment: Gravitational Doppler effect



The gedankenexperiment goes as follows:
We have a tower of height h, which is equipped with two ideal detectors.

[3] This needs to hold due to the consistency of nonrelativistic limits with classical mechanics.

- Event E_1: at t_1, detector A drops a particle of mass m.
- Event E_2: At t_2, detector B collects this particle and transfers its energy to a photon, which is emitted upwards.
- Event E_3: At t_3, detector A collects the photon and transfers all its energy to a particle of mass m and drops is

and so on ...
Without GR we would not assume any impact of gravity on light but Einstein's intuition was, that the **upward photon must lose energy on its way upwards, to avoid perpetual motion** (since the particle with mass gains energy by dropping).
The guess is, that the massive particle leaves at E_1 with

$$E^{(E_1)} = m \ .$$

It reaches E_2 with

$$E^{(E_2)} = \sqrt{m^2 + p^2}$$

Since the gain in velocity is very small ($\Delta v = g\Delta t$) with $h = g\frac{\Delta t^2}{2}$, so

$$\Delta v = \sqrt{2gh} \sim 10\,\mathrm{ms}^{-1} \ll c \ .$$

Therefore we can do a Taylor-expansion of the energy and get:

$$E^{(E_2)} \approx m + \frac{1}{2}mv^2 = m(1+gh) \qquad \left(+\mathcal{O}\left(\frac{v}{c}\right)^4 \right)$$

If we assume no Doppler shift the photon would be emitted with $E^{(E_2)} = m(1+gh)$ but since we need energy to be conserved the shift in energy for the photon needs to be

$$\boxed{\frac{E^{(E_3)}}{E^{(E_2)}} = \frac{m}{m(1+gh)} = \frac{1}{1+gh} \approx 1 - gh}$$

In fact this effect was experimentally confirmed by Pound & Rebka (1960) and Pound & Snider (1965), which is astonishing, since the effect is only of the order

$$\frac{E^{(E_3)}}{E^{(E_2)}} = \frac{\nu_3}{\nu_2} \overset{\text{SI-units}}{=} 1 - \frac{gh}{c^2}, \qquad \frac{gh}{c^2} \sim 10^{-14}$$

Now we want to calculate the Doppler shift **only using principle $\boldsymbol{\alpha}$**:
We consider two free falling frames, which are both falling vertically with the acceleration g:

- O momentarily at rest with detector B at t_2, i.e. with E_2

- $\bar{O}$ momentarily at rest with detector A at t_3, i.e. with E_3

Since both frames have the same acceleration they move with a constant relative velocity and are inertial frames. We know that at t_3, $\bar{O}$ is at rest, while O is falling at

$$v_{\bar{O}/O} = g\Delta t = g(t_3 - t_2) = g\frac{h}{c} \overset{nat.units}{=} \boxed{g \cdot h}$$

If we now use our Doppler formula from SR (see 1.6.8) we get (since O is at rest with the light pulse and if $\bar{O}$ looks in direction of O it is moving away from him)

$$\frac{E^{(\bar{O})}}{E^{(O)}} = \frac{\bar{\nu}}{\nu} = \sqrt{\frac{1 - v_{\bar{O}/O}}{1 + v_{\bar{O}/O}}} < 1$$

The energy of the massive particle in O is

$$E^{(O)} = E^{E_2} = m(1 + gh)$$

since the detector B was defined in O at t_2.
We thus get

$$\boxed{\sqrt{\frac{1 - v_{\bar{O}/O}}{1 + v_{\bar{O}/O}}} \approx 1 - v_{\bar{O}/O} \approx 1 - gh \quad (+\mathcal{O}(v^4))}$$

This confirms our guess! In particular we see, that **the gravitational Doppler effect follows from the Doppler effect in SRT**.

Note: Since at the beginning we assumed, that the gravitational field, in which we are is *weak*, this equation only holds in this limit. If we want to consider stronger gravitational fields we need to know the curvature of space and will get corrections of higher orders.

2.3. Metric tensor in spacetime

2.3.1. Generalization of Special relativity

In SR we have defined $\mathrm{d}s^2$ as

$$\mathrm{d}s^2 = \mathrm{d}x^\mu \eta_{\mu\nu} \mathrm{d}x^\nu$$

with

- $\mathrm{d}x^\mu$, $\mathrm{d}x^\nu$ are contravariant
- $\eta_{\mu\nu}$ is 2-covariant and two coordinate systems are connected by the Lorentz-transformation:

$$\frac{\partial x^\mu}{\partial \bar{x}^{\bar{\mu}}} \eta_{\mu\nu} \frac{\partial x^\nu}{\partial \bar{x}^{\bar{\nu}}} = \eta_{\bar{\mu}\bar{\nu}}$$

- x^μ are *physical, inertial coordinates*

In GR we want to define $\mathrm{d}s^2$ in a similar way with

$$\mathrm{d}s^2 = \mathrm{d}x^\mu \, g_{\mu\nu} \, \mathrm{d}x^\nu$$

Here we find, that

- The metric tensor $g_{\mu\nu}$ is 2-covariant and encodes the curvature of spacetime
- x^μ are not necessarily physical coordinates, they only need to be "valid coordinates"[4]
- The coordinate transformation needs to be a *bijection between events and coordinates* and it must be invertible
- At any point in curved spacetime we need a manifold, which has a tangent spacetime, whereas we demand, that tangent spacetimes represent the inertial frames of SR.

2.3.2. Mathematical formulation

> In any event E with coordinates x^α_E, there exists at least one transformation such that the metric transforms as
>
> $$g_{\mu\nu}(x^\alpha) \mapsto \bar{g}_{\mu\nu}(\bar{x}^\alpha) = n_{\mu\nu} + \mathcal{O}\left((\bar{x}^\alpha - \bar{x}^{*\alpha})^2\right) \tag{2.3.2.1}$$
>
> whereas $\eta_{\mu\nu}$ is the **metric of the tangent, flat spacetime**. We call this frame *Local, physical inertial frame* (LPIF)

The $\mathcal{O}\left((\bar{x}^\alpha - \bar{x}^{*\alpha})^2\right)$ is important, since this is the higher order correction to the *tangent space*, and correspondingly $\eta_{\mu\nu}$.
Hence we can neglect the $\mathcal{O}\left((\bar{x}^\alpha - \bar{x}^{*\alpha})^2\right)$ in a small region around $(\bar{x}^\alpha - \bar{x}^{*\alpha})$ and everything looks like flat spacetime and the laws of SR hold. This means, that around $\bar{x}^{*\alpha}$, the coordinates $\bar{x}^\alpha$ are **local physical and inertial coordinates**[5] of a LPIF[6] and GR automatically implies, that

> SR is an approximation of GR in a flat spacetime, which is a local tangent to curved spacetime.

[4] Coordinates, which satisfy the conditions, which we will demand in the following

[5] Inertial: The metric is the **metric of flat spacetime**, i.e. has the right signature and is *similar* to $\eta_{\mu\nu}$
Physical: The metric is **normed**, such that x^0 is the physical time of a clock at rest and $\sqrt{\sum(\mathrm{d}x^i)^2}$ is the distance of an object at rest.

[6] This frame experiences no gravity effects.

In curved spacetime we can generalize our findings from the curvature of euclidean space (see 2.1), which we expressed in formula (2.1.0.1) and get

$$\boxed{\mathrm{d}s^2 = \mathrm{d}x^\mu g_{\mu\nu} \mathrm{d}x^\nu} \tag{2.3.2.2}$$

whereas $g_{\mu\nu}$ is the *symmetric tensor of spacetime* and the generalization of the Minkowski-tensor $\eta_{\mu\nu}$ for curved spacetime. From this we see the following properties of curved spacetime:

- $\mathrm{d}s^2$ is invariant, just as in SR
- $\mathrm{d}x^\mu$ transforms like a vector:

  $$\mathrm{d}\bar{x}^{\bar{\mu}} = \frac{\partial \bar{x}^{\bar{\mu}}}{\partial x^\mu}$$

  whereas $\frac{\partial \bar{x}^{\bar{\mu}}}{\partial x^\mu}$ is a 4×4 tensor.
- Therefore $g_{\mu\nu}$ transforms like a 2-covariant tensor:

  $$\bar{g}_{\bar{\mu}\bar{\nu}} = \frac{\partial x^\mu}{\partial \bar{x}^{\bar{\mu}}} \frac{\partial x^\nu}{\partial \bar{x}^{\bar{\nu}}} g_{\mu\nu}$$

  In particular this means, that $g_{\mu\nu}$ transforms just like $\eta_{\mu\nu}$, which will be shown in ex. 4, task 2
- Since we have to be able to change back an fourth between coordinates, the transformation cannot be singular. This means, that the determinant of the jacobian cannot vanish (for a proof see [8]):

  $$\boxed{\det\left(\frac{\partial x^\mu}{\partial \bar{x}^{\bar{\mu}}}\right) \neq 0}$$

  Furthermore this inverse, which transforms back needs to fulfil

  $$g^{\mu\alpha} g_{\alpha\nu} = \delta^\mu_\nu$$
- $g_{\mu\nu}$ now plays the role of the *index-lowering tensor*

  $$v^\mu g_{\mu\nu} = v_\nu$$

  which switches a vector for a covector. In particular this means, that we can also find a transformation between $g^{\mu\nu}$ and $g_{\mu\nu}$[7]

  $$\boxed{g_{\mu\nu} g^{\nu\alpha} g^{\alpha\nu} = g^\alpha_\mu g^{\alpha\nu} = \delta^\alpha_\mu g^{\alpha\nu} = g^{\mu\nu}}$$

  (since g and δ are symmetric, we have no need to write ${g_\mu}^\nu$ or ${\delta_\mu}^\alpha$. In the future we will stick to this notation.)

[7] This is a very awkward notation which was chosen only to clarify the role of g. In the future we will just write such transformations as $g^{\alpha\beta} = g_{\mu\nu} g^{\mu\alpha} g^{\nu\beta}$ where we can just use Einstein's sum convention.

- The coordinate transformation, which links $g_{\mu\nu}$ to $\bar{g}_{\mu\nu} = \eta_{\mu\nu}$ is **not unique**! This can easily be shown if we look at three coordinate systems in a local, flat tangent space around x^*:
  - Original coordinates: $\{x^\mu\}: \quad g_{\mu\nu}(x^{*\mu}) \neq \eta_{\mu\nu}$
  - First possible transformation: $\{\bar{x}^\mu\}: \quad \bar{g}_{\mu\nu}(\bar{x}^{*\mu}) = \eta_{\mu\nu}$
  - Second possible transformation: $\{\bar{\bar{x}}^\mu\}: \quad \bar{\bar{g}}_{\mu\nu}(\bar{\bar{x}}^{*\mu}) = \eta_{\mu\nu}$

  If we then consider the transformation, which links $\bar{x}$ and $\bar{\bar{x}}$

  $$\bar{\bar{g}}_{\bar{\bar{\mu}}\bar{\bar{\nu}}} = \frac{\partial \bar{x}^{\bar{\mu}}}{\partial \bar{\bar{x}}^{\bar{\bar{\mu}}}} \frac{\partial \bar{x}^{\bar{\nu}}}{\partial \bar{\bar{x}}^{\bar{\bar{\nu}}}} \bar{g}_{\bar{\mu}\bar{\nu}}$$

  by plugging in the definition from above we see that

  $$\eta_{\bar{\bar{\mu}}\bar{\bar{\nu}}} = \left.\frac{\partial \bar{x}^{\bar{\mu}}}{\partial \bar{\bar{x}}^{\bar{\bar{\mu}}}}\right|_{x^*} \left.\frac{\partial \bar{x}^{\bar{\nu}}}{\partial \bar{\bar{x}}^{\bar{\bar{\nu}}}}\right|_{x^*} \eta_{\bar{\mu}\bar{\nu}}$$

  this means, that $\frac{\partial \bar{x}^{\bar{\mu}}}{\partial \bar{\bar{x}}^{\bar{\bar{\mu}}}}$ must be the **inverse of the Lorentz-transformation** $\Lambda^{\bar{\bar{\mu}}}_{\bar{\mu}}$, since this is (by definition) the most general transformation, which preserves $\eta_{\mu\nu}$. So (locally) all possible solutions to the equation above are linked by Lorentz-transformations.

2.4. Physical consequences of curved spacetime

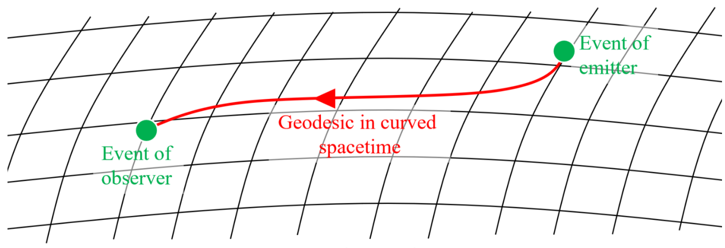
2.4.1. Change from curved to flat spacetime

We will look at an event E and two distinguishable particles, whose worldlines go through E. Then we can always find an inertial frame, such that one of the two particles is at rest (i.e. $\bar{g}_{\mu\nu} = \eta_{\mu\nu} + \mathcal{O}(\bar{x}^\alpha - \bar{x}^\alpha_E)$, $\bar{\boldsymbol{e}}_0$ is a tangent to the worldline and $\bar{\boldsymbol{e}}_i$ is orthogonal to the worldline).
Then this frame is (modulo spacial rotations) uniquely defined by a transformation discussed above. The second particle passing through E then has a different inertial frame, which is related to the frame of particle 1 by a Lorentz-Boost.

In fact we will solve many problems of GR by doing these steps:

- Start with a LPIF **around the event of the emitter** and characterize the physics with SR
- Use the laws of GR to compute the worldline in **curved spacetime**
- find a LPIF **around the observer** and again use SR to compute what the observer sees.

This means, that we can always define the properties of an object in a local physical inertial frame and then generalize this by a **change of coordinates** (similar to the MCRF in SR).



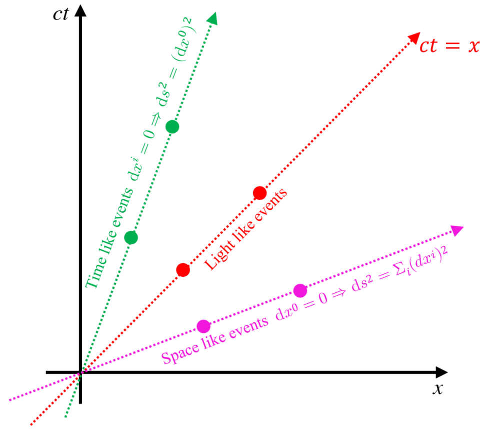
Curved spacetime

2.4.2. Length-elements

What is the meaning of $\mathrm{d}s^2$ in GR?
Like in SR it can represent an absolute measure of time or space but **not both at the same time**! We will look at the sign of $\mathrm{d}s^2$ to make this clear:

1. $\mathrm{d}s^2 = 0$:

   If the events are closeby we can use eq. (2.3.2.2) and directly see that $\mathrm{d}s^2 = (x^\mu_B - x^\mu_A) g_{\mu\nu} (x^\nu_B - x^\nu_A) = 0$, which means that the two events are on the **worldline of light**.



2. $\mathrm{d}s^2(A, B) < 0$ (for A, B close together):
   Then $\sqrt{-\mathrm{d}s^2}$ is the interval of *proper time* between A and B.
   This holds due to principle (α): In an MCRF of A and B we get

   $$\mathrm{d}s^2 = -(\mathrm{d}x^0)^2 = -(\mathrm{d}t)^2 \qquad \text{(see 1.6.1)}$$

   This can (obviously) be generalized to

   $$\mathrm{d}s^2 = \mathrm{d}x^\mu \eta_{\mu\nu} \mathrm{d}x^\nu$$

   in an arbitrary (flat) frame. We then use principle (α) and the fact, that $\mathrm{d}s^2$ is frame invariant to find the **proper time between A and B in GR**:

   $$\boxed{\underbrace{\mathrm{d}t_{=}\sqrt{-\mathrm{d}x^\mu \eta_{\mu\nu} \mathrm{d}x^\nu}}_{\text{flat space}} \quad \Rightarrow \quad \underbrace{\mathrm{d}t = \sqrt{-\mathrm{d}x^\mu g_{\mu\nu} \mathrm{d}x^\nu}}_{\text{curved space}}} \tag{2.4.2.1}$$

   In the special case of a frame which is locally at rest this becomes

   $$\mathrm{d}t = \sqrt{-\mathrm{d}x^0 g_{00} \mathrm{d}x^0}$$

   and if additionally $g_{00} = -1$ ($g_{\mu\nu} = \eta_{\mu\nu}$)

   $$\mathrm{d}t = \sqrt{(\mathrm{d}x^0)^2}$$

   which is the limit we wanted.

3. $\mathrm{d}s^2 > 0$:
   Then we call $\sqrt{\mathrm{d}s^2}$ the *proper distance* between A and B. The reasoning goes as above:
   If $\mathrm{d}s^2$ is spacelike we can find frames where A and B are simultaneous and find, that

$$\mathrm{d}s^2 = \sum_i (\mathrm{d}x^i)^2 = \underbrace{\mathrm{d}x^\mu \eta_{\mu\nu} \mathrm{d}x^\nu}_{\text{arbitrary flat frame}} =: \mathrm{d}l^2$$

   and we can call $\mathrm{d}l^2$ **proper distance (or proper length)**.
   In GR we then get:

$$\boxed{\mathrm{d}l = \sqrt{\mathrm{d}x^\mu g_{\mu\nu} \mathrm{d}x^\nu}} \tag{2.4.2.2}$$

   If $g_{\mu\nu} = \eta_{\mu\nu}$ we (obviously) get the limit from above, so our definition is consistent.

2.4.3. Volume-elements

We can follow the same logic to define a *proper volume*:
Since in GR the spacetime-volume element $\mathrm{d}^4x = \mathrm{d}^0x\mathrm{d}^1x\mathrm{d}^2x\mathrm{d}^3x$ is **coordinate dependent**. We therefore want to find a frame-invariant definition of a volume element, which we call proper volume $\mathrm{d}V$.

To find this we go to a LPIF $\{\bar{x}^\alpha\}$, and define our proper volume in this frame:

$$\mathrm{d}V = \mathrm{d}\bar{x}^{\bar{0}}\mathrm{d}\bar{x}^{\bar{1}}\mathrm{d}\bar{x}^{\bar{2}}\mathrm{d}\bar{x}^{\bar{3}}$$

As we know from the differential geometry if we change coordinates $\mathrm{d}V$ transforms like

$$\mathrm{d}V = \mathrm{d}\bar{x}^{\bar{0}}\mathrm{d}\bar{x}^{\bar{1}}\mathrm{d}\bar{x}^{\bar{2}}\mathrm{d}\bar{x}^{\bar{3}} = \det\left(\frac{\partial \bar{x}^{\bar{\mu}}}{\partial x^\mu}\right)\ \mathrm{d}x^0\mathrm{d}x^1\mathrm{d}x^2\mathrm{d}x^3$$

whereas $\frac{\partial x^\mu}{\partial \bar{x}^{\bar{\mu}}}$ is the **Jacobian** of the transformation $x^\mu \mapsto \bar{x}^{\bar{\mu}}$.

But since $\{\bar{x}^\alpha\}$ is locally inertial its metric is just $\eta_{\mu\nu}$, which means that we can express the jacobian in terms of the Minkowski-metric:

$$g_{\mu\nu} = \frac{\partial \bar{x}^{\bar{\mu}}}{\partial x^\mu}\frac{\partial \bar{x}^{\bar{\nu}}}{\partial x^\nu}\eta_{\bar{\mu}\bar{\nu}}$$

When we take the determinant we get:

$$g \equiv \det(g_{\mu\nu}) = \underbrace{\det\left(\frac{\partial \bar{x}^{\bar{\mu}}}{\partial x^\mu}\right)\det\left(\frac{\partial \bar{x}^{\bar{\nu}}}{\partial x^\nu}\right)}_{=\det\left(\frac{\partial \bar{x}^{\bar{\mu}}}{\partial x^\mu}\right)^2}\underbrace{\det(\eta_{\bar{\mu}\bar{\nu}})}_{=-1}$$

which means that the jacobian is simply given by

$$\det\left(\frac{\partial \bar{x}^{\bar{\mu}}}{\partial x^{\mu}}\right) = \text{Jacobian}(x^{\mu} \mapsto \bar{x}^{\bar{\mu}}) = \sqrt{-g}$$

which gives us the final (very important) formula

$$\boxed{\mathrm{d}V = \sqrt{-g}\ \mathrm{d}x^0\mathrm{d}x^1\mathrm{d}x^2\mathrm{d}x^3} \tag{2.4.3.1}$$

So if we want to do an integral over spacetime in GR we get

$$\int \mathrm{d}x^0\mathrm{d}x^1\mathrm{d}x^2\mathrm{d}x^3 \cdot \sqrt{-g}\ \{\text{density, e.g. } \mathcal{L}\}$$

2.4.4. Scalar product of basis vectors

If we take the scalar product of two basis vectors of coordinates $e^{\mu}_{\alpha} = \delta^{\mu}_{\alpha}$ we get:

$$\vec{e}_{\alpha} \cdot \vec{e}_{\beta} = \delta^{\mu}_{\alpha}\delta^{\nu}_{\beta}g_{\mu\nu} = g_{\alpha\beta} \tag{2.4.4.1}$$

which gives us a very handy way of computing the components of $g_{\mu\nu}$ in *any* given coordinate system:

$$\forall x^{\alpha}: \quad g_{\mu\nu} = \vec{e}_{\mu}(x^{\alpha}) \cdot \vec{e}_{\nu}(x^{\alpha}) \tag{2.4.4.2}$$

Of course if we look at locally inertial and physical coordinates $\{x^{\alpha}_{E}\}$ around $\{x^{\alpha}\}$ the unit vectors are invariant (up to $\mathcal{O}(x^{\alpha} - x^{\alpha}_{E})^2$), which we can write as

$$g_{\mu\nu}(x^{\alpha}) = \eta_{\mu\nu} + \mathcal{O}\left((x^{\alpha} - x^{\alpha}_{E})^2\right) \quad \Rightarrow \quad \vec{e}_{\mu}(x^{\alpha}) = \vec{e}_{\mu}(x^{\alpha}_{E}) + \mathcal{O}\left((x^{\alpha} - x^{\alpha}_{E})^2\right)$$

Then we have the following equivalences:

$$\begin{aligned}
&\Leftrightarrow \vec{e}_{\mu} \text{ is locally invariant up to order 2 terms} \\
&\Leftrightarrow \frac{\partial \vec{e}_{\mu}}{\partial x^{\beta}}(x^{\alpha}) = 0 + \mathcal{O}\left((x^{\alpha} - x^{\alpha}_{E})^2\right) \\
&\Leftrightarrow \frac{\partial \vec{e}_{\mu}}{\partial x^{\beta}}(x^{E}_{\alpha}) = 0
\end{aligned}$$

2.5. Covariant derivatives

We want to generalize the definition of the covariant derivative from 1.7.3 to curved manyfolds.

While one could naively assume that nothing changes except the metric, the expressions become much more complicated due to the fact, that for a vector field V^{α} we need to not only take the derivative of the V^{α} with respect to different coordinates but also **transform the basis vector(s)** accordingly. For this new derivative we use the notation

$$\frac{\partial \vec{V}}{\partial x^{\beta}} = V^{\gamma}{}_{;\beta}\, \vec{e}_{\gamma}$$

As we see, now we have a new basis $\vec{e}_{\gamma}$, which we will need to find in the following. To do this we look at different cases:

2.5.1. Scalar functions

In case of a scalar function $\phi(x^\alpha)$ we obtain the same as in SR (since scalars are independent of coordinate system):

$$\boxed{\phi_{;\alpha} = \phi_{,\alpha} = \frac{\partial \phi}{\partial x^\alpha}} \tag{2.5.1.1}$$

Therefore the total differential of course stays the same as well:

$$\mathrm{d}\phi = \phi_{,\alpha}\ \mathrm{d}x^\alpha (+\mathcal{O}(\mathrm{d}x^\alpha)^2)$$

2.5.2. Vector field

A vectorfield has a basis and thus $\vec{V}(x^\alpha) = V^\mu(x^\alpha)\vec{e}_\mu(x^\alpha)$. Then the total differential consists of the two terms:

$$\begin{aligned}
\mathrm{d}\vec{V} &= \frac{\partial \vec{V}}{\partial x^\beta}\ \mathrm{d}x^\beta = \frac{\partial}{\partial x^\beta}\left(V^\mu \vec{e}_\mu\right)\ \mathrm{d}x^\beta \\
&= \left(\vec{e}_\mu \frac{\partial V^\mu}{\partial x^\beta} + V^\mu \frac{\partial \vec{e}_\mu}{\partial x^\beta}\right)\ \mathrm{d}x^\beta
\end{aligned}$$

We see that we now have two terms, which are (at least explicitly) **not in the same basis**. But since $\frac{\partial \vec{e}_\mu}{\partial x^\beta}$ is a 4-vector in must have coordinates in the $\vec{e}_\mu$-Basis, which we will call *Christoffel symbols* $\Gamma^\gamma_{\mu\beta}$:

$$\frac{\partial \vec{e}_\mu}{\partial x^\beta} \equiv \Gamma^\gamma_{\mu\beta}\vec{e}_\gamma \tag{2.5.2.1}$$

They are **symmetric** (as shown in ex.5, task 2). With this definition we can factor $\vec{e}_\gamma$ out of our differential and get

$$\begin{aligned}
\mathrm{d}\vec{V} &= \left(\frac{\partial V^\mu}{\partial x^\beta}\vec{e}_\mu + V^\mu \Gamma^\gamma_{\mu\beta}\vec{e}_\gamma\right)\ \mathrm{d}x^\beta \\
&= \left(\frac{\partial V^\gamma}{\partial x^\beta}\vec{e}_\gamma + V^\alpha \Gamma^\gamma_{\alpha\beta}\vec{e}_\gamma\right)\ \mathrm{d}x^\beta \\
&= \underbrace{\left(\frac{\partial V^\gamma}{\partial x^\beta} + V^\alpha \Gamma^\gamma_{\alpha\beta}\right)}_{\equiv V^\gamma{}_{;\beta}}\vec{e}_\gamma\ \mathrm{d}x^\beta
\end{aligned}$$

(in the second step we renamed dummy indices to avoid confusion).
We see that $V^\gamma{}_{;\beta}$ is the γ-coordinate of the derivative of $\vec{V}$ with respect to x^β.
Our total differential then becomes

$$\mathrm{d}\vec{V} = V^\alpha{}_{;\beta}\,\vec{e}_\alpha\ \mathrm{d}x^\beta$$

Since $\mathrm{d}\vec{V}$ is an absolute geometrical quantity, $\vec{e}_\alpha$ and $\mathrm{d}x^\beta$ is a Lorentz-vector coordinate so logically (and by construction) $V^\alpha{}_{;\beta}$ is a **true** $\binom{1}{1}$ **Lorentz-tensor** which accordingly transforms like

$$\bar{V}^{\bar{\alpha}}{}_{;\bar{\beta}} = \Lambda^{\bar{\alpha}}{}_\alpha (\Lambda^{-1})^{\bar{\beta}}{}_\beta V^\alpha{}_{;\beta}$$

This is why we call it a *covariant derivative*:

$$\underbrace{V^\alpha{}_{;\beta}}_{\binom{1}{1}\text{-tensor}} = \underbrace{V^\alpha{}_{,\beta}}_{\substack{\text{Non-Lorentz} \\ \binom{1}{1}\text{-tensor!}}} + \underbrace{\Gamma^\alpha_{\gamma\beta}}_{\substack{\text{Non-Lorentz} \\ \binom{1}{2}\text{-tensor!}}} \underbrace{V^\gamma}_{\binom{1}{0}\text{-tensor}} \tag{2.5.2.2}$$

Since $\Gamma^\alpha_{\gamma\beta}$ is not a Lorentz-vector we will never raise or lower its indices and thus our notation of writing upper and lower indices on top of each other is justified.

2.5.3. Covector field

In case of a covector field we use the results from above and use the invariance of scalar products in different coordinate systems. With arbitrary $C_\alpha(x^\beta)$ and $V^\alpha(x^\beta)$ we then do

$$\begin{aligned} \phi(x^\beta) = C_\alpha(x^\beta) V^\alpha(x^\beta) \quad \Rightarrow \quad \phi_{,\gamma} &= C_{\alpha,\gamma} V^\alpha + C_\alpha V^\alpha_{,\gamma} \\ &= C_{\alpha,\gamma} V^\alpha + C_\alpha (V^\alpha_{;\gamma} - \Gamma^\alpha_{\delta\gamma} V^\delta) \end{aligned}$$

And by renaming the dummy index in the first term $C_{\alpha,\gamma} V^\alpha = C_{\delta,\gamma} V^\delta$ and factoring out V^δ we get

$$\cdots = (C_{\delta,\gamma} - \Gamma^\alpha_{\delta\gamma} C_\alpha) V^\delta + C_\alpha V^\alpha_{;\gamma}$$

Since the right term is the covector-field with the covariant derivative of the vector field we identify the left term in the brackets as the **covariant derivative of C_δ with respect to x^γ**:

$$\boxed{C_{\delta;\gamma} \equiv C_{\delta,\gamma} - \Gamma^\alpha_{\delta\gamma} C_\alpha} \tag{2.5.3.1}$$

This derivative is a $\binom{0}{2}$ tensor.

