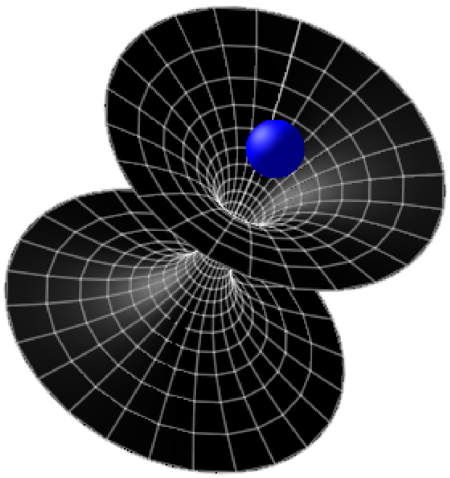
Christian Bär



# Theory of Relativity

Summer Term 2013



Version of August 26, 2013

# Contents

# Preface

These are the lecture notes of an introductory course on relativity theory that I gave in 2013. Ihe course was designed such that no prior knowledge of differential geometry was required. The course itself also did not introduce differential geometry (as it is often done in relativity classes). Instead, students unfamiliar with differential geometry had the opportunity to learn the subject in another course precisely set up for this purpose. This way, the relativity course could concentrate on its own topic. Of course, there is a price to pay; the first half of the course was dedicated to Special Relativity which does not require much mathematical background. Only the second half then deals with General Relativity. This gave the students time to acquire the geometric concepts.

The part on Special Relativity briefly recalls classical kinematics and electrodynamics emphasizing their conceptual incompatibility. It is then shown how Minkowski geometry is used to unite the two theories and to obtain what we nowadays call Special Relativity. Some famous phenomena like length contraction, time dilation, and the twin paradox are discussed. Relativistic velocity addition is investigated using hyperbolic geometry.

The part on General Relativity concentrates on the two most relevant models: Robertson-Walker spacetimes as models for the whole universe and the Schwarzschild model describing the vacuum neighborhood of a static star or black hole. The first model is used to discuss cosmic redshift, the expansion of the universe, big bang and big crunch. In the Schwarzschild model, we discuss the trajectories of massive particles and of light and see how they differ from the classical orbits.

It is my pleasure to thank all those who helped to improve the manuscript by suggestions, corrections or by work on the LaTeX code. My particular thanks go to Andrea Röser who wrote the first version in German language and created many pictures in wonderful quality, to Matthias Ludewig who translated the manuscript into English and to Ramona Ziese who improved the layout.

Potsdam, August 2013

Christian Bär

# 1 Special Relativity

Before starting with relativity theory we will briefly recall two older theories in physics, Newton's classical mechanics and Maxwell's electromagnetism theory. These two theories are incompatible in the sense that their laws transform differently under coordinate changes. This incompatibility was one of Einstein's main motivations to seek a theory that would combine the two. Einstein found a unification of mechanics and electromagnetism, now known under the term *special relativity theory*. In a way, Maxwell defeated Newton, the transformation laws of special relativity are those of electrodynamics. The laws of Newtonian mechanics are only valid approximately at low velocities.

## 1.1 Classical Kinematics

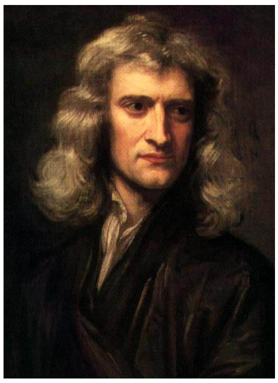
### Absolute Space

In Sir Isaac Newton's (1643-1727) world space exists independently of all the objects contained in it. In his own words:
*Absolute space, in its own nature, without regard to anything external, remains always similar and immovable.*
The geometry of space is assumed to be Euclidean, i.e., it is assumed that the laws of Euclidean geometry hold for measurements performed in physical space. In other words, we can introduce Cartesian coordinates to identify space with $\mathbb{R}^3$ and then apply the usual rules of Cartesian geometry,

$$\text{absolute space} \overset{\text{ident.}}{\longleftrightarrow} \mathbb{R}^3.$$

Such a coordinate system is not unique but the Euclidean structure is invariant under coordinate transformations of the form

$$\Phi : \mathbb{R}^3 \to \mathbb{R}^3, \quad \Phi(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b},$$



Godfrey Kneller's portrait of Isaac Newton (1689) [1]

with $\mathbf{A} \in \mathrm{O}(3)$, i.e., $\mathbf{A}^\top\mathbf{A} = \mathbf{I}$, and $\mathbf{b} \in \mathbb{R}^3$. The set of all such transformations is called the *Euclidean transformation group*.
It should be emphasized that such an assumption requires empirical justification. Indeed, measurements performed in every-day-life support Newton's ideas about space; if you measure the sum of angles in a triangle it will give 180 degrees to very high precision and is thus in accordance with Euclidean geometry.

[1] Source: `http://en.wikipedia.org/wiki/Isaac_Newton`

## Absolute Time

Newton's ideas about time are similar to those about space:
*Absolute, true and mathematical time, of itself, and from its own nature flows equably without regard to anything external.*
From a mathematical point of view, this means that we can measure time by a real parameter

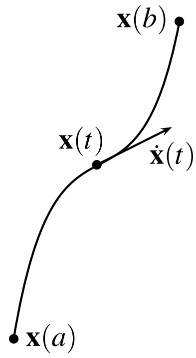
$$\text{absolute time} \overset{\text{ident.}}{\longleftrightarrow} \mathbb{R}.$$

More precisely, we fix a time interval, e.g., a second, and we then measure time in real multiples of this chosen time unit. The resulting identification of absolute time with $\mathbb{R}$ is unique up to transformations of the form

$$\mathbb{R} \to \mathbb{R}, \quad t \mapsto t + t_0,$$

with some fixed $t_0 \in \mathbb{R}$. Because we can distinguish future and past, we do not admit transformations of the form $\mathbb{R} \to \mathbb{R}$, $t \mapsto -t + t_0$, $t_0 \in \mathbb{R}$.
The trajectory of a point particle is described by a curve, i.e., by a map

$$\mathbf{x} : [a,b] \to \mathbb{R}^3,$$




where to each time coordinate $t$ we associate the corresponding space coordinates $\mathbf{x}(t) = (x^1(t), x^2(t), x^3(t))$ of the particle. Usually, we can and will assume that the curve $\mathbf{x}$ is smooth, $\mathbf{x} \in C^\infty([a,b], \mathbb{R}^3)$. The *velocity* of the particle is then given by

$$\dot{\mathbf{x}} : [a,b] \to \mathbb{R}^3$$

and the *acceleration* by

$$\ddot{\mathbf{x}} : [a,b] \to \mathbb{R}^3.$$

We measure the *mass* of the particle in real multiples of a fixed unit mass, like kilogram. Hence mass is mathematically given by a function

$$m : [a,b] \to \mathbb{R}.$$

The *momentum* is then given by

$$\mathbf{p} = m \cdot \dot{\mathbf{x}} : [a,b] \to \mathbb{R}^3$$

and the *kinetic energy* by

$$E = \frac{m \cdot \|\dot{\mathbf{x}}\|^2}{2} : [a,b] \to \mathbb{R}.$$

Finally, *the length of the trajectory* swept out by the particle can by calculated by the formula

$$\int_a^b \|\dot{\mathbf{x}}(t)\| \, dt.$$

A choice of space and time coordinates as described above will be called an *inertial frame*. According to Newton we can check whether or not our chosen coordinate system is "correct" as follows:

**Newton's First Law**
*In any inertial frame, particles that are not subject to any force, are characterized by*

$$\ddot{\mathbf{x}} = 0,$$

*which is equivalent to* $\mathbf{x}(t) = \mathbf{x}(0) + t \cdot \dot{\mathbf{x}}(0)$.

The transformations of space and time that were discussed above map inertial frames to other inertial frames. On top of that, we can admit an inertial frame the origin of which moves with constant velocity to the other. This leads to the following set of transformations mapping inertial frames to inertial frames, the so-called *Galilean transformations*.

$$\begin{aligned} \mathbb{R} \times \mathbb{R}^3 &\to \mathbb{R} \times \mathbb{R}^3, \\ \begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} &\mapsto \begin{pmatrix} t + t_0 \\ \mathbf{A}\mathbf{x} + \mathbf{b_0} + t\mathbf{b_1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \mathbf{b_1} & \mathbf{A} \end{pmatrix} \cdot \begin{pmatrix} t \\ \mathbf{x} \end{pmatrix} + \begin{pmatrix} t_0 \\ \mathbf{b_0} \end{pmatrix}, \end{aligned}$$

where $\mathbf{A} \in \mathrm{O}(3)$, $\mathbf{b_0}, \mathbf{b_1} \in \mathbb{R}^3$ and $t_0 \in \mathbb{R}$.
If $\mathbf{x}$ is the trajectory of a particle in one inertial frame, its trajectory, velocity and acceleration in another inertial frame take the form

$$\begin{aligned} \tilde{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{b_0} + t\mathbf{b_1}, \\ \dot{\tilde{\mathbf{x}}} &= \mathbf{A}\dot{\mathbf{x}} + \mathbf{b_1}, \\ \ddot{\tilde{\mathbf{x}}} &= \mathbf{A}\ddot{\mathbf{x}}. \end{aligned}$$

Observe that $\ddot{\mathbf{x}} = 0$ if and only if $\ddot{\tilde{\mathbf{x}}} = 0$, so indeed, Newton's first law is compatible with Galilean transformations.
In the special case $\mathbf{A} = \mathbf{I}$ and $t_0 = 0$, we have
$-\mathbf{v} := \mathbf{b_1}$ = velocity of observer 2, measured by observer 1
$\mathbf{v_1} := \dot{\mathbf{x}}$ = velocity of the particle, measured by observer 1
$\mathbf{v_2} := \dot{\tilde{\mathbf{x}}}$ = velocity of the particle, measured by observer 2
Hence we have derived the velocity-addition formula

$$\boxed{\mathbf{v_1} = \mathbf{v} + \mathbf{v_2}} \tag{1.1}$$

**Newton's second law**
*In any inertial system, if a particle is subject to the force* $\mathbf{F}$, *then*

$$\tfrac{d}{dt}(m(t)\dot{\mathbf{x}}(t)) = \mathbf{F}(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)).$$

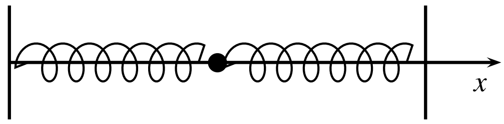
Here a force is described by a (smooth) mapping of the form $\mathbf{F} : \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$. In particular, for $m > 0$ we have

$$\ddot{\mathbf{x}}(t) = \frac{1}{m(t)} \big(\mathbf{F}(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)) - \dot{m}(t)\dot{\mathbf{x}}(t)\big).$$

A solution of such an ordinary differential equation is uniquely determined by its initial values

$$\mathbf{x}(t_0) \quad \text{and} \quad \dot{\mathbf{x}}(t_0).$$

Therefore, the theory is deterministic (we can predict the future, given initial values).




**Example 1.1.1.** A mass $m$ is suspended between two springs with spring constant $k > 0$. We want to find the equations of motion, given $x(0)$ and $\dot{x}(0)$.
By Hooke's law, the force is $F(t,x,y) = -kx$. From this it follows that $m\ddot{x}(t) = -kx(t)$, hence $\ddot{x}(t) = -\frac{k}{m}x(t)$. This ODE has the general solution

$$x(t) = A \cdot \sin\left(\sqrt{\frac{k}{m}}t\right) + B \cdot \cos\left(\sqrt{\frac{k}{m}}t\right),$$

where $B = x(0)$ and $\dot{x}(0) = A\sqrt{\frac{k}{m}}$. Therefore

$$x(t) = \dot{x}(0) \cdot \sqrt{\frac{m}{k}} \cdot \sin\left(\sqrt{\frac{k}{m}}t\right) + x(0) \cdot \cos\left(\sqrt{\frac{k}{m}}t\right).$$

### Energy Equation

Let us assume that the mass $m$ is constant. We differentiate the kinetic energy of a particle and obtain the energy equation

$$\frac{d}{dt}E = \frac{m}{2}\frac{d}{dt}\|\dot{\mathbf{x}}\|^2 = m\langle\ddot{\mathbf{x}},\dot{\mathbf{x}}\rangle = \langle\mathbf{F},\mathbf{v}\rangle \tag{1.2}$$

**Exercise 1.1.2.** A spacecraft travels from earth to a distant object $X$, its rear engine inducing constant acceleration (=gravitational acceleration) $g = 9.81ms^{-2}$. At half the distance, the spacecraft is turned over (so its rear engine now induces the same deceleration).
How long does the journey take and what is the maximal velocity for

(a) $X$ = moon (400.000 km),

(b) $X$ = mars (56-400 million km),

(c) $X$ = Proxima Centauri (4,3 light years) und

(d) $X$ = Andromeda galaxy (2 million lightyears)?

Remember: 1 light year $\approx 9.461 \cdot 10^{12}km$.

*Solution.* We write $x(t)$ for the distance from space craft to earth at time $t$. From $x(0) = 0$, $\dot{x}(0) = 0$ and $\ddot{x} = g$, we get

$$x(t) = \tfrac{1}{2}gt^2$$

for the first half of the journey. Is $D$ the distance between earth and the object $X$ and $T$ is the total time of travel, we obtain $D/2 = \frac{1}{2}g(T/2)^2$ and therefore

$$T = 2\sqrt{D/g}.$$

The maximal velocity $v_{\max}$ is achieved after the time $T/2$, just before initiating the deceleration. From $\dot{x}(t) = gt$, we obtain

$$v_{\max} = g\frac{T}{2} = \sqrt{gD}.$$

Plugging in the different values for $D$ results in

1. $X$ = moon: $T \approx 3,5$ h, $v_{\max} \approx 63$ km/s.
2. $X$ = Mars: $T \approx 42 - 112$ hours, $v_{\max} \approx 742 - 1980$ km/s.
3. $X$ = Proxima Centauri: $T \approx 4$ years, $v_{\max} \approx 2,1\ c$.
4. $X$ = Andromeda galaxy: $T \approx 2784$ years, $v_{\max} \approx 1434\ c$.

## 1.2 Electrodynamics

Now we turn to Maxwell's electrodynamics. Assume that we are in a vacuum without any electric charges present. In this case, electric and magnetic phenomena are described by functions

$$f : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R},$$

that solve the *wave equation*, i.e.,

$$\Box f := \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} - \Delta f = 0,$$

where $\Delta = \sum_{i=1}^{3} \frac{\partial^2}{\partial (x^i)^2}$ is the *Laplace operator* and $c$ the speed of light in vacuum (about $300,000$ km/s). As all observers in an inertial frame have equal right, the question arises which transformations preserve the wave equation. More precisely, which are the transformations

James Clerk Maxwell (1831–1879) [2]

$$\Phi : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R} \times \mathbb{R}^3, \quad \Phi(\mathbf{x}) = \mathbf{L}\mathbf{x} + \begin{pmatrix} t_0 \\ \mathbf{b_0} \end{pmatrix},$$

with $\mathbf{L} \in \mathrm{Mat}(4 \times 4, \mathbb{R})$ such that whenever $f$ solves the wave equation, so does $\tilde{f} := f \circ \Phi$? To find out, we set $x^0 := c \cdot t$ and $\mathbf{x} := (x^0, x^1, x^2, x^3)$. The wave equation then is

$$\Box f = \frac{\partial^2 f}{\partial (x^0)^2} - \sum_{i=1}^{3} \frac{\partial^2 f}{\partial (x^i)^2} = 0.$$

We now calculate $\Box \tilde{f}$. To this end, write

$$\mathbf{L} = (\mathbf{L_0}, \mathbf{L_1}, \mathbf{L_2}, \mathbf{L_3}) = \begin{pmatrix} L^0_0 & L^0_1 & L^0_2 & L^0_3 \\ L^1_0 & L^1_1 & L^1_2 & L^1_3 \\ L^2_0 & L^2_1 & L^2_2 & L^2_3 \\ L^3_0 & L^3_1 & L^3_2 & L^3_3 \end{pmatrix},$$

[2] Source: `http://de.wikipedia.org/wiki/James_Clerk_Maxwell`

where $\mathbf{L_i} \in \mathbb{R}^4$ for $i = 0, \dots, 3$ and $\mathbf{x_0} := (t_0, \mathbf{b_0})^\top$. We compute

$$\begin{aligned}\frac{\partial \tilde{f}}{\partial x^i} &= \frac{\partial}{\partial x^i} f(\mathbf{L_0}x^0 + \mathbf{L_1}x^1 + \mathbf{L_2}x^2 + \mathbf{L_3}x^3 + \mathbf{x_0}) \\ &= \sum_{m=0}^{3} \frac{\partial f}{\partial x^m}(\mathbf{Lx} + \mathbf{x_0}) \cdot L_i^m,\end{aligned}$$

hence

$$\frac{\partial^2 \tilde{f}}{\partial (x^i)^2} = \sum_{m,n=0}^{3} \frac{\partial^2 f}{\partial x^m \partial x^n}(\mathbf{Lx} + \mathbf{x_0}) L_i^m L_i^n. \tag{1.3}$$

Define

$$\mathbf{I}_{1,3} := \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The Hessian matrix of a twice continuously differentiable function $f$ is the symmetric matrix $\mathrm{hess} f = (\frac{\partial^2 f}{\partial x^i \partial x^j})_{i,j}$. On the vector space of the symmetric $(n \times n)$-matrices, we can define the following scalar product:

$$(\mathbf{A}, \mathbf{B})_S := \sum_{i,j=1}^{n} A_j^i B_j^i.$$

Then, by definition of the $\square$-operator,

$$\square f = (\mathrm{hess} f, \mathbf{I}_{1,3})_S.$$

By (1.3) we have

$$\square \tilde{f} = (\mathbf{L}^\top \cdot \mathrm{hess} f \cdot \mathbf{L}, \mathbf{I}_{1,3})_S = (\mathrm{hess} f, \mathbf{L}^\top \cdot \mathbf{I}_{1,3} \cdot \mathbf{L})_S.$$

We see that $\square f = 0$ means that $\mathrm{hess} f$ is perpendicular to $\mathbf{I}_{1,3}$ while $\square \tilde{f} = 0$ means that $\mathrm{hess} f$ is perpendicular to $\mathbf{L}^\top \cdot \mathbf{I}_{1,3} \cdot \mathbf{L}$. These two conditions are equivalent if and only if $\mathbf{L}^\top \cdot \mathbf{I}_{1,3} \cdot \mathbf{L} = \kappa \cdot \mathbf{I}_{1,3}$ for some $\kappa \in \mathbb{R}$. Without loss of generality we will assume $\kappa = 1$ for the scaling factor, because a transformation of the form $\kappa \cdot \mathbf{I}$ just corresponds to a change of the physical unit of length.

**Definition 1.2.1.** The set of transformations

$$\mathscr{L} := \{\mathbf{L} \in \mathrm{Mat}(4 \times 4, \mathbb{R}) \,|\, \mathbf{L}^\top \cdot \mathbf{I}_{1,3} \cdot \mathbf{L} = \mathbf{I}_{1,3}\}$$

is called the *Lorentz group*. The corresponding set of affine-linear transformations

$$\mathscr{P} := \{\Phi : \mathbb{R}^4 \to \mathbb{R}^4 \,|\, \Phi(\mathbf{x}) = \mathbf{Lx} + \mathbf{x_0},\, \mathbf{L} \in \mathscr{L},\, \mathbf{x_0} \in \mathbb{R}^4\}$$

is called the *Poincaré group*.

We have seen that the "admissible" coordinate transformations of Newtonian mechanics are the Galilean transformations while those for electrodynamics are the Poincaré transformations. These two groups are not contained in one another, in this sense classical kinematics and electrodynamics are incompatible.

**Exercise 1.2.2.** Determine all Poincaré transformations of $\mathbb{R}^4$ which are also Galilean transformations. Why is this subgroup of transformations not sufficient to derive (1.1)?

Henri Poincaré (1854–1912) [3]

Now the question is: Which theory is correct if any?
One should be able to answer this question by means of suitable experiments. For instance, look at the following situation: Observer 2 is located on a spacecraft that travels towards earth and sends a light signal to observer 1 on the earth. Observer 1 measures the velocity $c_1$ for the incoming light signal, observer 2 however measures another velocity $c_2$. According to (1.1), classical kinematics predicts $c_1 = c_2 + v$, where $v$ is the velocity of the spacecraft with respect to the earth. On the other hand, the theory of electrodynamics states that the speed of light in vacuum is a fixed value $c$, independently of the motions of the source and the observer.
In fact, experiments, such as the famous Michelson-Morley experiment, have confirmed the predictions of electrodynamics!

## 1.3 The Lorentz group and Minkowski geometry


Hermann Minkowski (1864 –1909) [4]

In order to develop a kinematic theory which is invariant under Poincaré transformations we first need to understand these Poincaré transformations better. The crucial part are the Lorentz transformations because adding translations then yields all Poincaré transformations. The resulting geometry of lightlike, timelike and spacelike vectors is known as Minkowski geometry, named after the mathematician Hermann Minkowski, a close friend of David Hilbert.

**Convention.** For $(x^0, x^1, x^2, x^3) \in \mathbb{R}^4$ write $(x^0, \hat{\mathbf{x}})$ with $\hat{\mathbf{x}} := (x^1, x^2, x^3)$. We write $\langle \cdot, \cdot \rangle$ for the usual scalar product in $\mathbb{R}^4$, i.e.,

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{4} x^i y^i.$$

We further define another inner product $\langle\langle \cdot, \cdot \rangle\rangle$ on $\mathbb{R}^4$ by

$$\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle := \langle \mathbf{x}, \mathbf{I}_{1,3} \cdot \mathbf{y} \rangle = -x^0 y^0 + \langle \hat{\mathbf{x}}, \hat{\mathbf{y}} \rangle .$$

[3] Source: `http://en.wikipedia.org/wiki/Henri_Poincare`
[4] Source: `http://en.wikipedia.org/wiki/Hermann_Minkowski`

The symmetric bilinear form $\langle\langle\cdot,\cdot\rangle\rangle$ is indefinite and non-degenerate. Recall that "non-degenerate" means that $\langle\langle \mathbf{x},\mathbf{y}\rangle\rangle = 0$ for all $\mathbf{x}\in\mathbb{R}^4$ implies that $\mathbf{y}=0$.

By definition, $\mathbf{L}\in\mathscr{L}$ if and only if $\mathbf{L}^\top\mathbf{I_{1,3}}\mathbf{L} = \mathbf{I_{1,3}}$. This is equivalent to

$$\left\langle \mathbf{x}, \mathbf{L}^\top\mathbf{I_{1,3}}\mathbf{L}\mathbf{y}\right\rangle = \langle \mathbf{x}, \mathbf{I_{1,3}}\mathbf{y}\rangle$$

for all $\mathbf{x},\mathbf{y}\in\mathbb{R}^4$. For the left-hand-side we get $\left\langle \mathbf{x}, \mathbf{L}^\top\mathbf{I_{1,3}}\mathbf{L}\mathbf{y}\right\rangle = \langle \mathbf{L}\mathbf{x}, \mathbf{I_{1,3}}\mathbf{L}\mathbf{y}\rangle = \langle\langle \mathbf{L}\mathbf{x},\mathbf{L}\mathbf{y}\rangle\rangle$ while the right-hand-side is $\langle \mathbf{x}, \mathbf{I_{1,3}}\mathbf{y}\rangle = \langle\langle \mathbf{x},\mathbf{y}\rangle\rangle$. Hence we have obtained another characterization of the Lorentz group as

$$\mathscr{L} = \{\mathbf{L}\in \mathrm{Mat}(4\times 4,\mathbb{R}) \,|\, \langle\langle \mathbf{L}\mathbf{x},\mathbf{L}\mathbf{y}\rangle\rangle = \langle\langle \mathbf{x},\mathbf{y}\rangle\rangle \text{ for all } \mathbf{x},\mathbf{y}\in\mathbb{R}^4\}.$$

This formally resembles the definition of the orthogonal group $\mathrm{O}(n)$, which by definition is

$$\mathrm{O}(n) = \{\mathbf{A}\in \mathrm{Mat}(n\times n,\mathbb{R}) \,|\, \langle \mathbf{A}\mathbf{x},\mathbf{A}\mathbf{y}\rangle = \langle \mathbf{x},\mathbf{y}\rangle \text{ for all } \mathbf{x},\mathbf{y}\in\mathbb{R}^n\}.$$

**Definition 1.3.1.** We call $(\mathbb{R}^4, \langle\langle\cdot,\cdot\rangle\rangle)$ the (4-dimensional) *Minkowski space*. The inner product $\langle\langle\cdot,\cdot\rangle\rangle$ is called the *Minkowski product*.

Any Lorentz transformation $\mathbf{L}\in\mathscr{L}$ has the following properties:

1. $\det(\mathbf{I_{1,3}}) = \det(\mathbf{L}^\top\mathbf{I_{1,3}}\mathbf{L}) = \det(\mathbf{L})^2\cdot\det(\mathbf{I_{1,3}})$, hence $\det(\mathbf{L}) = \pm 1$.
2. We have
$$-1 = (\mathbf{I_{1,3}})^0_0 = (\mathbf{L}^\top\mathbf{I_{1,3}}\mathbf{L})^0_0 = -(L^0_0)^2 + \sum_{i=1}^3 (L^0_i)^2,$$
thus
$$(L^0_0)^2 = 1 + \sum_{i=1}^3 (L^0_i)^2.$$
In particular, we have $(L^0_0)^2 \geq 1$, i.e., $L^0_0 \geq 1$ or $L^0_0 \leq -1$.

**Definition 1.3.2.** We define the following subsets of $\mathbf{L}$:

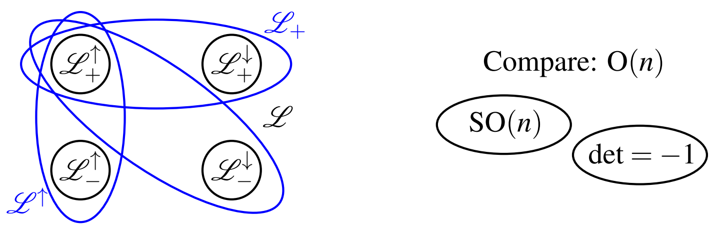
$$\begin{aligned}
\mathscr{L}^\uparrow_+ &:= \{\mathbf{L}\in\mathscr{L} \,|\, \det L = +1, L^0_0 \geq +1\},\\
\mathscr{L}^\uparrow_- &:= \{\mathbf{L}\in\mathscr{L} \,|\, \det L = -1, L^0_0 \geq +1\},\\
\mathscr{L}^\downarrow_+ &:= \{\mathbf{L}\in\mathscr{L} \,|\, \det L = +1, L^0_0 \leq -1\},\\
\mathscr{L}^\downarrow_- &:= \{\mathbf{L}\in\mathscr{L} \,|\, \det L = -1, L^0_0 \leq -1\}.
\end{aligned}$$

**Remark 1.3.3.** The subset $\mathscr{L}_+^\uparrow$ is a subgroup of $\mathscr{L}$, see Exercise 1.3.4, the other three subsets are not because they do not contain the identity matrix. We make the further assignments for subsets of $\mathscr{L}$:

| | |
|---|---|
| orientation preserving Lorentz tranformations: | $\mathscr{L}_+ := \mathscr{L}_+^\uparrow \sqcup \mathscr{L}_+^\downarrow,$ |
| time orientation preserving Lorentz tranformations: | $\mathscr{L}^\uparrow := \mathscr{L}_+^\uparrow \sqcup \mathscr{L}_-^\uparrow,$ |
| space orientation preserving Lorentz tranformations: | $\mathscr{L}_+^\uparrow \sqcup \mathscr{L}_-^\downarrow.$ |

**Exercise 1.3.4.** a) Show that $\mathscr{L}_+^\uparrow \cdot \mathscr{L}_+^\uparrow \subset \mathscr{L}_+^\uparrow$, $\mathscr{L}_+^\uparrow \cdot \mathscr{L}_+^\downarrow \subset \mathscr{L}_+^\downarrow$, and similarly for all other combinations.
b) Conclude from a) that $\mathscr{L}_+^\uparrow$, $\mathscr{L}_+$, $\mathscr{L}^\uparrow$, and $\mathscr{L}_+^\uparrow \sqcup \mathscr{L}_-^\downarrow$ are subgroups of $\mathscr{L}$.

**Remark 1.3.5.** The subsets $\mathscr{L}_+^\uparrow$, $\mathscr{L}_+^\downarrow$, $\mathscr{L}_-^\uparrow$ and $\mathscr{L}_-^\downarrow$ are the connected components of the Lorentz group.




The Lorentz group contaings the following special elements.

