General Relativity

Winter term 2015/2016

Lecture of Prof. Dr. J. J. van der Bij and Dr. C. Steinwachs
Transcript of M. Ruf, B. Rottler

March 11, 2016

Physikalisches Institut
Albert-Ludwigs-Universität
Freiburg im Breisgau

Contents

1. Newtonian Gravity

In NEWTONian physics we assume that we have absolute space and time that can be described by the a set of numbers x^1, x^2, x^3, t. We express the coordinates as functions of time.

1.1. Forces

The force $\mathbf{F}_{AB}$, which a massive body A with mass m_A exerts on another massive body B with mass m_B, is given by

$$\mathbf{F}_{AB} = -m_B \frac{G_\mathrm{N} m_A}{r^2} \mathbf{e}_r \,, \tag{1.1}$$

where G_N denotes NEWTON*'s constant*, numerically equal to $G_\mathrm{N} \approx 6.673 \cdot 10^{-11}\,\mathrm{m^3/kg\,s}$. Although there is no need for G_N to be constant over time, there is evidence that the relative variation is less than 10^{-12} per year. The force can be expressed in terms of *gravitational potential* Φ:

$$\mathbf{F}_{AB} = -m_B \nabla \left(-\frac{G_\mathrm{N} m_A}{r} \right) =: -m_B \nabla \Phi(\mathbf{r}_B) \,. \tag{1.2}$$

Given N particles labeled by n, the total force B experiences is

$$\mathbf{F}_B = -\sum_n \mathbf{F}_{nB} \,. \tag{1.3}$$

The potential at $\mathbf{r}$ is then easily found to be

$$\Phi(\mathbf{r}) = -G_\mathrm{N} \sum_n \frac{m_n}{|\mathbf{r} - \mathbf{r}_n|} \,. \tag{1.4}$$

In general, we assume a mass distribution $\varrho(\mathbf{r})$ and the sum is replaced by an integral:

$$\Phi(\mathbf{r}) = -G_\mathrm{N} \int_{\mathbb{R}^3} \mathrm{d}\mathbf{r}' \, \frac{\varrho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, . \tag{1.5}$$

1.2. Comparison with electrostatics

The classical theory of gravity bears a striking similarity to electrostatics. To make this clearer, we introduce the gravitational field $\mathbf{g}(\mathbf{r}) := -\nabla\Phi(\mathbf{r})$.

Table 1.1. – Comparison of electrostatics and NEWTONian gravity.

	NEWTONian Gravity	Electrostatics
Force	$\mathbf{F} = q\dfrac{kQ}{r^2}\mathbf{e}_r$	$\mathbf{F} = m\frac{G_\mathrm{N} M}{r^2}\mathbf{e}_r$
Potential	$\Phi_\mathrm{el}(\mathbf{r}) = q\frac{kQ}{r}$	$\Phi_\mathrm{g}(\mathbf{r}) = m\frac{G_\mathrm{N} M}{r}$
Field	$\mathbf{E}(\mathbf{r}) = -\nabla\Phi_\mathrm{el}(\mathbf{r})$	$\mathbf{g}(\mathbf{r}) = -\nabla\Phi_\mathrm{g}(\mathbf{r})$
LAPLACE equation	$\Delta\Phi_\mathrm{el} = -4\pi k \varrho_\mathrm{el}(\mathbf{r})$	$\Delta\Phi_\mathrm{g} = 4\pi G_\mathrm{N} \varrho_\mathrm{g}(\mathbf{r})$

Example 1 (Field of a spherical mass distribution). Assume we have a spherical mass distribution, i.e. $\varrho(\mathbf{r}) = \varrho(r)$. By symmetry considerations it follows, that the gravitational field can be expressed by

$$\mathbf{g}(\mathbf{r}) = g(r)\mathbf{e}_r \, . \tag{1.6}$$

We integrate the divergence of the field over a ball B of radius r

$$\int_B \mathrm{d}\mathbf{r} \, \nabla\mathbf{g} = -\int_B \mathrm{d}\mathbf{r} \, \Delta\Phi = -4\pi G_\mathrm{N} \int_B \mathrm{d}\mathbf{r} \, \varrho(r) = -4\pi G_\mathrm{N} M \, , \tag{1.7}$$

where M is the mass enclosed in B. On the other hand we can use Gauss's theorem to deduce

$$\int_B \mathrm{d}\mathbf{r} \, \nabla\mathbf{g} = \oint_{\partial B} \mathrm{d}\mathbf{A} \cdot \mathbf{g} = \oint_\Omega \mathrm{d}\Omega \, g(r) r^2 = 4\pi r^2 g(r) \, . \tag{1.8}$$

Together the gravitational field is given by

$$\mathbf{g}(r) = -\frac{G_{\mathrm{N}} M}{r^2} \mathbf{e}_r \,. \tag{1.9}$$

Weak Equivalence Principle (WEP)

NEWTON's first law reads

$$\mathbf{F} = m_{\mathrm{I}} \ddot{\mathbf{x}} \,, \tag{1.10}$$

where m_{I} is the inertial mass that works against the acceleration of the body. The force which a body with "active" mass $m_{\mathrm{g,a}}$ exerts on another body with mass $m_{\mathrm{g,p}}$ is given by

$$\mathbf{F} = m_{\mathrm{g,p}} \frac{G_{\mathrm{N}} m_{\mathrm{g,a}}}{r^2} \mathbf{e}_r \,. \tag{1.11}$$

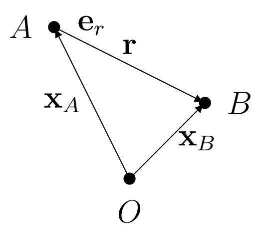


Figure 1.1.

A priori, there is no reason to assume any relation between this masses. The first question one might ask is whether the active and the passive mass are equal. Suppose we have two masses A and B. Using NEWTON's first law, we can explicitly write

$$m_{\mathrm{I}}^B \ddot{\mathbf{x}} = \mathbf{F}_{AB} = -m_{\mathrm{g,p}}^B \frac{G_{\mathrm{N}} m_{\mathrm{g,a}}^A}{r^2} \mathbf{e}_r \,, \tag{1.12}$$

$$m_{\mathrm{I}}^A \ddot{\mathbf{x}} = \mathbf{F}_{BA} = -m_{\mathrm{g,p}}^A \frac{G_{\mathrm{N}} m_{\mathrm{g,a}}^B}{r^2} \mathbf{e}_r \,. \tag{1.13}$$

By the third law $\mathbf{F}_{AB} = -\mathbf{F}_{BA}$ we have

$$\frac{m_{\mathrm{g,p}}^B}{m_{\mathrm{g,a}}^B} = \frac{m_{\mathrm{g,p}}^B}{m_{\mathrm{g,a}}^B} \,. \tag{1.14}$$

By proper choice of mass units we can set this quotient to one so that

$$m_{\mathrm{g,a}} = m_{\mathrm{g,p}} =: m_{\mathrm{g}} \tag{1.15}$$

Year	Experimenter	Experiment	Accuracy
1636	Galilei	inclined planes	10^{-2}
1689	Newton	pendulum	10^{-3}
1832	Bessel	pendulum	10^{-5}
1922	Eötvös	pendulum	10^{-9}
1922	Shapro et al.	pendulum	10^{-12}
1999	Baesler	torsion balance	10^{-14}

Table 1.2. – Experiments measuring the ratio $\frac{m_\mathrm{I}}{m_\mathrm{g}}$.

The next question is weather the inertial mass equivalent to the gravitational mass. By Newton's first law we have

$$m_\mathrm{I}\ddot{\mathbf{x}} = -m_\mathrm{g}\frac{G_\mathrm{N}M_\mathrm{g}}{r^2}\mathbf{e}_r = -m_\mathrm{g}\mathbf{g}\,. \tag{1.16}$$

As a experimental result that has been measured up to a high accuracy (compare tabular 1.2) all bodys recive the same acceleration due to gravity $\ddot{\mathbf{x}} \sim \mathbf{g}$. By a 'proper' choice of units of the flight-time $t = \sqrt{\frac{m_\mathrm{I}}{m_\mathrm{g}}}\sqrt{\frac{2h}{g}}$, we get

$$m_\mathrm{I} = m_\mathrm{g} =: m\,. \tag{1.17}$$

1.2.1. Tidal Forces

Assume we have a body of finite extension in a gravitational potential Φ, an example being the earth in the potential of the moon. On the center of the body we have

$$m\,\frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} = -\,\frac{\partial\Phi}{\partial x_i}\;. \tag{1.18}$$

If we consider a point shifted by χ^i from the center then the acceleration is given as

$$
\begin{aligned}
m\,\frac{\mathrm{d}^2}{\mathrm{d}t^2}\left(x^i+\chi^i\right) &= -\,\frac{\partial\Phi\big(x^i+\chi^i\big)}{\partial x_i} \\
&\simeq -\,\frac{\partial\Phi\big(x^i\big)}{\partial x_i}-\frac{\partial^2\Phi\big(x^i\big)}{\partial x_i\partial x_j}\,\chi^j\,.
\end{aligned}
\tag{1.19}
$$

Subtracting the previous equations yields the tidal force

$$
m\,\frac{\mathrm{d}^2\chi^i}{\mathrm{d}t^2}=-\,\frac{\partial^2\Phi\big(x^i\big)}{\partial x_i\partial x_j}\,\chi^j\,. \tag{1.20}
$$

The tidal force tensor $\dfrac{\partial^2\Phi}{\partial x_i\partial x_j}$ is of the form

$$
\frac{\partial^2\Phi}{\partial x_i\partial x_j}=\frac{G_{\mathrm{N}}M}{r^3}\left(\delta_{ij}-3\frac{x_i x_j}{r^2}\right). \tag{1.21}
$$

2. Special Relativity

2.1. Postulates and Definitions

Special relativity is based on two main postulates, namely

1. Principle of relativity:
   The laws of physics acquire the same form in all inertial systems.

2. Constancy of the speed of light:
   The speed of light in vacuum is constant $c \approx 3 \cdot 10^8\, {}^{\mathrm{m}}\!/\!_{\mathrm{s}}$.

We further define a series of objects, that will come handy when describing relativity:

Definition 1 (System of reference)**.** A *system of reference* K is a system of three spacial coordinates to indicate the position and one time coordinate to indicate the time.

Definition 2 (Inertial system)**.** An *inertial system* I belongs to a particular subspace of reference systems in which a freely moving bodies, i.e. that are not subject to any external force, move with constant velocity on an straight lines.

Definition 3 (Event)**.** An *event* or *world point* x is a point in spacetime. In any system of reference it can be described by four coordinates

$$x^\mu : (x^0, x^1, x^2, x^3)\,. \tag{2.1}$$

For example in Cartesian coordinates, we have

$$(x^0, x^1, x^2, x^3) = (ct, x, y, z)\,. \tag{2.2}$$

Definition 4 (Worldline)**.** A *worldline* γ is a parametrised curve in spacetime. To each parameter, there corresponds an event that lies on the worldline

$$\gamma^\mu(\lambda) = x^\mu(\lambda)\,. \tag{2.3}$$

2.1.1. Einstein Summation Convention

Since we will have to deal which many indices, we will find a way to keep notation as compact as possible. We agree that whenever we have a summation over an upper and a lower index like $\sum_\mu x_\mu x^\mu$, we drop the summation and assume that terms of type $x_\mu x^\mu$ are always summed over.

2.2. Propagation of Light Waves and the Line Element

We consider the propagation of a light ray, characterized by two events P and Q, in an Inertial System I:

- P : Event of emission of the ray (ct_1, x_1, y_1, z_1)
- Q : Event of absorption of the ray (ct_2, x_2, y_2, z_2)

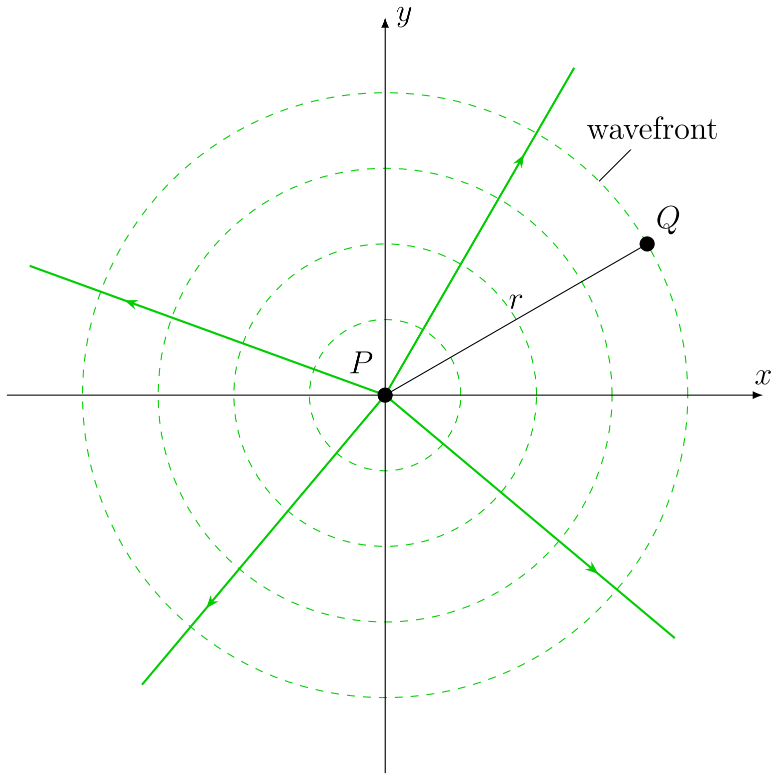


Figure 2.1. – A spherical light pulse.

The spatially distance between the events is

$$r = \left[(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2\right]^{\frac{1}{2}} . \tag{2.4}$$

Since we are tracking a light ray and the events are absorption and emission, we further have

$$r = c(t_2 - t_1) . \tag{2.5}$$

We construct the quantity

$$\Delta s^2 = -c^2(t_2 - t_1)^2 + (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 , \tag{2.6}$$

so that $\Delta s^2 = 0$ for light.

In another Inertial System I', the events are given by

- P : Event of emission $(ct'_1, x'_1, y'_1, z'_1)$
- Q : Event of absorption $(ct'_2, x'_2, y'_2, z'_2)$

By the same reasoning and the invariance of the physical laws and the constancy of the speed of light, we have $\Delta s'^2 = 0$ for light. We define an infinitesimal interval, or *line element*:

$$\mathrm{d}s^2 = -c^2\,\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 . \tag{2.7}$$

Our preliminary considerations lead to the following statement: if the interval is zero in one inertial system, it should be zero in all. To relate the lineelements of different systems to each other, we make the following thought: Suppose we have three inertial Systems I, I_1, I_2. The Systems I_1 and I_2 move at constant velocity $\mathbf{v}_1$ and $\mathbf{v}_2$ relative to I. Further I_2 moves with velocity $\mathbf{v}_{12}$ relative to I_1. The demand implies that there exists a function α, only dependent on the velocity $\mathbf{v}$, so that

$$\mathrm{d}s'^2 = \alpha(\mathbf{v})\,\mathrm{d}s^2 . \tag{2.8}$$

We have

$$\mathrm{d}s^2 = \alpha(\mathbf{v}_1)\,\mathrm{d}s_1^2 , \quad \mathrm{d}s^2 = \alpha(\mathbf{v}_2)\,\mathrm{d}s_2^2 , \quad \mathrm{d}s_1^2 = \alpha(\mathbf{v}_{12})\,\mathrm{d}s_2^2 . \tag{2.9}$$

Together they imply

$$\alpha(\mathbf{v}_{12}) = \frac{\alpha(\mathbf{v}_2)}{\alpha(\mathbf{v}_1)} , \tag{2.10}$$

but since the velocities where arbitary α has to be constant and we are free to chose $\alpha \equiv 1$. Therefore the line element has to stay invariant.

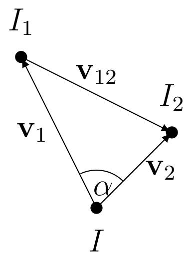


Figure 2.2.

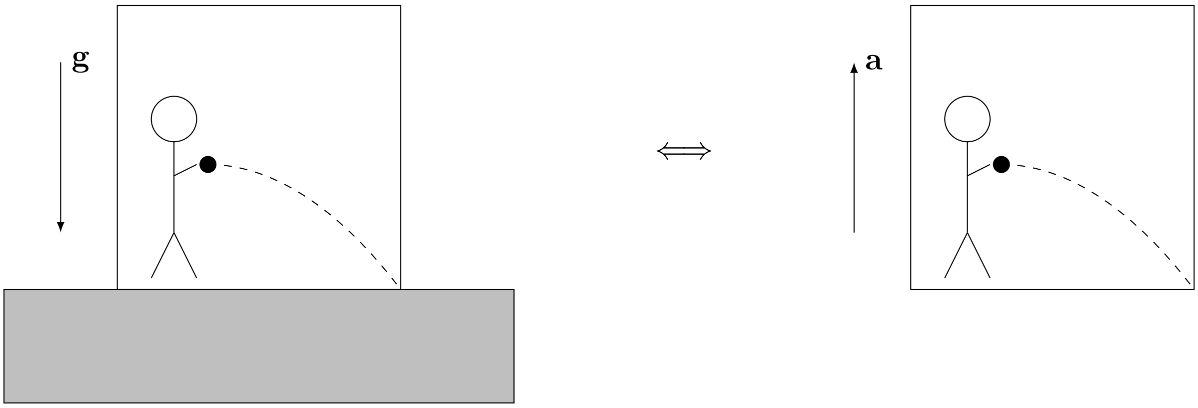


Figure 2.3.

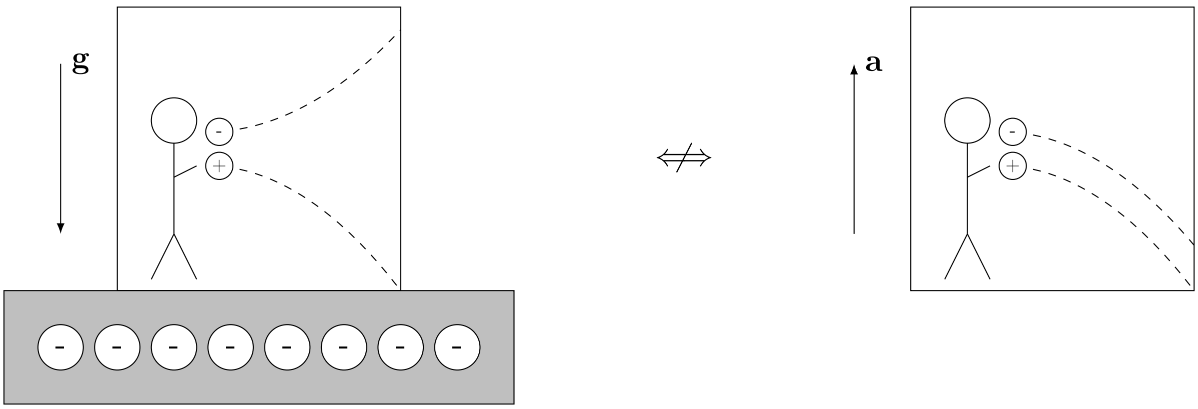


Figure 2.4.

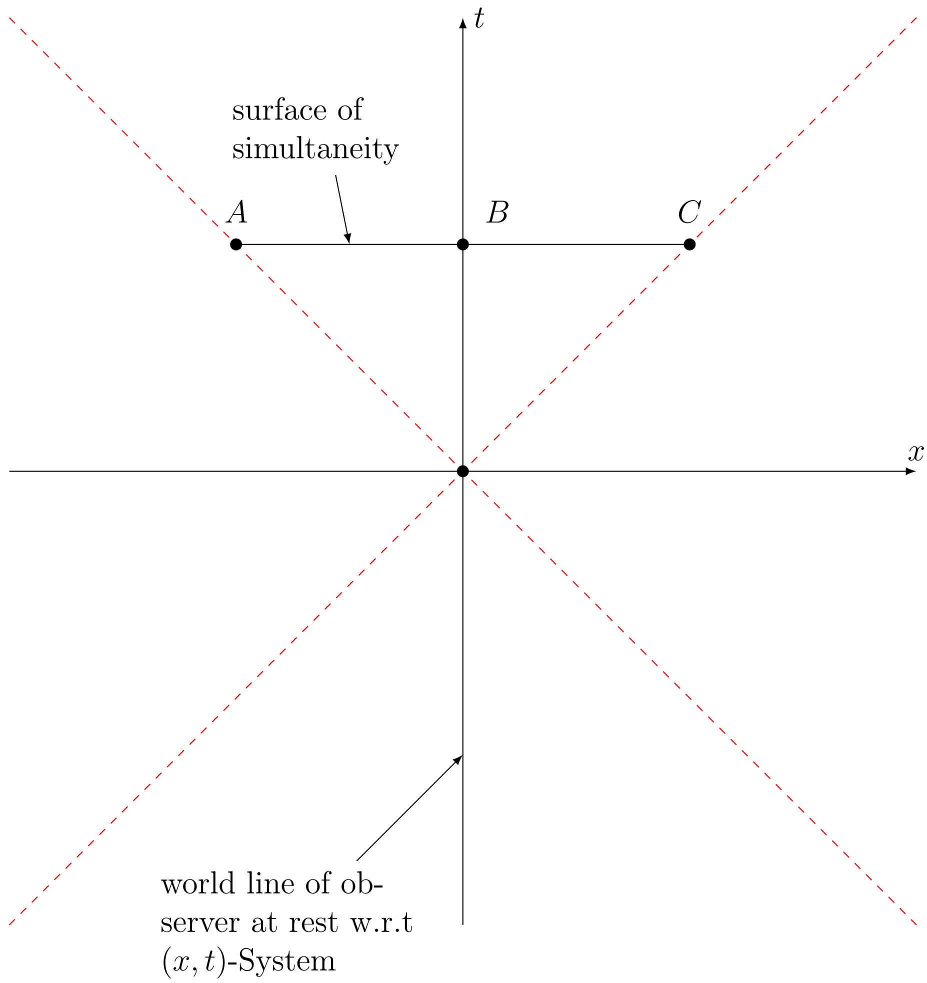

t
surface of simultaneity
A
B
C
x
world line of observer at rest w.r.t (x, t)-System

Figure 2.5.

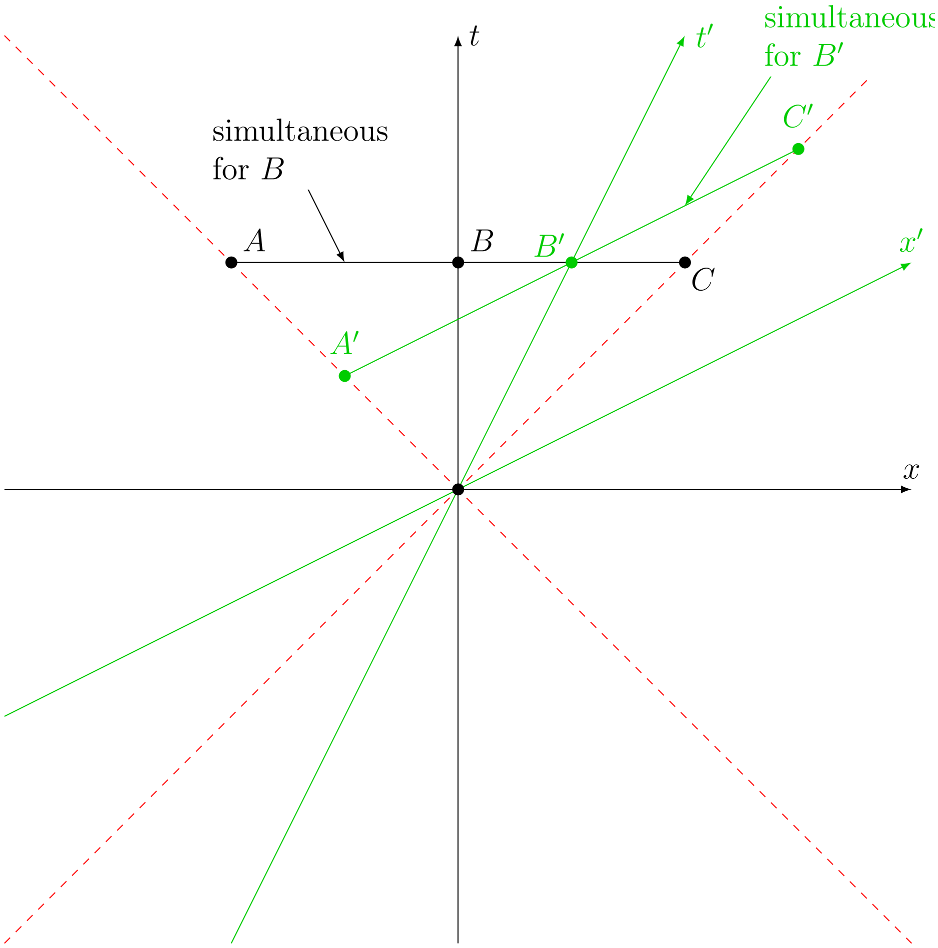


Figure 2.6.

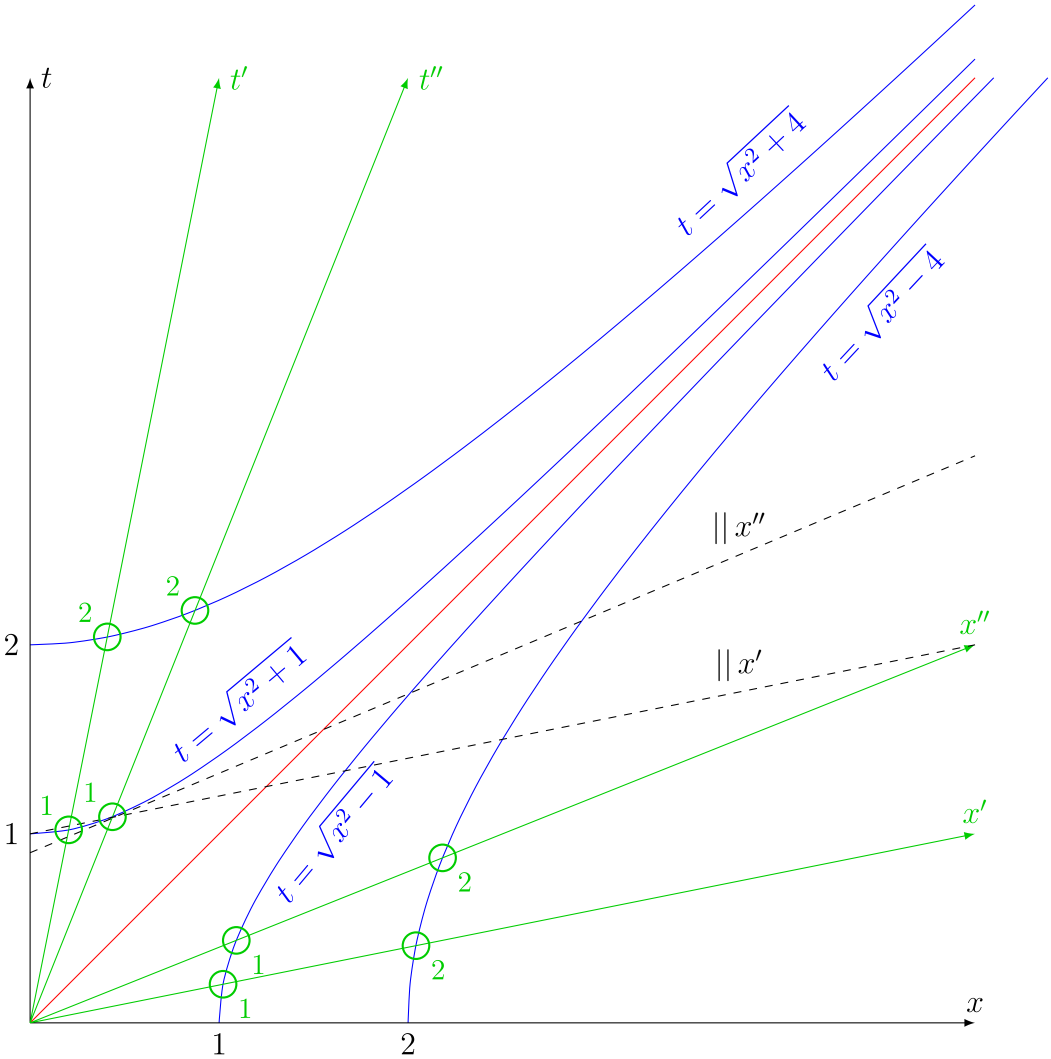


Figure 2.7.