2.5.4. General fields of arbitrary rank

In the case of a general tensor of any rank, we can just use what we found above while paying attention to the right usage of vectors and covectors. e.g.

$$\begin{aligned} T^{\mu\nu}{}_{;\alpha} &= T^{\mu\gamma}{}_{,\alpha} + T^{\mu\gamma}\Gamma^\nu_{\gamma\alpha} + T^{\gamma\nu}\Gamma^\mu_{\gamma\alpha} \\ T_{\mu\nu;\alpha} &= T_{\mu\nu,\alpha} - T_{\mu\gamma}\Gamma^\gamma_{\nu\alpha} - T_{\gamma\nu}\Gamma^\gamma_{\mu\alpha} \\ T^\mu{}_{\nu;\alpha} &= T^\mu{}_{\nu,\alpha} + T^\gamma{}_\nu \Gamma^\mu_{\gamma\alpha} - T^\mu{}_\gamma \Gamma^\gamma_{\nu\alpha} \end{aligned}$$

2.5.5. Relation between the metric and Christoffel symbols

Since the metric describes the nature of spacetime itself and we can **always** find a tangent space where $g_{\mu\nu} = \eta_{\mu\nu}$ we express the Christoffel symbols in terms of the metric tensor:

$$g_{\mu\nu}(x^\alpha) = \eta_{\mu\nu} + \mathcal{O}((x^\alpha - x^\alpha_E)^2) \quad \Rightarrow \quad \Rightarrow \quad \frac{\partial \vec{e}_\mu}{\partial x^\beta}(x^\alpha_E) = 0$$

Which directly implies that

$$\Gamma^\nu_{\mu\beta}(x^\alpha_E) = 0 \Rightarrow (\dots)_{;\beta} = (\dots)_{,\beta}$$

Since this holds in a LPIF and $g_{\mu\nu,\alpha}$ is a valid $\binom{0}{3}$-tensor it "normally" transforms with Λ and thus in any frame we get

$$g_{\bar{\mu}\bar{\nu};\bar{\alpha}} = \Lambda^\mu{}_{\bar{\mu}} \Lambda^\nu{}_{\bar{\nu}} \Lambda^\alpha{}_{\bar{\alpha}} g_{\mu\nu;\alpha} = 0$$

Since we can show this in any arbitrary point since in every point we can find a LPIF for every point this holds on our whole manyfold. The same argumentation can be done for the contravariant $g^{\mu\nu}$ and we obtain the very handy result that

$$\boxed{g_{\mu\nu;\alpha} = g^{\mu\nu}{}_{;\alpha} = 0 \quad \text{in any point}} \tag{2.5.5.1}$$

From this we get the formula (will be shown in ex.5, task 2):

$$\boxed{\Gamma^\alpha_{\mu\nu} = \frac{1}{2} g^{\alpha\beta}(g_{\mu\beta,\nu} + g_{\beta\nu,\mu} - g_{\mu\nu,\beta})} \tag{2.5.5.2}$$

2.6. Geodesics

2.6.1. Equation of geodesics

We already know that the worldline of an object describes its trajectory in spacetime, whereas massless particles follow a light-like ($\mathrm{d}s^2 = 0$), and massive particles a time-like ($\mathrm{d}s^2 < 0$) curve.
This worldline is an absolute geometrical quantity, but in a given coordinate system it is a function of its coordinates x^α. Since the worldline is a line, we can represent the functional dependence of its coordinates with a single parameter, which we will call λ, so $x^\alpha = x^\alpha(\lambda)$. We can then define a (local) tangent vector to the worldline around $(x^{*\beta})$:

$$\vec{T} = \left.\frac{\mathrm{d}x^\alpha(\lambda)}{\mathrm{d}\lambda}\right|_{\lambda^*} \vec{e}_\alpha(x^{*\beta})$$

and different choices of λ give us different $\vec{T}$, which are **colinear** to each other in $x^{*\beta}$.

Special case: $\lambda = t$
If we chose λ to be the proper time, we simply get back the 4-velocity which is clear if we take the scalar product:

$$\vec{T} \cdot \vec{T} = T^\alpha g_{\alpha\beta} T^\beta = \frac{\mathrm{d}x^\alpha}{\mathrm{d}t} g_{\alpha\beta} \frac{\mathrm{d}x^\beta}{\mathrm{d}t} \overset{(1.6.2.1)}{=} \frac{\mathrm{d}s^2}{\mathrm{d}t^2} = -1$$

We can therefore identify $\vec{T}$ as the **4-velocity along the worldline**:

$$\boxed{\vec{U} = \left.\frac{\mathrm{d}x^\alpha(t)}{\mathrm{d}t}\right|_{t^*} \vec{e}_\alpha(x^\beta(t^*))} \quad \text{for } \lambda = t \tag{2.6.1.1}$$

General case
We now want to find an equation describing the worldline in the general case. This evolution of the world line is what we call *geodesic*. We define it to be a **locally flat line**, i.e.

$$\boxed{\frac{\mathrm{d}\vec{T}}{\mathrm{d}\lambda} = \lim_{\mathrm{d}\lambda \to 0} \frac{\vec{T}(\lambda + \mathrm{d}\lambda) - \vec{T}(\lambda)}{\mathrm{d}\lambda} = 0}$$

so

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\vec{e}_\alpha\right) = \frac{\mathrm{d}^2x^\alpha}{\mathrm{d}\lambda^2}\vec{e}_\alpha + \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}\vec{e}_\mu}{\mathrm{d}\lambda} \\ &= \frac{\mathrm{d}^2x^\alpha}{\mathrm{d}\lambda^2}\vec{e}_\alpha + \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\left(\frac{\partial\vec{e}_\mu}{\partial x^\nu}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\right) \\ &\overset{(2.5.2.1)}{=} \frac{\mathrm{d}^2x^\alpha}{\mathrm{d}\lambda^2}\vec{e}_\alpha + \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\left(\Gamma^\alpha_{\mu\nu}\vec{e}_\alpha\right)\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} \end{aligned}$$

And by factoring out $\vec{e}_\alpha$ and renaming the dummy indices to the ones mostly used in literature we get the very important **equation of geodesics**

$$\boxed{\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} = 0} \tag{2.6.1.2}$$

We now want to look at the implications of this equation:

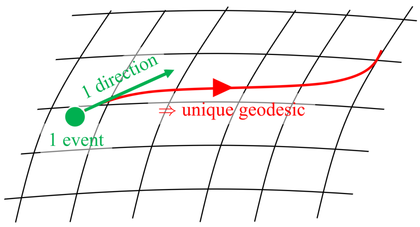
- The same geodesics can be described in any coordinate system and in each coordinate system by any choice of curvilinear[8] basis:
  - In the general case $\lambda \mapsto f(\lambda)$ gives us a **different** equation and thus also a different solution.
  - Any transformation, which is of the form

    $$\lambda \mapsto a\lambda + b$$

    leaves the shape of the equation and $x^\mu(\lambda)$ **invariant** (since this is a second order differential equation).

[8] curvilinear is just the fancy word for a curved coordinate system in a euclidian space (e.g. spherical coordinates)

- The fact that it is a second order DGL also implies, that we need *two initial conditions*, which are **1 event** and **1 direction**.
  This means that for every initial coordinate and velocity we get **one unique geodesic** (which is what we would expect from our intuition since trajectories need to be unique).



- On a geodesic, any segment $[A, B]$ gives the **shortest path** between A and B[9]. In particular this implies, that geodesics are the generalization of straight lines to curved manyfolds.

Alternative derivation using variation

While the derivation above may be the most intuitive approach if we want to derive geodesics from the tangent space known from SR, if we *assume* geodesics to be the generalization of straight lines we can also derive the equation of geodesics from the principle of variation:
Hence we want to use **the variation of** $\sqrt{\mathrm{d}s^2}$ between two points:

$$\mathrm{d}l = \sqrt{\mathrm{d}s^2} = \sqrt{g_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta}$$

The 4-length (which we can *choose* as the proper time t in an an appropriate frame but we will stick to l for an intuitive understanding) of the geodesic then becomes the integral of $\sqrt{\mathrm{d}s^2}$ from A to B. We can then parametrize $x^\alpha = x^\alpha(\lambda)$ with $\lambda \in [0, 1]$ and get:

$$l_{AB} = \int_A^B \sqrt{g_{\alpha\beta}\mathrm{d}x^\alpha\, \mathrm{d}x^\beta} = \int_0^1 \sqrt{g_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}}\, \mathrm{d}\lambda$$

We call the term in the integral the Lagrangian

$$\sqrt{g_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}} = L\left(\lambda, x^\gamma(\lambda), \frac{\mathrm{d}x^\gamma(\lambda)}{\mathrm{d}\lambda}\right) = L(\lambda, x^\gamma(\lambda), \dot{x}^\gamma(\lambda))$$

and see that

$$\frac{\mathrm{d}l}{\mathrm{d}\lambda} = L \tag{2.6.1.3}$$

If we now use variation theory $\delta l = 0$ we get the Euler-Lagrange equation, that we know from classical mechanics (derived in [1] and [2]):

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial L}{\partial \dot{x}^\gamma} = \frac{\mathrm{d}L}{\mathrm{d}x^\gamma}$$

We now carefully compute every term:

[9] This is not the entire truth:
If we look at a *compact* space, there are always two directions, which we can take to on a geodesic to go from A to B. Think for example of the sphere from 2.1, where we have two ways to move on a geodesic from A to B. Luckily all measurements imply that our universe is not compact.

1.

$$\frac{\mathrm{d}L}{\mathrm{d}x^\gamma} = \frac{1}{2L}\frac{\partial g_{\alpha\beta}}{\partial x^\gamma}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} + \frac{1}{2L}g_{\alpha\beta}\left(\underbrace{\frac{\partial^2 x^\alpha}{\partial x^\gamma \partial\lambda}}_{=0}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} + \underbrace{\frac{\partial^2 x^\beta}{\partial x^\gamma \partial\lambda}}_{=0}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\right)$$
$$\overset{(2.6.1.3)}{=} \frac{L}{2}\frac{\partial g_{\alpha\beta}}{\partial x^\gamma}\frac{\mathrm{d}x^\alpha}{\mathrm{d}l}\frac{\mathrm{d}x^\beta}{\mathrm{d}l}$$

2.

$$\frac{\partial L}{\partial \dot{x}^\gamma} = \frac{1}{2L}g_{\alpha\beta}\left(\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\delta^\beta_\gamma + \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}\delta^\alpha_\gamma\right)$$
$$= 2\frac{1}{2L}g_{\alpha\gamma}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} = \frac{1}{L}g_{\alpha\gamma}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}$$

where we used the symmetry of $g_{\alpha\beta}$ and renamed dummy indices.

3.

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\frac{\partial L}{\partial \dot{x}^\gamma}\right) = \frac{\mathrm{d}l}{\mathrm{d}\lambda}\frac{\partial}{\partial l}\left(\frac{1}{L}g_{\alpha\gamma}\frac{\partial x^\alpha}{\partial l}\frac{\mathrm{d}l}{\mathrm{d}\lambda}\right)$$
$$= L\frac{\partial}{\partial l}\left(g_{\alpha\gamma}\frac{\partial x^\alpha}{\partial l}\right)$$
$$= L\left(g_{\alpha\gamma}\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}l^2} + \frac{\partial g_{\alpha\gamma}}{\partial x^\beta}\frac{\mathrm{d}x^\beta}{\mathrm{d}l}\frac{\mathrm{d}x^\alpha}{\mathrm{d}l}\right)$$
$$= L\left(g_{\alpha\gamma}\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}l^2} + \frac{1}{2}\left(\frac{\partial g_{\alpha\gamma}}{\partial x^\beta} + \frac{\partial g_{\beta\gamma}}{\partial x^\alpha}\right)\frac{\mathrm{d}x^\beta}{\mathrm{d}l}\frac{\mathrm{d}x^\alpha}{\mathrm{d}l}\right)$$

In the last step we have again renamed dummy indices.

Altogether this gives us

$$0 = \frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial L}{\partial \dot{x}^\gamma} - \frac{\mathrm{d}L}{\mathrm{d}x^\gamma}$$
$$= g_{\alpha\gamma}\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}l^2} + \underbrace{\frac{1}{2}\left(\frac{\partial g_{\alpha\gamma}}{\partial x^\beta} + \frac{\partial g_{\beta\gamma}}{\partial x^\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x^\gamma}\right)}_{\overset{(2.5.5.2)}{=} g_{\alpha\gamma}\Gamma^\gamma_{\alpha\beta}}\frac{\mathrm{d}x^\beta}{\mathrm{d}l}\frac{\mathrm{d}x^\alpha}{\mathrm{d}l}$$

which (after renaming indices) altogether gives us the same equation of geodesics[10]

$$\frac{\mathrm{d}^2 x^\gamma}{\mathrm{d}l^2} + \Gamma^\gamma_{\alpha\beta}\frac{\mathrm{d}x^\beta}{\mathrm{d}l}\frac{\mathrm{d}x^\alpha}{\mathrm{d}l} = 0$$

[10] Note that the equation contains l instead of λ. We can generalize this equation to λ by reversing the transformation we did in the beginning.

Covectors

We have only looked at the equation of geodesics for vectors so far. Interestingly it turns out that the equation has a much simpler form if we use covectors instead. It is simply (will be shown in ex.6, task 3):

$$\boxed{\frac{\mathrm{d}^2 x_\alpha}{\mathrm{d}\lambda^2} = \frac{1}{2} g_{\mu\nu,\alpha} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}} \tag{2.6.1.4}$$

or in terms of the momentum

$$m \frac{\mathrm{d}p_\alpha}{\mathrm{d}t} = \frac{1}{2} g_{\mu\nu,\alpha} p^\mu p^\nu$$

2.6.2. Consequences of geodesics as worldlines

Massive particles

If we look at massive particles $m \neq 0$ with momentum p^μ our worldline is time-like and we can use what we derived above for $\lambda = t$. Then

$$p^\mu = mU^\mu = m \frac{\mathrm{d}x^\mu}{\mathrm{d}t}$$

which, if we plug it into equation (2.6.1.2) gives us the **equation of momentum conservation**

$$\boxed{m \frac{\mathrm{d}p^\mu}{\mathrm{d}t} + \Gamma^\mu_{\alpha\beta} p^\alpha p^\beta = 0} \tag{2.6.2.1}$$

Massless particles

Although we cannot just use proper time in the case of massless particles we can find a λ, so that $p^0 = \frac{\mathrm{d}x^0}{\mathrm{d}\lambda}$ is the energy of the particle. The reasoning then goes analogous to above and we obtain

$$\boxed{\frac{\mathrm{d}p^\mu}{\mathrm{d}\lambda} + \Gamma^\mu_{\alpha\beta} p^\alpha p^\beta = 0} \tag{2.6.2.2}$$

Physical interpretation

- **Local tangent spacetime:**
  From the formulae above we can extract a deeper physical understanding of the concept of inertial frames. In fact we have shown that

  $$g_{\mu\nu} = \eta_{\mu\nu} + \mathcal{O}((\delta x^\alpha)^2) \quad \Leftrightarrow \quad \text{being in an inertial/free falling frame.}$$

Indeed, if we use (2.6.2.1) and (2.3.2.1) we get that

$$\begin{aligned} g_{\mu\nu} &= \eta_{\mu\nu} + \mathcal{O}((\delta x^\alpha)^2) \\ &\Rightarrow \Gamma^\mu_{\alpha\beta} = 0 \\ &\Rightarrow \frac{\mathrm{d}p^\mu}{\mathrm{d}t} = 0 \\ &\Rightarrow p^\mu \quad \text{conserved (locally)} \end{aligned}$$

Which means, that **objects move in straight lines**. This implies that locally

– All geodesics are straight up to $\mathcal{O}(\Delta t^3)$:

$$x^\alpha = A^\alpha + B^\alpha(t - t^*) + \mathcal{O}((t - t^*)^3)$$

⇒ All velocities are constant: $U^\alpha = B^\alpha + \mathcal{O}((t - t^*)^2)$

⇒ All accelerations vanish: $a^\alpha = 0 + \mathcal{O}((t - t^*))$

which is just what we would expect from an inertial frame, which sees no effects from gravity.

- **General spacetime:**
  In general, the simple formulae we derived above do not hold and we get a nonlinear curvature of spacetime. This implies, that also velocities and accelerations vary. This is how gravity manifests itself! The deeper knowledge, that we can pull from this is that unlike all other forces, which we can quantify in the stress energy tensor (see 1.7.2) but which is not a part of spacetime itself[11], the curvature of space itself generates the effect, which we observe from gravity.

- **Presence of a force:**
  In the presence of a force (other than gravity) the equation of momentum conservation must be modified to include the 4-force and has the form

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}t} + \frac{1}{m}\Gamma^\mu_{\alpha\beta}p^\alpha p^\beta = F^\mu \neq 0$$

Summary

We now know how to compute the path of a test particle in a given metric, which corresponds to a given gravitational field. What we **do not** know is how to compute the dynamics of massive particles (which change the metric while they move) in spacetime. The connection between mass and curvature will be given by the *Einstein equation*, which we will derive later but to do so we first need a more mathematical chapter on the curvature of spacetime.

[11] We especially see this in the fact, that for **any** frame we can find a LPIF, so that $g_{\mu\nu} = \eta_{\mu\nu}$ but it is not possible to find a specific structure of $T^{\mu\nu}$ in an arbitrary frame.

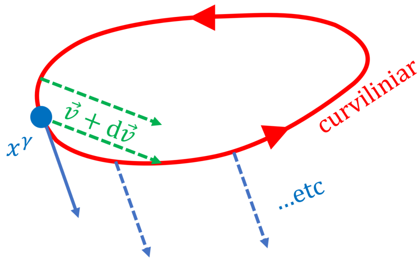
2.7. Curvature of spacetime

We know from the previous chapters:

- $g_{\mu\nu}$ encodes all properties of the spacetime but **does not** explicitly imply whether the manyfold is curved or not[12].
- The curvature of spacetime is encoded in the **second derivative** $f(x) \to f''(x)$ of the coordinates (which we would intuitively assume from analysis)
- The second derivative of $g_{\mu\nu}$ is **not a single number**, which means that several numbers characterize the curvature of space
- $\Gamma^{\gamma}_{\mu\nu}$ only contains first derivatives and thus **does not** show curvature explicitly.

2.7.1. The Riemann-Tensor

We remind ourselves of the methods we presented in 2.1 to test curvature of space. We need to chose one method and find a mathematical formulation for this test. We choose **parallel transport**.
We can formulate the concept of parallel transport mathematically as



$$\lim_{\mathrm{d}\lambda\to 0} \frac{\vec{V}(\lambda+\mathrm{d}\lambda) - \vec{V}(\lambda)}{\mathrm{d}\lambda} = \vec{0} \quad \text{everywhere.}$$

This means that

$$\mathrm{d}\vec{V} \neq \vec{0} \quad \Leftrightarrow \quad \text{spacetime is curved}$$

We now want to compute the components δV^{α} of $\mathrm{d}\vec{V}$. we simply rewrite

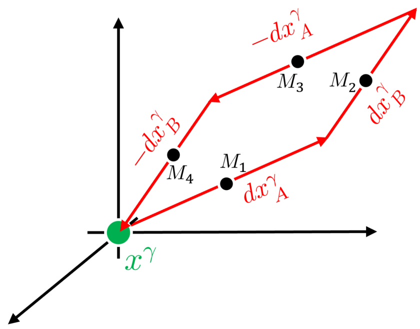
$$\vec{V} = V^{\alpha}\vec{e}_{\alpha}(x^{\gamma}) \qquad \vec{V} + \mathrm{d}\vec{V} = (V^{\alpha} + \delta V^{\alpha})\vec{e}_{\alpha}(x^{\gamma})$$

Since we look at small changes, $\mathrm{d}\vec{V}$ should be **linear** in $\vec{V}$, so

$$\delta V^{\alpha} = (\dots)^{\alpha}{}_{\beta} V^{\beta}$$

[12]Think for example of curvilinear ("*soft*" coordinates.)

Furthermore we see that $(\dots)^\alpha{}_\beta$ should depend on the path. Take for example a small parallelogram, which is generated by two small vectors $\mathrm{d}\vec{x}_A$, $\mathrm{d}\vec{x}_B$ (see picture). For small $\mathrm{d}x^\mu_A$, $\mathrm{d}x^\nu_B$ we get only linear terms and finally obtain



$$\delta V^\alpha = (\dots)^\alpha{}_{\beta\mu\nu} \mathrm{d}x^\nu_A \mathrm{d}x^\mu_B V^\beta$$

This tensor has the following properties:

- It is a valid Lorentz-tensor of rank $\binom{1}{3}$ (the argument is like always: Dx^μ_A, $\mathrm{d}x^\nu_B$, V^β) are all valid Lorentz-vectors so $(\dots)^\alpha{}_{\beta\mu\nu}$ needs to be a valid Lorentz-tensor
- If **fully describes the curvature in a given event**

We call this tensor the *Riemann Tensor* $R^\alpha{}_{\beta\mu\nu}(x^\gamma)$.

We can relate $R^\alpha{}_{\beta\mu\nu}$ to $\Gamma^\alpha_{\mu\nu}$ and thus to the metric. We will do a brief derivation of the connection by using thinking about the parallelogram from above:

- We know that on each of the four sides $\frac{\mathrm{d}\vec{V}}{\mathrm{d}\lambda} = 0$ (by construction of the parallel transport), which means that

$$\begin{aligned}
&\Rightarrow \frac{\mathrm{d}\vec{V}}{\mathrm{d}\lambda} = V^\alpha{}_{,\mu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \vec{e}_\alpha = \vec{0} \\
&\Rightarrow V^\alpha{}_{;\mu} = 0 \quad \Rightarrow \quad V^\alpha{}_{,\mu} = -\Gamma^\alpha_{\mu\beta} V^\beta
\end{aligned}$$

- We then write down the coordinate difference between the initial and final vector in the point x^γ after going around the parallelogram:

$$\delta V^\alpha = \underbrace{V^\alpha{}_{,\nu}\, \mathrm{d}x^\nu_A}_{\text{in } M_1} + \underbrace{V^\alpha{}_{,\mu}\, \mathrm{d}x^\mu_B}_{\text{in } M_2} - \underbrace{V^\alpha{}_{,\nu}\, \mathrm{d}x^\nu_A}_{\text{in } M_3} - \underbrace{V^\alpha{}_{,\mu}\, \mathrm{d}x^\mu_B}_{\text{in } M_4}$$

  (We evaluate the derivatives in the middle of the sides of the parallelogram). We then rewrite the derivatives in terms of the Christoffel symbols and get:

$$\delta V^\alpha = -\underbrace{\Gamma^\alpha_{\nu\beta} \mathrm{d}x^\nu_A V^\beta}_{\text{in } M_1} - \underbrace{\Gamma^\alpha_{\mu\beta} \mathrm{d}x^\mu_B V^\beta}_{\text{in } M_2} + \underbrace{\Gamma^\alpha_{\nu\beta} \mathrm{d}x^\nu_A V^\beta}_{\text{in } M_3} + \underbrace{\Gamma^\alpha_{\mu\beta} \mathrm{d}x^\mu_B V^\beta}_{\text{in } M_4}$$

- We rearrange the terms, so that M_1 and M_3 are together as well as M_2 and M_4:

$$\dots = \left[\underbrace{\Gamma^\alpha_{\nu\beta} V^\beta}_{\text{in } M_3} - \underbrace{\Gamma^\alpha_{\nu\beta} V^\beta}_{\text{in } M_1} \right] \mathrm{d}x^\nu_A + \left[\underbrace{\Gamma^\alpha_{\mu\beta} V^\beta}_{\text{in } M_4} - \underbrace{\Gamma^\alpha_{\mu\beta} V^\beta}_{\text{in } M_2} \right] \mathrm{d}x^\mu_B$$

and since $\mathrm{d}x_B^\mu$ describes the coordinates of $\overrightarrow{M_1M_3}$ and $-\mathrm{d}x_A^\nu$ describes the coordinates of $\overrightarrow{M_2M_4}$ (becomes clear from the way we go through the parallelogram) we can rewrite this to

$$
\begin{aligned}
&= \left[\frac{\partial}{\partial x^\mu}\left(\Gamma^\alpha_{\nu\beta}V^\beta\right)\right]\mathrm{d}x_B^\mu \mathrm{d}x_A^\nu - \left[\frac{\partial}{\partial x^\nu}\left(\Gamma^\alpha_{\mu\beta}V^\beta\right)\right]\mathrm{d}x_A^\nu \mathrm{d}x_B^\mu \\
&= \left[\Gamma^\alpha_{\nu\beta,\mu}V^\beta + \Gamma^\alpha_{\nu\beta}V^\beta_{\ \ ,\mu}\right]\mathrm{d}x_B^\mu \mathrm{d}x_A^\nu + \left[\Gamma^\alpha_{\mu\beta,\nu}V^\beta - \Gamma^\alpha_{\mu\beta}V^\beta_{\ \ ,\nu}\right]\mathrm{d}x_B^\mu \mathrm{d}x_A^\nu \\
&= \left[\Gamma^\alpha_{\nu\beta,\mu}V^\beta - \Gamma^\alpha_{\nu\beta}\Gamma^\beta_{\mu\sigma}V^\sigma\right]\mathrm{d}x_B^\mu \mathrm{d}x_A^\nu - \left[\Gamma^\alpha_{\mu\beta,\nu}V^\beta - \Gamma^\alpha_{\mu\beta}\Gamma^\beta_{\nu\sigma}V^\sigma\right]\mathrm{d}x_B^\mu \mathrm{d}x_A^\nu \\
&= \underbrace{\left[\Gamma^\alpha_{\beta\nu,\mu} - \Gamma^\alpha_{\nu\sigma}\Gamma^\sigma_{\mu\beta} - \Gamma^\alpha_{\mu\beta,\nu} + \Gamma^\alpha_{\mu\sigma}\Gamma^\sigma_{\nu\beta}\right]}_{\equiv R^\alpha{}_{\beta\mu\nu}}\mathrm{d}x_B^\mu \mathrm{d}x_A^\nu V^\beta
\end{aligned}
$$

Which means, that the explicit representation of the Riemann tensor is given by

$$\boxed{R^\alpha{}_{\beta\mu\nu} = \Gamma^\alpha_{\beta\nu,\mu} - \Gamma^\alpha_{\beta\mu,\nu} + \Gamma^\alpha_{\sigma\mu}\Gamma^\sigma_{\nu\beta} - \Gamma^\alpha_{\sigma\nu}\Gamma^\sigma_{\mu\beta}} \tag{2.7.1.1}$$

Since the Christoffel symbols depend on the first derivative of $g_{\mu\nu}$, $R^\alpha{}_{\beta\mu\nu}$ has to depend on the **second derivative of the metric**. The relation to $g_{\mu\nu}$ though is not trivial.

The only particular case, in which we can easily see an explicit form of the Riemann tensor is for the LPIF, in which

$$g_{\mu\nu} = \eta_{\mu\nu} + \mathcal{O}(\delta x^2) \quad \Rightarrow \quad \Gamma^\alpha_{\mu\nu} = 0$$

$$\Rightarrow R^\alpha{}_{\beta\mu\nu} = \frac{1}{2}g^{\alpha\sigma}\left(g_{\sigma\nu,\beta\mu} - g_{\sigma\mu,\beta\nu} + g_{\beta\mu,\sigma\nu} - g_{\beta\nu,\sigma\mu}\right)$$

Note: This result might at first seem counter intuitive, since the LPIF is defined as **locally flat** tangent metric. It is important though to remind ourselves, that the flatness exclusively applies to the tangent space but not to space itself. From our findings we can extract the following properties of $R^\alpha{}_{\beta\mu\nu}$:

- From our geometrical construction directly follows, that $R^\alpha{}_{\beta\mu\nu} = 0$ in flat spacetime (even in complicated coordinates)
- Since the $\binom{1}{3}$ Riemann-tensor is a proper Lorentz tensor we can lower one index (or raise/lower any number of indices for that matter) and can define the $\binom{0}{4}$ Riemann tensor:

$$R^\alpha{}_{\beta\mu\nu} = g^{\alpha\delta}R_{\delta\beta\mu\nu}$$

- Since R is a 4×4-tensor it has $4^4 = 256$ components! This is a lot but luckily we do not need to calculate this many components, since only 20 of them are independent. This follows from symmetry considerations

$$R_{\alpha\beta\,\mu\nu} = -R_{\underbrace{\beta\alpha}\,\mu\nu} = -R_{\alpha\beta\,\underbrace{\nu\mu}} = +R_{\underbrace{\underline{\mu\nu}\,\underline{\alpha\beta}}}$$

and from the *first Bianchi identity* (for proof see A.1):

$$R^\alpha{}_{[\beta\mu\nu]} := R^\alpha{}_{\beta\mu\nu} + R^\alpha{}_{\mu\nu\beta} + R^\alpha{}_{\nu\beta\mu} = 0$$

For a full proof see A.2

2.8. The Einstein equation

With the Riemann tensor we now have a mean to measure the curvature of spacetime and know how to relate this curvature to the metric. We furthermore know, that matter generates curvature. But what is the relation between matter and curvature?

$$\text{curvature} \quad \overset{?}{\Leftrightarrow} \quad \text{matter}$$

We know:

- Our theory can only involve valid Lorentz scalars/vectors/tensors.
- In the limit of small masses (small energies) we have to get back our Newtonian theory with

$$m\boldsymbol{a} = m\nabla\phi \quad \Rightarrow \quad \underbrace{\vec{a} = \nabla\phi}_{\text{geodesic}}$$

  Thus we know that the Einstein equation must have the *Poisson equation* as newtonian limit:

$$\Delta\phi = 4\pi G\rho_m$$

  with the *gravitational potential* ϕ and the *mass density* ρ_m.