1. $\mathbf{L} = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \mathbf{A} \end{array}\right)$, $\mathbf{A} \in \mathrm{O}(3)$, for example space-like rotations:

$$\underset{\text{rotation about } x^3\text{-axis}}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi & 0 \\ 0 & \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}} \quad \underset{\text{rotation about } x^2\text{-axis}}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\varphi & 0 & -\sin\varphi \\ 0 & 0 & 1 & 0 \\ 0 & \sin\varphi & 0 & \cos\varphi \end{pmatrix}}$$

   or space-like reflections

$$\underset{\text{reflection at } x^2,x^3\text{-plane}}{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}$$

2. *Boosts* are special Lorentz transformations which mix space and time components, for example

$$\mathbf{L_1} = \begin{pmatrix} \cosh\varphi & \sinh\varphi & 0 & 0 \\ \sinh\varphi & \cosh\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \mathbf{L_2} = \begin{pmatrix} \cosh\varphi & 0 & \sinh\varphi & 0 \\ 0 & 1 & 0 & 0 \\ \sinh\varphi & 0 & \cosh\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.4}$$

**Lemma 1.3.6 (Hyperbolic identities)**
*For all $\varphi_1, \varphi_2 \in \mathbb{R}$, we have*

1. $\cosh(\varphi_1 + \varphi_2) = \cosh(\varphi_1)\cosh(\varphi_2) + \sinh(\varphi_1)\sinh(\varphi_2)$.
2. $\sinh(\varphi_1 + \varphi_2) = \cosh(\varphi_1)\sinh(\varphi_2) + \sinh(\varphi_1)\cosh(\varphi_2)$.

*Proof.* By definition, $\cosh\varphi = \frac{1}{2}(e^{\varphi} + e^{-\varphi})$ and $\sinh\varphi = \frac{1}{2}(e^{\varphi} - e^{-\varphi})$, hence

$$e^{\varphi} = \cosh\varphi + \sinh\varphi. \tag{1.5}$$

Inserting (1.5) into

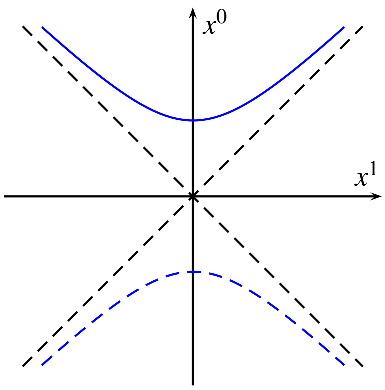
$$\cosh(\varphi_1 + \varphi_2) = \frac{1}{2}(e^{\varphi_1+\varphi_2} + e^{-(\varphi_1+\varphi_2)}) = \frac{1}{2}(e^{\varphi_1}e^{\varphi_2} + e^{-\varphi_1}e^{-\varphi_2})$$

yields (a) and similarly for (b). □

**Remark 1.3.7.** We have $\cosh\varphi - \sinh\varphi = \cosh(-\varphi) + \sinh(-\varphi) = e^{-\varphi}$. Multiplication with (1.5) gives

$$1 = e^{\varphi} \cdot e^{-\varphi} = (\cosh\varphi + \sinh\varphi) \cdot (\cosh\varphi - \sinh\varphi) = \cosh^2\varphi - \sinh^2\varphi.$$

Geometrically, this means that for each $\varphi \in \mathbb{R}$ the point $(\cosh\varphi, \sinh\varphi)^\top$ lies on the upper branch of the hyperbola in $\mathbb{R}^2$ given by $(x^0)^2 = 1 + (x^1)^2$. In fact, $\varphi \mapsto (\cosh\varphi, \sinh\varphi)^\top$ maps $\mathbb{R}$ bijectively onto this curve.



**Lemma 1.3.8 (Hyperbolic angular identities)**
*For all $\varphi_1, \varphi_2 \in \mathbb{R}$, we have*

$$\begin{pmatrix} \cosh\varphi_1 & \sinh\varphi_1 \\ \sinh\varphi_1 & \cosh\varphi_1 \end{pmatrix} \cdot \begin{pmatrix} \cosh\varphi_2 & \sinh\varphi_2 \\ \sinh\varphi_2 & \cosh\varphi_2 \end{pmatrix} = \begin{pmatrix} \cosh(\varphi_1+\varphi_2) & \sinh(\varphi_1+\varphi_2) \\ \sinh(\varphi_1+\varphi_2) & \cosh(\varphi_1+\varphi_2) \end{pmatrix}.$$

*Proof.* This follows directly from Lemma 1.3.6. □

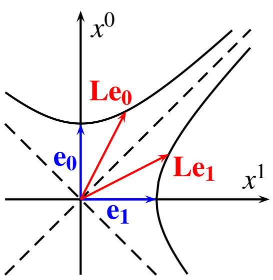
This lemma tells us for instance that the boosts of the form

$$\begin{pmatrix} \cosh\varphi & \sinh\varphi & 0 & 0 \\ \sinh\varphi & \cosh\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

form a subgroup of the Lorentz group. More precisely,

$$\mathbb{R} \to \mathscr{L}, \quad \varphi \mapsto \begin{pmatrix} \cosh\varphi & \sinh\varphi & 0 & 0 \\ \sinh\varphi & \cosh\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is a group homomorphism. The first and second column traces a hyperbola when $\varphi$ runs through $\mathbb{R}$.




$$\mathbf{L} = \begin{pmatrix} \cosh\varphi & \sinh\varphi \\ \sinh\varphi & \cosh\varphi \end{pmatrix}$$

$\mathbf{Le_0}$ perpendicular to $\mathbf{Le_1}$ with respect to $\langle\langle\cdot,\cdot\rangle\rangle$.

**Definition 1.3.9.** A vector $\mathbf{v} \in \mathbb{R}^4$ is called

- *timelike* iff $\langle\langle \mathbf{v}, \mathbf{v} \rangle\rangle < 0$,

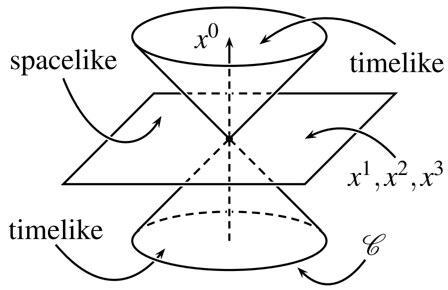
- *lightlike* iff $\langle\langle \mathbf{v},\mathbf{v}\rangle\rangle = 0$ and $\mathbf{v} \neq 0$,
- *spacelike* iff $\langle\langle \mathbf{v},\mathbf{v}\rangle\rangle > 0$ or $\mathbf{v} = 0$.

The set $\mathscr{C} := \{\mathbf{v} \in \mathbb{R}^4 \,|\, \mathbf{v} \text{ lightlike } \}$ is called the *light cone*.

We observe that $\mathbf{v} = (v^0, \hat{\mathbf{v}})$ is lightlike if and only if

$$0 = -(v^0)^2 + \|\hat{\mathbf{v}}\|^2 \iff |v^0| = \|\hat{\mathbf{v}}\|.$$

This is the equation of a cone, hence the terminology “light cone”.




**Remark 1.3.10.** The set $\mathscr{T} := \{\mathbf{v} \in \mathbb{R}^4 \,|\, \mathbf{v} \text{ timelike } \}$ is open (the ”interior” of the light cone) and decomposes into two components

$$\mathscr{T}^{\uparrow} := \{\mathbf{v} \in \mathscr{T} \,|\, v^0 > 0\} \quad \text{and} \quad \mathscr{T}^{\downarrow} := \{\mathbf{v} \in \mathscr{T} \,|\, v^0 < 0\}.$$

**Remark 1.3.11.** Since $\langle\langle \mathbf{Lv},\mathbf{Lv}\rangle\rangle = \langle\langle \mathbf{v},\mathbf{v}\rangle\rangle$ for all vectors and Lorentz tranformations, the type (time-, light- or spacelike) of a vector $\mathbf{v} \in \mathbb{R}^4$ is left invariant under Lorentz transformations.

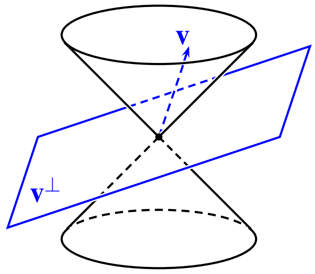
**Remark 1.3.12.** In fact,

$$\mathscr{L}^{\uparrow} \cdot \mathscr{T}^{\uparrow} = \mathscr{T}^{\uparrow} \quad \text{and} \quad \mathscr{L}^{\uparrow} \cdot \mathscr{T}^{\downarrow} = \mathscr{T}^{\downarrow}.$$

This can be seen as follows: Let $\mathbf{L} \in \mathscr{L}^{\uparrow}$. Then the special vector $\mathbf{e_0} \in \mathscr{T}^{\uparrow}$ fulfills $\mathbf{Le_0} = (\underbrace{L^0_0}_{\geq 1}, L^1_0, L^2_0, L^3_0)^{\top} \in \mathscr{T}^{\uparrow}$. To any other vector $\mathbf{v} \in \mathscr{T}^{\uparrow}$, we associate the vector $\mathbf{v}_t := (1-t)\mathbf{v} + t\mathbf{e_0}$ for $t \in [0,1]$. It is easy to see that $\mathbf{v}_t$ is always timelike and that $\mathbf{v}_0 = \mathbf{v}$ as well as $\mathbf{v}_1 = \mathbf{e_0}$. As $\mathbf{L}$ leaves the type of $\mathbf{v}$ invariant, $\mathbf{Lv}_t$ is timelike as well for any $t \in [0,1]$. Because the map $t \mapsto \mathbf{Lv}_t$ is continuous, $\mathbf{Lv}_t$ is contained in just one of the two components for all $t$ and since $\mathbf{Lv}_1 = \mathbf{Le_0} \in \mathscr{T}^{\uparrow}$, we also have $\mathbf{Lv}_0 = \mathbf{Lv} \in \mathscr{T}^{\uparrow}$.

This shows $\mathbf{L}\mathscr{Z}^{\uparrow} \subset \mathscr{Z}^{\uparrow}$. In a similar fashion one shows that $\mathbf{L}\mathscr{Z}^{\downarrow} \subset \mathscr{Z}^{\downarrow}$. Because $\mathbf{L}$ is invertible, we can use the same argument for $\mathbf{L}^{-1}$ and conclude $\mathbf{L}\mathscr{Z}^{\uparrow} = \mathscr{Z}^{\uparrow}$ and $\mathbf{L}\mathscr{Z}^{\downarrow} = \mathscr{Z}^{\downarrow}$.
The same argument shows $\mathscr{L}^{\downarrow}_{\pm} \cdot \mathscr{Z}^{\uparrow} = \mathscr{Z}^{\downarrow}$ and $\mathscr{L}^{\downarrow}_{\pm} \cdot \mathscr{Z}^{\downarrow} = \mathscr{Z}^{\uparrow}$.

**Lemma 1.3.13**
*Let $\mathbf{v} \in \mathbb{R}^4$ be timelike. Then every $\mathbf{w} \in \mathbb{R}^4$ with $\langle\langle \mathbf{v}, \mathbf{w} \rangle\rangle = 0$ is spacelike.*




*Proof.* Write $\mathbf{v} = (v^0, \hat{\mathbf{v}})$. Choose a matrix $\mathbf{A} \in \mathrm{O}(3)$ such that $\mathbf{A}\hat{\mathbf{v}} = \alpha \cdot \mathbf{e_1} = (\alpha, 0, 0)^\top$ for some $\alpha \in \mathbb{R}$. For $\mathbf{L_1} = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{A} \end{pmatrix} \in \mathscr{L}$, we have $\mathbf{L_1 v} = (v^0, \alpha, 0, 0)^\top$. Choose a boost $\mathbf{L_2} \in \mathscr{L}$ with

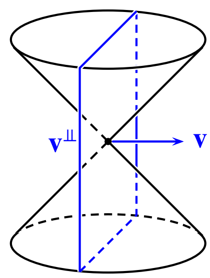
$$\mathbf{L_2} \cdot \begin{pmatrix} v^0 \\ \alpha \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \beta \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

For $\mathbf{L} = \mathbf{L_2 L_1} \in \mathscr{L}$ we have

$$0 = \langle\langle \mathbf{v}, \mathbf{w} \rangle\rangle = \langle\langle \mathbf{Lv}, \mathbf{Lw} \rangle\rangle = \left\langle\!\!\left\langle \begin{pmatrix} \beta \\ 0 \\ 0 \\ 0 \end{pmatrix}, \mathbf{Lw} \right\rangle\!\!\right\rangle = -\beta (\mathbf{Lw})^0.$$

This shows $(\mathbf{Lw})^0 = 0$, hence $\mathbf{Lw} \in \{0\} \times \mathbb{R}^3$ is spacelike. Therefore $\mathbf{w} = \mathbf{L}^{-1}\mathbf{Lw}$ is also spacelike. □

**Remark 1.3.14.** If $\mathbf{w}$ is $\langle\langle \cdot, \cdot \rangle\rangle$-perpendicular to a spacelike $\mathbf{v}$, it can be of any type.

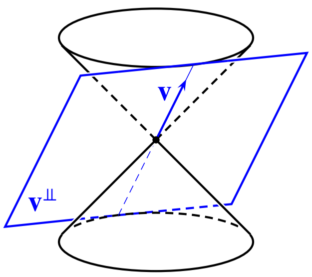



From now on we use the notation

$$\mathbf{v}^{\perp\!\!\!\perp} := \{\mathbf{w} \in \mathbb{R}^4 \mid \langle\langle \mathbf{v}, \mathbf{w} \rangle\rangle = 0\}$$

for the $\langle\langle \cdot, \cdot \rangle\rangle$-orthogonal complement of a vector $\mathbf{v} \in \mathbb{R}^4$. If $\mathbf{v} \neq 0$, then $\mathbf{v}^{\perp\!\!\!\perp}$ is a 3-dimensional vector subspace of $\mathbb{R}^4$.

Let $\mathbf{v} \in \mathbb{R}^4$ be lightlike. Then $\mathbf{v}^{\perp\!\!\!\perp}$ is the tangent space to the light cone at $\mathbf{v}$.



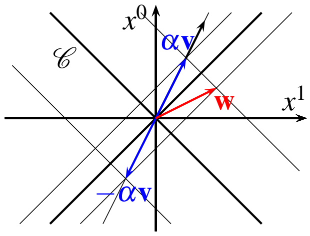

To see this, choose a differentiable curve $c : I \to \mathscr{C}$ with $c(0) = \mathbf{v}$. Then $\dot{c}(0)$ is a tangent vector to $\mathscr{C}$ and all tangent vectors are of this form. We compute

$$\begin{aligned} 0 &= \frac{d}{dt} \langle\langle c(t), c(t) \rangle\rangle |_0 \\ &= \langle\langle \dot{c}(0), c(0) \rangle\rangle + \langle\langle c(0), \dot{c}(0) \rangle\rangle \\ &= 2 \langle\langle \mathbf{v}, \dot{c}(0) \rangle\rangle . \end{aligned}$$

This shows $\dot{c}(0) \in \mathbf{v}^{\perp\!\!\!\perp}$ and we conclude that the tangent space to $\mathscr{C}$ at $p$ is contained in $\mathbf{v}^{\perp\!\!\!\perp}$. Since both, the tangent space and $\mathbf{v}^{\perp\!\!\!\perp}$, are 3-dimensional vector subspaces, they must be equal. In particular, $\mathbf{v}^{\perp\!\!\!\perp}$ contains lightlike and spacelike vectors but no timelike vectors.

**Lemma 1.3.15**
*Let $\mathbf{v} \in \mathbb{R}^4$ be timelike. Then*

$$\mathbf{v}^{\perp\!\!\!\perp} = \{\mathbf{w} \in \mathbb{R}^4 \mid \exists\, \alpha \in \mathbb{R} \setminus \{0\}, \textit{ such that } \alpha\mathbf{v} + \mathbf{w} \textit{ and } \alpha\mathbf{v} - \mathbf{w} \textit{ lightlike}\} \cup \{0 \in \mathbb{R}^4\}.$$




*Proof.* We show both inclusions.
We start with "$\supset$": Let $\mathbf{w} \in \mathbb{R}^4$ be such that there exists $\alpha \in \mathbb{R} \setminus \{0\}$ with $\alpha\mathbf{v} + \mathbf{w}$ and $\alpha\mathbf{v} - \mathbf{w}$ lightlike. Then

$$0 = \langle\langle \alpha\mathbf{v} \pm \mathbf{w}, \alpha\mathbf{v} \pm \mathbf{w} \rangle\rangle = \alpha^2 \langle\langle \mathbf{v}, \mathbf{v} \rangle\rangle \pm 2\alpha \langle\langle \mathbf{v}, \mathbf{w} \rangle\rangle + \langle\langle \mathbf{w}, \mathbf{w} \rangle\rangle$$

and hence

$$4\alpha \langle\langle \mathbf{v}, \mathbf{w} \rangle\rangle = 0.$$

Since $\alpha \neq 0$ this shows $\langle\langle \mathbf{v}, \mathbf{w} \rangle\rangle = 0$, i.e., $\mathbf{w} \in \mathbf{v}^{\perp\!\!\perp}$.
Now we show "$\subset$": Let $\mathbf{w} \in \mathbf{v}^{\perp\!\!\perp} \setminus \{0\}$. Then we have for all $\alpha \in \mathbb{R}$:

$$\langle\langle \alpha\mathbf{v} \pm \mathbf{w}, \alpha\mathbf{v} \pm \mathbf{w} \rangle\rangle = \alpha^2 \langle\langle \mathbf{v}, \mathbf{v} \rangle\rangle \pm 2\alpha \underbrace{\langle\langle \mathbf{v}, \mathbf{w} \rangle\rangle}_{=0} + \langle\langle \mathbf{w}, \mathbf{w} \rangle\rangle = \alpha^2 \langle\langle \mathbf{v}, \mathbf{v} \rangle\rangle + \langle\langle \mathbf{w}, \mathbf{w} \rangle\rangle .$$

Since $\langle\langle \mathbf{v}, \mathbf{v} \rangle\rangle < 0$ and $\langle\langle \mathbf{w}, \mathbf{w} \rangle\rangle \geq 0$ we can choose

$$\alpha = \sqrt{-\frac{\langle\langle \mathbf{w}, \mathbf{w} \rangle\rangle}{\langle\langle \mathbf{v}, \mathbf{v} \rangle\rangle}}$$

which does the job. $\square$

**Lemma 1.3.16**
*Let* $\mathbf{x}, \mathbf{y} \in \mathscr{Z}^{\uparrow}$ *with* $\langle\langle \mathbf{x}, \mathbf{x} \rangle\rangle = \langle\langle \mathbf{y}, \mathbf{y} \rangle\rangle = -1$. *Then*

$$\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle \leq -1$$

*and equality holds if and only if* $\mathbf{x} = \mathbf{y}$.

*Proof.* We choose $\mathbf{A} \in \mathrm{O}(3)$ such that $\mathbf{A}\hat{\mathbf{x}} = (\alpha, 0, 0)^\top$. Then we have for $\mathbf{L_1} := \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \mathbf{A} \end{array}\right) \in \mathscr{L}$ that

$$\mathbf{L_1 x} = \begin{pmatrix} \beta \\ \alpha \\ 0 \\ 0 \end{pmatrix}.$$

From

$$-1 = \langle\langle \mathbf{x}, \mathbf{x} \rangle\rangle = \langle\langle \mathbf{L_1 x}, \mathbf{L_1 x} \rangle\rangle = -\beta^2 + \alpha^2$$

we see that the point $(\beta, \alpha)^\top$ lies on the hyperbola as in Remark 1.3.7. Because of $\mathbf{x} \in \mathscr{X}^\uparrow$ it lies on the upper branch. Therefore there exists $\varphi \in \mathbb{R}$ such that $(\beta, \alpha)^\top = (\cosh\varphi, \sinh\varphi)^\top$. Putting $\mathbf{L_2} := \begin{pmatrix} \cosh(-\varphi) & \sinh(-\varphi) & 0 & 0 \\ \sinh(-\varphi) & \cosh(-\varphi) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathscr{L}$ and $\mathbf{L} := \mathbf{L_2 L_1} \in \mathscr{L}$ we obtain $\mathbf{Lx} = \mathbf{L_2 L_1 x} = \mathbf{e_0}$.

Next we observe

$$-1 = \langle\langle \mathbf{y}, \mathbf{y} \rangle\rangle = \langle\langle \mathbf{Ly}, \mathbf{Ly} \rangle\rangle = -((\mathbf{Ly})^0)^2 + \|\widehat{\mathbf{Ly}}\|^2 \geq -((\mathbf{Ly})^0)^2$$

with equality if and only if $\widehat{\mathbf{Ly}} = \mathbf{0}$. Hence $|(\mathbf{Ly})^0| \geq 1$ with equality if and only if $\widehat{\mathbf{Ly}} = \mathbf{0}$. Both $\mathbf{L_1}$ and $\mathbf{L_2}$ preserve time orientation, hence $\mathbf{Ly} \in \mathscr{X}^\uparrow$. In other words, $(\mathbf{Ly})^0 > 0$. Therefore we know $(\mathbf{Ly})^0 \geq 1$ with equality if and only if $\widehat{\mathbf{Ly}} = \mathbf{0}$.

Now we see

$$\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle = \langle\langle \mathbf{Lx}, \mathbf{Ly} \rangle\rangle = \langle\langle \mathbf{e_0}, \mathbf{Ly} \rangle\rangle = -(\mathbf{Ly})^0 \leq -1$$

with equality if and only if $\widehat{\mathbf{Ly}} = \mathbf{0}$. Since $\mathbf{Ly} \in \mathscr{X}^\uparrow$ with $\langle\langle \mathbf{Ly}, \mathbf{Ly} \rangle\rangle = -1$ the condition $\widehat{\mathbf{Ly}} = \mathbf{0}$ is equivalent to $\mathbf{Ly} = \mathbf{e_0} = \mathbf{Lx}$ and hence to $\mathbf{y} = \mathbf{x}$. □

Recall that cosh maps $[0, \infty)$ bijectively onto $[1, \infty)$.

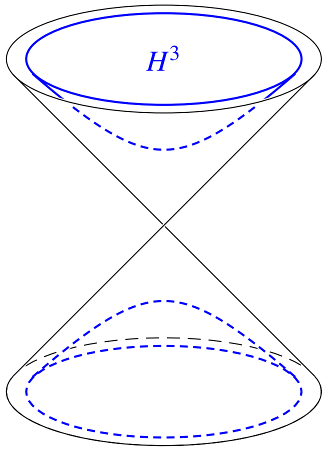
**Definition 1.3.17.** We set

$$H^3 := \{\mathbf{x} \in \mathscr{X}^\uparrow \mid \langle\langle \mathbf{x}, \mathbf{x} \rangle\rangle = -1\}.$$

The unique function $d_H : H^3 \times H^3 \to [0, \infty)$ satisfying

$$\cosh(d_H(\mathbf{x}, \mathbf{y})) = -\langle\langle \mathbf{x}, \mathbf{y} \rangle\rangle$$

is called *hyperbolic distance*. The pair $(H^3, d_H)$ is called the (3-dimensional) *hyperbolic space*.




**Remark 1.3.18.** Hyperbolic space $(H^3, d_H)$ is a metric space, i.e., for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in H^3$, we have

1. $d_H(\mathbf{x}, \mathbf{y}) \geq 0$ and $d_H(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$,
2. $d_H(\mathbf{x}, \mathbf{y}) = d_H(\mathbf{y}, \mathbf{x})$,
3. $d_H(\mathbf{x}, \mathbf{z}) \leq d_H(\mathbf{x}, \mathbf{y}) + d_H(\mathbf{y}, \mathbf{z})$.

Assertion (b) is clear and (a) is a consequence of Lemma 1.3.16. A proof of the triangle inequality can be found in [1, Satz 4.2.6].
Lorentz transformations preserving time orientation act on $H^3$ and preserve the hyperbolic distance. In other words, $\mathscr{L}^{\uparrow}(H^3) = H^3$ and $d_H(\mathbf{Lx}, \mathbf{Ly}) = d_H(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in H^3$ and $\mathbf{L} \in \mathscr{L}^{\uparrow}$.

**Remark 1.3.19.** For any $\mathbf{x} \in H^3$, the orthogonal complement $\mathbf{x}^{\perp\!\!\!\perp}$ coincides with the tangent space $T_{\mathbf{x}}H^3$ to $H^3$ at the point $\mathbf{x}$. To see this, take a smooth curve $c : (-\varepsilon, \varepsilon) \to H^3$ with $c(0) = \mathbf{x}$. Differentiating the equation

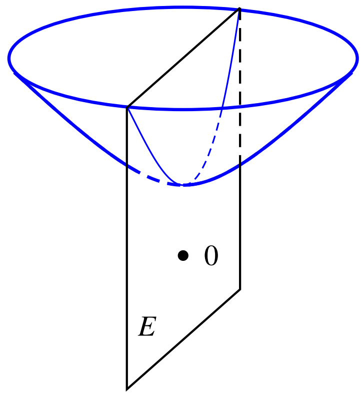
$$\langle\langle c(t), c(t) \rangle\rangle \equiv -1$$

at $t = 0$ yields

$$0 = \langle\langle \dot{c}(0), c(0) \rangle\rangle + \langle\langle c(0), \dot{c}(0) \rangle\rangle = 2\,\langle\langle \dot{c}(0), \mathbf{x} \rangle\rangle$$

and hence $\dot{c}(0) \in \mathbf{x}^{\perp\!\!\!\perp}$. Thus $T_{\mathbf{x}}H^3 \subset \mathbf{x}^{\perp\!\!\!\perp}$ and since both spaces have the same dimension three, $T_{\mathbf{x}}H^3 = \mathbf{x}^{\perp\!\!\!\perp}$.
By Lemma 1.3.13, $T_{\mathbf{x}}H^3$ contains only spacelike vectors. Hence the restriction of $\langle\langle \cdot, \cdot \rangle\rangle$ to $T_{\mathbf{x}}H^3$ is positive definite. Restricted to any tangent space of $H^3$, the Minkowski inner product becomes a Euclidean scalar product.

**Remark 1.3.20.** Two points $\mathbf{x}$ and $\mathbf{y} \in H^3$, $\mathbf{x} \neq \mathbf{y}$, determine a *great hyperbola* $G = G_{\mathbf{x},\mathbf{y}}$ as follows: Take the plane $E$ that is spanned by $\mathbf{0}$, $\mathbf{x}$ and $\mathbf{y}$. The intersection of $E$ with $H^3$ defines the great hyperbola $G = E \cap H^3$.



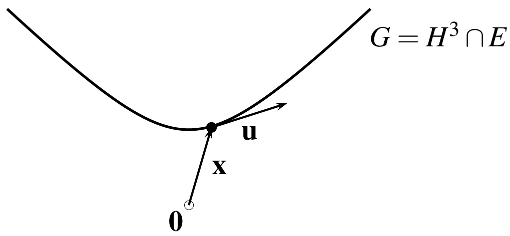

We can parametrize the great hyperbola as follows: Choose $\mathbf{u} \in E \cap T_{\mathbf{x}}H^3$ with $\langle\langle u,u \rangle\rangle = 1$. The plane $E$ contains with $\mathbf{x}$ a timelike vector, $T_{\mathbf{x}}H^3 = \mathbf{x}^{\perp\!\!\perp}$ however contains only spacelike vectors. Hence $E$ is not contained in $T_{\mathbf{x}}H^3$ and $E \cap T_{\mathbf{x}}H^3$ must be one-dimensional. This means that there are only two possibilities to choose $\mathbf{u}$; we can only replace $\mathbf{u}$ by $-\mathbf{u}$. Both choices are equally valid. The curve parametrized by

$$c(t) = \cosh(t) \cdot \mathbf{x} + \sinh(t) \cdot \mathbf{u},$$

is contained in $E$, as $c(t)$ is always a linear combination of $\mathbf{x}$ and $\mathbf{u} \in E$. The curve $c(t)$ is also contained in $H^3$, because

$$\begin{aligned}
\langle\langle c(t), c(t) \rangle\rangle &= \langle\langle \cosh(t) \cdot \mathbf{x} + \sinh(t) \cdot \mathbf{u}, \cosh(t) \cdot \mathbf{x} + \sinh(t) \cdot \mathbf{u} \rangle\rangle \\
&= \cosh(t)^2 \underbrace{\langle\langle \mathbf{x},\mathbf{x} \rangle\rangle}_{=-1} + 2\cosh(t)\sinh(t) \underbrace{\langle\langle \mathbf{x},\mathbf{u} \rangle\rangle}_{=0} + \sinh(t)^2 \underbrace{\langle\langle \mathbf{u},\mathbf{u} \rangle\rangle}_{=1} \\
&= -\cosh(t)^2 + \sinh(t)^2 \\
&= -1.
\end{aligned}$$

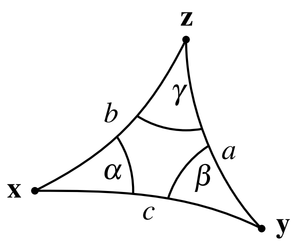
In fact, $c$ passes exactly once through the great hyperbola $G$, when $t$ traverses the real numbers. Furthermore, $c(0) = \mathbf{x}$ and $c(\pm d_H(\mathbf{x},\mathbf{y})) = \mathbf{y}$, where the sign depends on the choice of $\mathbf{u}$ (whether $\mathbf{u}$ points in direction of $\mathbf{y}$ or not).