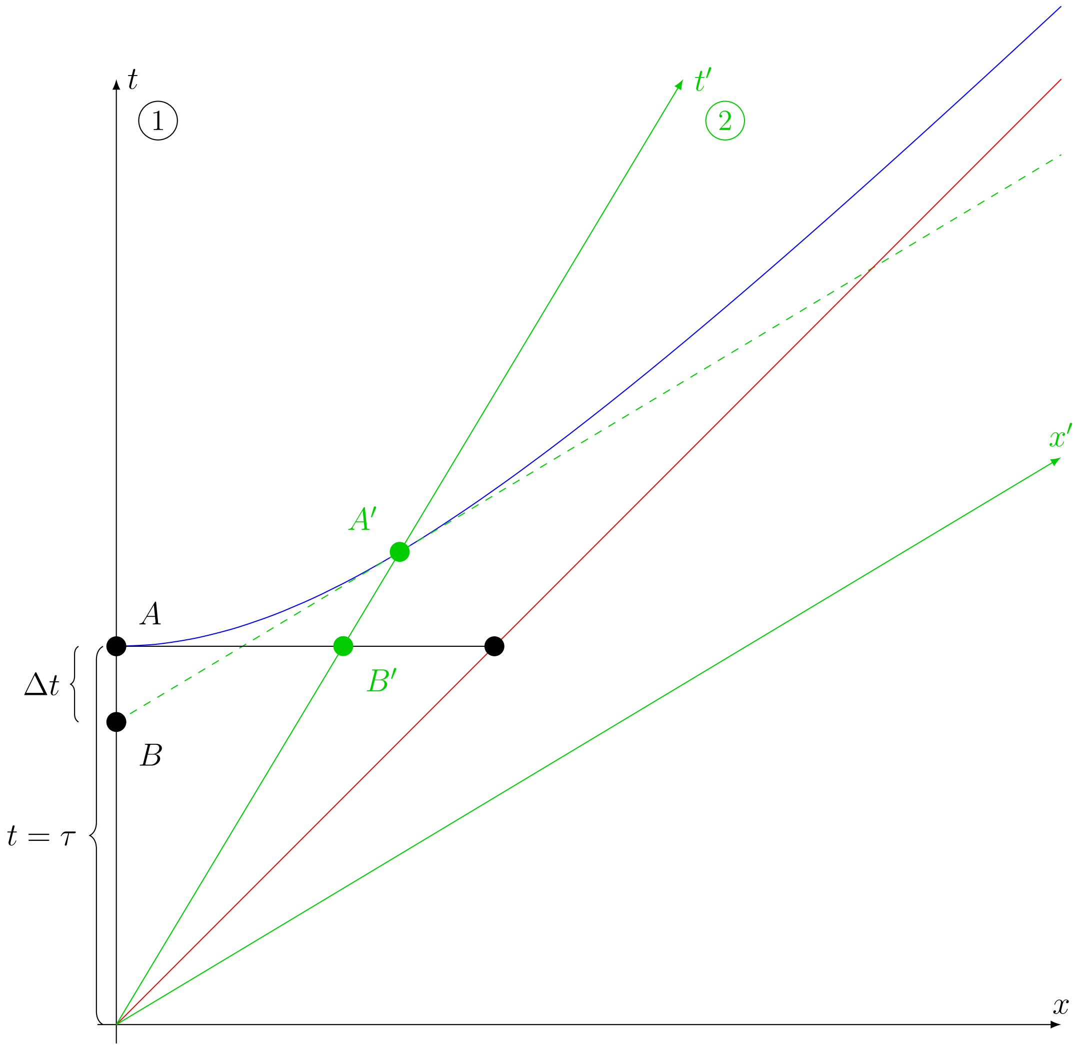


Figure 2.8.

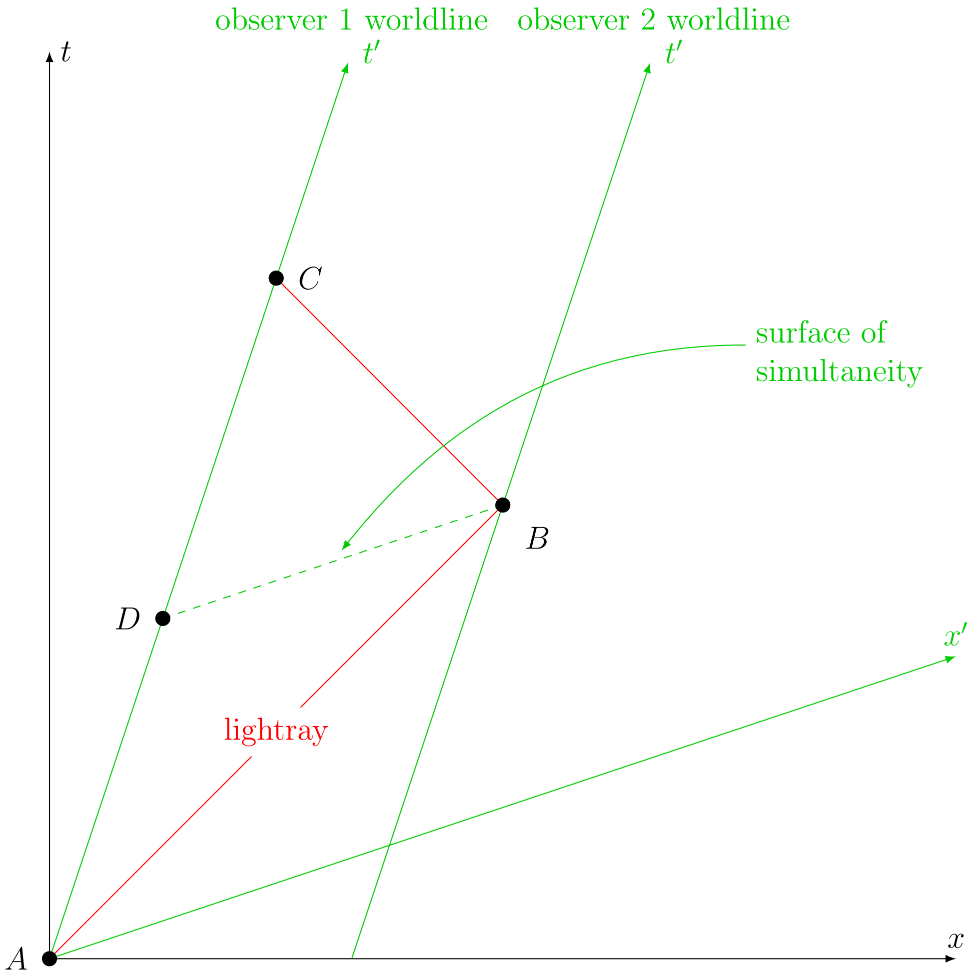

observer 1 worldline
observer 2 worldline
t
t'
t'
C
surface of
simultaneity
B
D
x'
lightray
A
x

Figure 2.9.

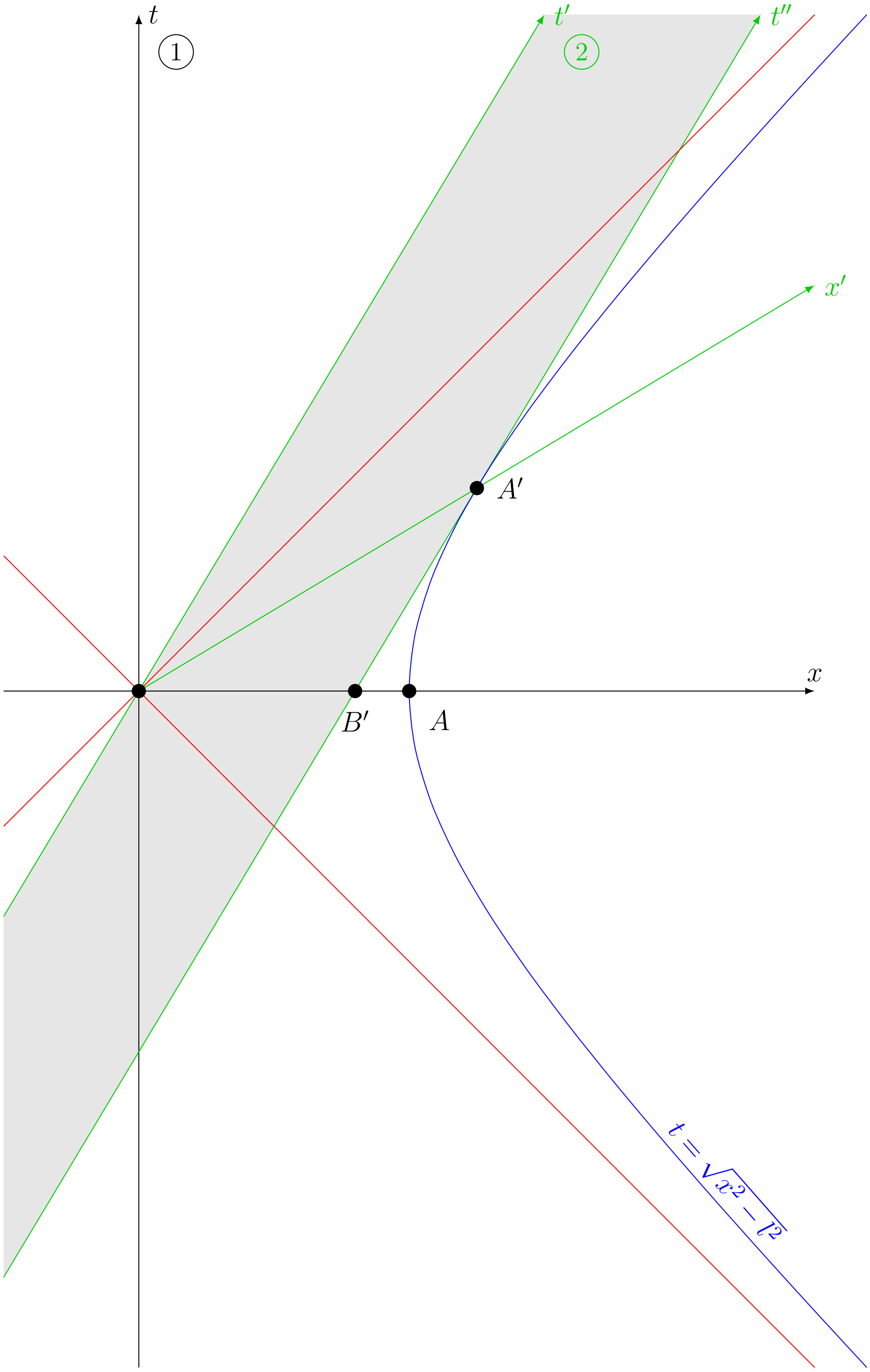


Figure 2.10.

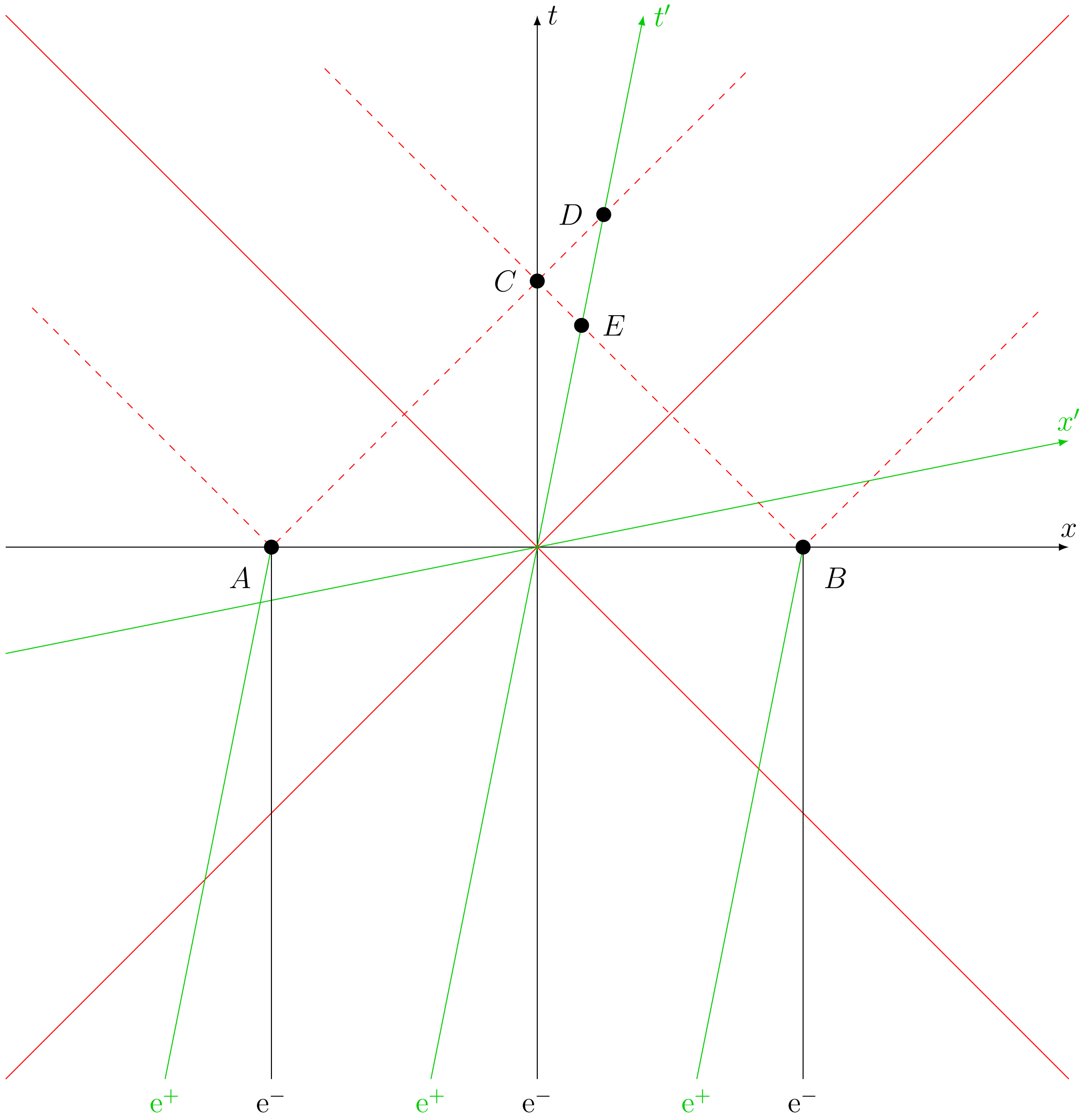


Figure 2.11.

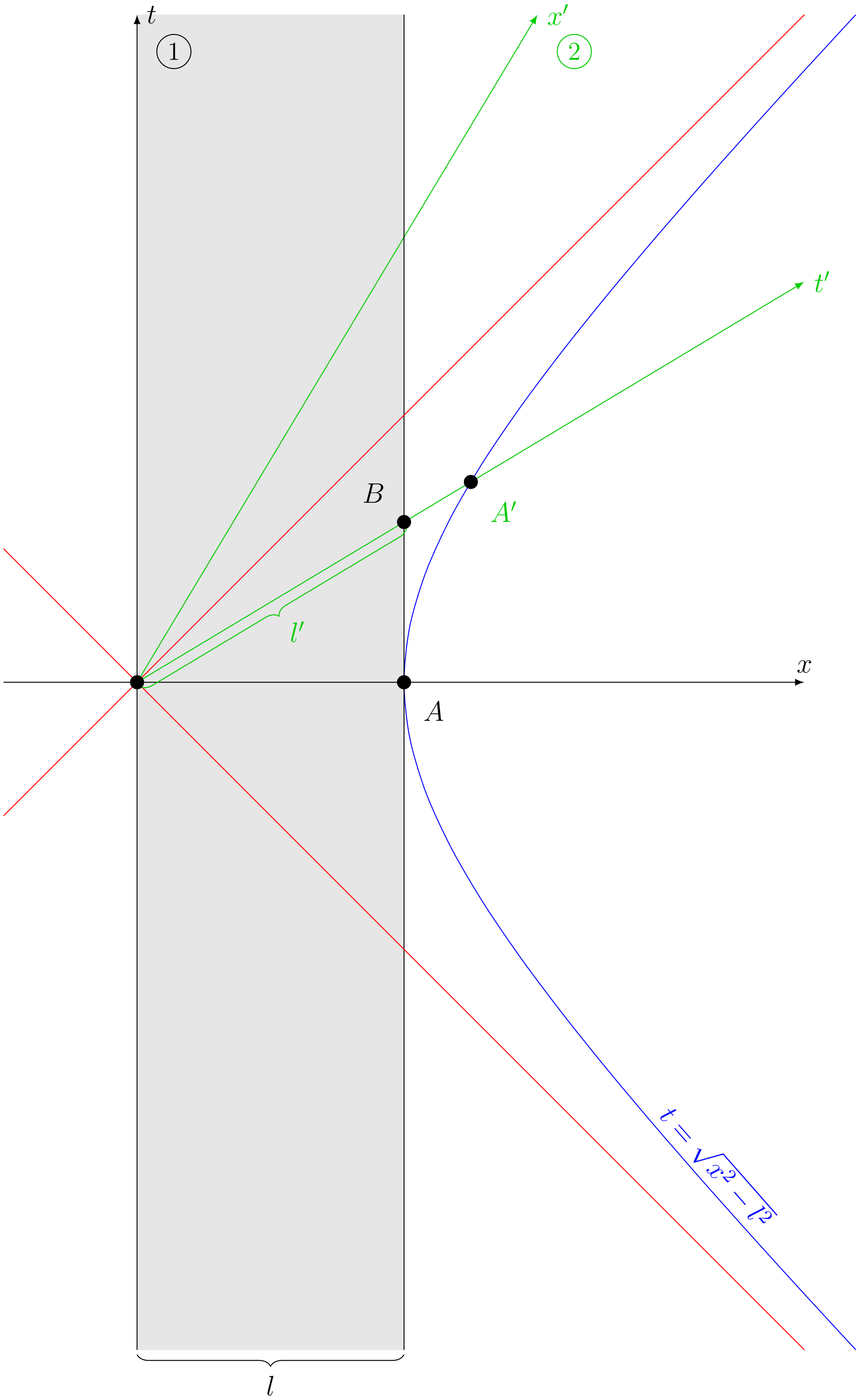


Figure 2.12.

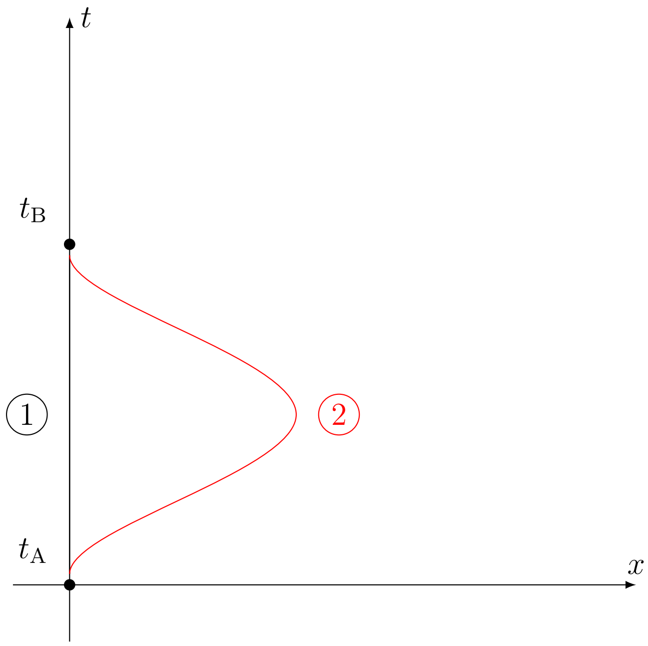


Figure 2.13.

2.3. Invariant Distances, Metric and Signature

In flat space we can calculate the distance Δl between two points (x_1, y_1) and (x_2, y_2) by the PYTHAGORAS' theorem

$$\Delta l^2 = \Delta x^2 + \Delta y^2\,, \quad \Delta x = x_2 - x_1\,, \quad \Delta y = y_2 - y_1\,. \tag{2.11}$$

For an infinitesimal distance $\mathrm{d}l$ we recover

$$\mathrm{d}l^2 = \mathrm{d}x^2 + \mathrm{d}y^2 = \delta_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j\,. \tag{2.12}$$

If we introduce

$$(\mathrm{d}x^i) = \begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix}, \quad (\delta_{ij}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{2.13}$$

We can expand the formula for the infinitesimal element and obtain

$$\mathrm{d}l^2 = \begin{bmatrix} \mathrm{d}x & \mathrm{d}y \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathrm{d}x \\ \mathrm{d}y \end{bmatrix}. \tag{2.14}$$

Which we do to point out the similarity to the line element $\mathrm{d}s^2$

$$\mathrm{d}s^2 = \begin{bmatrix} \mathrm{d}t & \mathrm{d}x & \mathrm{d}y & \mathrm{d}z \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathrm{d}t \\ \mathrm{d}x \\ \mathrm{d}y \\ \mathrm{d}z \end{bmatrix} =: \eta_{\mu\nu}\, \mathrm{d}x^\mu\, \mathrm{d}x^\nu\,. \tag{2.15}$$

We call the matrix $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ the MINKOWSKI-*metric*. It has some obvious properties:

- constancy: $\eta_{\mu\nu,\varrho} = 0$
- symmetry: $\eta_{\mu\nu} = \eta_{\nu\mu}$
- self inverse: $\eta^{\mu\nu} := {\eta^{-1}}_{\mu\nu} = \eta_{\mu\nu}$

We can use the metric to raise and lower indices:

$$x_\nu = \eta_{\mu\nu} x^\mu\,, \quad x^\mu = \eta^{\mu\nu} x_\nu\,. \tag{2.16}$$

One has to pay attention with the matrix form, for example, the *trace* of the metric is given by

$$\eta^\mu{}_\mu = \eta^{\mu\nu}\eta_{\nu\mu} = \delta^\mu{}_\mu = 4\,. \tag{2.17}$$

We say $\eta_{\mu\nu}$ has signature $(-,+,+,+)$ or $(1,3)$, because it has one negative and three positive eigenvalues.

2.3.1. Poincaré-transformation

To be consistent with a constant speed of light a transformation between two coordinate systems must leave the line element invariant, i.e.

$$\mathrm{d}s'^2 = \eta_{\mu\nu}\, \mathrm{d}x'^\mu\, \mathrm{d}x'^\nu = \eta_{\mu\nu}\, \mathrm{d}x^\mu\, \mathrm{d}x^\nu = \mathrm{d}s^2\,. \tag{2.18}$$

We consider afine coordinate transformations

$$x'^\mu = f^\mu(x^\nu) = L^\mu{}_\nu x^\nu + a^\mu\,. \tag{2.19}$$

Where $L^\mu{}_\nu$ is a a LORENTZ transformation and a^μ is a constant shift.
We can now inspect the transformation properties of an allowed transformation. The

invariance of the line element implies

$$\mathrm{d}s'^2 = \eta_{\mu\nu}\,\mathrm{d}x'^\mu\,\mathrm{d}x'^\nu = \eta_{\varrho\sigma}L^\varrho{}_\mu L^\sigma{}_\nu\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \overset{!}{=} \mathrm{d}s^2\,, \tag{2.20}$$

independent of the shift a^μ. We will therefore focus our attention towards the linear transformation $L^\mu{}_\nu$. Equation (2.18) implies that

$$\eta_{\mu\nu} = L^\varrho{}_\nu \eta_{\varrho\sigma} L^\sigma{}_\nu\,. \tag{2.21}$$

We can take a look at an infinitimal transformation, that is up to an order of ε^2

$$L^\mu{}_\nu = \delta^\mu{}_\nu + \varepsilon\omega^\mu{}_\nu\,. \tag{2.22}$$

The $\omega^\mu{}_\nu$ are called the *generators* of the transformation. Plugging into (2.21), ignoring higher powers in ε^2 results in

$$\begin{aligned} \eta_{\mu\nu} &= \eta_{\varrho\sigma}\Big(\delta^\varrho{}_\mu + \varepsilon\omega^\varrho{}_\mu\Big)(\delta^\sigma{}_\nu + \varepsilon\omega^\sigma{}_\nu) \\ &= \eta_{\mu\nu} + \varepsilon\Big(\omega_{\mu\nu} + \omega_{\nu\mu}\Big)\,. \end{aligned} \tag{2.23}$$

Since this must hold true for arbitrary ε, the generators $\omega_{\nu\mu}$ must satisfy $\omega_{\nu\mu} = -\omega_{\mu\nu}$, i.e. be antisymmetric. In general we have $\frac{n(n-1)}{2}$ of these objects, where n is the dimension of the underlying space. Since we are considering four dimensional spacetime, there are six generators. The finite transformations by an generalized angle $\alpha \in \mathbb{C}$ can be obtained from the generators via

$$L^\mu{}_\nu = \exp(\alpha\omega^\mu{}_\nu)\,. \tag{2.24}$$

We can classify the resulting transformations into two classes:

Rotations

Rotations mix the spatial coordinates, but leave the time fixed. Consider for example an rotation around z-axis. Then we have

$$x' = x\cos\alpha + y\sin\alpha\,,\quad y' = -x\sin\alpha + y\cos\alpha\,. \tag{2.25}$$

The infinitesimal generator is

$$\omega^{\mu}{}_{\nu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.26}$$

and the transformation can be expressed in matrix form

$$\begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} \tag{2.27}$$

Boosts

Boosts can be thought of as rotations between time and spacial coordinates. We consider a boost in x-direction. Since the line element must be invariant and a linear transformation keeps the origin fixed we have

$$-t^2 + x^2 = -t'^2 + x'^2 \,. \tag{2.28}$$

This is an hyperbolic equation and can be parametrised by

$$t = t' \cosh\psi + x' \sinh\psi \,, \quad x = t' \sinh\psi + x' \cosh\psi \,. \tag{2.29}$$

Where ψ is called the *rapidity*. The generator is

$$\omega^{\mu}{}_{\nu} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} . \tag{2.30}$$

The matrix form of the transformation is

$$\begin{bmatrix} t \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cosh\psi & \sinh\psi & 0 & 0 \\ \sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t' \\ x' \\ y' \\ z' \end{bmatrix} . \tag{2.31}$$

The rapidity ψ is not related to a physical quantity yet. We may ask to which velocity v does this boost correspond. Therefore we consider the line $x' = 0$. Then

$$t = t' \cosh\psi \,, \quad x = t' \sinh\psi \,. \tag{2.32}$$

We can express the velocity v, as seen by an observer in the primed system

$$v = \frac{x}{t} = \tanh\psi \,. \tag{2.33}$$

It is convenient to introduce the parameters β and γ defined as follows

$$\beta := v \,, \quad \gamma := \frac{1}{\sqrt{1-\beta^2}} \,. \tag{2.34}$$

In terms of these we find

$$\sinh\psi = \beta\gamma \,, \quad \cosh\psi = \gamma \,. \tag{2.35}$$

Therefore an alternative form is given by

$$L^\mu{}_\nu = \begin{bmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} . \tag{2.36}$$

And the coordinates are therefore related by

$$t = \gamma\big(t' + \beta x'\big) \,, \quad x = \gamma\big(x' + \beta t'\big) \,. \tag{2.37}$$

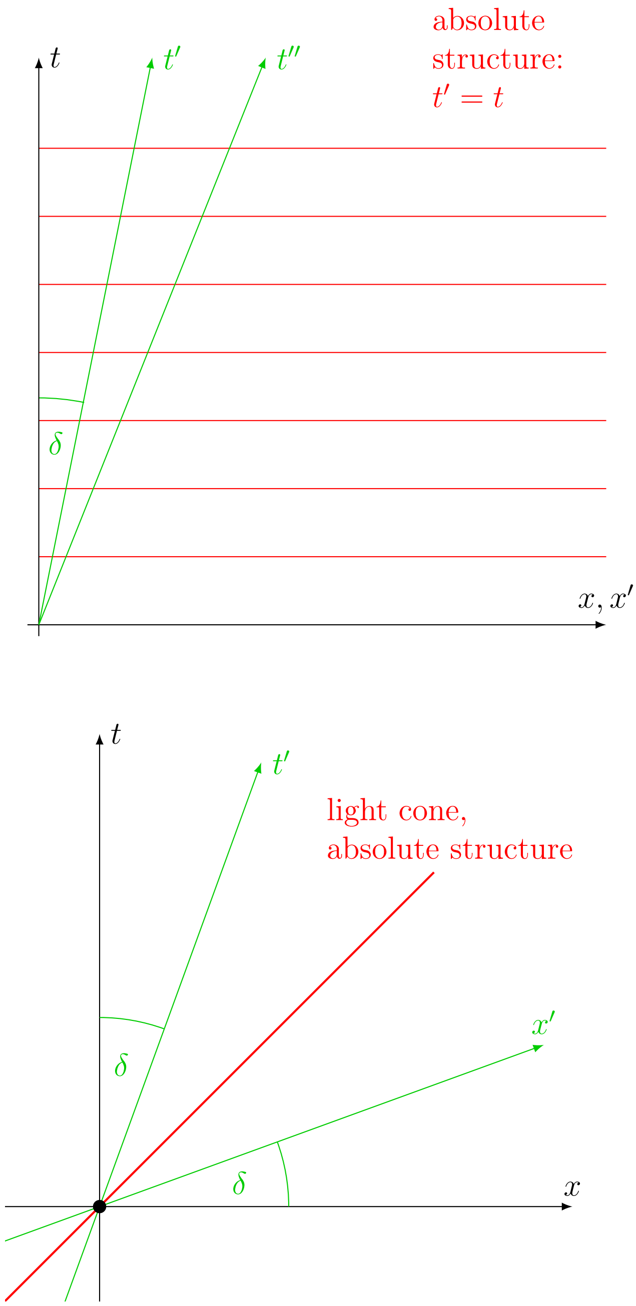
All together we have ten independent (affine) transformations that leave the line element invariant. Namely:

- four shifts by $a^\mu = (a^0, a^1, a^2, a^3)^\top$
- three rotations in space corresponding to three real angles $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)^\top$ (Euler angles)
- three boosts with constant velocity $\mathbf{v} = (v_1, v_2, v_3)^\top$

The Poincaré transformations form a ten parameter group.

Table 2.1. – Comparison of NEWTONian theory and special relativity.