We will derive the Einstein equation by argumenting. We do this in steps:

2.8.1. Requirements to a covariant relation between curvature and matter

We know from SR, that the *energy density* ρ, which is the T^{00}-component of the stress-energy tensor (see 1.7.2) becomes the mass density in the limit of small velocities[13]:

$$\rho = T^{00} \overset{v \ll c}{\to} \rho_m$$

Furthermore we know, that the curvature must depend on some form of second order differentiation, so that we get the shape of the Poisson equation[14]

$$(\text{curvature})^{\mu\nu} = (\ldots)T^{\mu\nu} = (g^{\mu\nu})_{,??} = (\ldots)T^{\mu\nu}$$

Now what can be the left hand side? We need the curvature to have **two indices**, it needs to depend on the curvature of spacetime and it needs to contain a second order derivative.
We look at the variables, which we have introduced in the last chapters and see, that we have

[13] $E^2 = p^2 + m^2 \overset{p \to 0}{\to} m^2$

[14] We need the form of the Poisson equation, since we require, that although the Einstein equation is *local*, the curvature needs to depend on some matter outside the area we are looking at.

- $g_{\mu\nu}$ Depends on the geometry and is a true Lorentz-tensor but **does not contain derivatives**
- $\Gamma^{\gamma}_{\mu\nu}$ Depends on the geometry and contains derivatives but is **not a true Lorentz-tensor**
- ${R^{\alpha}}_{\beta\mu\nu}$ Depends on the geometry, is a true Lorentz-tensor and contains derivatives.

We thus need to build our equation **with the Riemann-tensor**.

The problem we immediately see, is that the Riemann-tensor **has the wrong rank**, so we need to contract it with something. We use the metric tensor and try contracting different indices:

- Try $g^{\alpha\beta}R_{\alpha\beta\mu\nu} = {R^{\beta}}_{\beta\mu\nu}$. We see, that this has the right rank, but **vanishes** due to the antisymmetry under the exchange of indices.
- Try $g^{\alpha\mu}R_{\alpha\beta\mu\nu}$. This tensor **does not vanish** and thus is a valid choice for our tensor. We call it the *Ricci-tensor*:

  $$\boxed{g^{\alpha\mu}R_{\alpha\beta\mu\nu} = {R^{\mu}}_{\beta\mu\nu} = \mathcal{R}_{\beta\nu}} \tag{2.8.1.1}$$

  Of course we can also raise its indices, to make a $\binom{2}{0}$-tensor:

  $$\mathcal{R}^{\alpha\beta} = g^{\alpha\alpha'}g^{\beta\beta'}\mathcal{R}_{\alpha'\beta'}$$

  Furthermore the tensor is **symmetric**, which we can see from

  $$\mathcal{R}_{\nu\beta} = g^{\alpha\mu}R_{\alpha\nu\mu\beta} = g^{\mu\alpha}R_{\mu\beta\alpha\nu} = \mathcal{R}_{\beta\nu}$$

- Try $g^{\alpha\nu}R_{\alpha\beta\mu\nu} = -g^{\alpha\nu}R_{\alpha\beta\nu\mu} = -\mathcal{R}_{\beta\mu}$. So this gives us no extra information.

So altogether we see, that

> All contractions of Riemann-tensor leading to a $\binom{0}{2}$-tensor are either 0 or ±Ricci.

Can we go even further and contract the last two remaining indices of the Ricci Tensor:

$$\boxed{g^{\alpha\beta}\mathcal{R}_{\alpha\beta} = {\mathcal{R}^{\beta}}_{\beta} = g^{\alpha\beta}g^{\mu\nu}R_{\mu\alpha\nu\beta} \equiv \mathcal{R}} \tag{2.8.1.2}$$

which we call the *Ricci*-scalar[15].

We put our knowledge together and find for a possible equation:

$$\boxed{(\dots)\mathcal{R}^{\mu\nu} + (\dots)\mathcal{R}g^{\mu\nu} + (\dots)g^{\mu\nu} = (\dots)T^{\mu\nu}}$$

Is this equation unique and the only solution? **No**! It is just the simplest equation fulfilling the requirements. We could add terms like $f(\mathcal{R})g^{\mu\nu}$ and $g(\mathcal{R})\mathcal{R}^{\mu\nu}$ but these only make our problem more complicated and so far the Einstein equation describes all observations perfectly. We can therefore proceed.

[15] The physical interpretation of the Ricci scalar is a bit more subtle, one could say, that it is a kind of "average curvature", but this is very sloppy and not really well defined.

2.8.2. Equations of motion

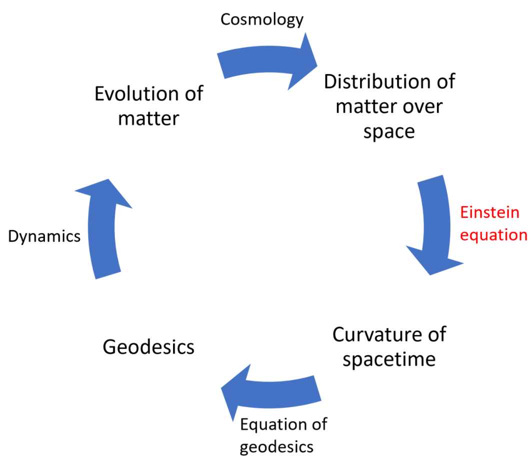
We know from SR, that

$$T^{\mu\nu}{}_{;\nu} = 0$$

which is true in **any** frame and thus also applies in GR. This means, that we can impose the constraint, that

$$[(\dots)\mathcal{R}^{\mu\nu} + (\dots)\mathcal{R}g^{\mu\nu} + (\dots)g^{\mu\nu}]_{;\nu} = [(\dots)T^{\mu\nu}]_{;\nu} = 0$$

Is this a real constraint or does this follow directly from the construction of our theory?
In fact we see, that this is **not a real constraint**, since we already built our theory to be self-consistent and incorporate the evolution of matter (and thus energy) itself.



This means, that the Einstein equation (by construction) needs to take the form

$$\boxed{\left(\mathcal{R}^{\alpha\beta} + \mathcal{R}\mu g^{\alpha\beta} + \lambda g^{\alpha\beta}\right)_{;\beta} = 0} \tag{2.8.2.1}$$

We look at the appearing terms one by one and look if they vanish:

- $(\lambda g^{\alpha\beta})_{;\beta} = 0$, since we showed, that $g^{\alpha\beta}{}_{,\beta} = 0$ in any frame (see eq. (2.5.5.1)). So this term can stay as it is.
- $(\mu\mathcal{R}g^{\alpha\beta})_{;\beta} = \mu\mathcal{R}_{;\beta}g^{\alpha\beta} + \underbrace{\mu\mathcal{R}g^{\alpha\beta}{}_{;\beta}}_{=0} = \mu\mathcal{R}_{;\beta}g^{\alpha\beta} \neq 0$ in general. So either this term cannot exist or the first term needs to cancel it out.
- $R^{\alpha\beta}{}_{;\beta}$ is the hardest to calculate. The problem can be largely simplified though. We can show, that (for proof see A.1)

  $$R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\lambda\mu;\nu} + R_{\alpha\beta\nu\lambda;\mu} = 0$$

  which is called the *second Bianchi-identity*.
  We can contract this tensor twice with the metric $g^{\alpha\mu}$ and get a $\binom{0}{3}$-tensor, which is called the *contracted Bianchi identity* (also proven in A.1).
  we contract two more indices with $g^{\beta\lambda}$ and rename dummy indices and finally get the *twice contracted Bianchi identity* (proven in A.1):

  $$2 \cdot \mathcal{R}^{\alpha}{}_{\nu;\alpha} - \mathcal{R}_{;\nu} = 2 \cdot g^{\beta\epsilon}\mathcal{R}_{\beta\nu;\epsilon} - \mathcal{R}_{;\nu}$$

We can apply the metric $g^{\nu\mu}$ to both terms in order to raise all indices and rename dummy indices, which gives us

$$2(g^{\nu\mu}g^{\beta\epsilon}\mathcal{R}_{\beta\nu;\epsilon}) - g^{\mu\nu}\mathcal{R}_{;\nu} = 0$$

which we can simplify by swapping the dummy indices $\epsilon \leftrightarrow \nu$ which finally yields[16]

$$\left(\mathcal{R}^{\mu\nu} - \frac{1}{2}g^{\mu\nu}\mathcal{R}\right)_{;\nu} = 0$$

We call the term in brackets the *Einstein tensor* $G^{\mu\nu}$:

$$\boxed{G^{\mu\nu} \equiv \mathcal{R}^{\mu\nu} - \frac{1}{2}g^{\mu\nu}\mathcal{R}} \tag{2.8.2.2}$$

This means, that we **have to choose** $\mu = -\frac{1}{2}$ **in the Einstein equation** (2.8.2.1) and the equation will be automatically fulfilled. λ remains unconstrained. We call it the *Cosmological constant* Λ and view it as a physical constant[17].

The **Einstein equation** then becomes:

$$\boxed{G^{\alpha\beta} + \Lambda g^{\alpha\beta} = \kappa T^{\alpha\beta}} \tag{2.8.2.3}$$

where we have introduced the constant κ, which we will get by requiring the correct Newtonian limit. This will be done in the next part.

2.8.3. Weak field limit

Weak field limit (w.f. limit) is what we have called *Newtonian limit in the past.* In GR we can identify this weak field limit by perturbing the flat Minkowksi metric with a small parameter $h_{\mu\nu}$ with $|h_{\mu\nu}|1$. We only look at the first order and get:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} + (\text{higher orders})$$

Then the key to working out the correct equations characterizing the weak field metric, we need to take a close look at the form of $h_{\mu\nu}$ and derive its properties:

[16] The notation taking the covariant derivative with respect to μ is more common but of course this is the same since the Ricci tensor and the metric are both symmetric.

[17] A constant is the only possibility in this case, since a dependence on coordinates would give us additional derivatives. The importance of Λ will become clear later in the course

- $h_{\mu\nu}$ is **not a $\binom{0}{2}$-Lorentz tensor**. We see that from the fact, that $h_{\mu\nu}$ is a proper metric minus the Minkowski metric, so if we look at a coordinate transformation $x^\alpha \mapsto \tilde{x}^{\tilde{\alpha}}$:

$$\tilde{g}_{\tilde{\alpha}\tilde{\beta}} = \Lambda^\alpha{}_{\tilde{\alpha}}\Lambda^\beta{}_{\tilde{\beta}} g_{\alpha\beta} = \Lambda^\alpha{}_{\tilde{\alpha}}\Lambda^\beta{}_{\tilde{\beta}}\eta_{\alpha\beta} + \Lambda^\alpha{}_{\tilde{\alpha}}\Lambda^\beta{}_{\tilde{\beta}} h_{\alpha\beta}$$

so we see that we have two possibilities:

  - If $\Lambda^\alpha{}_{\tilde{\alpha}}$ is the Lorentz transformation of SR, then $\Lambda^\alpha{}_{\tilde{\alpha}}\Lambda^\beta{}_{\tilde{\beta}}\eta_{\alpha\beta} = \tilde{\eta}_{\tilde{\alpha}\tilde{\beta}}$ which means, that we can write

$$\tilde{g}_{\tilde{\alpha}\tilde{\beta}} = \Lambda^\alpha{}_{\tilde{\alpha}}\Lambda^\beta{}_{\tilde{\beta}} g_{\alpha\beta} = \underbrace{\Lambda^\alpha{}_{\tilde{\alpha}}\Lambda^\beta{}_{\tilde{\beta}}\eta_{\alpha\beta}}_{=\eta_{\tilde{\alpha}\tilde{\beta}}} + \Lambda^\alpha{}_{\tilde{\alpha}}\Lambda^\beta{}_{\tilde{\beta}} h_{\alpha\beta}$$

    which means, that

$$\tilde{h}_{\tilde{\alpha}\tilde{\beta}} = \tilde{g}_{\tilde{\alpha}\tilde{\beta}} = \Lambda^\alpha{}_{\tilde{\alpha}}\Lambda^\beta{}_{\tilde{\beta}} h_{\alpha\beta}$$

    Wee see, that in this case $h_{\alpha\beta}$ in fact transforms like a proper $\binom{0}{2}$ Lorentz vector.

  - $\Lambda^\alpha{}_{\tilde{\alpha}} = \frac{\partial x^\alpha}{\partial \tilde{x}^{\tilde{\alpha}}}$ is a general coordinate transformation. Then

$$\Lambda^\alpha{}_{\tilde{\alpha}}\Lambda^\beta{}_{\tilde{\beta}}\eta_{\alpha\beta} \neq \eta_{\tilde{\alpha}\tilde{\beta}}$$

    so we generally get:

$$\tilde{h}_{\tilde{\alpha}\tilde{\beta}} = \Lambda^\alpha{}_{\tilde{\alpha}}\Lambda^\beta{}_{\tilde{\beta}}\eta_{\alpha\beta} + \Lambda^\alpha{}_{\tilde{\alpha}}\Lambda^\beta{}_{\tilde{\beta}} h_{\alpha\beta} - \eta_{\tilde{\alpha}\tilde{\beta}}$$

    which is **not** a LT of a $\binom{0}{2}$-tensor

- **Inverse metric** of the w.f. limit:
We can just simply raise the index of $h_{\alpha\beta}$ with the Minkowski metric (in $\mathcal{O}(h)$). So we can construct $g^{\mu\nu}$, so that $g^{\mu\alpha}g_{\alpha\nu} = \delta^\mu_\nu$ with (proof in ex. 7, task 2)

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} \quad \text{with} \quad h^{\mu\nu} = \eta^{\mu\alpha}\eta^{\nu\beta}h_{\alpha\beta}$$

so $\eta^{\alpha\beta}$ is just simply the index raising tensor of $h_{\alpha\beta}$.
This is actually not a surprising result, since $\eta^{\alpha\beta}$ must be the index raising tensor of **any** perturbation, since if we take $g^{\alpha\beta}$ as index raising tensor we would immediately get:

$$g^{\alpha\beta}h_{\beta\nu} = \eta^{\alpha\beta}h_{\beta\nu} + \mathcal{O}(h^2)$$

- **Trace of** h:
We see, that we can generally find the trace of a tensor of rank 2:

$$T^\mu{}_\nu = g^{\mu\beta}T_{\beta\nu} = g_{\nu\beta}T^{\mu\beta}$$

Which means, that we can contract μ with ν and we get

$$T \equiv T^{\mu}{}_{\mu} = T_{\mu}{}^{\mu} = g^{\mu\nu}T_{\mu\nu} = g_{\mu\nu}T^{\mu\nu}$$

This means, that the trace of h is given by

$$\begin{aligned} h \equiv h^{\mu}{}_{\mu} &= g^{\mu\nu}h_{\mu\nu} = g_{\mu\nu}h^{\mu\nu} \\ &\approx \eta^{\mu\nu}h_{\mu\nu} = \eta_{\mu\nu}h^{\mu\nu} \quad \text{at order 1 in } h \end{aligned}$$

We have to note though, that since $h^{\alpha\beta}$ is not a true $\binom{2}{0}$ Lorentz-tensor, h is **not a Lorentz-scalar!**

- **Trace reverse tensor of h:**
  We *define* the trace reverse tensor as

$$\boxed{\bar{h}_{\alpha\beta} = h_{\alpha\beta} - \frac{1}{2}\eta_{\alpha\beta}h} \tag{2.8.3.1}$$

  This is also **not a $\binom{0}{2}$ Lorentz tensor**.
  Why do we call it trace reverse tensor? We see, that $\bar{h} = -h$:

$$\bar{h} = \eta^{\alpha\beta}\left(h_{\alpha\beta} - \frac{1}{2}\eta_{\alpha\beta}h\right) = h - \frac{1}{2}\underbrace{\left(\eta^{\alpha\beta}\eta_{\alpha\beta}\right)}_{=\mathrm{Tr}(\mathbb{1}_4)=4}h = -h \tag{2.8.3.2}$$

  This looks familiar to the Einstein tensor from eq. (2.8.2.2). This is **not a coincidence**. In fact $G_{\alpha\beta}$ is the trace-reverse tensor of the Ricci tensor.

- **Riemann-tensor** in the weak field limit
  We know from eq. (2.7.1.1), that we can represent the Riemann tensor in terms of the Christoffel symbols:

$$R_{\alpha\beta\mu\nu} = g_{\alpha\gamma}\left(\Gamma^{\gamma}_{\beta\nu,\mu} - \Gamma^{\gamma}_{\beta\mu,\nu} + \Gamma^{\gamma}_{\sigma\mu}\Gamma^{\sigma}_{\nu\beta} - \Gamma^{\gamma}_{\sigma\nu}\Gamma^{\sigma}_{\mu\beta}\right)$$

  But we know, that for $h_{\alpha\beta} = 0$ the Christoffel symbol needs to vanish, so the lowest non vanishing order of $\Gamma^{\alpha}_{\mu\nu}$ needs to be linear in h:

$$\Gamma^{\alpha}_{\mu\nu} = 0 + \mathcal{O}(h) + \dots$$

  This means, that we can neglect all terms, which are quadratic in Γ and by using eq. (2.5.5.2) we finally get

$$\begin{aligned} R_{\alpha\beta\mu\nu} &= \frac{1}{2}\left(g_{\alpha\nu,\beta\mu} + g_{\beta\mu,\alpha\nu} - g_{\alpha\mu,\beta\nu} - g_{\beta\nu,\alpha\mu}\right) \qquad (\eta_{\alpha\beta,\mu\nu} = 0) \\ &= \frac{1}{2}\left(h_{\alpha\nu,\beta\mu} + h_{\beta\mu,\alpha\nu} - h_{\alpha\mu,\beta\nu} - h_{\beta\nu,\alpha\mu}\right) \end{aligned}$$

- **Einstein tensor** in weak field limit
  It turns out, that if we want to express $G_{\alpha\beta}$ in this metric it is easier to express it in terms of $\bar{h}_{\alpha\beta}$ instead of $h_{\alpha\beta}$ (shown in ex. 7, task 1, *optional task*):

$$G_{\alpha\beta} = -\frac{1}{2}\left(\eta^{\mu\nu}\bar{h}_{\alpha\beta,\mu\nu} + \eta_{\alpha\beta}\eta^{\mu\gamma}\eta^{\nu\delta}h_{\mu\nu,\gamma\delta} - \eta^{\mu\nu}\bar{h}_{\mu,\beta\nu} - \eta^{\mu\nu}\bar{h}_{\beta\mu,\alpha\nu}\right)$$

  which we will shorten down using the more compact notation:

$$(\dots)^{,\mu} \equiv g^{\mu\nu}(\dots)_{,\nu}$$

  Then we get

$$\boxed{G_{\alpha\beta} = -\frac{1}{2}\left(\bar{h}_{\alpha\beta,\mu}{}^{,\mu} + \eta_{\alpha\beta}\bar{h}_{\mu\nu}{}^{,\mu\nu} - \bar{h}_{\alpha\mu,\beta}{}^{,\mu} - \bar{h}_{\beta\mu,\alpha}{}^{,\mu}\right)} \tag{2.8.3.3}$$

  Note, that we have kept the order of the derivatives here, in fact this is **not necessary** since we are looking at a small perturbation and locally the index raising tensor is just the Minkowski metric, which has a vanishing partial derivative:

$$(\dots)_{,\mu}{}^{,\mu} = \eta^{\mu\nu}(\dots)_{,\mu\nu} = \eta^{\mu\nu}(\dots)_{,\nu\mu} = (\eta^{\mu\nu}(\dots)_{,\nu})_{,\mu} = (\dots)^{,\mu}{}_{,\mu}$$

- Remarks on the **differential operator** $(\dots)_{,\mu}{}^{,\mu}$
  This operator in general has a non trivial solution, since

$$(\dots)_{,\mu}{}^{,\mu} = g^{\mu\nu}(\dots)_{,\mu\nu} = g^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu}(\dots)$$

  But for a small perturbation $g^{\mu\nu} = \eta^{\mu\nu} + \mathcal{O}(h^2)$ since all non-vanishing derivatives **need to be at least** $\mathcal{O}(h)$. With this we get

$$\begin{aligned}(\dots)_{,\mu}{}^{,\mu} &= \eta^{\mu\nu}\frac{\partial}{\partial x^\mu}\frac{\partial}{\partial x^\nu}(\dots) + \mathcal{O}(h^2)\\ &= (-\frac{\partial^2}{\partial t^2} + \Delta)(\dots) \equiv \Box(\dots)\end{aligned}$$

  with the *Laplacian* Δ and the *d'Alembertian* $\Box$ known from vector analysis.
  Since $\Box$ is the usual differential operator in free wave equations, in flat spacetime this means, that we get **plane waves**:

$$\Box f = 0 \quad \Rightarrow \quad f = A \cdot e^{\mathrm{i}k_\alpha x^\alpha}$$

  which gives us a condition for k_α:

$$\begin{aligned}(k_0^2 - k_1^2 - k_2^2 - k_3^2)f &= 0\\ \Rightarrow k_0^2 &= \sum_i k_i^2 = \omega^2\end{aligned}$$

2.8.4. Gauge transformations

We have seen from deriving the properties of our weak field limit, that the Einstein tensor generally does not take a very simple form.
Although this perturbed manifold has an absolute geometrical existence we can of course describe it in **any system of coordinates**, so we can try to find a coordinate system which fulfils $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ and which makes our calculations as easy as possible. We call this a *gauge transformation*:

> A gauge transformation is an infinitesimal change of coordinates, such that a **small** $h_{\mu\nu}$ transforms into a small $\tilde{h}_{\mu\nu}$

Mathematically we can express this with an infinitesimal vector $\xi^{\tilde{\alpha}}(x^{\beta})$:

$$x^{\alpha} \mapsto \underbrace{\tilde{x}^{\tilde{\alpha}}}_{\text{new coordinates}} = \underbrace{\delta^{\tilde{\alpha}}_{\alpha} x^{\alpha}}_{\text{old coordinates}} + \underbrace{\xi^{\tilde{\alpha}}(x^{\alpha})}_{\text{difference "new}\to\text{old" in each point}}$$

So we can write the **transformation matrix** and its inverse as

$$\frac{\partial \tilde{x}^{\tilde{\alpha}}}{\partial x^{\alpha}} = \delta^{\tilde{\alpha}}_{\alpha} + \xi^{\tilde{\alpha}}{}_{,\alpha} \tag{2.8.4.1}$$

$$\frac{\partial x^{\alpha}}{\partial \tilde{x}^{\tilde{\alpha}}} = \delta^{\alpha}_{\tilde{\alpha}} - \xi^{\alpha}{}_{,\tilde{\alpha}} \tag{2.8.4.2}$$

because in that case we get back the identity in linear order if we do a change back and fourth:

$$\begin{aligned}
\frac{\partial \tilde{x}^{\tilde{\alpha}}}{\partial x^{\alpha}} \frac{\partial x^{\beta}}{\partial \tilde{x}^{\tilde{\alpha}}} &= (\delta^{\tilde{\alpha}}_{\alpha} + \xi^{\alpha}{}_{,\tilde{\alpha}})(\delta^{\beta}_{\tilde{\alpha}} - \xi^{\beta}{}_{,\tilde{\alpha}}) \\
&= \delta^{\tilde{\alpha}}_{\alpha}\delta^{\beta}_{\tilde{\alpha}} + \cancel{\delta^{\beta}_{\tilde{\alpha}}\xi^{\tilde{\alpha}}_{,\alpha}} - \cancel{\delta^{\beta}_{\tilde{\alpha}}\xi^{\tilde{\alpha}}_{,\alpha}} + \mathcal{O}(\xi^2) \\
&= \delta^{\beta}_{\alpha} + \mathcal{O}(\xi^2)
\end{aligned}$$

This means, that we can write the transformation $g_{\alpha\beta} \to \tilde{g}_{\tilde{\alpha}\tilde{\beta}}$ as

$$\begin{aligned}
\tilde{g}_{\tilde{\alpha}\tilde{\beta}} &= (\delta^{\alpha}_{\tilde{\alpha}} - \xi^{\alpha}{}_{,\tilde{\alpha}})(\delta^{\beta}_{\tilde{\beta}} - \xi^{\beta}{}_{,\tilde{\beta}})(\eta_{\alpha\beta} + h_{\alpha\beta}) \\
&= \eta_{\tilde{\alpha}\tilde{\beta}} - \eta_{\alpha\tilde{\beta}}\xi^{\alpha}{}_{,\tilde{\alpha}} - \eta_{\tilde{\alpha}\beta}\xi^{\beta}{}_{,\tilde{\beta}} + h_{\tilde{\alpha}\tilde{\beta}} + \mathcal{O}(\xi^2, \xi \cdot h)
\end{aligned}$$

So by inserting $\tilde{h}_{\tilde{\alpha}\tilde{\beta}} = g_{\tilde{\alpha}\tilde{\beta}} - \eta_{\tilde{\alpha}\tilde{\beta}}$ we see that

$$\boxed{\tilde{h}_{\tilde{\alpha}\tilde{\beta}} = h_{\tilde{\alpha}\tilde{\beta}} - \xi_{\tilde{\beta},\tilde{\alpha}} - \xi_{\tilde{\alpha},\tilde{\beta}}} \tag{2.8.4.3}$$

Hence we have found the **gauge transformation of a metric perturbation**.
We can use this knowledge to **eliminate degrees of freedom**:

$$\begin{cases} h_{\mu\nu} \to 10 \text{ d.o.f.} \\ \xi^{\alpha} \to 4 \text{ d.o.f.} \end{cases}$$

So we can eliminate four degrees of freedom with a gauge transformation.

Lorenz gauge

Imagine an arbitrary coordinate system with a small $h_{\mu\nu}$. We then compute $\bar{h}^{\alpha\nu}{}_{,\nu}$ and **impose**[18] a gauge transformation using a ξ^α, which obeys (similar to the Lorenz gauge in electrodynamics)

$$\Box\xi^\alpha = \xi^{\alpha,\mu}{}_{,\mu} = \bar{h}^{\alpha\nu}{}_{,\nu}$$

where the ${}^{,\alpha}$ is a *contravariant derivative*. We then get a new $\tilde{h}_{\mu\nu}, \tilde{g}_{\mu\nu}$ where (this will be shown in ex. 7, task 2)

$$\bar{\tilde{\bar{h}}}^{\mu\nu}{}_{,\nu} = 0 = \bar{\tilde{\bar{h}}}_{\mu\nu}{}^{,\nu}$$

which also implies, that

$$\tilde{G}_{\alpha\beta} = -\frac{1}{2}\bar{\tilde{\bar{h}}}_{\alpha\beta}{}^{,\mu}{}_{,\mu}$$

We can then use the d'Alembertian to write

$$\boxed{\tilde{G}_{\alpha\beta} = -\frac{1}{2}\cdot\Box\bar{\tilde{\bar{h}}}_{\alpha\beta}} \quad \text{in the Lorenz gauge} \tag{2.8.4.4}$$

Note: Since from now on we will work in the Lorentz-gauge we omit the   in order to simplify notations.

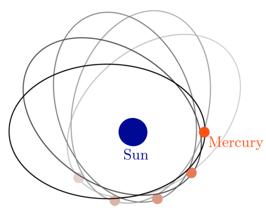
With this the Einstein equation in the weak field limit with the Lorenz gauge simplifies to:

$$-\frac{1}{2}\Box\bar{h}^{\alpha\beta} + \Lambda g^{\alpha\beta} = \kappa T^{\alpha\beta}$$

2.8.5. The Newtonian limit of General Relativity

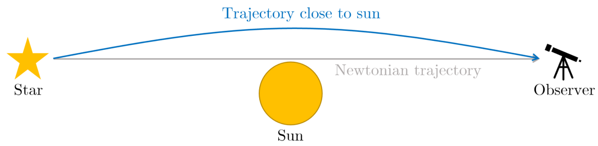
While for many problems in nature, the Newtonian theory of gravity gives a first order description already in the 19th century there were some deviations in measurements due to GR corrections. We will look at a few examples:

- The measured **shift in the perihelion angle of Mercury** was observed to deviate 43″/century from the predictions from Newtonian gravity even after considering the perturbations from other planets. This deviation could be explained by Einstein.