Let $c$ and $\tilde{c}$ be two great hyperbolic arcs starting at $\mathbf{x}$, parametrized by $c(t) = \cosh(t) \cdot \mathbf{x} + \sinh(t) \cdot \mathbf{u}$ and $\tilde{c}(t) = \cosh(t) \cdot \mathbf{x} + \sinh(t) \cdot \mathbf{v}$ with $\mathbf{u},\mathbf{v} \in T_{\mathbf{x}}H^3$ and $\langle\langle \mathbf{u},\mathbf{u} \rangle\rangle = \langle\langle \mathbf{v},\mathbf{v} \rangle\rangle = 1$. The angle $\alpha \in [0, \pi)$ between the two great hyperbolas is characterized by

$$\cos(\alpha) = \langle\langle \mathbf{u}, \mathbf{v} \rangle\rangle .$$

We have the following **trigonometric identities of hyperbolic geometry**:
Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in H^3$ three different points and let $\alpha$ be the angle between the great hyperbolas running from $\mathbf{x}$ to $\mathbf{y}$ and to $\mathbf{z}$, respectively. Similarly, let $\beta$ be the angle at $\mathbf{y}$ and $\gamma$ the angle at $\mathbf{z}$. Let furthermore $a = d_H(\mathbf{y}, \mathbf{z})$, $b = d_H(\mathbf{x}, \mathbf{z})$ and $c = d_H(\mathbf{x}, \mathbf{y})$.




**Theorem 1.3.21**
*The following identities hold:*
**Law of sines:**

$$\frac{\sinh(a)}{\sin(\alpha)} = \frac{\sinh(b)}{\sin(\beta)} = \frac{\sinh(c)}{\sin(\gamma)},$$

**Law of cosines for angles:**

$$\cos(\alpha) = \cosh(a)\sin(\beta)\sin(\gamma) - \cos(\beta)\cos(\gamma),$$

**Law of cosines for sides:**

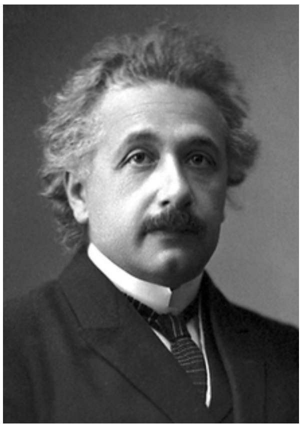
$$\cosh(a) = \cosh(b)\cosh(c) - \sinh(b)\sinh(c)\cos(\alpha).$$

The law of cosines for sides will be helpful in the investigation of the relativistic addition of velocities. It allows to determine the length of a side of a hyperbolic triangle, given the other lengths and the opposite angle.
For proofs of these laws see [1, Section 4.2].

**Exercise 1.3.22.** Use the law of cosines for sides to show

$$\alpha + \beta + \gamma < \pi.$$

## 1.4 Relativistic Kinematics

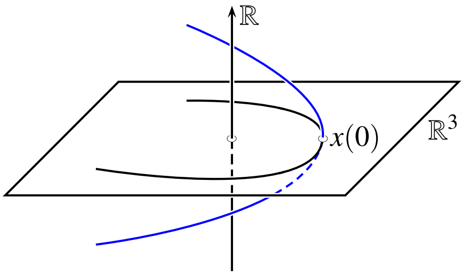


Albert Einstein (1879–1955) [5]

We now start to develop relativistic kinematics as introduced by Albert Einstein in 1905. We merge space and time to the 4-dimensional *spacetime*. The elements of spacetime are called *events*. To model particles, we use their *world lines* in spacetime

$$\{(t, \mathbf{x}(t)) \in \mathbb{R} \times \mathbb{R}^3 \,|\, t \in \mathbb{R}\}.$$

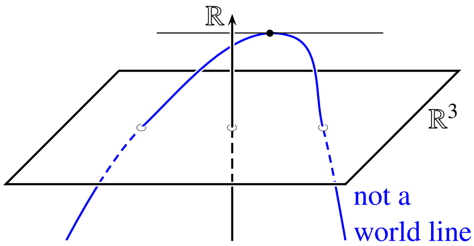
instead of their parametrizations $x : \mathbb{R} \to \mathbb{R}^3$ in space. Observe that the world line is a subset of $\mathbb{R}^4$ while $t \mapsto \mathbf{x}(t)$ is a parametrized curve. Both contain the same information, they determine each other.




We use the canonical parametrization of the world line $t \mapsto (t, \mathbf{x}(t))$ to compute its tangent:

$$\frac{d}{dt}(t, \mathbf{x}(t)) = (1, \dot{\mathbf{x}}(t)).$$

Therefore the tangents to a world line are never parallel to $\{0\} \times \mathbb{R}^3$.




Conversely, by the implicit function theorem, any smooth curve in $\mathbb{R} \times \mathbb{R}^3$ with tangents never parallel to $\{0\} \times \mathbb{R}^3$ can be parametrized in the form

$$t \mapsto (t, \mathbf{x}(t)).$$

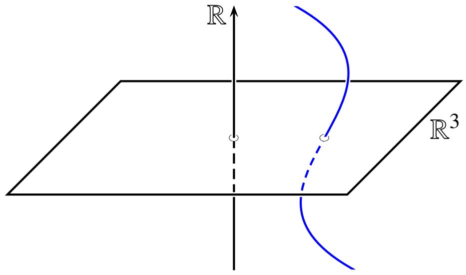
Hence it is a world line.

### The Postulate of Special Relativity

**Inertial frames.** *There exist distinguished coordinate systems for spacetime (i.e., identifications of physical spacetime with $\mathbb{R} \times \mathbb{R}^3$) called* inertial frames. *In an inertial frame the world lines of particles not subject to any forces are straight lines.*
*A coordinate system is an inertial frame if and only if it can be mapped to an inertial frame by a time orientation preserving Poincaré transformation.*

We compute the velocity of a particle $X$ in an inertial frame from the view of an observer not in motion with respect to this frame, i.e., one that has the world line $\mathbb{R} \cdot \mathbf{e_0}$.

[5] Source: `http://en.wikipedia.org/wiki/Albert_Einstein`




To this end, parametrize the world line of $X$ in the form

$$t \mapsto (ct, \hat{\mathbf{x}}(t)),\ x^0 := ct.$$

Here $c$ is the vacuum speed of light. The *physical velocity* of the particle $X$ is then given by

$$\mathbf{v}_{\text{phys}} := \frac{d\hat{\mathbf{x}}}{dt} = c\frac{d\hat{\mathbf{x}}}{dx^0}.$$

The *mathematical velocity* is

$$\mathbf{v} = \frac{d\hat{\mathbf{x}}}{dx^0} = \frac{1}{c}\mathbf{v}_{\text{phys}}.$$

For a reparametrization $\sigma \mapsto (\varphi(\sigma), \hat{\mathbf{x}}(\varphi(\sigma))) = \mathbf{x}(\varphi(\sigma))$ of the world line, we have

$$\frac{d}{d\sigma}(\mathbf{x}\circ\varphi) = \left(\varphi'(\sigma), \frac{d\hat{\mathbf{x}}}{dx^0}(\varphi(\sigma))\cdot\varphi'(\sigma)\right) = \varphi'(\sigma)\left(1, \frac{d\hat{\mathbf{x}}}{dx^0}(\varphi(\sigma))\right).$$

This implies the invariance of the mathematical velocity under reparametrizations:

$$\mathbf{v} = \frac{d\hat{\mathbf{x}}}{dx^0} = \frac{d(\hat{\mathbf{x}}\circ\varphi)}{d\sigma}\Big/\frac{d(x^0\circ\varphi)}{d\sigma}.$$

The mathematical velocity of the particle $X$ is determined by the slope of the tangent: For a tangent vector $\dot{\mathbf{x}} = (\dot{x}^0(s), \dot{\hat{\mathbf{x}}}(s))$, we have

$$\begin{aligned}\langle\langle \dot{\mathbf{x}}(s), \dot{\mathbf{x}}(s)\rangle\rangle &= \left\langle\left\langle (\dot{x}^0(s), \dot{\hat{\mathbf{x}}}), (\dot{x}^0(s), \dot{\hat{\mathbf{x}}})\right\rangle\right\rangle \\ &= -\dot{x}^0(s)^2 + \left\|\dot{\hat{\mathbf{x}}}(s)\right\|^2 \\ &= \dot{x}^0(s)^2(-1 + \|\mathbf{v}(s)\|^2).\end{aligned}$$

We observe:

$$\begin{array}{lclclcl} \dot{\mathbf{x}}(s) \text{ is timelike} & \Leftrightarrow & -1+\|\mathbf{v}(s)\|^2 < 0 & \Leftrightarrow & \|\mathbf{v}(s)\| < 1 & \Leftrightarrow & \left\|\mathbf{v}_{\text{phys}}(s)\right\| < c, \\ \dot{\mathbf{x}}(s) \text{ is lightlike} & \Leftrightarrow & -1+\|\mathbf{v}(s)\|^2 = 0 & \Leftrightarrow & \|\mathbf{v}(s)\| = 1 & \Leftrightarrow & \left\|\mathbf{v}_{\text{phys}}(s)\right\| = c, \\ \dot{\mathbf{x}}(s) \text{ is spacelike} & \Leftrightarrow & -1+\|\mathbf{v}(s)\|^2 > 0 & \Leftrightarrow & \|\mathbf{v}(s)\| > 1 & \Leftrightarrow & \left\|\mathbf{v}_{\text{phys}}(s)\right\| > c. \end{array}$$

We measured velocity of $X$ with respect to an observer B1 with world line $\mathbb{R}\cdot\mathbf{e_0}$ in a given inertial frame. Which velocity of $X$ will be measured by a second observer B2 moving with constant velocity with respect to B1?

We choose a time orientation preserving Poincaré transformation which maps the world line of B2 to $\mathbb{R}\mathbf{e_0}$. By the postulate of special relativity, this yields another inertial frame. We write $P(\mathbf{x}) = \mathbf{L}\mathbf{x} + \mathbf{p}$ with $\mathbf{L} \in \mathscr{L}^{\uparrow}$ and $\mathbf{p} \in \mathbb{R}^4$. In the new coordinates B2 has the world line $\mathbb{R}\mathbf{e_0}$ and the world line of $X$ is parametrized by $t \mapsto P(ct, \hat{\mathbf{x}}(t)) = P(\mathbf{x}(t)) =: \mathbf{y}(t)$. This means that in these coordinates, B2 observes the tangent vector $\dot{\mathbf{y}}(t) = \mathbf{L}\dot{\mathbf{x}}(t)$ to the world line of $X$ and therefore measures the mathematical velocity $\frac{\dot{\hat{\mathbf{y}}}(t)}{\dot{y}^0(t)}$. Here we have $\dot{\mathbf{y}} = \dot{y}^0\cdot\mathbf{e_0} + \dot{\hat{\mathbf{y}}}$, the splitting of $\dot{\mathbf{y}}$ into the part tangential to $e_0$ and the normal part corresponding to the factorization $\mathbb{R}^4 = \mathbb{R}\mathbf{e_0} \oplus \mathbf{e_0}^{\perp\!\!\perp} = \mathbb{R}\mathbf{e_0} \oplus T_{\mathbf{e_0}}H^3$. Reversing the transformation yields

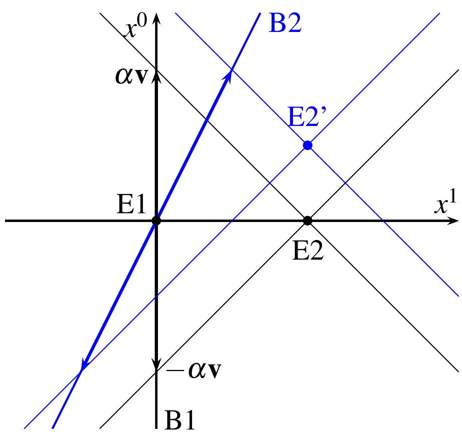
$$\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{L}^{-1}\dot{\mathbf{y}} \\
&= \mathbf{L}^{-1}(\dot{y}^0 \cdot \mathbf{e_0} + \dot{\hat{\mathbf{y}}}) \\
&= \dot{y}^0 \cdot \mathbf{L}^{-1}\mathbf{e_0} + \mathbf{L}^{-1}\dot{\hat{\mathbf{y}}} \\
&= -\langle\langle \dot{\mathbf{y}}, \mathbf{e_0} \rangle\rangle \cdot \mathbf{L}^{-1}\mathbf{e_0} + \mathbf{L}^{-1}\dot{\hat{\mathbf{y}}} \\
&= -\left\langle\left\langle \dot{\mathbf{x}}, \mathbf{L}^{-1}\mathbf{e_0} \right\rangle\right\rangle \cdot \mathbf{L}^{-1}\mathbf{e_0} + \mathbf{L}^{-1}\dot{\hat{\mathbf{y}}} \\
&= -\langle\langle \dot{\mathbf{x}}, \mathbf{z} \rangle\rangle \cdot \mathbf{z} + \mathbf{L}^{-1}\dot{\hat{\mathbf{y}}}
\end{aligned}$$

where we put $\mathbf{z} := \mathbf{L}^{-1}\mathbf{e_0} \in H^3$ for the tangent vector to the world line of B2 (with respect to the original inertial frame before transformation). We conclude that in the original coordinates, B2 observes the mathematical velocity vector

$$\frac{\mathbf{L}^{-1}\dot{\hat{\mathbf{y}}}}{-\langle\langle \dot{\mathbf{x}}, \mathbf{z} \rangle\rangle} = \frac{\dot{\mathbf{x}} + \langle\langle \dot{\mathbf{x}}, \mathbf{z} \rangle\rangle \cdot \mathbf{z}}{-\langle\langle \dot{\mathbf{x}}, \mathbf{z} \rangle\rangle} \in \mathbf{z}^{\perp\!\!\perp} = T_{\mathbf{z}}H^3. \tag{1.6}$$

## Simultaneity

How does our inertial observer B1 determine if two events happen simultaneously?
Observer B1 sends out a light signal at the event $-\alpha v$, which is reflected at the event E2 and is again received by B1 at the event $\alpha v$. Because the light took the same time for the way to E2 and the way back, the event E2 must happen at the same time as the event E1$=\mathbf{0}$.
For observer B2 having constant velocity with respect to B1, the event E1 happens at the same time as E2'. But for B1, the events E1 and E2' do *not* happen simultaneously.



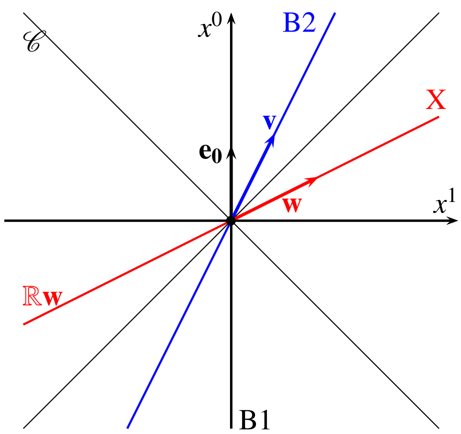

Lemma 1.3.15 can now be interpreted as follows: The set of events that are simultaneous to $\mathbf{0} \in \mathbb{R} \times \mathbb{R}^3$ for an observer with world line $\mathbb{R}\mathbf{v}$, is precisely $\mathbf{v}^{\perp\!\!\perp}$.
There is a different way to establish this statement: For observer B1, two events are simultaneous if and only if they have the same $x^0$ component, as the $x^0$ component was introduced to by the time component from the view of B1.
The hyperplanes $\{x^0\} \times \mathbb{R}^3$ in $\mathbb{R}^4$ with fixed $x^0$ are exactly those perpendicular to the world line of B1 (with respect to the Mikowski product). Using a Lorentz transformation that converts B1 to B2, this converts events which are simultaneous for B1 into events which are simultaneous for B2, by the postulate of special relativity. On the other hand, we know that our Lorentz transformation respects the Minkowski product, in particular, it maps $\mathbf{e_0}^{\perp\!\!\perp}$ to $\mathbf{v}^{\perp\!\!\perp}$.

This shows that simultaneity of events is seen differently by different inertial observers. But who is right? Since no inertial observer is distinguished from another, all are equally right. We have to abandon the idea that simultaneity of two events is a property of these events only; simultaneity is not an *absolute concept*. Simultaneity is a *relative* concept in the sense that it depends on the observer.

**Remark 1.4.1.** "Being at the same place" is already a relative concept in classical mechanics. If observer B2 is moving with constant velocity with respect to the inertial observer B1, then, from the point of view of B1, B2 occupies different locations at different times, while B2 considers itself as staying in the same place for all times.

### Superluminal Velocity

Consider the world line $\mathbb{R}\mathbf{w}$ of a hypothetical particle $X$ moving with constant speed higher than that of light with respect to an inertial observer B1. In other words, the tangent vector $\mathbf{w}$ to the world line of $X$ is spacelike. We may choose the inertial frame such that B1 has the world line $\mathbb{R}\mathbf{e_0}$ and $\mathbf{w} = \begin{pmatrix} w^0 \\ w^1 \\ 0 \\ 0 \end{pmatrix}$.



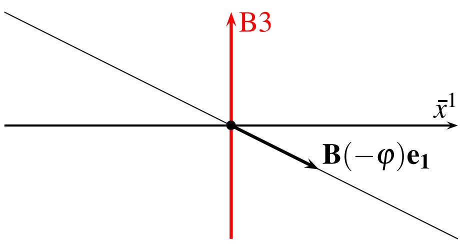

The vector $\mathbf{v} := \begin{pmatrix} w^1 \\ w^0 \\ 0 \\ 0 \end{pmatrix}$ is timelike and perpendicular to $\mathbf{w}$ with respect to the Minkowski product, $\mathbf{v} \perp\!\!\!\perp \mathbf{w}$. Let now B2 by the inertial observer with the world line $\mathbb{R}\mathbf{v}$.

This means that the whole world line of $\mathbb{R}\mathbf{w}$ is perpendicular to the world line of B2, i.e., B2 observes the particle $X$ to be at all places at the same time. Nothing like this has ever been observed.

Even worse: Denote the boost matrix $\mathbf{L_1}$ in (1.4) by $\mathbf{B}(\varphi)$. We choose $\varphi$ such that $\mathbf{v} = \mathbf{B}(\varphi)\mathbf{e_0}$. For the third observer B3 with world line $\mathbb{R} \cdot \mathbf{B}(2\varphi)\mathbf{e_0}$ we find:

After the transformation $\mathbf{B}(-2\varphi)$, the observer B3 gets the world line $\mathbb{R}\mathbf{e_0}$. The world line of X gets the tangent vector

$$\mathbf{B}(-2\varphi)\mathbf{w} = \mathbf{B}(-2\varphi)\mathbf{B}(\varphi)\mathbf{e_1} = \mathbf{B}(-\varphi)\mathbf{e_1}.$$




From the point of view of B3, the world line of X moves into the past! This leads to causality problems. If one were able to send signals to the past one could influence the past and thus change the present.

These considerations lead to the conclusion that nothing can move faster than light. Hypothetical particles moving at superluminous velocity are sometimes called *tachyons*.

### Absolute Velocity and Hyperbolic Distance

Let $X$ be a particle with world line $\mathbb{R}\mathbf{x}+\mathbf{p}$ and $B$ an inertial observer with world line $\mathbb{R}\mathbf{y}+\mathbf{q}$. Here $\mathbf{p},\mathbf{q}\in\mathbb{R}^4$ and, without loss of generality, we can choose $\mathbf{x}$ and $\mathbf{y}$ normalized such that $\mathbf{x},\mathbf{y}\in H^3$. By (1.6) $B$ observes the particle $X$ to have the mathematical velocity $\mathbf{v}=\frac{\mathbf{x}+\langle\langle\mathbf{x},\mathbf{y}\rangle\rangle\cdot\mathbf{y}}{-\langle\langle\mathbf{x},\mathbf{y}\rangle\rangle}$. For the square of the absolute velocity, we calculate

$$\begin{aligned}
\langle\langle\mathbf{v},\mathbf{v}\rangle\rangle &= \frac{\overbrace{\langle\langle\mathbf{x},\mathbf{x}\rangle\rangle}^{=-1}+2\langle\langle\mathbf{x},\mathbf{y}\rangle\rangle^2+\langle\langle\mathbf{x},\mathbf{y}\rangle\rangle^2\overbrace{\langle\langle\mathbf{y},\mathbf{y}\rangle\rangle}^{=-1}}{\langle\langle\mathbf{x},\mathbf{y}\rangle\rangle^2}\\
&= \frac{-1+\langle\langle\mathbf{x},\mathbf{y}\rangle\rangle^2}{\langle\langle\mathbf{x},\mathbf{y}\rangle\rangle^2}\\
&= \frac{-1+\cosh(d_H(\mathbf{x},\mathbf{y}))^2}{\cosh(d_H(\mathbf{x},\mathbf{y}))^2}\\
&= \frac{\sinh(d_H(\mathbf{x},\mathbf{y}))^2}{\cosh(d_H(\mathbf{x},\mathbf{y}))^2}\\
&= \tanh(d_H(\mathbf{x},\mathbf{y}))^2.
\end{aligned}$$

Therefore we get

$$\boxed{\|\mathbf{v}\|=\tanh(d_H(\mathbf{x},\mathbf{y}))}$$

for the absolute velocity.

### Addition of Velocity

As preparation we need a little lemma on hyperbolic functions.

**Lemma 1.4.2**
*For any $x\in(-1,1)$ we have*

*(a)* $\cosh(\operatorname{artanh}(x))=\frac{1}{\sqrt{1-x^2}}$;

*(b)* $\sinh(\operatorname{artanh}(x))=\frac{x}{\sqrt{1-x^2}}$;

*(c)* $e^{\operatorname{artanh}(x)}=\frac{1+x}{\sqrt{1-x^2}}=\frac{\sqrt{1+x}}{\sqrt{1-x}}$.

*Proof.* (a) Set $y := \operatorname{artanh}(x)$. Then

$$x^2 = \tanh(y)^2 = \frac{\sinh(y)^2}{\cosh(y)^2} = \frac{\cosh(y)^2 - 1}{\cosh(y)^2} = 1 - \frac{1}{\cosh(y)^2},$$

which implies $\cosh(y)^2 = \frac{1}{1-x^2}$. Because cosh is positive, we are allowed to take the positive square root, which gives the statement.

(b) From (a), we get

$$\sinh(\operatorname{artanh}(x))^2 = \cosh(\operatorname{artanh}(x))^2 - 1 = \frac{1}{1-x^2} - 1 = \frac{x^2}{1-x^2}.$$

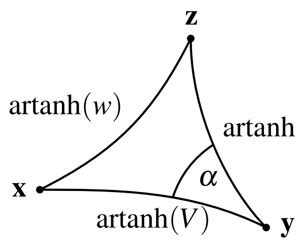
Here we have to be careful with the sign, namely we have

$$x > 0 \Leftrightarrow \operatorname{artanh}(x) > 0 \Leftrightarrow \sinh(\operatorname{artanh}(x)) > 0.$$

Taking the square root with the correct sign yields the claim.

(c) follows from $e^y = \cosh(y) + \sinh(y)$. □

Let us now consider the following situation: We have an inertial observer with world line $\mathbb{R}\mathbf{x} + \mathbf{p}$, and inertial observer B2 with world line $\mathbb{R}\mathbf{y} + \mathbf{q}$ and an object $X$ with world line $\mathbb{R}\mathbf{z} + \mathbf{r}$. Let $v = \tanh(d_H(\mathbf{y}, \mathbf{z}))$ the absolute velocity of $X$ in the view of $B2$ and $V = \tanh(d_H(\mathbf{x}, \mathbf{y}))$ the absolute velocity of B2 in the view of B1. We want to determine the absolute velocity $w = \tanh(d_H(\mathbf{x}, \mathbf{z}))$ of X in the view of B1. This is a problem of hyperbolic trigonometry.




Write $\alpha$ for the angle at the vertex $\mathbf{y}$ in this hyperbolic triangle. This is the angle between the two velocity vectors of B1 and of X in the view of B2. The law of cosines for sides of the hyperbolic geometry (Theorem 1.3.21) now states

$$\begin{aligned}\cosh(\operatorname{artanh}(w)) = &\cosh(\operatorname{artanh}(v))\cosh(\operatorname{artanh}(V)) \\ &- \sinh(\operatorname{artanh}(v))\sinh(\operatorname{artanh}(V))\cos(\alpha),\end{aligned}$$

and Lemma 1.4.2 gives

$$\begin{aligned}\frac{1}{\sqrt{1-w^2}} &= \frac{1}{\sqrt{1-v^2}}\frac{1}{\sqrt{1-V^2}} - \frac{v}{\sqrt{1-v^2}}\frac{V}{\sqrt{1-V^2}}\cos(\alpha) \\ &= \frac{1 - vV\cos(\alpha)}{\sqrt{(1-v^2)(1-V^2)}}\end{aligned}$$

Hence

$$1 - w^2 = \frac{(1-v^2)(1-V^2)}{(1-vV\cos(\alpha))^2}$$

and therefore

$$\begin{aligned} w^2 &= 1 - \frac{(1-v^2)(1-V^2)}{(1-vV\cos(\alpha))^2} \\ &= \frac{(1-vV\cos(\alpha))^2 - (1-v^2)(1-V^2)}{(1-vV\cos(\alpha))^2} \\ &= \frac{1-2vV\cos(\alpha)+v^2V^2\cos(\alpha)^2 - (1-v^2-V^2+v^2V^2)}{(1-vV\cos(\alpha))^2} \\ &= \frac{v^2+V^2-2vV\cos(\alpha)-v^2V^2\sin(\alpha)^2}{(1-vV\cos(\alpha))^2} \end{aligned}$$

This gives the general formula for relativistic addition of velocities:

$$\boxed{w = \frac{\sqrt{v^2+V^2-2vV\cos(\alpha)-v^2V^2\sin(\alpha)^2}}{1-vV\cos(\alpha)}}$$

Let us look at two special cases. For $\alpha = \pi$, we have $\cos(\alpha) = -1$ and $\sin(\alpha) = 0$. Hence we get

$$w = \frac{\sqrt{v^2+V^2+2vV}}{1+vV} = \frac{v+V}{1+vV}.$$

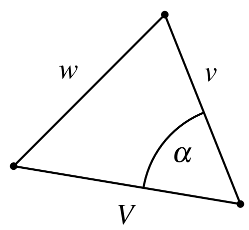
The is a deviation from the classical result $w = v+V$ by the factor $\frac{1}{1+vV}$. For velocities that are small compared to the speed of light, $vV$ is very small and the difference is barely measurable. Now look at the case that the velocities are perpendicular to each other. For $\alpha = \pi/2$, we have $\cos(\alpha) = 0$ and $\sin(\alpha) = 1$. Therefore we get

$$w = \sqrt{v^2+V^2-v^2V^2}.$$

In classical mechanics, the Pythagorean theorem would have given the result $w = \sqrt{v^2+V^2}$. For general $\alpha$ the law of cosines for the Euclidean geometry yields

$$w = \sqrt{v^2+V^2-2vV\cos(\alpha)}$$

for classical mechanics.



It is also interesting to consider that case $v = 1$, i.e., $X$ moves with the speed of light relative to $B2$. Relativistic velocity addition gives us

$$w = \frac{\sqrt{1+V^2-2V\cos(\alpha)-V^2\sin(\alpha)^2}}{1-V\cos(\alpha)} = \frac{\sqrt{1+V^2\cos(\alpha)^2-2V\cos(\alpha)}}{1-V\cos(\alpha)} = 1.$$

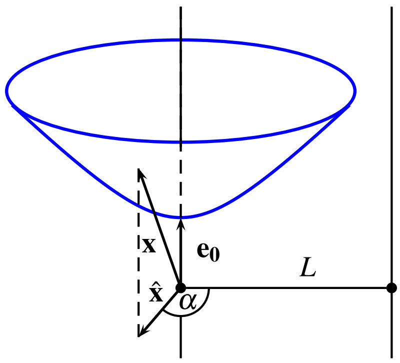
Thus $X$ also moves with the same speed of light relative of $B1$, independently of the relative motion of $B1$ and $B2$.

