Physik-Institut der Universität Zürich
in conjunction with ETH Zürich

General Relativity

Autumn semester 2016

Prof. Philippe Jetzer

Original version by Arnaud Borde

Revision: Antoine Klein, Raymond Angélil, Cédric Huwyler, Simone Balmelli, Yannick Boetzel

Last revision of this version: January 4, 2017

Sources of inspiration for this course include

- S. Carroll, Spacetime and Geometry, Pearson, 2003
- S. Weinberg, Gravitation and Cosmology, Wiley, 1972
- N. Straumann, General Relativity with applications to Astrophysics, Springer Verlag, 2004
- C. Misner, K. Thorne and J. Wheeler, Gravitation, Freeman, 1973
- R. Wald, General Relativity, Chicago University Press, 1984
- T. Fliessbach, Allgemeine Relativitätstheorie, Spektrum Verlag, 1995
- B. Schutz, A first course in General Relativity, Cambridge, 1985
- R. Sachs and H. Wu, General Relativity for mathematicians, Springer Verlag, 1977
- J. Hartle, Gravity, An introduction to Einstein's General Relativity, Addison Wesley, 2002
- H. Stephani, General Relativity, Cambridge University Press, 1990, and
- M. Maggiore, Gravitational Waves: Volume 1: Theory and Experiments, Oxford University Press, 2007.
- A. Zee, Einstein Gravity in a Nutshell, Princeton University Press, 2013

As well as the lecture notes of

- Sean Carroll (http://arxiv.org/abs/gr-qc/9712019)
- Matthias Blau (http://www.blau.itp.unibe.ch/Lecturenotes.html)
- Gian Michele Graf

Contents

Part I

Introduction

1 Newton's theory of gravitation

In his book *Principia* in 1687, Isaac Newton laid the foundations of classical mechanics and made a first step in unifying the laws of physics.

The trajectories of N point masses, attracted to each other via gravity, are the solutions to the equation of motion

$$m_i \frac{\mathrm{d}^2\vec{r}_i}{\mathrm{d}t^2} = -G \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{m_i m_j (\vec{r}_i - \vec{r}_j)}{|\vec{r}_i - \vec{r}_j|^3} \qquad i = 1 \ldots N, \tag{1.1}$$

with $\vec{r}_i(t)$ being the position of point mass m_i at time t. Newton's constant of gravitation is determined experimentally to be

$$G = 6.6743 \pm 0.0007 \times 10^{-11}\ \mathrm{m}^3\,\mathrm{kg}^{-1}\,\mathrm{s}^{-2} \tag{1.2}$$

The scalar gravitational potential $\phi(\vec{r})$ is given by

$$\phi(\vec{r}) = -G \sum_{j=1}^{N} \frac{m_j}{|\vec{r} - \vec{r}_j|} = -G \int \mathrm{d}^3r' \frac{\rho(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|}, \tag{1.3}$$

where it has been assumed that the mass is smeared out in a small volume d^3r. The mass is given by $\mathrm{d}m = \rho(\vec{r}\,')\mathrm{d}^3r'$, $\rho(\vec{r}\,')$ being the mass density. For point-like particles we have $\rho(\vec{r}\,') \sim m_j\delta^{(3)}(\vec{r}\,' - \vec{r}_j)$. The gradient of the gravitational potential can then be used to produce the equation of motion:

$$m \frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = -m\nabla\phi(\vec{r}). \tag{1.4}$$

According to (1.3), the field $\phi(\vec{r})$ is determined through the mass of the other particles. The corresponding field equation derived from (1.3) is given by[1]

$$\Delta\phi(\vec{r}) = 4\pi G\rho(\vec{r}) \tag{1.5}$$

The so called *Poisson equation* (1.5) is a linear partial differential equation of 2^{nd} order. The source of the field is the mass density. Equations (1.4) and (1.5) show the same structure as the field equation of electrostatics:

$$\Delta\phi_e(\vec{r}) = -4\pi\rho_e(\vec{r}), \tag{1.6}$$

and the non-relativistic equation of motion for charged particles

$$m \frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = -q\nabla\phi_e(\vec{r}). \tag{1.7}$$

Here, ρ_e is the charge density, ϕ_e is the electrostatic potential and q represents the charge which acts as coupling constant in (1.7). m and q are independent characteristics of the considered body. In

[1] $\Delta \frac{1}{|\vec{r}-\vec{r}\,'|} = -4\pi\delta^{(3)}(\vec{r} - \vec{r}\,')$

analogy one could consider the “gravitational mass” (right side) as a charge, not to be confused with the “inertial mass” (left side). Experimentally, one finds to very high accuracy ($\sim 10^{-13}$) that they are the same. As a consequence, *all bodies fall at a rate independent of their mass* (Galileo Galilei). This appears to be just a chance in Newton’s theory, whereas in GR it will be an important starting point.