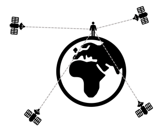


[18] This does not follow trivially from the equation but was rather obtained from thinking a lot about the structure of the Einstein equation.

- The **gravitational lensing of the sun** leads to deviations between the apparent and true position of stars when they pass close to the sun.



- **GPS** uses high precision measurements of time differences in received light signals. Without correcting effects of GR these measurements give completely wrong results.



How do we connect our weak field limit and the Newtonian limit mathematically? We know, that in the Newtonian limit $|h_{\mu\nu}|$ is small (small curvature) and all v are small (slow objects) $v \ll 1$. Then we can look at the components of the Einstein equation

$$-\frac{1}{2}\Box\bar{h}^{\alpha\beta} = \kappa T^{\alpha\beta} \tag{2.8.5.1}$$

We see that

- $$\begin{aligned} T^{00} &= \text{rest energy density} + \underbrace{\text{kinetic energy density}}_{\text{small}} \\ &\approx \text{rest energy density} \\ &\overset{1.7.2}{=} m \cdot n \quad \text{for identical particles} \end{aligned}$$

  So T^{00} is **not necessarily small**. In fact it is $\mathcal{O}(v^0)$ in an expansion in v. In contrast

- $T^{0i} = T^{i0} = \mathcal{O}(T^{00} \cdot v) = \mathcal{O}(v)$

- $T^{ij} = T^{ji} = \mathcal{O}(T^{00} \cdot v^2) = \mathcal{O}(v^2)$

So in v we have a hierarchy:

$$|T^{00}| \gg |T^{0i}| \gg |T^{ij}|$$

Note, that this is just the expansion in terms of v, but we still need to do the expansion in terms of $h_{\mu\nu}$. We only keep the leading order terms in both expansions and since the Einstein equation has the form (2.8.5.1) we need to have the same hierarchy in $\bar{h}^{\mu\nu}$[19]:

$$\begin{aligned} &\Box\bar{h}^{00} \gg \Box\bar{h}^{i0} \gg \Box\bar{h}^{ij} \\ \Rightarrow\ &\bar{h}^{00} \gg \bar{h}^{i0} \gg \bar{h}^{ij} \end{aligned}$$

[19] The step from the first relation to the second is in fact a bit subtle. Although we could in principle add a constant term α, a first order term $\alpha_\mu x^\mu$ or $\alpha((x^1)^2 - (x^2)^2)$ and thus change the order, these kind of functions are not allowed physically since $\bar{h}$ is a **small perturbation**, hence terms which include x^μ are not allowed since they would diverge at $x^\mu \to \infty$ and a constant α would be so small, that it would not matter.

Then we can explicitly write out

$$\bar{h} = \eta_{\mu\nu}\bar{h}^{\mu\nu} = -\bar{h}^{00} + \bar{h}^{11} + \bar{h}^{22} + \bar{h}^{33} = -\bar{h}^{00} + \mathcal{O}(\bar{h}^{00}v^2)$$

but also

$$\bar{h}^{\mu\nu} = h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h \overset{(2.8.3.2)}{=} h^{\mu\nu} + \frac{1}{2}\eta^{\mu\nu}\bar{h}$$

so altogether we get

$$h^{\mu\nu} = \bar{h}^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}\bar{h} \tag{2.8.5.2}$$

This enables us to explicitly calculate the **coefficients of $h^{\mu\nu}$ in leading order**:

- $h^{00} = \bar{h}^{00} - \frac{1}{2}(-1)\bar{h} = -\bar{h} + \frac{1}{2}\bar{h} + \mathcal{O}(\bar{h}v^2) \approx -\frac{1}{2}\bar{h}$
- $h^{0i} = \bar{h}^{0i} = \mathcal{O}(\bar{h}v) \ll \bar{h}^{00}$
- $h^{i\neq j} = \bar{h}^{i\neq j} = \mathcal{O}(\bar{h}v^2) \ll \bar{h}^{00}$
- $h^{ii} = \bar{h}^{ii} - \frac{1}{2}\bar{h} = \mathcal{O}(\bar{h}v^2) - \frac{1}{2}\bar{h} \approx -\frac{1}{2}\bar{h}$

which means, that the only remaining terms are the diagonal terms and the metric $g^{\alpha\beta} = \eta^{\alpha\beta} + h^{\alpha\beta}$ in first order becomes:

$$g^{\alpha\beta} = \eta^{\alpha\beta} - \begin{pmatrix} -\frac{1}{2}\bar{h} & 0 & 0 & 0 \\ 0 & -\frac{1}{2}\bar{h} & 0 & 0 \\ 0 & 0 & -\frac{1}{2}\bar{h} & 0 \\ 0 & 0 & 0 & -\frac{1}{2}\bar{h} \end{pmatrix}$$

and analogously the metric with the index lowered gives us

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}$$

with

$$h_{\alpha\beta} = \eta_{\alpha\mu}\eta_{\beta\nu}h^{\mu\nu} = \begin{pmatrix} -\frac{1}{2}\bar{h} & 0 & 0 & 0 \\ 0 & -\frac{1}{2}\bar{h} & 0 & 0 \\ 0 & 0 & -\frac{1}{2}\bar{h} & 0 \\ 0 & 0 & 0 & -\frac{1}{2}\bar{h} \end{pmatrix}$$

We can shorten down this notation in the line elements

$$\boxed{\mathrm{d}s^2 = \left(-1 - \frac{1}{2}\bar{h}\right)(\mathrm{d}x^0)^2 + \left(1 - \frac{1}{2}\bar{h}\right)\left((\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2\right)} \tag{2.8.5.3}$$

Which reduces the Einstein equation further to only include the 00-component and we get

$$-\frac{1}{2} \cdot \Box\bar{h}^{00} = \kappa T^{00} \quad \Leftrightarrow \quad \frac{1}{2} \cdot \Box\bar{h} = \kappa\rho$$

where ρ is the *rest energy density*.
We proceed by doing an approximation of $\Box$:

$$\Box \equiv - \underbrace{\frac{\partial^2}{\partial t^2}}_{\frac{\partial}{\partial t} = \mathcal{O}(v \frac{\partial}{\partial x})} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \approx \Delta + \mathcal{O}(v^2 \Delta)$$

so we finally get for the Einstein equation:

$$\boxed{\Delta \bar{h} = 2\kappa\rho}$$

How do we get the connection between $\bar{h}$ and the usual gravitational potential Φ known from Newtonian dynamics, so that we can get a value for κ? We already know, that

$$\frac{\bar{h}}{2\kappa} = \frac{\Phi}{4\pi G} \tag{2.8.5.4}$$

since $\Delta\phi = 4\pi G\rho$. So in order to draw the connection we need to prove, that $-\nabla\bar{h}$ gives the **acceleration** (up to a normalization factor).
We look at the *spacial part of the* equation of geodesics (2.6.2.1)

$$m\frac{\mathrm{d}p^i}{\mathrm{d}t} + \Gamma^i_{\alpha\beta}p^\alpha p^\beta = 0$$

and see, that

$$p^\alpha = (m\gamma, mv^1\gamma, mv^2\gamma, mv^3\gamma) = (m, 0, 0, 0) + \mathcal{O}(v)$$

which means, that in leading order we obtain

$$p^i{}_{,0} = -\Gamma^i_{00}p^0p^0 = -m^2\Gamma^i_{00}$$

but with the metric from (2.8.5.3) we get (shown in ex. 7, task 2):

$$\Gamma^i_{00} = \delta^{ij}\left(\frac{\bar{h}}{4}\right)_{,j}$$

which means, that we finally have found the relation between $\bar{h}$ and the acceleration:

$$m \cdot p^i{}_{,0} = m^2\frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} = m^2 \underbrace{a^i}_{\substack{\text{usual}\\\text{acceleration}}} = -\frac{m^2}{4}\delta^{ij}\bar{h}_{,j} = -\frac{1}{4}\nabla\bar{h}$$

This means, that

$$\boxed{\frac{\bar{h}}{4} = \Phi \quad \overset{(2.8.5.4)}{\Rightarrow} \quad \kappa = 8\pi G} \tag{2.8.5.5}$$

Remark on the relation between x^0 and t:
Why do we write $\frac{\mathrm{d}p^\mu}{\mathrm{d}x^0} = \frac{\mathrm{d}p^\mu}{\mathrm{d}t}$ in the geodesic equation, i.e. use the x^0-component as proper time?
Normally we would have to take the proper time, since in the geodesic equation eq. (2.6.1.2) the derivatives are taken with respect to proper time.
But if use the definition of the proper time

$$\mathrm{d}t = \sqrt{-\mathrm{d}x^\mu g_{\mu\nu}\mathrm{d}x^\nu}$$

and use the weak field-metric we see, that

$$\mathrm{d}t^2 = -\mathrm{d}x^\mu(\eta_{\mu\nu} + \underbrace{h_{\mu\nu}}_{\overset{(2.8.5.2)}{=} -\frac{1}{2}\bar{h}\delta_{\mu\nu}})\mathrm{d}x^\nu$$

But for slowly moving particles, we saw that $\mathrm{d}x = \mathcal{O}(v\ \mathrm{d}x^0) \ll \mathrm{d}x^0$, so

$$\mathrm{d}t^2 \approx -(\mathrm{d}x^0)^2\left(\eta_{00} - \frac{1}{2}\bar{h}\right) = (\mathrm{d}x^0)^2\left(1 + \frac{1}{2}\bar{h}\right)$$

$$\Rightarrow \mathrm{d}t \approx \left(1 + \frac{1}{2}\bar{h}\right)^{1/2}\mathrm{d}x^0 \approx \left(1 + \frac{1}{4}\bar{h}\right)\mathrm{d}x^0 = (1+\phi)\mathrm{d}x^0$$

This means, that the (small) difference between $\mathrm{d}t$ and $\mathrm{d}x^0$ is a **second order correction**, which we call *post-Newtonian* corrections.
In particular we find

$$\frac{\mathrm{d}p^i}{\mathrm{d}x^0} = \frac{\partial p^i}{\partial t}\frac{\mathrm{d}t}{\mathrm{d}x^0} = -\frac{m}{4}\delta^{ij}\left(1 + \frac{\bar{h}}{4}\right)\bar{h}_{nj} = \mathcal{O}(\bar{h}^2)$$

Remark on units:
If we would not use natural units, we would find

$$G^{\alpha\beta} = \Lambda g^{\alpha\beta} = \frac{8\pi G}{c^4}T^{\alpha\beta}$$

and since in SI $\rho = (mc^2)\cdot n$ we get $\rho = \rho_m c^2$ this altogether gives

$$\Delta\phi = \frac{4\pi G}{c^2}\rho_m \quad \Leftrightarrow \quad 4\pi G\rho_m$$

which indeed gives us the usual gravitational potential from Newtonian dynamics if we define $\phi c^2 \equiv \phi_{\mathrm{grav}}$. Indeed this gives the right units for the gradient:

$$\underbrace{\boldsymbol{a}}_{\mathrm{ms}^{-2}} = -\underbrace{\nabla}_{\mathrm{m}^{-1}}\underbrace{\phi_{\mathrm{grav}}}_{\mathrm{m}^2\mathrm{s}^{-2}}$$

Another choice is to extend natural units to

$$c = G = 1$$

by a proper choice of mass. Then we can use "meters", "meters of time", "meters of mass" and in these units the Einstein equation further simplifies to

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

2.9. The Schwarzschild metric

In this section we will try to explicitly compute a solution of the Einstein equation in the case of **spherical symmetry**.
We already know, that spherical coordinates imply

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2(\theta)\mathrm{d}\varphi^2)$$

where we have just used the metric of a 2-sphere and added the $x^0 =: t$ coordinate.
We can then find the components of the metric by using the properties of spherical symmetry:

- Each element belonging to a given 2-sphere with fixed radius should have identical properties:

  $$\mathrm{d}s^2 = g_{00}(\mathrm{d}x^0)^2 + 2g_{0i}\mathrm{d}x^0\mathrm{d}x^i + g_{11}(\mathrm{d}x^1)^2 + 2g_{1i}\mathrm{d}x^1\mathrm{d}x^i + f(r,t)((\mathrm{d}x^2)^2 + \sin^2(\theta)(\mathrm{d}x^3)^2)$$

  with $i = 2, 3$. With a more convenient naming by using t, r, θ, φ as coordinate-names this becomes

  $$\mathrm{d}s^2 = g_{tt}\mathrm{d}t^2 + 2g_{ti}\mathrm{d}t\mathrm{d}x^i + g_{rr}\mathrm{d}r^2 + 2g_{ri}\mathrm{d}r\mathrm{d}x^i + f(r,t)(\mathrm{d}\theta^2 + \sin^2(\theta)\mathrm{d}\varphi^2)$$

- Since $\vec{e}_r$ and $\vec{e}_t$ need to be orthogonal to the 2-sphere (otherwise spherical symmetry would be violated) and $\vec{e}_\theta$ and $\vec{e}_\varphi$ are tangents to the 2-sphere we see, that

  $$\vec{e}_r \cdot \vec{e}_\varphi = 0 \overset{(2.4.4.1)}{=} g_{r\varphi}$$

  and analogously

  $$g_{r\theta} = g_{t\theta} = g_{t\varphi} = 0$$

  which altogether leaves us with

  $$\mathrm{d}s^2 = g_{tt}(t,r)\mathrm{d}t^2 + 2g_{tr}(t,r)\mathrm{d}t\mathrm{d}r + g_{rr}(t,r)\mathrm{d}r^2 + f(t,r)(\mathrm{d}\theta^2 + \sin^2(\theta)\mathrm{d}\varphi^2)$$

- We know, that the components of the metric cannot depend on θ or φ since this would also violate spherical symmetry.

- Finally we see, that if we define a transformation

  $$r \to r' = (f(t,r))^{1/2}$$

  and use our symmetry arguments (no component can depend on θ, φ) the metric in this system simplifies to[20]

  $$\mathrm{d}s^2 = g_{tt}\mathrm{d}t^2 + g_{r't}\mathrm{d}t\mathrm{d}r + g_{r'r'}\mathrm{d}r'^2 + r'^2\underbrace{(\mathrm{d}\theta^2 + \sin^2(\theta)\mathrm{d}\varphi^2)}_{=:\mathrm{d}\Omega^2}$$

[20] We can simply use this, since in our transformation $g_{\mu'\nu'} = \frac{\partial x^\alpha}{\partial x'^\mu} g_{\alpha\beta} \frac{\partial x^\beta}{\partial x'^\nu}$ all derivatives except $\frac{\partial r}{\partial r'} = 1$ vanish.

- To find the remaining components g_{tt}, g_{rt}, g_{rr} we solve $G^{\alpha\beta} = T^{\alpha\beta}$
  We proceed analogously as in electrodynamics and look at different cases:
  - **Interior solution**: $T^{\mu\nu}$ takes a general ($T \in SO(3)$) form, so $g_{\mu\nu}$ depends on $\rho(r)$ and $\mathfrak{p}(r)$
  - **Exterior solution in vacuum**: The stress-energy-tensor vanishes (locally) so $G^{\alpha\beta} = 0$ and with the assumption of asymptotic flatness $R_{\alpha\beta\mu\nu} \overset{r\to\infty}{\to} 0$ we get to the Schwarzschild solution.
- **Schwarzschild solution**: With static boundary conditions we get for the metric (**Herleitung?**):

$$\boxed{\mathrm{d}s^2 = -\left(1 - \frac{2M}{r}\right)\mathrm{d}t^2 + \left(1 - \frac{2M}{r}\right)^{-1}\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega^2}$$

In fact Birkhoff proved, that this solution is **always true**, even if we do not assume a static distribution of mass. The only requirements, which remain are **spherical symmetry, vacuum** outside the star and **asymptotic flatness** of spacetime. Note, that spherical symmetry is not only required for the mass distribution but also for the movement of particles, i.e. a rotating mass distribution only has axial symmetry due to angular momentum.

Comparison between Schwarzschild and Newtonian metric

From the first term of the metric we can immediately identify $\phi = -\frac{M}{r}$, since the metric of the Newtonian metric read

$$\mathrm{d}s^2 = -(1 + 2\phi)\mathrm{d}t^2 + (1 - 2\phi)\underbrace{\mathrm{d}\boldsymbol{x}^2}_{\substack{\text{spacial} \\ \text{3-vector}}}$$

but if we do this in the Newtonian metric we get:

$$\mathrm{d}s^2 = -\left(1 - \frac{2M}{r}\right)\mathrm{d}t^2 + \left(1 + \frac{2M}{r}\right)(\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2)$$

We see two differences with respect to the Schwarzschild metric:

- $g_{rr} = (1 + \frac{2M}{r})$ instead of $(1 + \frac{2M}{r})^{-1}$
- $(1 + \frac{2M}{r})$ is also a factor of $r^2\mathrm{d}\Omega^2$

So we see, that the two metrics **are in fact different**. But how do we recover the Newtonian metric from the Schwarzschild metric?

We see, that we need to take the limit $|\phi| = |\frac{M}{r}| \ll 1$, i.e. M small or r large. In that case the Schwarzschild metric becomes

$$\mathrm{d}s^2 = -\left(1 - \frac{2M}{r}\right)\mathrm{d}t^2 + \left(1 + \frac{2M}{r} + \mathcal{O}\left(\frac{M^2}{r^2}\right)\right)\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2$$

We then change variables $r \to \bar{r} = r - M$ ($\frac{\partial x^\mu}{\partial x^{\bar{\mu}}} P = \delta^\mu_{\bar{\mu}}$):

$$\begin{aligned} \mathrm{d}s^2 &\approx 1\left(1 - \frac{2M}{\bar{r}+M}\right)\mathrm{d}t^2 + \left(1 + \frac{2M}{\bar{r}+M}\right)\mathrm{d}\bar{r}^2 + (\bar{r}+M)^2\mathrm{d}\Omega^2 \\ &= -\left(1 - \frac{2M}{\bar{r}}\left(1+\frac{M}{\bar{r}}\right)^{-1}\right)\mathrm{d}t^2 + \left(1 + \frac{2M}{\bar{r}}\left(1+\frac{M}{\bar{r}}\right)^{-1}\right)\mathrm{d}\bar{r}^2 + \bar{r}^2\left(1+\frac{M}{\bar{r}}\right)^2\mathrm{d}\Omega^2 \\ &\approx -\left(1 - \frac{2M}{\bar{r}}\right)\mathrm{d}t^2 + \left(1 + \frac{2M}{\bar{r}}\right)(\mathrm{d}\bar{r}^2 + \bar{r}^2\mathrm{d}\Omega^2) \end{aligned}$$

so we see, that the Schwarzschild solution returns the Newtonian metric with spherical symmetry for $|\frac{M}{r}| \to 0$.
In particular the Kepler orbital velocity is asymptotically the same at large r:

$$v(r) \to \frac{M}{r}$$

Furthermore we see, that the Schwarzschild metric has a **singularity at** $r = 2M$, since

$$g_{rr} \xrightarrow{r \to 2M} \infty$$

For more informations on this case see 2.10. For "normal" objects though, this singularity is never reached, e.g. for the sun $R_\odot \sim 10^6\,\mathrm{km} \gg M_\odot \sim 1\,\mathrm{km}$

2.9.1. Conserved quantities

To derive the equations of motion for particles in the Schwarzschild solution we will heavily rely on conserved quantities, so we will take some time to take a closer look at the implications of spherical symmetry and a static metric on conservation laws.

General remarks on the connection between the metric and conservation laws

We remind ourselves of the geodesic equation from 2.6.1:

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} = 0$$

Furthermore we have seen, that for a massive particle we can chose $\lambda = t$ so $\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} = \frac{p^\alpha}{m}$ and

$$m\frac{\mathrm{d}}{\mathrm{d}t}p^\mu = -\Gamma^\mu_{\alpha\beta}p^\alpha p^\beta$$

and analogously we can chose λ for a photon so that $E_{\mathrm{obs}} = h\nu_{\mathrm{obs}} = -\vec{p}\cdot\vec{U}_{\mathrm{obs}} = -\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}g_{\alpha\beta}U^\beta_{\mathrm{obs}}$, which is exactly the definition for the energy from 1.6.5 which is equivalent to the condition, that an observer at rest in a LPIF measures $h\nu = \frac{\mathrm{d}x^0}{\mathrm{d}\lambda}$. So we get

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}p^\mu = -\Gamma^\mu_{\alpha\beta}p^\alpha p^\beta$$

In both cases we can lower the index (as shown in ex. 6, task 3) and get:

$$\left.\begin{array}{ll}\text{massive:} & m\dfrac{\mathrm{d}p_\beta}{\mathrm{d}t} \\ \text{photon:} & \dfrac{\mathrm{d}p_\beta}{\mathrm{d}\lambda}\end{array}\right\} = \frac{1}{2} g_{\mu\nu,\beta} p^\mu p^\nu$$

so we immediately see, that **symmetries of the metric imply conserved quantities**:

$$\boxed{\text{if } g_{\mu\nu,\beta} = 0 \quad \forall\, \mu,\nu \quad \Rightarrow \quad p_\beta = \text{const}} \text{ along } \textbf{all} \text{ geodesics}$$

Conserved quantities in the Schwarzschild metric

In case of the Schwarzschild metric we have three conserved quantities, which follow from the conditions *static* (1) and *spherical symmetry* (2):

1. **Spherical symmetry:**
   To see the consequences of spherical symmetry it is easier to rotate the frame (without loss of generality), so that the geodesic of a given particle is along the *equatorial plane*, i.e. $\theta = \frac{\pi}{2}$. Then $x^\mu(\lambda) = (t(\lambda), r(\lambda), 0, \varphi(\lambda))$. Then the two conservation laws follow from

   $$\begin{cases} \frac{\mathrm{d}\theta}{\mathrm{d}\lambda} = 0 & \Rightarrow \quad p_\theta = 0 \\ \forall\, \mu,\nu \quad g_{\mu\nu,\varphi} = 0 & \Rightarrow \quad p_\varphi = \text{const} \quad \text{along all geodesics} \end{cases}$$

   the second condition is equivalent to the **conservation of angular momentum**, which we know from Newtonian mechanics.

2. **Static metric:**

   $$\forall\, \mu,\nu \quad g_{\mu\nu,0} = 0 \quad \Rightarrow \quad p_0 = \text{const} \quad \text{along all geodesics}$$

General remarks on the connection between E, p^0 and p_0:

In general $E \neq -p^0 \neq -p_0$, as these are in principle very different values:

- $E = E_{\text{obs}} = -\vec{p} \cdot \vec{U}_{\text{obs}}$ depends both on the observer and the frame where the measurement is performed
- p^0 = 0-component of the "true" 4-momentum p^μ in a given frame. It increases (decreases) when the particle is accelerated (decelerated).
- p_0 is "something" which encodes information on the particle *and* the gravitational environment, since $p_0 = g_0\mu p^\mu$. From above we know that this term is constant if the metric is static.

We can also use a simple example to show, that they are in fact different:
Let $\vec{U}_{\text{obs}} = (v^0, 0, 0, 0)$ with $\vec{U}_{\text{obs}} \cdot \vec{U}_{\text{obs}} = (v^0)^2 g_{00} = -1$, so $v^0 = (-g_{00})^{-1/2}$

$$E_{\text{obs}} = -\vec{p} \cdot \vec{U}_{\text{obs}} = -p^\mu g_{\mu 0}(-g_0 0)^{-1/2} = p_0(-g_{00})^{-1/2} = -(p^0 g_{00} + p^i g_{i0})(-g_{00})^{-1/2}$$

In particular we see, that the three quantities are only equal in a LPIF, since there $g_{00} = 1$ and $g_{0i} = 0$.

2.9.2. Evolution of one massive particle

Again we set $\theta = \frac{\pi}{2}$, so that the particle only moves in the equatorial plane. We can then use our knowledge about conserved quantities to write down p_0, p_θ, p_ϕ:

$$\begin{cases} p_0 & = \text{const} = -\tilde{K}m \\ p_\theta & = \text{const} = 0 \\ p_\varphi & = \text{const} = \tilde{L}m \end{cases}$$

with the constants $\tilde{K}$ and $\tilde{L}$, which we will need to determine from the equation of geodesics. We can then write the contravariant 4-momentum by using the Schwarzschild metric and get:

$$\begin{cases} p^0 = g^{0\nu}p_\nu & = g^{00}p_0 = \left(1 - \frac{2M}{r}\right)^{-1} m\tilde{K} \\ p^r & = m\frac{\mathrm{d}r}{\mathrm{d}\tau} \quad \text{as usual} \\ p^\theta & = 0 \\ p^\varphi & = g^{\varphi\nu}p_\nu = g^{\varphi\varphi}p_\varphi = \frac{m\tilde{L}}{r^2\sin^2\theta} = \frac{m\tilde{L}}{r^2} \end{cases}$$

Normally we would now use the geodesic equation to solve the equation of motion for each component but since we already have three integrals solved $(0, \theta, \varphi)$, so we can directly find the equation for r by just using:

$$\vec{p} \cdot \vec{p} = -m^2 \Leftrightarrow p^0 g_{00} p^0 + p^r g_{rr} p^r + p^\varphi g_{\varphi\varphi} g^\varphi = -m^2$$

so we directly obtain the differential equation:

$$-m^2\left(1 - \frac{2M}{r}\right)^{-1}\tilde{K} + m^2\left(1 - \frac{2M}{r}\right)^{-1}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 + \frac{m^2\tilde{L}}{r^2} = -m^2$$

so after simplifying the remaining terms are

$$\boxed{\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 = \tilde{K}^2 - \left(1 - \frac{2M}{r}\right)\left(1 + \frac{\tilde{L}^2}{r^2}\right)} \tag{2.9.2.1}$$

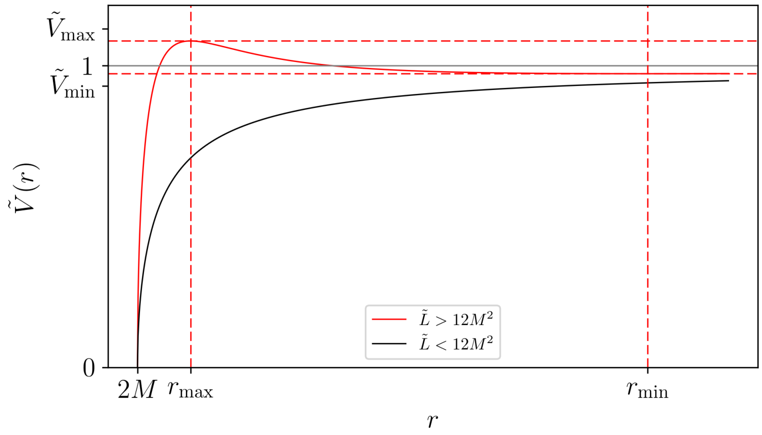
and together with

$$\boxed{p^\varphi = m\frac{\mathrm{d}\varphi}{\mathrm{d}\tau} = \frac{m\tilde{L}}{r^2}} \tag{2.9.2.2}$$

we have characterized the motion of our particle completely. Unluckily this is not a trivial differential equation, so we will not try to derive solutions but rather look at different cases for $\tilde{K}$,$\tilde{L}$ and M:

We notice, that the left hand side of (2.9.2.1) is always positive, which means, that

$$\tilde{V}^2(r) \equiv \left(1 - \frac{2M}{r}\right)\left(1 + \frac{\tilde{L}^2}{r^2}\right) < \tilde{K}^2$$



Furthermore we can extract some information from the shape of $\tilde{V}(r)$. The asymptote of this function is 1 for every value of M and $\tilde{L}$ and we see the singularity at $r = 2M$. Furthermore we see from calculating the extrema of $\tilde{V}(r)$ (for full calculation see B.1), that there are two cases:

- $\tilde{L}^2 < 12M^2$: $\tilde{V}(r)$ is monotonic. We then have the following possibilities for the geodesics:
  - → A particle arriving from $r \to \infty$ requires $\tilde{K} > 1$ and just falls onto the star at some radius $R_{\text{star}} > 2M$
  - → A particle, which is ejected by the star can go to ∞ if $\tilde{K} > 1$ or "hit the curve $\tilde{V}(r)$" and fall back if $\tilde{K} < 1$
- $\tilde{L}^2 > 12M^2$: $\tilde{V}(r)$ has a local maximum and a local minimum, which we will call V_{max} and V_{min}. The possibilities for the geodesics are:
  - → A particle arriving from $r \to \infty$ requires $\tilde{K} > 1$ and it
    - ∗ falls on the star if $\tilde{K} > \tilde{V}_{\text{max}}$
    - ∗ "hits the curve $\tilde{V}(r)$" and goes back to ∞ if $\tilde{K} < \tilde{V}_{\text{max}}$ similar to the hyperbolic orbits in Newtonian gravitation (it is not an exact hyperbola though)
  - → A particle, which is ejected by the star can
    - ∗ go to ∞ if $\tilde{K} > \tilde{V}_{\text{max}} > 1$
    - ∗ fall back if $1 < \tilde{K} < \tilde{V}_{\text{max}}$

again similar to Newtonian gravity

→ Finally particles can be trapped in a *close orbit* if $\tilde{V}_{\min} < \tilde{K} < 1$ with the special case of a circular orbit if $\tilde{K} = \tilde{V}_{\min}$. Otherwise we get a nearly elliptic trajectory. This solution is not an exact ellipse though:
If one solves $r(\tau)$ for this case one finds, that $r(\tau)$ is **periodic** but $\varphi(r)$ is **monotonic** with different periodicity than the radius, which means that a fixed point, i.e. the point where r is minimal in each rotation, called the *perihelion* rotates by some angle $\delta\varphi$

2.9.3. Calculation of perihelion shift

To calculate the precession angle of a planet (in this case mercury) we need to do two steps:
First we need to integrate

$$\frac{\mathrm{d}r}{\mathrm{d}\tau} = \sqrt{\tilde{K}(r)^2 - \tilde{V}(r)^2}$$

to find the period $\Delta\tau$ between two minima of r.
We furthermore know, that

$$\frac{\mathrm{d}\varphi}{\mathrm{d}\tau} = \frac{p^\varphi}{m} = \frac{g^{\varphi\varphi}p_\varphi}{m} = \frac{\tilde{L}}{r^2}$$

so we need to compute

$$\Delta\varphi = \int_0^{\Delta\tau} \frac{\mathrm{d}\varphi}{\mathrm{d}\tau}\,\mathrm{d}\tau = \int_0^{\Delta\tau} \frac{\tilde{L}}{(r(\tau))^2}\,\mathrm{d}\tau$$

unluckily we again have **no easy analytic solution** but an expansion in $\frac{M}{\tilde{L}}$ gives us the leading post-Newtonian correction for a *nearly circular* orbit. The proof for why this gives this limit can be done in two ways:

1. We know that in the weak field limit $\frac{M}{r} \ll 1$ but for a bound orbit r also stays in the order of $r_{\min}$ (see (B.1.0.1)):

$$r_{\min} = \frac{\tilde{L}^2}{2M}\left(1 + \sqrt{1 - \frac{12M^2}{\tilde{L}^2}}\right)^{-1}$$

so

$$\frac{M}{r(\tau)} \sim \frac{M}{r_{\min}} = \frac{2M^2}{\tilde{L}^2}\left(1 + \sqrt{1 - \frac{12M^2}{\tilde{L}^2}}\right)^{-1} \overset{!}{\ll} 1$$

which is only fulfilled for $\frac{M^2}{\tilde{L}^2} \ll 1$

2. Use the newtonian limit where the angular momentum is just $L = mrv$. Then for a circular orbit

$$v = \sqrt{\frac{M}{r}}, \qquad |v| \ll 1$$

In GR we know, that

$$p_\varphi = g_{\varphi\varphi} p^\varphi = r^2 m \frac{\mathrm{d}\varphi}{\mathrm{d}\tau} = mr(r\frac{\mathrm{d}\varphi}{\mathrm{d}\tau}) = mrv$$

so

$$\tilde{L} = rv \quad \Rightarrow \quad \frac{M^2}{\tilde{L}^2} = \frac{M^2}{r^2 v^2} = \frac{M^2}{r^2\left(\frac{M}{r}\right)} = \frac{M}{r}$$

This means, that if we again take the weak field limit $\frac{M}{r} \ll 1$ we also get $\frac{M^2}{\tilde{L}^2} \ll 1$

Using this approximation we can calculate the perihelion shift in leading order in $\frac{M}{r}$[21]:

$$\boxed{\Delta\varphi = 2\pi + 6\pi\frac{M^2}{\tilde{L}^2} + \mathcal{O}\left(\frac{M^4}{\tilde{L}^4}\right) = 2\pi + 6\pi\frac{M}{r} + \mathcal{O}\left(\frac{M^2}{r^2}\right)}$$

In the special (historical) case of mercury if we plug in $M = M_\odot$ and mercury's orbiting radius $R \sim 0.4\,\mathrm{AU}$ we get

$$\Delta\varphi - 2\pi \approx 5 \cdot 10^{-7}\,\mathrm{rad}$$

and after computing the number of orbits per century, one finds a perihelion shift of 43″ per century, which is exactly what astronomers had found in 1882 and which had no physical explanation so far.

2.9.4. Evolution of a photon

In the case of a photon we do analogous considerations to a massive particle. The equations

$$\begin{cases} p_0 = \mathrm{const} \equiv -K \\ p_\theta = 0 \\ p_\varphi = \mathrm{const} \equiv L \end{cases}$$

still holds, so we get

$$\begin{cases} p^0 = g^{00} p_0 = \left(1 - \frac{2M}{r}\right)^{-1} K \\ p^r = \frac{\mathrm{d}r}{\mathrm{d}\lambda} \\ p^\varphi = g^{\varphi\varphi} p_\varphi = \frac{L}{r^2} = \frac{\mathrm{d}\varphi}{\mathrm{d}\lambda} \end{cases}$$

[21] We will skip the explicit calculation here, since it is quite lengthy.

Again we can take the shortcut not by using $\vec{p} \cdot \vec{p} = 0$ instead of solving the geodesic equation:

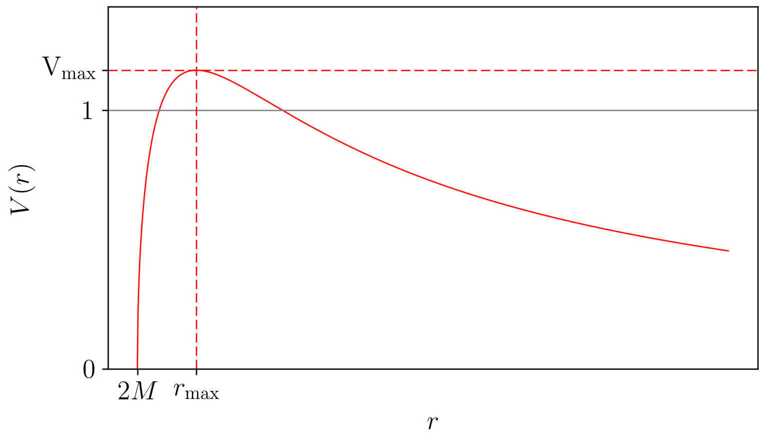
$$0 = p^0 g_{00} p^0 + p^r g_{rr} p^r + p^\varphi g_{\varphi\varphi} p^\varphi$$

$$\Rightarrow \left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 = K^2 - \left(1 - \frac{2M}{r}\right)\frac{L^2}{r^2} \tag{2.9.4.1}$$

so we can define $V^2(r)$ analogously to above:

$$V^2(r) \equiv \left(1 - \frac{2M}{r}\right)\frac{L^2}{r^2} < K^2$$

again we look at the shape of $V(r)$ (for a detailed calculation see B.1):



→ A photon (or any other massless particle) arriving from $r \to \infty$

* Falls on the star if $K > V_{\max}$, i.e. if $L < \sqrt{27}KM$[22]
* "hits curve $V(r)$" and goes back to $r \to \infty$ if $K < V_{\max}$. This deflection of light is called *gravitational lensing* (see 2.8.5).

→ A photon which is ejected from the star

* goes to ∞ if $K > V_{\max}$.
* Falls back in the (very rare) case $2M < R_{\text{star}} < r < 3M$[23]

We furthermore see, that there is **no closed orbit** for light rays[24].

[22] Later we will see, that $\frac{L}{K}$ is the *impact parameter* d, so we can rephrase this condition to $d < \sqrt{27}M$
[23] While this is a very strong constraint *neutron stars* fulfil this condition.
[24] $r = 3M$ would be a closed orbit but it is unstable.

2.9.5. Deflection of light

We now want to **explicitly calculate** the deflection angle for gravitational lensing. We know (rename $K \to E$)

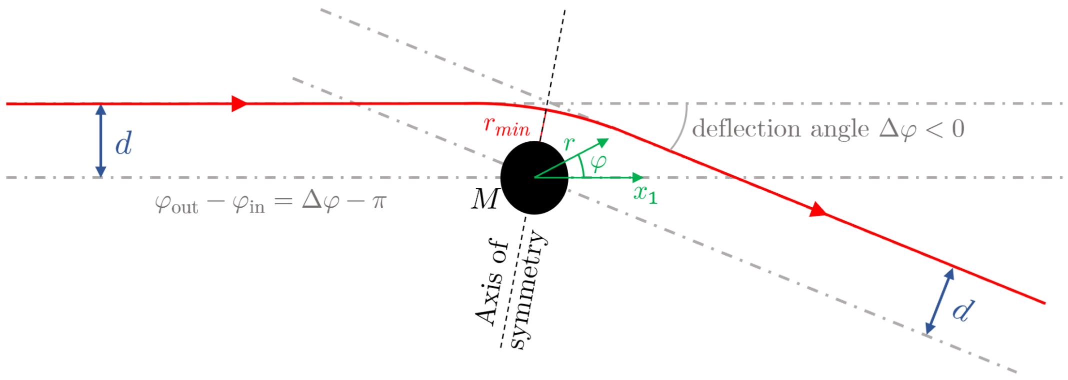
$$\frac{\mathrm{d}r}{\mathrm{d}\lambda} = \pm\sqrt{E^2 - \left(1 - \frac{2M}{r}\right)\frac{L^2}{r^2}} \qquad \frac{\mathrm{d}\varphi}{\mathrm{d}\lambda} = \frac{L}{r^2}$$

so

$$\frac{\mathrm{d}\varphi}{\mathrm{d}r} = \pm\frac{1}{r^2}\left[\left(\frac{E}{L}\right)^2 - \frac{1}{r^2}\left(1 - \frac{2M}{r}\right)\right]^{-1/2}$$

We notice that when $r \to \infty$ (only approximating the term in the square root):

$$\lim_{r\to\infty}\frac{\mathrm{d}\varphi}{\mathrm{d}r} = \pm\frac{1}{r^2}\frac{L}{E}$$



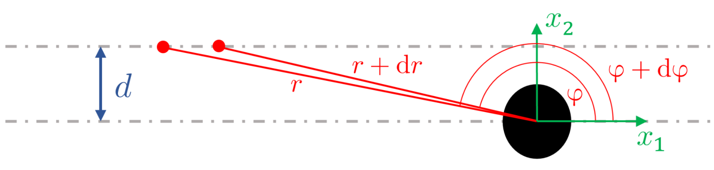
Furthermore we know, that when the photon

- arrives: $\frac{\mathrm{d}\varphi}{\mathrm{d}\lambda} < 0$, $\frac{\mathrm{d}r}{\mathrm{d}\lambda} < 0 \Rightarrow \frac{\mathrm{d}\varphi}{\mathrm{d}r} > 0$
- leaves: $\frac{\mathrm{d}\varphi}{\mathrm{d}\lambda} < 0$, $\frac{\mathrm{d}r}{\mathrm{d}\lambda} > 0 \Rightarrow \frac{\mathrm{d}\varphi}{\mathrm{d}r} < 0$

So we can associate the signs for $\frac{\mathrm{d}\varphi}{\mathrm{d}r}$ to the incoming and outgoing photon:

$$\begin{cases} t \to -\infty & \Rightarrow \quad \frac{\mathrm{d}\varphi}{\mathrm{d}r} \to +\frac{1}{r^2}\frac{L}{E} \\ t \to +\infty & \Rightarrow \quad \frac{\mathrm{d}\varphi}{\mathrm{d}r} \to -\frac{1}{r^2}\frac{L}{E} \end{cases}$$

We can now use a **geometrical reasoning** at $t = \pm\infty$ to show, that $(\frac{L}{E}) = d$:



In the small angle approximation we get

$$\begin{aligned} d = r(\pi - \varphi) \quad &\Rightarrow \quad \pi - \varphi = \frac{d}{r} \\ d = (r + \mathrm{d}r)(\pi - (\varphi + \mathrm{d}\varphi)) \quad &\Rightarrow \quad \pi - \varphi - \mathrm{d}\varphi = \frac{d}{r + \mathrm{d}r} \end{aligned}$$

So altogether we get

$$\mathrm{d}\varphi = \frac{d}{r} - \frac{d}{r + \mathrm{d}r} = \frac{d}{r}\left(1 - \frac{1}{1 + \frac{\mathrm{d}r}{r}}\right) \overset{\text{taylor}}{\approx} \frac{d}{r}\left(1 - 1 + \frac{\mathrm{d}r}{r}\right) = \frac{d \cdot \mathrm{d}r}{r^2}$$

so

$$\frac{\mathrm{d}\varphi}{\mathrm{d}r} = \frac{d}{r^2} = \frac{L}{r^2 K} \Rightarrow \boxed{d = \frac{L}{K}}$$

We can then calculate the change in angle $\varphi_{\text{out}} - \varphi_{\text{in}}$ by integrating:

$$\begin{aligned} \varphi_{\text{out}} - \varphi_{\text{in}} = \int_{\varphi_{\text{in}}}^{\varphi_{\text{out}}} \mathrm{d}\varphi = \int_{+\infty}^{r_{\min}} \underbrace{\frac{\mathrm{d}\varphi}{\mathrm{d}r}}_{>0} \mathrm{d}r + \int_{R_{\min}}^{+\infty} \underbrace{\frac{\mathrm{d}\varphi}{\mathrm{d}r}}_{<0} \mathrm{d}r \\ = 2\int_{r_{\min}}^{+\infty} -\frac{1}{r^2}\left[d^{-2} - \frac{1}{r^2}\left(1 - \frac{2M}{r}\right)\right]^{-1/2} \mathrm{d}r \\ \equiv \Delta\varphi - \pi \end{aligned}$$

Again we can take the weak field limit $|\Delta\varphi| \ll 1$ which requires $r_{\min} \sim d$ which can be rewritten to

$$\frac{M}{r} \ll 1 \Leftrightarrow \frac{M}{r_{\min}} \ll 1 \Leftrightarrow \frac{M}{d} \ll 1$$

in this limit we find (not derived here) **Herleitung?**

$$|\Delta\varphi| = 4\frac{M}{d} + \mathcal{O}\left(\frac{M^2}{d^2}\right)$$

Remark: Before GR there were already attempts to incorporate the deflection of light in Newton's theory but there already some problems occured:

- If one assumes a photon to be a massive particle at speed c and takes the limit $m \to 0$ one finds $\Delta\varphi \sim \frac{2M}{d}$
- If one assumes the photon to be massless from the beginning the deflection is $\Delta\varphi = 0$

The gravitational deflection of light was first observed by A. Eddington during a total solar eclipse in 1919. He measured the difference in angle when stars passed by the sun and the result corresponded to the predicted difference of $\Delta\varphi = 4\frac{M_\odot}{R_\odot} = 1.74''$

While this effect is very small in case of the sun, there are spectacular clusters and galaxies, which cause *strong lensing* so one gets arclets, multiple images etc.

2.10. Black holes

The prediction, that objects can exist from which light cannot escape is not really new. In Newtonian mechanics if one assumes the speed of light to be finite and and uses the classical escape velocity

$$v^2 = \frac{2GM}{r}$$

we see, that no light can escape if

$$\frac{2GM}{r} > c^2$$

which in geometrized units simplifies to

$$\boxed{r < 2M}$$

with the mass M and radius r of the object[25].
This was already postulated in the 18th century, since it was already known by then that the speed of light is finite. Since we can also express M and r by the density

$$\rho = \frac{3M}{4\pi r^3}$$

so we can express the limit in terms of M, r and ρ (where we need two of the above). In particular this does not impose a limit on ρ. For example:

- If $\rho = \rho_{\text{water}}$ we get a black hole if $M \geq 10^9 \cdot M_\odot$
- If $R = R_\odot$ we get a black hole if $\rho \geq 10^6 \cdot \rho_\odot$
- if $M = M_\odot$ we get a black hole if $\rho \geq 10^{18} \cdot \rho_\odot$

2.10.1. The Schwarzschild metric at $r \to 2M$

We look at the limit $r \to 2M$ and look at the behaviour of the metric $g_{rr} \to \infty$. Consider a radial infall:

$$\frac{\mathrm{d}\varphi}{\mathrm{d}r} = 0, \quad \tilde{L} = 0 \quad \text{(we consider a massive particle)}$$

We look at the coordinate- and proper time:

1. We already know that

$$\frac{\mathrm{d}r}{\mathrm{d}\tau} = -\left(\tilde{K}^2 - 1 + \frac{2M}{r}\right)^{-1/2} \qquad \text{where } \tau = \text{proper time}$$

[25] By coincidence this is also the solution we get in GR.

Then the time to go from some (fixed) $r_{\text{ini}} > 2M$ to $r = 2M$ is

$$\Delta\tau = \int_{r_{\text{ini}}}^{2M} \left(\frac{\tau}{r}\right) \, \mathrm{d}r = \int_{r_{\text{ini}}}^{2M} -\frac{\mathrm{d}r}{\sqrt{\tilde{K}^2 - 1 + \frac{2M}{r}}}$$

which gives a **finite, well defined result**.

2. The 4-velocity of the object is $\vec{U} = (\frac{\mathrm{d}t}{\mathrm{d}\tau}, \frac{\mathrm{d}r}{\mathrm{d}\tau}, 0, 0)$ so

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = U^0 = g^{00}U_0 = g^{00}\frac{p_0}{m} = -g^{00}\tilde{K} = \left(1 - \frac{2M}{r}\right)^{-1}\tilde{K}$$

The problem that arises from this is, that for $r \to 2M$:

$$\Delta t = \int_{r_{\text{ini}}}^{2M} \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)\left(\frac{\mathrm{d}\tau}{\mathrm{d}r}\right) \, \mathrm{d}r = \int_{r_{\text{ini}}}^{2M} -\frac{\tilde{K} \, \mathrm{d}r}{\left(1 - \frac{2M}{r}\right)\sqrt{\tilde{K} - 1 + \frac{2M}{r}}} \overset{r \to 2M}{\to} \infty$$

So we see, that the physical proper time is finite, just the coordinate time diverges. This implies, that **nothing is singular**, just our choice of coordinates is inappropriate for studying black holes.

2.10.2. Kruskal-Szekeres coordinates

The first alternative choice of coordinates, which circumvents this problem was presented in 1960 by M. Kruskal and G. Szekeres and thus has the name **Kruskal-Szekeres coordinates** (KS-coordinates). We start from the Schwarzschild metric and redefine the coordinates t, r to[26]

$$\boxed{\begin{aligned} r > 2M : &\begin{cases} R = \sqrt{\frac{r}{2M} - 1}\, e^{\frac{r}{4M}} \cosh\left(\frac{t}{4M}\right) \\ T = \sqrt{\frac{r}{2M} - 1}\, e^{\frac{r}{4M}} \sinh\left(\frac{t}{4M}\right) \end{cases} \\ r < 2M : &\begin{cases} R = \sqrt{1 - \frac{r}{2M}}\, e^{\frac{r}{4M}} \sinh\left(\frac{t}{4M}\right) \\ T = \sqrt{1 - \frac{r}{2M}}\, e^{\frac{r}{4M}} \cosh\left(\frac{t}{4M}\right) \end{cases} \end{aligned}}$$

with $T > -R$. We see, that in these equations we still have the divergences for $r \to 2M$ but the singularities here are **not physical** since they *compensate* the divergences in the Schwarzschild solution.
In fact we can rewrite these coordinates such that at the Schwarzschild radius T and R simply switch:

$$\text{outside:} R > 2M \left\{ \begin{aligned} R &= \sqrt{\left|\frac{r}{2M} - 1\right|}\, e^{\frac{r}{4M}} \cosh\left(\frac{t}{4M}\right) = T \\ T &= \sqrt{\left|\frac{r}{2M} - 1\right|}\, e^{\frac{r}{4M}} \sinh\left(\frac{t}{4M}\right) = R \end{aligned} \right\} \text{inside:} R < 2M$$

[26] We will leave out the explicit proof here since it is quite lengthy.

With this transformation we can rewrite the line element to

$$\boxed{\mathrm{d}s^2 = \frac{32M^3}{r(T,R)} e^{-\frac{r(T,R)}{2M}} (-\mathrm{d}T^2 + \mathrm{d}R^2) + r^2(T,R)\mathrm{d}\Omega^2} \tag{2.10.2.1}$$

which has **no singularities**. Additionally in the case of photons with radial trajectories we get $\mathrm{d}R = \pm \mathrm{d}T$ like in the Minkowski metric.

Remarks

- There is still a singularity at $r \to 0$ because we assume all mass to be concentrated in one single point (this is of course not realistic)
- KS coordinates are better for studying black holes but Schwarzschild coordinates in practice are often still better for other problems because of the obvious Newtonian limit, the limit $t \to \tau$ for $r \to \infty$, obvious asymptotic flatness and the generally simpler form of equations.
- (T, R) is defined in a non-trivial region only. To find this region we look at the behaviour of T and R. We find:

$$\frac{T}{R} = \tanh\left(\frac{t}{4M}\right) \qquad R > 2M$$

$$\frac{T}{R} = \tanh\left(\frac{t}{4M}\right)^{-1} \qquad R < 2M$$

and

$$R^2 - T^2 = \left|\frac{r}{2M} - 1\right| e^{\frac{r}{2M}} \tag{2.10.2.2}$$

which means, that $t = \text{const} \Leftrightarrow \frac{T}{R} = \text{const}$ and $r = \text{const} \leftrightarrow R^2 - T^2 = \text{const}$. This means, that we can plot T and R:

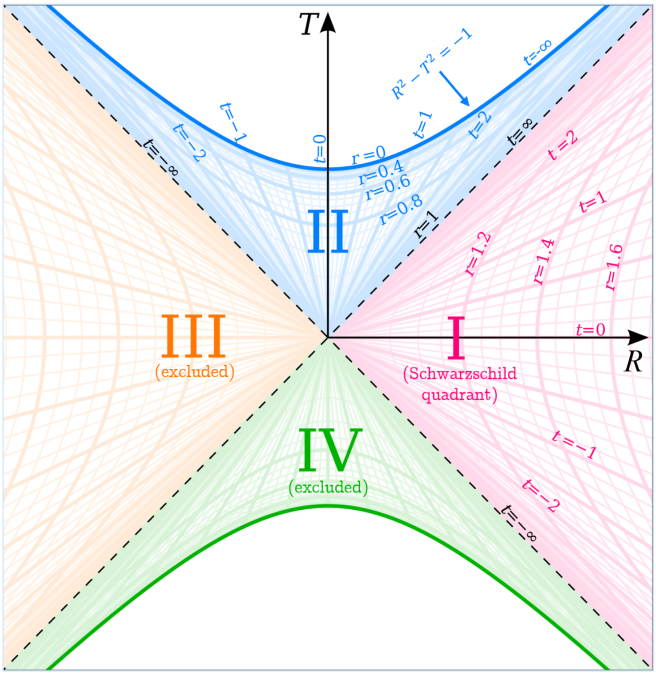


Figure 2.1.: KS coordinates with $2M = r_{\text{Schwarzschild}} = 1$ (taken from Wikipedia)

We see, that quadrant I is the *exterior region* of the Black hole and II the *interior region*, whereas the boundary between *exterior* and *interior* is the Schwarzschild radius.
We exclude quadrants III and IV since they give the same results as I and II just with negative t and r and this is unphysical. There are theories though, which also consider these parts and call quadrant III the *parallel exterior region* and IV a *white hole*.

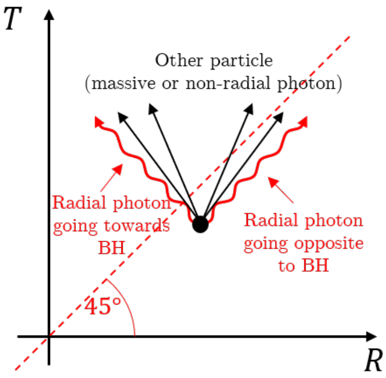
- We know, that for massive/massless particles we always have $\mathrm{d}s^2 \leq 0$, which means, that

$$\left(\frac{T}{R}\right)^2 \geq 1 + \frac{r^3}{32M^3} e^{\frac{r}{2M}} \left(\frac{\mathrm{d}\Omega}{\mathrm{d}R}\right)^2 \quad \Rightarrow \quad \left|\frac{\mathrm{d}T}{\mathrm{d}R}\right| \geq 1$$

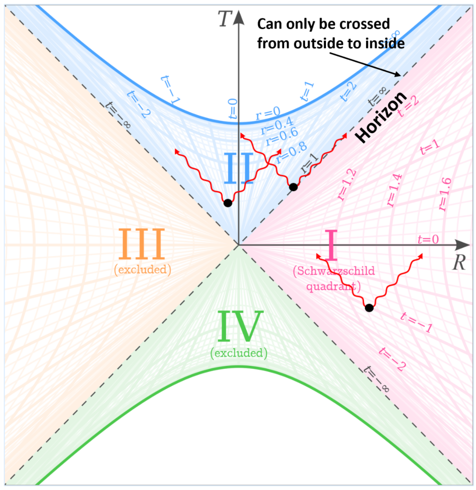
or more precisely:

$$\begin{aligned} \left|\frac{\mathrm{d}T}{\mathrm{d}R}\right| &= 1 \quad \text{for radial photons} \\ &> 1 \quad \text{for non-radial photons and massive particles} \end{aligned}$$

This means, that we can draw radial photons as 45° lines in our diagram:



We immediately conclude from this, that **nothing can cross** the divide from $r < 2M$ $(T > R)$ to $r > 2M$ $(T < R)$:



Furthermore we see, that any geodesic with a least one event in the region $r < 2M$ will necessarily end up on the Black hole at $r \to 0$, so the sphere with $r = 2M$ is the **Horizon of the Schwarzschild BH**[27].

Remarks:

- We only discussed the Schwarzschild BH, which is **spherically symmetric, static and with constant** R. This is not a realistic assumption since a general BH is **rotating** and thus only has axial symmetry (Kerr BH), more complicated in the sense that it is asymmetric, dynamical, accreting mass etc.
- The Notion of a BH horizon is **loosely defined**:
  For a dynamical BH we would need to consider the **entire spacetime** until $t \to \infty$,

[27] Horizon in this sense means, that we **cannot see** past this point similar to the horizon on earth.

since only actually calculating the geodesic will show us whether a particle can escape the BH or not.

- In any case the horizon has no special physical property. It is like any other place around a BH[28].

- Mathematically the KS coordinates do a very peculiar thing to achieve their special properties. They are **time dependent**! We can see this if we look at any point with fixed r (e.g. the Schwarzschild horizon $r = 2M$) where we can use (2.10.2.2)

$$R^2 - T^2 = \text{const} \qquad (R^2 - T^2 = 0 \text{ in the Schwarzschild case})$$

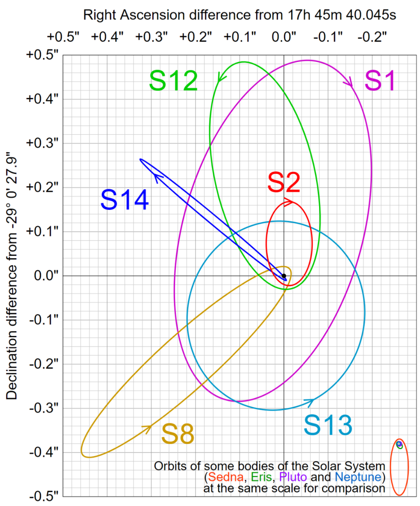
so as T increases R has to increase as well. Hence the radial coordinate R is **moving outwards** with increasing time T. This expansion happens slower than c outside, and faster than c inside the Schwarzschild horizon.

2.10.3. Observing Black holes

Since Black holes do not emit any radiation[29] we can only observe BH *indirectly*:

- **Stellar mass BH**: $1 - 10\,M_\odot$
  We know that these BH are formed out of stars when a star becomes a white dwarf (delocalised electrons), which then becomes a neutron star (only neutrons) and finally collapses into a black hole.

  Since the BH is invisible when isolated we only see them when another object orbits them (*binary systems*, e.g. pulsar and BH).