	NEWTON	EINSTEIN (SR)
	law of inertia	law of inertia
NEWTON's Laws	$\mathbf{F} = m\ddot{\mathbf{x}} = m\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}$	$\mathbf{F} = m\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}\tau}$
	$\mathbf{F}_{12} = -\mathbf{F}_{21}$	momentum conservation
		(postulate)
transformations	POINCARÉ transformations	GALILEI transformations
	absolute structure	absolute time
	$c = \text{const.}$	$t' = t$
force	$\mathbf{F} = \frac{G_\mathrm{N} m_1 m_2}{r^2}$	$F^\alpha = \frac{q}{m} U_\beta F^{\alpha\beta},\ F = \gamma(f_0, \mathbf{f})^\top$
	$\Delta\Phi = 4\pi\varrho$	MAXWELL's equations



absolute
structure:
$t' = t$
t
t'
t''
δ
x, x'
t
t'
light cone,
absolute structure
x'
δ
δ
x

Figure 2.14.

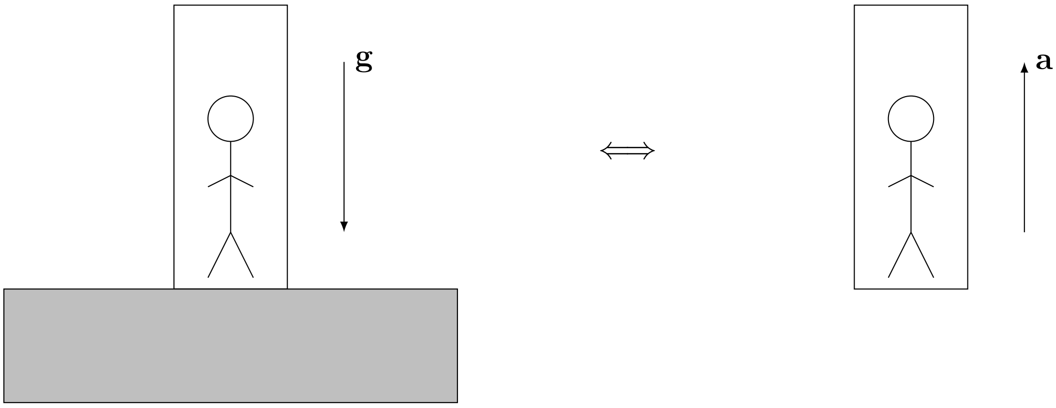


Figure 2.15.

2.4. Vectors and Tensors in SRT

At this point we will just list some objects, that we call vectors and tensors and delay the definition of those to a later chapter. A bit sloppy we refer to a vector V with components V^μ, as a object that transforms as

$$V^\mu \to L^\mu{}_\alpha V^\alpha \tag{2.38}$$

and to a tensor as a 'vector with more indices' that transforms accordingly, i.e.

$$T^{\mu\nu} \to L^\mu{}_\alpha L^\nu{}_\beta T^{\alpha\beta}\,, \tag{2.39}$$

and so forth. An example of a vector is the four current J^μ:

$$J^\mu = \gamma(\varrho, \mathbf{j})^\top\,, \tag{2.40}$$

with ϱ the charge density and $\mathbf{j}$ the 3-current density. A typical example is the field strength tensor $F^{\mu\nu}$ of electrodynamics given in terms of the vector potential A_μ:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,. \tag{2.41}$$

The matrix form is

$$F_{\mu\nu} = \begin{bmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{bmatrix}. \tag{2.42}$$

Therefore the tensor $F^{\mu\nu}$ takes the role of the fields $\mathbf{E}, \mathbf{B}$. In tensor form Maxwell equations take the particular simple form, namely

$$\partial_\nu F^{\mu\nu} = 4\pi J^\mu \,, \quad \partial_\alpha F_{\mu\nu} + \partial_\nu F_{\alpha\mu} + \partial_\mu F_{\nu\alpha} = 0 \,. \tag{2.43}$$

The Lorentz force is given by

$$F^\mu = qF^{\mu\nu}u_\nu \tag{2.44}$$

Where $u^\mu = \gamma(1, \mathbf{v})^\top$ is the four velocity. As another example we can take a look at plane waves. The vector potential of a plane wave is

$$A^\mu = \hat{A}^\mu \exp\big[\mathrm{i}(\mathbf{k}\cdot\mathbf{r} - \omega t)\big] = \hat{A}^\mu \exp\Big(\mathrm{i}K^\mu X_\mu\Big) \,. \tag{2.45}$$

Where $K^\mu = (\omega, \mathbf{k})^\top$ must be a four vector, because otherwise we could use plane waves to distinguish between inertial systems. The dispersion relation reads as

$$K^\mu K_\mu = -\omega^2 + \mathbf{k}^2 = 0 \,. \tag{2.46}$$

It is equivalent to the on shell condition (2.72), when we identify the photons energy $E = \omega$, momentum $\mathbf{p} = \mathbf{k}$ and mass $m = 0$.

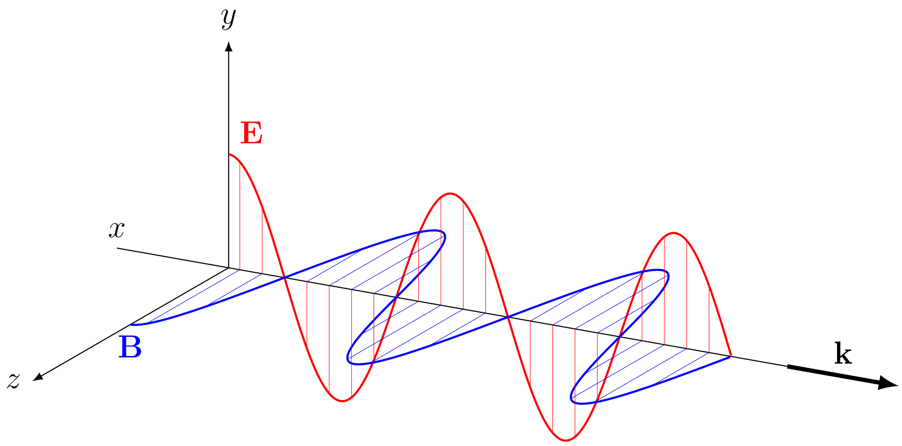


Figure 2.16.

2.4.1. Fourvectors

There are three types of vectors[1]:

[1] We could also put up a fourth class, containing solely the zero vector.

Table 2.2. – Examples of fourvectors and normalisation.

fourvector	definition	normalisation
velocity	$u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \gamma(1, \mathbf{v})^\top$	$u_\mu u^\mu = -1$
momentum	$p^\mu := m u^\mu$	$p_\mu p^\mu = -m^2$
acceleration	$a^\mu := \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau}$	$a_\mu a^\mu = 0$

1. *spacelike* $\Delta x^\mu \Delta x_\mu > 0$
2. *lightlike* $\Delta x^\mu \Delta x_\mu = 0$
3. *timelike* $\Delta x^\mu \Delta x_\mu < 0$

All classes are transformed into themselves by Lorentz transformation. There is a set of transformations that is forbidden by physical considerations: namely the parity change $\mathbf{x} \to -\mathbf{x}$ and time reversal $t \to -t$ because this are no real symmetries of nature[2]. We will mainly deal with timelike and lightlike vectors for which the time order does not change. For spacelike intervals the order can change. To remain causal all particles have to move at velocity $v \leq c$. This is in contrast to Newtonian theory where everything can affect everything, because the velocities are unbounded. A direct consequence of this is for example that even the two body problem has no exact solution in special relativity. Tabular 2.2 shows a list of the dynamic fourvectors in SR, notice that in contrast to the classical, theory the fouraccelaration a_μ has no importance in SR.

2.4.2. The Energy Momentum Tensor

Aside 1. Above a certain limit a, system cannot be stabilized by pressure, because the gravity couples to the pressure and therefore increasing the pressure would lead to a stronger attraction due to gravity and hence unfetchable collapse.

The energy momentum tensor is symmetric

$$T^{\mu\nu} = T^{\nu\mu} \,. \tag{2.47}$$

[2] both violated in weak interactions.

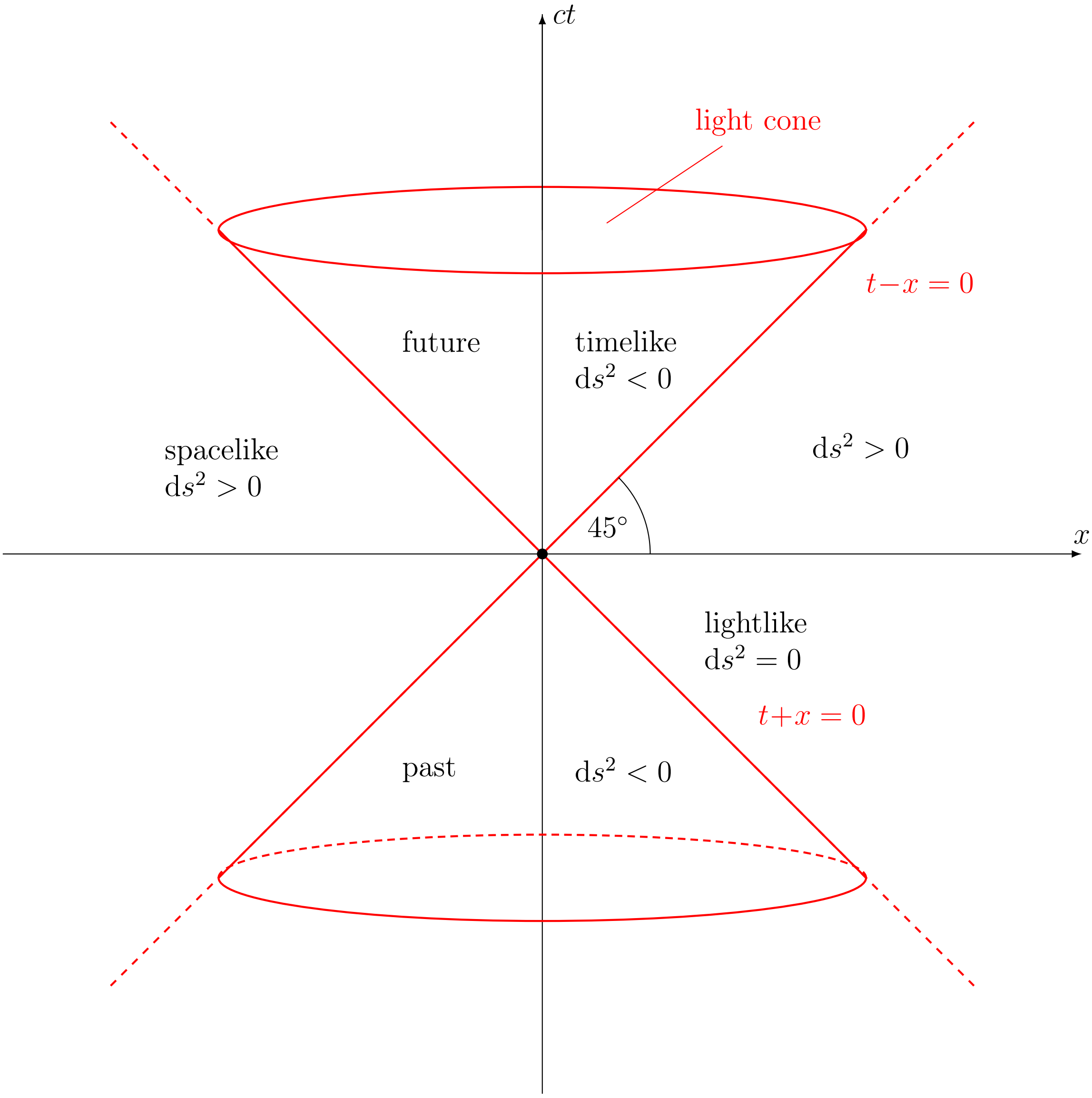


Figure 2.17. – Lightcone structure in special relativity.

Its trace is given by

$$T^{\mu}{}_{\mu} = (\varrho - p) + 4p = -\varrho + 3p\,. \tag{2.48}$$

It is traceless for photons. If an external field is present the divergence is given by

$$T^{\mu\nu}{}_{,\nu} = D^{\mu}\,. \tag{2.49}$$

Where the external momentum is itself given as the divergence

$$D^{\mu} = -S^{\mu\nu}{}_{,\nu}\,. \tag{2.50}$$

Putting both terms together gives

$$\partial_{\mu}(T^{\mu\nu} + S^{\mu\nu}) = 0\,, \tag{2.51}$$

Which can be interpreted as the total energy conservation.

2.4.3. Noether's Theorem

We introduce a tensorial generalisation of the angular momentum

$$T^{\lambda\mu\nu} := x^{\lambda}T^{\mu\nu} - x^{\mu}T^{\lambda\nu}\,. \tag{2.52}$$

The angular momentum is given by

$$L^{ij} = \int \mathrm{d}^3x M^{ij0}\,. \tag{2.53}$$

It can be shown that it is a conserved quantity

$$\partial_{\nu}\left(x^{\lambda}T^{\mu\nu} - x^{\nu}T^{\lambda\nu}\right) = T^{\mu\lambda} - T^{\lambda\mu} + x^{\lambda}T^{\mu\nu}{}_{,\nu} - x^{\nu}T^{\lambda\nu}{}_{,\nu} = 0\,. \tag{2.54}$$

Remark 1. The existence of a conserved symmetric tensor $T^{\mu\nu}$ is necessary in order to build a relativistic theory

Example 2 (Electrodynamics). We start by inspecting the LORENTZ Force F^{μ} which is given by

$$F^{\mu} = eF^{\mu\nu}u_{\nu} = F^{\mu\nu}J_{\nu} = D^{\mu}\,. \tag{2.55}$$

We now ask whether there is a potential $S^{\mu\nu}$ so that $D^{\mu} = -S^{\mu\nu}{}_{,\nu}$. Indeed a potential is given by

$$S^{\mu\nu} = F^{\mu}{}_{\alpha} F^{\nu\alpha} - \frac{1}{4}\eta^{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \,. \tag{2.56}$$

We can check

$$S^{\mu\nu}{}_{,\nu} = F^{\mu}{}_{\alpha,\nu} F^{\nu\alpha} + F^{\mu}{}_{\alpha} F^{\nu\alpha}{}_{,\nu} - \frac{1}{2}\eta^{\mu\nu} F_{\alpha\beta,\nu} F^{\alpha\beta} \,. \tag{2.57}$$

$$F^{\mu}{}_{\alpha,\nu} F^{\nu\alpha} = F^{\mu}{}_{\nu,\alpha} F^{\alpha\nu} = F^{\nu}{}_{\mu,\alpha} F^{\nu\alpha} \tag{2.58}$$

$$\eta^{\mu\nu} F_{\alpha\beta,\nu} F^{\alpha\beta} = F_{\alpha\nu}{}^{,\mu} F^{\alpha\nu} = -F_{\alpha\nu}{}^{,\mu} F^{\nu\alpha} \,. \tag{2.59}$$

$$S^{\mu\nu}{}_{,\nu} = F^{\nu}{}_{\mu,\alpha} F^{\nu\alpha} + \frac{1}{2} F_{\alpha\nu}{}^{,\mu} F^{\nu\alpha} \tag{2.60}$$

The Lagrangian of electrodynamics is given by

$$\mathcal{L} = -\frac{1}{2} F_{\mu\nu} F^{\mu\nu} + J_{\mu} A^{\mu} - \frac{1}{2} m^2 A_{\mu} A^{\mu} \,. \tag{2.61}$$

The tensor S is known as *electromagnetic stress–energy tensor* and given in contravariant form as

$$S^{\mu\nu} = \begin{pmatrix} u & \mathbf{S}^{\top} \\ \mathbf{S} & -\sigma_{ij} \end{pmatrix} . \tag{2.62}$$

Where $\sigma_{ij} = E_i E_j + E_i E_j - \frac{1}{2}\delta_{ij}\big(\mathbf{E}^2 + \mathbf{B}^2\big)$ is the MAXWELL *stress tensor*, $\mathbf{S} = \mathbf{E} \times \mathbf{B}$ POYNTING *vector* and energy density $u = \frac{1}{2}\big(\mathbf{E}^2 + \mathbf{B}^2\big)$.

2.5. Lagrangian Formalism

We want to formulate special relativity in terms of a variation principle. Therefore we consider massive particles, i.e. timelike paths. As a postulate we take that the action S is proportional to the proper time (as a generalized “distance”) $\mathrm{d}\tau := \sqrt{-\,\mathrm{d}s^2}$

$$S = -\alpha \int_a^b \mathrm{d}\tau = -\alpha \int_{t_1}^{t^2} \sqrt{1 - \mathbf{v}^2}\,\mathrm{d}t =: \int_{t_1}^{t^2} L \,\mathrm{d}t \,, \tag{2.63}$$

with α a constant, that has to be determined. The Lagrangian L is given by

$$L = -\alpha\sqrt{1 - \mathbf{v}^2} \simeq -\alpha\left(1 - \frac{1}{2}\mathbf{v}^2 + \dots\right) . \tag{2.64}$$

We demand that we recover the classical theory in the limit $\mathbf{v} \to 0$. The the lowest order kinetic term is

$$T = \frac{1}{2}\alpha\mathbf{v}^2 \,. \tag{2.65}$$

Comparing with the classical kinetic energy $T = \frac{1}{2}m\mathbf{v}^2$ yields $\alpha = m$. If we substitute α we recover the Lagrangian of special relativity

$$L_{\mathrm{SR}} = -m\sqrt{1-\mathbf{v}^2} = -m\gamma^{-1} \,. \tag{2.66}$$

We can proceed calculate the generalized momenta

$$p_i = \frac{\partial L}{\partial v_i} = \frac{mv_i}{\sqrt{1-\mathbf{v}^2}} = \gamma m v_i \,. \tag{2.67}$$

The energy can be calculated via the Hamiltonian H

$$E = H = \mathbf{p}\cdot\mathbf{v} - L = \gamma m\mathbf{v}^2 + m\gamma^{-1} = \gamma m \,. \tag{2.68}$$

Expanded in $\mathbf{v}^2$ the energy reads as

$$E = m + \frac{1}{2}m\mathbf{v}^2 + \dots \,. \tag{2.69}$$

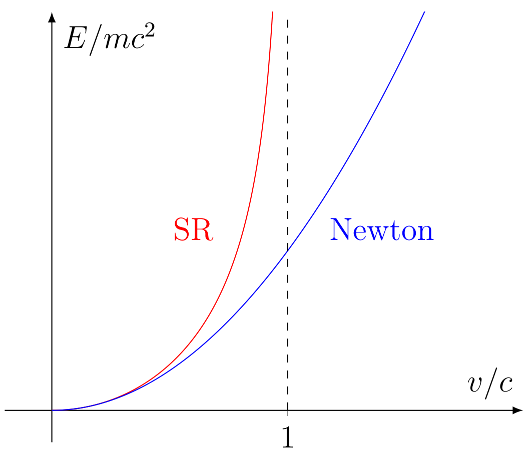


Figure 2.18. – Comparison of the kinetic energies in Newtonian physics and SRT.

If we restore units of c for a moment we get that the constant term is equal to mc^2. This is EINSTEIN's famous $E = mc^2$. We can further relate energy and momentum to each other. Therefore consider the square of the momentum $\mathbf{p}$

$$\mathbf{p}^2 = \frac{m\mathbf{v}^2}{1-\mathbf{v}^2}\,. \tag{2.70}$$

Solving for $\mathbf{v}^2$ yields

$$\mathbf{v}^2 = \frac{\mathbf{p}^2}{\mathbf{p}^2+m^2}\,. \tag{2.71}$$

Which we can insert in the expression for the energy

$$E^2 = \frac{m^2}{1-\mathbf{v}^2} = \mathbf{p}^2 + m^2\,. \tag{2.72}$$

The equation (2.72) is called the *on shell condition*.

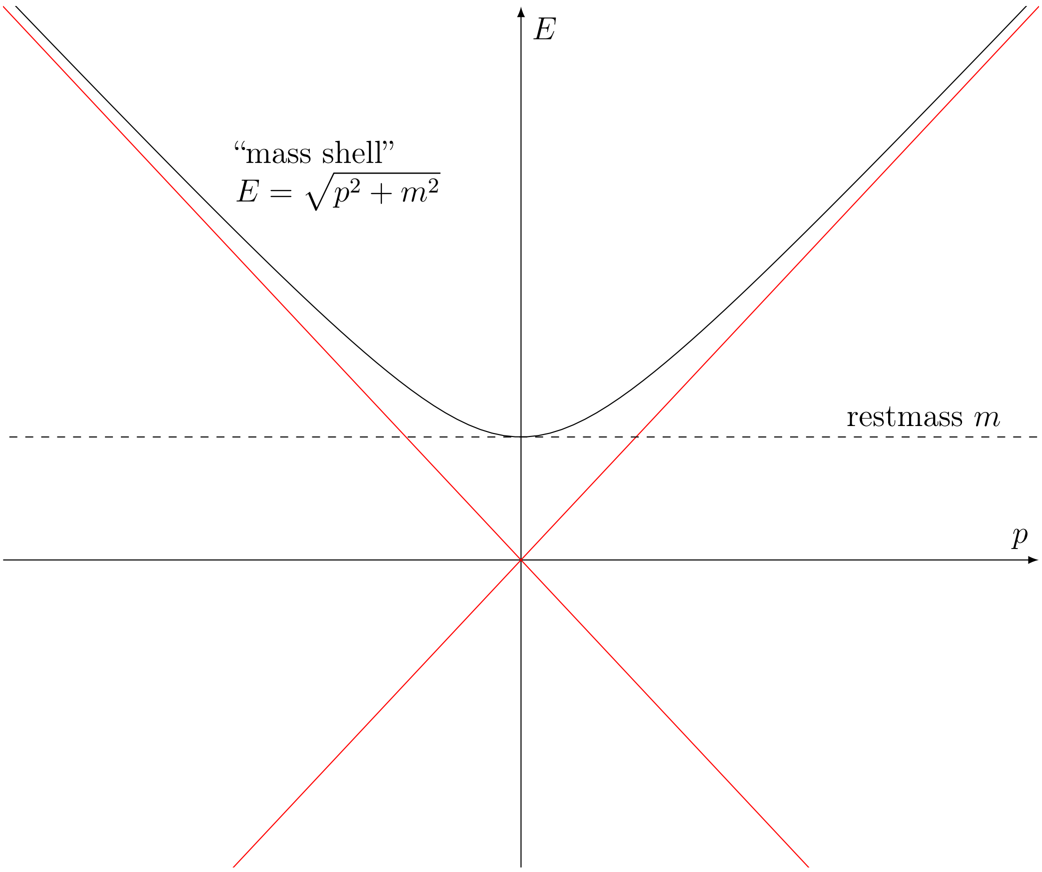


Figure 2.19. – Energy E of a particle with mass m, as a function of the momentum p.

Aside 2 (On massive photons). There is nothing in the theory that predicts that the photon is massless. It could be possible that its mass is only really small and that in fact that the photon does not travel at the speed of light. (Of course this would make it convenient to rename the constant c) Even though the mass should be so small that the de Broglie wavelength is of the order of the size of the universe, massive photons would have a huge impact.

Aside 3 (Complex Time). We could in principle make a substitution $t \to \mathrm{i}t$. Which would free us from the need to distinguish between co- and contravariant vectors, because the inner product would be given as

$$g(x,y) = \left(\mathrm{i}x^0\right)\left(\mathrm{i}y^0\right) + \mathbf{x}\cdot\mathbf{y} = -x^0y^0 + \mathbf{x}\cdot\mathbf{y} \tag{2.73}$$

and hence $g_{ij} = \delta_{ij}$, which basically means we can raise and lower indices at will. This is practical whenever you do calculations, e.g. in computer simulations.

3. Gravity and Geometry

The observable universe is stable. There are two obvious configurations in which this is possible:

1. Static universe, masses are arranged in a grid, all nett forces cancel. However small fluctuations cause the system to collapse therefore this is no possible description for the universe.

2. Expanding universe, all masses move away from each other, overcoming the gravitational attraction. Theoretically such a system can be described by using Newtonian Physics introducing additional energy contributions. This turns out to be inconsistent.

Since in the second description all particles are accelerated relative to each other, there are no inertial systems. A theory in which all observers are equal must therefore be local and thus be described by means of differential geometry. We claim that the laws of physics are the same in every system. If we assume that the MAXWELL's equations are right, the Newtonian theory of gravity must be wrong. Implications: All free falling systems are equivalent (i.e. indistinguishable by the observer). Light must bend, otherwise a beam could be used to deduce whether your system is inertial. The following example illustrates that Euclidean geometry is no adequate description of space-time.

Example 3 (Rotating Sphere). see Introduction to tensor calculus

3.1. Coordinate Systems

We will start by studying coordinate systems in the flat space $\mathbb{R}^2$, which should be familiar.

Cartesian Coordinates

Cartesian coordinates are described by two coordinates x, y that are measured in two orthogonal directions from the origin. The distance s between two arbitrary points

(x_1, y_1) and (x_2, y_2) can be calculated using PYTHAGORAS' theorem

$$s^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 \,. \tag{3.1}$$

An infinitesimal distance is likewise given by

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 \,. \tag{3.2}$$

Polar Coordinates

If we describe a point in flat space by an angle φ and an distance r from the origin, we get polar coordinates. The conversion between the systems reads

$$x = r\cos\varphi \quad y = r\sin\varphi \,. \tag{3.3}$$

A infinitival change in the polar coordinates therefore results in

$$\mathrm{d}x = \frac{\partial x}{\partial r}\,\mathrm{d}r + \frac{\partial x}{\partial \varphi}\,\mathrm{d}\varphi = \cos\varphi\,\mathrm{d}r - r\sin\varphi\,\mathrm{d}\varphi \,, \tag{3.4}$$

$$\mathrm{d}y = \frac{\partial x}{\partial r}\,\mathrm{d}r + \frac{\partial y}{\partial \varphi}\,\mathrm{d}\varphi = \sin\varphi\,\mathrm{d}r + r\cos\varphi\,\mathrm{d}\varphi \,. \tag{3.5}$$

Plugging this into (3.2) gives the line element in polar coordinates

$$\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\,\mathrm{d}\varphi^2 \tag{3.6}$$

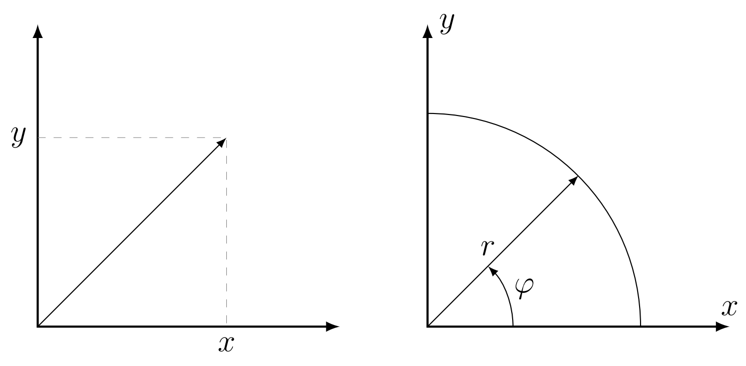


Figure 3.1.

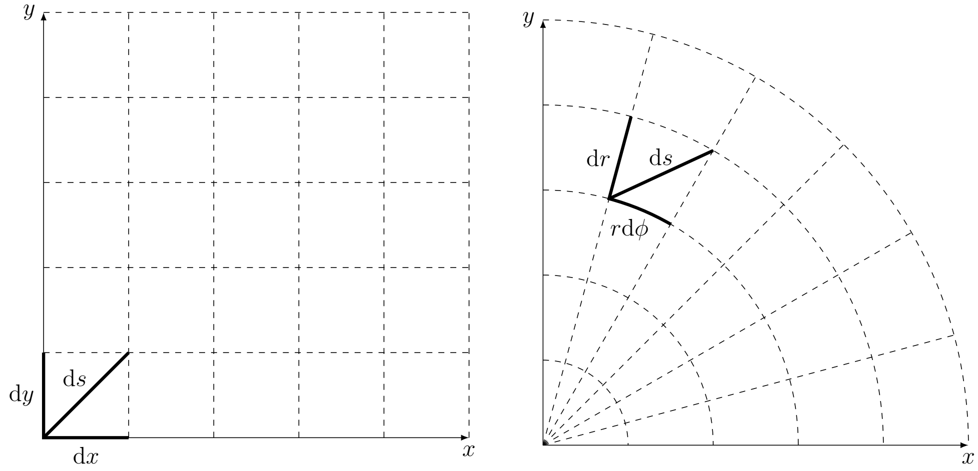


Figure 3.2. – Coordinate grids.

In matrix form

$$\mathrm{d}s^2 = \begin{bmatrix} \mathrm{d}r & \mathrm{d}\varphi \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix} \begin{bmatrix} \mathrm{d}r \\ \mathrm{d}\varphi \end{bmatrix} . \tag{3.7}$$

The matrix

$$g(\mathbf{r}) = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix} , \tag{3.8}$$

is called the *metric*. In general we have

$$\mathrm{d}s^2 = g_{ij}\, \mathrm{d}x^i\, \mathrm{d}x^j . \tag{3.9}$$

The idea is to keep the law of inertia, i.e. particles still move on straight line. However, we need to generalize the concept of a 'straight' line, in a curved space.