For many applications, (1.4) and (1.5) are good enough. It must however be clear that these equations cannot be always valid. In particular (1.5) implies an instantaneous action at a distance, what is in contradiction with the predictions of special relativity. We therefore have to suspect that Newton’s theory of gravitation is only a special case of a more general theory.

2 Goals of general relativity

In order to get rid of instantaneous interactions, we can try to perform a relativistic generalization of Newton’s theory (eqs. (1.4) and (1.5)), similar to the transition from electrostatics (eqs. (1.6) and (1.7)) to electrodynamics.

The Laplace operator Δ is completed such as to get the D’Alembert operator (wave equation)

$$\Delta \Rightarrow \Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta \tag{2.1}$$

Changes in ρ_e travel with the speed of light to another point in space. If we consider inertial coordinate frames in relative motion to each other it is clear that the charge density has to be related to a current density. In other words, charge density and current density transform into each other. In electrodynamics we use the current density j^α ($\alpha = 0, 1, 2, 3$):

$$\rho_e \to (\rho_e c, \rho_e v^i) = j^\alpha, \tag{2.2}$$

where the v^i are the cartesian components of the velocity $\vec{v}$ ($i = 1, 2, 3$). An analogous generalization can be performed for the potential:

$$\phi_e \to (\phi_e, A^i) = A^\alpha. \tag{2.3}$$

The relativistic field equation is then

$$\Delta\phi_e = -4\pi\rho_e \to \Box A^\alpha = \frac{4\pi}{c} j^\alpha. \tag{2.4}$$

In the static case, the 0-component reduces to the equation on the left.

Equation (2.4) is equivalent to Maxwell’s equations (in addition one has to choose a suitable gauge condition). Since electrostatics and Newton’s theory have the same mathematical structure, one may want to generalize it the same way. So in (1.5) one could introduce the change $\Delta \to \Box$. Similarly one generalizes the mass density. But there are differences with electrodynamics. The first difference is that the charge q of a particle is independent on how the particle moves; this is not the case for the mass: $m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$.

As an example, consider a hydrogen atom with a proton (rest mass m_p, charge $+e$) and an electron (rest mass m_e, charge $-e$). Both have a finite velocity within the atom. The total charge of the atom is $q = q_e + q_p = 0$, but for the total mass we get $m_H \neq m_p + m_e$ (binding energy). Formally this means that charge is a Lorentz scalar (does not depend on the frame in which the measurement is performed). Therefore we can assign a charge to an elementary particle, and not only a "charge at rest", whereas for the mass we must specify the rest mass.

Since charge is a Lorentz scalar, the charge density ($\rho_e = \frac{\delta q}{\delta V}$) transforms like the 0-component of a Lorentz vector ($\frac{1}{\delta V}$ gets a factor $\gamma = \frac{1}{\sqrt{1-v^2/c^2}}$ due to length contraction). The mass density ($\rho = \frac{\delta m}{\delta V}$) transforms instead like the 00-component of a Lorentz tensor, which we denote as the energy-momentum tensor $T_{\alpha\beta}$. This follows from the fact that the energy (mass is energy $E = mc^2$) is the 0-component of a 4-vector (energy-momentum vector p^α) and transforms as such. Thus, instead of (2.2), we shall have

$$\rho \Rightarrow \begin{pmatrix} \rho c^2 & \rho c v^i \\ \rho c v^i & \rho v^i v^j \end{pmatrix} \sim T^{\alpha\beta} \qquad i, j = 1, 2, 3 \tag{2.5}$$

This implies that we have to generalize the gravitational potential ϕ to a quantity depending on 2 indices which we shall call the metric tensor $g^{\alpha\beta}$. Hence we get

$$\Delta\phi = 4\pi G\rho \Rightarrow \Box g^{\alpha\beta} \sim GT^{\alpha\beta}. \tag{2.6}$$

In GR one finds (2.6) for a weak gravitational field (linearized case), e.g. used for the description of *gravitational waves.*