## Length Contraction

Regard a bar not subject to any acceleration. Choose the coordinate system such that one end of the bar has the world line $\mathbb{R}\mathbf{e_0}$ and the other end has the world line $\mathbb{R}\mathbf{e_0} + L\mathbf{e_1}$. An inertial observer B1 sitting at the first end of the bar (i.e. having world the world line $\mathbb{R}\mathbf{e_0}$) measures $L$ for the length of the bar. To see this, note that in the view of B1, the events $\mathbf{0}$ and $(0,L,0,0)^\top$ are simultaneous events on the world lines of the two ends and there distance in space is

$$\sqrt{\langle\langle \mathbf{0}-(0,L,0,0)^\top, \mathbf{0}-(0,L,0,0)^\top\rangle\rangle} = \sqrt{\langle\langle (0,L,0,0)^\top, (0,L,0,0)^\top\rangle\rangle} = L.$$

Let now B2 be a second inertial observer with world line $\mathbb{R}\mathbf{x}$. Which length $\tilde{L}$ will by measured by B2?




To calculate this, we have to determine the event on the world line $\mathbb{R}\mathbf{e_0} + L\mathbf{e_1}$ that is simultaneous to $\mathbf{0}$. Since in the view of B2, the events simultaneous to $\mathbf{0}$ are exactly the points on $\mathbf{x}^{\perp\!\!\!\perp}$. We solve

$$0 = \left\langle\!\!\left\langle (t,L,0,0)^\top, \mathbf{x} \right\rangle\!\!\right\rangle = -tx^0 + Lx^1$$

for $t$ and we obtain

$$t = L\frac{x^1}{x^0} = L\frac{\langle \hat{\mathbf{x}}, \mathbf{e_1}\rangle}{x^0} = L\frac{\cos(\alpha)\cdot\|\hat{\mathbf{x}}\|}{x^0} = L\cdot\cos(\alpha)\cdot V,$$

where $V$ is the absolute velocity between B1 and B2 and $\alpha$ is the angle between $\mathbf{e_1}$ and the velocity vector. Hence in the view of B2, the events $\mathbf{0}$ and $(L\cdot\cos(\alpha)\cdot V, L, 0, 0)^\top$ are simultaneous events lying on the world lines of the ends of the bar. B2 measures the distance in space

$$\begin{aligned}\tilde{L}^2 &= \left\langle\!\!\left\langle \mathbf{0}-(L\cdot\cos(\alpha)\cdot V,L,0,0)^\top, \mathbf{0}-(L\cdot\cos(\alpha)\cdot V,L,0,0)^\top \right\rangle\!\!\right\rangle \\ &= \left\langle\!\!\left\langle (L\cdot\cos(\alpha)\cdot V,L,0,0)^\top, (L\cdot\cos(\alpha)\cdot V,L,0,0)^\top \right\rangle\!\!\right\rangle \\ &= -L^2\cdot\cos(\alpha)^2\cdot V^2+L^2 \\ &= L^2\cdot(1-\cos(\alpha)^2V^2)\end{aligned}$$

and therefore

$$\boxed{\tilde{L} = L\cdot\sqrt{1-\cos(\alpha)^2V^2}}$$

From the point of view of the inertial observer B2 moving towards the bar, the length of the bar is shortened by the factor $\sqrt{1-\cos(\alpha)^2V^2}$. For motion in direction of the bar ($\alpha = 0$ or $\alpha = \pi$), we have the strongest length contraction, namely by the factor $\sqrt{1-V^2}$. For motion perpendicular to the bar ($\alpha = \pm\pi/2$) there is no contraction.
The length of an object measured by an observer in rest relatively to the object is called *proper length* of the object. It is the maximal length of the object that an observer can measure.
Hence the length of an object has also become a relative concept in the sense that it depends on the observer. This leads to a number of puzzling questions. Here is an example:
**The Tunnel Paradox.** A train with proper length $L$ is travelling through a tunnel that also has proper length $L$. Is the train contained completely in the tunnel at some point?
*From an outside view:*
Because of length contraction, the train is shorter than the tunnel. Therefore, for some time, the train is completely contained in the tunnel.
*From the traindriver's view:*
Because of length contraction, the tunnel is shorter than the train. Therefore the train is never completely contained in the tunnel.
Who is right?
"Being completely contained in the tunnel" means that both ends of the train are in the tunnel *simultaneously*. Simultaneity, however, is a relative concept (depending on the inertial observer) and hence this is also the case for the concept of "being completely contained in the tunnel". Both observers are right from there respective points of view.

## Time Dilation

An inertial observer B1 with world line $\mathbb{R}\cdot\mathbf{e_0}$ observes the elapsed time $T$ between the events $\mathbf{0}$ and $T\cdot\mathbf{e_0}$. More generally, if B1 has the world line $\mathbb{R}\cdot\mathbf{x}$ with $\mathbf{x}\in H^3$, then B1 observes the elapsed time $T$ between the events $\mathbf{0}$ and $T\cdot\mathbf{x}$. Let now B2 be another inertial observer with world line $\mathbb{R}\cdot\mathbf{y}$ where $\mathbf{y}\in H^3$. Which is the time $\tilde{T}$ elapsed between the events $\mathbf{0}$ und $T\cdot\mathbf{x}$ from the viewpoint of B2?
To answer this question, we have to find the time $\tilde{T}$ for which the event $\tilde{T}\cdot\mathbf{y}$ is simultaneous to the event $T\cdot\mathbf{x}$ in the view of B2. This is the case when the difference vector $T\cdot\mathbf{x}-\tilde{T}\cdot\mathbf{y}$ is perpendicular to the world line of B2, or equivalently, to $\mathbf{y}$:

$$0 = \langle\!\langle \mathbf{y}, T\cdot\mathbf{x}-\tilde{T}\cdot\mathbf{y}\rangle\!\rangle = T\cdot\langle\!\langle\mathbf{y},\mathbf{x}\rangle\!\rangle - \tilde{T}\cdot\langle\!\langle\mathbf{y},\mathbf{y}\rangle\!\rangle = -T\cdot\cosh(d_H(\mathbf{y},\mathbf{x}))+\tilde{T},$$

hence, by Lemma 1.4.2,

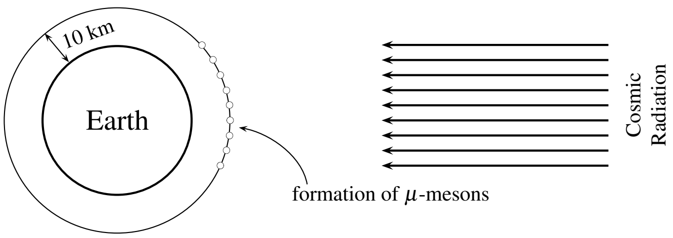
$$\tilde{T} = T \cdot \cosh(d_H(\mathbf{y},\mathbf{x})) = T \cdot \cosh(\operatorname{artanh}(V)) = T \cdot \frac{1}{\sqrt{1-V^2}},$$

where $V$ is the velocity between B1 and B2. Because of the correction factor $\frac{1}{\sqrt{1-V^2}} > 1$, the time elapsed is longer in the view of B2. This phenomenon is known as *time dilation*. In the view of observer B2, the clock of B1 runs slower than his own. Exchanging roles of B1 and B2, we analogously obtain that the clock of B2 runs slower than his own in the view of B1.
In physical units, we have

$$\boxed{\tilde{T} = \frac{T}{\sqrt{1-\frac{{V_{phys}}^2}{c^2}}}}$$

For velocities much below the speed of light, $V_{phys} \ll c$, i.e. $V \ll 1$, the correction factor is very close to 1. For this reason, time dilation is not noticed in daily life.

**Example 1.4.3.** *Cosmic radiation and $\mu$-mesons*

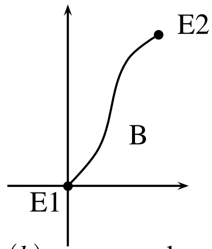



Cosmic radiation generates certain elementary particles, so-called $\mu$-mesons, on impact with the outer atmosphere. These $\mu$-mesons have a mean lifetime of $2 \cdot 10^{-6}$ s. Even with light speed, the $\mu$-mesons can cover a distance of only

$$3 \cdot 10^5 \frac{km}{s} \cdot 2 \cdot 10^{-6} s = 6 \cdot 10^{-1} km$$

on average. Therefore one would expect that only very few $\mu$-mesons reach the surface of the earth because the distance between the outer atmosphere and the surface of the earth is roughly 10 km. It is a fact however, that $\mu$-mesons can be detected on the earth's surface in great numbers. What is the explanation for this?
*Explanation from our point of view on earth:* Time dilation implies that time goes by much slower for the $\mu$-mesons moving with very high speed towards the earth. For this reason, from our point of view, the lifetime of $\mu$-mesons is much longer than $2 \cdot 10^{-6}$ s.
*Explanation from the $\mu$-meson's point of view:* Because of length contraction the distance between the outer atmosphere and the surface of the earth is much less than 10 km. Therefore the distance to the surface can be overcome even in the short time at disposal.
This example shows nicely that length contraction and time dilation are really two sides of the same medal.

We now consider an observer B that is subject to acceleration. We assume that B always has velocity below light speed with respect to inertial observers, i.e. its world line is timelike. Parametrize the world line of B by $\mathbf{x} : [a,b] \to \mathbb{R}^4$. After possibly using the parameter transform $s \mapsto -s$, we can assume that $\mathbf{x}$ is future directed, i.e. that $\frac{d\mathbf{x}}{ds}(s) \in \mathscr{Z}^\uparrow$ for all $s \in [a,b]$.




What is the time elapsed between two events E1$=\mathbf{x}(a)$ and E2$=\mathbf{x}(b)$, measured on a clock taken along by observer B? In the special case that B moves with constant velocity (with respect to inertial observers), we already know that the time elapsed between E1 and E2 is given by

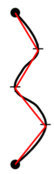
$$\sqrt{-\langle\langle E2-E1, E2-E1\rangle\rangle}.$$

We reduce the general case to this one. For a sufficiently fine partition $a = s_0 < s_1 < \ldots < s_n = b$ we have $\mathbf{x}(s_i) - \mathbf{x}(s_{i-i}) \in \mathscr{Z}^\uparrow$, $i = 1,\ldots,n$, because

$$\frac{\mathbf{x}(s_i) - \mathbf{x}(s_{i-i})}{s_i - s_{i-i}} \to \underbrace{\frac{d\mathbf{x}}{ds}}_{\in \mathscr{Z}^\uparrow},$$

as the mesh of the partition tends to 0. Since $\mathscr{Z}^\uparrow$ is open, $\frac{\mathbf{x}(s_i)-\mathbf{x}(s_{i-i})}{s_i - s_{i-i}}$ has to be in $\mathscr{Z}^\uparrow$ if $s_i - s_{i-1}$ is small enough[6].

This partition leads to the approximation of an accelerated observer B by a "piecewiese inertial observer". This approximation becomes better as the mesh of the partition gets smaller. The time elapsed between to subsequent events $\mathbf{x}(s_{i-1})$ and $\mathbf{x}(s_i)$ in the view of the corresponding inertial observer with world line $\mathbb{R}\cdot(\mathbf{x}(s_i) - \mathbf{x}(s_{i-1})) + \mathbf{x}(s_{i-1})$ is



$$\sqrt{-\langle\langle \mathbf{x}(s_i) - \mathbf{x}(s_{i-1}), \mathbf{x}(s_i) - \mathbf{x}(s_{i-1})\rangle\rangle}.$$

Summation gives an approximate value for the time elapsed between E1$=\mathbf{x}(a)$ and E2$=\mathbf{x}(b)$ from the viewpoint of B:

$$\begin{aligned}
&\sum_{i=1}^{n} \sqrt{-\langle\langle \mathbf{x}(s_i) - \mathbf{x}(s_{i-1}), \mathbf{x}(s_i) - \mathbf{x}(s_{i-1})\rangle\rangle} \\
&\quad = \sum_{i=1}^{n} \sqrt{-\left\langle\!\!\left\langle \frac{\mathbf{x}(s_i) - \mathbf{x}(s_{i-1})}{s_i - s_{i-1}}, \frac{\mathbf{x}(s_i) - \mathbf{x}(s_{i-1})}{s_i - s_{i-1}} \right\rangle\!\!\right\rangle} \cdot (s_i - s_{i-1}) \\
&\quad \to \int_a^b \sqrt{-\left\langle\!\!\left\langle \frac{d\mathbf{x}}{ds}(s), \frac{d\mathbf{x}}{ds}(s) \right\rangle\!\!\right\rangle}\, ds,
\end{aligned}$$

[6] In fact, it is not hard to see that for every future-directed timelike curve $\mathbf{x}$, we always have $\mathbf{x}(t_2) - \mathbf{x}(t_1) \in \mathscr{Z}^\uparrow$ even if $t_2$ is much larger than $t_1$. However, we will not need this.

as the mesh tends to 0 (theorem on Riemann sums). We summarize: From the view of an accelerated observer B with word line $\mathbf{x}$, the time elapsed between the events $\mathbf{x}(a)$ and $\mathbf{x}(b)$ is given by

$$\boxed{\int_a^b \sqrt{-\left\langle\!\!\left\langle \frac{d\mathbf{x}}{ds}(s), \frac{d\mathbf{x}}{ds}(s) \right\rangle\!\!\right\rangle}\, ds}$$

**Definition 1.4.4.** A future-directed parametrization $\mathbf{x} : [a,b] \to \mathbb{R}^4$ of a timelike world line is called a *parametrization by proper time* if

$$\left\langle\!\!\left\langle \frac{d\mathbf{x}}{d\tau}, \frac{d\mathbf{x}}{d\tau} \right\rangle\!\!\right\rangle \equiv -1.$$

In other words, we have for all $\tau$:

$$\frac{d\mathbf{x}}{d\tau}(\tau) \in H^3.$$

**Remark 1.4.5.** If the world line of an observer is parametrized by proper time, the parameter $\tau$ always gives the time elapsed from the view of this observer:

$$\int_a^{\tau_0} \sqrt{-\left\langle\!\!\left\langle \frac{d\mathbf{x}}{d\tau}, \frac{d\mathbf{x}}{d\tau} \right\rangle\!\!\right\rangle}\, d\tau = \tau_0 - a.$$

**Lemma 1.4.6**
*Every timelike world line can be parametrized by proper time. The parametrization by proper time is unique up to parameter transformations of the form $\tau \mapsto \tau + \tau_0$ for fixed $\tau_0 \in \mathbb{R}$.*

*Proof. Existence:* Let $s \mapsto \mathbf{x}(s)$ a parametrization of the world line. Without loss of generality let $\dfrac{dx^0}{ds} > 0$ (otherwise replace $s$ by $-s$). For fixed $t_0 \in \mathbb{R}$ set

$$\psi(s) := \int_{t_0}^{s} \sqrt{-\left\langle\!\!\left\langle \frac{d\mathbf{x}}{ds}(t), \frac{d\mathbf{x}}{ds}(t) \right\rangle\!\!\right\rangle}\, dt, \quad \text{so that} \quad \psi'(s) = \sqrt{-\left\langle\!\!\left\langle \frac{d\mathbf{x}}{ds}(s), \frac{d\mathbf{x}}{ds}(s) \right\rangle\!\!\right\rangle} > 0.$$

Hence $\psi$ is strictly increasing and for the inverse $\varphi := \psi^{-1}$ we have

$$\frac{d\varphi}{d\tau}(\tau) = \frac{1}{\psi'(\varphi(\tau))} = \frac{1}{\sqrt{-\left\langle\!\!\left\langle \frac{d\mathbf{x}}{ds}(\varphi(\tau)), \frac{d\mathbf{x}}{ds}(\varphi(\tau)) \right\rangle\!\!\right\rangle}}.$$

This implies

$$\begin{aligned}
\left\langle\!\!\left\langle \frac{d(\mathbf{x}\circ\varphi)}{d\tau}(\tau), \frac{d(\mathbf{x}\circ\varphi)}{d\tau}(\tau) \right\rangle\!\!\right\rangle &= \left\langle\!\!\left\langle \frac{d\mathbf{x}}{ds}(\varphi(\tau)) \cdot \frac{d\varphi}{d\tau}(\tau), \frac{d\mathbf{x}}{ds}(\varphi(\tau)) \cdot \frac{d\varphi}{d\tau}(\tau) \right\rangle\!\!\right\rangle \\
&= \left(\frac{d\varphi}{d\tau}(\tau)\right)^2 \cdot \left\langle\!\!\left\langle \frac{d\mathbf{x}}{ds}(\varphi(\tau)), \frac{d\mathbf{x}}{ds}(\varphi(\tau)) \right\rangle\!\!\right\rangle \\
&= -1.
\end{aligned}$$

*Uniqueness:* Let $\mathbf{x}$ and $\mathbf{x}\circ\varphi$ be parametrizations by proper time. Then

$$-1 = \left\langle\!\!\left\langle \frac{d(\mathbf{x}\circ\varphi)}{d\tau}, \frac{d(\mathbf{x}\circ\varphi)}{d\tau} \right\rangle\!\!\right\rangle = \left(\frac{d\varphi}{d\tau}\right)^2 \cdot \underbrace{\left\langle\!\!\left\langle \frac{d\mathbf{x}}{ds}, \frac{d\mathbf{x}}{ds} \right\rangle\!\!\right\rangle}_{=-1}.$$

This implies $\left|\frac{d\varphi}{d\tau}\right| = 1$ and hence $\varphi(\tau) = \pm\tau + \tau_0$ for some fixed $\tau_0 \in \mathbb{R}$. Since both parametrizations are future directed, we have

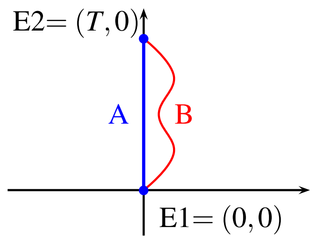
$$0 < \frac{d(\mathbf{x}\circ\varphi)^0}{d\tau} = \frac{d\varphi}{d\tau} \cdot \underbrace{\frac{dx^0}{ds}}_{>0} \quad \text{and hence} \quad \frac{d\varphi}{d\tau} > 0.$$

Thus $\varphi(\tau) = \tau + \tau_0$. □

**The Twin Paradox**

Suppose A and B are twins. Twin B decides to go on a round trip in a space craft while twin A remains at home in an inertial frame. On his return, B is younger than A! This can be seen as follows: In the inertial frame of A, let E1= $\mathbf{0}$ the event of B's departure and E2= $(T,0,0,0)^\top$ the event of his or her return. This means that A has aged by time $T$ during the separation of the twins.

To compute the aging of B let $s \mapsto \mathbf{x}(s) = (s, \hat{\mathbf{x}}(s))$ be a parametrization of the worldline of B in the inertial system of A. We compute the time that has passed for B:




$$\int_0^T \sqrt{-\left\langle\!\left\langle \frac{d\mathbf{x}}{ds}, \frac{d\mathbf{x}}{ds} \right\rangle\!\right\rangle}\, ds = \int_0^T \underbrace{\sqrt{1 - \left\|\frac{d\hat{\mathbf{x}}}{ds}\right\|^2}}_{\substack{\leq 1 \text{ and } = 1 \text{ only,} \\ \text{if } \left\|\frac{d\hat{\mathbf{x}}}{ds}\right\| = 0}} ds < T.$$

We conclude: traveling keeps you young!
In fact, this was verified experimentally. In the *Hafele-Keating experiment (1971)*, one compared two atomic clocks, one on board of a Boeing 747, the other one remaining on earth. Here one had to take into account the rotation of the earth. We will see later, in the part about general relativity, that another effect also plays a role, namely the influence of gravitation. Gravitation is weaker on board of the airplane while in high altitude.

**Exercise 1.4.7.** Redo the computations of Exercise 1.1.2 using special relativity instead of Newtonian mechanics. Calculate both earth times and proper times as well as the maximal velocities (as seen from the earth).

**Definition 1.4.8.** Let $\mathbf{x} : I \to \mathbb{R}^4$ by a parametrization by proper time of the world line of a timelike particle. The vector

$$\mathbf{u} := \frac{d\mathbf{x}}{d\tau}$$

is called *four-velocity* of the particle (at $\mathbf{x}(\tau)$) and

$$\mathbf{a} := \frac{d^2\mathbf{x}}{d\tau^2}$$

is called its *four-acceleration*.

**Remark 1.4.9.** By definition of a proper-time parametrization, the four-velocity is a curve in $H^3$. Hence its derivative, the four-acceleration, is always tangent to $H^3$,

$$\mathbf{a}(\tau) = \frac{d\mathbf{u}}{d\tau}(\tau) \in T_{\mathbf{u}(\tau)} H^3 = \mathbf{u}(\tau)^{\perp\!\!\perp}.$$

In particular, by Lemma 1.3.13, four-acceleration is always spacelike.

Write $\mathbf{x} = (x^0, \hat{\mathbf{x}})$ for the world line of a particle. The *observed velocity* of $\mathbf{x}$ from the view of an inertial observer with world line $\mathbb{R}\mathbf{e_0}$ is given by

$$\frac{\hat{\mathbf{u}}}{u^0},$$

as discussed before. The *observed acceleration* from the viewpoint of this inertial observer is the change of velocity per change of time, which is

$$\frac{d}{dx^0}\left(\frac{\hat{\mathbf{u}}}{u^0}\right) = \frac{1}{u^0}\frac{d}{d\tau}\left(\frac{\hat{\mathbf{u}}}{u^0}\right) = \frac{1}{u^0}\frac{\frac{d}{d\tau}\hat{\mathbf{u}}\cdot u^0 - \frac{d}{d\tau}u^0\cdot\hat{\mathbf{u}}}{(u^0)^2} = \frac{\hat{\mathbf{a}}}{(u^0)^2} - \frac{a^0}{(u^0)^3}\hat{\mathbf{u}}.$$

If the inertial frame is the rest frame of the particle at time $\tau = \tau_0$, i.e., $\mathbf{u}(\tau_0) = \mathbf{e_0}$, then the four-acceleration satisfies $\mathbf{a}(\tau_0) = (0, \hat{\mathbf{a}}(\tau_0))$, because $\mathbf{a}(\tau_0) \perp\!\!\!\perp \mathbf{u}(\tau_0) = \mathbf{e_0}$ and hence $a^0(\tau_0) = 0$. Since $u^0(\tau_0) = 1$ and $a^0(\tau_0) = 0$, the observed acceleration is just $\hat{\mathbf{a}}(\tau_0)$. The absolute value of the observed acceleration in the rest frame is therefore

$$\|\hat{\mathbf{a}}(\tau_0)\| = \sqrt{\langle\langle \mathbf{a}(\tau_0), \mathbf{a}(\tau_0)\rangle\rangle}.$$

*Solution to Exercise 1.4.7.* We perform the computations in the inertial frame of the earth. The earth's world line is then $\mathbb{R}\mathbf{e_0}$ and we choose the coordinates such that the destination X has the world line $\mathbb{R}\mathbf{e_0} + (0, D, 0, 0)$. Let $\tau \mapsto \mathbf{x}(\tau)$ the world line of the space ship, parametrized by proper time. The four-velocity is then

$$\mathbf{u}(\tau) = (\cosh(\varphi(\tau)), \sinh(\varphi(\tau)), 0, 0)$$

for a function $\varphi(\tau)$ yet to be determined. For the four-acceleration, we have

$$\mathbf{a}(\tau) = \varphi'(\tau)(\sinh(\varphi(\tau)), \cosh(\varphi(\tau)), 0, 0).$$

The absolute value of the four-acceleration is

$$g^2 = \langle\langle \mathbf{a}(\tau_0), \mathbf{a}(\tau_0)\rangle\rangle = (\varphi'(t))^2 \cdot 1,$$

therefore $\varphi'(\tau) = \pm g$ and hence $\varphi(\tau) = \pm g\tau + \varphi_0$. From $\mathbf{e_0} = \mathbf{u}(0) = (\cosh(\varphi_0), \sinh(\varphi_0), 0, 0)$ we conclude $\varphi_0 = 0$ and therefore $\varphi(\tau) = \pm g\tau$.
During the first half of the travel, the spacecraft accelerates in direction X. Thus $\varphi'(\tau) > 0$, hence $\varphi(\tau) = g\tau$. We conclude $\mathbf{u}(\tau) = (\cosh(g\tau), \sinh(g\tau), 0, 0)$ and therefore

$$\mathbf{x}(\tau) = \frac{1}{g}(\sinh(g\tau), \cosh(g\tau), 0, 0) + \mathbf{x_0}.$$

From

$$(0,0,0,0) = \mathbf{x}(0) = \frac{1}{g}(0,1,0,0) + \mathbf{x_0}$$

we get $\mathbf{x_0} = -\frac{1}{g}\mathbf{e_1}$. Summarizing we have the proper-time parametrization of the world line of the spacecraft:

$$\mathbf{x}(\tau) = \frac{1}{g}(\sinh(g\tau), \cosh(g\tau) - 1, 0, 0).$$

For the time of travel $T_{\text{ship}}$ from the viewpoint of the crew we obtain

$$\frac{D}{2} = x^1\left(\frac{T_{\text{ship}}}{2}\right) = \frac{1}{g}\left(\cosh\left(g\frac{T_{\text{ship}}}{2}\right) - 1\right)$$

and hence

$$T_{\text{ship}} = \frac{2}{g}\operatorname{arcosh}\left(\frac{Dg}{2} + 1\right).$$

For the time of travel $T_{\text{earth}}$ from the viewpoint of the earth we get

$$\begin{aligned}
\frac{T_{\text{earth}}}{2} &= x^0\left(\frac{T_{\text{ship}}}{2}\right) \\
&= x^0\left(\frac{1}{g}\operatorname{arcosh}\left(\frac{Dg}{2} + 1\right)\right) \\
&= \frac{1}{g}\sinh\left(\operatorname{arcosh}\left(\frac{Dg}{2} + 1\right)\right) \\
&= \frac{1}{g}\sqrt{\frac{g^2D^2}{4} + gD} \\
&= \sqrt{\frac{D^2}{4} + \frac{D}{g}}
\end{aligned}$$

hence

$$T_{\text{earth}} = \sqrt{D^2 + \frac{4D}{g}}.$$

Furthermore, the maximal velocity $v_{\max}$ can be calculated by

$$\begin{aligned}
v_{\max} &= \frac{u^1(T_{\text{ship}}/2)}{u^0(T_{\text{ship}}/2)} \\
&= \tanh\left(\operatorname{arcosh}\left(\frac{gD}{2} + 1\right)\right) \\
&= \frac{\sqrt{\frac{g^2D^2}{4} + gD}}{\frac{gD}{2} + 1} \\
&= \frac{\sqrt{g^2D^2 + 4gD}}{gD + 2}.
\end{aligned}$$

Here, we always calculated in terms of the dimensionless mathematical velocity $v$, which is related to the physical velocity by the speed of light,

$$v = \frac{v_{\text{phys}}}{c}.$$

Mathematical length and time therefore have to have the same dimension. We choose the convention to calculate in units of length, i.e.

$$D = D_{\text{phys}}, \quad T = c \cdot T_{\text{phys}}, \quad g = \frac{g_{\text{phys}}}{c^2}.$$

Then we get

$$
\begin{aligned}
T_{\text{ship, phys}} &= \frac{2c}{g_{\text{phys}}} \operatorname{arcosh}\left(\frac{D g_{\text{phys}}}{2c^2} + 1\right) \\
T_{\text{earth, phys}} &= \sqrt{\frac{D^2}{c^2} + \frac{4D}{g_{\text{phys}}}} \\
v_{\text{max, phys}} &= \frac{c\sqrt{D^2 {g_{\text{phys}}}^2 + 4D {g_{\text{phys}}}^2 c^2}}{D g_{\text{phys}} + 2c^2}
\end{aligned}
$$

Inserting the values for our destinations we get the following table:

| | | *classical* | | *relativistic* | | |
|---|---|---|---|---|---|---|
| object X | distance $D$ | time $T$ | $v_{\max}$ | $T_{\text{earth}}$ | $T_{\text{ship}}$ | $v_{\max}$ |
| moon | 400.000 km | 3,5 h | 63 km/s | 3,5 h | 3,5 h | 63 km/s |
| Mars (near) | 56 mill. km | 42 h | 742 km/s | 42 h | 42 h | 741 km/s |
| Mars (far) | 400 mill. km | 112 h | 1980 km/s | 112 h | 112 h | 1980 km/s |
| Proxima Centauri | 4,3 light years | 4 years | 2,1 $c$ | 5,9 years | 3,6 years | 0,95 $c$ |
| Andromeda galaxy | 2 mill. light years | 2784 years | 1434 $c$ | 2 mill. light years | 28 years | almost $c$ |

For a convenient way to compute such travel values, see the applet at `http://mobius.maplesoft.com/maplenet/mobius/application.jsp?appId=11532144`. Java support must be activated in the browser for the applet to work.