3.2. Variation Principle

We know that straight lines are curves minimizing the distance between two points. We generalize this concept to curved space by an variation principle. Again we take a look at flat space, but with curved coordinates. The length S of a curve γ with $\gamma^i(\lambda) = x^i(\lambda)$ is given by the integral

$$S = \int_\gamma \sqrt{\mathrm{d}s^2} = \int_\gamma \sqrt{g_{ij}\, \mathrm{d}x^i\, \mathrm{d}x^j} = \int_a^b \sqrt{g_{ij} \frac{\mathrm{d}x^i}{\mathrm{d}\lambda} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}\, \mathrm{d}\lambda . \tag{3.10}$$

As stated above generalised straight lines satisfy $\delta S = 0$. If we define $L := \left(g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right)^{1/2}$, S takes a form familiar from classical mechanics:

$$S = \int_a^b L\,\mathrm{d}\lambda\,. \tag{3.11}$$

The extremal condition implies the Euler Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial L}{\partial\left(\frac{\partial x^i}{\partial\lambda}\right)} - \frac{\partial L}{\partial x^i} = 0\,. \tag{3.12}$$

We can calculate the relevant terms to

$$\frac{\partial L}{\partial x^i} = \frac{1}{2\sqrt{g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}} g_{jk,i}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\frac{\mathrm{d}x^k}{\mathrm{d}\lambda}\,, \tag{3.13}$$

$$\frac{\partial L}{\partial\left(\frac{\partial x^i}{\partial\lambda}\right)} = \frac{1}{\sqrt{g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}} g_{ji}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\,. \tag{3.14}$$

If we choose the parameter λ so that we are parametrised by the arc length[1] i.e.

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\sqrt{g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}\right) = 0\,, \tag{3.15}$$

the Euler Lagrange equations simplify to

$$0 = \frac{1}{\sqrt{g_{ij}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(g_{ji}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right) - \frac{1}{2\sqrt{g_{ij}\frac{\mathrm{d}x^i}{\mathrm{d}\lambda}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}}} g_{jk,i}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\frac{\mathrm{d}x^k}{\mathrm{d}\lambda}\,, \tag{3.16}$$

or equivalently

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}\lambda}\left(g_{ji}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\right) - \frac{1}{2}g_{jk,i}\frac{\mathrm{d}x^a}{\mathrm{d}\lambda}\frac{\mathrm{d}x^k}{\mathrm{d}\lambda} \\ &= g_{ji,k}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\frac{\mathrm{d}x^k}{\mathrm{d}\lambda} + g_{ji}\frac{\mathrm{d}^2x^j}{\mathrm{d}\lambda^2} - \frac{1}{2}g_{jk,i}\frac{\mathrm{d}x^j}{\mathrm{d}\lambda} \\ &= g_{ji}\frac{\mathrm{d}^2x^j}{\mathrm{d}\lambda^2} + \frac{1}{2}\left(g_{ji,k} + g_{ij,k} - g_{jk,i}\right)\frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\frac{\mathrm{d}x^k}{\mathrm{d}\lambda}\,. \end{aligned} \tag{3.17}$$

[1]this is impossible for null i.e. lightlike geodesics, it can be shown however, that the resulting equation also holds true for null geodesics.

The term invoking derivatives of the metric defines the *Christoffel symbols of the first kind*

$$[jk,\, i] := \frac{1}{2}\Big(g_{ji,k} + g_{ij,k} - g_{jk,i}\Big)\,. \tag{3.18}$$

It is convenient to multiply (3.17) by the inverse metric g^{li} so that we obtain the *geodesic equation*

$$0 = \frac{\mathrm{d}^2 x^l}{\mathrm{d}\lambda^2} + \begin{Bmatrix} l \\ jk \end{Bmatrix} \frac{\mathrm{d}x^j}{\mathrm{d}\lambda}\frac{\mathrm{d}x^k}{\mathrm{d}\lambda}\,. \tag{3.19}$$

Where $\begin{Bmatrix} l \\ jk \end{Bmatrix}$ are the *Christoffel symbols of the second kind*

$$\begin{Bmatrix} l \\ jk \end{Bmatrix} := g^{li}[jk,\, i] = \frac{1}{2} g^{li}\Big(g_{ji,k} + g_{ij,k} - g_{jk,i}\Big)\,. \tag{3.20}$$

In flat space we have $g_{ij} = \eta_{ij}$ and can easily check that all Christoffel symbols vanish. We therefore recover the ordinary equation of motion for a free particle

$$0 = \frac{\mathrm{d}^2 x^i}{\mathrm{d}\lambda^2}\,. \tag{3.21}$$

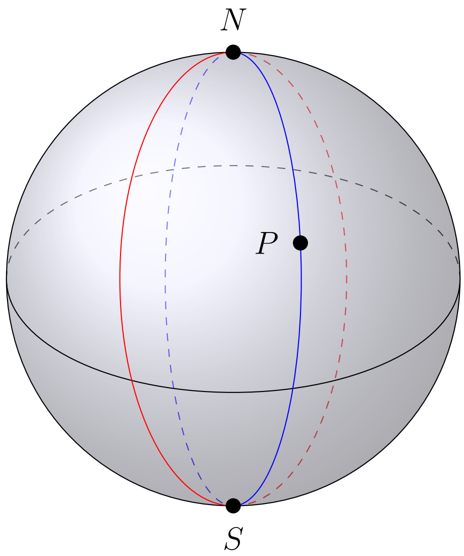


Figure 3.3. – Great circles are geodesics, i.e. shortest connections of points, on a sphere.

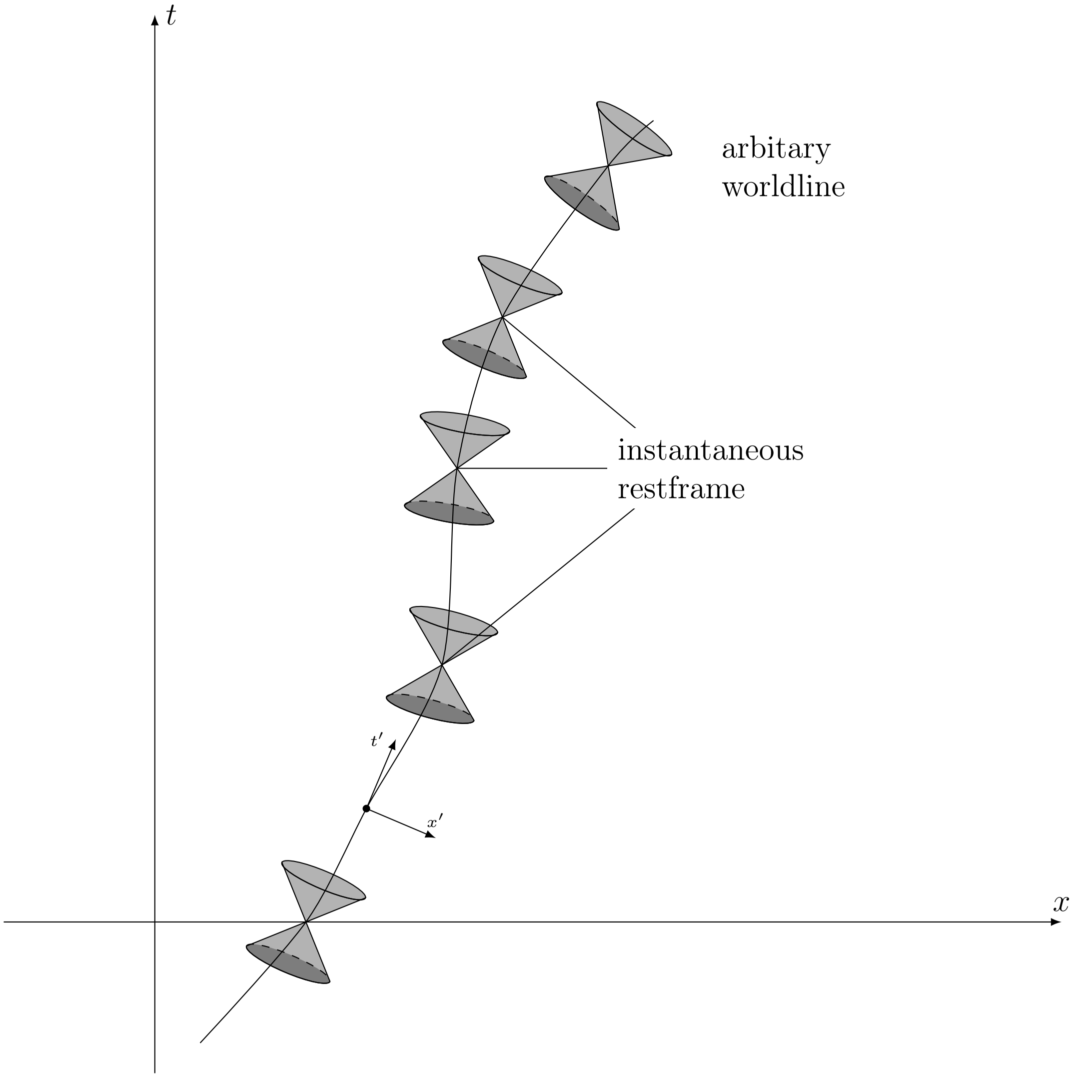


Figure 3.4.

4. Differential Geometry

As we have noted before, general relativity is a inherent local theory. It is convenient to formulate it in terms of differential geometry.

4.1. Manifolds

Definition 5. A n dimensional manifold M is a Hausdorff space with countable basis, that is locally homeomorphic to $\mathbb{R}^n$.

We will give a short introduction to the most important terms.

Remark 2. The requirements Hausdorff and countable basis are of a more technical nature and are satisfied for most of the objects one can imagine except some pathological examples (we won't go into the details on this).

Locally homeomorphic to $\mathbb{R}^n$ means there exists a set of *charts* (φ, U^φ) called an *atlas* $\mathcal{A}$ with $\cup_{\varphi \in \mathcal{A}} U^\varphi = M$, i.e. the charts cover the whole manifold. The maps $\varphi : U^\varphi \to \varphi(U^\varphi) \subset \mathbb{R}^n$ are homeomorphisms, meaning that U^φ is open, φ is onto and both φ and φ^{-1} are continuous. Further for any two $\varphi, \psi \in \mathcal{A}$, the coordinate changes $\varphi \circ \psi^{-1} : \psi(U^\psi \cup U^\varphi) \to \phi(U^\psi \cup U^\varphi)$ be smooth[1].

$$\begin{tikzcd}
& U_{\alpha\beta} = U_\alpha \cap U_\beta \arrow[dl, "h_\alpha"'] \arrow[dr, "h_\beta"] & \\
\mathbb{R}^n \supseteq h_\alpha(U_{\alpha\beta}) \arrow[rr, "h_{\alpha\beta} = h_\beta \circ h_\alpha^{-1}"'] & & h_\beta(U_{\alpha\beta}) \subseteq \mathbb{R}^n
\end{tikzcd}$$

Figure 4.1. – Coordinate change

[1]Infinitely often differentiable or short C^∞.

We can now reduce differentiation on the manifold to the ordinary differentiation in $\mathbb{R}^n$. Since physical laws are described in terms of differential equations, we can formulate them on M. The fact that the coordinate changes are smooth ensures that differentiability is well defined (and thus the physical laws are).

Aside 4 (Differential Structures)*.* There can be different *differential structures* on a manifold, which means there are multiple (maximal)alases, which could not be merged because the coordinate changes would not be C^∞. Those differentiable structures therefore imply different notions of differentiability. Remarkably this may even play a role in some physical theories. As an example an 11d-supergravity can be described as a product $\mathbb{R}^{3+1} \times \mathbb{S}^7$. Where $\mathbb{S}^7$ is the seven-sphere and $\mathbb{R}^{3+1}$ Minkovski space. This means on every point in the $\mathbb{R}^{3+1}$ there is a (small) $\mathbb{S}^7$ located that contains additional spatial dimensions. The $\mathbb{S}^7$ has 28 different differential structures, so the choice of such a structure affects the theory for the above reasons.

All simple examples we come of can be embedded in a higher space. The Whitney embedding theorem states that every real n-dimensional Manifold can be embedded to $\mathbb{R}^{2n}$ (This is however not true for complex, i.e. analytic manifolds). For example the $\mathbb{S}^2$ can be interpreted as submanifold of the $\mathbb{R}^3$. However manifolds are objects that exists independent of such embeddings. For example a torus can be thought of as a square with the opposite sides identified (leaving to the left results in re-entering in the left).

Aside 5 (Topology of the Universe)*.* In addition to the local structure, we may question the global, i.e. the topological structure of the universe. On may for example imagine that we live on the surface of a three-sphere (finite but boundless universe). However this might be observable in cross-corelation in the cosmic microwave background from photons reaching us from different directions but coming from the same event. There is no evidence of such phenomena so far. Most models can be excluded to some certainty. A cylindrical universe is still possible (finite in one, infinite in the other directions).

4.2. Vectors

Vectors are important objects describing physics. The naive view as an 'arrow pointing frow one point to another' is flawed. For example on a sphere an arrow connecting two points does not make much sense. We want to find a description of vectors as objects that are naturally related to the structure of the manifold independent of the embedding.

4.2.1. Definitions

There are three equivalent definitions for a (contravariant) vector:

1. algebraic
2. physical
3. geometrical.

We start by giving the algebraic definition which is the most abstract and preferred by mathematicians, because it is suitable for proofs. Vectors are identified with derivatives, which are formally defined by

Definition 6 (Derivation)**.** A derivation D satisfies the following rules, for all $f, g \in C^\infty(M, \mathbb{R})$ and $\lambda \in \mathbb{R}$:

$$D(af + bg) = Df + Dg\,, \tag{4.1}$$
$$D(\lambda f) = \lambda f\,, \tag{4.2}$$
$$D(fg) = (Df)g + f(Dg)\,. \tag{4.3}$$

We then define a vector by

Definition 7 (Vector, algebraic)**.** A vector in p is a derivation on the germ at p.

The germ is the set of all functions $f \in C^\infty(M, \mathbb{R})$, where we identify all functions that are equal in some neighbourhood of p, i.e. vectors are local objects. Given two vectors we can construct a new one, the *Lie bracket*

$$[X, Y]f := X(Yf) - Y(Xf)\,. \tag{4.4}$$

The only property that has to be checked is that it satisfies the Leibniz rule.

$$XY(fg) = X[(Yf)g + f(Yg)] = (XYf)g + (Yf)(Xg) + (Yg)(Xf) + (XYg)f \tag{4.5}$$

Subtracting $YX(fg)$ proves that $[X, Y]$ is indeed a vector. The fact that we have a natural vector space structure on the set of vectors at p motivates the following

Definition 8. The tangent space T_pM is the space of all vectors in $p \in M$.

A basis of T_pM is given by the derivation along the coordinates ∂_i, therefore its dimension is equal to that of the manifold M. Proof sketch:

1. Show $f(x^i) = f(0) + x^i \tilde{f}(x^i)$
2. Write $X = a^i \partial_i$
3. Show $Xf = 0 \quad \forall f \iff X = 0$

Every vector A can be written as $A = A^i \frac{\partial}{\partial x^i}$, where A^i are the components of the vector. We can now look how the components of the vector transform under a change of coordinates [2]. We usually denote the elements of the transformed systems with a bar. By the chain rule we have

$$A = a^k \frac{\partial}{\partial x^k} = a^k \frac{\partial \overline{x}^i}{\partial x^k} \frac{\partial}{\partial \overline{x}^k} \,. \tag{4.6}$$

We can also express A directly in the new basis

$$A = \overline{a}^i \frac{\partial}{\partial \overline{x}^i} \,. \tag{4.7}$$

Comparing the coefficients gives the vector transformation law

$$\overline{a}^i = a^k \frac{\partial \overline{x}^i}{\partial x^k} \,. \tag{4.8}$$

Definition 9 (Vector, physical)**.** A vector with components A^i is a object that transforms according to 4.8 under a change of coordinates.

Consider a curve on a Manifold M, i.e. a map $\gamma : \mathbb{R} \to M$, with $\gamma(0) = p$, $\dot{\gamma}(0) = X$. Then $D_X f = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \gamma)(0)$ is a derivative, namely the directional derivative along X. Consider the special curves $\gamma_i(t) = \varphi(p + te_i)$, with φ a chart of M. Then $D_{\dot{\gamma}_i} f = \partial_i f$, so $D_{\dot{\gamma}_i}$ represents a the directional derivative and we can relate derivatives to the geometrical tangent space.

Since we have a basis, we can work in (local) coordinates and will do so most of the time.

Example 4 (Lie brackets in local coordinates). Let $A = A^i \partial_i$, $B = B^i \partial_i$ be vectors, then the Lie bracket (4.4) in local coordinates is given as

$$[A, B]^j = A^i \partial_i B^j - B^i \partial_i A^i \,. \tag{4.9}$$

Since the tangent space is a vector space, we can define its dual space

[2]The vector itself is invariant!

Definition 10 (Cotangent space)**.** The cotangent space T_pM^* is the set of linear maps from T_pM to $\mathbb{R}$, i.e.

$$T_pM^* := \{L : T_pM \to \mathbb{R} \,|\, L \text{ linear}\} . \tag{4.10}$$

The cotangent space is again a vector space of the same dimension. Its elements are called *dual* or *covariant* vectors. We can define a basis on T_pM^*, which we denote by $\mathrm{d}x^i$ and which acts on T_pM via

$$\mathrm{d}x^i(\partial_j) = \delta^i_j . \tag{4.11}$$

It can easily deduced by (4.11) that the components of a dual vector transform as

$$\overline{a}_i = \frac{\partial x^k}{\partial \overline{x}^i} a_k . \tag{4.12}$$

Remark 3 (Dual vectors in euclidean space)*.* If $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ contain the components of a vector and a dual vector respectively, then the transformation can be written in matrix form

$$\mathbf{a} \to \overline{\mathbf{a}} = V\mathbf{a} , \tag{4.13}$$

$$\mathbf{b} \to \overline{\mathbf{b}} = \left(V^\top\right)^{-1} \mathbf{b} , \tag{4.14}$$

with $V_{ij} = \dfrac{\partial \overline{x}^i}{\partial x^j}$ the Jacobian of the transformation. In normal calculus we restrict ourselves to orthogonal transformations (i.e. mapping orthonormal bases onto each other) for which $\left(O^\top\right)^{-1} = O$. Which is the reason why we do not bother to distinguish between vectors and dual vectors because they transform identically. In special relativity we have e.g. $\left(\Lambda^\top\right)^{-1} \neq \Lambda$ for a boost and the difference becomes even more important in general relativity where the relation can become arbitrarily complicated.

4.3. Tensors

From vectors A ,B we can construct new objects with multiple indices that posses well defined transformation behaviour. For example consider

$$\overline{T}^{ij} = a^i b^j , \tag{4.15}$$

which transforms as

$$T^{ij} = \frac{\partial \overline{x}^i}{\partial x^k}\frac{\partial \overline{x}^j}{\partial x^l} a^k b^l = \frac{\partial \overline{x}^i}{\partial x^k}\frac{\partial \overline{x}^j}{\partial x^l} T^{kl} \,. \tag{4.16}$$

We call an object that transforms in this way a *tensor*. As with vectors, it is possible to define tensors in a coordinate independent way. At this point we will make things easier and only consider the physical definition, i.e. classify tensors by a transformation according to (4.16).

A tensor is said to be symmetric in two indices if it stays invariant when exchanging those indices, e.g.

$$T_{ab} = T_{ba} \,. \tag{4.17}$$

Remark 4. We have not yet established a relation between upper and lower indices, i.e. we have no metric. Expressions of the form

$$T^a{}_b = T^b{}_a \tag{4.18}$$

therefore make no sense.

4.4. The Metric

So far we have not defined a length scale on manifolds yet. We will do so now by introducing a *metric*

Definition 11 (Metric)**.** A metric g on a manifold M, is a non-degenerate ($\det(g) \neq 0$), symmetric covariant two tensor.

We have already seen examples of metrics for the flat space, e.g. in spherical coordinates g was given as

$$g = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix} . \tag{4.19}$$

A metric that is positive definite is called RIEMANN*ian metric*. In relativity we deal with LORENTZ*ian metrics*, for which there are vectors beside the zero vector which have zero 'length'. In flat space we have $g_{ij} = \eta_{ij}$. A metric gives two natural notions on the tangent space of the manifold. Inner product

$$g(A, B) = g_{ij} A^i B^j \tag{4.20}$$

Pseudo[3] norm

$$g(A,A) = g_{ij}A^iA^j \tag{4.21}$$

Remark 5 (About raising and lowering indices). Suppose we have given a vector $A = A^i\partial_i$ with coordinates A^i, then we can relate it in a natural way to a linear form $A^{\flat 4} := g(A,\cdot)$,

$$\begin{aligned} A^\flat : T_pM &\to \mathbb{R} \\ X &\mapsto g(A,X)\,, \end{aligned} \tag{4.22}$$

i.e. a dual vector. The components of this dual vector are given by its action on the basis elements of the tangent space

$$A_i := (A^\flat)_i = g(A,\partial_i) = A^j g(\partial_j,\partial_i) = g_{ji}A^j\,, \tag{4.23}$$

which is exactly the law for lowering indices. Given the inverse metric we can multiply this equation by it to obtain A^i in terms of A_i.

4.5. Affine Connections

To derive a vector field, we have to relate different tangent spaces. The idea is again to generalize from flat space. The connection is established by introducing a affine connection.

4.5.1. Parallel Transport

We consider a the parallel transport of a vector. If we express a vector in non-Cartesian coordinates and shift it it's coordinates do not change. We take a look on two operations:

1. the change of the vector itself
2. the change of its coordinates.

Let A_i be the coordinates of a vector in a system x^i and B_i in a system associated with coordinates y^i respectively. They are therefore related by

$$A_i = \frac{\partial y^j}{\partial x^i} B_j\,, \quad B_i = \frac{\partial x^j}{\partial y^i} A_j\,. \tag{4.24}$$

[3]Not positive definite.
[4]The symbols $\flat$ and $\sharp$ are borrowed from musical notation.

We look at vectors whose coordinates in the system y^i do not change i.e. $\delta B_i = 0$. The variation of A_i is given by

$$\delta A_i = \delta\left(\frac{\partial y^j}{\partial x^i}\right) B_j = \frac{\partial^2 y^j}{\partial x^i \partial x^k} \delta x^k B_j \,. \tag{4.25}$$

Expressing B_i in terms of A_i yields

$$\delta A_i = \frac{\partial^2 y^j}{\partial x^i \partial x^k} \frac{\partial x^l}{\partial y^j} A_l \delta x^k =: \Gamma^l_{ik} A_l \delta x^k \tag{4.26}$$

Γ^l_{ik} is called *affine connection* or short affinity.

Remark 6. We can always find a coordinate system in which $\Gamma^l_{ik} \equiv 0$, this system is called RIEMANN*ian normal coordinate system* (RNCS).

We notice that if

$$\left(\frac{\partial A_i}{\partial x^j} - \Gamma^l_{ik} A_l\right) \delta x^k = 0 \,, \tag{4.27}$$

The vector A does not change its coordinates. We define a *covariant derivative*

$$A_{i;j} := \nabla_j A_i := \frac{\partial A_i}{\partial x^j} - \Gamma^l_{ik} A_l \,. \tag{4.28}$$

It can easily be seen that the covariant derivative of a tensor transforms as a tensor, by inspecting (4.27) and applying the quotient theorem.

Remark 7 (The Covariant Derivative in Electrodynamics). Example from Electrodynamics concerning the covariant derivative. The theory is invariant under transformations $\phi \to e^{\mathrm{i}\alpha}\phi$, because $\phi^*\phi$ and $\phi^*\nabla\phi - \phi\nabla\phi^*$ do not change. In order to make the Lagrangian gauge invariant we exchange

$$\partial_\mu \to D_\mu + \mathrm{i}A_\mu \,, \tag{4.29}$$

which effectively produces additional terms in the Lagrangian, namely

$$A^\mu\left(\phi^*\partial_\mu\phi - \phi\partial_\mu\phi^*\right) = A^\mu J_\mu \,, \quad A_\mu A^\mu \phi^2 \,. \tag{4.30}$$

The commutator between the covariant derivatives calculates to

$$\left[D_\mu, D_\nu\right] = \mathrm{i}F_{\mu\nu} \,. \tag{4.31}$$

So the noncommutativity is associated with the presence of a field. This is similar to GR where it was related with curvature.

Since we have now established a relation between vectors and dual vectors, we can also determine the covariant derivative of a dual vector. Therefore we consider the scalar $A_i B^i$. Since the covariant derivative satisfies the Leibniz rule we get

$$(A_i B^i)_{;j} = A_{i;j} B^i + A_i B^i{}_{;j} \,. \tag{4.32}$$

But for scalars the covariant derivative is identical to the normal derivative so that

$$(A_i B^i)_{;j} = (A_i B^i)_{,j} = A_{i,j} B^i + A_i B^i{}_{,j} \tag{4.33}$$

If we put in the covariant derivative of a we get

$$A_i B^i{}_{;j} = A_i \Big(B^i{}_{,j} + \Gamma^i_{kj} B^k \Big) \tag{4.34}$$

Since A was arbitrary, we can deduce that

$$B^i{}_{;j} = \Big(B^i{}_{,j} + \Gamma^i_{kj} B^k \Big) \tag{4.35}$$

for a (1,1)-tensor we get:

$$A^i{}_{j;k} = A^i{}_{j,k} - \Gamma^a_{jk} A^i{}_a + \Gamma^i_{ak} A^a{}_j \,. \tag{4.36}$$

Similar expressions hold for tensors of arbitrary rank, where each index gives an additional term containing a contraction with the affinity Γ^i_{jk}. We now want to consider curved spaces. This can not immediately be determined by the metric, for example the polar coordinates do not look flat even though they describe the ordinary $\mathbb{R}^2$.

4.5.2. Geodesics

A curve is a map $\gamma : \mathbb{R} \to M$. The parametrisation is arbitrary e.g.

$$\begin{aligned} \gamma : \mathbb{R} &\to \mathbb{R}^2 \\ t &\mapsto \frac{1}{2} t^2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} , \end{aligned} \tag{4.37}$$

clearly describes a straight line with $y = x$. However, we notice that $\frac{\mathrm{d}\gamma^i}{\mathrm{d}t}$ is parallel to $\frac{\mathrm{d}^2\gamma^i}{\mathrm{d}t^2}$. This gives rise to another possible generalisation of a straight line. In curved coordinates we have

$$\nabla\left(\frac{\mathrm{d}x^j}{\mathrm{d}t}\right) = \frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} + \Gamma^i_{jk}\frac{\mathrm{d}x^j}{\mathrm{d}t}\frac{\mathrm{d}x^k}{\mathrm{d}t} = \lambda(t)\frac{\mathrm{d}x^j}{\mathrm{d}t}\,. \tag{4.38}$$

Suppose we have a different parametrisation $s(t)$

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = \frac{\mathrm{d}x^i}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}t}\,, \quad \frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} = \frac{\mathrm{d}^2x^i}{\mathrm{d}s^2}\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 + \frac{\mathrm{d}x^i}{\mathrm{d}s}\frac{\mathrm{d}^2s}{\mathrm{d}t^2}\,, \tag{4.39}$$

then the equation for a straight line reads as

$$\left(\frac{\mathrm{d}^2x^i}{\mathrm{d}s^2} + \Gamma^i_{jk}\frac{\mathrm{d}x^j}{\mathrm{d}s}\frac{\mathrm{d}x^k}{\mathrm{d}s}\right)\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 + \frac{\mathrm{d}^2s}{\mathrm{d}t^2}\frac{\mathrm{d}x^i}{\mathrm{d}s} = \lambda(t)\frac{\mathrm{d}x^j}{\mathrm{d}t}\frac{\mathrm{d}s}{\mathrm{d}t}\,. \tag{4.40}$$

To get to the usual form of the geodesic equation we choose s so that

$$\frac{\mathrm{d}^2s}{\mathrm{d}t^2} = \lambda(t)\frac{\mathrm{d}s}{\mathrm{d}t} \tag{4.41}$$

Which is an ordinary differential equation of type $\ddot{s} = \lambda\dot{s}$ that should posses a solution. For this special choice of parametrisation we have

$$\frac{\mathrm{d}^2x^i}{\mathrm{d}s^2} + \Gamma^i_{jk}\frac{\mathrm{d}x^j}{\mathrm{d}s}\frac{\mathrm{d}x^k}{\mathrm{d}s} = 0\,. \tag{4.42}$$

We therefore have a preferred set of parameters called affine parameters that satisfy (4.41). The character of the differential equation implies that the affine parameter is only defined modulo affine transformations $s \to as + b$. This freedom once more reflects some kind of gauge invariance.

Remark 8. Only the symmetric part of the affinity does contribute to the geodesic equation (4.42). As long as you are on a geodesic you can always compare lengths without needing a metric, by the affine parameter. The 'length' defined this way does not have to concede with the length given by the metric and is also defined for example for lightlike curves (which cannot be parametrised by the arc length).

If we demand that the two definitions of a geodesic (4.42) is identical to (3.19) coincide, we get a preferred connection. This corresponds to the choice $\Gamma^k_{ij} = \left\{\begin{matrix} k \\ ij \end{matrix}\right\}$. This is called the *metric*, *Levi-Civita* or *Christoffel connection*, and we will always choose it in the following. A general metric compatible connection[5] satisfies

$$\Gamma^k_{ij} = \begin{Bmatrix} k \\ ij \end{Bmatrix} + T^k{}_{ij} + g^{ir}\Big(T^s{}_{jr}\, g_{si} + T^s{}_{ir}\, g_{sj}\Big)\,, \tag{4.43}$$

with the torsion tensor

$$T^k{}_{ij} = \frac{1}{2}\Big(\Gamma^k_{ij} - \Gamma^k_{ji}\Big)\,. \tag{4.44}$$

Therefore the Christoffel connection is the unique metric compatible symmetric connection.