Due to the equivalence between mass and energy, the energy carried by the gravitational field is also mass and thus also a source of the gravitational field itself. This leads to non-linearities. One can note that photons do not have a charge and thus Maxwell's equations can be linear.
To summarize:

1. GR is the relativistic generalization of Newton's theory. Several similarities between GR and electrodynamics exist.

2. GR requires tensorial equations (rather than vectorial as in electrodynamics).

3. There are non-linearities which will lead to non-linear field equations.

Part II

Special Relativity

3 Lorentz transformations

A reference system with a well defined choice of coordinates is called a *coordinate system. Inertial* reference systems (IS) are (from a "practical" point of view) systems which move with constant speed with respect to distant (thus fixed) stars in the sky. Newton's equations of motion are valid in IS. Non-IS are reference systems which are accelerated with respect to an IS. In this chapter we will establish how to transform coordinates between different inertial systems.

3.1 Galilean invariance

Galilei stated that "all IS are equivalent", i.e. all physical laws are valid in any IS: the physical laws are *covariant* under transformations from an IS to another IS'. Covariant means here form invariant. The equations should have the same form in all IS.

With the coordinates x^i $(i = 1, 2, 3)$ and t, an event in an IS can be defined. In another IS', the same event has different coordinates x'^i and t'. A general Galilean transformation can then be written as:

$$x'^i = \alpha^i{}_k x^k + v^i t + a^i, \tag{3.1}$$

$$t' = t + \tau, \tag{3.2}$$

where

- x^i, v^i and a^i are cartesian components of vectors
- $\vec{v} = v^i \vec{e}_i$ where $\vec{e}_i$ is a unit vector
- we use the summation rule over repeated indices: $\alpha^i{}_k x^k = \sum_k \alpha^i{}_k x^k$
- latin indices run on 1,2,3
- greek indices run on 0,1,2,3
- $\vec{v}$ is the relative velocity between IS and IS'
- $\vec{a}$ is a constant vector (translation)
- $\alpha^i{}_k$ is the relative rotation of coordinates systems, $\alpha = (\alpha^i{}_k)$ is defined by

$$\alpha^i{}_n (\alpha^T)^n_k = \delta^i{}_k \quad \text{or} \quad \alpha\alpha^T = \mathbb{I} \quad \text{i.e.} \quad \alpha^{-1} = \alpha^T \tag{3.3}$$

The condition $\alpha\alpha^T = \mathbb{I}$ ensures that the line element

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \tag{3.4}$$

remains invariant. α can be defined by giving 3 Euler angles. Eqs. (3.1) and (3.2) define a 10 ($a = 3$, $v = 3$, $\tau = 1$ and $\alpha = 3$) parametric group of transformations, the so-called *Galilean group*.

The laws of mechanics are left invariant under transformations (3.1) and (3.2). But Maxwell's equations are not invariant under Galilean transformations, since they contain the speed of light c. This led Einstein to formulate a new relativity principle (special relativity, SR): *All physical laws, including Maxwell's equations, are valid in any inertial system.* This leads us to Lorentz transformations (instead of Galilean), thus the law of mechanics have to be modified.

3.2 Lorentz transformations

We start by introducing 4-dimensional vectors, glueing time and space together to a *spacetime*. The *Minkowski* coordinates are defined by

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z. \tag{3.5}$$

x^α is a vector in a 4-dimension space (or 4-vector). An event is given by x^α in an IS and by x'^α in an IS'. Homogeneity of space and time imply that the transformation from x^α to x'^α has to be linear:

$$x'^\alpha = \Lambda^\alpha{}_\beta x^\beta + a^\alpha, \tag{3.6}$$

where a^α is a space and time translation. The relative rotations and boosts are described by the 4×4 matrix Λ. Linear means in this context that the coefficients $\Lambda^\alpha{}_\beta$ and a^α do not depend on x^α. In order to preserve the speed of light appearing in Maxwell's equations as a constant, the Λ^α_β have to be such that the square of the line element

$$\mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta = c^2\mathrm{d}t^2 - \mathrm{d}\vec{r}^2 \tag{3.7}$$

remains unchanged, with the *Minkowski metric*

$$\eta_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{3.8}$$

Because of $\mathrm{d}s^2 = \mathrm{d}s'^2 \Leftrightarrow c^2\mathrm{d}\tau^2 = c^2\mathrm{d}\tau'^2$, the proper time is an invariant under Lorentz transformations. Indeed for light $\mathrm{d}\tau^2 = \mathrm{d}t^2 - \frac{\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2}{c^2} = 0$. Thus, $c^2 = \left|\frac{\mathrm{d}\vec{x}}{\mathrm{d}t}\right|^2$ and $c = \left|\frac{\mathrm{d}\vec{x}}{\mathrm{d}t}\right|$. Applying a Lorentz transformation results in $c = \left|\frac{\mathrm{d}\vec{x}'}{\mathrm{d}t'}\right|$. This has the important consequence that the speed of light c is the same in all coordinate systems (what we intended by the definition of (3.7)).