## 1.5 Mass and Energy

**Definition 1.5.1.** A *force field* $\mathbf{F}$ is a smooth mapping $\mathbf{F} : \mathbb{R}^4 \times H^3 \to \mathbb{R}^4$ such that for all $\mathbf{x} \in \mathbb{R}^4$ and $\mathbf{u} \in H^3$ we have

$$\langle\langle \mathbf{F}(\mathbf{x}, \mathbf{u}), \mathbf{u} \rangle\rangle = 0.$$

This means $\mathbf{F}(\mathbf{x}, \mathbf{u}) \in T_{\mathbf{u}} H^3$. We impose this condition because we know already that it is fulfilled for the four-acceleration and we want to demand later that the force is proportional to acceleration, as in Newton's second law.
Let $m_0 > 0$ be a constant which we interpret as the *rest mass* of a particle. Let the particle have world line $\mathbf{x}$ in an inertial system. The analog to the Newton's second law is then:

*If the world line of a particle is parametrized by proper time and the particle is subject to the force* $\mathbf{F}$*, then*

$$\frac{d}{d\tau}(m_0\mathbf{u}(\tau)) = \mathbf{F}(\mathbf{x}(\tau),\mathbf{u}(\tau)), \tag{1.7}$$

or, in slightly different words,

$$m_0\frac{d^2}{d\tau^2}\mathbf{x}(\tau) = \mathbf{F}\big(\mathbf{x}(\tau),\frac{d\mathbf{x}}{d\tau}(\tau)\big).$$

This is an ordinary differential equation of second order. Given any initial conditions $\mathbf{x}(\tau_0)$ and $\mathbf{u}(\tau_0) = \frac{d\mathbf{x}}{d\tau}(\tau)$ it has a unique solution. Hence special relativity is, as the theory of classical mechanics, a deterministic theory.
In the rest frame of the particle, i.e. if $\mathbf{u}(\tau_0) = \mathbf{e_0}$, the relation $\mathbf{F}(\mathbf{x}(\tau_0),\mathbf{u}(\tau_0)) \perp\!\!\!\perp \mathbf{u}(\tau_\mathbf{0})$ means

$$\mathbf{F}(\mathbf{x}(\tau_0),\mathbf{u}(\tau_0)) = (0,\hat{\mathbf{F}}(\mathbf{x}(\tau_0),\mathbf{u}(\tau_0))).$$

Hence in the rest frame, we are left with the classical Newtonian equation of motion $m_0\,\hat{\mathbf{a}}(\tau_0) = \hat{\mathbf{F}}(\mathbf{x}(\tau_0),\mathbf{u}(\tau_0))$.
Without the assumption that the given inertial frame is the rest system of the particle, we define the *relativistic mass*

$$m(\tau) := \frac{m_0}{\sqrt{1-\left\|\frac{\hat{\mathbf{u}}(\tau)}{u^0(\tau)}\right\|^2}}.$$

The inertial frame is the rest frame of the particle at the event $\mathbf{x}(\tau_0)$ if and only if $\mathbf{u}(\tau_0) = \mathbf{e_0}$, i.e. if and only if $m(\tau_0) = m_0$. Otherwise, $m(\tau) > m_0$. We then have

$$\frac{d}{dx^0}\left(m\frac{\hat{\mathbf{u}}}{u^0}\right) = m_0\frac{d}{dx^0}\underbrace{\frac{\hat{\mathbf{u}}}{\sqrt{(u^0)^2-\|\hat{\mathbf{u}}\|^2}}}_{=1} = \frac{m_0}{u^0}\frac{d}{d\tau}\hat{\mathbf{u}} = \frac{m_0}{u^0}\hat{\mathbf{a}} = \frac{\hat{\mathbf{F}}}{u^0}.$$

This is the classical Newtonian equation of motion with mass $m$ and force $\frac{\hat{\mathbf{F}}}{u^0}$. Therefore, the relativistic mass is interpreted as the mass of the particle from the viewpoint of our inertial observer and $\frac{\hat{\mathbf{F}}}{u^0}$ as the *observed force* acting on the particle from the viewpoint of this observer.

**Definition 1.5.2.** For a particle with constant rest mass $m_0$ and a world line parametrized by proper time, the *four-momentum* is given by

$$\mathbf{p} := m_0 \cdot \mathbf{u}.$$

rest mass

four-velocity

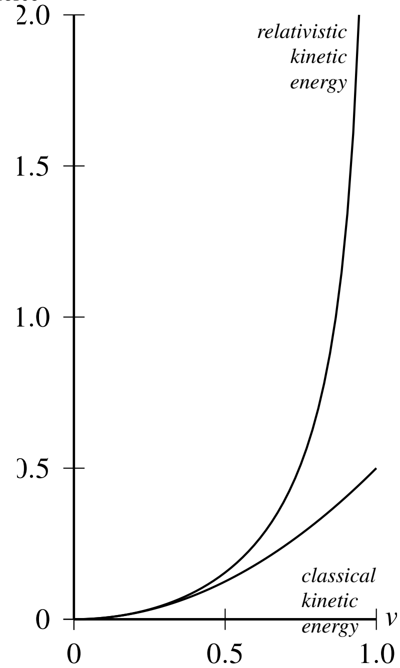
Equation (1.7) then takes the form

$$\frac{d}{d\tau}\mathbf{p}(\tau) = \mathbf{F}(\mathbf{x}(\tau), \mathbf{u}(\tau)).$$

As we have seen,

$$\frac{d}{dx^0}\Bigg(\underbrace{m}_{\text{relativistic mass}} \cdot \underbrace{\frac{\hat{\mathbf{u}}}{u^0}}_{\text{observed velocity}}\Bigg) = \underbrace{\frac{\hat{\mathbf{F}}}{u^0}}_{\text{observed force}}.$$

where $m = \frac{m_0}{\sqrt{1-\|\hat{\mathbf{u}}/u^0\|^2}}$. We write $u^0$ in terms of the observed velocity $\hat{\mathbf{u}}/u^0$:




$$\begin{aligned}
&-1 = \langle\langle \mathbf{u}, \mathbf{u} \rangle\rangle = -(u^0)^2 + \|\hat{\mathbf{u}}\|^2 \\
\Rightarrow\quad &\frac{1}{(u^0)^2} = 1 - \left\|\frac{\hat{\mathbf{u}}}{u^0}\right\|^2 \\
\Rightarrow\quad &u^0 = \frac{1}{\sqrt{1 - \left\|\frac{\hat{\mathbf{u}}}{u^0}\right\|^2}}
\end{aligned}$$

Hence we can write $m = u^0 m_0$ for the relativistic mass. The time component of the vector equation (1.7) is then

$$u^0 \frac{d}{dx^0} m = \frac{d}{d\tau}(m_0 \cdot u^0) \overset{(1.7)}{=} F^0(\mathbf{x}, \mathbf{u}) = \frac{1}{u^0}\left\langle \hat{\mathbf{F}}, \hat{\mathbf{u}} \right\rangle,$$

because $0 = \langle\langle \mathbf{F}(\mathbf{x}, \mathbf{u}), \mathbf{u} \rangle\rangle = -F^0 u^0 + \left\langle \hat{\mathbf{F}}, \hat{\mathbf{u}} \right\rangle$. This implies

$$\frac{d}{dx^0} m = \left\langle \frac{\hat{\mathbf{F}}}{u^0}, \frac{\hat{\mathbf{u}}}{u^0} \right\rangle.$$

This is the classical energy equation (1.2) with the relativistic mass $m$ instead of the kinetic energy $E$. Therefore we can interpret the relativistic mass as the energy of the particle as well.

$$E = m = \frac{m_0}{\sqrt{1 - \left\|\frac{\hat{\mathbf{u}}}{u^0}\right\|^2}} = \underbrace{m_0}_{\textit{rest energy}} + \underbrace{\underbrace{\frac{m_0}{2}\left\|\frac{\hat{\mathbf{u}}}{u^0}\right\|^2}_{\textit{classical kinetic energy}} + \frac{3m_0}{8}\left\|\frac{\hat{\mathbf{u}}}{u^0}\right\|^4 + \mathrm{O}\left(\left\|\frac{\hat{\mathbf{u}}}{u^0}\right\|^6\right)}_{\textit{relativistic kinetic energy}},$$

where we used that $\frac{1}{\sqrt{1-x}} = 1 + \frac{1}{2}x + \frac{3}{8}x^2 + \mathrm{O}(x^3)$. Since energy has has the same physical units as mass $\cdot$ velocity$^2$ we get the famous formula

$$\boxed{E_{\text{phys}} = m_{\text{phys}} \cdot c^2}$$

**Example 1.5.3 (Electromagnetic field).** Suppose we are given

$$\begin{aligned}\hat{\mathbf{E}} &= (E^1,E^2,E^3):\mathbb{R}\times\mathbb{R}^3\to\mathbb{R}^3 \text{ an electric field and}\\ \hat{\mathbf{B}} &= (B^1,B^2,B^3):\mathbb{R}\times\mathbb{R}^3\to\mathbb{R}^3 \text{ a magnetic field.}\end{aligned}$$

We combine the two fields to the *electromagnetic field*

$$\mathscr{F}:\mathbb{R}\times\mathbb{R}^3\to\{\text{skew-symmetric bilinear forms on }\mathbb{R}^4\}=\{\text{skew-symmetric }4\times 4\text{ matrices}\}$$

by

$$\mathscr{F}:=\begin{pmatrix} 0 & -E^1 & -E^2 & -E^3\\ E^1 & 0 & B^3 & -B^2\\ E^2 & -B^3 & 0 & B^1\\ E^3 & B^2 & -B^1 & 0\end{pmatrix},$$

or, more accurately,

$$\mathscr{F}|_{\mathbf{x}}=\begin{pmatrix} 0 & -E^1(\mathbf{x}) & -E^2(\mathbf{x}) & -E^3(\mathbf{x})\\ E^1(\mathbf{x}) & 0 & B^3(\mathbf{x}) & -B^2(\mathbf{x})\\ E^2(\mathbf{x}) & -B^3(\mathbf{x}) & 0 & B^1(\mathbf{x})\\ E^3(\mathbf{x}) & B^2(\mathbf{x}) & -B^1(\mathbf{x}) & 0\end{pmatrix}.$$

The electromagnetic field is not to be confused with a force field. It exposes a particle with charge $q$ to the force $\mathbf{F}$ that is characterized by

$$\langle\langle \mathbf{F}(\mathbf{x},\mathbf{u}),\mathbf{y}\rangle\rangle=-q\cdot\mathscr{F}|_{\mathbf{x}}(\mathbf{u},\mathbf{y})$$

for all $\mathbf{y}\in\mathbb{R}^4$. Observe that for $\mathbf{x}$ and $\mathbf{u}$ fixed, the right hand side is linear in $\mathbf{y}$. Because the Minkowski product is non-degenerate, there exists a unique vector $\mathbf{F}(\mathbf{x},\mathbf{u})$ such that the equation is fulfilled for all $\mathbf{y}$. Hence $\mathbf{F}$ is well defined. We then have

$$\langle\langle \mathbf{F}(\mathbf{x},\mathbf{u}),\mathbf{u}\rangle\rangle=-q\cdot\mathscr{F}(\mathbf{u},\mathbf{u})=0$$

as desired because $\mathscr{F}$ is skew-symmetric. Let us explicitly calculate $\mathbf{F}$.

$$\begin{aligned}\mathscr{F}\left(\begin{pmatrix}u^0\\ \hat{\mathbf{u}}\end{pmatrix},\begin{pmatrix}y^0\\ \hat{\mathbf{y}}\end{pmatrix}\right) &= \Big(\langle\hat{\mathbf{u}},\hat{\mathbf{E}}\rangle,-E^1u^0-u^2B^3+u^3B^2,-u^0E^2+u^1B^3-u^3B^1,\\ &\qquad -u^0E^3-u^1B^2+u^2B^1\Big)\begin{pmatrix}y^0\\ \hat{\mathbf{y}}\end{pmatrix}\\ &= \Big(\langle\hat{\mathbf{u}},\hat{\mathbf{E}}\rangle,-u^0\hat{\mathbf{E}}+\hat{\mathbf{B}}\times\hat{\mathbf{u}}\Big)\begin{pmatrix}y^0\\ \hat{\mathbf{y}}\end{pmatrix}\\ &= \left\langle\!\!\left\langle\begin{pmatrix}-\langle\hat{\mathbf{u}},\hat{\mathbf{E}}\rangle\\ -u^0\hat{\mathbf{E}}+\hat{\mathbf{B}}\times\hat{\mathbf{u}}\end{pmatrix},\begin{pmatrix}y^0\\ \hat{\mathbf{y}}\end{pmatrix}\right\rangle\!\!\right\rangle\end{aligned}$$

and therefore

$$\mathbf{F}(\mathbf{x},\mathbf{u})=q\begin{pmatrix}\langle\hat{\mathbf{u}},\hat{\mathbf{E}}(\mathbf{x})\rangle\\ u^0\hat{\mathbf{E}}(\mathbf{x})-\hat{\mathbf{B}}(\mathbf{x})\times\hat{\mathbf{u}}\end{pmatrix}.$$

The space coordinates result in the observable force

$$\frac{\hat{\mathbf{F}}}{u^0} = q\left(\hat{\mathbf{E}} - \hat{\mathbf{B}} \times \frac{\hat{\mathbf{u}}}{u^0}\right).$$

This is the formula for the *Lorentz force* acting on a charged particle, as established in electrodynamics. The energy equation reads

$$\frac{dm}{dx^0} = q\left\langle \hat{\mathbf{E}}, \frac{\hat{\mathbf{u}}}{u^0}\right\rangle.$$

Note that the magnetic field does not enter into the energy equation.

So far, we only discussed point particles. For extended bodies we have in classical mechanics
*Mass density*: $\rho : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}$. The total mass of a body at time $t$ is then given by

$$\int_{\mathbb{R}^3} \rho(t, \hat{\mathbf{x}})\, dx^1\, dx^2\, dx^3.$$

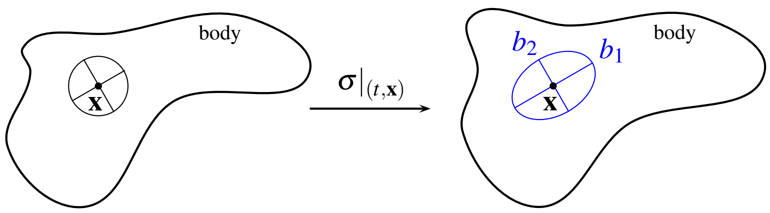
*Momentum density*: $\mathbf{p} : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$. The total momentum of the body at time $t$ is then given by

$$\int_{\mathbb{R}^3} \mathbf{p}(t, \hat{\mathbf{x}})\, dx^1\, dx^2\, dx^3.$$

*Stress tensor*: $\sigma : \mathbb{R} \times \mathbb{R}^3 \to \{\text{symmetric bilinear forms on } \mathbb{R}^3\}$. The physical interpretation is the following:
For a diagonalizing orthonormal basis $\mathbf{b_1}, \mathbf{b_2}, \mathbf{b_3}$ of $\sigma|_{(t,\hat{\mathbf{x}})}$, we have

$$\sigma|_{(t,\hat{\mathbf{x}})}(\mathbf{b_i}, \mathbf{b_j}) = \begin{cases} \lambda_i, \, i = j \\ 0, \, i \neq j \end{cases}$$

At time $t$ and at the point $\hat{\mathbf{x}}$, the body exerts pressure of strength $\lambda_i$ in direction $\mathbf{b_i}$.




In relativity theory, these entities are combined to the *stress-energy tensor*

$$\mathbf{T} : \mathbb{R}^4 \to \{\text{symmetric bilinear forms on } \mathbb{R}^4\},$$

where $\mathbf{T}$ is given by the matrix

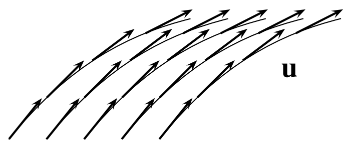
$$\left(\begin{array}{c|c} \rho & \mathbf{p}^\top \\ \hline \mathbf{p} & \sigma \end{array}\right).$$

To every extended body we assign such a $\mathbf{T}$, with the following physical interpretation: For an observer B whose world line has the four-velocity $\mathbf{u}$ at event $\mathbf{x}$,

$$
\begin{aligned}
\mathbf{T}|_{\mathbf{x}}(\mathbf{u},\mathbf{u}) &= \text{mass density of the body at the event } \mathbf{x}\text{, observed by B} \\
&= \text{energy density of the body at the event } \mathbf{x}\text{, observed by B} \\
\text{For } \mathbf{e} \perp\!\!\!\perp \mathbf{u}:\ \mathbf{T}|_{\mathbf{x}}(\mathbf{u},\mathbf{e}) &= \langle \text{momentum density of the body observed by B at event } \mathbf{x}, \mathbf{e}\rangle \\
\text{For } \mathbf{e},\mathbf{e}' \perp\!\!\!\perp \mathbf{u}:\ \mathbf{T}|_{\mathbf{x}}(\mathbf{e},\mathbf{e}') &= (\text{stress tensor of the body at event } \mathbf{x}\text{, observed by B})(\mathbf{e},\mathbf{e}')
\end{aligned}
$$

**Example 1.5.4.** (1) *Vacuum:* $\mathbf{T} = 0$.

(2) *Dust:* Positive mass density $\rho > 0$, no pressure. The four-velocities of the dust particles define a timelike, future-directed unit vector field $\mathbf{u}$, the momentum density vanishes. Hence




$$\mathbf{T}|_{\mathbf{x}}(\mathbf{y},\mathbf{z}) = \rho(\mathbf{x})\,\langle\langle \mathbf{y},\mathbf{u}(\mathbf{x})\rangle\rangle \cdot \langle\langle \mathbf{z},\mathbf{u}(\mathbf{x})\rangle\rangle\,.$$

(3) *Ideal Liquid:* Positive mass density $\rho > 0$, isotropic pressure, i.e. $\sigma|_{\mathbf{x}} = \lambda(\mathbf{x})\,\langle\cdot,\cdot\rangle$, momentum density vanishes. Again, the four-velocities of the liquid molecules define a timelike, future-directed unit vector field $\mathbf{u}$. Hence

$$\mathbf{T}|_{\mathbf{x}}(\mathbf{y},\mathbf{z}) = (\rho(\mathbf{x}) + \lambda(\mathbf{x}))\,\langle\langle \mathbf{y},\mathbf{u}(\mathbf{x})\rangle\rangle \cdot \langle\langle \mathbf{z},\mathbf{u}(\mathbf{x})\rangle\rangle + \lambda(\mathbf{x})\,\langle\langle \mathbf{y},\mathbf{z}\rangle\rangle\,.$$

(4) *Electromagnetic field:* Here one finds

$$
\begin{aligned}
T^{00} &= \frac{1}{8\pi}(\|\hat{\mathbf{E}}\|^2 + \|\hat{\mathbf{B}}\|^2), \\
T^{0j} = T^{j0} &= \frac{1}{4\pi}(\hat{\mathbf{E}} \times \hat{\mathbf{B}})^j, \\
T^{jk} &= \frac{1}{4\pi}\left[-(E^jE^k + B^jB^k) + \frac{1}{2}(\|\hat{\mathbf{E}}\|^2 + \|\hat{\mathbf{B}}\|^2)\delta^{jk}\right].
\end{aligned}
$$

Later we will find a more conceptual way of finding the energy stress tensor for the different kinds of matter.

## 1.6 Closing Remarks about Special Relativity

Let us summarize briefly the structure of special relativity, now making use of differential geometric language. Space and time are joined to the 4-dimensional spacetime. The Postulate of Special Relativity states that the mathematical model for spacetime is a time-oriented Lorentz manifold $M$ which is isometric to $(\mathbb{R}^4, g_{\mathrm{Mink}})$. An isometry $M \to (\mathbb{R}^4, g_{\mathrm{Mink}})$ preserving the time orientation is called an inertial frame. The coordinates $(x^0, \hat{\mathbf{x}})$ that are assigned to an event by such an isometry are the time and space coordinates from the point of view of an observer with world line $\mathbb{R} \cdot \mathbf{e_0}$.

The world lines of particles slower than light are the timelike smooth curves in $M$. The world lines of particles moving at the speed of light are null curves. The world lines of particles not subject to any acceleration are geodesics (straight lines) in $M$.
Let $\mathscr{H} := \{\xi \in TM \,|\, g(\xi,\xi) = -1 \text{ and } \xi \text{ is future directed}\}$. An external force is given by a vector field $\mathbf{F}$ along the footpoint map $\pi : \mathscr{H} \to M$ with $g(\mathbf{F}(\xi),\xi) = 0$ for all $\xi \in \mathscr{H}$. We have an analog to Newton's second law,

$$m_0 \frac{\nabla}{d\tau}\frac{d\mathbf{x}}{d\tau} = m_0 \frac{d^2}{d\tau^2}\mathbf{x}(\tau) = \frac{d}{d\tau}(m_0\mathbf{u})(\tau) = F(\mathbf{x}(\tau),\mathbf{u}(\tau)).$$

This equation can be studied in arbitrary coordinate systems, not only in inertial frames.
All relevant physical objects possess a stress-energy tensor containing information about the mass density, momentum density and stress density.

# 2 General Relativity

The goal is now to include gravitation into relativity theory. From now on the reader will be assumed to be familiar with differential geometry. We start by quickly reviewing classical Newtonian gravity theory.

## 2.1 Classical theory of gravitation

In a Galilean inertial frame *Newton's law of gravitation* holds: Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ the position vectors of two point particles with masses $m$ and $M$, respectively. Then $\mathbf{y}$ exerts the force

$$\mathbf{F} = -\frac{GmM}{\|\mathbf{x}-\mathbf{y}\|^2}\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|}$$

on $\mathbf{x}$. Here $G = 6{,}673 \cdot 10^{-11}\,\mathrm{m}^3\mathrm{kg}^{-1}\mathrm{s}^{-2}$ is the *gravitational constant*.
We assume that $M \gg m$ so that the gravitational force of $\mathbf{x}$ exerted on $\mathbf{y}$ is negligible. Hence we may assume that $\mathbf{y} = \mathbf{0} \in \mathbb{R}^3$ is constant. Combining the law of gravitation and Newton's second law $\mathbf{F} = m\ddot{\mathbf{x}}$ we get

$$\ddot{\mathbf{x}} = -\frac{GM}{\|\mathbf{x}\|^3}\mathbf{x}. \tag{2.1}$$

**Remark 2.1.1.** The mass $m$ of $\mathbf{x}$ has canceled in (2.1), so the orbit of $\mathbf{x}$ does not depend on its mass. A priori, one would have to distinguish between the inertial mass $m_{\mathrm{inert}}$ occurring in Newton's second law $\mathbf{F} = m_{\mathrm{inert}} \cdot \ddot{\mathbf{x}}$ and the gravitational mass $m_{\mathrm{grav}}$ in

$$\mathbf{F} = -\frac{Gm_{\mathrm{grav}}M_{\mathrm{grav}}}{\|\mathbf{x}-\mathbf{y}\|^2}\frac{\mathbf{x}-\mathbf{y}}{\|\mathbf{x}-\mathbf{y}\|}.$$

Equation (2.1) and hence the equality $m_{\mathrm{inert}} = m_{\mathrm{grav}}$ of these two concepts of mass is experimentally well tested (see `http://www.youtube.com/watch?v=5C5_dOEyAfk`) and is therefore considered an empirical fact.

Define the *angular momentum per mass* by $\mathbf{L}(t) := \mathbf{x}(t) \times \dot{\mathbf{x}}(t)$.

**Lemma 2.1.2 (Preservation of angular momentum)**
*If* $\mathbf{x}$ *satisfies equation* (2.1) *then* $\mathbf{L}$ *is constant.*

*Proof.* We compute

$$\frac{d}{dt}\mathbf{L} = \underbrace{\dot{\mathbf{x}} \times \dot{\mathbf{x}}}_{=\mathbf{0}} + \mathbf{x} \times \ddot{\mathbf{x}} \overset{(2.1)}{=} -\frac{GM}{\|\mathbf{x}\|^3} \underbrace{\mathbf{x} \times \mathbf{x}}_{=\mathbf{0}} = \mathbf{0}. \qquad \square$$

**Remark 2.1.3.** Assume that $\mathbf{x}$ satisfies (2.1) so that $\mathbf{L}$ is constant. If $\mathbf{L} \neq \mathbf{0}$, then $\mathbf{x}(t) \perp \mathbf{L}$ for all $t$. Hence $\mathbf{x}$ is confined to the plane perpendicular to $\mathbf{L}$.
If $\mathbf{L} = \mathbf{0}$, then $\mathbf{x}(t) = \mathbf{0}$ or $\dot{\mathbf{x}}(t) = \lambda(t)\mathbf{x}(t)$, that is

$$\mathbf{x}(t) = \mathbf{x}(t_0) \cdot e^{\int_{t_0}^{t} \lambda(s)\,ds}.$$

This means $\mathbf{x}(t)$ lies on the straight line through $\mathbf{0}$ and $\mathbf{x}(t_0)$ (with $t_0$ fixed). In this case $\mathbf{x}$ is even confined to a one-dimensional subspace.

Let $\mathbf{L} \neq \mathbf{0}$. We choose the coordinate system such that $\mathbf{L} = \|\mathbf{L}\| \cdot \mathbf{e_3}$. Hence $\mathbf{x}$ stays in the $x^1$-$x^2$-plane. We introduce polar coordinates $(r, \varphi)$ in the $x^1$-$x^2$-plane:

$$x^1 = r\cos\varphi \quad \text{and} \quad x^2 = r\sin\varphi.$$

We express (2.1) in polar coordinates:

$$\begin{aligned}
\ddot{\mathbf{x}} &= \underset{\substack{\text{covariant Derivative} \\ \text{w.r.t. } g_{\text{eukl}}}}{\frac{\nabla}{dt}}\dot{\mathbf{x}} = \frac{\nabla}{dt}\left(\dot{r}\frac{\partial}{\partial r} + \dot{\varphi}\frac{\partial}{\partial\varphi}\right) \\
&= \ddot{r}\frac{\partial}{\partial r} + \dot{r}\frac{\nabla}{dt}\frac{\partial}{\partial r} + \ddot{\varphi}\frac{\partial}{\partial\varphi} + \dot{\varphi}\frac{\nabla}{dt}\frac{\partial}{\partial\varphi} \\
&= \ddot{r}\frac{\partial}{\partial r} + \dot{r}\nabla_{\dot{r}\frac{\partial}{\partial r}+\dot{\varphi}\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial r} + \ddot{\varphi}\frac{\partial}{\partial\varphi} + \dot{\varphi}\nabla_{\dot{r}\frac{\partial}{\partial r}+\dot{\varphi}\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial\varphi} \qquad (2.2)
\end{aligned}$$

In polar coordinates the metric coefficients of the Euclidean metric $g_{\text{eukl}}$ are

$$(g_{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$$

and the Christoffel symbols are easily computed to be

$$\Gamma_{12}^2 = \Gamma_{21}^2 = \frac{1}{r}, \quad \Gamma_{22}^1 = -r, \text{ and all other } \Gamma_{ij}^k = 0.$$

Therefore

$$\nabla_{\dot{r}\frac{\partial}{\partial r}+\dot{\varphi}\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial r} = \dot{r}\underbrace{\nabla_{\frac{\partial}{\partial r}}\frac{\partial}{\partial r}}_{=\mathbf{0}} + \dot{\varphi}\underbrace{\nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial r}}_{=\frac{1}{r}\frac{\partial}{\partial\varphi}} = \frac{\dot{\varphi}}{r}\frac{\partial}{\partial\varphi}$$

and

$$\nabla_{\dot{r}\frac{\partial}{\partial r}+\dot{\varphi}\frac{\partial}{\partial \varphi}}\frac{\partial}{\partial \varphi} = \dot{r}\underbrace{\nabla_{\frac{\partial}{\partial r}}\frac{\partial}{\partial \varphi}}_{=\frac{1}{r}\frac{\partial}{\partial \varphi}} + \dot{\varphi}\underbrace{\nabla_{\frac{\partial}{\partial \varphi}}\frac{\partial}{\partial \varphi}}_{=-r\frac{\partial}{\partial r}} = \frac{\dot{r}}{r}\frac{\partial}{\partial \varphi} - \dot{\varphi} r \frac{\partial}{\partial r}.$$

Inserting this into (2.2) yields

$$\frac{\nabla}{dt}\dot{\mathbf{x}} = (\ddot{r} - \dot{\varphi}^2 r)\frac{\partial}{\partial r} + \left(\ddot{\varphi} + 2\dot{\varphi}\frac{\dot{r}}{r}\right)\frac{\partial}{\partial \varphi}.$$

Now (2.1) reads

$$\frac{\nabla}{dt}\dot{\mathbf{x}} = -\frac{GM}{r^3} r \frac{\partial}{\partial r} = -\frac{GM}{r^2}\frac{\partial}{\partial r}.$$

so that (2.1) takes the form

$$\ddot{r} - \dot{\varphi}^2 r = -\frac{GM}{r^2}, \quad \ddot{\varphi} + 2\dot{\varphi}\frac{\dot{r}}{r} = 0 \tag{2.3}$$

in polar coordinates.