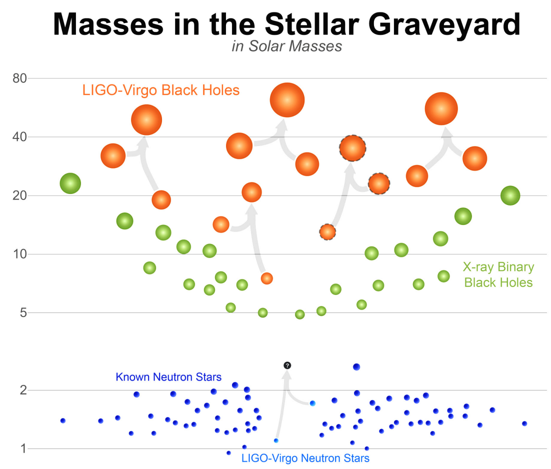


- **Supermassive BH**: $\sim 10^6\,M_\odot$
  These are found near the center of almost every galaxy. They are detected through the motion of mass.

- **Intermediate BH**: $10 - 10^6\,M_\odot$
  These were only found in 2015 with observation of graviational waves, which found BH merging with masses $M \sim 10 - 70\,M_\odot$. There is no astrophysical explanation for them yet.

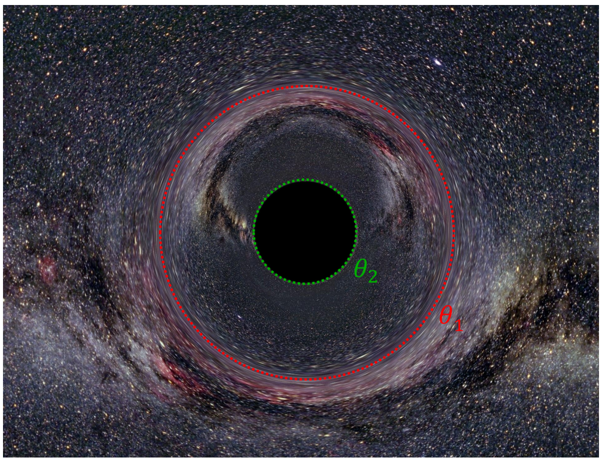
[28] Unlike in most science-fiction where weird stuff happens at the Schwarzschild radius. We will explicitly calculate this in ex. 10, task 1

[29] This is not entirely correct. There is a postulated radiation, the *Hawking radiation*.



2.10.4. Gravitational lensing around black holes

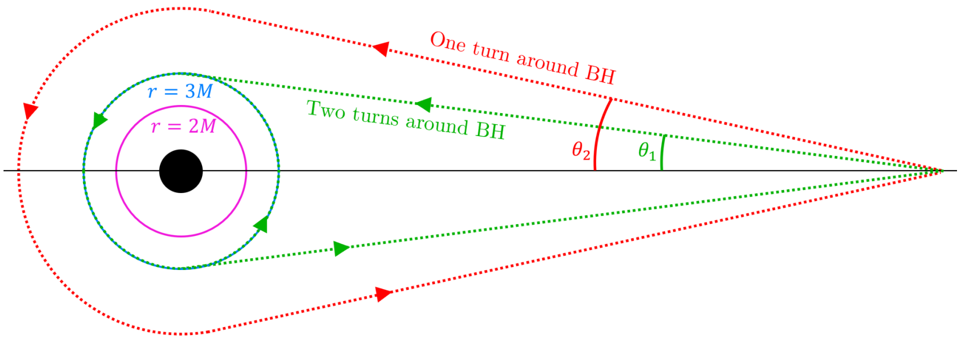
We want to understand the lensing patterns that would be visible, if we were looking directly at a black hole. While no telescope has yet taken any direct images of the lensing pattern of a black hole, we can simulate it by just calculating the geodesics of photons around it. Such a simulation is shown in the figure below.



We have drawn in two distinctive angles θ_1 and θ_2 whereas

- below θ_2: **completely black** region
- around θ_2: A **fuzzy ring**
- around θ_1: Another fuzzy ring
- otherwise we see more or less distorted images of the stars/clouds surrounding the BH.

The interpretation of this pattern is, that θ_1 and θ_2 correspond to the geodesics shown in the image below.



More turns are also possible but $\theta_2 \approx \theta_3 \approx \theta_4 \ldots$. This means, that We have three regions:

- $\theta > \theta_1$: Image of the **full sky** in all directions (even behind the observer)
- $\theta_2 < \theta < \theta_1$: Second image of the sky, but inverted (like a mirror)
- $\theta < \theta_2$: No possible geodesics/all geodesics end in the BH. So the black circle is an **"image" of the sphere with radius** $3M$**!** (not the Schwarzschild radius at 2M)

2.11. Gravitational waves

We have not yet looked at the question whether for a given $T_{\mu\nu}$, $g_{\mu\nu}$ is unique. It turns out, that from the Einstein equation

$$G_{\mu\nu}[g_{\mu\nu}] = 8\pi T_{\mu\nu}$$

the answer is **no**. The solutions differ through gravitational waves (GWs), which are solutions to the equation

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

with a small amplitude ($\delta g_{\mu\nu} \ll 1$).

2.11.1. Gravitational waves in nearly flat spacetime

We will first look at the concept of gravitational waves in the weak field limit:

$$-\frac{1}{2}\Box\bar{h}^{\mu\nu} = 8\pi T^{\mu\nu}$$

which has the advantage, that this limit turns our second order DEQ into a first order one. We will furthermore look at GWs propagating in a Minkowski background:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$$

and with $\square = \partial_\mu \eta^{\mu\nu} \partial_\nu$ gravitational waves are just the solutions of

$$\boxed{\square \bar{h}^{\mu\nu} = 0} \tag{2.11.1.1}$$

Note: In contrast to the newtonian limit from 2.8.5 where we take the weak field limit with $\bar{h}^{00} \gg \bar{h}^{0i} \gg \bar{h}^{ij}$ here we **only** take the weak field limit, so

$$\bar{h}^{00} \sim \bar{h}^{0i} \sim \bar{h}^{ij}$$

The solution to our differential equation are then **plane waves**:

$$\bar{h}^{\mu\nu} = \mathrm{Re}\left[A^{\mu\nu} e^{\mathrm{i} k_\alpha x^\alpha}\right] \tag{2.11.1.2}$$

with the *4-wave covector* $\vec{k}$ and the *complex tensor* $A^{\mu\nu}$ which also contains the phase information.

We look at some properties of waves, that we already from electrodynamics:

$$\psi = A e^{\mathrm{i}(-\omega t + \boldsymbol{k} \cdot \boldsymbol{x})}$$

with the *frequency* ω and the *3 wavevector* $\boldsymbol{k}$.
This gives the dispersion relation:

$$v(k) = \frac{\omega(k)}{|\boldsymbol{k}|} \quad \Rightarrow \quad \omega(k) = v(k)|\boldsymbol{k}|$$

and if v is independent of k then

$$\omega(k) = v|\boldsymbol{k}| = v(\boldsymbol{k} \cdot \boldsymbol{k})^{1/2}$$

For gravitational waves this becomes

$$\begin{aligned} \square \bar{h}_{\mu\nu} &= \eta^{\alpha\beta} \partial_\alpha \partial_\beta \bar{h}_{\mu\nu} = 0 \\ \Rightarrow \eta^{\alpha\beta} (\mathrm{i}k_\alpha)(\mathrm{i}k_\beta) \bar{h}_{\mu\nu} &= 0 \end{aligned}$$

and from this directly follows

$$k_\alpha \eta^{\alpha\beta} k_\beta = 0 = k_\alpha k^\alpha = \text{null vector}$$

which implies, that k^α **must be tangent to the worldline of photons** (see 1.6.7). In turn this implies, that

$$k^\alpha \eta_{\alpha\beta} k^\beta = -(k^0)^2 + (k^1)^2 + (k^2)^2 + (k^3)^2 = 0 \quad \Rightarrow \quad \omega^2 = |\boldsymbol{k}|^2$$

which means, that **all gravitational waves propagate at the speed of light**.

Note, that to obtain

$$-\frac{1}{2}\Box\bar{h}^{\alpha\beta} = 8\pi G T^{\alpha\beta}$$

and

$$\Box\bar{h}^{\alpha\beta} = 0$$

for gravitational waves, we worked in the **Lorenz gauge** which restricts $A^{\alpha\beta}$ since

$$\bar{h}^{\alpha\beta}{}_{,\beta} = 0 \quad \Rightarrow \quad (\mathrm{i}k_\beta)A^{\alpha\beta}\exp[\ldots] = 0$$

Thus

$$A^{\alpha\beta}k_\beta = 0$$

which means, that $A^{\alpha\beta}$ **must be orthogonal** to k^β.

As we can see though, the Lorenz gauge is **not unique**:
We constructed it using ξ^α such that

$$\Box\xi^\alpha = \bar{h}^{\alpha\beta}{}_{,\beta}$$

but since we can always add a ξ^α with $\tilde{\xi}^\alpha = \xi^\alpha + \zeta^\alpha$ which fulfils

$$\zeta^\alpha = B^\alpha \exp(\mathrm{i}k_\mu x^\mu), \quad k_\mu k^\mu \overset{!}{=} 0 \qquad \text{(We omitted the Re}(\ldots))$$
$$\Rightarrow \Box\zeta^\alpha = \eta^{\alpha\beta}\frac{\partial}{\partial x^\alpha}\frac{\partial}{\partial x^\beta}\xi^\beta = k_\mu k^\mu B^\alpha \exp(\mathrm{i}k_\mu x^\mu) = 0$$

we have this degree of freedom to impose an **additional restriction**. But what is the most convenient way to chose it?
This leads us to the following theorem (will be proven in ex.11, task 1):

> If we take the same k^μ for the GW and for ξ^α, we can always choose B^α in such a way, that
>
> $$\underbrace{A^\alpha{}_\alpha = 0}_{\mathrm{Tr}(A^{\alpha\beta})=0} \quad \text{and} \quad \underbrace{A_{\alpha\beta}U^\beta = 0}_{A^{\alpha\beta}\perp U^\beta} \quad \text{for any fixed } U^\beta$$
>
> Additionally the condition from the beginning still holds:
>
> $$A_{\alpha\beta}k^\beta = 0$$
>
> So this type of gauge is a special type of Lorenz gauge. We get the simplest equations by chosing $\vec{U} = (1,0,0,0)$, which then is called the *traceless transverse (TT) gauge*. Most calculations with GW are done in this gauge.

In the TT gauge it is easy to compute $A_{\alpha\beta}$. For a GW going in x^3-direction ($\vec{k} = (\omega, 0, 0, \omega)$) we get for example:

$$A_{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{11} & A_{12} & 0 \\ 0 & A_{12} & -A_{11} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

We get this result if we use our three conditions:

$$A_{\alpha\beta}U^\beta = 0 \Leftrightarrow A_{\alpha\beta}\delta^\beta_0 = 0 \Leftrightarrow A_{\alpha 0} = 0$$

$$A_{\alpha\beta}k^\beta = 0 \Leftarrow A_{\alpha 3}k^3 = 0 \Leftarrow \omega A_{\alpha 3} = 0$$

$$A^\alpha{}_\alpha = 0 \Leftrightarrow \eta^{\alpha\beta}A_{\alpha\beta} = 0 \Leftarrow \eta^{11}A_{11} + \eta^{22}A_{22} \Leftrightarrow A_{22} = -A_{11}$$

This means, that we have **2 independent modes**.

Remarks:

- If we chose a different $\vec{k}$ we get different non-zero components and if we chose a different $\vec{U}$ we also get non-zero component in the time part but still the number of independent modes **remains as two**[30].
- Here we are studying GWs propagating in the Minkowski background (flat space+vacuum everywhere). Of course this does not have to be always the case. Generally the metric is of the form

$$g_{\mu\nu} = \bar{g}_{\mu\nu} + \underbrace{h_{\mu\nu}}_{\text{GWs}}$$

  where $\bar{g}_{\mu\nu}$ is any *background metric* describing a physical problem. The solution for $h_{\mu\nu}$ are then not simply plane waves but take up a more complicated form. Nevertheless they still have the **same two independent degrees of freedom** for every k^μ, i.e. for every frequency ω and direction $\left(\frac{\vec{k}}{\omega}\right)$
- In the TT gauge since

$$A^\alpha{}_\alpha = 0 \quad \Rightarrow \quad h = -\bar{h} = h^\alpha{}_\alpha = 0$$

  We can easily see that $h^{\alpha\beta}$ and $\bar{h}^{\alpha\beta}$ are the same

$$h^{\alpha\beta} = \bar{h}^{\alpha\beta} - \frac{1}{2}\eta^{\alpha\beta}\underbrace{\bar{h}}_{=0} = \bar{h}^{\alpha\beta}$$

  which means, that the **GW solution directly gives the full perturbed metric**:

$$g_{\mu\nu} = \eta_{\mu\nu} + A_{\mu\nu}e^{\mathrm{i}k_\alpha x^\alpha}$$

[30] This number of independent modes will be important in 2.11.5

2.11.2. Effect of gravitational waves on test particles

Intuitively we would expect that if gravitational waves are measurable they need to exert some kind of **force** (since we can only measure effects if they induce some kind of interaction which is only possible through forces). To check this we work in the TT gauge and place a massive particle which is **initially at rest** at $t = t_0$, so $\frac{\mathrm{d}x^i}{\mathrm{d}t} = 0$. We can compute the motion of the particle with the equation of geodesics:

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}t^2} = -\Gamma^\alpha_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}t} \frac{\mathrm{d}x^\nu}{\mathrm{d}t}$$

Since we are in almost flat spacetime $x^0 = t$ =proper time and the initial 4-velocity is $\vec{U} = \frac{\mathrm{d}\vec{x}}{\mathrm{d}t} = (1, 0, 0, 0)$ which means, that

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}t^2} = -\Gamma^\alpha_{00} \cdot 1 \cdot 1$$

with

$$\begin{aligned}
\Gamma^\alpha_{00} =& \frac{1}{2} g^{\alpha\beta} (g_{\beta 0,0} + g_{0\beta,0} - g_{00,\beta}) \\
=& \frac{1}{2} \eta^{\alpha\beta} \underbrace{(h_{\beta 0,0} + h_{0\beta,0} - h_{00,\beta})}_{\text{vanish in the TT gauge}} \\
=& \, 0
\end{aligned}$$

So we get

$$\left.\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}t^2}\right|_{t=t_0} = 0$$

which means, that the test particle feels **no acceleration** and remains at rest forever. The conclusion is, that GW generate **no apparent force** on particles. Their effect (if it exists) must be truly different from any possible effect from Newtonian gravity or known waves (such as EM-waves, mechanical waves, ...).

To actually see the effect of GW we need to take *two* particles M and N, which we place at the locations

$$\vec{M} = (t, 0, 0, 0) \qquad \vec{N} = (t, \varepsilon^1, \varepsilon^2, \varepsilon^3)$$

where we assume ε^i to be small for simplicity.
Like before we assume,that the GWs propagate along the $\boldsymbol{e}_3$-direction but this time we compute the **distance** between $\vec{M}$ and $\vec{N}$:

$$\begin{aligned}
\mathrm{d}s^2 &= \mathrm{d}x^\mu g_{\mu\nu} \mathrm{d}x^\nu \quad \text{with} \quad \mathrm{d}x^\mu = \vec{N} - \vec{M} = (0, \varepsilon^1, \varepsilon^2, \varepsilon^3) \\
&= \underbrace{\mathrm{d}x^i (\eta_{ij}) \mathrm{d}x^j}_{=(\varepsilon^1)^2 + (\varepsilon^2)^2 + (\varepsilon^3)^2 = L^2} + \underbrace{\mathrm{d}x^i h_{ij} \mathrm{d}x^j}_{=[(\varepsilon^1)^2 A_{11} + (\varepsilon^2)^2 A_{22} + 2\varepsilon^1 \varepsilon^2 A_{12}] e^{-\mathrm{i}k_\mu x^\mu}}
\end{aligned}$$

where L is the distance in Minkowski spacetime. If we assume a GW in x^3-direction, i.e. $k^\mu = (\omega, 0, 0, \omega)$ and $\mathrm{i}k^\mu x^\mu = -\mathrm{i}\omega(x^0 - x^3)$ we get

$$\text{proper distance} = \mathrm{d}s = \left(L^2 + [A_{11}((\varepsilon^1)^2 - (\varepsilon^2)^2) + 2A_{12}\varepsilon^1\varepsilon^2]e^{-\mathrm{i}\omega(x^0-x^3)}\right)^{1/2}$$

which we can expand to

$$\boxed{\mathrm{d}s = \left(L^2 + \mathrm{Re}\left(\left[A_{11}\left((\varepsilon^1)^2 - (\varepsilon^2)^2\right) + A_{12}\left(\left(\frac{\varepsilon^1 + \varepsilon^2}{\sqrt{2}}\right)^2 - \left(\frac{\varepsilon^1 - \varepsilon^2}{\sqrt{2}}\right)^2\right)\right] e^{\mathrm{i}\omega(x^0-x^3)}\right)\right)^{1/2}}$$

Physical interpretation

We consider the two possible modes $A_{11} \neq 0, A_{12} = 0$ and $A_{12} \neq 0, A_{11} = 0$:

- $A_{11} \neq 0, A_{12} = 0$: We assume, that A_{11} is **real** and **positive**. In this case when

  $$e^{\mathrm{i}\omega(x^0-x^3)} = 1, \quad \text{i.e.} \quad -\omega(x^0 - x^3) = 2\pi N \quad N \in \mathbb{Z}$$

  $\mathrm{d}s$ simplifies to

  $$\mathrm{d}s = (L^2 + A_{11}((\varepsilon^1)^2 - (\varepsilon^2)^2))^{1/2}$$

  so for any ε^i we can say that distances **increase** along x^1, **decrease** along x^2 and remain **invariant** along x^3.
  Analogously when we look half a period later such that

  $$e^{-\mathrm{i}\omega(x^0-x^3)} = -1, \quad \text{i.e.} \quad -\omega(x^0 - x^3) = (2N+1)\pi \quad N \in \mathbb{Z}$$

  Distances **decrease** along x^1 and increase along x^2

- $A_{12} \neq 0, A_{11} = 0$: Again we assume, that A_{12} is real and positive. Then for

  $$-\omega(x^0 - x^3) = 2\pi N \quad N \in \mathbb{Z}$$

  distances **increase** along $\frac{x^1+x^2}{\sqrt{2}}$ and **decrease along** $\frac{x^1-x^2}{\sqrt{2}}$, so the axis of motion is **rotated by 45°**.
  Analogously for

  $$-\omega(x^0 - x^3) = (2N+1)\pi$$

  we get the opposite.

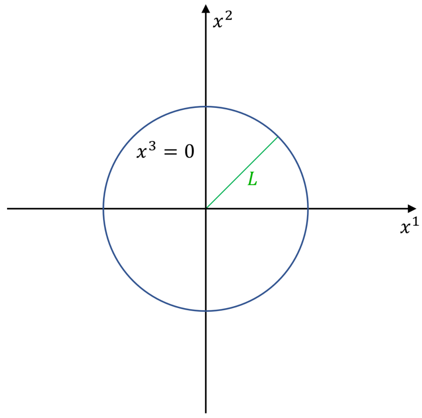
We can visualize this effect if we take the scenario above and imagine particles forming a perfect circle along the x^1, x^2 plane, which is centered on the origin.
Such a circle can be described by the vector

$$\vec{v} = (t, \varepsilon^1, \varepsilon^2, x^3)$$

which fulfils

$$(\varepsilon^1)^2 + (\varepsilon^2)^2 = L^2$$

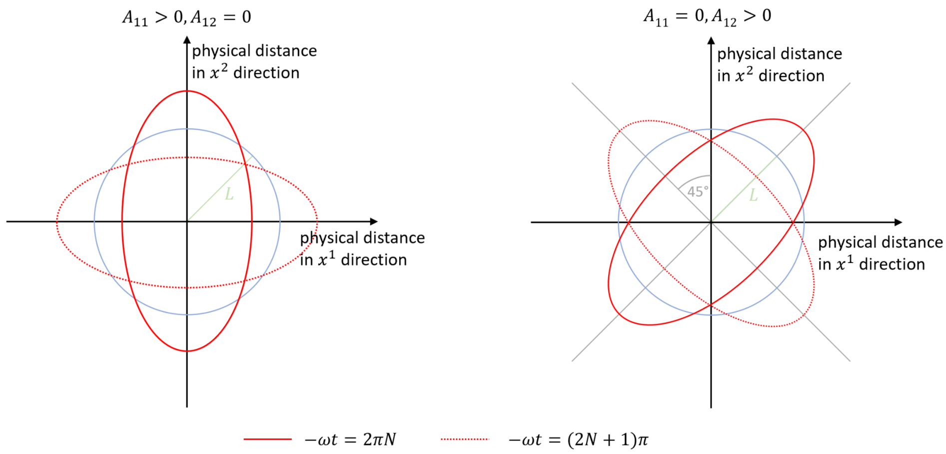
$$x^3 = \text{const (in our case we choose } x^3 = 0)$$



Although these particles keep the same *coordinates* when the GW passes by, their *physical distances* change. Hence we need to look at the physical distance, where we compute the ratio of the two axes of our ellipse. For the maximum elongation ($\omega t = 2\pi N$) we get the ratio

$$\begin{aligned}\frac{\text{large axis}}{\text{small axis}} &= \frac{(L^2 + A_{11}L^2)^{1/2}}{(l^2 - A_{11})^{1/2}} = \frac{\sqrt{1 + A_11}}{\sqrt{1 - A_{11}}} \\ &\approx 1 + \frac{1}{2}A_{11} + \frac{1}{2}A_{11} + \mathcal{O}(A_{11}^2) = 1 + A_{11} + \mathcal{O}(A_{11}^2)\end{aligned}$$

(the result for A_{12} is the same). Hence we get two *modes* which are **rotated by 45°**, **in phase** and have the **same amplitude**.



The modes are often referred as **+-Mode** for $A_{11} \neq 0$ and **×-Mode** for $A_{12} \neq 0$.

Comment: All this looks familiar to the study of the polarization of light. So we can call the two modes the *two degrees of polarization of GWs*[31].

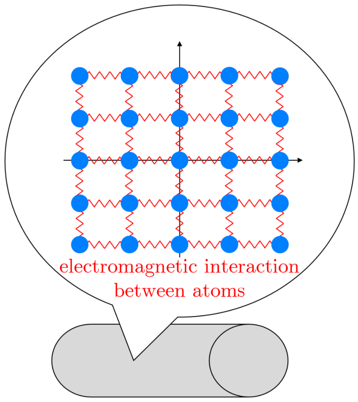
2.11.3. Detection of Gravitational waves

There are two major methods of detecting gravitational waves of which only one has worked so far.

[31] But unlike EM-waves, which are oscillations of fields GW are the polarized oscillations of spacetime.

Resonant bars: Imagine a massive cylinder out of metal (or another crystal).

Without EM interactions each atom of the solid would remain at fixed coordinates when a GW passes by but EM interactions want to keep the atoms at a **fixed physical distance**. This effect is very small but if the GW are oriented correctly w.r.t the bar and if their frequency coincides with the proper oscillation frequency of the (slightly elastic) solid then the oscillations of the bar's physical size is amplified by the resonance, which in principle could be detected. Unluckily the effect is so small that no current experiment has achieved sufficient sensitivity to have measured any GW.

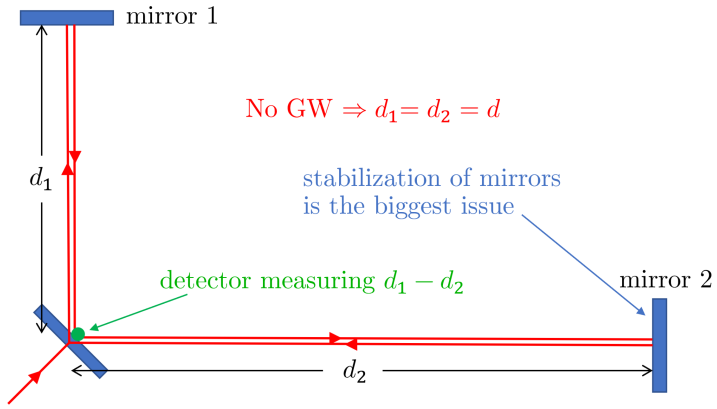


Interferometers: GW interferometers are in principle just large Michelson-Morley interferometers which measure the difference in length of two arms. The advantage of this principle is, that differences in length are a lot easier to measure than absolute lengths. Nevertheless the difference Δd, that a GW causes in a detector of 4 km length is $\sim 10^{-21}$ m (see ex. 10, task 2) so a lot of technical challenges have to be overcome. Several experiments of this type are in operation or in the design phase:

- *LIGO, VIRGO*: $d \sim 3-4$ km, sensitive to $h_{\mu\nu} \sim 10^{-23}$, has detected 10 events in $2015-2018$ which resulted in the Nobel prize in 2017 (Weiss, Barish & Thorne). We will do some calculations regarding the sensitivity of these detectors in ex. 11 task 2.

- *Space projects*: Have the advantage, that longer distances can be covered by the arms and that no geological noise (which is very hard to filter out) can disturb the measurements. *LISA pathfinder* tested technology for space based interferometers until 2017. *LISA* is a planned space based interferometer which will use this technology and which is set to launch in 2034 with an arm length of $\sim 5 \cdot 10^6$ km.

- Several GW interferometers on earth have been proposed so far, the largest project of this kind in the planning phase is the underground *Einstein telescope*, which is set to be in a triangular shape (to allow for two independent interferometers, one for low- and one for high frequencies). Three locations in Europe have been proposed for the telescope, one of which is just a few kilometres from Aachen at the border between the Netherlands and Belgium!



Indirect detection: Some astrophysical systems like binary pulsars emit GWs. Hence they lose energy, slowing the system down by a tiny amount over time. This can be detected through through pulsar's beams. Several of these binary pulsars have been observed and the results match the theoretical prediction of a slow down due to GWs very well. *Hulse & Taylor* were awarded the Nobel prize for this indirect confirmation of the existence of GW in 1993.
The effect is called *pulsar timing effect*.

2.11.4. Production of gravitational waves

We have already looked at the almost flat case where we get plane waves:

$$h_{\mu\nu} = \mathrm{Re}\left(\int \mathrm{d}^3\boldsymbol{k} A_{\mu\nu}(\boldsymbol{k}) e^{\mathrm{i}k_\alpha x^\alpha}\right)$$

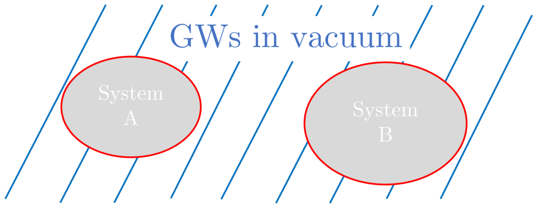
In general if we have a vacuum around arbitrary objects we get a curvature and thus a background metric which is not Minkowski ($g_{\mu\nu} \neq \eta_{\mu\nu}$). Then $\Box \equiv g^{\mu\nu} D_\mu D_\nu$ has eigenfunctions, which are more complicated than just $e^{\mathrm{i}k_\alpha k^\alpha}$. $h_{\mu\nu}$ is then given by

$$h_{\mu\nu} = \int \mathrm{d}^3\boldsymbol{k} A_{\mu\nu}(\boldsymbol{k}) \cdot \mathrm{Eigenfunction}(k_\alpha, x^\alpha)$$

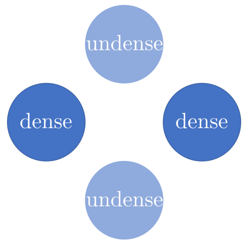
Furthermore to find $A_{\mu\nu}(\boldsymbol{k})$ we can only allow solutions with specific, non-trivial boundary conditions. To demonstrate this we look at two systems A and B which are surrounded by vacuum. Each non-zero contribution of GWs in the vacuum then induces a non-zero contribution on the boundaries of A and B. To find the correct GWs one needs to follow these steps:

1. **Specify** $T^{\alpha\beta} \neq 0$ inside A and B.
2. Solve $G^{\alpha\beta} = 8\pi T^{\alpha\beta}$ to find the **metrics inside** A and B.
3. Solve $G^{\alpha\beta} = 0$ to find the **all solutions outside** where $g_{\mu\nu}^{\text{outside}}$ =background metric + arbitrary superposition of GW modes $A_{\mu\nu}(\boldsymbol{k})$.

4. Match $g_{\mu\nu}^{\text{inside}}$ and $g_{\mu\nu}^{\text{outside}}$ on the boundaries. This implies a **unique solution** for $A_{\mu\nu}(\boldsymbol{k})$[32].