A 4-dimensionial space together with this metric is called a *Minkowski space*. Inserting (3.6) into the invariant condition $\mathrm{d}s^2 = \mathrm{d}s'^2$ gives

$$\begin{aligned} \mathrm{d}s'^2 &= \eta_{\alpha\beta}\mathrm{d}x'^\alpha \mathrm{d}x'^\beta \\ &= \eta_{\alpha\beta}\Lambda^\alpha{}_\gamma \mathrm{d}x^\gamma \Lambda^\beta{}_\delta \mathrm{d}x^\delta \\ &= \eta_{\gamma\delta}\mathrm{d}x^\gamma \mathrm{d}x^\delta = \mathrm{d}s^2. \end{aligned} \tag{3.9}$$

Then we get

$$\Lambda^\alpha{}_\gamma \Lambda^\beta{}_\delta \eta_{\alpha\beta} = \eta_{\gamma\delta} \quad \text{or} \quad \Lambda^T \eta \Lambda = \eta. \tag{3.10}$$

Rotations are special subcases incorporated in Λ: $x'^\alpha = \Lambda^\alpha{}_\beta x^\beta$ with $\Lambda^i{}_k = \alpha^i{}_k$, and $\Lambda^0{}_0 = 1$, $\Lambda^i{}_0 = \Lambda^0{}_i = 0$. The entire group of Lorentz transformations (LT) is the so called *Poincaré* group (and has 10 parameters). The case $a^\alpha \neq 0$ corresponds to the Poincaré group or *inhomogeneous Lorentz group*, while the subcase $a^\alpha = 0$ can be described by the *homogeneous Lorentz group*. Translations and rotations are subgroups of Galilean and Lorentz groups.

Consider now a Lorentz 'boost' in the direction of the x-axis: $x'^2 = x^2$, $x'^3 = x^3$. v denotes the relative velocity difference between IS and the boosted IS'. Then

$$\Lambda^\alpha{}_\beta = \begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & 0 & 0 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.11}$$

Evaluating eq. (3.10):

$$(\gamma, \delta) = (0,0) \qquad (\Lambda^0{}_0)^2 - (\Lambda^1{}_0)^2 = 1 \tag{3.12a}$$

$$= (1,1) \qquad (\Lambda^0{}_1)^2 - (\Lambda^1{}_1)^2 = -1 \tag{3.12b}$$

$$= (0,1) \text{ or } (1,0) \qquad \Lambda^0{}_0\Lambda^0{}_1 - \Lambda^1{}_0\Lambda^1{}_1 = 0 \tag{3.12c}$$

The solution to this system is

$$\begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 \\ \Lambda^1{}_0 & \Lambda^1{}_1 \end{pmatrix} = \begin{pmatrix} \cosh\Psi & -\sinh\Psi \\ -\sinh\Psi & \cosh\Psi \end{pmatrix}. \tag{3.13}$$

For the origin of IS' we have $x'^1 = 0 = \Lambda^1{}_0 ct + \Lambda^1{}_1 vt$. This way we find

$$\tanh\Psi = -\frac{\Lambda^1{}_0}{\Lambda^0{}_0} = \frac{v}{c}, \tag{3.14}$$

and as a function of velocity:

$$\Lambda^0{}_0 = \Lambda^1{}_1 = \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}, \tag{3.15a}$$

$$\Lambda^0{}_1 = \Lambda^1{}_0 = \frac{-v/c}{\sqrt{1-\frac{v^2}{c^2}}}. \tag{3.15b}$$

A Lorentz transformation (called a *boost*) along the x-axis can then be written explicitly as