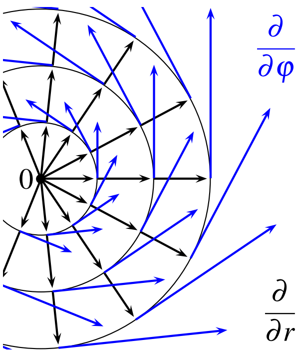
**Lemma 2.1.4 (Kepler's Second Law)**
*Let* $\mathbf{x}$ *satisfy* (2.1). *Then*

$$r^2\dot{\varphi} = \pm\|\mathbf{L}\|$$

*is constant.*

*Proof.* We compute

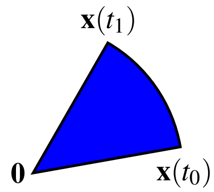
$$\begin{aligned} \|\mathbf{L}\| &= \|\mathbf{x} \times \dot{\mathbf{x}}\| \\ &= \left\| r\frac{\partial}{\partial r} \times \left(\dot{r}\frac{\partial}{\partial r} + \dot{\varphi}\frac{\partial}{\partial \varphi}\right)\right\| \\ &= r|\dot{\varphi}| \cdot \left\|\frac{\partial}{\partial r} \times \frac{\partial}{\partial \varphi}\right\| \\ &= r|\dot{\varphi}| \cdot \underbrace{\left\|\frac{\partial}{\partial r}\right\|}_{=1} \cdot \underbrace{\left\|\frac{\partial}{\partial \varphi}\right\|}_{=r} \\ &= r^2|\dot{\varphi}|. \end{aligned}$$




□

After possibly applying a reflection, we can w.l.o.g. assume $r^2\dot{\varphi} = \|\mathbf{L}\|$.

**Remark 2.1.5.** Kepler's second law is often formulated in a more geometrical way as follows: The line segment from $\mathbf{0}$ to the point $\mathbf{x}(t)$ sweeps out equal areas during equal intervals of time. To see this, we compute the area of the surface that is bordered by the line segments from $\mathbf{0}$ to $\mathbf{x}(t_0)$ and $\mathbf{x}(t_1)$ respectively ($t_0 < t_1$) and the corresponding segment of the orbit.




From differential geometry, it is known that the area element is given in polar coordinates by

$$\sqrt{\det(g_{ij})}\,dr\,d\varphi = \sqrt{\det\begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}}\,dr\,d\varphi = r\,dr\,d\varphi.$$

Employing the substitution rule for integration, we find for the area

$$\int_{\varphi(t_0)}^{\varphi(t_1)} \int_0^{r(\varphi)} r\,dr\,d\varphi = \int_{\varphi(t_0)}^{\varphi(t_1)} \frac{r(\varphi)^2}{2}\,d\varphi = \frac{1}{2}\int_{t_0}^{t_1} r(t)^2\,\dot{\varphi}(t)\,dt = \frac{\|L\|}{2}(t_1 - t_0).$$

So indeed, the area swept out is proportional to the time elapsed.

Now we restrict ourselves to the interesting case $\mathbf{L} \neq \mathbf{0}$. By Lemma 2.1.4 we have $r(t) > 0$ and $\dot{\varphi}(t) \neq 0$ for all $t$. Consider the auxiliary function $u : I \to \mathbb{R}$, given by

$$u(s) := \frac{1}{r(\varphi^{-1}(s))} \qquad \text{i.e.,} \qquad u(\varphi(t)) = \frac{1}{r(t)}.$$

Such a $u$ exists and is smooth because $\dot{\varphi} \neq 0$. For the sake of brevity we write a dot for $\frac{d}{dt}$ and a prime for $\frac{d}{ds}$.

**Lemma 2.1.6 (Orbit Equation)**
*Let* $\mathbf{x}$ *satisfy* (2.1). *Then we have*

$$u'' + u = \frac{GM}{\|\mathbf{L}\|^2}.$$

*Proof.* From $r^2\dot{\varphi} = \|\mathbf{L}\|$ we see

$$\dot{\varphi}(t) = \frac{\|\mathbf{L}\|}{r(t)^2} = \|\mathbf{L}\| \cdot u(\varphi(t))^2.$$

Hence

$$\dot{r}(t) = -\frac{u'(\varphi(t))}{u(\varphi(t))^2} \cdot \dot{\varphi}(t) = -\|\mathbf{L}\| \cdot u'(\varphi(t))$$

and therefore

$$\ddot{r}(t) = -\|\mathbf{L}\| \cdot u''(\varphi(t)) \cdot \dot{\varphi}(t) = -\|\mathbf{L}\|^2 \cdot u''(\varphi(t)) \cdot u(\varphi(t))^2.$$

Inserting this into (2.3) yields

$$-GMu^2 = -\|\mathbf{L}\|^2 u''u^2 - \|\mathbf{L}\|^2 u^4 \frac{1}{u} = -\|\mathbf{L}\|^2 u^2(u'' + u). \qquad \square$$

The orbit equation can be solved explicitly. Its general solution is

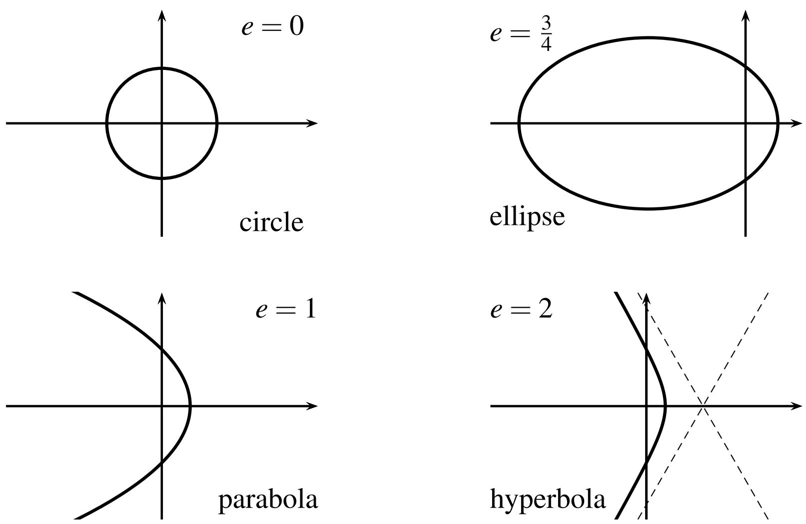
$$u(\varphi) = \frac{GM}{\|\mathbf{L}\|^2} + A\cos(\varphi - \varphi_0)$$

where $A, \varphi_0 \in \mathbb{R}$ are constants. After applying a rotation in the $x^1$-$x^2$-plane if necessary, we can assume w.l.o.g. that $\varphi_0 = \varphi(t_0) = 0$ and $A \geq 0$. We then have

$$r(t) = \frac{1}{\frac{GM}{\|\mathbf{L}\|^2} + A\cos(\varphi(t))} = \frac{\|\mathbf{L}\|^2/GM}{1 + e \cdot \cos(\varphi(t))},$$

where $e := \frac{A\|\mathbf{L}\|^2}{GM}$ is called the *eccentricity*. Geometrically, the solution is

1. an ellipse[1] for $0 \leq e < 1$ (Kepler's first law),
2. a parabola for $e = 1$,
3. a hyperbola for $e > 1$.




[1] As a special case we get a circle for $e = 0$.

The *gravitational potential* $V : \mathbb{R}^3 \setminus \{\mathbf{0}\} \to \mathbb{R}$ is defined by

$$V(\mathbf{x}) := -\frac{GM}{\|\mathbf{x}\|}.$$

We have

$$-\mathrm{grad} V = -\frac{GM}{r^2}\frac{\partial}{\partial r} = -\frac{GM}{r^3}\mathbf{x} = \frac{1}{m}\mathbf{F}.$$

For the energy we have

$$\begin{aligned} \text{kinetic energy:} \quad & E_{\mathrm{kin}} = \frac{m}{2}\|\dot{\mathbf{x}}\|^2 \\ \text{potential energy:} \quad & E_{\mathrm{pot}} = m \cdot V(\mathbf{x}) \\ \text{total energy:} \quad & E = E_{\mathrm{kin}} + E_{\mathrm{pot}}. \end{aligned}$$

**Lemma 2.1.7 (Energy Equation)**
*Let* $\mathbf{x}$ *satisfy* (2.1). *Then*

$$\frac{2}{m}E = \dot{r}^2 + \frac{\|\mathbf{L}\|^2}{r^2} - \frac{2GM}{r}$$

*is constant.*

*Proof.* We have

$$\|\dot{\mathbf{x}}\|^2 = \left\|\dot{r}\frac{\partial}{\partial r} + \dot{\varphi}\frac{\partial}{\partial \varphi}\right\|^2 = \dot{r}^2 \cdot 1 + \dot{\varphi}^2 \cdot r^2 = \dot{r}^2 + \frac{\|\mathbf{L}\|^2}{r^2}$$

because of $\dot{\varphi} = \frac{\|\mathbf{L}\|}{r^2}$. This implies

$$\frac{2}{m}E = \|\dot{\mathbf{x}}\|^2 - \frac{2GM}{\|\mathbf{x}\|} = \dot{r}^2 + \frac{\|\mathbf{L}\|^2}{r^2} - \frac{2GM}{r}.$$

Therefore

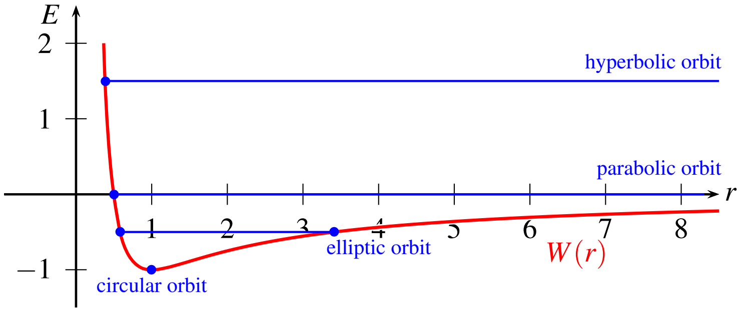
$$\frac{2}{m}\frac{d}{dt}E = 2\dot{r}\ddot{r} - 2\frac{\|\mathbf{L}\|^2\,\dot{r}}{r^3} + \frac{2GM\dot{r}}{r^2} = 0$$

by the equation of motion (2.3). □

We now define the *effective potential*

$$W(r) := \frac{\|\mathbf{L}\|^2}{r^2} - \frac{2GM}{r}.$$

By the energy equation $\dot{r}^2 + W(r)$ is constant. From $\dot{r}^2 \geq 0$ we get that $W(r) \leq \mathrm{const}$. The energy diagram then takes the following form:

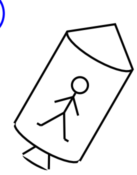
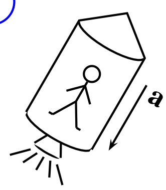
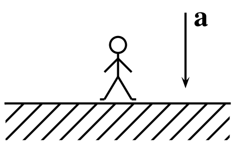
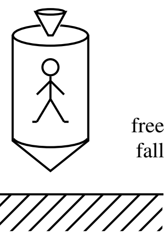



**Exercise 2.1.8.** Prove *Kepler's third law* which states that for elliptic orbits,

$$GM \cdot \text{orbital period}^2 = \frac{4\pi^2}{(1+e)^3} \cdot r_{\max}^3.$$

## 2.2 Equivalence Principle and the Einstein Field Equations

**Problem.** In classical mechanics, the gravitational field carries signals with infinite velocity. A similar gravitational law with a suitable Four-Force $\mathbf{F}$ is therefore problematic in relativity theory.
The *equivalence principle* wants to explain why $m_{\text{inert}} = m_{\text{grav}}$.

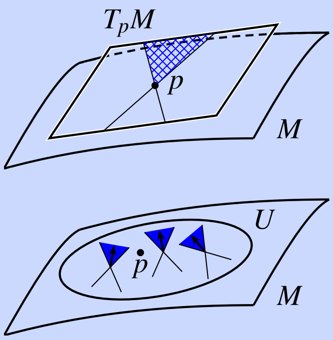
| | not accelerated | accelerated |
|---|---|---|
| in weight-lessness | A  $\mathbf{F} = \mathbf{0}$ | B   $\mathbf{F} = m\mathbf{a}$ |
| in a graviational field | C   $\mathbf{F} = m\mathbf{a}$ | D   $\mathbf{F} = \mathbf{0}$ |

By means of physical experiments, an observer (under complete isolation from the outside world) cannot distinguish A from D nor B from C. On the other hand, A and C are inertial

observer, but B and D are not.
Therefore, from now on we no longer demand the existence of global inertial frames. Indeed, realistic coordinate systems are usually only local and do not describe the whole universe. They locally approximate inertial frames. A spacetime will be modeled by a four-dimensional Lorentz manifold which is not necessarily Minkowski space.

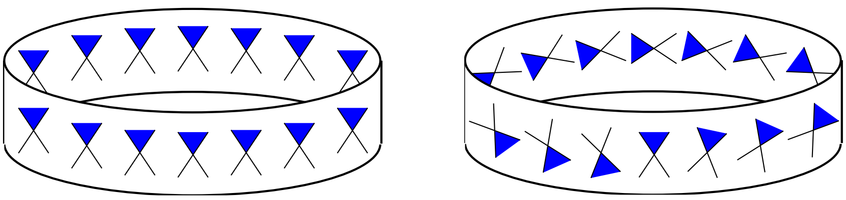
**Definition 2.2.1.** A *time orientation* on a Lorentz manifold $M$ is a mapping that assigns to each $p \in M$ one of the two connected components of

$$\mathscr{Z}_p := \{v \in T_pM \,|\, g|_p(v,v) < 0\}.$$




Write $\mathscr{Z}_p^{\uparrow}$ for this component. We require a time orientation to be continuous in the following sense: For all $p \in M$ there is an open neighborhood $U$ of $p$ and a continuous vector field $v$ on $U$ such that $v(q) \in \mathscr{Z}_q^{\uparrow}$ for all $q \in U$.

**Remark 2.2.2.** Not every Lorentz manifold is time orientable, i.e., there are Lorentz manifolds which do not have a time orientation. The following picture shows two different Lorentz metrics on the *same* manifold $\mathbb{R} \times S^1$ so that the first Lorentz manifold is time orientable whereas the second isn't.



time orientable    not time orientable

**Definition 2.2.3.** Once a time orientation is chosen, timelike tangent vectors $v \in \mathscr{Z}_p^{\uparrow}$ will be called *future directed*. If $-v \in \mathscr{Z}_p^{\uparrow}$, then $v$ will be called *past directed*.

From now on, gravitation will no longer be considered as an external force (as opposed to the electromagnetic force, for instance), but will be modeled by the geometry of the spacetime $M$. In special relativity the world lines of point particles subject to no force are straight lines. These are the geodesics of Minkowski space. Generalizing this, the world lines of point particles that move only under the influence of gravitation will be the geodesics of the spacetime.
**Question.** Which Lorentz manifold should be taken? What is the connection between geometry and physics?
On the physical side we will use the stress-energy tensor $T$ of the matter generating the gravitation. On the geometric side we will use the Einstein tensor $G$.
*What is this?*
Recall the basic curvature tensors of a Lorentz manifold:

1. The *Riemann curvature tensor*
$$R : T_pM \times T_pM \times T_pM \times T_pM \to \mathbb{R}.$$
2. The *Ricci curvature*
$$\mathrm{ric} : T_pM \times T_pM \to \mathbb{R}, \quad \mathrm{ric}(\xi,\eta) := \sum_{i=1}^{n} \varepsilon_i R(\xi,\mathbf{e_i},\mathbf{e_i},\eta),$$
where $\mathbf{e_1},\dots,\mathbf{e_n}$ is a generalized orthonormal basis, that is
$$g(\mathbf{e_i},\mathbf{e_j}) = \varepsilon_i \delta_{i,j} \text{ with } \varepsilon_i = \pm 1.$$
The map ric is a symmetric bilinear form on $T_pM$.
3. The *scalar curvature*
$$\mathrm{scal}(p) := \sum_{i=1}^{n} \varepsilon_i \mathrm{ric}(\mathbf{e_i},\mathbf{e_i}).$$
Then $\mathrm{scal} : M \to \mathbb{R}$ is a function on $M$.
4. The *Einstein tensor*
$$G := \mathrm{ric} - \tfrac{1}{2}\mathrm{scal}\cdot g.$$

Why do we use $G$ and not simply ric?

**Lemma 2.2.4**
*On any semi-Riemannian manifold we have* $2\mathrm{div}(\mathrm{ric}) = d\mathrm{scal}$ *and hence* $\mathrm{div}(G) = 0$.

*Proof.* The divergence of a symmetric $(0,2)$-tensor field like ric is a 1-form defined by

$$\mathrm{div}(\mathrm{ric})(X) = \sum_j \varepsilon_j \nabla_{\mathbf{e_j}} \mathrm{ric}(\mathbf{e_j}, X)$$

where $\mathbf{e_j}$ is a generalized orthonormal tangent frame, $g(\mathbf{e_j}, \mathbf{e_k}) = \varepsilon_j \delta_{jk}$ with $\varepsilon_j = \pm 1$. We now check the formula $2\mathrm{div}(\mathrm{ric}) = d\mathrm{scal}$ at a fixed point $p$ in the manifold and we may assume that $X$ and the tangent frame are synchronous at $p$, i.e., $\nabla X = \nabla \mathbf{e_j} = 0$ at $p$. Using the second Bianchi identity we get

$$\begin{aligned}
d\mathrm{scal}(X) &= \partial_X \sum_{jk} \varepsilon_j \varepsilon_k g(R(\mathbf{e_j}, \mathbf{e_k})\mathbf{e_k}, \mathbf{e_j}) \\
&= \sum_{jk} \varepsilon_j \varepsilon_k g(\nabla_X R(\mathbf{e_j}, \mathbf{e_k})\mathbf{e_k}, \mathbf{e_j}) \\
&= -\sum_{jk} \varepsilon_j \varepsilon_k g((\nabla_{\mathbf{e_j}} R(\mathbf{e_k}, X) + \nabla_{\mathbf{e_k}} R(X, \mathbf{e_j}))\mathbf{e_k}, \mathbf{e_j}) \\
&= -\sum_{jk} \varepsilon_j \varepsilon_k (g(\nabla_{\mathbf{e_j}} R(\mathbf{e_k}, X)\mathbf{e_k}, \mathbf{e_j}) + g(\nabla_{\mathbf{e_j}} R(X, \mathbf{e_k})\mathbf{e_j}, \mathbf{e_k})) \\
&= 2\sum_{jk} \varepsilon_j \varepsilon_k g(\nabla_{\mathbf{e_j}} R(\mathbf{e_k}, X)\mathbf{e_j}, \mathbf{e_k}) \\
&= 2\sum_{j} \varepsilon_j \nabla_{\mathbf{e_j}} \mathrm{ric}(X, \mathbf{e_j}) \\
&= 2\mathrm{div}(\mathrm{ric})(X).
\end{aligned}$$

The formula $\mathrm{div}(G) = 0$ follows readily. □

For physical reasons the stress-energy tensor $T$ is divergence free and hence so should be its geometric counterpart. This is the reason for preferring $G$ over ric.
We now postulate the **Einstein field equation.**

$$\boxed{\kappa \cdot T = G} \tag{EFE}$$

Here $\kappa$ is a universal constant. The value of $\kappa$ is determined by transition to the Newtonian limit. If

**(1)** $T$ is the stress-energy tensor of dust (only mass density),

**(2)** the Einstein field equation is replaced by its linearization and

**(3)** $c$ tends to infinity,

then the geodesic equations become Newton's equations of motions with

$$\kappa = 2,07 \cdot 10^{-48} \, \frac{\mathrm{s}^2}{\mathrm{g} \cdot \mathrm{cm}}.$$

It is possible to derive the Einstein field equation from a variational principle and one can give various heuristic arguments for it. Ultimately however, one has to verify it by checking the predicted results experimentally.

**Definition 2.2.5.** A Lorentz manifold $M$ is called *vacuum solution*, if $T \equiv 0$ and hence (by the Einstein field equation) $G \equiv 0$.

**Example 2.2.6.** Let $M$ be Minkowski space. Here we even have $R \equiv 0$.

**Lemma 2.2.7**
*In general, on* 4*-dimensional semi-Riemannian manifolds, we have*

$$\mathrm{ric} = G - \frac{1}{2} \sum_{i=1}^{4} \varepsilon_i \cdot G(\mathbf{e_i}, \mathbf{e_i}) \cdot g.$$

**Corollary 2.2.8**
*We have* $G = 0$ *if and only if* $\mathrm{ric} = 0$*. Hence by* (EFE) *the vacuum solutions are exactly the Ricci-flat Lorentz manifolds.*

*Proof of Lemma 2.2.7.*

$$\begin{aligned}\sum_{1=1}^{4}\varepsilon_i\cdot G(\mathbf{e_i},\mathbf{e_i}) &= \sum_{i=1}^{4}\varepsilon_i(\mathrm{ric}(\mathbf{e_i},\mathbf{e_i})-\tfrac{1}{2}\mathrm{scal}\cdot g(\mathbf{e_i},\mathbf{e_i}))\\ &= \mathrm{scal}-\tfrac{1}{2}\mathrm{scal}\cdot 4\\ &= -\mathrm{scal}\end{aligned}$$

$$\begin{aligned}\Longrightarrow\quad G-\tfrac{1}{2}\sum_{i=1}^{4}\varepsilon_i\cdot G(\mathbf{e_i},\mathbf{e_i})\cdot g &= \mathrm{ric}-\tfrac{1}{2}\mathrm{scal}\cdot g-\tfrac{1}{2}(-\mathrm{scal}\cdot g)\\ &= \mathrm{ric}\end{aligned}$$

□

Since the metric itself is divergence free, i.e., $\mathrm{div}g = 0$, the tensor $G+\Lambda g$ is also divergence free for any constant $\Lambda$. Therefore $G+\Lambda g$ could replace $G$ in the field equation which leads to the **Einstein field equation with cosmological constant**:

$$G+\Lambda\cdot g = \kappa\cdot T, \qquad (\mathrm{EFE}_\Lambda)$$

where $\Lambda\in\mathbb{R}$ is called the *cosmological constant*. The general opinion whether or not one should allow a nonzero cosmological constant has changed various times. Einstein once considered its introduction as the "greatest stupidity of his life" but changed his mind later. At the moment, a nonzero cosmological constant is often considered.

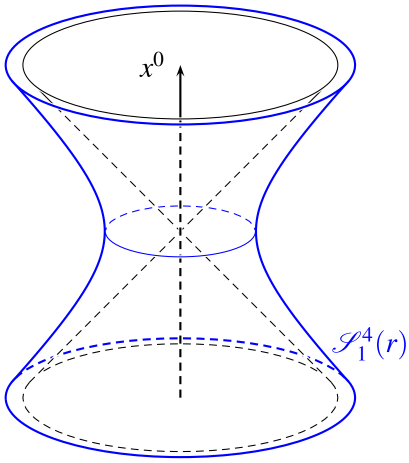
**Example 2.2.9 (deSitter spacetime).** Let $r>0$. Set

$$\mathrm{S}^4_1(r) := \{\mathbf{x}\in(\mathbb{R}^5, g_{\mathrm{Mink}})\,|\,\langle\langle\mathbf{x},\mathbf{x}\rangle\rangle = r^2\}.$$

A Lorentz metric is obtained by restricting $\langle\langle\cdot,\cdot\rangle\rangle$ to the tangent spaces of $\mathrm{S}^4_1(r)$. This way, one gets a four-dimensional Lorentz manifold. A time orientation is defined by requiring $x^0>0$. Calculation shows

$$G = -\frac{3}{r^2}g.$$

Hence $\mathrm{S}^4_1(r)$ is a vacuum solution of $(\mathrm{EFE}_\Lambda)$ with $\Lambda = \frac{3}{r^2}$.




**Convention.** From now on, choose physical units in such a way that the speed of light and the gravitational constant are equal to 1. This leads to $\kappa = 8\pi$.

## 2.3 Robertson-Walker spacetime

**Goal.** Find a simple model for the whole spacetime.
**Ansatz.** Describe the "spacial part" of the universe by a three-dimensional Riemannian manifold $(S, g_S)$ which is connected and complete (i.e. geodesics are defined for all times). Let us assume that the space part of the universe looks the same in whatever direction one looks (at least when observing objects not too far away). This property is called (local) isotropy and is formulated mathematically as follows.

**Definition 2.3.1.** A manifold $S$ is called *locally isotropic*, if for all $p \in S$ and for all $X, Y \in T_pS$ with $\|X\| = \|Y\|$, there exists an open neighborhood $U$ of $p$ and an isometry $\Phi : U \to U$ with $\Phi(p) = p$ and $d\Phi|_p(X) = Y$.

On a 3-dimensional Riemannian manifold we have for the sectional curvature that $K(E) = K(E')$ whenever $E, E' \subset T_pS$ are 2-dimensional subspaces. This can be seen as follows:
Given planes $E, E' \subset T_pS$ choose $X \in T_pS$ with $\|X\| = 1$ and $X \perp E$ as well as $Y \in T_pS$ with $\|Y\| = 1$ and $Y \perp E'$. Then an appropriate local isometry $\Phi$ takes $X$ to $Y$, i.e., $d\Phi|_p(X) = Y$. Hence $\Phi(E) = E'$. This implies

$$K(E') = K(d\Phi|_p(E)) = K(E),$$

as claimed.

If the manifold is connected, Schur's theorem (compare the differential geometry lecture) now implies

$$K \equiv \varepsilon,$$

i.e., the sectional curvature is constant.

**Example 2.3.2.** Here is a table for the candidates of our Riemannian manifold $S$ for the different signs of $\varepsilon$.

| $\varepsilon$ | $-1$ | $0$ | $1$ |
|---|---|---|---|
| model spaces | $(H^3, g_{\text{hyp}})$ | $(\mathbb{R}^3, g_{\text{eucl}})$ | $(S^3, g_{\text{std}})$ |
| | $\vdots$ | $T^3 = S^1 \times S^1 \times S^1$<br>$S^1 \times \mathbb{R}^2$<br>$T^2 \times \mathbb{R}$<br>$\vdots$ | $\mathbb{R}P^3$<br>$\vdots$ |
| | infinitely many, not completely understood | essentially finitely many, all known | infinitely many, all known |
| | some compact, others not | some compact, others not | all compact |

Now set

$$M := I \times S,$$

for our spacetime, where $I \subset \mathbb{R}$ is an open interval. For the Lorentz metric we make the ansatz

$$\boxed{g = -dt \otimes dt + f(t)^2 \cdot g_S}$$

where $t \in I$ and $f : I \to \mathbb{R}$ is a positive smooth function (*warped product*). Put differently: For $\xi = \alpha \frac{\partial}{\partial t} + X, \eta = \beta \frac{\partial}{\partial t} + Y \in T_{(t,p)}M$ with $X, Y \in T_pS$, we have

$$g(\xi, \eta) = -\alpha\beta + f(t)^2 \cdot g_S(X, Y).$$

**Example 2.3.3.** For $(S, g_S) = (\mathbb{R}^3, g_{\text{eukl}})$, $I = \mathbb{R}$ and $f = 1$ we obtain the Minkowski space $(M, g) = (\mathbb{R}^4, g_{\text{Mink}})$.

**Remark 2.3.4.** If $(N, g)$ is a Riemannian manifold with sectional curvature $K$ and if $c > 0$, then the Riemannian manifold $(N, c^2 \cdot g)$ has the sectional curvature $\frac{1}{c^2}K$. For this reason it suffices to consider the cases $\varepsilon = -1, 0, 1$.

### Geodesics of the spacetime

A straightforward computation shows: A curve $s \mapsto c(s) = (t(s), \gamma(s))$ is a geodesic in $M$ if and only if

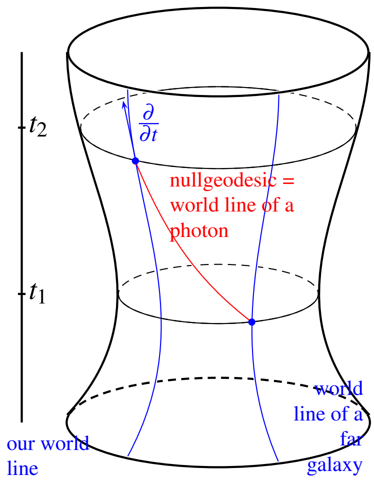
(i) $\dfrac{d^2t}{ds^2} + f(t)\dot{f}(t)g_S(\gamma'(s), \gamma'(s)) = 0$ and

(ii) $\dfrac{\nabla}{ds}\gamma'(s) + 2\dfrac{\dot{f}(t)}{f(t)} \cdot \dfrac{dt}{ds} \cdot \gamma'(s) = 0.$

**Example 2.3.5.** The curve $c(s) = (s, \gamma_0)$, where $\gamma_0 \in S$ is constant, is a timelike geodesic. We interpret it as the world line of a galaxy.