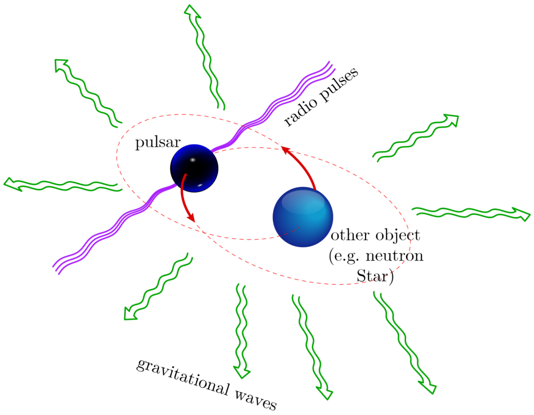
Implications of Birkhoffs theorem on the production of GWs:



Birkhoffs theorem directly implies, that spherically symmetric systems cannot produce any gravitational waves (since their metric is static). In fact one can show, that a **quadrupole** is needed (i.e. a distribution of mass with the symmetry of a quadrupole, for proof see [9]) with **periodic motion**. This also implies, that rotating disks do not generate GWs.
Systems which fulfil the requirements are for example *binary systems* (two objects rotating around each other). For the GW generated to have any significant effect these objects need to be as massive as possible (white dwarfs, neutron stars, black holes) and need to be in close proximity to each other.
GW produced by such systems can in principle travel infinite distances if not interrupted by matter[33].
Thes types of GWs are called *astrophysical GWs* and have been detected by LIGO (2015), VIRGO (2017) and Hulse & Taylor (1974).
In the next semester we will talk about another mechanism which produces *primordial GWs*.

[32] This procedure is similar to the methods used in electrostatics and electrodynamics to find the electric potential and field. In fact we can often use the same mathematical tools, that we already used there (Gauss law etc.).

[33] As we have seen above the electromagnetic interaction in matter generates heat when a GW passes through it. This means, that the GW has to **lose energy** (and thus amplitude) when going through matter.



2.11.5. Remark on the degrees of freedom in GR

In general we call a field obeying to a wave equation of the form

$$\Box\phi + \cdots = 0$$

a *physical degree of freedom* or *propagating d.o.f.*
In quantum field theory each independent propagating d.o.f. generates a physical particle (e.g. Higgs field→Higgs boson, $A_{\mu\nu}$ →photon, ...).
In GR the field we are looking at is $g_{\mu\nu}$ with **10 independent d.o.f.** (if the gauge is fixed) but **only two are propagating d.o.f.**: The two polarization states of GWs (for a given k^μ). In attempts to quantize gravity[34] these d.o.f. generate new particles called *gravitons* which can exist even when there is vacuum everywhere ($g_{\mu\nu} = \eta_{\mu\nu}, T^{\alpha\beta} = 0$).
All the other d.o.f. are not propagating d.o.f. but just used to compute the trajectories of particles. Thus they **do not generate new particles**. This also becomes clear in the fact, that these d.o.f. vanish when there is vacuum everywhere.

Furthermore one can show, that since GWs/gravitons are described by the tensor $h_{\mu\nu}$ with **two Lorentz indices** they are **spin 2 fields/particles**.
Intuitively we can understand this if we recall, that spin N objects are **invariant under rotations by an angle of** $\frac{2\pi}{N}$. Hence gravitons are invariant $\frac{2\pi}{2} = \pi$ rotations. Indeed if we think of the solutions + and × we see, that our ellipses are in fact invariant under rotations by π.

[34] Unfortunately so far only the quantization of the perturbation $h_{\mu\nu}$ (and not of the background metric $g_{\mu\nu}$) and thus of the linear theory has been possible. This is due to the fact that a nonlinear theory like GR leads to **non-renormizable divergences** for which no adequate solution has been found yet.

2.12. Einstein-Hilbert action

In classical field theory we have already seen (for more information on this topic see [10])

$$\underbrace{S}_{\text{action}} = \int \underbrace{\mathrm{d}^4 x}_{\substack{\text{4D volume}\\ \text{element}}} \underbrace{\mathcal{L}}_{\text{Lagrangian}} (\underbrace{\Phi^a}_{\text{fields}}, \underbrace{\partial_\mu \Phi^a}_{\substack{\text{their}\\ \text{derivatives}}}, \underbrace{\dots}_{\substack{\text{possibly higher}\\ \text{derivatives}}})$$

In the following we will look at the case where the action **only contains first derivatives** of the fields.

A small variation of the field is then

$$\begin{cases} \phi^a \to \Phi^a + \delta\Phi^a \\ \partial_\mu \Phi^a \to \partial_\mu \Phi^a = \partial_\mu \Phi^a + \delta(\partial_\mu \Phi^a) \end{cases}$$

This means that a small variation of the action is

$$\begin{aligned} \delta S &= \int \mathrm{d}^4 x \left[\frac{\partial \mathcal{L}}{\partial \Phi^a} \delta\Phi^a + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi^a)} \delta(\partial_\mu \Phi^a) \right] \\ &= \int \mathrm{d}^4 x \left[\frac{\partial \mathcal{L}}{\partial \Phi^a} \delta\Phi^a + \underbrace{\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi^a)} \delta\Phi^a \right)}_{\substack{\text{only depends on } \mathcal{L} \text{on boundary,}\\ \text{fixed by assumptions}}} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi^a)} \right) \delta\Phi^a \right] \end{aligned}$$

where we treat Φ^a and $\partial_\mu \Phi^a$ as independent functions. We can chose the boundary conditions in such a way, that the second term vanishes and factor out $\delta\Phi^a$ which leaves us with

$$\delta S = \int \mathrm{d}^4 x \underbrace{\left[\frac{\partial \mathcal{L}}{\partial \Phi^a} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi^a)} \right) \right]}_{\text{total variation of } \mathcal{L} \text{ w.r.t. } \Phi^a} \delta\Phi^a$$

Classical mechanics: Among all possible trajectories the valid ones minimize the action i.e. $\delta S = 0$ which leads to the Euler Lagrange equation as discussed in 2.6.1:

$$\frac{\partial \mathcal{L}}{\partial \Phi^a} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi^a)} \right) = 0$$

Quantum mechanics: All trajectories are possible but weighted according to their action (path integral formalism).

We look at two classical examples:

1. **Scalar field in flat spacetime**:
$$\mathcal{L} = \frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi - V(\Phi) = \frac{1}{2}\eta^{\mu\nu}\partial_\mu\Phi\partial_\nu\Phi - V(\Phi)$$
Using the Euler Lagrange equation this gives us
$$\Box\Phi + \frac{\mathrm{d}V}{\mathrm{d}\Phi} = 0$$
with $\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu$. This is the well known *Klein-Gordon equation.*

2. **Scalar field in curved spacetime**: Since we know, that
$$S = \int \mathrm{d}V\,\mathcal{L} \overset{(2.4.3.1)}{=} \int \mathrm{d}^4x\sqrt{-g}\mathcal{L}$$
for curved spacetime we get
$$S = \underbrace{\int \mathrm{d}^4x\sqrt{-g}}_{\substack{\text{comoving volume}\\ \text{element}}}\left(\frac{1}{2}D_\mu\Phi D^\mu\Phi - V(\Phi)\right) = \underbrace{\int \mathrm{d}^4x\sqrt{-g}}_{\textit{Lagrangian density}}\ \underbrace{\left(\frac{1}{2}g^{\mu\nu}D_\mu\Phi D_\nu\Phi - V(\Phi)\right)}_{\textit{Lagrangian}}$$
where D_μ is the **covariant derivative**. Sometimes the word *Lagrangian* is used for the Lagrangian density as well, especially in flat spacetime where the two are equal.
In this case we get back the same Euler-Lagrange equation and thus the same Klein-Gordon equation
$$\Box\Phi + \frac{\mathrm{d}V}{\mathrm{d}\Phi} = 0$$
if we define $\Box = g^{\mu\nu}D_\mu D_\nu$. In this case the **derivative contains the impact of gravity** on the trajectories.

3. **GR in vacuum**: In this case one can show, that
$$\begin{aligned}\mathcal{L} = \sqrt{-g}\mathcal{R} \quad &\Rightarrow \quad S_{\text{EH}} = \int \mathrm{d}^4x\sqrt{-g}\mathcal{R}\\ &\Rightarrow \quad \delta S_{\text{EH}} = \int \mathrm{d}^4x\,\sqrt{-g}\left(\mathcal{R}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R}\right)\delta g^{\mu\nu}\end{aligned}$$
where S_{EH} is the *Einstein Hilbert action* which was derived by Hilbert in 1915. From here we directly see the Euler-Lagrange equation which is
$$G_{\mu\nu} = 0$$
just as we expect in vacuum.

Full theory From this we can finally construct a full (classical) theory of gravitation and other forces with the action

$$S = \int \mathrm{d}^4x \sqrt{-g} \left(\frac{\mathcal{R}}{16\pi G} + L_M \right) \tag{2.12.0.1}$$

where L_M is the Lagrangian of all matter fields (fermions, gauge fields, Higgs, ...). The variation w.r.t $g^{\mu\nu}$ then gives

$$\delta S = \int \mathrm{d}^4x \left(\sqrt{-g} \frac{G_{\mu\nu}}{16\pi G} + \frac{\delta(\sqrt{-g} L_M)}{\delta g^{\mu\nu}} \right) \delta g^{\mu\nu}$$

Comparing the term in brackets with the Einstein equation gives us a **fundamental definition of the energy momentum tensor**:

$$\boxed{T_{\mu\nu} \equiv -\frac{2}{\sqrt{-g}} \frac{\delta(-\sqrt{-g} L_M)}{\delta g^{\mu\nu}}} \tag{2.12.0.2}$$

Then the Euler Lagrange equation is just the Einstein equation.

Conclusion: We can express the full classical theory of gravity and other forces through the action. In this case we get

- $\frac{\delta S}{\delta \Phi^a}$ →equations of motion for fields
- $\frac{\delta S}{\delta g^{\mu\nu}}$ →equation of motion for the metric

3. Homogeneous Cosmology

In this section we will explore the evolution of the universe as a whole with time. Until $\sim$ 20 years ago this field of physics had been (almost) purely theoretical since no observations with sufficient sensitivity had been performed.

3.1. Newtonian Cosmology

Newton's law already implied some kind of dynamic of the universe. If we assume homogeneity this results in a uniform expansion or contraction of the universe due to the initial ratio of kinetic and potential energy which is contained in the universe.
It turns out, that the concept of homogeneity is only compatible with either a **static distribution** or with a **linear velocity field** of the form

$$\vec{v} = H\vec{r}$$

with a fixed parameter H which is called the *Hubble parameter*. Note, that this equation just holds for a **fixed time**.
This velocity can also be redefined in terms of the redshift z which gives the famous *Hubble law*:

$$\boxed{zc = Hr}$$

But what is the time evolution of expansion? We can calculate this from the Newtonian theory using Gauss' theorem if we assume either a **finite distribution of mass** or a **radius of interaction** where inside there are gravitational interactions and outside not. Then we can take Newton's law

$$\ddot{r}(t) = -\frac{GM(r(t))}{r^2(t)}$$

use the conservation of mass inside our Gaussian sphere ($\frac{\mathrm{d}M(r(t))}{\mathrm{d}t} = 0$) and integrate the equation (first multiplying it by $\dot{r}(t)$):

$$\frac{\dot{r}^2(t)}{2} = \frac{GM(r(t))}{r(t)} - \frac{k}{2}$$

where k is the constant of integration. We can then replace the mass $M(r(t))$ by the volume of the sphere times the homogeneous mass density $\rho_{\text{mass}}(t)$ and rearrange the equation to

$$\left(\frac{\dot{r}(t)}{r(t)}\right)^2 = \frac{8\pi G}{3}\rho_{\text{mass}}(t) - \frac{k}{r^2(t)} \tag{3.1.0.1}$$

The quantity $\dot{r}/r$ is called the *rate of expansion.* Since $M(r(t))$ is time-independent this means, that the mass density needs to evolve as $\rho_{\text{mass}}(t) \propto \frac{1}{r^3(t)}$.
The behaviour of $r(t)$ is dependent on the sign of k.

- $k > 0$: $r(t)$ can grow at early times but always decreases at some point.
- $k \leq 0$: $r(t)$ expands forever.

So there is a *critical value* for the homogeneous mass density $\rho_{\text{mass}}(t)$ which is

$$\rho_{\text{crit}} = \frac{3(\dot{r}(t)/r(t))^2}{8\pi G}$$

If $\rho_{\text{mass}}(t)$ is **bigger** than this value, the universe will **re-collapse**, otherwise it will keep expanding forever.

The limitations of the Newtonian predictions In our previous calculations we assumed that the universe is isotropic around us but did not check whether it is isotropic everywhere (and thus homogeneous). Earlier we saw though, that homogeneous expansion requires a linear law connecting distance and speed which is only fulfilled for $k = 0$. So it seems like only the solution $k = 0$ is compatible with the cosmological principle.
Another big problem arises from the fact, that we used the additivity of speed for the construction of our linear law. This cannot be applied at large distances though where $v \sim c$ which happens around the characteristic scale called the *Hubble radius* R_H:

$$R_H = cH^{-1}$$

at which the Newtonian expansion gives $v = HR_H = c$

3.2. The Friedmann-Lemaître-Robertson-Walker (FLRW) metric

3.2.1. Cosmological backgrounds and perturbation

To describe the universe as a whole we first need to think about the distribution of matter and energy in it. The cosmological principle then leads us to the following assumption:

> The exact description of the universe can be decomposed into two independent problems:
>
> - The *background problem*
> - The *inhomogeneity problem*

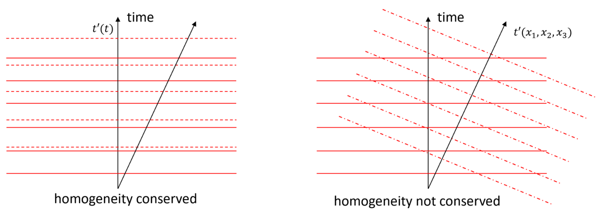
In the background problem one assumes, that in first approximation we can see the universe as a **smooth distribution of matter** (i.e. we can average over the small inhomogeneities like stars etc.). Then we can view all matter in the universe as a *cosmological fluid* and compute its dynamics.
The background problem consists of first order (or sometimes second order) perturbations, which can describe the large-scale structure or the CMB.
For the formation of small scale structures (like galaxies, stars etc.) this approach does not work since these are fully non-linear problems. They are usually solved through many-body Newtonian simulations or similar approaches.

3.2.2. Coordinate choice

The FLRW model is the most general solution in GR under the assumption of a background universe which is **homogeneous** and **isotropic**[1].
Under this assumption we can restrict the energy density ρ in such a way, that it is **only a function of time**, but not of space:

$$\rho(x^\mu) = \rho(x^0)$$

From this we immediately notice, that this means, that the notion of *homogeneous* and *isotropic* is **not a coordinate-independent property**. In particular this means, that a particular definition of time is preferred or more precisely a particular *time-slicing*.
A rescaling of $t \to t'(t)$ does not violate homogeneity but a mixing of time and space does.



The easiest way to construct such a system of coordinates in a homogeneous universe is to start from an initial homogeneous hypersurface and to assign to it a time coordinate t_1 and an arbitrary spatial direction. We can then use three arguments to map the whole spacetime:

1. If we place an observer at rest in our coordinate system $\frac{\mathrm{d}x^i}{\mathrm{d}x^0} = 0$ since due to homogeneity there is no *bulk velocity* due to some kind of force.

2. The time basis vector $\vec{e}_0 = \vec{e}_t$ must be **orthogonal** to the initial hypersurface in each point since otherwise there would be a preferred direction.

[1] Originally Einstein and later De Sitter considered **static** solutions of the Einstein equation but these were found to not describe the universe correctly.

3. If we let each observer be **free falling** and measure their proper time, we can define the coordinates at a different time in such a way, that at these coordinates the clocks of all observers show the same proper time t_2

The first argument allows us to map the whole space while not mixing space and time, the second one shows, that

$$e_0 \cdot \vec{e}_i = 0 \quad \Leftrightarrow \quad g_{0i} = \delta_0^\mu g_{\mu\nu} \delta_i^\nu = 0, \quad i = 1, 2, 3$$

The third argument can then be used to assign spacial coordinates to any point in time. This type of set of coordinates is called *comoving coordinates*.

Using proper time t we can then write down the explicit form of the metric (in non-natural units):

$$\boxed{\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + g_{ij}\mathrm{d}x^i\mathrm{d}x^j} \tag{3.2.2.1}$$

g_{ij} is furthermore restricted in the sense, that it must preserve homogeneity and isotropy. Its explicit form will be given by eq. (3.2.3.2).

Comments:

- As mentioned above a redefinition of time $t \to t'(t)$ preserves homogeneity although the time coordinate does not correspond to the proper time. This is sometimes useful to calculate specific things, although if we want to calculate the actual time an observer measures we always need to take proper time.
- An internal redefinition of spacial coordinates $x^i \to x^{i'}(x^i)$ will of course preserve homogeneity. In the following we will mostly stick to spherical coordinates.
- A general change of coordinates mixing space and time does not preserve homogeneity. In particular this means, that an observer which moves w.r.t. the comoving coordinates does not see the universe as homogeneous (he sees a blueshift in one and a redshift in the other direction). Therefore there is a **global** comoving frame.

3.2.3. Curvature

To specify g_{ij} we need to use the consequences which we obtain when assuming homogeneity. Let us consider two small comoving sticks S_A and S_B with arbitrary rotation and location. We assume that the ends of these sticks coincide with the locations of comoving observers. We call their position coordinates x_A^μ, x_B^μ and their vector coordinates $\mathrm{d}x_A^\mu, \mathrm{d}x_B^\mu$, i.e. the sticks stretch from $x_A^i - \mathrm{d}x_A^i/2$ to $x_A^i + \mathrm{d}x_A^i/2$ and $x_B^i - \mathrm{d}x_B^i/2$ to $x_B^i + \mathrm{d}x_B^i/2$ respectively. Hence their squared proper lengths just read:

$$\mathrm{d}s_A^2 = g_{ij}(x_A^\mu)\mathrm{d}x_A^i\mathrm{d}x_A^j \qquad \mathrm{d}s_B^2 = g_{ij}(x_B^\mu)\mathrm{d}x_B^i\mathrm{d}x_B^j$$

Then between two arbitrary times t_1 and t_2 the relative change in the proper length should vary by a constant factor independent of orientation and location:

$$\frac{g_{ij}(t_2, x_A^k)\mathrm{d}x_A^i\mathrm{d}x_A^j}{g_{ij}(t_1, x_A^k)\mathrm{d}x_A^i\mathrm{d}x_A^j} = \frac{g_{ij}(t_2, x_B^k)\mathrm{d}x_B^i\mathrm{d}x_B^j}{g_{ij}(t_1, x_B^k)\mathrm{d}x_B^i\mathrm{d}x_B^j} \equiv f(t_2, t_1)$$

which simplifies to

$$\frac{g_{ij}(t_2, x_A^k)}{g_{ij}(t_1, x_A^k)} = \frac{g_{ij}(t_2, x_B^k)}{g_{ij}(t_1, x_B^k)} = f(t_2, t_1)$$

so at any arbitrary times t_1, t_2 for fixed coordinates x^k we get the simple relation

$$g_{ij}(t_2, x^k) = f(t_2, t_1)g_{ij}(t_1, x^k)$$

From this we can conclude, that g_{ij} can only depend on time through a global scale factor $f(t, t_1)$ (where we interpret t_2 as the "running" time t and t_1 as a fixed reference time). Since f appears in factor of squared quantities and **needs to be positive** to guarantee, that proper lengths remain positive at all times it is preferable to refer to the square root $a(t) \equiv \sqrt{f(t, t_1)}$ which we call *scale factor*.
Hence we have proven, that the spatial part of our metric can only depend on time through a global scale factor $a(t)$:

$$g_{ij} = a(t)^2\tilde{g}_{ij}$$

where $\tilde{g}_{ij}$ is time independent.
The only thing left to do now is to find a form for $\tilde{g}_{ij}$.
This is fairly easy since there are only 3 types of metrics with constant curvature everywhere[2]:

1. Flat, euclidian space
2. 3-Sphere
3. 3-Hyperboloid

Conveniently choosing polar coordinates we can express all three cases through a single parameter $k \in \mathbb{R}$:

$$\mathrm{d}l^2 = \tilde{g}_{ij}\mathrm{d}x^i\mathrm{d}x^j = \frac{\mathrm{d}r^2}{1 - kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2(\theta)\mathrm{d}\varphi^2) \qquad (3.2.3.1)$$

where

1. $k = 0$ Euclidian universe (which is called *flat* universe)

[2]This is required to preserve homogeneity.

2. $k > 0$ The universe is a 3-sphere and thus has a finite volume. It is only defined in a finite range $0 \leq r < r_c$. This type of universe is called *closed* universe

3. $k < 0$ The universe is a 3-hyperboloid and negatively curved. This is called an *open* universe.

In the last two cases we can define the radius of curvature (of spacetime) as

$$R_c = \frac{1}{\sqrt{|k|}}$$

Hence we have found the most general solution of a universe, which obeys the cosmological principle, the *FLRW-metric*:

$$\boxed{\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a(t)^2\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2(\theta)\mathrm{d}\Omega)\right]} \tag{3.2.3.2}$$

If we look at the spatial part of this metric, we immediately see, that it just corresponds to the line element which we computed in (3.2.3.1) times the scale factor $a(t)$. Since we can express all lengths by integrating the line-element over a path in three-dimensional space we can also express a 3-dimensional, physical radius of curvature which is just

$$R_{c,\text{physical}}(t) = \frac{a(t)}{\sqrt{|k|}}$$

Comments:

- by rescaling $r \to r\sqrt{|k|}$ and $a(t) \to \frac{a(t)}{\sqrt{|k|}}$ the metric has the exact form given in (3.2.3.2) but we can restrict k to the three values $\pm 1, 0$

- As we will show in ex. 12, task 2 observers at rest with the *cosmological fluid* **stay at rest**, i.e. all trajectories parametrized by$x^i = x^i_M =$ const are solutions of the geodesic equation. Nevertheless the **proper distances** between points are still increasing or decreasing. Thus the universe expansion is not described by the proper motion of particles any more as it was the case in Newtonian cosmology but rather by the evolution of spacetime itself.

- As we have seen the FLRW metric describes a curved spacetime with **two different kinds of curvature**:
  - The spatial curvature $\pm\frac{a(t)}{\sqrt{k}}$ at each time
  - The spacetime curvature described by the evolution of $a(t)$

  While the first curvature is very intuitive the second one is a bit more subtle but we will see the effect of both in effects like the trajectories of photons.

- Using the scale factor we can define an actual radius of curvature. The *Hubble radius*

$$R_H(t) = c\frac{a(t)}{\dot{a}(t)}$$

- For the case $k = 0$ and $a =$ const we could just redefine the coordinate system with $r, \theta, \phi \to ar, \theta, \phi$ we just get back the Minkowski metric, which underlines the fact, that the curvature manifests itself as $k \neq 0$ for spatial curvature and $\dot{a} \neq 0$ for the remaining spacetime curvature.

- Sometimes equations become simpler when we redefine the time to $\mathrm{d}t = a(t)\mathrm{d}\tau$, so the metric reads:

$$\mathrm{d}s^2 = a^2(\tau)\left(-c^2\mathrm{d}\tau^2 + \left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2(\theta)\mathrm{d}\varphi^2)\right]\right)$$

  This metric exhibits conformal symmetry, thus τ is called *conformal time* in opposition to the proper time t which is sometimes also called the *cosmological time*.

3.3. Light propagation in the FLRW universe

3.3.1. Photon geodesics

We know, that in GR photons move along geodesics with the speed of light. Hence in an infinitesimal interval $\mathrm{d}t$ the photons move by $\mathrm{d}l^2 = c^2\mathrm{d}t^2$. Hence we can integrate $\mathrm{d}l = \pm c\mathrm{d}t$ to get macroscopic distances.
If we consider a photon along a straight geodesic (a free photon) and we assume, that we are a comoving observer, we can choose the origin of our spherical FLRW universe to coincide with our position[3]. Since we require isotropy the photon then still needs to travel along a straight line, so in $\mathrm{d}r$-direction, which means, that we can express the distance travelled by

$$\int_{r_e}^{r} -\frac{\mathrm{d}r}{\sqrt{1-kr^2}} = \int_{t_e}^{t} \frac{c\ \mathrm{d}t}{a(t)}$$

The solutions to this equation are indeed a solution of the geodesic equation. This means that if we put ourselves at $r = 0$ in the FLRW universe we can further simplify the equation above. Additionally the observer sees the photon at a time t_0 which can be implicitly deduced from r_e and t_e through the important equation

$$\boxed{\int_{r_e}^{0} -\frac{\mathrm{d}r}{\sqrt{1-kr^2}} = \int_{t_e}^{t_0} \frac{c\ \mathrm{d}t}{a(t)}} \tag{3.3.1.1}$$

Comments:

[3] This is just done to simplify calculations, it **does not** put us in a special position in our universe (apart from the fact, that we are at rest with the cosmological fluid).

- The ensemble of points (t_e, r_e, θ, ϕ) for which (3.3.1.1) holds define our **past-light cone**.
- We can already see the physical interpretation of the conformal time here since the right side corresponds to $\tau_r - \tau_e$

3.3.2. Redefining the redshift

Again we place ourselves as a comoving observer at the origin of coordinates. We pretend to observe a galaxy located at (r_e, θ_e, ϕ_e) emitting light at a given frequency $\nu_e = c/\lambda_e$. with a period $\mathrm{d}t_e \equiv 1/\nu_e$. When we compute their trajectory with (3.3.1.1) we receive it with a frequency $\nu_r = c/\lambda_r = 1/\mathrm{d}t_r$ so

$$\int_{r_e}^{0} -\frac{\mathrm{d}r}{\sqrt{1-kr^2}} = \int_{t_e}^{t_r} \frac{\mathrm{d}t}{a(t)} = \int_{t_e+\mathrm{d}t_e}^{t_r+\mathrm{d}t_r} \frac{\mathrm{d}t}{a(t)}$$

Rearranging the second equality gives

$$\int_{t_e}^{t_e+\mathrm{d}t_e} \frac{\mathrm{d}t}{a(t)} = \int_{t_r}^{t_r+\mathrm{d}t_r} \frac{\mathrm{d}t}{a(t)}$$

so we get in very good approximation:

$$\frac{\mathrm{d}t_e}{a(t_e)} = \frac{\mathrm{d}t_r}{a(t_r)}$$

Remapping this to the emission and reception of wavelengths gives

$$\frac{\lambda_r}{\lambda_e} = \frac{\mathrm{d}t_r}{\mathrm{d}t_e} = \frac{a(t_r)}{a(t_e)}$$

This is in fact the answer which we would intuitively expect since we already know, that wavelengths become longer when the scale factor increases. Hence in the FLRW universe the redshift is given by

$$z = \frac{\Delta\lambda}{\lambda} = \frac{\lambda_r - \lambda_e}{\lambda_e} = \frac{a(t_r)}{a(t_e)} - 1$$

In other words if we observe an object at our time t_0 the observed redshift is

$$\boxed{z = \frac{a(t_0)}{a(t_e)} - 1} \tag{3.3.2.1}$$

Comments:

- Unlike in Newtonian cosmology where $z = v/c$ here the redshift is **not restricted to values** < 1 since the scale factor a can be arbitrarily large without violating the fundamental principles of GR

- Since there is a global comoving frame this equation only holds exactly if the observer is exactly at rest w.r.t. the comoving frame. In reality this is never the case, so we need to take this into account as an additional contribution coming from the Doppler-formula from SR (see 1.6.8). These velocities rarely exceed values of $c/1000$ so the correction is $\mathcal{O}(10^{-3})$ and can be neglected for large distances. It is however dominant for small distances.

3.3.3. Redefining the Hubble parameter

The Hubble parameter in the FLRW universe can be derived by taking the line-element and looking at the Newtonian limit, which should read

$$z = \frac{v}{c} = \frac{HL}{c}$$

where L denotes to the *physical distance* to the object. To show this let us assume, that we receive a light ray from a nearby point at t_0 so it was emitted at $t_0 - \mathrm{d}t$. In the limit of small $\mathrm{d}t$ the equation of propagation of light then gives

$$L \approx \frac{a(t_0)|\mathrm{d}r|}{\sqrt{1-kr^2}} = c\ \mathrm{d}t$$

while the redshift of the galaxy is

$$z = \frac{a(t_0)}{a(t_0 - \mathrm{d}t)} - 1 \approx \frac{a(t_0)}{a(t_0) - \dot{a}(t_0)\mathrm{d}t} = \frac{1}{1 - \frac{\dot{a}(t_0)}{a(t_0)}\mathrm{d}t} - 1 \approx \frac{\dot{a}(t_0)}{a(t_0)}\mathrm{d}t$$

and by combining the two equations above this gives

$$z \approx \frac{\dot{a}(t_0)}{a(t_0)} \cdot \frac{L}{c}$$

Hence $\frac{\dot{a}(t_0)}{a(t_0)}$ plays the role of the Hubble parameter at any given time. This means, that

$$\boxed{H(t) = \frac{\dot{a}(t)}{a(t)}} \tag{3.3.3.1}$$

Comments:

- We will often use the **current value** of the Hubble constant which we will call H_0.
- While at small distances we recover the linear Hubble law, which links distance and velocity (or redshift) but this approximation is only valid for small distances. At larger distances we have to find a **new definition of distance**, which will be discussed in the following.