Remark 9. Non-Christoffel connections play a role, for example when dealing with spinors.

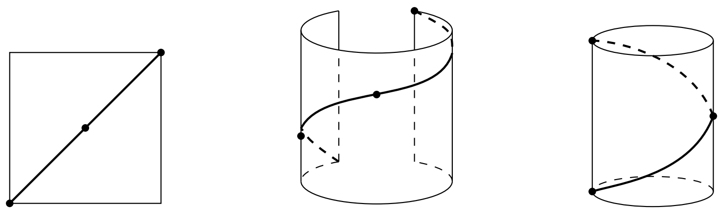


Figure 4.2. – The geodesics on a cylinder can be obtained by "folding" flat space.

4.6. The Riemannian Curvature Tensor

We can contract indices on the Riemann Tensor

$$R^i{}_{ikl} = \frac{\partial \Gamma^i_{il}}{\partial x^k} - \frac{\partial \Gamma^i_{ik}}{\partial x^l}\,, \tag{4.45}$$

which is zero for a metric affinity. The Ricci tensor is given by

$$R_{jk} = R^i{}_{jki}\,. \tag{4.46}$$

Because of the symmetry this are all independent contractions. Notice that at this point we cannot raise or lower indices to contract different indices. Given a metric the curvature or Ricci skalar is defined as

$$R^i{}_i = g^{ij} R_{ji}\,. \tag{4.47}$$

[5] $\nabla_k g_{ij} = 0$

Since we can express the RIEMANN tensor in terms of commutators there is a also a symmetry that can be derived from the Jacobi identity:

$$\big[A[B,C]\big] + \big[C[A,B]\big] + \big[B[C,A]\big] = 0\,. \tag{4.48}$$

Physics can be described in terms of differential equations. We for example would like to have a object similar to the Laplacian in curved coordinates. However $\partial_i\partial_i$ is not coordinate invariant. We therefore introduce a metric. The equivalence principle implies that space is locally Minkovski.

Aside 6. There are certain theories that can be formulated without a metric. An example being three dimensional gravity, which is non-dynamic.

Since the connection and the metric are not related, we still don't have a length scale on the manifold

4.6.1. Symmetries of the Riemann Tensor

$$R^i{}_{jkl} = -R^i{}_{jlk}\,. \tag{4.49}$$

Bianci identities for the Riemann tensor

$$R^i{}_{jkl;m} + R^i{}_{jmk;l} + R^i{}_{jlm;k} = 0 \tag{4.50}$$

Proof: If we would write it all out we wold have to write 22 terms. Instead we use a RNCS so that the Riemann tensor simplifies to

$$R^i{}_{jkl} = \partial_k\Gamma^i_{jl} - \partial_l\Gamma^i_{jk} \tag{4.51}$$

And with $\nabla_k = \partial_k$ we get

$$R^i{}_{jkl;m} = \partial_m R^i{}_{jkl} = \partial_k\partial_m\Gamma^i_{jl} - \partial_l\partial_m\Gamma^i_{jk}\,. \tag{4.52}$$

Plugging in gives the result in the RNCS system, but since the equation is tensorial, it holds in all systems. Notice the similarity to

$$\partial_m F_{ab} = \partial_b\partial_m A^a - \partial_a\partial_m A^b \tag{4.53}$$

Therefore the second set of Maxwell's equations (2.43) is in fact a the Bianchi identity.

If the affinity is symmetrical $\Gamma^k_{ij} = \Gamma^k_{ji}$ we further have the identity

$$R^i{}_{jkl} + R^i{}_{ljk} + R^i{}_{klj} = 0 \tag{4.54}$$

Given a metric g_{ij}, we can raise and lower indices

$$\begin{aligned}
R_{ijkl} &= g_{ia} R^a{}_{jkl} \\
&= g_{ia} \left(\partial_k \Gamma^i_{jl} - \partial_l \Gamma^i_{jk} \right) \\
&= g_{ia} \left(\partial_k g^{as} [jl,\, s] - \partial_l g^{as} [jk,\, s] \right) \\
&= \partial_k [jl,\, i] - \partial_l [jk,\, i] \\
&= \frac{1}{2} \left(\frac{\partial^2 g_{il}}{\partial x^j \partial x^k} + \frac{\partial^2 g_{jk}}{\partial x^i \partial x^l} - \frac{\partial^2 g_{ik}}{\partial x^j \partial x^l} - \frac{\partial^2 g_{jl}}{\partial x^i \partial x^j} \right)
\end{aligned} \tag{4.55}$$

Immediately we can extract symmetries

$$R_{ijkl} = -R_{jikl} \tag{4.56}$$

$$R_{ijkl} = -R_{ijlk} \tag{4.57}$$

If we further introduce multi-indices $A = (i, j)$, $B = (k, l)$ (by antisymmetry there are six independent components for A, B each)

$$R_{AB} = R_{BA} \tag{4.58}$$

So R can be thought as a 6×6 matrix. Table 4.1 lists the number of independent components of the curvature tensor and the Ricci tensor in various dimensions. Thereby a one dimensional space is always flat, a two dimensional is characterised by the curvature scalar R alone and in three dimensions the Ricci tensor is sufficient to know the Riemann tensor. Therefore four is the lowest dimension in which the Riemann tensor contains additional information of the curvature of space.

Table 4.1. – Number of independent components of Riemann and Ricci tensor.

dimension	Riemann tensor	Ricci tensor
1	0	0
2	1	1
3	6	6
4	20	10
n	$\frac{1}{12}n^2(n^2-1)$	$\frac{1}{2}n(n+1)$

5. Einstein's Field Equations

We will derive Einsteins Equations by physical considerations. Remember that the Poisson equation reads

$$\Delta\Phi = 4\pi\varrho\,. \tag{5.1}$$

So matter (energy) is the source of the gravitational field Φ. From SR we know that the energy momentum tensor $T_{\mu\nu}$ is an adequate generalisation of energy. We therefore put $T_{\mu\nu}$ as the right hand side of a yet to be found equation, and ask for the left hand side. We would like a tensor $S_{\mu\nu}$, related to the geometry, so that we can express

$$S_{\mu\nu} = T_{\mu\nu}\ . \tag{5.2}$$

In SR, the energy momentum tensor $T_{\mu\nu}$ is conserved i.e. $T_{\mu\nu}{}^{,\nu} = 0$. As a natural extension, we demand that the energy momentum tensor of general relativity is *covariantly* conserved

$$\nabla^\nu T_{\mu\nu} = T_{\mu\nu}{}^{;\nu} = 0 \tag{5.3}$$

Theorem 1 (Lovelock)**.** *For a four dimensional space* [1] *the most general divergence free tensor* $A_{\mu\nu}$ *is given by*

$$A_{\mu\nu} = c_1 G_{\mu\nu} + c_2 g_{\mu\nu}\ . \tag{5.4}$$

Where $G_{\mu\nu}$ *is the* Einstein tensor $G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R$.

The theorem immediately implies *Einstein's field equations*

$$R_{\mu\nu} - \frac{1}{2}R g_{\mu\nu} - \Lambda g_{\mu\nu} = \kappa T_{\mu\nu}\ , \tag{5.5}$$

with some constants κ, Λ. Of course we identify Einsteins constant $\kappa = \frac{8\pi G_\mathrm{N}}{c^2}$ and the

[1] this does certainly not hold true for $d > 4$

cosmological constant Λ. As a slight variation, we can rewrite equation (5.5) as

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \kappa \left(T_{\mu\nu} - \frac{\Lambda}{\kappa} g_{\mu\nu} \right) \tag{5.6}$$

so that the left hand side represents the geometrical part and the right hand side the matter content and we identify Λ with an vacuum energy. Wheeler condenses this in the statement:

Geometry tells matter how to move, matter tells geometry how to curve.

Aside 7. In the time of the inflation the cosmological constant must have been large. Since it is small today it has do decay with time.

There is also an variational derivation dating back to Hilbert that is simpler than Einsteins initial derivation. We start by considering a general action

$$S_{\mathrm{g}} = \int \mathrm{d}^4 x \, \tilde{\mathcal{L}} \tag{5.7}$$

where $\tilde{\mathcal{L}}$ must transform as a (scalar) density. Therefore we define a scalar $\mathcal{L} = \frac{\tilde{\mathcal{L}}}{\sqrt{-g}}$

$$S_{\mathrm{g}} = \int \mathrm{d}^4 x \, \sqrt{-g} \mathcal{L} \tag{5.8}$$

One can think of various contributions to $\mathcal{L}$, e.g.

$$R, \, \Box R, \, \nabla^\mu \nabla^\mu R_{\mu\nu}, \, R_{\mu\nu} R^{\mu\nu}, \, R_{\mu\nu\sigma\varrho} R^{\mu\nu\sigma\varrho} \dots ,$$

which are contractions, so that the resulting quantity becomes a scalar. We have no contributions of the metric alone, because $g_{\mu\nu;\sigma} = 0$. From Yang-Mills theory one would expect a structure

$$\mathcal{L} \sim F_{\mu\nu} F^{\mu\nu} , \tag{5.9}$$

but Γ is not the fundamental field but g is. If we demand that we only have up to second derivatives of g the only allowed term in the Lagrangian is R.

Aside 8 (On higher derivatives). If we include higher order derivatives of g in the right way we can make the resulting theory renormalizable. However we violate unitarity and introduce so-called ghost fields which are associated with the additional degrees of freedom we get.

Remark 10 (Dimensions). In natural units[2] the line element $\mathrm{d}s^2$ has dimension, $[\mathrm{d}s^2] = \mathrm{M}^{-2}$.[3] Since further $[x^\mu] = \mathrm{M}^{-1}$, the Lagrange density must have Dimension $[\mathcal{L}] = \mathrm{M}^4$.

This constraint leads to the *Einstein-Hilbert-action*

$$S_{\mathrm{EH}} = \frac{1}{2\kappa} \int \mathrm{d}^4x \, \sqrt{-g}(R - 2\Lambda) \tag{5.10}$$

We now check that its variation indeed reproduces Einstein's equations. To do so we introduce the formalism of *functional derivation*. Let therefore $\Phi = \{\varphi, A^\mu, \Psi, \dots\}$ be a collection of fields. $F[\Phi]$ a functional. We define the variation of F as

$$\delta F := \int \mathrm{d}x \, \frac{\delta F}{\delta \Phi^i} \delta \Phi^i \, . \tag{5.11}$$

Typically the functionals are given in the form

$$S[\Phi] = \int \mathrm{d}x \, L(x, \Phi) \, , \tag{5.12}$$

where L is some local function.

$$\frac{\delta g_{\varrho\sigma}(x)}{\delta g_{\mu\nu}(x')} = \delta^{\mu\nu}_{\varrho\sigma} \delta(x, x') \tag{5.13}$$

Where $\delta^{\mu\nu}_{\varrho\sigma} = \frac{1}{2}\left(\delta^\nu_\varrho \delta^\mu_\sigma + \delta^\mu_\varrho \delta^\nu_\sigma\right)$ is the unity of the space of symmetric rank two tensors

Remark 11. In general $\delta(x, x') \neq \delta(x - x')$

We transform to the origin of an RNCS, so that the Christoffel symbols vanish. In that coordinate system the covariant and the partial derivative coincide: $\partial_\mu = \nabla_\mu$.

$$\begin{aligned} \delta R^\rho{}_{\mu\nu\sigma} &= \delta \partial_\nu \begin{Bmatrix} \rho \\ \mu\sigma \end{Bmatrix} - \delta \partial_\mu \begin{Bmatrix} \rho \\ \nu\sigma \end{Bmatrix} \\ &= \partial_\nu \delta \begin{Bmatrix} \rho \\ \mu\sigma \end{Bmatrix} - \partial_\mu \delta \begin{Bmatrix} \rho \\ \nu\sigma \end{Bmatrix} \end{aligned} \tag{5.14}$$

[2] So that length has dimension of inverse mass.
[3] Where M refers to the dimension of mass.

Attention: i.A. $\delta\partial_\mu \neq \partial_\mu \delta$

$$
\begin{aligned}
\delta R_{\mu\nu} &= \delta R^{\rho}{}_{\mu\rho\nu} \\
&= \partial_\rho \delta \begin{Bmatrix} \rho \\ \mu\nu \end{Bmatrix} - \partial_\mu \delta \begin{Bmatrix} \rho \\ \rho\nu \end{Bmatrix} \\
&= \nabla_\rho \delta \begin{Bmatrix} \rho \\ \mu\nu \end{Bmatrix} - \nabla_\mu \delta \begin{Bmatrix} \rho \\ \rho\nu \end{Bmatrix}
\end{aligned}
\tag{5.15}
$$

This holds in a general frame since it is a tensor equation.

$$
\begin{aligned}
\delta R &= \delta\Big(g^{\mu\nu} R_{\mu\nu}\Big) \\
&= R_{\mu\nu}\delta g^{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu}
\end{aligned}
\tag{5.16}
$$

Use (A.1)

$$
\begin{aligned}
2\kappa\delta S_{\mathrm{EH}} &= \int \mathrm{d}^4x \left[(R-2\Lambda)\delta\sqrt{-g} + \sqrt{-g}\delta R\right] \\
&= \int \mathrm{d}^4x \left[\frac{1}{2}\sqrt{-g}g^{\mu\nu}\delta g_{\mu\nu}(R-2\Lambda) + \sqrt{-g}\Big(R_{\mu\nu}\delta g^{\mu\nu} + g^{\mu\nu}\delta R_{\mu\nu}\Big)\right] \\
&= \int \mathrm{d}^4x \, \sqrt{-g}\left[\frac{1}{2}g^{\mu\nu}(R-2\Lambda) + R^{\mu\nu}\right]\delta g_{\mu\nu} + \int \mathrm{d}^4x \, \sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu}
\end{aligned}
\tag{5.17}
$$

We treat both occurring terms separately

$$
\begin{aligned}
\int \mathrm{d}^4x \, \sqrt{-g}g^{\mu\nu}\delta R_{\mu\nu} &= \int \mathrm{d}^4x \, \sqrt{-g}g^{\mu\nu}\left(\nabla_\rho \delta \begin{Bmatrix} \rho \\ \mu\nu \end{Bmatrix} - \nabla_\mu \delta \begin{Bmatrix} \rho \\ \rho\nu \end{Bmatrix}\right) \\
&= \int \mathrm{d}^4x \, \nabla_\rho \left(\sqrt{-g}g^{\mu\nu}\delta \begin{Bmatrix} \rho \\ \mu\nu \end{Bmatrix}\right) \\
&\quad - \int \mathrm{d}^4x \nabla_\mu \left(\sqrt{-g}g^{\mu\nu}\delta \begin{Bmatrix} \rho \\ \rho\nu \end{Bmatrix}\right)
\end{aligned}
\tag{5.18}
$$

The Integrals vanish by GAUSS law (neglecting surface Terms). We are left with

$$
2\kappa\delta S_{\mathrm{EH}} = \int \mathrm{d}^4x \, \sqrt{-g}\left[\frac{1}{2}g^{\mu\nu}(R-2\Lambda) + R^{\mu\nu}\right]\delta g_{\mu\nu}
\tag{5.19}
$$

So that we can now finally calculate the variation with respect to the metric field

$$\frac{\delta S_{\mathrm{EH}}[g_{\mu\nu}(x)]}{\delta g_{\mu\nu}(x')} = \sqrt{-g}\left[\frac{1}{2}g^{\mu\nu}(R-2\Lambda)+R^{\mu\nu}\right] \tag{5.20}$$

It vanishes if

$$R^{\mu\nu}+\frac{R}{2}g^{\mu\nu}-\Lambda g^{\mu\nu}=0\,. \tag{5.21}$$

so we have finally derived Einsteins field equations from an variational principle.

5.1. Introduction of Matter

In the gravitational context we mean be *matter* any non gravitational fields this include scalar fields φ, spinor fields Ψ, gauge fields A^{μ},.... We collect all of them in a multivariable Φ. A local action can be written as

$$S_{\mathrm{m}}[\Phi,g] = \int \mathrm{d}^4x\,\sqrt{-g}L_{\mathrm{m}}\Big(\Phi,\nabla_{\mu}\Phi,g\Big) \tag{5.22}$$

$g^{\mu\nu}$ appears in L_{m} because the derivatives $\nabla_{\mu},\partial_{\mu}$ must be contracted. Additionally it enters via $\sqrt{-g}$.

Example 5 (Free scalar field). The action of a free scalar field in Minkovski space has the form

$$S_{\mathrm{m}} = \int \mathrm{d}^4x\left(-\frac{1}{2}\eta^{\mu\nu}\partial_{\mu}\varphi\partial_{\nu}\varphi-\frac{1}{2}m^2\varphi^2\right). \tag{5.23}$$

The minus sign in front of the partial derivative should come as no surprise since we have $\eta^{00}=-1\,\dot{\varphi}^2>0$. In a non inertial frame we have to make the usual replacements

$$\eta_{\mu\nu}\to g_{\mu\nu}\,,\quad \partial_{\mu}\to\nabla_{\mu}\,,\quad \mathrm{d}^4x\to\mathrm{d}^4x\,\sqrt{-g}\,, \tag{5.24}$$

which is also known as a *minimal coupling description.* The action for a scalar φ in the presence of gravity, i.e. a dynamical $g_{\mu\nu}(x)$, reads as

$$S_{\mathrm{m}} = \int \mathrm{d}^4x\,\sqrt{-g}\left(-\frac{1}{2}g^{\mu\nu}\nabla_{\mu}\varphi\nabla_{\nu}\varphi-\frac{1}{2}m^2\varphi^2\right). \tag{5.25}$$

The combined action of scalar field and gravity is given as

$$S[g,\varphi]=S_{\mathrm{g}}[g]+S_{\mathrm{m}}[g,\varphi]\,. \tag{5.26}$$

The variation with respect to the field φ is

$$\begin{aligned}\frac{\delta S[g,\varphi]}{\delta\varphi(x')} &= \frac{\delta S_{\mathrm{m}}[g,\varphi]}{\delta\varphi(x')} \\ &= \int \mathrm{d}^4x\,\sqrt{-g}\left[-g^{\mu\nu}\nabla_\mu\varphi\nabla_\nu\left(\frac{\delta\varphi(x)}{\delta\varphi(x')}\right) - m^2\varphi\frac{\delta\varphi(x)}{\delta\varphi(x')}\right],\end{aligned} \tag{5.27}$$

where we used that the δ and ∇_μ commute. Partial integration yields

$$\begin{aligned}\frac{\delta S_{\mathrm{m}}[g,\varphi]}{\delta\varphi(x')} &= \int \mathrm{d}^4x\,\sqrt{-g}\big(\Box_g - m^2\big)\varphi\delta(x,x') \\ &= \sqrt{-g}\big(\Box_g - m^2\big)\varphi\,,\end{aligned} \tag{5.28}$$

with $\Box_g := g^{\mu\nu}\nabla_\mu\nabla_\nu$ the LAPLACE–BELTRAMI *operator*, a generalisation of the ordinary Laplacian to curved space. Demanding that the variation with respect to φ vanishes implies the *Klein-Gordon equation*

$$\big(\Box_g - m^2\big)\varphi = 0 \tag{5.29}$$

We can also vary the action with respect to the metric field $g_{\mu\nu}$ resulting in

$$\frac{\delta S[g,\varphi]}{\delta g_{\mu\nu}(x')} = \frac{\delta S_{\mathrm{g}}[g]}{\delta g_{\mu\nu}(x')} + \frac{\delta S_{\mathrm{m}}[g,\varphi]}{\delta g_{\mu\nu}(x')} = \frac{\sqrt{-g}}{2\kappa}(G^{\mu\nu} + \Lambda g^{\mu\nu}) + \frac{\delta S_{\mathrm{m}}[g,\varphi]}{\delta g_{\mu\nu}(x')}\,. \tag{5.30}$$

To recover the Einstein equations it is convenient to define the energy-momentum tensor

$$T^{\mu\nu} := -\frac{2}{\sqrt{-g}}\frac{\delta S_{\mathrm{m}}[g,\varphi]}{\delta g_{\mu\nu}(x')}\,. \tag{5.31}$$

We can now proceed in calculating the quantity we have just introduced for a scalar field

$$\begin{aligned}\frac{\delta S_{\mathrm{m}}[g,\varphi]}{\delta g_{\mu\nu}(x')} &= \int \mathrm{d}^4x\frac{\delta\sqrt{-g}}{\delta g_{\mu\nu}}\left(-\frac{1}{2}g^{\mu\nu}\nabla_\mu\varphi\nabla_\nu\varphi - \frac{1}{2}m^2\varphi^2\right) \\ &\quad + \sqrt{-g}\left(-\frac{1}{2}g^{\alpha\varrho}g^{\beta\sigma}\nabla_\alpha\varphi\nabla_\beta\varphi\frac{\delta g_{\varrho\sigma}}{\delta g_{\mu\nu}}\right) \\ &= \frac{1}{2}\int \mathrm{d}^4x\sqrt{-g}\left(-\frac{1}{2}g^{\mu\nu}\nabla_\varrho\varphi\nabla^\varrho\varphi - \frac{1}{2}g^{\mu\nu}m^2\varphi^2 + \nabla^\mu\varphi\nabla^\nu\varphi\right)\delta(x,x') \\ &= \frac{1}{2}\sqrt{-g}\left(-\frac{1}{2}g^{\mu\nu}\nabla_\varrho\varphi\nabla^\varrho\varphi - \frac{1}{2}g^{\mu\nu}m^2\varphi^2 + \nabla^\mu\varphi\nabla^\nu\varphi\right)\end{aligned} \tag{5.32}$$

So that

$$T^{\mu\nu}(\varphi) = -\frac{1}{2}g^{\mu\nu}\nabla_\varrho\varphi\nabla^\varrho\varphi + \nabla^\mu\varphi\nabla^\nu\varphi - \frac{1}{2}g^{\mu\nu}m^2\varphi^2 \tag{5.33}$$

As we have noticed, the Einstein Tensor is covariantly conserved (contracted Bianci identities). The Einstein equation then implies that also ${T^{\mu\nu}}_{;\nu} = 0$ this can be checked for the given Tensor

$$\begin{aligned}\nabla_\mu T^{\mu\nu} &= g^{\mu\nu}\nabla_\mu\nabla_\varrho\varphi\nabla^\varrho\varphi + \Box\varphi\nabla^\nu\varphi + \nabla^\mu\varphi\nabla_\mu\nabla^\nu\varphi - g^{\mu\nu}m^2\varphi\nabla_\mu\varphi \\ &= \nabla^\nu\varphi\big(\Box - m^2\big)\varphi \\ &= 0\end{aligned} \tag{5.34}$$

Where the last equality holds because φ satisfies the Klein-Gordon equation.

The Einstein equations are ten quasi linear, i.e. the highest order derivative appears only linear, differential equations for the metric field $g_{\mu\nu}$. Strictly speaking the Einstein equations are *nonlinear*. How do we find a solution to this equations?

1. Prescribe $T_{\mu\nu}$. This is only possible for high symmetry problems, e.g. the Schwazschild solution and the cosmological solutions (Friedmann's equations)

2. Assume $g_{\mu\nu}$, then compute $T_{\mu\nu}$ and (try!) to interpret this.

5.1.1. ADM-Decomposition

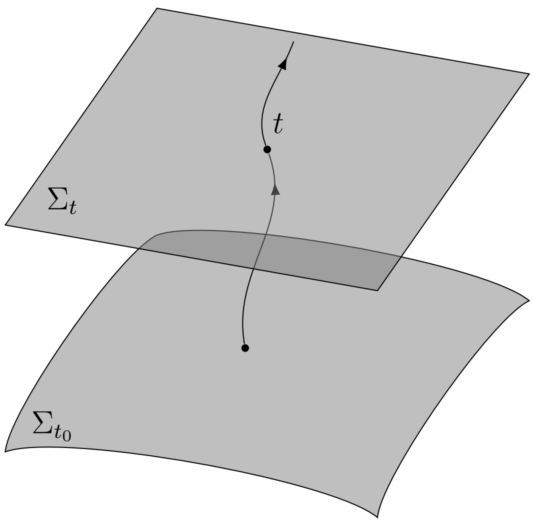


Figure 5.1. – Foliation of spacetime into spatial hypersurfaces Σ_t.

The formulation of initial value problems is not as easy as it is in classical physics.[4] Assume we know either

- $g_{\mu\nu}$ on Σ_{t_0}
- $g_{\mu\nu;j}$, $g_{\mu\nu;0}$ on Σ_{t_0}

We then see the spacetime as a collection of spacelike hypersurfaces at time t $\Sigma_t = \{x^0 = t\}$. For simplicity we consider a vacuum solution to the Einstein equations, i.e.

$$0 = G = R - 2R \implies R_{\mu\nu} = 0\,. \tag{5.35}$$

Divided into the respective parts the field equations are

$$0 = R_{00} = -\frac{1}{2}g^{ij}g_{ij,00} + M_{00}\,, \tag{5.36}$$
$$0 = R_{0i} = -\frac{1}{2}g^{0j}g_{ij,00} + M_{0i}\,, \tag{5.37}$$
$$0 = R_{ij} = -\frac{1}{2}g^{00}g_{ij,00} + M_{ij}\,. \tag{5.38}$$

Where $M_{\mu\nu}$ is a rest term containing lower order time derivatives. This shows that there are no second order time derivatives of $g_{0\mu}$. We have 10 equations and 6 undetermined functions. The DOFs can be used for a coordinate transformation, so that $g_{0\mu,00} = 0$ on Σ_{t_0}. This is always possible but we will not proof this. It can be further shown, by means of the contracted Bianci identities, that this implies $g_{0\mu,00} = 0$ on *all* hypersurfaces Σ_t. Since we have to much freedom the solution will not be unique. We have the freedom to choose four coordinates

$$x^{\mu'} = f^{\mu'}(x^\mu)\,. \tag{5.39}$$

One typical choice is the *harmonic*[5] *gauge*

$$\Box x^\mu = 0\,. \tag{5.40}$$

[4] In fact, depending on the setting, a well defined formulation can be impossible.
[5] A function f is said to be harmonic if it satisfies $\Box f = 0$.

We can expand the d'Alembertian, using (A.7), to

$$\begin{aligned}\Box x^\mu &= g^{-1/2}\partial_\varrho\left(g^{1/2}g^{\varrho\sigma}\partial_\sigma x^\mu\right)\\ &= g^{-1/2}\partial_\varrho\left(g^{1/2}g^{\varrho\sigma}\delta_\sigma{}^\mu\right)\\ &= g^{-1/2}\partial_\varrho\left(g^{1/2}g^{\varrho\mu}\right),\end{aligned} \tag{5.41}$$

the harmonic gauge is therefore equivalent to

$$\partial_\varrho\left(g^{1/2}g^{\varrho\mu}\right) = 0\,. \tag{5.42}$$

The equation can be divided into spatial and time components and derive by the zero component, so that

$$\partial_0^2\left(g^{1/2}g^{0\mu}\right) = -\partial_i\left[\partial_0\left(g^{1/2}g^{0\mu}\right)\right], \tag{5.43}$$

which fixes the second order time derivatives of the relevant components $g^{0\mu}$. Therefore now the time evolution can be solved.

Degrees of freedom

- **10** componnents for every spacetime point from the symmetric $g_{\mu\nu}$
- **-4** from the constraint equation $G_{\mu\nu}{}^{;\nu} = 0$
  - $G^{00} = \kappa T^{00}$ ensures that the evolution is independent of the choice of spatial coordinates on Σ_{t_0}.
  - $G^{i0} = \kappa T^{i0}$ ensures that the time evolution is independent of the way we foliated spacetime into spacial hypersurfaces Σ_t.
- **-4** due to the freedom to choose coordinates (i.e. a gauge).

We are left with two physical degrees of freedom which may be interpreted as the polarisation states of the graviton field.

Comparison with electrodynamics in flat spacetime

In electrodynamics instead of EINSTEIN's equations we have the field equations for the four potential A_μ:

$$\Box A_\mu - \partial_\mu(\partial_\nu A^\nu) = 0\,. \tag{5.44}$$

As we did for the gravitational field, we take a look at the zero component. We find

$$\begin{aligned} 0 &= -\partial_0^2 A_0 + \partial_i \partial^i A_0 - \partial_0 \Big(-\partial_0 A_0 + \partial_i A^i \Big) \\ &= \partial_i \partial^i A_0 - \partial_0 \partial_i A^i \end{aligned} \tag{5.45}$$

This equation is equivalent to $\nabla \mathbf{E} = 0$ and the Bianchi identities. So once again A_0 is *not* determined by the dynamical evolution equation because there is no second order time derivative analogous to g_{00}. Since A_0 is not determined and cannot be specified on the initial time slice. This reflects some internal redundancy namely gauge invariance of the theory. For any scalar function Λ, the transformation

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \Lambda \,, \tag{5.46}$$

leaves the physics invariant. It is trivial to check that the field strength tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ stays invariant. Perhaps more interesting the field equation is also gauge invariant:

$$\begin{aligned} \Box A'_\mu - \partial_\mu \big(\partial_\nu A^{\nu\prime} \big) &= \Box A_\mu + \Box \partial_\mu \Lambda - \partial_\mu (\partial_\nu A^\nu) - \partial_\mu \Box \Lambda \\ &= \Box A_\mu - \partial_\mu (\partial_\nu A^\nu) \,. \end{aligned} \tag{5.47}$$

Thus if A_μ is a solution to the field equation A'_μ is and therefore both are physically indistinguishable. We can also fix a gauge for example the *Lorentz gauge*:

$$\partial_\mu A^\mu = 0 \,. \tag{5.48}$$

If we derive this by the zero component we get

$$\partial_0^2 A^0 = -\partial_i \partial_0 A^i \,, \tag{5.49}$$

so as with g_{00} the evolution of the zero component is now related to the other components. There is still one residual gauge condition, namely we can still transform

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu \Lambda \,, \tag{5.50}$$

but to keep the gauge, we have to demand that $\Box \Lambda = 0$. Again we count the DOFs:

4 components of the potential A_μ.