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{3.16a}$$

$$y' = y, \tag{3.16b}$$

$$z' = z, \tag{3.16c}$$

$$ct' = \frac{ct - x\frac{v}{c}}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{3.16d}$$

which is valid only for $|v| < c$. For $|v| \ll c$, (3.16) recovers the special (no rotation) Galilean transformation $x' = x - vt$, $y' = y$, $z' = z$ and $t' = t$. The parameter

$$\Psi = \operatorname{arctanh} \frac{v}{c} \tag{3.17}$$

is called the *rapidity*. From this we find for the addition of parallel velocities:

$$\Psi = \Psi_1 + \Psi_2$$

$$\Rightarrow v = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}} \tag{3.18}$$

3.3 Proper time

The time coordinate t in IS is the time shown by clocks at rest in IS. We next determine the *proper time* τ shown by a clock which moves with velocity $\vec{v}(t)$. Consider a given moment t_0 an IS', which moves with respect to IS with a constant velocity $\vec{v}_0(t_0)$. During an infinitesimal time interval $\mathrm{d}t'$ the clock can be considered at rest in IS', thus:

$$\mathrm{d}\tau = \mathrm{d}t' = \sqrt{1 - \frac{v_0^2}{c^2}}\,\mathrm{d}t. \tag{3.19}$$

Indeed (3.16) with $x = v_0 t$ gives $t' = \frac{t(1-v_0^2/c^2)}{\sqrt{1-\frac{v_0^2}{c^2}}} = t\sqrt{1 - \frac{v_0^2}{c^2}}$ and thus (3.19).

At the next time $t_0 + \mathrm{d}t$, we consider an IS" with velocity $\vec{v}_0 = \vec{v}(t_0 + \mathrm{d}t)$ and so on. Summing up all infinitesimal proper times $\mathrm{d}\tau$ gives the proper time interval:

$$\tau = \int_{t_1}^{t_2} \mathrm{d}t \sqrt{1 - \frac{v^2(t)}{c^2}} \tag{3.20}$$

This is the time interval measured by an observer moving at a speed $v\,(t)$ between t_1 and t_2 (as given by a clock at rest in IS). This effect is called *time dilation*.

4 Relativistic mechanics

Let us now perform the relativistic generalization of Newton's equation of motion for a point particle.

4.1 Equations of motion

The velocity $\vec{v}$ can be generalized to a 4-velocity vector u^α:

$$v^i = \frac{\mathrm{d}x^i}{\mathrm{d}t} \to u^\alpha = \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau} \tag{4.1}$$

Since $\mathrm{d}\tau = \frac{\mathrm{d}s}{c}$, $\mathrm{d}\tau$ is invariant. With $\mathrm{d}x'^\alpha = \Lambda^\alpha_\beta \mathrm{d}x^\beta$ it follows that u^α transforms like $\mathrm{d}x^\alpha$:

$$u'^\alpha = \Lambda^\alpha{}_\beta u^\beta \tag{4.2}$$

All quantities which transforms this way are Lorentz vectors or form-vectors. The generalized equation of motion is then

$$m\frac{\mathrm{d}u^\alpha}{\mathrm{d}\tau} = f^\alpha. \tag{4.3}$$

Both $\frac{\mathrm{d}u^\alpha}{\mathrm{d}\tau}$ and f^α are Lorentz vectors, therefore, (4.3) is a Lorentz vector equation: if we perform a Lorentz transformation, we get $m\frac{\mathrm{d}u'^\alpha}{\mathrm{d}\tau} = f'^\alpha$. Eq. (4.3) is covariant under Lorentz transformations and for $v \ll c$ it reduces to Newton's equations. (left hand side becomes $m\left(0, \frac{\mathrm{d}\vec{v}}{\mathrm{d}t}\right)$ and the right hand side $\left(f^0, \vec{f}\right) = \left(0, \vec{K}\right)$). The *Minkowski force* f'^α is determined in any IS through a corresponding LT: $f'^\alpha = \Lambda^\alpha{}_\beta f^\beta$. For example $\vec{v} = -v\vec{e}_1$ with $\gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2}$, leads to $f'^0 = \frac{\gamma v K^1}{c}$, $f'^1 = \gamma K^1$, $f'^2 = K^2$ and $f'^3 = K^3$. For a general direction of velocity $(-\vec{v})$ we get:

$$f'^0 = \gamma\frac{\vec{v}\cdot\vec{K}}{c}, \qquad \vec{f'} = \vec{K} + (\gamma - 1)\,\vec{v}\,\frac{\vec{v}\cdot\vec{K}}{v^2}. \tag{4.4}$$