Let now $s \mapsto c(s) = (t(s), \gamma(s))$ be a null geodesic. Then

$$\begin{aligned} 0 &= g\left(\frac{dc}{ds}, \frac{dc}{ds}\right) \\ &= g\left(\frac{dt}{ds}\frac{\partial}{\partial t} + \gamma'(s), \frac{dt}{ds}\frac{\partial}{\partial t} + \gamma'(s)\right) \\ &= -\left(\frac{dt}{ds}\right)^2 + f^2 \cdot g_S(\gamma'(s), \gamma'(s)). \end{aligned}$$




This implies

$$\begin{aligned} \frac{d}{ds}\left(f \cdot \frac{dt}{ds}\right) &= \dot{f} \cdot \left(\frac{dt}{ds}\right)^2 + f \cdot \frac{d^2t}{ds^2} \\ &= \dot{f} \cdot f^2 \cdot g_S(\gamma', \gamma') + f \cdot \frac{d^2t}{ds^2} \overset{\text{(i)}}{=} 0. \end{aligned}$$

Hence $f \cdot \frac{dt}{ds}$ is constant.
From the point of view of an observer with world line $s \mapsto (s, \gamma_0)$, we get for the energy of a photon

$$E = g\left(\frac{\partial}{\partial t}, \frac{dt}{ds}\frac{\partial}{\partial t} + \gamma'(s)\right) = -\frac{dt}{ds} \implies \frac{E(t_1)}{E(t_2)} = \frac{f(t_2)}{f(t_1)}.$$

**Definition 2.3.6.** The quantity

$$z := \frac{f(t_2) - f(t_1)}{f(t_1)} = \frac{f(t_2)}{f(t_1)} - 1$$

is called *redshift* (of the null geodesic).

Redshift can be observed and measured very precisely because chemical elements emit light at specific energy levels (frequencies, colors). This light, emitted by other other galaxies arrives with a color shift that is easy to determine.
Taylor expansion of $f$ in the variable $t_2$ yields

$$\begin{aligned} f(t_1) &= f(t_2) + \dot{f}(t_2)(t_1 - t_2) + O(|t_1 - t_2|^2) \\ &= f(t_2)\big(1 + H(t_2)(t_1 - t_2) + O(|t_1 - t_2|^2)\big), \end{aligned}$$

where $H(t) = \frac{\dot{f}(t)}{f(t)}$. This implies

$$\begin{aligned} z &= \frac{1}{1 + H(t_2)(t_1 - t_2) + O(|t_1 - t_2|^2)} - 1 \\ &= 1 - H(t_2)(t_1 - t_2) + O(|t_1 - t_2|^2) + O(|t_1 - t_2|^2) - 1 \\ &= H(t_2)(t_1 - t_2) + O(|t_1 - t_2|^2). \end{aligned}$$

Hence if we observe light from galaxies not too far away, such that the term $O(|t_1 - t_2|^2)$ is negligible compared to the term $H(t_2)(t_1 - t_2)$, then the redshift is essentially proportional to the time difference $|t_1 - t_2|$, hence to the distance to the other galaxy. The constant $H(\text{now})$ is called the *Hubble constant*. In fact, one observes $z > 0$, so the Hubble constant is positive. Therefore $\dot{f}(\text{now})$ is positive, i.e., the universe is currently extending.
The formulas for warped products give

$$\begin{aligned} \mathrm{ric}\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial t}\right) &= -3\frac{\ddot{f}}{f}, \\ \mathrm{ric}\left(\frac{\partial}{\partial t}, X\right) &= \mathrm{ric}\left(X, \frac{\partial}{\partial t}\right) = 0 \text{ and} \\ \mathrm{ric}(X,Y) &= \left\{2\left(\frac{\dot{f}}{f}\right)^2 + 2\frac{\varepsilon}{f^2} + \frac{\ddot{f}}{f}\right\} g(X,Y), \end{aligned}$$

where $X$ and $Y$ are tangent to $S$. This implies

$$\mathrm{scal} = 3\frac{\ddot{f}}{f} + 3\left\{2\left(\frac{\dot{f}}{f}\right)^2 + 2\frac{\varepsilon}{f^2} + \frac{\ddot{f}}{f}\right\} = 6\left(\frac{\dot{f}}{f}^2 + \frac{\varepsilon}{f^2} + \frac{\ddot{f}}{f}\right)$$

and we get for the Einstein tensor

$$\begin{aligned} G\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial t}\right) &= 3\left\{\left(\frac{\dot{f}}{f}\right)^2 + \frac{\varepsilon}{f^2}\right\}, \\ G\left(\frac{\partial}{\partial t}, X\right) &= G\left(X, \frac{\partial}{\partial t}\right) = 0, \\ G(X,Y) &= -\left\{\left(\frac{\dot{f}}{f}\right)^2 + \frac{\varepsilon}{f^2} + 2\frac{\ddot{f}}{f}\right\} g(X,Y). \end{aligned}$$

So in this setting, the Einstein field equations are

$$\frac{8\pi}{3}\rho = \left(\frac{\dot f}{f}\right)^2 + \frac{\varepsilon}{f^2}, \tag{6}$$

$$-8\pi p = \left(\frac{\dot f}{f}\right)^2 + \frac{\varepsilon}{f^2} + 2\frac{\ddot f}{f}, \tag{7}$$

where $\rho$ is the energy/mass density and $p$ the (isotropic) pressure.

**Definition 2.3.7.** Let $(S, g_S)$ by a complete, 3-dimensional Riemannian manifold with constant sectional curvature $K \equiv \varepsilon \in \{-1, 0, 1\}$. Let $f : I \to \mathbb{R}$ be a positive smooth function. Then $(M, g)$ with

$$M = I \times S, \quad g = -dt \otimes dt + f(t)^2 g_S,$$

is called a *Robertson-Walker spacetime*.

**Example 2.3.8.** We already saw that the four-dimensional Minkowski space is a Robertson-Walker spacetime. Here $(S, g_S) = (\mathbb{R}^3, g_{\text{eukl}})$, $\varepsilon = 0$ and $f \equiv 1$.

Subtracting (6) from (7) gives

$$3\frac{\ddot f}{f} = -4\pi(\rho + 3p). \tag{8}$$

Differentiation of (6) and insertion of (8) and (6) yields

$$\begin{aligned}
\frac{8\pi}{3}\dot\rho &= 2\cdot\frac{\dot f}{f}\cdot\frac{\ddot f f - \dot f^2}{f^2} - 2\cdot\frac{\varepsilon \dot f}{f^3} \\
&= \left(2\frac{\ddot f}{f} - 2\left(\left(\frac{\dot f}{f}\right)^2 + \frac{\varepsilon}{f^2}\right)\right)\cdot\frac{\dot f}{f} \\
&= \left(-\frac{8\pi}{3}(\rho + 3p) - \frac{16\pi}{3}\rho\right)\cdot\frac{\dot f}{f} \\
&= (-8\pi\rho - 8\pi p)\cdot\frac{\dot f}{f},
\end{aligned}$$

hence

$$\dot\rho = -3(\rho + p)\cdot\frac{\dot f}{f}. \tag{9}$$

### Singularities

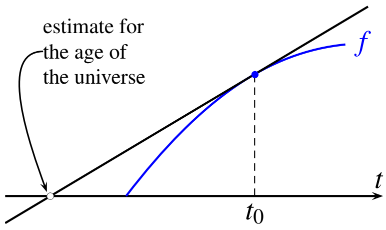
Let the domain $I = (t_*, t^*)$ of $f$ be maximal in the sense that $f$ cannot be extended beyond $I$ as a positive smooth function. Here $-\infty \leq t_* < t^* \leq \infty$.

**Definition 2.3.9.** (1) $t_*$ or $t^*$ is called a *physical singularity*, if $\rho \to \infty$ for $t \searrow t_*$ or $t \nearrow t^*$, respectively.

(2) $t_*$ is called a *big bang*, if $f(t) \to 0$ and $\dot{f}(t) \to \infty$ for $t \searrow t_*$.

(3) $t^*$ is called a *big crunch* or *collapse*, if $f(t) \to 0$ and $\dot{f}(t) \to -\infty$ for $t \nearrow t^*$.

**Remark 2.3.10.** If $\rho + 3p \geq 0$ and $H_0 = H(t_0) > 0$, then $M$ has a starting singularity, i.e., $t_* > -\infty$. To see this, notice that $f$ is concave because $\ddot{f} = -\frac{4\pi}{3}(\rho + 3p)f \leq 0$. From $H(t_0) > 0$ we see $\dot{f}(t_0) > 0$ and, by concavity, $\dot{f} \geq \dot{f}(t_0)$ on $(t_*, t_0]$. This implies that for any $t_1 \in (t_*, t_0]$,




$$f(t_0) > f(t_0) - f(t_1) = \int_{t_1}^{t_0} \dot{f}(t)\, dt \geq (t_0 - t_1) \cdot \dot{f}(t_0)$$

and thus

$$(t_0 - t_1) \leq \frac{f(t_0)}{\dot{f}(t_0)} = \frac{1}{H_0}.$$

If we let $t_1$ tend to $t_*$, we obtain

$$(t_0 - t_*) \leq \frac{1}{H_0}.$$

This way, we did not only show $t_* > -\infty$, but also derived an upper bound for the age of the universe in terms of the Hubble constant. Remember that the Hubble constant can be quite well determined experimentally via the observation of redshift. Current estimates for the age of the universe give a value of about 13.8 billion years.

**Proposition 2.3.11**
*Suppose $t_*$ and $t^*$ are physical singularities if they are finite. Let $H_0 > 0$, $\rho > 0$ and suppose there are constants $-\frac{1}{3} < a < A$, such that $a \leq \frac{p}{\rho} \leq A$. Then*

*(1) The initial singularity $t_*$ is a big bang.*

*(2) If $\varepsilon = 0$ or $\varepsilon = -1$, then $I = (t_*, \infty)$ and $f \to \infty$, $\rho \to 0$ for $t \to \infty$.*

*(3) If $\varepsilon = 1$, then $I = (t_*, t^*)$ and $t^* < \infty$ is a big crunch.*

*Proof.* Set $\delta := 3a + 1$. Because of $a > -\frac{1}{3}$ we have $\delta > 0$ and

$$\begin{aligned} -\frac{1}{3} + \frac{\delta}{3} = a \leq \frac{p}{\rho} \quad &\Longrightarrow \quad 3p + \rho \geq \delta\rho > 0 \\ &\overset{(8)}{\Longrightarrow} \quad \ddot{f} < 0 \\ \big(\text{with } \dot{f}(t_0) > 0\big) \quad &\Longrightarrow \quad \dot{f} > 0 \text{ on } (t_*, t_0]. \end{aligned}$$

Furthermore,

$$\dot{\rho} \overset{(9)}{=} -3(\rho + p)\frac{\dot{f}}{f} \geq -3(\rho + A\rho)\frac{\dot{f}}{f} = -C \cdot \rho \cdot \frac{\dot{f}}{f}$$

with $C := 3(1 + A) > 0$. This implies

$$\begin{aligned} &(\ln \rho)^{\cdot} \geq -C(\ln f)^{\cdot} = (\ln f^{-C})^{\cdot} \\ \Longrightarrow \quad &(\ln(\rho f^C))^{\cdot} \geq 0 \\ \Longrightarrow \quad &(\rho f^C)^{\cdot} \geq 0 \\ \Longrightarrow \quad &\rho \cdot f^C \leq \rho(t_0) f(t_0)^C \quad \text{on} \quad (t_*, t_0] \\ \Longrightarrow \quad &f \to 0 \quad \text{for} \quad t \to t_* \end{aligned}$$

Furthermore we have

$$\dot{\rho} \overset{(9)}{=} -3(\rho + p)\frac{\dot{f}}{f} \leq -(\delta\rho + 2\rho)\frac{\dot{f}}{f} = -(2 + \delta)\rho \cdot \frac{\dot{f}}{f}$$

and in a similar fashion to before we obtain

$$(\rho f^{2+\delta})^{\cdot} \leq 0 \quad \Longrightarrow \quad \underbrace{\rho \cdot f^{2+\delta}}_{=\rho f^2 \cdot f^\delta} \geq \rho(t_0) f(t_0)^{2+\delta} \text{ on } (t_*, t_0].$$

With $f \to 0$, we get $f^\delta \to 0$ and thus $\rho f^2 \to \infty$ for $t \to t_*$. With (6) we then get

$$\infty \leftarrow \frac{8\pi}{3}\rho f^2 = \dot{f}^2 + \varepsilon \quad \Longrightarrow \quad \dot{f} \to \infty.$$

This shows (1).
*Case 1:* The function $f$ has a maximum at $t_m \in I$.

$$0 < \rho(t_m) \overset{(6)}{=} \frac{3}{8\pi}\Big\{\overbrace{\frac{\dot{f}(t_m)^2}{f(t_m)^2}}^{=0} + \frac{\varepsilon}{f(t_m)^2}\Big\} \quad \Longrightarrow \quad \varepsilon > 0, \quad \text{which means} \quad \varepsilon = 1.$$

With $\ddot{f} < 0$ this implies $\dot{f} < 0$ on $(t_m, t^*)$. A discussion similar to the one before shows that $t^*$ is a big crunch. This is the situation in (3).
*Case 2:* The function $f$ does not have a maximum on $I$.
This implies $\dot{f} > 0$ on $I$ because $\ddot{f} < 0$. From $\rho > 0$ and $3p + \rho > 0$ it follows that $3(p+\rho) > 0$. Hence

$$\dot{\rho} = -3(\rho + p)\frac{\dot{f}}{f} < 0.$$

Thus $I$ does not have an ending singularity, $t^* = \infty$.
*Subcase A:* $f \to \infty$ for $t \to \infty$.
We have $(\rho f^{2+\delta})^{\cdot} \leq 0$ so that $\rho f^{2+\delta}$ is bounded on $(t_0, \infty)$. With $f \to \infty$, this implies $\rho f^2 \to 0$ for $t \to \infty$. Now (6) implies

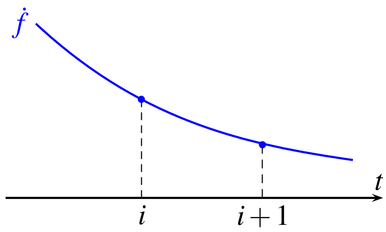
$$0 \leftarrow \frac{8\pi}{3}\rho f^2 \overset{(6)}{=} \dot{f}^2 + \varepsilon \quad \Longrightarrow \quad \varepsilon \leq 0$$

i.e., $\varepsilon = -1$ or $e = 0$. This is the situation in (2).
*Subcase B:* $f \to b < \infty$ for $t \to \infty$.
This implies $\dot{f} \to 0$ for $t \to \infty$.

$$\Longrightarrow \quad \frac{8\pi}{3}\rho f^2 = \dot{f}^2 + \varepsilon \overset{t\to\infty}{\longrightarrow} \varepsilon \quad \Longrightarrow \quad \varepsilon \geq 0$$

which means $\varepsilon = 0$ or $\varepsilon = 1$.

Because $\rho f^C$ is increasing, we have $\rho f^2 \nrightarrow 0$. This implies $\varepsilon \neq 0$, i.e. $\varepsilon = 1$. This shows $\rho f^2 \to \frac{3}{8\pi}$ for $t \to \infty$. By the mean value theorem, there is a sequence $t_i \in (i, i-1)$ with $\ddot{f}(t_i) = \dot{f}(i+1) - \dot{f}(i)$. Because of $\dot{f} \to 0$ we then get $\ddot{f}(t_i) \to 0$.




This implies

$$3\frac{\ddot{f}}{f} \overset{(8)}{=} -4\pi(\rho + 3p) \quad \Longrightarrow \quad \rho(t_i) + 3p(t_i) \to 0 \text{ for } i \to \infty,$$

i.e., $\rho(t_i) + 3p(t_i) \to 0$ for $i \to \infty$. On the other hand,

$$0 \leftarrow \rho(t_i) + 3p(t_i) \geq \delta\rho(t_i) \quad \Longrightarrow \quad \rho(t_i) \to \infty \quad \Longrightarrow \quad \rho(t_i)f(t_i)^2 \to 0$$

This is a contradiction, so Subcase B does not occur. □

**Remark 2.3.12.** In the literature, the case $\varepsilon \leq 0$ is often called *open* case as opposed to the *closed* case $\varepsilon > 0$, because for the model spaces we have

$$S \text{ is noncompact for } \varepsilon \leq 0 \text{ and } S \text{ is compact for } \varepsilon > 0.$$

Indeed, $S$ is always compact when $\varepsilon = 1$. However, when $\varepsilon = 0$ or $\varepsilon = -1$, the manifold $S$ can be compact as well, for example $S = T^3$. Therefore this terminology is somewhat misleading.

**Definition 2.3.13.** The constant $\rho_c := \dfrac{3{H_0}^2}{8\pi}$ is called the *critical energy density*.

The reason for this terminology is given by

**Proposition 2.3.14**
*We have*

$$\begin{aligned}
\rho(t_0) < \rho_c &\iff \varepsilon = -1 \\
\rho(t_0) = \rho_c &\iff \varepsilon = 0 \\
\rho(t_0) > \rho_c &\iff \varepsilon = 1
\end{aligned}$$

*Proof.* From the equation

$$\rho(t_0) - \rho_c \stackrel{(6)}{=} \frac{3}{8\pi}\left\{{H_0}^2 + \frac{\varepsilon}{f(t_0)} - {H_0}^2\right\} = \frac{3}{8\pi} \cdot \frac{\varepsilon}{f(t_0)}$$

we see that the sign of $\varepsilon$ is the same as that of $\rho(t_0) - \rho_c$. □

**Definition 2.3.15.** A *dust cosmos* is a Robertson-Walker spacetime with $p = 0$. A dust cosmos with $H_0 > 0$ is called *Friedmann cosmos*.

**Proposition 2.3.16**
*For a Robertson-Walker spacetime $M$ with nonconstant $f$, the following statements are equivalent.*

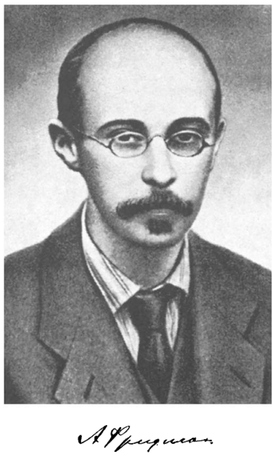
*(i)* *$M$ is a dust cosmos.*

*(ii)* $\rho \cdot f^3 =: m$ *is constant.*

*(iii)* *The* Friedmann equation

$$\dot{f}^2 + \varepsilon = \frac{A}{f}$$

*holds with* $A = \frac{8\pi}{3} m > 0$.



Alexander Alexandrovich Friedmann (1888–1925) [2]

*Proof.* “(ii) ⇔ (iii)” is clear because of

$$\frac{8\pi}{3}\rho f^3 \overset{(6)}{=} f \cdot (\dot{f}^2 + \varepsilon).$$

“(i)⇒(ii)” We have $\dot{\rho} \overset{(9)}{=} -3\rho \frac{\dot{f}}{f}$ which implies

$$(\ln \rho)^{\cdot} - 3(\ln f)^{\cdot} = 0.$$

Hence $(\ln \rho f^3)^{\cdot} = 0$ and $(\rho f^3)^{\cdot} = 0$.
“(ii)⇒(i)” On the one hand,

$$(\rho f^3)^{\cdot} = 0 \Longrightarrow \dot{\rho} = -3\rho \frac{\dot{f}}{f}.$$

On the other hand,

$$\dot{\rho} = -(3\rho + p)\frac{\dot{f}}{f}.$$

Therefore

$$p \cdot \dot{f} = 0.$$

Set $J := \{t \in I \mid p(t) \neq 0\}$. We have to show $J = \emptyset$. Suppose $J \neq \emptyset$ and let $J_0$ be a connected component of $J$ which must be an open interval. It follows that $\dot{f} \equiv 0$ on $J_0$, hence $f \equiv c > 0$ on $J_0$. By (7) this means

$$-8\pi p \equiv \frac{\varepsilon}{c^2} \neq 0 \quad \text{on} \quad J_0.$$

Because $p$ is continuous,

$$p \equiv -\frac{\varepsilon}{8\pi c^2} \neq 0 \quad \text{on} \quad \bar{J}_0$$

[2]Source: `http://en.wikipedia.org/wiki/Alexander_Friedmann`

where $\overline{J}_0$ is the closure of $J_0$ in $I$. This implies $\overline{J_0} \subset J$, i.e., $J_0 = \overline{J}_0$. This means that either $J_0 = \emptyset$ or $J_0 = I$. The latter implies that $f$ is constant, contradictory to the assumption. $\square$

We now determine the solutions of the Friedmann equation. Without loss of generality, let $t_* = 0$.

(1) $\varepsilon = 0$: $f(t) = (\frac{3}{2})^{\frac{2}{3}} A^{\frac{1}{3}} t^{\frac{2}{3}}$ (semicubical parabola).

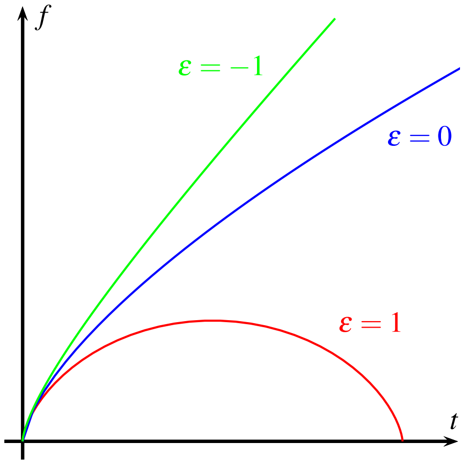
(2) $\varepsilon = -1$: Set $T := \int_0^t \frac{dt'}{f(t')}$. Then

$$t = \frac{A}{2}(\sinh(T) - T) \text{ and } f = \frac{A}{2}(\cosh(T) - 1).$$

(3) $\varepsilon = 1$: Set $T$ as before. Then

$$t = \frac{A}{2}(T - \sin(T)) \text{ and } f = \frac{A}{2}(1 - \cos(T))$$

(cycloid).



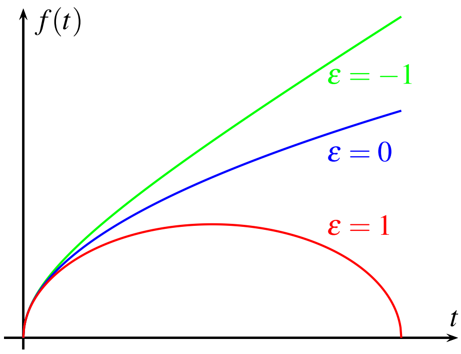

**Definition 2.3.17.** A Robertson-Walker spacetime is called a *radiation cosmos*, if

$$p = \frac{\rho}{3}.$$

**Exercise 2.3.18.** Show

(1) $\rho \cdot f^4 =: A$ is constant.

(2) We have $f(t)^2 = -\varepsilon(t - t_*)^2 + 4\sqrt{\frac{2\pi}{3} A} \cdot (t - t_*)$.




## Horizons

Let $M = I \times S$ be a Robertson-Walker spacetime with distortion function $f$. Let $\gamma : [s_0, \infty) \to M$,

$$\gamma(s) = (\underbrace{\gamma^0(s)}_{\in I}, \underbrace{\hat{\gamma}(s)}_{\in S})$$

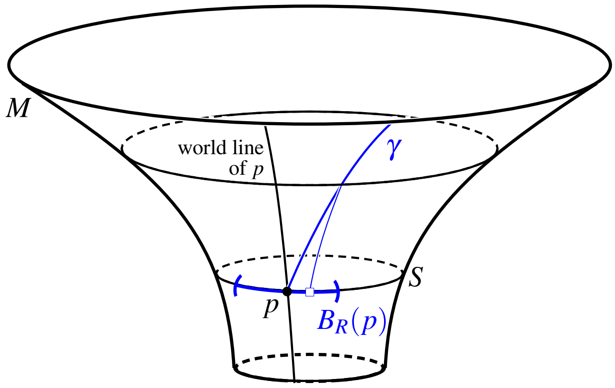
a future-directed null curve, for example a lightlike geodesic (i.e. the world line of a photon). Then

$$0 = -((\gamma^0)')^2 + f(\gamma^0)^2 \cdot \|\hat{\gamma}'\|_S^2 \quad \Longrightarrow \quad \|\hat{\gamma}'\|_S = \frac{(\gamma^0)'}{f(\gamma^0)}$$

For the length of $\hat{\gamma}$ we have

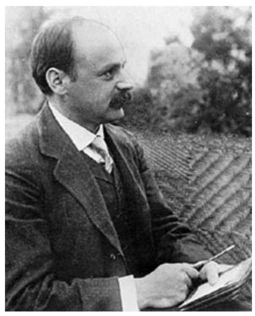
$$\mathscr{L}[\hat{\gamma}] = \int_{s_0}^{\infty} \|\hat{\gamma}'(s)\|_S \, ds = \int_{s_0}^{\infty} \frac{(\gamma^0)'(s)}{f(\gamma^0(s))} \, ds = \int_{\gamma^0(s_0)}^{\infty} \frac{d\gamma^0}{f(\gamma^0)}$$

If $f$ growth fast enough, for example $f(t) = t^2$ or $f(t) = e^t$, then $R := \mathscr{L}[\hat{\gamma}] < \infty$. This shows that photons starting in a point $p$ cannot leave the ball around $p$ of radius $R$ in $S$. This means that parts of the universe cannot be observed. This is called the *horizon problem*.




**Remark 2.3.19.** The existence of big bang singularities is not that much dependent on the particular ansatz used here but can be derived in great generality. This is the content of *singularity theorem* by Hawking, see e.g. [2, Satz 2.8.1].

## 2.4 The Schwarzschild solution



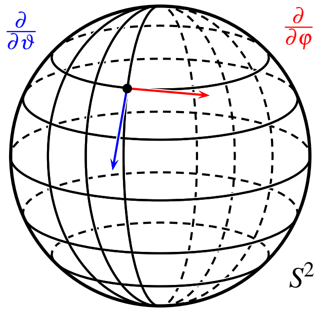
Karl Schwarzschild (1873–1916) [3]

**Goal.** We want to find a model for a vacuum spacetime outside a static, radially symmetric astronomical object.
**Ansatz.** Set $M := \mathbb{R} \times J \times S^2$, where $J \subset \mathbb{R}$, and for $t \in \mathbb{R}$, $\tilde{r} \in J$ set

$$g := -F(\tilde{r})^2 dt \otimes dt + H(\tilde{r})^2 d\tilde{r} \otimes d\tilde{r} + G(\tilde{r})^2 g_{S^2}$$

with positive smooth functions $F, G, H : J \to \mathbb{R}$. W.l.o.g. assume $H \equiv 1$, otherwise substitute $\tilde{\tilde{r}} = h(\tilde{r})$ with $h' = H$.