-1 from constraint $\nabla \mathbf{E} = 0$.

-1 from gauge freedom Λ.

This leaves two physical degrees of freedom, the polarisation states of a photon.

Remark 12. As we have have seen there is a direct correspondence between the gauge freedom in electrodynamics and the freedom of choice of coordinates of coordinates in GR.

6. The Energy Momentum Tensor

In special relativity we have seen that the energy momentum tensor $T^{\mu\nu}$ is conserved or divergence free respectively, i.e.

$$\partial_\mu T^{\mu\nu} = 0\,. \tag{6.1}$$

The problem we are faced in general relativity is the very definition of local energy. Because gravity surely contributes to the energy, a problem arises as we can always transform to local flat space. We start by revisiting the example of dust

Example 6 (Dust).

$$T^{\mu\nu} = \rho_0 u^\mu u^\nu\,. \tag{6.2}$$

In special relativity: $u^\mu = \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} = \gamma(1, \mathbf{v})^\top$, $T^{00} = \rho_0 \left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 = \gamma^2\rho_0 := \rho$. Where ρ is the density with respect to an observer at rest. For the Volume we have $V = \gamma^{-1}V_0$. For the Energy $E = \gamma\omega_0$. Then the density is given by $\rho = \frac{E}{V} = \gamma^2\rho_0$. The conservation of $T^{0\nu}$ implies

$$\begin{aligned} 0 &= T^{0\nu}{}_{,\nu} \\ &= T^{00}{}_{,0} + T^{0i}{}_{,i} \\ &= \partial_t\left(\rho_0\gamma^2\right) + \partial_i\left(\rho_0\gamma^2 v^i\right) \\ &= \partial_t\rho + \partial_i\left(\rho v^i\right) \\ &= \dot\rho + \boldsymbol{\nabla}(\rho\mathbf{v})\,, \end{aligned} \tag{6.3}$$

the *continuity equation*. For the remaining spatial components we get

$$\begin{aligned} 0 &= T^{i\nu}{}_{,\nu} \\ &= T^{i0}{}_{,0} + T^{ji}{}_{,i} \\ &= \dot\rho v^i + \rho \dot v^i + v^i \partial_j \left(\rho v^j\right) + \dot v^j \rho \partial_j v^i \\ &= v^i \left[\dot\rho + \partial_j \left(\rho \dot v^j\right)\right] + \rho\left(\dot v^i + \dot v^j \partial_j v^i\right) \\ &= \rho\left(\dot v^i + \dot v^j \partial_j v^i\right), \end{aligned} \tag{6.4}$$

the *Euler equation* for vanishing pressure (which was the key assumption for dust). It is natural to generalize equation (6.1) to curved space

$$T^{\mu\nu}{}_{;\nu} = 0\,. \tag{6.5}$$

In expanded form

$$\begin{aligned} 0 &= \nabla_\nu (\rho_0 u^\mu u^\nu) \\ &= \rho_{0;\nu} u^\mu u^\nu + \rho_0 u^\mu{}_{;\nu} u^\nu + \rho_0 u^\mu u^\nu{}_{;\nu} \\ &= u^\mu \left(\rho_0 u^\nu{}_{;\nu} + \rho_{0;\nu} u^\nu\right) + \rho_0 u^\mu{}_{;\nu} u^\nu \end{aligned} \tag{6.6}$$

We multiply both sides with u_μ

$$\begin{aligned} 0 &= -\left(\rho_0 u^\nu{}_{;\nu} + \rho_{0;\nu} u^\nu\right) + \rho_0 u^\nu u_\mu u^\mu{}_{;\nu} \\ &= -\left(\rho_0 u^\nu{}_{;\nu} + \rho_{0;\nu} u^\nu\right) \end{aligned} \tag{6.7}$$

If we plugg this back into equation (6.6) we get

$$\begin{aligned} 0 &= u^\nu \nabla_\nu u^\mu \\ &= u^\nu \partial_\nu u^\mu + u^\nu \begin{Bmatrix} \mu \\ \nu\rho \end{Bmatrix} u^\rho \\ &= \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \frac{\partial u^\mu}{\partial x^\nu} + \begin{Bmatrix} \mu \\ \nu\rho \end{Bmatrix} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \\ &= \frac{\partial^2 x^\mu}{\partial \tau^2} + \begin{Bmatrix} \mu \\ \nu\rho \end{Bmatrix} \frac{\mathrm{d}x^\nu}{\mathrm{d}\nu} \frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \end{aligned} \tag{6.8}$$

The geodesic equation (3.19)

Remark 13. This is a difference between electrodynamics and general relativity; dust moves on geodesics, i.e. the path is determined by the field equations alone. In contrast in electrodynamics an additional Force (Lorentz force) has to be *postulated* to describe the motion of test particles. The case is not settled however e.g. it is unclear whether the paths of spin particles is also determined by the field equations.

7. Linearized theory and Newtonian limit

7.1. Linearized theory

Consider a weak gravitational field. Then we can split the full spacetime metric $g_{\mu\nu}(x)$ into two parts.

Definition 12 (Linearization of the metric field.)**.**

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x) + \mathcal{O}(h^2)\,. \tag{7.1}$$

Thereby $h_{\mu\nu}$ is the flat, constant "background" metric of Minkowski space, i.e. there is no gravitational field present. The field $h_{\mu\nu}(x)$ can be interpreted as a perturbation on the fixed background $\eta_{\mu\nu}$. One can identify a spin-2 particle, the so-called *graviton*, with the excitations (quantized fluctuations) of this field. Because only the linear order of h is considered, the nonlinearity of Einstein's equations is lost. We can raise and lower indices with $h_{\mu\nu}$ and $h^{\mu\nu}$.

Remark 14. This works only for a weak gravitational field, since a strong gravitational field produces a strong back reaction of "matter" on the geometry, which follows from the nonlinearity of Einstein's equations. Exactly this back reaction is neglected in the linearized theory.

7.1.1. Derivation of the linearized Einstein's equations

In the following we neglect all terms with $\mathcal{O}(h^2)$. Our goal is to express Einstein's field equations in the linearized approximation. For this we need to calculate the Christoffel symbols, the Riemann tensor, the Ricci tensor and the Ricci scalar.

Christoffel symbols

$$[\mu\nu,\, \varrho] = \frac{1}{2}\Big(h_{\mu\varrho,\nu} + h_{\nu\varrho,\mu} - h_{\mu\nu,\varrho}\Big) + \mathcal{O}(h^2)\,. \tag{7.2}$$

Christoffel symbols of the second kind:

$$\begin{Bmatrix} \varrho \\ \mu\nu \end{Bmatrix} = g^{\varrho\sigma}[\mu\nu,\, \sigma] = \eta^{\varrho\sigma}[\mu\nu,\, \sigma] + \mathcal{O}(h^2) = \frac{1}{2}\Big(h_{\mu}{}^{\varrho}{}_{,\nu} + h_{\nu}{}^{\varrho}{}_{,\mu} - h_{\mu\nu}{}^{,\varrho}\Big) + \mathcal{O}(h^2)\,. \tag{7.3}$$

Riemann tensor

The Riemann tensor can be calculated to

$$\begin{aligned} R^{\varrho}{}_{\sigma\mu\nu} &= \partial_\mu \begin{Bmatrix} \varrho \\ \nu\sigma \end{Bmatrix} - \partial_\nu \begin{Bmatrix} \varrho \\ \mu\sigma \end{Bmatrix} + \underbrace{\begin{Bmatrix} \varrho \\ \mu\lambda \end{Bmatrix}\begin{Bmatrix} \lambda \\ \nu\sigma \end{Bmatrix} - \begin{Bmatrix} \varrho \\ \nu\lambda \end{Bmatrix}\begin{Bmatrix} \lambda \\ \mu\sigma \end{Bmatrix}}_{\mathcal{O}(h^2)} \\ &= \frac{1}{2}\Big(h_{\nu}{}^{\varrho}{}_{,\sigma\mu} + h_{\sigma}{}^{\varrho}{}_{,\nu\mu} - h_{\nu\sigma}{}^{,\varrho}{}_{\mu} - h^{\varrho}{}_{\mu,\sigma\nu} - h_{\sigma}{}^{\varrho}{}_{,\mu\nu} + h_{\mu\sigma}{}^{,\varrho}{}_{\nu}\Big) + \mathcal{O}(h^2) \\ &= \frac{1}{2}\Big(h_{\nu}{}^{\varrho}{}_{,\sigma\mu} - h_{\nu\sigma}{}^{,\varrho}{}_{\mu} - h^{\varrho}{}_{\mu,\sigma\nu} + h_{\mu\sigma}{}^{,\varrho}{}_{\nu}\Big) + \mathcal{O}(h^2)\,. \end{aligned} \tag{7.4}$$

By contracting we get the Ricci tensor

$$R_{\sigma\nu} = R^{\varrho}{}_{\sigma\varrho\nu} = \frac{1}{2}\Big(h_{\nu}{}^{\varrho}{}_{,\sigma\varrho} - h_{\nu\sigma}{}^{,\varrho}{}_{\varrho} - h_{,\sigma\nu} + h_{\varrho\sigma}{}^{,\varrho}{}_{\nu}\Big) + \mathcal{O}(h^2)\,, \tag{7.5}$$

where for convenience the trace of h is denoted with $h \coloneqq h_{\mu\nu}\eta^{\mu\nu}$. Lastly the Ricci scalar is given by

$$R = g^{\sigma\nu}R_{\sigma\nu} = \eta^{\sigma\nu}R_{\sigma\nu} + \mathcal{O}(h^2) = h^{\sigma\nu}{}_{,\sigma\nu} - h_{,\sigma}{}^{\sigma} + \mathcal{O}(h^2) \tag{7.6}$$

We define

$$\overline{h}_{\mu\nu} \coloneqq h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h\,. \tag{7.7}$$

The trace is given by

$$\overline{h} \coloneqq \overline{h}_{\mu\nu}\eta^{\mu\nu} = h - \frac{h}{2} = \frac{h}{2}\,. \tag{7.8}$$

If we repeat the procedure we arrive at the initial metric:

$$\overline{\overline{h}}_{\mu\nu} = \overline{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\overline{h} = \overline{h}_{\mu\nu} + \frac{1}{2}\eta_{\mu\nu}h = h_{\mu\nu}\,. \tag{7.9}$$

Linearized Einstein tensor $G_{\mu\nu}$ in terms of $\overline{h}_{\mu\nu}$

The linearized Einstein Tensor is given as

$$\begin{aligned}
G^{(\mathrm{L})}_{\mu\nu} &= R^{(\mathrm{L})}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}R^{(\mathrm{L})} \\
&= \frac{1}{2}\partial_\mu\partial_\varrho h_\nu{}^\varrho + \frac{1}{2}\partial_\nu\partial_\varrho h_\mu{}^\varrho - \frac{1}{2}\Box h_{\mu\nu} - \frac{1}{2}\partial_\mu\partial_\nu h - \frac{1}{2}\eta_{\mu\nu}\partial_\varrho\partial_\sigma h^{\varrho\sigma} + \frac{1}{2}\eta_{\mu\nu}\Box h \\
&= \frac{1}{2}\partial_\mu\partial_\varrho \overline{h}_\nu{}^\varrho - \frac{1}{4}\partial_\mu\partial_\nu\overline{h} + \frac{1}{2}\partial_\nu\partial_\varrho\overline{h}_\mu{}^\varrho - \frac{1}{4}\partial_\nu\partial_\mu\overline{h} - \frac{1}{2}\Box\overline{h}_{\mu\nu} \\
&\quad + \frac{1}{2}\eta_{\mu\nu}\Box\overline{h} + \frac{1}{2}\partial_\mu\partial_\nu\overline{h} - \frac{1}{2}\eta_{\mu\nu}\partial_\varrho\partial_\sigma\overline{h}^{\varrho\sigma} + \frac{1}{4}\eta_{\mu\nu}\Box\overline{h} - \frac{1}{2}\eta_{\mu\nu}\Box\overline{h} \\
&= -\frac{1}{2}\Box\overline{h}_{\mu\nu} + \partial_\varrho\partial_{(\mu}\overline{h}_{\nu)}{}^\varrho - \frac{1}{2}\eta_{\mu\nu}\partial_\varrho\partial_\sigma\overline{h}^{\varrho\sigma}\,,
\end{aligned} \tag{7.10}$$

with the (linearized) d'Alembert operator

$$\Box^{(\mathrm{L})} = \Box = \partial_\mu\partial_\nu\eta^{\mu\nu} = \partial_\mu\partial^\mu\,. \tag{7.11}$$

Definition 13 (Linearized Einstein equations).

$$-\frac{1}{2}\Box\overline{h}_{\mu\nu} + \partial_\varrho\partial_{(\mu}\overline{h}_{\nu)}{}^\varrho - \frac{1}{2}\eta_{\mu\nu}\partial_\varrho\partial_\sigma\overline{h}^{\varrho\sigma} = \kappa T_{\mu\nu} \tag{7.12}$$

Gauge transformations

Usually field equations are in the form of

$$\Box\text{"field"} = \text{"source"}\,. \tag{7.13}$$

Equation (7.12) can be written in this form:

$$\underbrace{\Box\overline{h}_{\mu\nu}}_{\Box\text{"field"}} \underbrace{-2\partial_\varrho\partial_{(\mu}\overline{h}_{\nu)}{}^\varrho + \eta_{\mu\nu}\partial_\varrho\overline{h}^{\varrho\sigma}}_{\text{ensures gauge invariance of equation}} = \underbrace{-2\kappa T_{\mu\nu}}_{\text{"source"}} \tag{7.14}$$

We are now considering infinitesimal diffeomorphisms, which are given by affine transformations

$$x^\mu = x'^\mu + \xi^\mu(x'^\mu), \qquad \xi^\mu \ll 1 \tag{7.15}$$

In the following we neglect terms with $\mathcal{O}(\xi^2)$, $\mathcal{O}(\xi h)$, and $\mathcal{O}(h^2)$ and higher order terms, which we denote by $\mathcal{O}$. The transformed metric reads

$$\begin{aligned}
\eta_{\mu\nu} + h'_{\mu\nu}(x') &= g'_{\mu\nu} \\
&= \frac{\partial x^\varrho}{\partial x'^\mu}\frac{\partial x^\sigma}{\partial x'^\nu} g_{\varrho\sigma}(x) \\
&= \frac{\partial (x'^\varrho + \xi^\varrho)}{\partial x'^\mu}\frac{\partial (x'^\sigma + \xi^\sigma)}{\partial x'^\nu}\big(\eta_{\varrho\sigma} + h_{\varrho\sigma}(x)\big) + \mathcal{O} \\
&= \Big(\delta_\mu{}^\varrho + \xi^\varrho{}_{,\mu}\Big)\Big(\delta_\nu{}^\sigma + \xi^\sigma{}_{,\nu}\Big)\big(\eta_{\varrho\sigma} + h_{\varrho\sigma}(x)\big) + \mathcal{O} \\
&= \Big(\delta_\mu{}^\varrho + \xi^\varrho{}_{,\mu}\Big)\Big(\eta_{\varrho\nu} + h_{\varrho\nu} + \xi_{\varrho,\nu}\Big) + \mathcal{O} \\
&= \eta_{\mu\nu} + h_{\mu\nu} + \xi_{\mu,\nu} + \xi_{\nu,\mu} + \mathcal{O}\,.
\end{aligned} \tag{7.16}$$

The perturbation $h_{\mu\nu}$ therefore transforms under infinitesimal diffeomorphisms in the following way

$$\begin{aligned}
h'_{\mu\nu}(x) &= h_{\mu\nu}(x) + \xi_{\mu,\nu} + \xi_{\nu,\mu} \\
&= h_{\mu\nu}(x) + \big(\mathcal{L}_\xi \eta\big)_{\mu\nu}\,.
\end{aligned} \tag{7.17}$$

Definition 14 (Lie derivative)**.** The Lie derivative off a tensor field T with k contravariant and l covariant indices along the vector ξ is defined as

$$\begin{aligned}
\big(\mathcal{L}_\xi T\big)^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l} := \xi^\mu\partial_\mu T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_l} &- \big(\partial_\mu\xi^{\alpha_1}\big)T^{\mu\alpha_2\ldots\alpha_k}_{\beta_1\ldots\beta_l} - \ldots - \big(\partial_\mu\xi^{\alpha_k}\big)T^{\alpha_1\ldots\alpha_{k-1}\mu}_{\beta_1\ldots\beta_l} \\
&+ \big(\partial_{\beta_1}\xi^\mu\big)T^{\alpha_1\ldots\alpha_k}_{\mu\beta_2\ldots\beta_l} + \ldots + \big(\partial_{\beta_l}\xi^\mu\big)T^{\alpha_1\ldots\alpha_k}_{\beta_1\ldots\beta_{l-1}\mu}\,.
\end{aligned} \tag{7.18}$$

Therefore

$$\big(\mathcal{L}_\xi \eta\big)_{\mu\nu} = \underbrace{\xi^\varrho\partial_\varrho\eta_{\mu\nu}}_{=0} + \xi_{\mu,\nu} + \xi_{\nu,\mu} = \xi_{\mu,\nu} + \xi_{\nu,\mu}\,. \tag{7.19}$$

If the derivative of a metric vanishes for a given ξ^μ, then one obtains the killing equations

$$\xi_{\mu,\nu} + \xi_{\nu,\mu} = 0 \tag{7.20}$$

for ξ^μ and the solutions are referred to as *killing vector fields.* In Minkowski-space the ten infinitesimal killing vectors correspond to the Poincaré-generators.

Aside 9. We can use the Lie-derivative on metric to detect symmetries of the Manifold.

Invariance of the linearized field equations under infinitesimal diffeomorphisms

We now check that linearized field equations are invariant under infinitesimal diffeomorphism

$$h'_{\mu\nu} = h_{\mu\nu} + \xi_{\mu,\nu} + \xi_{\nu,\mu}\,. \tag{7.21}$$

The barred metric transforms as

$$\begin{aligned}\overline{h'}_{\mu\nu} &= h'_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h' \\ &= h_{\mu\nu} + \xi_{\mu,\nu} + \xi_{\nu,\mu} - \frac{1}{2}\eta_{\mu\nu}h - \frac{1}{2}\eta_{\mu\nu}\partial^\varrho\xi_\varrho - \frac{1}{2}\eta_{\mu\nu}\partial^\varrho\xi_\varrho \\ &= \overline{h}_{\mu\nu} + \xi_{\mu,\nu} + \xi_{\nu,\mu} - \eta_{\mu\nu}\xi^\varrho{}_{,\varrho}\,.\end{aligned} \tag{7.22}$$

We proceed by plugging this into Einstein's equations, the relevant terms are

$$-\frac{1}{2}\Box\overline{h'}_{\mu\nu} = -\frac{1}{2}\Box\overline{h}_{\mu\nu} - \frac{1}{2}\Box\xi_{\mu,\nu} - \frac{1}{2}\Box\xi_{\nu,\mu} + \frac{1}{2}\eta_{\mu\nu}\Box\xi^\varrho{}_{,\varrho}\,, \tag{7.23}$$

$$-\frac{1}{2}\eta_{\mu\nu}\partial_\varrho\partial_\sigma\overline{h'}^{\varrho\sigma} = -\frac{1}{2}\eta_{\mu\nu}\partial_\varrho\partial_\sigma\left(\xi^{\varrho,\sigma} + \xi^{\sigma,\varrho} - \eta^{\varrho\sigma}\xi^\alpha{}_{,\alpha}\right) - \frac{1}{2}\eta_{\mu\nu}\partial_\varrho\partial_\sigma\overline{h}^{\varrho\sigma}\,, \tag{7.24}$$

$$\partial^\varrho\partial_{(\mu}\overline{h'}_{\nu)\varrho} = \partial^\varrho\partial_{(\mu}\overline{h}_{\nu)\varrho} + \frac{1}{2}\Box\xi_{\nu,\mu} + \frac{1}{2}\Box\eta_{\mu\nu}\,. \tag{7.25}$$

Therefore

$$\begin{aligned}& -\frac{1}{2}\Box\overline{h'}_{\mu\nu} + \partial_\varrho\partial_{(\mu}\overline{h'}_{\nu)}{}^\varrho - \frac{1}{2}\eta_{\mu\nu}\partial_\varrho\partial_\sigma\overline{h'}^{\varrho\sigma} \\ =\; & -\frac{1}{2}\Box\overline{h}_{\mu\nu} - \frac{1}{2}\Box\xi_{\mu,\nu} - \frac{1}{2}\Box\xi_{\nu,\mu} + \frac{1}{2}\eta_{\mu\nu}\Box\xi^\varrho{}_{,\varrho} + \partial^\varrho\partial_{(\mu}\overline{h}_{\nu)\varrho} \\ & + \Box\xi_{(\mu,\nu)} - \frac{1}{2}\eta_{\mu\nu}\Box\xi^\varrho{}_{,\varrho} - \frac{1}{2}\eta_{\mu\nu}\partial_\varrho\partial_\sigma\overline{h}^{\varrho\sigma} \\ =\; & -\frac{1}{2}\Box\overline{h}_{\mu\nu} + \partial_\varrho\partial_{(\mu}\overline{h}_{\nu)}{}^\varrho - \frac{1}{2}\eta_{\mu\nu}\partial_\varrho\partial_\sigma\overline{h}^{\varrho\sigma}\end{aligned} \tag{7.26}$$

This shows that the Einstein equations are invariant under an infinitesimal diffeomorphisms. Therefore $\overline{h}_{\mu\nu}$ and $\overline{h'}_{\mu\nu}$ are the same *physical* field.

Harmonic gauge in linearized gravity

As described above we want to bring the field equation in the form $\Box$"field" = "source", i.e. a wave equation. This can ge done with the gauge condition

$$\chi_\nu\left[\overline{h}\right] := \partial^\mu\overline{h}_{\mu\nu} = 0. \tag{7.27}$$

In therms of the original field this condition reads

Definition 15 (de Donder gauge, harmonic gauge)**.**

$$\chi_\nu[h] = \partial^\mu h_{\mu\nu} - \frac{1}{2} h_{\mu\nu} \partial^\mu h = 0 \tag{7.28}$$

Proof:

$$\begin{aligned} \partial^\mu \overline{h'}_{\mu\nu} &= \partial^\mu \overline{h}_{\mu\nu} + \Box \xi_\nu + \partial_\nu \partial^\mu \xi_\mu - \eta_{\mu\nu} \partial^\mu \partial_\varrho \xi^\varrho \\ &= \partial^\mu \overline{h}_{\mu\nu} + \Box \xi_\nu = 0 \end{aligned} \tag{7.29}$$

Solve for $\Box \xi_\nu$

$$\implies \Box \overline{h'}_{\mu\nu} = -2\kappa T_{\mu\nu} \tag{7.30}$$

Since $\overline{h}_{\mu\nu}$ and $\overline{h'}_{\mu\nu}$ correspond to the same physical field configuration, we can drop the prime.

Definition 16. Linearized field equations in de Donder gauge.

$$\Box \overline{h}_{\mu\nu} = -2\kappa T_{\mu\nu} \tag{7.31}$$

Remark 15 (Fierz-Pauli action, 1939).

$$\mathcal{L}_{\text{FP}} = \frac{1}{2} \big(\partial_\mu h^{\mu\nu}\big)\big(\partial_\nu h\big) - \partial_\mu h^{\rho\sigma} \partial_\rho h^\mu{}_\sigma + \frac{1}{2} \eta^{\mu\nu} \big(\partial_\mu h^{\rho\sigma}\big)\big(\partial_\nu h_{\rho\sigma}\big) - \frac{1}{2} \eta^{\mu\nu} \big(\partial_\mu h\big)\big(\partial_\nu h\big) \tag{7.32}$$

For vacuum this is the Lagrangian of a massless spin-2 field $h_{\mu\nu}(x)$ ("the graviton") in flat spacetime $h^{\mu\nu}$.
Problem: non-linearity (in electrodynamics: linear coupling)

$$T_{\mu\nu} h^{\mu\nu} \to h^{(2)}_{\mu\nu} \propto \left(h^{(1)}_{\mu\nu}\right)^2 \tag{7.33}$$

$\to$ Deser 1970: Iterative procedure
$\hookrightarrow$ including gravitational self energy and resuming one recovers the full nonlinear Einstein equations.

Table 7.1. – Comparison between linearized gravity and electrodynamics.

	linearized gravity	electrodynamics
basic field	$\overline{h}_{\mu\nu}$ $(h_{\mu\nu})$, spin-2, *graviton*	A_μ, spin-1, *gauge boson*, *gauge potential*
field equations	$\underbrace{\Box\overline{h}_{\mu\nu}}_{\Box\text{“field”}} - \underbrace{2\partial_\rho\partial_{(\mu}\overline{h}_{\nu)}{}^{\rho} - \eta_{\mu\nu}\partial_\rho\partial_\sigma\overline{h}^{\rho\sigma}}_{\text{ensures gauge inv.}} = -\underbrace{2\kappa T_{\mu\nu}}_{\text{“source”}}$	$\underbrace{\Box A_\mu}_{\Box\text{“field”}} - \underbrace{\partial_\mu(\partial_\nu A^\nu)}_{\text{ensures gauge inv.}} = -\underbrace{4\pi j_\mu}_{\text{source}}$
transf. under inf. gauge trafos	$\overline{h'}_{\mu\nu} = h_{\mu\nu} + \xi_{\mu,\nu} + \xi_{\nu,\mu} - \eta_{\mu\nu}\xi^\rho{}_{,\rho}$	$A'_\mu = A_\mu + \partial_\mu\lambda(x)$
	inf. coordinate transformation	internal symmetry
inv. of field eqs	yes	yes
specific gauges	de Donder gauge, $\partial_\mu\overline{h}^{\mu\nu} = 0$	Lorentz gauge, $\partial_\mu A^\mu = 0$
field eqs. in specific gauges	$\Box\overline{h}_{\mu\nu} = -2\kappa T_{\mu\nu}$	$\Box A_\mu = -4\pi j_\mu$
inv. tensors under gauge trafo	$R^{(L)\prime}_{\mu\nu} = R^{(L)}_{\mu\nu}$	$F'_{\mu\nu} = F_{\mu\nu}$

7.2. Newtonian Limit

Empirically we know

1. Newtonian gravity describes the dynamics in our solar system to a high accuracy
2. On earth, we can measure the gravitation constant G_{N} e.g. by Cavendish-type experiments

If General Relativity is a more fundamental gravitational theory than Newton's theory it should

1. recover NEWTON's theory in appropriate limit, i.e. in the domain where Newtonian Gravity is a good description
2. be more accurate than Newton's theory, i.e. it should predict small corrections to Newtonian Gravity.