3.3.4. Distances

Comoving distance

As above we assume, that we are at the origin of a FLRW universe with spherical coordinates at time t_0 and observe an object at $(t_e, r_e, \theta_e, \phi_e)$.
The easiest way to define a distance is by assuming, that the object we observe is comoving so at the time when we observe it it should be at $(t_0, r_e, \theta_e, \phi_e)$. We can then compute the distance on a constant time-hypersurface with $t = t_0$ and just integrate the length-element:

$$d = \int_0^{r_e} \mathrm{d}l = a(t_0) \frac{\mathrm{d}r}{\sqrt{1 - kr^2}}$$

This is a unique definition of distance up to the normalization factor $a(t_0)$ which we can **choose to be** 1. In this case the distance d coincides with the *comoving distance* χ:

$$\chi(r_e) \equiv \int_0^{r_e} \frac{\mathrm{d}r}{\sqrt{1 - kr^2}} \tag{3.3.4.1}$$

which can be explicitly integrated for the three possible cases for k:

$$\chi(r) = \begin{cases} \sin^{-1}(r) & \text{if} \quad k = 1 \\ r & \text{if} \quad k = 0 \\ \sinh^{-1}(r) & \text{if} \quad k = -1 \end{cases}$$

Comments

- We can also find an explicit form for χ if we do not restrict k to the values $\pm 1, 0$ (see ex. 12, task 3).
- It is useful to define

  $$f_k(x) = \begin{cases} \sin(x) & \text{if} \quad k = 1 \\ x & \text{if} \quad k = 0 \\ \sinh(x) & \text{if} \quad k = -1 \end{cases}$$

  so that $r = f_k(\chi)$.
- From (3.3.1.1) follows, that $\chi(r)$ is equal to the **conformal age** of the object:

  $$\chi(r) = \int_{t_e}^{t_0} \frac{c \ \mathrm{d}t}{a(t)} = c(\tau_0 - \tau_e)$$

  which means, that in natural units conformal time and comoving distance are **the same**.

- Although the comoving distance is a very nice definition of distance since it is independent of time it is **not something we can measure**. This motivates the definition of other, measurable distances.

The three things which enable us to measure distances, since we cannot determine them directly are the **redshift**, the **angular diameter** of objects of known size and the **luminosity** of standard candles. In the next two sections we will find relations between the redshift and the other two, so effectively connections between all three types of measurements.

3.3.5. Angular diameter distance

For objects of known physical size $\mathrm{d}l$ we can measure their angular diameter $\mathrm{d}\theta$ and calculate the distance through $\mathrm{d}l = d \times \mathrm{d}\theta$ which we call the *angular diameter distance*:

$$\mathrm{d}_A \equiv \frac{\mathrm{d}l}{\mathrm{d}\theta}$$

In Euclidian space with linear expansion we can easily find a relation between angular diameter distance and redshift:

$$d_A = d = \frac{c}{H_0} z$$

For the FLRW universe we need to additionally take into account the bending of light rays which implies, that the physical size $\mathrm{d}l$ (evaluated at t_e) of an object orthogonal to the line of sight is

$$\mathrm{d}l = a(t_e) r_e \, \mathrm{d}\theta$$

where t_e is the time of emission and r_e the comoving coordinate of the emitter.
Using (3.3.2.1) this give

$$d_A = a(t_e) r_e = a(t_0) \frac{r_e}{1 + z_e}$$

We can then replace r_e with (3.3.4.1) which yields:

$$d_A = \frac{a(t_0)}{1 + z_e} f_k \left(\int_{t_e}^{t_0} \frac{c \, \mathrm{d}t}{a(t)} \right)$$

which if we replace $\mathrm{d}t$ by $\mathrm{d}z$ gives the important *angular diameter-redshift relation*[4]

$$d_A = \frac{a(t_0)}{1 + z_e} f_k \left(\int_0^{z_e} \frac{c \, \mathrm{d}z}{a(t_0) H(z)} \right) \tag{3.3.5.1}$$

[4] This is a handy relation to test the validity of cosmological models if we already know the physical size of objects. Luckily there are such objects called *standard rulers*

3.3.6. Luminosity distance

If we know the absolute luminosity of objects we can infer their distance from their apparent luminosity. In a Euclidian universe the relation between apparent and absolute luminosity ist just $l = \frac{L}{2\pi d^2}$. Although geometry is not Euclidian in cosmology we will stick to that definition:

$$d_L \equiv \sqrt{\frac{L}{2\pi l}} \tag{3.3.6.1}$$

the effect of curvature manifests itself through the different definition of the relation between apparent and absolute luminosity in cosmology:

$$l = \frac{L}{4\pi a^2(t_0) r_e^2 (1+z_e)^2}$$

so

$$\boxed{d_L = a(t_0) r_e (1+z_e)} \tag{3.3.6.2}$$

We can again rewrite this equation in terms of the redshift which gives the *luminosity distance-redshift relation*:

$$\boxed{d_L = a(t_0)(1+z_e) f_k \left(\int_0^{z_e} \frac{c\,\mathrm{d}z}{a(t_0)H(z)} \right)} \tag{3.3.6.3}$$

Comments

- We can easily see the relation between angular distance and luminosity distance:

$$d_L = a(t_0) r_e (1+z_e) = a(t_e) r_e (1+z_e)^2 = (1+z_e)^2 d_A$$

- In the limit $z \to 0$ the three definitions of distance are equivalent and reduce to the usual definition $d = \frac{c}{H_0} z$, so at small redshifts we gain no additional information from the measurement of any of these quantities (apart from measuring H_0)

3.4. The Friedmann Law

The goal of this section is to link the curvature k and the scale factor $a(t)$ to the source of curvature: The matter density.

3.4.1. Einstein's equation

From 2.8 we already know the relation between curvature and matter in any metric:

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}$$

The Einstein tensor $G_{\mu\nu}$ can be directly computed from the metric using its Christoffel symbols. This yields the (unsurprisingly simple) result, that the Einstein tensor is **diagonal** and $G^1_1 = G^2_2 = G^3_3$. This is a direct consequence of the invariance under rotations which is a requirement of isotropy (see ex. 3, task 1).
As we have already seen the stress energy tensor of a perfect fluid takes exactly this form. In particular for such a fluid we can use (1.7.2.2):

$$\boxed{T^{\alpha\beta} = (\rho_{\mathrm{MCRF}} + \mathfrak{p}_{\mathrm{MCRF}})U^\alpha U^\beta + \mathfrak{p}_{\mathrm{MCRF}}\eta^{\alpha\beta}}$$

where we know, that in an MCRF (comoving frame) $U^\mu = (U^0, 0, 0, 0)$. Furthermore in the case of the FLRW $g_{00} = -1$ so

$$-1 = \vec{U} \cdot \vec{U} = U^0 g_{00} U^0 \quad \Rightarrow \quad U^0 = 1, \quad U_\nu = -\delta^0_\nu$$

which allows us to express the SET explicitly:

$$T^\mu{}_\nu = \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & \mathfrak{p} & 0 & 0 \\ 0 & 0 & \mathfrak{p} & 0 \\ 0 & 0 & 0 & \mathfrak{p} \end{pmatrix}$$

This underlines the fact, that the FLRW model is only compatible with a **perfect cosmological fluid** inhabiting the universe.

The first component of the Einstein tensor reads (not derived here)

$$G_{00} = 3\left[\frac{k}{a^2} + \left(\frac{\dot{a}}{a}\right)^2\right]$$

This is an interesting expression to study. It is in fact the sum of the squared spatial radius $R_c(t) = \pm\frac{a}{\sqrt{|k|}}$ and the inverse squared Hubble radius $R_H(t) = \frac{\dot{a}}{a}$. This emphasizes what we already learned in 3.2.3 about the different types of curvature abundant in the universe[5].
Using the Einstein equation (with one index up and one down):

$$-G^0{}_0 = -8\pi G T^0{}_0$$

we can rewrite the above:

$$3\left[\frac{k}{a^2} + \left(\frac{\dot{a}}{a}\right)^2\right] = 8\pi G\rho$$

which when rearranged gives the famous *Friedmann equation*

$$\boxed{H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3} - \frac{k}{a^2}} \tag{3.4.1.1}$$

[5]The additional factor 3 appears due to the three spatial dimensions.

Comments:

- We can rewrite ρ after thinking about the implications of the energy content contained in ρ. We will do this in the next section.
- If we assume that all mass in the universe is non-relativistic, so $p^2 \ll m^2$ the Friedmann law looks exactly like the Newtonian expansion law (3.1.0.1) with $a(t)$ having the role of $r(t)$.
  Of course they are very different though since
  1. In the Newtonian model the expansion leads to velocities $> c$, in the FLRW model not
  2. The Newtonian model forbids $k \neq 0$ since it violates isotropy, in the FLRW model it is allowed.

3.4.2. Energy conservation

From the Bianchi identities the Einstein equation implies

$$G^{\mu}{}_{\nu;\mu} = T^{\mu}{}_{\nu;\mu} = 0$$

For the first component this is just an equation of energy conservation.
By explicitly computing the Christoffel symbols and using (2.5.2.2) and (2.5.3.1) one finds, that this just yields:

$$\boxed{\dot{\rho} = -3\frac{\dot{a}}{a}(\mathfrak{p} + \rho)} \tag{3.4.2.1}$$

Hence the *dilution* of energy as the universe expands depends on the pressure[6]. In homogeneous cosmology we usually look at the two limiting cases:

- **Non-relativistic matter:** In case of slow moving matter we neglect any kinetic energies which implies $\mathfrak{p} = 0$ (a comoving box enclosing the fluid would not feel any pressure), hence[7]

$$\boxed{\dot{\rho} = -3\frac{\dot{a}}{a}\rho \quad \Rightarrow \quad \rho \propto a^{-3}}$$

- **Ultra-relativistic matter:** In this case we know from statistical mechanics (see [1] or [5]), that $\mathfrak{p} = \frac{\rho}{3}$. Hence

$$\dot{\rho} = -3\frac{\dot{a}}{a}\left(1 + \frac{1}{3}\right) = -4\frac{\dot{a}}{a}\rho \quad \Rightarrow \quad \rho \propto a^{-4}$$

  This can be understood if we remind ourselves, that the *energy density* of photons is E/V. The dilution is $V \propto a^3$ and $E \propto a^{-1}$. Hence $\rho \propto a^{-4}$.

[6] More precisely on the *equation of state* $\mathfrak{p}(\rho)$.
[7] This result is obvious since $V \propto a^3 \Rightarrow \rho \propto a^{-3}$

3.4.3. Cosmological constant

The last thing to discuss in the Friedmann equation is the cosmological constant. This is just an integration constant which can be added on the left side of the Einstein equation without violating any principle:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}$$

The constant Λ **cannot depend on time or space** and is called the *cosmological constant*. If we move it from the left side of the equation to the right we see, that the cosmological constant is equivalent to a homogeneous fluid with the SET

$$\tilde{T}^{\mu}{}_{\nu} = -\frac{\Lambda}{8\pi G}\delta^{\nu}_{\mu}$$

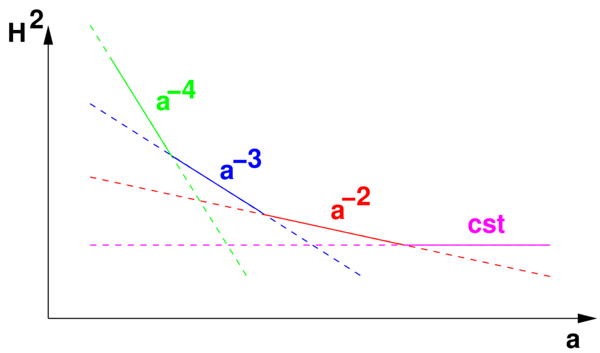
In opposition to matter or radiation this fluid has $\rho = -p$ which implies, that $\dot{\rho} = 0$ so the energy **does not dilute**.
While this had no good explanation in Einstein's times the **vacuum energy of QFT** is a good candidate for this this predicts a way larger cosmological constant then measured though, so it remains a bit of a mystery.

3.4.4. Possible scenarios for the history of the universe

If we write the Friedmann law with all possible contributions discussed so far (sorted by the speed of dilution) we get

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho_{\mathrm{R}} + \frac{8\pi G}{3}\rho_{\mathrm{M}} - \frac{k}{a^2} + \frac{\Lambda}{3}$$

where ρ_R denotes to the radiation- and ρ_M to the matter density. since they evolve w.r.t. a with a^{-4}, a^{-3}, a^{-2} and a^0 if the scale factor keeps growing and assuming that all four terms contribute we get a distinct **order of domination** of each term, which is best illustrated in the graph below:



If we furthermore assume that one of the parameters strongly dominates over the others we can compute the behaviour of the scale factor, Hubble parameter and Hubble radius in each of the stages:

Dominating contribution	$\left(\frac{\dot{a}}{a}\right)^2 \propto$	$a(t) \propto$	$H(t)$	$R_H(t)$	Type of curvature
Radiation	a^{-4}	$t^{1/2}$	$\frac{1}{2t}$	$2t$	decelerated expansion
Matter	a^{-3}	$t^{2/3}$	$\frac{2}{3t}$	$\frac{3}{2}t$	decelerated expansion
Negative curvature	a^{-2}	t	$\frac{1}{t}$	t	linear expansion
Positive Curvature	a^{-2}	t	$\frac{1}{t}$	t	recollapse
Cosmological constant	$a^0 = \text{const}$	$\exp(Ht)$	$\frac{1}{R_H} = \sqrt{\Lambda/3}$		acclerated expansion

3.4.5. Cosmological parameters

Since we can only measure the quantities in the Friedmann equation today it is more useful to rewrite it in terms of H_0 instead of an arbitrary $H(t)$. If we then divide by H_0^2 we get

$$1 = \frac{8\pi G}{3H_0^2}(\rho_{\mathrm{R},0} + \rho_{\mathrm{M},0}) - \frac{k}{a_0^2 H_0^2} + \frac{\Lambda}{3H_0^2} \tag{3.4.5.1}$$

where 0 indicates, that we evaluate the quantities today. From this we can compute the whole evolution of the universe.

A nice side-effect of this formulation is, that by construction the four terms in this equation add up to 1 and hence are the relative contribution to the present universe expansion. Hence it is useful to give each term a symbol:

$$\begin{aligned}
\Omega_R &= \frac{8\pi G}{3H_0^2}\rho_{\mathrm{R},0} \\
\Omega_M &= \frac{8\pi G}{3H_0^2}\rho_{\mathrm{M},0} \\
\Omega_k &= -\frac{k}{a_0^2 H_0^2} \\
\Omega_\Lambda &= \frac{\Lambda}{3H_0^2}
\end{aligned}$$

which simplifies the *matter budget equation* to

$$\Omega_{\mathrm{R}} + \Omega_{\mathrm{M}} + \Omega_k + \Omega_\Lambda = 1$$

An important consequence of this is, that the universe is flat if

$$\Omega_0 \equiv \Omega_{\mathrm{R}} + \Omega_{\mathrm{M}} + \Omega_\Lambda = 1$$

which means, that at any time the total density of matter, radiation and Λ must be equal to the *critical density*

$$\rho_c(t) = \frac{3H(t)^2}{8\pi G}$$

A. Proofs

A.1. Bianchi identities

There are two Bianchi identities[1] and two contractions of the second Bianchi identity, which we call *contracted-* and *twice contracted Bianchi identity*. While the first one is very straight forward to prove, the second one is a bit more subtle. We will present a way of proving both:

First Bianchi identity:
we want to prove, that

$$R^{\alpha}{}_{[\beta\mu\nu]} := R^{\alpha}{}_{\beta\mu\nu} + R^{\alpha}{}_{\mu\nu\beta} + R^{\alpha}{}_{\nu\beta\mu} = 0$$

The proof becomes trivial if we simply write out the representation of the Riemann-tensor by the Christoffel symbols:

$$\begin{aligned}
R^{\alpha}{}_{\beta\mu\nu} &= \Gamma^{\alpha}_{\beta\nu,\mu} - \Gamma^{\alpha}_{\beta\mu,\nu} + \Gamma^{\alpha}_{\sigma\mu}\Gamma^{\sigma}_{\nu\beta} - \Gamma^{\alpha}_{\sigma\nu}\Gamma^{\sigma}_{\mu\beta} \\
R^{\alpha}{}_{\mu\nu\beta} &= \Gamma^{\alpha}_{\mu\beta,\nu} - \Gamma^{\alpha}_{\mu\nu,\beta} + \Gamma^{\alpha}_{\sigma\nu}\Gamma^{\sigma}_{\beta\mu} - \Gamma^{\alpha}_{\sigma\beta}\Gamma^{\sigma}_{\nu\mu} \\
R^{\alpha}{}_{\nu\beta\mu} &= \Gamma^{\alpha}_{\nu\mu,\beta} - \Gamma^{\alpha}_{\nu\beta,\mu} + \Gamma^{\alpha}_{\sigma\beta}\Gamma^{\sigma}_{\mu\nu} - \Gamma^{\alpha}_{\sigma\mu}\Gamma^{\sigma}_{\beta\nu}
\end{aligned}$$

which altogether gives us for the sum:

$$\begin{aligned}
R^{\alpha}{}_{[\beta\mu\nu]} &= (\Gamma^{\alpha}_{\beta\nu,\mu} - \Gamma^{\alpha}_{\nu\beta,\mu}) + (\Gamma^{\alpha}_{\mu\beta,\nu} - \Gamma^{\alpha}_{\beta\mu,\nu}) + (\Gamma^{\alpha}_{\nu\mu,\beta} - \Gamma^{\alpha}_{\mu\nu,\beta}) \\
&+ (\Gamma^{\alpha}_{\sigma\mu}\Gamma^{\sigma}_{\nu\beta} - \Gamma^{\alpha}_{\sigma\mu}\Gamma^{\sigma}_{\beta\nu}) + (\Gamma^{\alpha}_{\sigma\nu}\Gamma^{\sigma}_{\beta\mu} - \Gamma^{\alpha}_{\sigma\nu}\Gamma^{\sigma}_{\mu\beta}) + (\Gamma^{\alpha}_{\sigma\beta}\Gamma^{\sigma}_{\mu\nu} - \Gamma^{\alpha}_{\sigma\beta}\Gamma^{\sigma}_{\nu\mu}) \\
&= 0 \qquad \square
\end{aligned}$$

where in the last step we have used the symmetry of the Christoffel symbols.

Second Bianchi identity:
We want to show, that

$$R_{\alpha\beta[\mu\nu;\lambda]} := R_{\alpha\beta\mu\nu;\lambda} + R_{\alpha\beta\lambda\mu;\nu} + R_{\alpha\beta\nu\lambda;\mu} = 0 \tag{A.1.0.1}$$

This identity can be shown using the first Bianchi identity but this is a very tedious task, so we will show it using some considerations of the Riemann-tensor in a LPIF. For this we first switch to a LPIF, so that

$$g_{\alpha\beta} = \eta_{\alpha\beta} + \mathcal{O}(\delta(x^{\alpha})^2)$$

[1] although the first one was discovered by Ricci.

In this frame all Christoffel symbols vanish, so

$$0 \overset{(2.5.5.1)}{=} g_{\alpha\beta;\gamma} = g_{\alpha\beta,\gamma} - g_{\alpha\sigma}\Gamma^{\sigma}_{\beta\gamma} - g_{\sigma\beta}\Gamma^{\sigma}_{\alpha\gamma} = g_{\alpha\beta,\gamma}$$

$$\Rightarrow g_{\alpha\beta;\gamma} = g_{\alpha\beta,\gamma}$$

Note, that however $g_{\alpha\beta,\gamma\epsilon} \neq 0$ in general. In particular this also means, that in the LPIF, the covariant derivative of the Riemann-tensor is equal to its partial derivative:

$$\begin{aligned} R_{\alpha\beta\mu\nu;\epsilon} &= R_{\alpha\beta\mu\nu,\epsilon} + \text{terms involving } \Gamma \\ &= R_{\alpha\beta\mu\nu,\epsilon} \end{aligned}$$

This means, that the only thing left to do is to represent the Riemann-tensor in terms of the metric and do the permutations:

$$\begin{aligned} R_{\alpha\beta\mu\nu} &= g_{\alpha\sigma}\left(\Gamma^{\sigma}_{\beta\nu,\mu} - \Gamma^{\sigma}_{\beta\mu,\nu} + \Gamma^{\sigma}_{\gamma\mu}\Gamma^{\gamma}_{\nu\beta} - \Gamma^{\sigma}_{\gamma\nu}\Gamma^{\gamma}_{\mu\beta}\right) \\ \Rightarrow R_{\alpha\beta\mu\nu,\epsilon} &= \underbrace{g_{\alpha\sigma,\epsilon}}_{=0}(\dots) + g_{\alpha\sigma}(\Gamma^{\sigma}_{\beta\nu,\mu\epsilon} - \Gamma^{\sigma}_{\beta\mu,\nu\epsilon}) \\ &= \frac{1}{2}g_{\alpha\sigma}\left(g^{\sigma\xi}(g_{\beta\xi,\nu} + g_{\xi\nu,\beta} - g_{\beta\nu,\xi})_{,\mu\epsilon} - g^{\sigma\xi}(g_{\beta\xi,\mu} + g_{\xi\mu,\beta} - g_{\beta\mu,\xi})_{,\nu\epsilon}\right) \end{aligned}$$

This means, that we have expressed the Riemann-tensor solely through *partial* derivatives of the metric, which commute. We can thus write

$$\begin{aligned} R_{\alpha\beta[\mu\nu;\epsilon]} = R_{\alpha\beta[\mu\nu,\epsilon]} &= g_{\alpha\sigma}\left[\underbrace{(\Gamma^{\sigma}_{\beta\nu,\mu\epsilon} - \Gamma^{\sigma}_{\beta\nu,\epsilon\mu})}_{=0} + \underbrace{(\Gamma^{\sigma}_{\beta\mu,\epsilon\nu} - \Gamma^{\sigma}_{\beta\mu,\nu\epsilon})}_{=0} + \underbrace{(\Gamma^{\sigma}_{\beta\epsilon,\nu\mu} - \Gamma^{\sigma}_{\beta\epsilon,\mu\nu})}_{=0}\right] \\ &= 0 \qquad \square \end{aligned}$$

Contracted Bianchi identity

We want to show, that

$$\mathcal{R}_{\beta\nu;\epsilon} + R^{\alpha}{}_{\beta\nu\epsilon;\alpha} - \mathcal{R}_{\beta\epsilon;\nu} = 0$$

This is simply done by contracting the second Bianchi identity with the metric $g^{\alpha\mu}$ and use, that we can pull the metric through the derivative since $g^{\alpha\beta}{}_{;\epsilon} = 0$:

$$\begin{aligned} 0 &= \underbrace{g^{\alpha\mu}R_{\alpha\beta\mu\nu;\epsilon}}_{=\mathcal{R}_{\beta\nu;\epsilon}} + \underbrace{g^{\alpha\mu}R_{\alpha\beta\epsilon\mu;\nu}}_{=-g^{\alpha\mu}R_{\alpha\beta\mu\epsilon;\nu}=-\mathcal{R}_{\beta\epsilon;\nu}} + \underbrace{g_{\alpha\mu}R_{\alpha\beta\nu\epsilon;\mu}}_{=R^{\mu}{}_{\beta\nu\epsilon;\mu}=R^{\alpha}{}_{\beta\nu\epsilon;\alpha}} \\ &= \mathcal{R}_{\beta\nu,\epsilon} + R^{\alpha}{}_{\beta\nu\epsilon;\alpha} - \mathcal{R}_{\beta\epsilon;\nu} \qquad \square \end{aligned}$$

Twice contracted Bianchi identity We do another contraction of the indices in the contracted Bianchi identity and show, that

$$2 \cdot \mathcal{R}^{\alpha}{}_{\nu;\alpha} - \mathcal{R}_{;\nu} = 0$$

We multiply the contracted Bianchi identity with $g^{\beta\epsilon}$ and get:

$$\begin{aligned} 0 &= g^{\beta\epsilon}\mathcal{R}_{\beta\nu;\epsilon} + g^{\beta\epsilon}R^{\alpha}{}_{\beta\nu\epsilon;\alpha} - g^{\beta\epsilon}\mathcal{R}_{\beta\epsilon;\nu} \\ &= g^{\beta\epsilon}\mathcal{R}_{\beta\nu;\epsilon} + g^{\beta\epsilon}g^{\alpha\sigma}R_{\sigma\beta\nu\epsilon;\alpha} - \mathcal{R}_{;\nu} \\ &= g^{\beta\epsilon}\mathcal{R}_{\beta\nu;\epsilon} + g^{\alpha\sigma}\mathcal{R}_{\sigma\nu;\alpha} - \mathcal{R}_{;\nu} \\ &= 2 \cdot g^{\beta\epsilon}\mathcal{R}_{\beta\nu;\epsilon} - \mathcal{R}_{;\nu} \\ &= 2 \cdot \mathcal{R}^{\alpha}{}_{\nu;\alpha} - \mathcal{R}_{;\nu} \qquad \square \end{aligned}$$

A.2. Number of independent components of the Riemann-tensor

We want to prove, that the Riemann tensor has only 20 independent components. To do this we use the symmetries

$$R_{\alpha\beta\,\mu\nu} = -R_{\underbrace{\beta\alpha}\,\mu\nu} = -R_{\alpha\beta\,\underbrace{\nu\mu}} = +R_{\underbrace{\underline{\mu\nu}\,\underline{\alpha\beta}}}$$

and the first Bianchi-identity (see A.1). We look at the distinctive components:

- All components, with the first two or last two indices being the same vanish, since they are their own negative:

$$R_{\underbrace{\alpha\alpha}\,\mu\nu} = -R_{\underbrace{\alpha\alpha}\,\mu\nu} = 0$$

  So all components of the form $R_{\alpha\alpha\alpha\alpha}$, $R_{\alpha\alpha\alpha\beta}$, $R_{\alpha\alpha\mu\nu}$ vanish.

- There are 6 independent components of the form $R_{\alpha\beta\alpha\beta}$, which are the number of possibilities to arrange 2 indices which can take 4 values ($\binom{4}{2} = 6$)

- There are 12 independent components of the form $R_{\alpha\beta\alpha\mu}$, which are

R_{0102}	R_{0103}	R_{0203}
R_{1213}	R_{1013}	R_{1012}
R_{2123}	R_{2021}	R_{2023}
R_{3132}	R_{3031}	R_{3032}

- Lastly we have 3 independent components if all indices are different:

$$R_{0123};\ R_{0312};\ R_{0231}$$

  But because of the first Bianchi-identity $R_{\alpha[\beta\mu\nu]} = 0$ these three degrees of freedom reduce to two.

Therefore we have $6 + 12 + 2 = 20$ independent components. $\square$

B. The Schwarzschild solution

B.1. Trajectories of test particles

Massive particle

While there is no general, analytical solution to the differential equation (2.9.2.1) we can look at the behaviour of

$$\tilde{V}^2(r) \equiv \left(1 - \frac{2M}{r}\right)\left(1 + \frac{\tilde{L}^2}{r^2}\right)$$

We differentiate with respect to r and find

$$\frac{\mathrm{d}\tilde{V}^2}{\mathrm{d}r} = \frac{\tilde{L}^2(6M - 2r) + 2Mr^2}{r^4}$$

and find the extrema at

$$r = \frac{\tilde{L}^2}{2M} \pm \sqrt{\frac{L^4}{4M^2} - 3\tilde{L}^2} \tag{B.1.0.1}$$

which means, that for $L^2 < 12M^2$ we get **no** solution and for $L^2 > 12M^2$ we get **two solutions**. From the second derivative we see, that the extreme with the $-$ sign is a **maximum** and the other one a **minimum**.
We can then plug r_{min} and r_{max} in the original equation and get:

$$V_{\text{max}}^2 = 2\,L^2\left(L^2 - 4\,M^2 + \sqrt{\frac{L^2\,(L^2 - 12\,M^2)}{M^2}}M\right)^2 \cdot \left(L^2 + \sqrt{\frac{L^2\,(L^2 - 12\,M^2)}{M^2}}M\right)^{-3}$$
$$V_{\text{min}}^2 = 2\,L^2\left(L^2 - 4\,M^2 - \sqrt{\frac{L^2\,(L^2 - 12\,M^2)}{M^2}}M\right)^2 \cdot \left(L^2 - \sqrt{\frac{L^2\,(L^2 - 12\,M^2)}{M^2}}M\right)^{-3}$$

Massless particle

We start with the equation (2.9.4.1):

$$V^2(r) = \left(1 - \frac{2M}{r}\right)\frac{L^2}{r^2}$$

If we take the derivative we get

$$\frac{\mathrm{d}V^2}{\mathrm{d}r} = \frac{2}{r^3} + \frac{6M}{r^4}$$

which means, that there is **a single maximum** at

$$r_{\max} = 3M$$

with

$$V_{\max}^2 = \frac{L^2}{27M^2} \quad \Rightarrow \quad V_{\max} = \frac{L}{\sqrt{27}M}$$