[3] Source: `http://en.wikipedia.org/wiki/Karl_Schwarzschild`




After introducing polar coordinates $\varphi, \vartheta$ on $S^2$, the metric takes the form

$$g_{S^2} = \sin^2\vartheta d\varphi \otimes d\varphi + d\vartheta \otimes d\vartheta,$$

hence

$$\begin{aligned} g &= -F(\tilde{r})^2 dt \otimes dt + d\tilde{r} \otimes d\tilde{r} \\ &\quad + G(\tilde{r})^2(\sin^2\vartheta d\varphi \otimes d\varphi + d\vartheta \otimes d\vartheta). \end{aligned}$$

The mapping $(t, \tilde{r}, \varphi, \vartheta) \mapsto (2t_0 - t, \tilde{r}, \varphi, \vartheta)$ is an isometry. This implies that the fixed point set $\{t_0\} \times J \times S^2 =: N_1(t_0)$ is a totally geodesic hypersurface whose unit normal field is given by

$$\nu_1 = \frac{1}{F(\tilde{r})}\frac{\partial}{\partial t}.$$

Since $N_1(t_0)$ is totally geodesic, we have $\nabla_\xi \nu_1 = 0$ for all $\xi$ tangent to $N_1(t_0)$ and thus

$$\begin{aligned} \nabla_{\frac{\partial}{\partial\tilde{r}}}\frac{\partial}{\partial t} &= \nabla_{\frac{\partial}{\partial\tilde{r}}}((F(\tilde{r})\nu_1) = F'(\tilde{r})\frac{1}{F(\tilde{r})}\frac{\partial}{\partial t} + 0 = \frac{F'(\tilde{r})}{F(\tilde{r})}\frac{\partial}{\partial t}, \\ \nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial t} &= \nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial t} = 0. \end{aligned}$$

The mapping $(t, \tilde{r}, \varphi, \vartheta) \mapsto (t, \tilde{r}, 2\varphi_0 - \varphi, \vartheta)$ is an isometry as well, so once again, its fixed point set $\mathbb{R} \times J \times \{\sigma \in S^2 | \varphi(\sigma) = \varphi_0\} =: N_2(\varphi_0)$ is a totally geodesic hypersurface. In this case, its unit normal field is given by

$$\nu_2 = \frac{1}{G(\tilde{r})\sin(\vartheta)}\frac{\partial}{\partial\varphi}.$$

Once again, for all $\xi$ tangent to $N_2(\varphi_0)$, we have $\nabla_\xi \nu_2 = 0$ and

$$\begin{aligned} \nabla_{\frac{\partial}{\partial\tilde{r}}}\frac{\partial}{\partial\varphi} &= \frac{G'(\tilde{r})}{G(\tilde{r})}\frac{\partial}{\partial\varphi}, \\ \nabla_{\frac{\partial}{\partial\vartheta}}\frac{\partial}{\partial\varphi} &= \cot(\vartheta)\frac{\partial}{\partial\varphi}. \end{aligned}$$

For the covariant derivative of $\frac{\partial}{\partial\varphi}$ in direction $\frac{\partial}{\partial\varphi}$, therefore we get

$$\begin{aligned} \left\langle \nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial\varphi}, \frac{\partial}{\partial\varphi} \right\rangle &= \frac{1}{2}\frac{\partial}{\partial\varphi}\underbrace{\left\langle \frac{\partial}{\partial\varphi}, \frac{\partial}{\partial\varphi} \right\rangle}_{=G(\tilde{r})^2\sin(\vartheta)^2} = 0, \\ \left\langle \nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial\varphi}, \frac{\partial}{\partial t} \right\rangle &= \frac{\partial}{\partial\varphi}\underbrace{\left\langle \frac{\partial}{\partial\varphi}, \frac{\partial}{\partial t} \right\rangle}_{=0} - \left\langle \frac{\partial}{\partial\varphi}, \underbrace{\nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial t}}_{=0} \right\rangle = 0, \\ \left\langle \nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial\varphi}, \frac{\partial}{\partial\tilde{r}} \right\rangle &= -\left\langle \frac{\partial}{\partial\varphi}, \nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial\tilde{r}} \right\rangle = -\left\langle \frac{\partial}{\partial\varphi}, \frac{G'}{G}\frac{\partial}{\partial\varphi} \right\rangle = -G'G\sin(\vartheta)^2, \\ \left\langle \nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial\varphi}, \frac{\partial}{\partial\vartheta} \right\rangle &= -\left\langle \frac{\partial}{\partial\varphi}, \nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial\vartheta} \right\rangle = -\cot(\vartheta)\left\langle \frac{\partial}{\partial\varphi}, \frac{\partial}{\partial\varphi} \right\rangle = -\sin(\vartheta)\cos(\vartheta)G^2. \end{aligned}$$

This shows

$$\nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial\varphi} = -G'G\sin(\vartheta)^2\frac{\partial}{\partial\tilde{r}} - \sin(\vartheta)\cos(\vartheta)\frac{\partial}{\partial\vartheta}.$$

The other covariant derivatives of the coordinate fields can be derived in a similar fashion. Collecting all derivatives we have

$$\begin{aligned}
\nabla_{\frac{\partial}{\partial\tilde{r}}}\frac{\partial}{\partial t} &= \nabla_{\frac{\partial}{\partial t}}\frac{\partial}{\partial\tilde{r}} &&= \frac{F'}{F}\frac{\partial}{\partial t}\\
\nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial t} &= \nabla_{\frac{\partial}{\partial t}}\frac{\partial}{\partial\varphi} &&= 0\\
\nabla_{\frac{\partial}{\partial\vartheta}}\frac{\partial}{\partial t} &= \nabla_{\frac{\partial}{\partial t}}\frac{\partial}{\partial\vartheta} &&= 0\\
\nabla_{\frac{\partial}{\partial\tilde{r}}}\frac{\partial}{\partial\varphi} &= \nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial\tilde{r}} &&= \frac{G'}{G}\frac{\partial}{\partial\varphi}\\
\nabla_{\frac{\partial}{\partial\vartheta}}\frac{\partial}{\partial\varphi} &= \nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial\vartheta} &&= \cot(\vartheta)\frac{\partial}{\partial\varphi}\\
&\phantom{=} \nabla_{\frac{\partial}{\partial\vartheta}}\frac{\partial}{\partial\vartheta} &&= -G'G\frac{\partial}{\partial\tilde{r}}\\
&\phantom{=} \nabla_{\frac{\partial}{\partial\varphi}}\frac{\partial}{\partial\varphi} &&= -\sin(\vartheta)^2G'G\frac{\partial}{\partial\tilde{r}} - \sin(\vartheta)\cos(\vartheta)\frac{\partial}{\partial\vartheta}\\
&\phantom{=} \nabla_{\frac{\partial}{\partial\tilde{r}}}\frac{\partial}{\partial\tilde{r}} &&= 0\\
&\phantom{=} \nabla_{\frac{\partial}{\partial t}}\frac{\partial}{\partial t} &&= F'F\frac{\partial}{\partial\tilde{r}}
\end{aligned}$$

For the Ricci curvature, we obtain

$$0 \stackrel{!}{=} \operatorname{ric}\left(\frac{\partial}{\partial t},\frac{\partial}{\partial t}\right) = F\left(F'' + 2F'\frac{G'}{G}\right) \tag{10}$$

$$0 \stackrel{!}{=} \operatorname{ric}\left(\frac{\partial}{\partial\tilde{r}},\frac{\partial}{\partial\tilde{r}}\right) = -\left(\frac{F''}{F} + 2\frac{G''}{G}\right) \tag{11}$$

$$0 \stackrel{!}{=} \operatorname{ric}\left(\frac{\partial}{\partial\varphi},\frac{\partial}{\partial\varphi}\right) = -\sin^2\vartheta\left(\frac{F'}{F}GG' + GG'' - 1 + (G')^2\right) \tag{12}$$

$$0 \stackrel{!}{=} \operatorname{ric}\left(\frac{\partial}{\partial\vartheta},\frac{\partial}{\partial\vartheta}\right) = -\left(\frac{F'}{F}GG' + GG'' - 1 + (G')^2\right) \tag{13}$$

Equation $-(11)\cdot F^2G - (10)\cdot G$ yields

$$0 = FGF'' + 2F^2G'' - FGF'' - 2F'G'F = 2F(FG'' - F'G')$$

and hence

$$\left(\frac{G'}{F}\right)' = \frac{G''F - G'F'}{F^2} = 0.$$

This means that $\frac{G'}{F} =: a$ is constant and non-zero, for otherwise $G' \equiv 0$ and also $G'' = 0$, a contradiction to (13). It follows that $G$ is strictly monotonic and we can make the parameter transformation

$$r := G(\tilde{r}).$$

Then $\frac{dr}{d\tilde{r}} = G'$ hence $dr = G'd\tilde{r}$ and (abbreviating $dx^2 := dx \otimes dx$)

$$dr^2 = (G')^2 d\tilde{r}^2 = a^2 F^2 d\tilde{r}^2.$$

We have shown that the metric has the following form (with a "new" $F$):

$$\begin{aligned} g &= -F(r)^2 dt^2 + \frac{1}{a^2 F(r)^2} dr^2 + r^2(\sin(\vartheta)^2 d\varphi^2 + d\vartheta^2) \\ &= -F(r)^2 dt^2 + \frac{1}{a^2 F(r)^2} dr^2 + r^2 g_{S^2}. \end{aligned}$$

We make the following *physical assumption*: Far from our astronomical object, the spacetime should look approximately like Minkowski space.

$$g_{\text{Mink}} = -dt^2 + dr^2 + r^2 g_{S^2}.$$

More precisely, this means $\lim\limits_{r\to\infty} F(r) = 1$ and $a^2 = 1$. Hence the metric must have the form

$$g = -F(r)^2 dt^2 + \frac{1}{F(r)^2} dr^2 + r^2(\sin(\vartheta)^2 d\varphi^2 + d\vartheta^2)$$

For the Ricci curvature, we now obtain

$$\begin{aligned} 0 &= \operatorname{ric}\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial t}\right) &&= F^2\left((F')^2 + FF'' + 2\frac{FF'}{r}\right) \qquad (14) \\ 0 &= \operatorname{ric}\left(\frac{\partial}{\partial \tilde{r}}, \frac{\partial}{\partial \tilde{r}}\right) &&= -\left(\left(\frac{F'}{F}\right)^2 + \frac{F''}{F} + 2\frac{F'}{rF}\right) \\ 0 &= \operatorname{ric}\left(\frac{\partial}{\partial \varphi}, \frac{\partial}{\partial \varphi}\right) &&= -\sin^2\vartheta\left(2FF'r - 1 + F^2\right) \\ 0 &= \operatorname{ric}\left(\frac{\partial}{\partial \vartheta}, \frac{\partial}{\partial \vartheta}\right) &&= -2FF'r + 1 - F^2 \end{aligned}$$

Hence

$$\begin{aligned} (rF^2)'' &= (F^2 + 2rFF')' = 2FF' + 2FF' + 2r(F')^2 + 2rFF'' \\ &= 2(2FF' + r(F')^2 + rFF'') \overset{(14)}{=} 0 \end{aligned}$$

Therefore $rF^2 = br - 2m$ with $b, m \in \mathbb{R}$. In other words,

$$F^2 = b - \frac{2m}{r}.$$

Taking the limit shows

$$1 = \lim_{r\to\infty} F(r)^2 = b$$

Hence $F^2 = 1 - 2m/r$ and

$$g = -(1 - \frac{2m}{r})dt^2 + \frac{1}{1-\frac{2m}{r}}dr^2 + r^2 g_{S^2}.$$

Then we indeed have $\mathrm{ric} \equiv 0$ and

$$\nabla_{\frac{\partial}{\partial t}} \frac{\partial}{\partial t} = \frac{(r-2m)m}{r^2} \frac{\partial}{\partial r}.$$

**Definition 2.4.1.** For any $m \geq 0$, the manifold $\mathbb{R} \times ((0,2m) \cup (2m,\infty)) \times S^2$ with the metric

$$g = -(1 - \frac{2m}{r})dt^2 + \frac{1}{1-\frac{2m}{r}}dr^2 + r^2 g_{S^2}$$

is called a *Schwarzschild spacetime*.

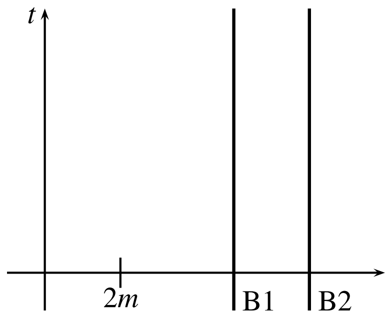
**Remark 2.4.2.** The Schwarzschild spacetime has the following properties:

| *physical formulation* | *mathematical formulation* |
|---|---|
| radially symmetric | $\mathrm{SO}(3)$ acts isometrically on $S^2$, trivially on the $r$- and $t$-axis. |
| static | $\mathbb{R}$ acts isometrically by translation on the $t$-axis. |
| vacuum solution | $\mathrm{ric} \equiv 0$ |
| asymptotic to Minkowski spacetime | $g - g_{\mathrm{Mink}} \xrightarrow{r\to\infty} 0$ |

**Definition 2.4.3.** A curve

$$\gamma : s \mapsto (t(s), r(s), \varphi(s), \vartheta(s))$$

is called a *Schwarzschild observer*, if $r \equiv r_0$, $\varphi \equiv \varphi_0$, $\vartheta \equiv \vartheta_0$, and if $\gamma$ is future directed and parametrized by proper time, i.e., $t' > 0$, and $g(\gamma', \gamma') = -1$.




For any Schwarzschild observer (with $\gamma' = t' \frac{\partial}{\partial t}$) we have

$$-1 = g(\gamma', \gamma') = (t')^2 g(\frac{\partial}{\partial t}, \frac{\partial}{\partial t}) = -(t')^2 (1 - \frac{2m}{r_0}).$$

$$\Rightarrow \quad t' = \frac{1}{\sqrt{1 - \frac{2m}{r_0}}}.$$

$$\Rightarrow \quad \gamma(s) = \Big(t_0 + \frac{s}{\sqrt{1-\frac{2m}{r_0}}}, r_0, \varphi_0, \vartheta_0\Big).$$

This parametrization by proper time shows that for a Schwarzschild observer B1 with small $r_0 > 2m$, less time elapses to traverse the same cosmic time interval (measured in the coordinate $t$) than for a distant Schwarzschild observer with big $r_0$. Hence clocks run slower when under the influence of gravitation. The *Global Positioning System (GPS)* was the first technical installation where this effect had to be taken into account.

A Schwarzschild observer is subject to the acceleration

$$\frac{\nabla}{ds}\gamma' = \nabla_{\left(\frac{1}{\sqrt{1-\frac{2m}{r_0}}}\frac{\partial}{\partial t}\right)} \left(\frac{1}{\sqrt{1 - \frac{2m}{r_0}}} \frac{\partial}{\partial t}\right) = \frac{1}{1 - \frac{2m}{r_0}} \frac{(r_0 - 2m)m}{{r_0}^3} \frac{\partial}{\partial r} = \frac{m}{{r_0}^2} \frac{\partial}{\partial r}.$$

This acceleration has the absolute value

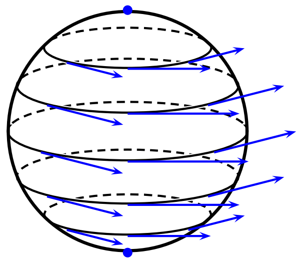
$$\frac{m}{{r_0}^2} \frac{1}{\sqrt{1 - \frac{2m}{r_0}}} \overset{r_0 \to \infty}{\sim} \frac{m}{{r_0}^2},$$

which approximates that of a central star of mass $m$, see Section 2.1. Hence $m$ is interpreted as the mass of the astronomical object.

**Definition 2.4.4.** Let $M$ be a semi-Riemannian manifold and let $\Phi : (-\varepsilon, \varepsilon) \to \mathrm{Isom}(M)$ be such that $\Phi(0) = \mathrm{id}_M$ and assume that $(-\varepsilon, \varepsilon) \times M \to M$ defined by $(s, p) \mapsto \Phi(s)(p)$ is smooth. Then the vector field $\xi$, defined by

$$\xi|_p := \frac{d}{ds} \Phi(s)(p)|_{s=0},$$

is called a *Killing vector field.*



**Example 2.4.5.** Let $M = S^2$ and

$$\Phi(s) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(s) & -\sin(s) \\ 0 & \sin(s) & \cos(s) \end{pmatrix}.$$

The corresponding Killing vector field is $\dfrac{\partial}{\partial \varphi}$ in polar coordinates.

**Lemma 2.4.6**
*Let $M$ be a semi-Riemannian manifold, let $\xi$ be a Killing vector field on $M$ and let $\gamma$ be a geodesic in $M$. Then the function*

$$t \mapsto g(\gamma'(t), \xi|_{\gamma(t)})$$

*is constant.*

*Sketch of proof.* One can check that Killing vector fields $\xi$ satisfy

$$\langle \nabla_X \xi, X \rangle = 0$$

for all tangent vectors $X$. We then compute

$$\frac{d}{dt}\langle \gamma'(t), \xi(\gamma(t)) \rangle = \langle \frac{\nabla}{dt}\gamma'(t), \xi(\gamma(t)) \rangle + \langle \gamma'(t), \nabla_{\gamma'(t)}\xi(\gamma(t)) \rangle = 0 + 0 = 0. \qquad \square$$

This lemma may be regarded as a version of Noether's theorem; infinitesimal symmetries (Killing vector fields) give rise to conservation laws.
In the Schwarzschild model $M$, $\frac{\partial}{\partial t}$ is a Killing vector field because $M$ is static and $\frac{\partial}{\partial \varphi}$ is a Killing vector field because $M$ is radially symmetric. Lemma 2.4.6 implies that for geodesics

$$\gamma(s) = (t(s), r(s), \varphi(s), \vartheta(s))$$

with $\vartheta \equiv \frac{\pi}{2}$ (i.e. in particular $\gamma' = t' \frac{\partial}{\partial t} + r' \frac{\partial}{\partial r} + \varphi' \frac{\partial}{\partial \varphi}$),

the *energy* $E := \left\langle \gamma', \dfrac{\partial}{\partial t} \right\rangle = -t'h$ and the *angular momentum* $L := \left\langle \gamma', \dfrac{\partial}{\partial \varphi} \right\rangle = \varphi' r^2$

are constant. Here, $h(r) := 1 - 2m/r$. This means that for light, we have

$$0 = \langle \gamma', \gamma' \rangle = -(t')^2 \cdot h + \frac{(r')^2}{h} + r^2(\varphi')^2 = t' \cdot E + \frac{(r')^2}{h} + \varphi' \cdot L.$$

This implies the *energy equation for light particles*

$$E^2 = (r')^2 + \varphi' \cdot Lh = (r')^2 + \frac{L^2}{r^2} h.$$

For massive particles, we obtain

$$-1 = \langle \gamma', \gamma' \rangle = t'E + \frac{(r')^2}{h} + \varphi'.$$

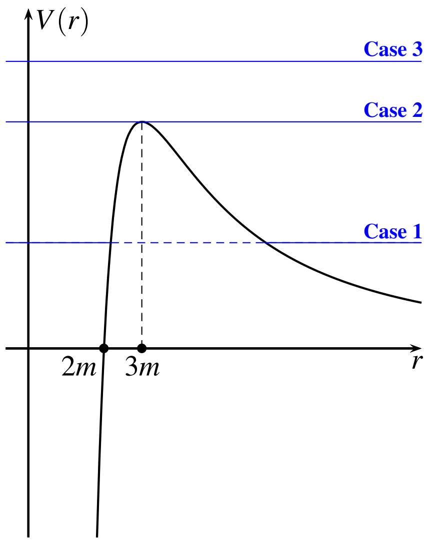
This implies the *energy equation for massive particles*

$$E^2 = (r')^2 + \varphi' \cdot Lh + h = (r')^2 + \left( \frac{L^2}{r^2} + 1 \right) h.$$

### Trajectories of Light Particles ($L \neq 0$)

Set

$$V(r) := \frac{L^2}{r^2} h(r) = \frac{L^2}{r^2} \left( 1 - \frac{2m}{r} \right).$$




We have $V(2m) = 0$, $\lim_{r \to \infty} V(r) = 0$ and $\lim_{r \to 0} V(r) = -\infty$. We determine the extrema

$$\begin{aligned} 0 &\stackrel{!}{=} V'(r) \\ &= -2\frac{L^2}{r^3} \left( 1 - \frac{2m}{r} \right) + \frac{L^2}{r^2} \frac{2m}{r^2} \\ &= \frac{L^2}{r^4} (-2r + 4m + 2m) \\ &= \frac{2L^2}{r^4} (-r + 3m) \end{aligned}$$

This implies that the only extremum is at $r = 3m$. Because of the behavior of $V$ for large $r$, $r = 3m$ must be a maximum with $V(3m) = \frac{L^2}{27m^2}$.

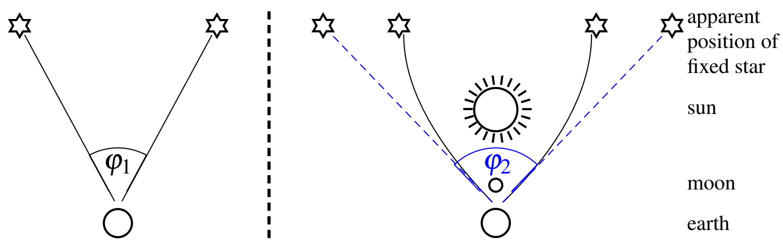
The energy equation takes the form $E^2 = (r')^2 + V(r)$. In particular, $V(r) \leq E^2$.

*Case 1*: $E^2 < \dfrac{L^2}{27m^2}$.

(a) $r_0 < 3m$: collision-collision orbit.

(b) $r_0 > 3m$: fly-by orbit.

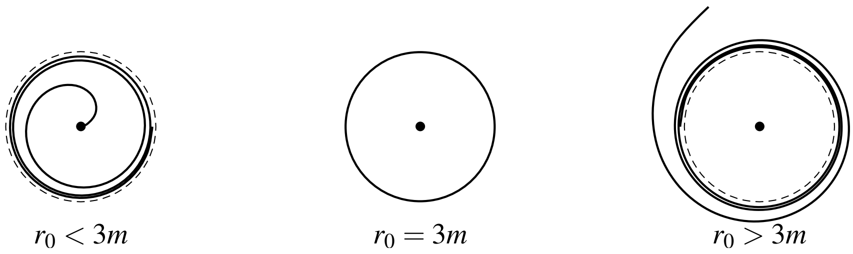
*light deflection* (First experimental verification of general relativity):



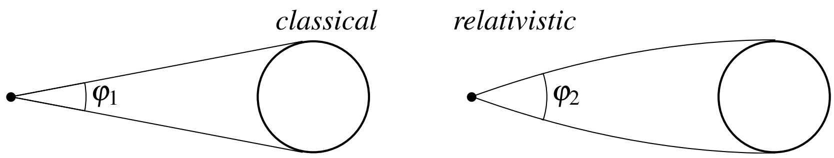
*Case 2*: $E^2 = \dfrac{L^2}{27m^2}$.

(a) $r \equiv 3m$: Exceptional orbit.

(b) $r_0 < 3m$ or $r_0 > 3m$: spiral orbits.




Case 3: $E^2 > \frac{L^2}{27m^2}$: Collision-escape orbit.
**Sight angle.** We have $\varphi_2 > \varphi_1$, i.e. astronomical objects seem bigger than they are.




### Orbits of Massive Particles

Now set

$$V(r) := \left(\frac{L^2}{r^2} + 1\right) h(r).$$

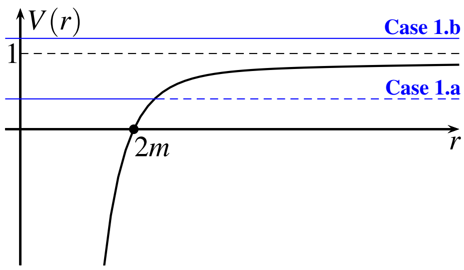
We have $V(2m) = 0$, $\lim_{r\to\infty} V(r) = 1$ and $\lim_{r\to 0} V(r) = -\infty$. The local extrema are at

$$r_{1,2} = \frac{L^2}{2m^2} \pm L\sqrt{\frac{L^2}{4m^2} - 3}.$$

*Case 1*: $L^2 < 12m^2$, i.e. there are no local extrema

(a) $E^2 < 1$: collision-collision orbit.

(b) $E^2 \geq 1$: collision-escape orbit.



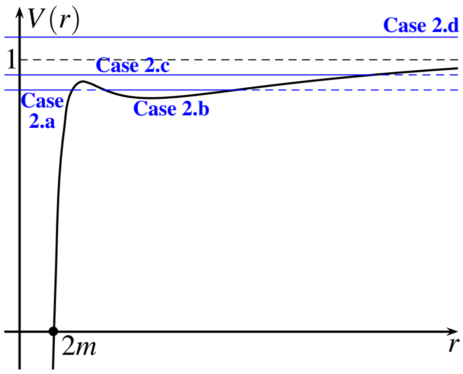
*Case 2*: $12m^2 \leq L^2 < 16m^2$.

(a) $E^2 < V(r_1)$ and $r_0 < r_1$: collision-collision orbit.

(b) $E^2 < V(r_1)$ and $r_0 > r_1$: bounded orbit.

(c) $V(r_1) < E^2 < 1$: collision-collision orbit.

(d) $E^2 \geq 1$: collision-escape orbit.



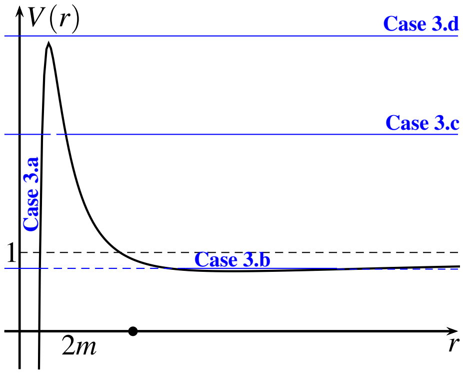

*Case 3*: $L^2 \geq 16m^2$.

(a) $E^2 < V(r_1)$ and $r_0 < r_1$: collision-collision orbit.

(b) $V(r_2) < E^2 < 1$ and $r_0 > r_1$: bounded orbit.

(c) $1 \leq E^2 < V(r_1)$ and $r_0 > r_1$: fly-by orbit.

(d) $E^2 > V(r_1)$: collision- escape orbit.




**Definition 2.4.7.** The constant $2m$ is called the *Schwarzschild radius*.

Now we want to remove the singularity at $r = 2m$. Set

$$f(r) := (r-2m)e^{r/2m-1}.$$

Then $f : (0,\infty) \to (-\frac{2m}{e}, \infty)$ is a diffeomorphism, because

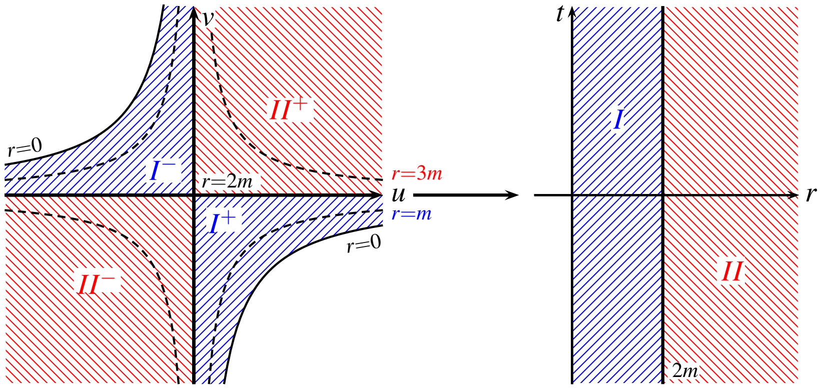
$$f'(r) = e^{\frac{r}{2m}-1} + (r-2m)\frac{1}{2m}e^{\frac{r}{2m}-1} = \frac{r}{2m}e^{\frac{r}{2m}-1} > 0.$$

We have $f((0,2m)) = (-\frac{2m}{e}, 0)$ and $f((2m,\infty)) = (0,\infty)$. We introduce the *Kruskal coordinates* $u$ and $v$ by

$$f(r) = uv, \quad t = 2m \cdot \ln\left(\left|\frac{v}{u}\right|\right).$$

For the metric, we obtain

$$g = \frac{4m^2}{r}e^{1-\frac{r}{2m}}(du \otimes dv + dv \otimes du) + r^2(\sin(\vartheta)^2 d\varphi^2 + d\vartheta^2).$$




The metric is also smooth at $v = 0$ corresponding to $r = 2m$. On the area $II$, the time orientation is characterized by $t' > 0$. We have

$$0 < t' = 2m\frac{\frac{v'u-u'v}{u^2}}{\frac{v}{u}} = \frac{v'u-u'v}{uv} \qquad \Longleftrightarrow \qquad v'u-u'v > 0.$$

This implies that $-\frac{\partial}{\partial u}$ and $\frac{\partial}{\partial v}$ are positive and time oriented on $II^+$, and because of continuity, this holds on $I^+$ as well. Converting to $I$ then gives

$$\text{positive time oriented} \qquad \Longleftrightarrow \qquad r' > 0.$$

In particular, this means that no future oriented causal curve can leave the area $I$. Not even light can leave the region $r < 2m$. Therefore, astronomical objects with radius smaller than the Schwarzschild radius are called *black holes*.
The singularity at $r = 0$ on the other hand cannot be removed for $m > 0$. This leads to the Penrose singularity theorem.

# Bibliography


[1] C. BÄR: *Elementargeometrie*, Lecture Notes, University of Potsdam, Potsdam 2008, `http://geometrie.math.uni-potsdam.de/index.php/de/lehre2/lehrmaterialien`

[2] C. BÄR: *Lorentzgeometrie*, Lecture Notes, University of Potsdam, Potsdam 2006, `http://geometrie.math.uni-potsdam.de/index.php/de/lehre2/lehrmaterialien`

[3] A. EINSTEIN: *Zur Elektrodynamik bewegter Körper*, Ann. der Phys. (4) 17, 891–921 (1905)

[4] C. W. MISNER, K. S. THORNE AND J. A. WHEELER: *Gravitation*, W. H. Freeman and Company, New York 1973

[5] H. STEPHANI: *Relativity. An Introduction to Special and General Relativity*, 3rd Edition, Cambridge University Press, Cambridge 2004

[6] R. M. WALD: *General relativity*, The University of Chicago Press, Chicago-London (1984)

# Index