Conditions for the Newtonian limit:

1. $v \ll c$ (sources move slowly)
   slowly changing geometry $\approx$ static: no $\mathrm{d}x^i\,\mathrm{d}t$ terms in $\mathrm{d}s^2$ (would violate $t \to -t$ invariance)
2. $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ with $\left|h_{\mu\nu}\right| \ll 1$ (weak gravitational field)
3. $p \ll \rho$ (sources have low internal pressure)

ad 1. $v \ll c$ is required as (special) relativistic effects must be small

ad 2. Consider the solar system as a closed system: Then a particle in the outer region with $v \ll c$ initially, will fall into the inner region (center of mass) and it will be accelerated by gravity. It will be then have a kinetic energy $E_{\mathrm{kin}} = \frac{1}{2}mv^2 \sim |m\Phi|$, where $\Phi < 0$ is the gravitational Newtonian potential with boundary condition $\lim\limits_{x\to\infty} \Phi(x) = 0$. Small velocities of the sources imply weak gravitational fields.

ad 3. Speed of sound

$$c_s := \left|\frac{T_{ij}}{T_{00}}\right| \quad \text{with} \quad T_{\mu\nu} = \mathrm{diag}\left(\rho, \frac{p}{c^2}, \frac{p}{c^2}, \frac{p}{c^2}\right) \quad \text{(perfect fluid)} \tag{7.34}$$

$$c_s \sim \left(\frac{p}{\rho}\right)^{1/2} \tag{7.35}$$

The internal pressure of the sources must be small, otherwise they would also create (fast) motion of sound waves.
$\Longrightarrow$ $p \ll \rho$
$\Longrightarrow$ energy-momentum tensor of dust

$$T^{\mu\nu} = \rho_0 t^\mu t^\nu \tag{7.36}$$
$$t^\mu = \delta^\mu_0 = \left(\frac{\partial}{\partial x^0}\right)^\mu \tag{7.37}$$

t^μ is the "direction" of an internal coordinate system of time

$$\Box \overline{h}_{\mu\nu} \approx \Delta \overline{h}_{\mu\nu} \tag{7.38}$$

alternatively $\Box = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \Delta \approx \Delta$ as $\frac{1}{c}\frac{\partial}{\partial t} = \frac{1}{c}\frac{\partial}{\partial x}\frac{\partial x}{\partial t} \sim \frac{v}{c}\frac{\partial}{\partial x} \ll \frac{\partial}{\partial x}$

For the Newtonian limit, we must look for solutions to $\Box \overline{h}_{\mu\nu} = -2\kappa T_{\mu\nu}$, where time-derivatives are negligible and where the energy-momentum tensor is the one of dust.

$$\Delta \overline{h}_{\mu\nu} = \begin{cases} -2\kappa\rho_0 & \mu = \nu = 0 \\ 0 & \text{else} \end{cases} \tag{7.39}$$

Consider first the Poisson equation with vanishing sources. The unique solution is $\overline{h}_{\mu\nu} =$ const.. The const. can be always adjusted to zero by a residual gauge transformation:

$$\overline{h}_{\mu\nu} = 0, \quad \mu \neq \nu = 0 \tag{7.40}$$

Residual gauge transformation

$$\partial^\mu \overline{h'}_{\mu\nu} = \underbrace{\partial^\mu \overline{h}_{\mu\nu}}_{=0} + \Box \xi_\nu = 0 \tag{7.41}$$

This means that all gauge transformations ξ_μ with $\Box \xi_\mu = 0$ are compatible with the de Donder gauge (i.e. that doesn't lead out of the de Donder gauge).
As far we know:

$$\overline{h}_{\mu\nu} = 0 \qquad \text{for the 0-0 component} \tag{7.42}$$
$$\Delta \overline{h}_{00} = -2\kappa\rho_0 \qquad \mu \neq \nu = 0 \tag{7.43}$$

We identify the gravitational potential as

$$\Phi := -\frac{1}{4}\overline{h}_{00} \tag{7.44}$$

We obtain the Poisson equation

$$\Delta\Phi = \frac{\kappa}{2}\rho_0 = 4\pi G_{\mathrm{N}}\rho_0 \tag{7.45}$$

Solution in terms of the original field $h_{\mu\nu}$:

$$h_{\mu\nu} = \overline{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\overline{h} = \overline{h}_{00}\left(\delta_\mu{}^0\delta_\nu{}^0 + \frac{1}{2}\eta_{\mu\nu}\right) = -4\Phi\left(\delta_\mu{}^0\delta_\nu{}^0 + \frac{1}{2}\eta_{\mu\nu}\right) \tag{7.46}$$

We have used $\overline{h} = \eta^{\rho\sigma}\overline{h}_{\rho\sigma} = -\overline{h}_{00} = 4\Phi$

$$\implies h_{00} = -2\Phi \qquad h_{ij} = -2\Phi\delta_{ij} \qquad h_{0\mu} = 0 \tag{7.47}$$

$$\implies \mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = -(1+2\Phi)\,\mathrm{d}t^2 + (1-2\Phi)\delta_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j \tag{7.48}$$

This is the Newtonian geometry.

7.2.1. Motion of test particles in Newtonian Geometry

Test particles carry clocks that read universal time in Newton Geometry.

$$\frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} = \frac{\mathrm{d}^2x^i}{\mathrm{d}\tau^2} = -\begin{Bmatrix} i \\ \alpha\beta \end{Bmatrix}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}\frac{\mathrm{d}x^\beta}{\mathrm{d}\tau} = -\begin{Bmatrix} i \\ 00 \end{Bmatrix} = -[00,\, i] = \frac{1}{2}h_{00,i} - h_{0i,0} = \frac{1}{2}h_{00,i} = -\Phi_{,i} \tag{7.49}$$

$$\implies \frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} = -\Phi_{,i} \qquad (\ \widehat{=}\quad \mathbf{a} = -\nabla\Phi) \tag{7.50}$$

where we used $\frac{\mathrm{d}\tau}{\mathrm{d}t} = 1$,$v^i \sim \left|\frac{\mathrm{d}x^i}{\mathrm{d}\tau}\right| \ll 1$, $g_{\mu\nu} = \eta_{\mu\nu}$, and $\mathrm{d}_t \sim \frac{v^i}{c}\,\mathrm{d}_{x_i} \ll 1$. In Newtonian Gravity, we have two equations

1. "field equations": $\Delta\Phi = 4\pi G_{\mathrm{N}}\rho_0$ (Poisson equation), describes how the gravitational potential (geometry in General Relativity) reacts on matter ρ_0

2. "geodesic equation": $\frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} = -\Phi_{,i}$, describes how matter (test particles, dust) moves under the influence of the gravitational potential Φ (in General Relativity: moves in curved geometry)

7.2.2. Geodesic deviation in Newtonian Geometry

From the last analysis, we know

$$\begin{Bmatrix} i \\ 00 \end{Bmatrix} = \Phi^{,i} \qquad \text{(all other components are zero).} \tag{7.51}$$

Insert this in the Riemannian curvature tensor:

$$R^i{}_{0j0} = -R^i{}_{00j} = \Phi^{,i}{}_j \qquad \text{(all other components are zero).} \tag{7.52}$$

Ricci-tensor:

$$R_{00} = \Delta\Phi = 4\pi G_{\mathrm{N}}\rho_0 \tag{7.53}$$

geodesic deviation

$$\frac{\mathrm{D}^2\,\eta^i}{\mathrm{D}\,\tau^2} \approx \frac{\mathrm{d}^2\eta^i}{\mathrm{d}\tau^2} = -R^i{}_{0j0}\eta^j = -\Phi^{,i}{}_j\eta^j \tag{7.54}$$

where we used that

$$\frac{\mathrm{D}\,\eta^i}{\mathrm{D}\,\tau} = \nabla_j\eta^i\frac{\mathrm{d}x^j}{\mathrm{d}\tau} = \partial_j\eta^i x^j = \frac{\mathrm{d}\eta^i}{\mathrm{d}t} \tag{7.55}$$

Compare to the deviation equation due to tidal forces in Newtonian Gravity.

$$\begin{aligned} \frac{\mathrm{d}^2\eta^i}{\mathrm{d}\tau^2} &= \frac{\mathrm{d}^2\left(x^i + \eta^i\right)}{\mathrm{d}t^2} - \frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} \\ &= -\left.\frac{\partial\Phi}{\partial x_i}\right|_{x+\eta} + \left.\frac{\partial\Phi}{\partial x_i}\right|_x \\ &= -\left.\frac{\partial\Phi}{\partial x_i}\right|_x - \left.\frac{\partial^2\Phi}{\partial x_i\partial x^j}\right|_x\eta^j + \left.\frac{\partial\Phi}{\partial x_i}\right|_x \\ &= -\Phi^{,i}{}_j\eta^j \end{aligned} \tag{7.56}$$

This shows again that tidal forces are a genuine gravitational effect that is related to curvature of spacetime (in General Relativity) and cannot be transformed away.

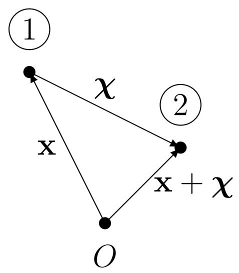


Figure 7.1.

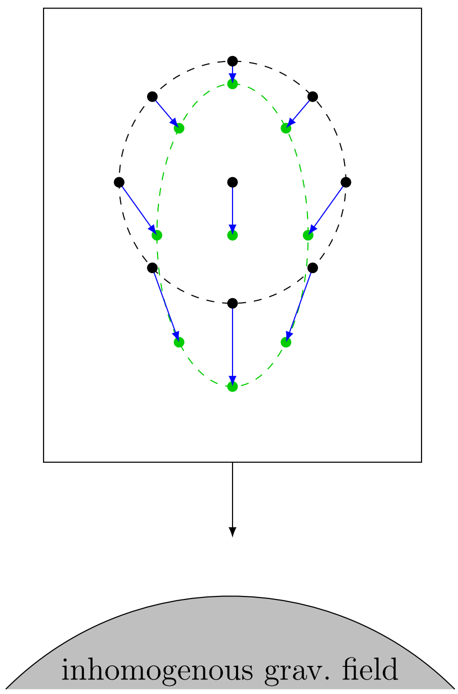
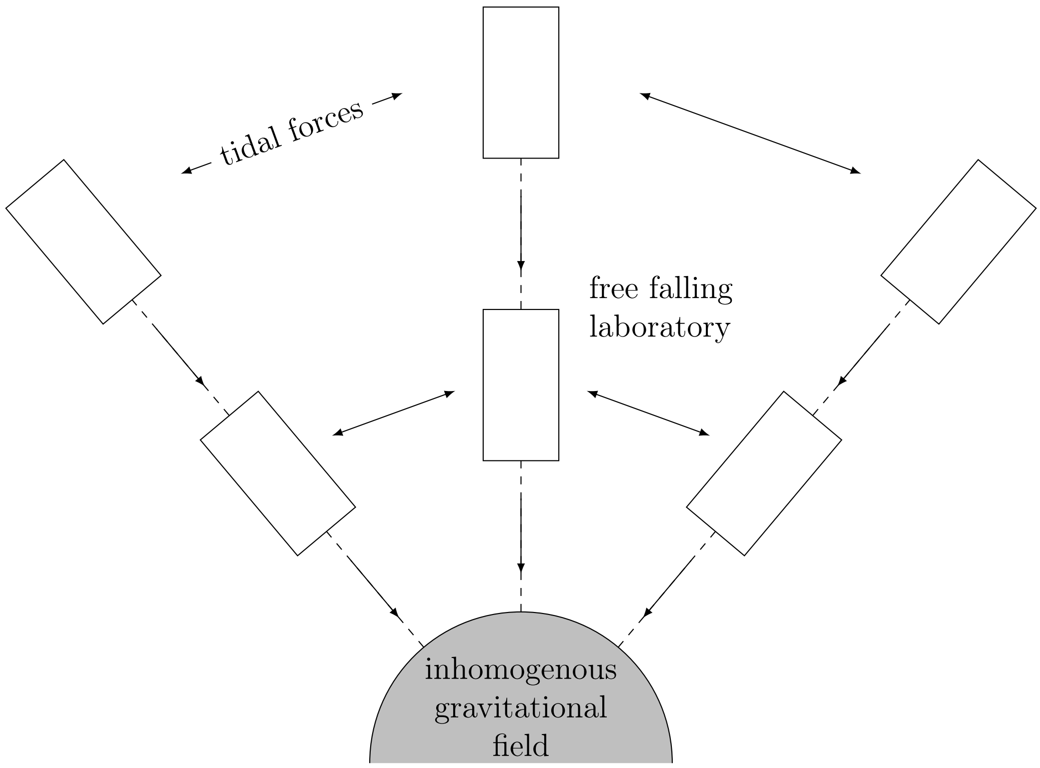
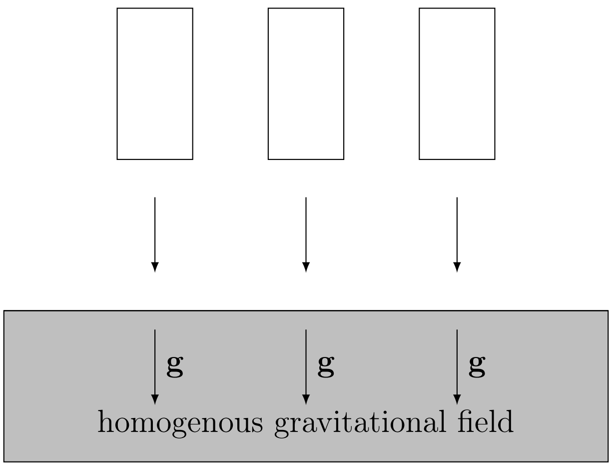


Figure 7.2. – Tidal forces acting on particles on a ring


tidal forces
free falling
laboratory
inhomogenous
gravitational
field


g
g
g
homogenous gravitational field

8. Gravitational Waves

- vacuum solution of linearized Einstein equations

$$\Box \overline{h}_{\mu\nu} = 0 \qquad \text{in de Donder gauge} \tag{8.1}$$

- describes *weak* gravitational waves *only* $\to$ linearized treatment justified
- description breaks down for strong gravitational fields as the theory becomes essentially *non-linear* (e.g. two black holes merge)
- analysis similar to electrodynamics, but here $h_{\mu\nu}$: spin-2 field, A_μ: spin-1 field
- vacuum equations

$$\Box \overline{h}_{\mu\nu} = 0 \qquad \overline{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h \qquad \partial^\mu \overline{h}_{\mu\nu} = 0 \tag{8.2}$$

- Gauge freedom not yet completely exhausted by de Donder gauge.

$$\partial^\mu \overline{h'}_{\mu\nu} = \underbrace{\partial^\mu \overline{h'}_{\mu\nu}}_{=0} + \Box \xi_\nu = 0 \tag{8.3}$$

$\Longrightarrow$ All gauge transformations generated by ξ_ν that satisfy $\Box \xi_\nu = 0$ do not lead out of the de Donder gauge.
Compare with electromagnetic field:

$$A_\mu \to A'^\mu = A^\mu + \partial_\mu \lambda(x) \tag{8.4}$$

$$\partial_\mu A^\mu = 0 \qquad \text{Lorentz gauge} \tag{8.5}$$

$$\partial_\mu A'^\mu = \underbrace{\partial_\mu A^\mu}_{=0} + \partial_\mu \partial^\mu \lambda(x) = 0 \implies \Box \lambda = 0 \tag{8.6}$$

Exploit this remaining gauge freedom to make perturbations $h_{\mu\nu}$ *transverse* and *traceless*

- transversality

$$\partial^\mu h'_{\mu\nu} = \partial^\mu h_{\mu\nu} + \Box\xi_\nu + \partial_\nu\partial^\mu\xi_\mu \overset{!}{=} 0 \tag{8.7}$$

Since only gauge transformations that satisfy $\Box\xi_\nu$ are allowed (do not lead out of de Donder gauge), the equation that shoud be solved for ξ_ν is

$$\partial_\nu\partial^\mu\xi_\mu = -\partial^\mu h_{\mu\nu} \tag{8.8}$$

- In order for the perturbations to be traceless, we need to find a solution to

$$h' = h + \partial_\mu\xi^\mu = 0 \implies \partial_\mu\xi^\mu = -\frac{1}{2}h \tag{8.9}$$

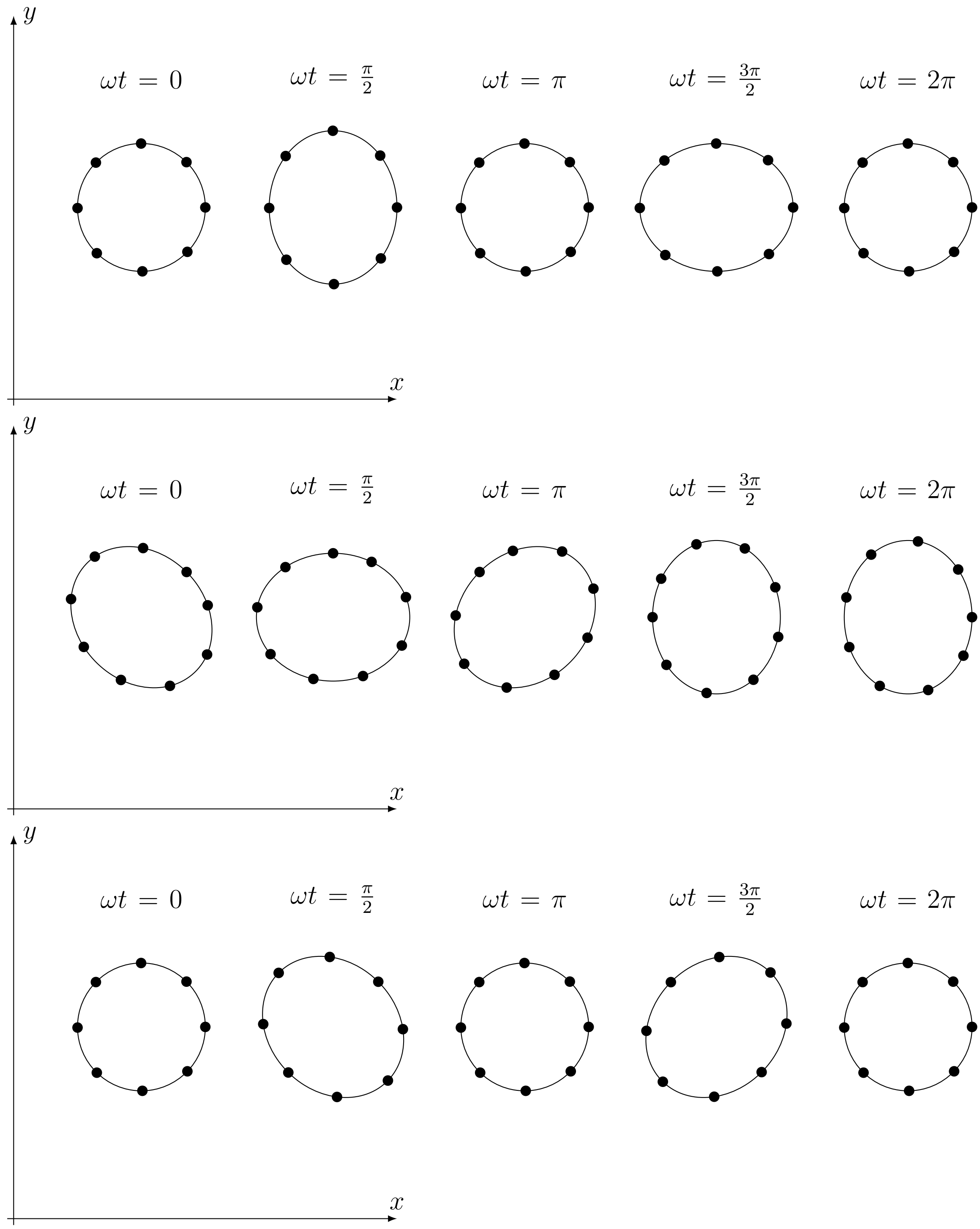


Figure 8.1. – Polarisations of gravitational waves. From top to bottom: +,× and mixed (circular) polarised waves.

8.1. Degrees of Freedom (DoF) and Scalar Vector Tensor (SVT) Decomposition in Space/Time

Parametrize the line element as

$$\mathrm{d}s^2 = -(1+2\Phi)\,\mathrm{d}t^2 + v_i\Big(\mathrm{d}t\,\mathrm{d}x^i + \mathrm{d}t\,\mathrm{d}x^i\Big) + \Big(\delta_{ij} + h_{ij}\Big)\,\mathrm{d}x^i\,\mathrm{d}x^j\,. \tag{8.10}$$

Block matrix

$$h_{\mu\nu} = \begin{bmatrix} h_{00} & h_{0j} \\ h_{i0} & h_{ij} \end{bmatrix} = \begin{bmatrix} -2\Phi & v_j \\ v_i & h_{ij} \end{bmatrix}, \quad |\Phi|\,,|v_i|\,,\Big|h_{ij}\Big| \ll 1\,. \tag{8.11}$$

In general algebraic decomposition: symmetric/antisymmetric

$$T_{\mu\nu} = T^{\mathrm{s}}_{\mu\nu} + T^{\mathrm{as}}_{\mu\nu} = T_{(\mu\nu)} + T_{[\mu\nu]} = \frac{1}{2}\Big(T_{\mu\nu} + T_{\nu\mu}\Big) + \frac{1}{2}\Big(T_{\mu\nu} - T_{\nu\mu}\Big)\,. \tag{8.12}$$

Can we decompose $T^{\mathrm{s}}_{\mu\nu}$ further?
Wake the trace

$$T^{\mathrm{s}} = g^{\mu\nu}T^{\mathrm{s}}_{\mu\nu}\,, \tag{8.13}$$

so we can write $T^{\mathrm{s}}_{\mu\nu}$ as follows:

$$T^{\mathrm{s}}_{\mu\nu} = T^{\mathrm{tf}}_{\mu\nu} + \frac{1}{d}\,g_{\mu\nu}T^{\mathrm{s}}\,, \tag{8.14}$$

where d denotes the dimension of spacetime and $T^{\mathrm{tf}}_{\mu\nu}$ is a tracefree, symmetric tensor. Can we decompose $T^{\mathrm{tf}}_{\mu\nu}$ any further? Possible for tensor fields $T_{\mu\nu}(x)$ as this involves derivatives. Decompose metric perturbations:

$$h_{00} = -2\Phi \tag{8.15}$$

$$h_{0i} = v_i \tag{8.16}$$

$$h_{ij} = 2s_{ij} - 2\Psi s_{ij} \qquad \begin{cases} \Psi := -\frac{1}{6}\delta^{ij}h_{ij} & \text{“trace”} \\ s_{ij} := \frac{1}{2}\Big(h_{ij} - \frac{1}{3}\delta^{kl}h_{kl}\delta_{ij}\Big) & \text{“strain”} \end{cases} \tag{8.17}$$

$$\mathrm{d}s^2 = -(1+2\Phi)\,\mathrm{d}t^2 + v_i\Big(\mathrm{d}t\,\mathrm{d}x^i + \mathrm{d}t\,\mathrm{d}x^i\Big) + \Big[(1-2\Psi)\delta_{ij} + 2S_{ij}\Big]\,\mathrm{d}x^i\,\mathrm{d}x^j \tag{8.18}$$

$\partial_i = k_i$ (momentum space) Decompose vector ω_i into transverse and longitudinal components.

$$v_i = S_i + \partial_i B \tag{8.19}$$

$$3 \; + \quad 1 \quad \underbrace{-1}_{\partial_i S^i = 0} \qquad \text{DoF} \tag{8.20}$$

$$S_{ij} = \partial_{(i} F_{j)} + \left(\partial_i \partial_j - \frac{1}{3}\,\delta_{ij} \Delta\right) E + h_{ij}^{\mathrm{TT}} \tag{8.21}$$

4	4 scalars
+	
$2\,(3-1)$	2 transverse vectors
+	
$6-1-3$	1 symmetric transverse traceless tensor
10	independent components of $h_{\mu\nu}$

valid for (reduce free components):

$$\partial_i F^i = 0\,, \quad h_{ij}^{\mathrm{TT}} = h_{ji}^{\mathrm{TT}}\,, \quad h_{ij}^{\mathrm{TT}} \delta^{ij} = 0\,, \quad \partial^i h_{ij}^{\mathrm{TT}} = 0 \tag{8.22}$$

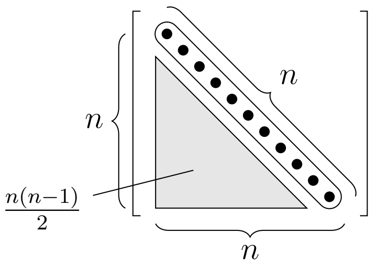


Figure 8.2.

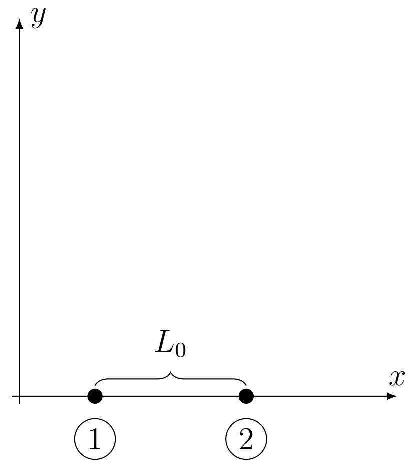


Figure 8.3.

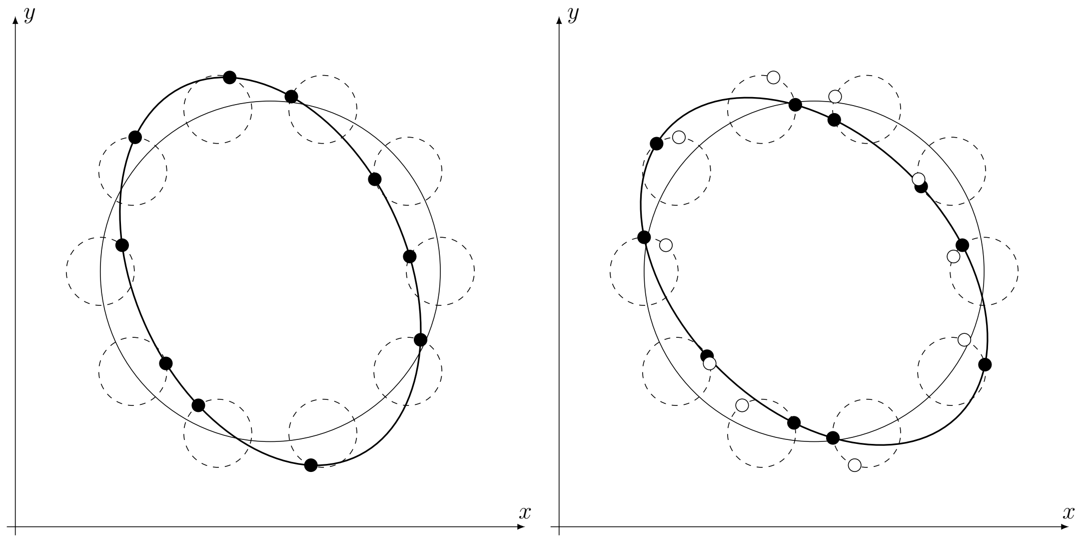


Figure 8.4.

9. The Schwarzschild Solution

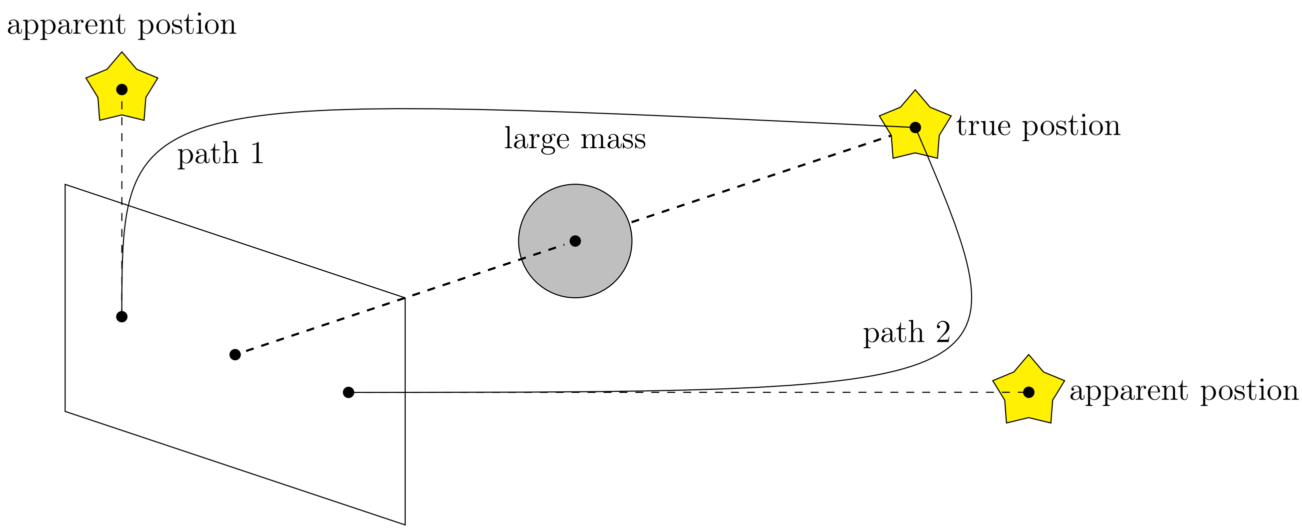
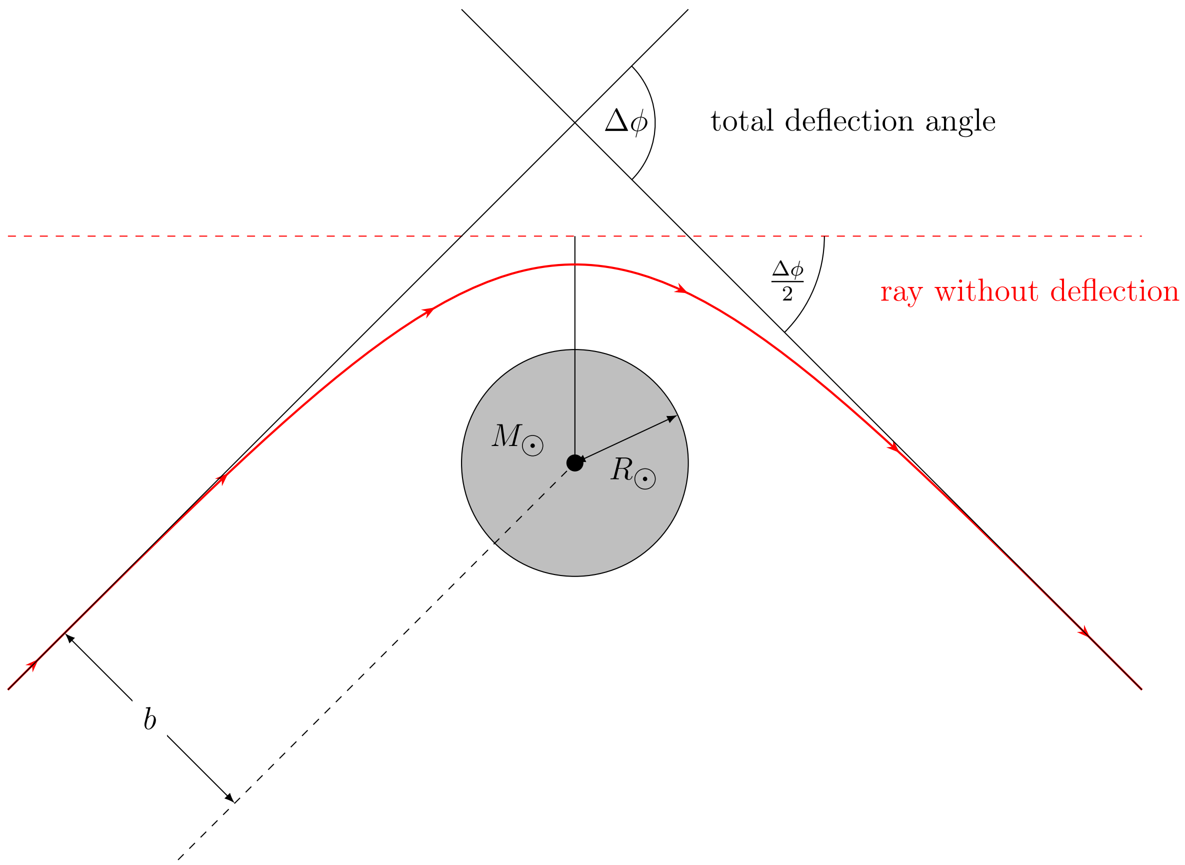


Figure 9.1.



$\Delta\phi$
total deflection angle
$\frac{\Delta\phi}{2}$
ray without deflection
$M_\odot$
$R_\odot$
b

Figure 9.2.

10. Experimental Tests in the Solar System

How can we test relativistic effects in our solar system?

- The sum of the mass of all planets is much smaller than the solar mass $M_\odot \approx 2 \cdot 10^{30}$ kg. The heaviest planet is Jupiter with a mass of $M_{\mathrm{Jup}} \approx 2 \cdot 10^{27}$ kg, therefore we can assume the planets to be testparticles.
- The sun is in good approximation a spherically symmetric object. We can therefore use the Schwarzschild metric.

We define $K := -g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu$, which is conserved along geodesics and it holds true that

$$K = \begin{cases} -1 & \text{timelike geodesics} \\ 0 & \text{lightlike geodesics} \end{cases} . \tag{10.1}$$

Using the Schwarzschild metric, we can explicitly write

$$K = -e^{2a(r)}\dot{t}^2 + e^{2b(r)}\dot{r}^2 + r^2\left(\dot{\vartheta} + \sin^2\vartheta\dot{\phi}\right) \tag{10.2}$$

A Killing-vector ξ^μ satisfies

$$\xi_\mu \dot{x}^\mu = \text{ const.} \tag{10.3}$$

For the Schwarzschild metric, there are four independent Killing-vectors, corresponding to 3 rotations, and staticity (time independence). Conservation of angular momentum leads to a motion in a plane, w.l.o.g. we can chose a coordinate system in which $\vartheta = {}^{\pi}/_{2}$. The Killing-Vectors are given by

$$\xi^\mu_{(\varphi)} = (\partial_\varphi)^\mu = \delta^\mu_\varphi\,, \tag{10.4}$$

$$\xi^\mu_{(t)} = (\partial_t)^\mu = \delta^\mu_t\,. \tag{10.5}$$

The associated conserved quantities are

$$E := \xi^{\mu}_{(\varphi)} g_{\mu\nu} \dot{x}^{\nu} = g_{tt} \dot{t} = e^{2a} \dot{t} = \left(1 - \frac{2M}{r}\right) \dot{t} \tag{10.6}$$

$$L := \xi^{\mu}_{(t)} g_{\mu\nu} \dot{x}^{\nu} = g_{\varphi\varphi} \dot{\varphi} = r^2 \dot{\varphi}\,. \tag{10.7}$$

For massless particles, we can think of E and L as conserved energy and angular momentum. It follows

$$K = -\left(1 - \frac{2M}{r}\right) \dot{t}^2 + \left(1 - \frac{2M}{r}\right)^{-1} \dot{r}^2 + r^2 \varphi^2 \tag{10.8}$$

If we insert the conserved quantities L, E and multiply by $\frac{1}{2}\left(1 - \frac{2M}{r}\right)$, we get

$$\frac{E^2}{2} = \frac{\dot{r}^2}{2} + \left(1 - \frac{2M}{r}\right) \left(\frac{L^2}{2r} - \frac{K}{2}\right). \tag{10.9}$$

This expression can be rearranged to the Form

$$\frac{\dot{r}^2}{2} + V_{\text{eff}}(r) = \varepsilon\,, \tag{10.10}$$

with the convenient definitions

$$V_{\text{eff}}(r) := \frac{MK}{r} + \frac{L^2}{2r^2} - \frac{ML^2}{r^3}\,, \quad \varepsilon := \frac{E^2 + K}{2}\,. \tag{10.11}$$

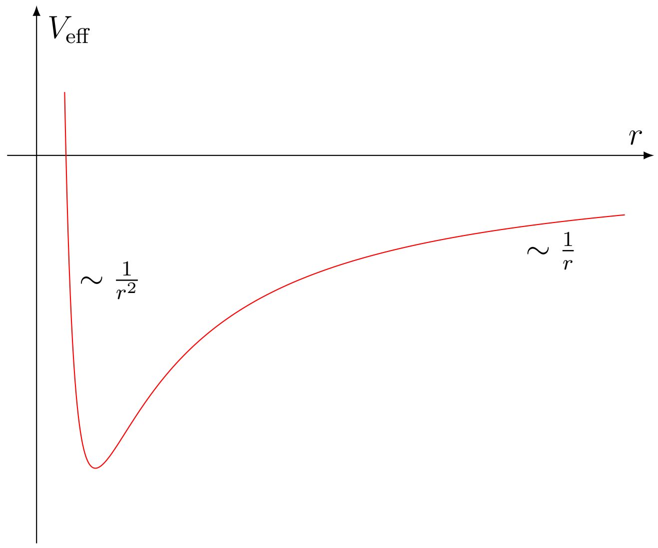


Figure 10.1. – Effective Potential in Newtonian physics.

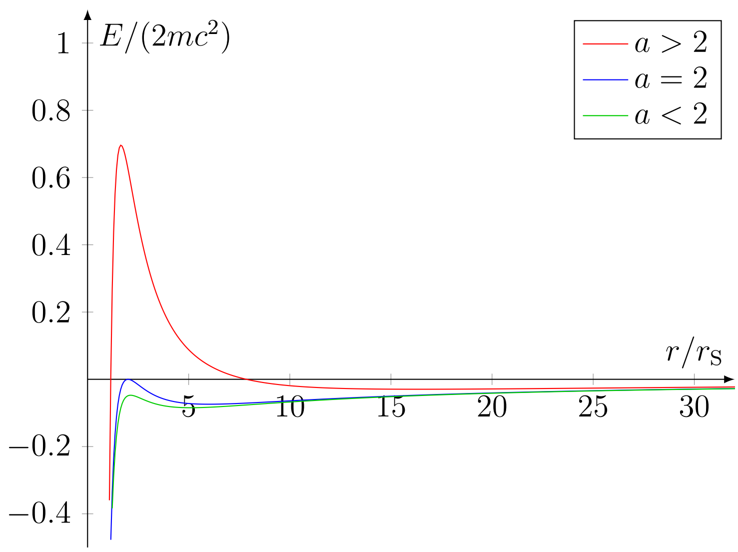


Figure 10.2. – Effective Potential in GR for various $a = L/(mr_S)$.

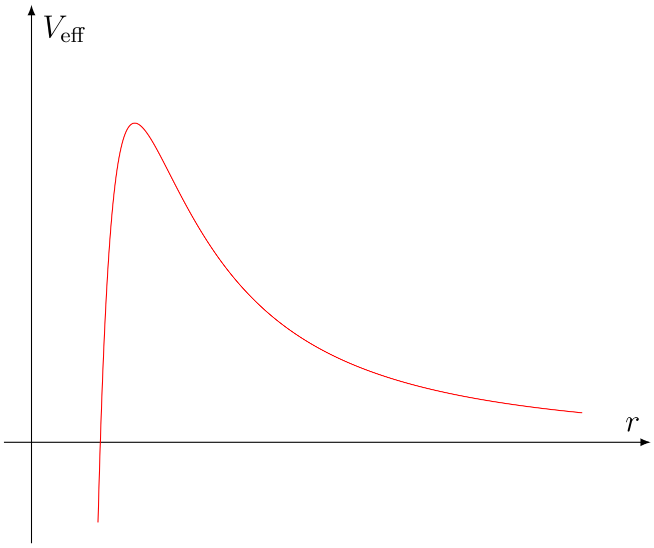


Figure 10.3. – Effective Potential for a photon.

10.1. Perihelion Shift of Mercury

It has been known for a long time that the perihelion of the mercury moves by roughly $5061''$[1]/century, if on subtracts other effects such a fraction of $43''$/century.

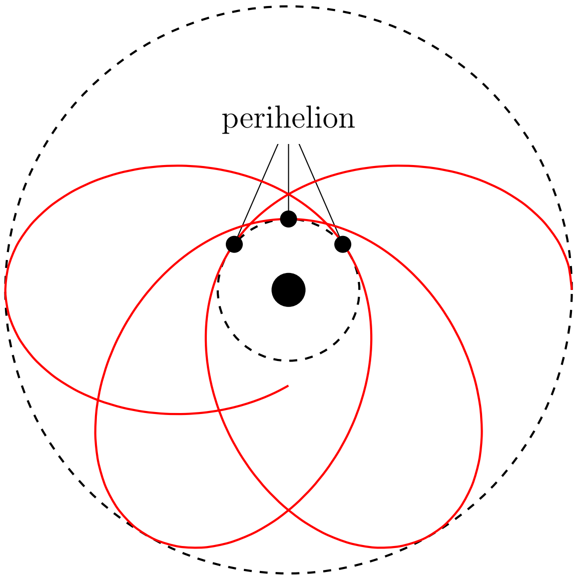


Figure 10.4. – Perihelion shift of mercury. (exaggerated)

There where several proposed explanations for this including

- a new planet called Vulcan between mercury and the sun,

[1]Where an arcsec $''$ denotes the 3600 part of an degree.

- a change to newtons $1/r^2$-law,
- effects due the suns quadrupole moment.

We will now look at what is predicted by GR. The trajectories of planets ($K = -1$) are given by

$$\dot{r}^2 - \frac{2M}{r} + \frac{L^2}{r^2} - \frac{2ML^2}{r^3} = E^2 - 1\,. \tag{10.12}$$

If we think of the radius as a function of the angle, we can write

$$\dot{r} = \frac{\mathrm{d}r}{\mathrm{d}\lambda} = \frac{\mathrm{d}r}{\mathrm{d}\phi}\frac{\mathrm{d}\phi}{\mathrm{d}\lambda}\,. \tag{10.13}$$

Multiplying (10.12) with $\left(\frac{\mathrm{d}\phi}{\mathrm{d}\lambda}\right)^{-2} = \frac{1}{\dot{\phi}} = \frac{r^4}{L^2}$ yields

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 - \frac{2Mr^3}{L^2} + r^2 - 2Mr = \left(E^2 - 1\right)\frac{r^4}{L^2}\,. \tag{10.14}$$

Similar to the treatment of the Keppler problem in classical mechanics, we define

$$u(\phi) := \frac{L^2}{Mr(\phi)}\,. \tag{10.15}$$

This implies

$$\mathrm{d}r = -\frac{L^2}{Mu^2}\,\mathrm{d}u\,, \quad \left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)^2 = \frac{L^4}{M^2u^4}\left(\frac{\mathrm{d}u}{\mathrm{d}\phi}\right)^2. \tag{10.16}$$

In terms of the new variables (10.14) reads

$$\left(\frac{\mathrm{d}u}{\mathrm{d}\phi}\right)^2 - 2u + u^2 - \frac{2M^2}{L^2}u^3 = (E^2 - 1)\frac{L^2}{M^2}\,. \tag{10.17}$$

We differentiate with respect to ϕ and get [2]

$$2u'u'' - 2u' + 2uu' - \frac{6M^2}{L^2}u^2u' = 0\,. \tag{10.18}$$

Dividing by $2u'$ yields

$$u'' - 1 + u = \frac{3M^2}{L^2}u^2\,, \tag{10.19}$$

which is an equation similar to the Keppler problem. The difference is in the right hand side, which is equal zero in the classical case. We may calculate an approximation by

[2] primes denote ϕ derivatives

pertubative corrections [3]

$$u = u_0 + u_1 + u_2 + \dots \,. \tag{10.20}$$

At zeroth order, we neglect the quadratic term, so that we are left with

$$u_0'' - 1 + u_0 = 0\,, \tag{10.21}$$

which is equal to the Newtonian problem, with exact solution

$$u_0(\phi) = 1 + \varepsilon\cos\phi\,. \tag{10.22}$$

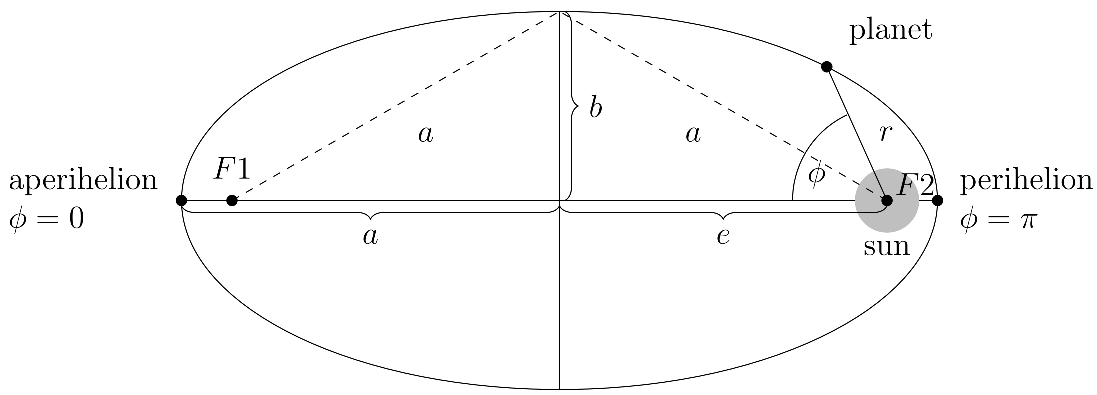


Figure 10.5. – An ellipse is the solution to the classical unperturbed Kepler problem.

At first order we have

$$\begin{aligned} u_1'' - 1 + u_1 &= \frac{3M^2}{L^2}u_0^2 \\ &= \frac{3M^2}{L^2}(1+\varepsilon\cos\phi)^2 \\ &= \frac{3M^2}{L^2}\left(1 + \frac{\varepsilon^2}{2} + 2\varepsilon\cos\phi + \frac{1}{2}\varepsilon^2\cos 2\phi\right)\,, \end{aligned} \tag{10.23}$$

Where we made use of trigonometric identities in the last step. The first order solution reads

$$u_1(\phi) = 1 + \varepsilon\cos\phi + \frac{3M^2}{L^2}\phi\sin\phi\,. \tag{10.24}$$

[3] The quantity relevant for the error of a given order is $\frac{3M^2}{L^2}$.

which can be checked by using

$$\frac{\mathrm{d}^2}{\mathrm{d}\phi^2}(\phi\sin\phi) + \phi\sin\phi = 2\cos\phi\,, \tag{10.25}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}\phi^2}(\cos 2\phi) + \cos 2\phi = -3\cos 2\phi\,. \tag{10.26}$$

Assuming that we still have approximately a movement along an ellipse and Taylor expansion yields

$$u(\phi) = 1 + \varepsilon\cos[(1-\delta)\phi] \simeq 1 + \varepsilon\cos\phi + \varepsilon\delta\phi\sin\phi + \mathcal{O}(\delta^2)\,. \tag{10.27}$$

Comparing this to the expression for u_1 we find

$$\delta = \frac{3M^2}{L^2} \tag{10.28}$$

and the error is of order δ^2. If we assume that there was a perihelion at ϕ_p, the next one will occure at $\phi + \Delta\Phi$ with

$$\Delta\phi = 2\pi\delta = \frac{6\pi M^2}{L^2} \tag{10.29}$$

The zeroth order solution gives

$$r_0(\phi) = \frac{L^2}{M(1+\varepsilon\cos\phi)} \stackrel{!}{=} \frac{a(1-\varepsilon^2)}{1+\varepsilon\cos\phi}\,. \tag{10.30}$$

So $L^2 \approx M(1-\varepsilon^2)a$ with an error of order δ, that can be neglected. With this approximation and restoring factors of c and G we arrive at the following expression for the perihelion shift:

$$\Delta\phi = \frac{6\pi G M_\odot}{c^2(1-\varepsilon^2)a}\,. \tag{10.31}$$

As expected the effect scalses with $1/a$ and thus the planet nearest to the sun recives the strongest effect. For mercury we have

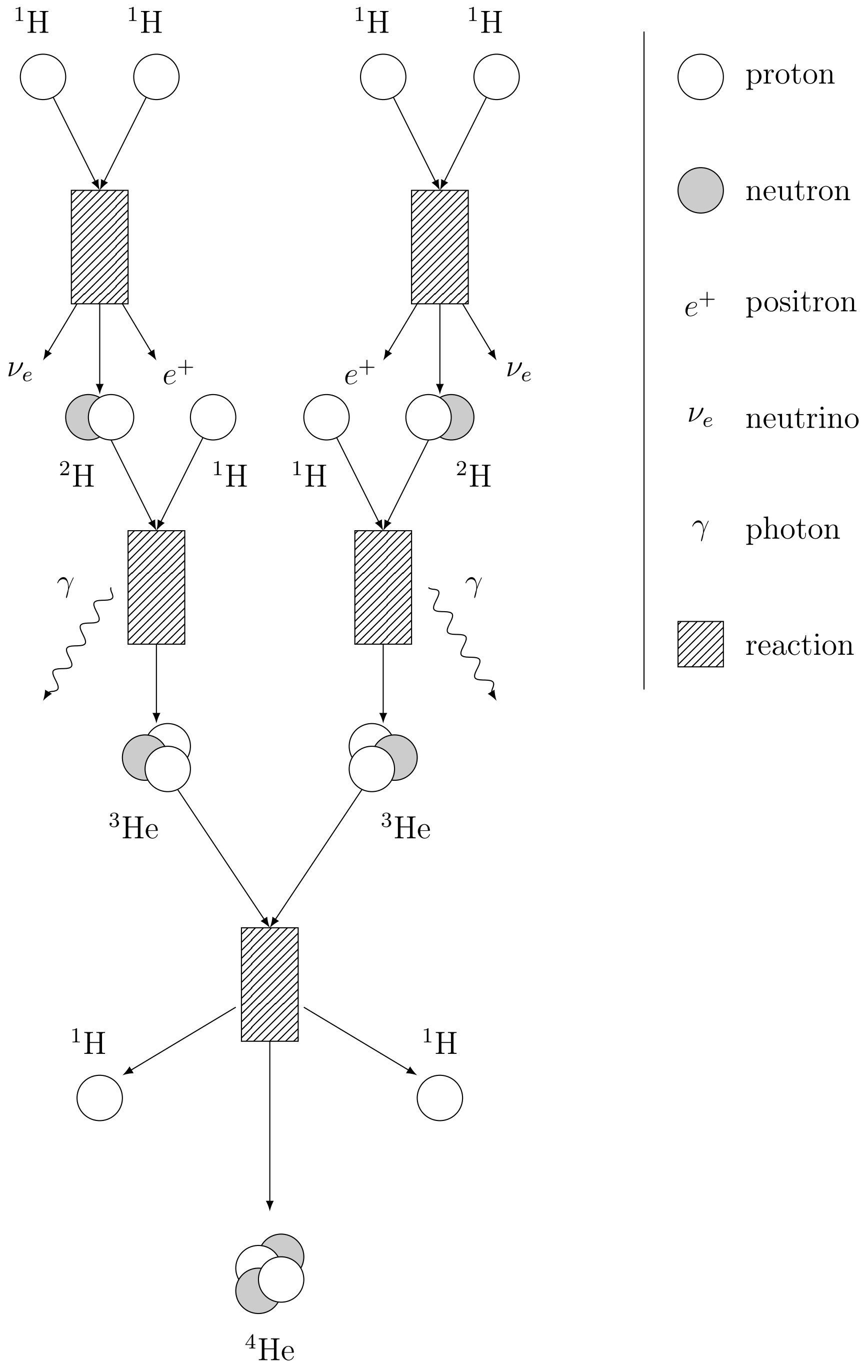
$$\frac{GM_\odot}{c^2} = 1.48\,\mathrm{km}\,, \quad a = 5.79\cdot 10^7\,\mathrm{km}\,, \quad \varepsilon = 0.2056\,, \tag{10.32}$$

Which results in an perihelionshift

$$\Delta\phi_{\mathrm{mer}} = 5.03\cdot 10^{-7}\,{}^{\mathrm{rad}}/_{\mathrm{orbit}} = 43\,{}^{\prime\prime}/_{\mathrm{century}}\,, \tag{10.33}$$

which fits the observation in the error margin.

11. Black Holes



^{1}H
^{1}H
^{1}H
^{1}H
ν_e
e^+
e^+
ν_e
^{2}H
^{1}H
^{1}H
^{2}H
γ
γ
^{3}He
^{3}He
^{1}H
^{1}H
^{4}He
proton
neutron
e^+ positron
ν_e neutrino
γ photon
reaction

Figure 11.1.

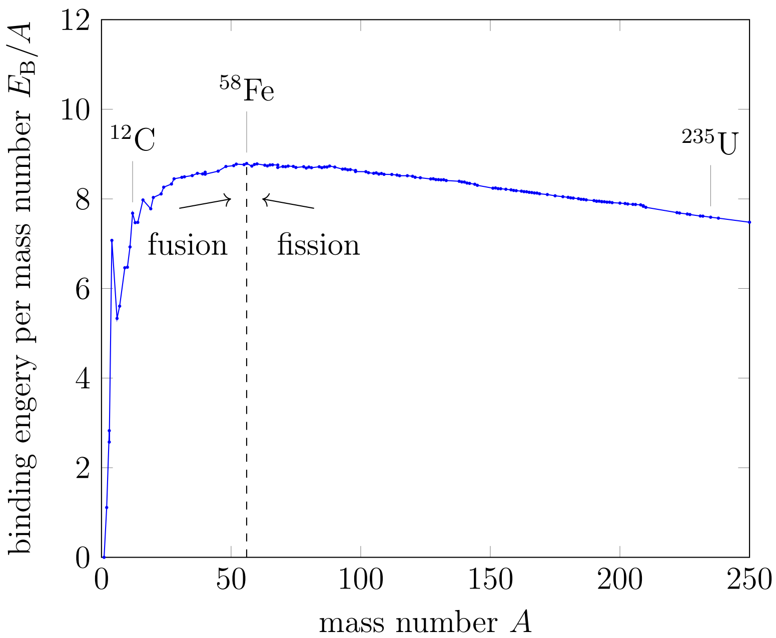


Figure 11.2.

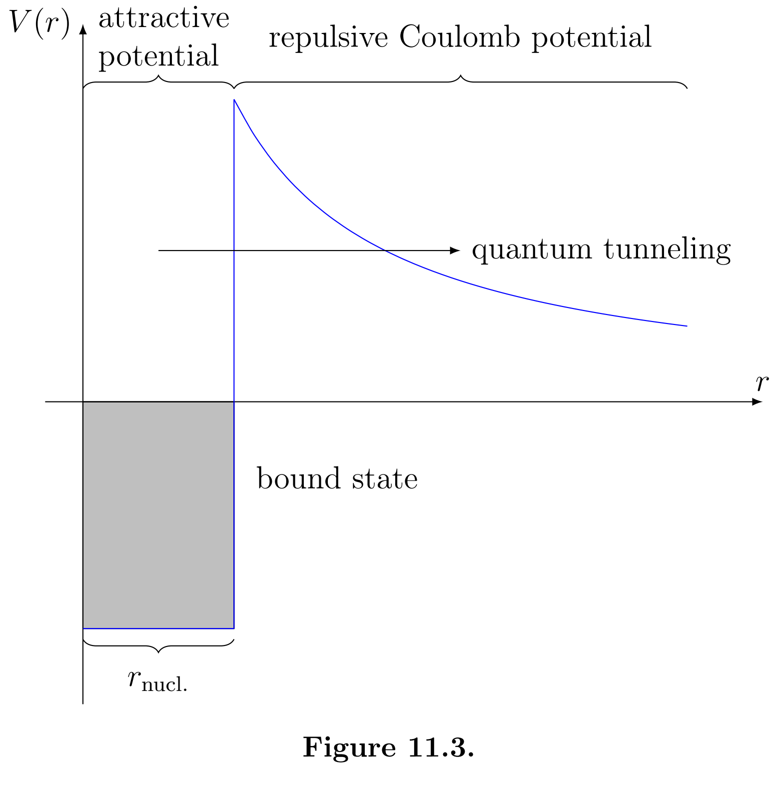


Figure 11.3.

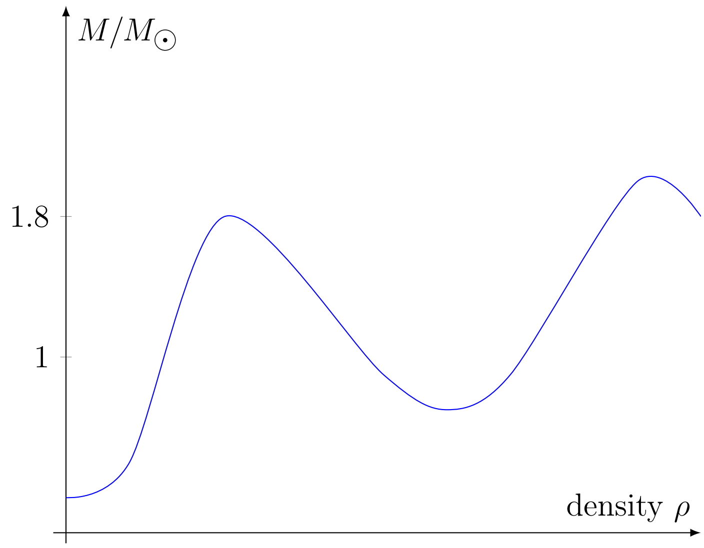


Figure 11.4.

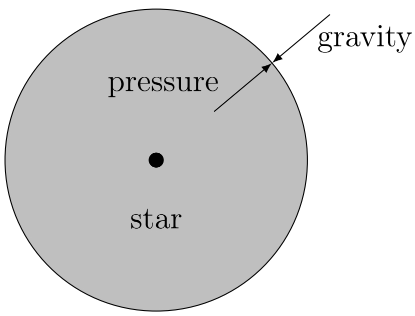


Figure 11.5.

12. Cosmology

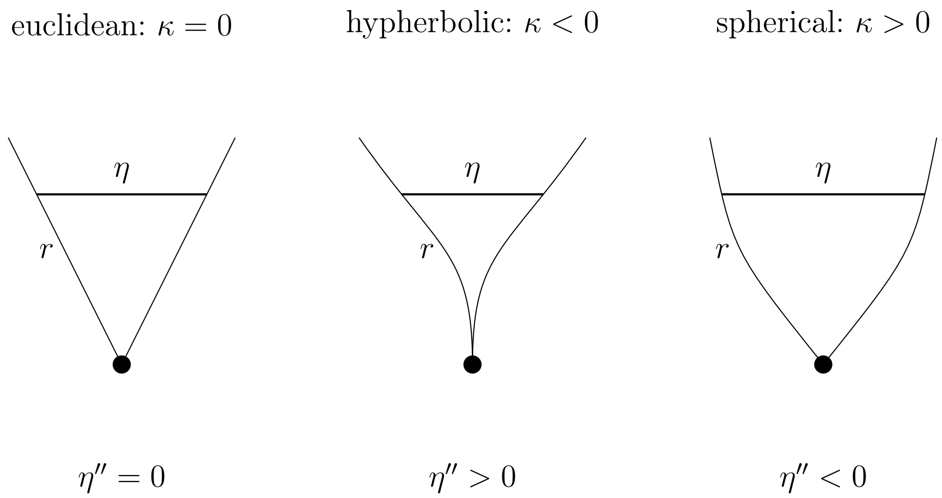


Figure 12.1. – Sectional curvature κ and shapes of the spatial part of spacetime.

A. Formulae

$$\delta g^{\mu\nu} = -g^{\varrho(\mu}g^{\nu)\sigma}\delta g_{\mu\nu} \tag{A.1}$$

$$\delta \det g = \det g g^{\mu\nu}\delta g_{\mu\nu} \tag{A.2}$$

$$\mathrm{div} A = \partial_i A^i = \begin{Bmatrix} i \\ ik \end{Bmatrix} A^k \tag{A.3}$$

$$\partial_k g = 2g \begin{Bmatrix} i \\ ik \end{Bmatrix} \tag{A.4}$$

$$\begin{Bmatrix} i \\ ik \end{Bmatrix} = \frac{1}{\sqrt{g}}\partial_k \sqrt{g} \tag{A.5}$$

$$\mathrm{div} A = \frac{1}{\sqrt{g}}\partial_k\left(\sqrt{g}A^k\right) \tag{A.6}$$

If further A is derived from a potential $A_i = \partial_i V$

$$\Box V = \mathrm{div} A = \frac{1}{\sqrt{g}}\partial_k\left(\sqrt{g}g^{ki}\partial_i V\right) \tag{A.7}$$