4.2 Energy and momentum

The 4-momentum $p^\alpha = mu^\alpha = m\frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}$ is a Lorentz vector. With (3.19) we get

$$p^\alpha = \left(\frac{mc}{\sqrt{1-\frac{v^2}{c^2}}}, \frac{mv^i}{\sqrt{1-\frac{v^2}{c^2}}}\right) = \left(\frac{E}{c}, \vec{p}\right). \tag{4.5}$$

This yields the relativistic

energy : $$E = \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} = \gamma mc^2 \tag{4.6a}$$

momentum : $$\vec{p} = \frac{m\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}} = \gamma m\vec{v} \xrightarrow{v\ll c} \vec{p} = m\vec{v}. \tag{4.6b}$$

With (4.4), the 0-component of (4.3) becomes (in the case $v \ll c$)

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \underbrace{\vec{v}.\vec{K}}_{\substack{\text{power given}\\ \text{to the particle}}}. \tag{4.7}$$

This justifies to call the quantity $E = \gamma mc^2$ an *energy*. From $\mathrm{d}s^2 = c^2\mathrm{d}\tau^2 = \eta_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta$ it follows $\eta_{\alpha\beta}p^\alpha p^\beta = m^2c^2$ and thus

$$E^2 = m^2c^4 + c^2\vec{p}^2, \tag{4.8}$$

the relativistic energy-momentum relation. The limit cases are

$$E = \sqrt{m^2c^4 + c^2\vec{p}^2} \approx \begin{cases} mc^2 + \frac{p^2}{2m} & v \ll c \text{ or } p \ll mc^2 \\ cp & v \sim c \text{ or } p \gg mc^2 \end{cases} \tag{4.9}$$

with $p = |\vec{p}|$. For particles with no rest mass (photons): $E = cp$ (exact relation).

4.3 Equivalence between mass and energy

One can divide the energy into the rest energy

$$E_0 = mc^2 \tag{4.10}$$

and the kinetic energy $E_{kin} = E - E_0 = E - mc^2$. The quantities defined in (4.6) are conserved when more particles are involved. Due to the equivalence between energy and mass, the mass or the mass density becomes a source of the gravitational field.

5 Tensors in Minkowski space

Let us discuss the transformation properties of physical quantities under a Lorentz transformation. We have already seen how a 4-vector is transformed:

$$V^\alpha \to V'^\alpha = \Lambda^\alpha{}_\beta V^\beta. \tag{5.1}$$

This is a so-called *contravariant* 4-vector (indices are up). The coordinate system transforms according to $X^\alpha \to X'^\alpha = \Lambda^\alpha{}_\beta X^\beta$. A *covariant* 4-vector is defined through

$$V_\alpha = \eta_{\alpha\beta}V^\beta. \tag{5.2}$$

Let us now define the matrix $\eta^{\alpha\beta}$ as the inverse matrix to $\eta_{\alpha\beta}$:

$$\eta^{\alpha\beta}\eta_{\beta\gamma} = \delta^\alpha_\gamma. \tag{5.3}$$

In our case

$$\eta^{\alpha\beta} = \eta_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{5.4}$$

With (5.3) we can express (5.2) equally as

$$V^\alpha = \eta^{\alpha\beta}V_\beta. \tag{5.5}$$

The transformation of a covariant vector is then given by

$$V'_\alpha = \eta_{\alpha\beta} V'^\beta = \eta_{\alpha\beta} \Lambda^\beta{}_\gamma V^\gamma = \eta_{\alpha\beta} \Lambda^\beta{}_\gamma \eta^{\gamma\delta} V_\delta = \bar{\Lambda}^\delta{}_\alpha V_\delta, \tag{5.6}$$

with

$$\bar{\Lambda}^\delta{}_\alpha = \eta_{\alpha\beta} \Lambda^\beta{}_\gamma \eta^{\gamma\delta} \tag{5.7}$$

(one can use $\Lambda_\alpha{}^\beta$ instead of $\bar{\Lambda}^\beta{}_\alpha$ but one should be very careful in writings since $\Lambda_\alpha{}^\beta \neq \Lambda^\beta{}_\alpha$). Thanks to (3.10), we find

$$\bar{\Lambda}^\gamma{}_\alpha \Lambda^\alpha{}_\beta = \eta_{\alpha\delta} \eta^{\gamma\epsilon} \Lambda^\delta{}_\epsilon \Lambda^\alpha{}_\beta = \eta^{\gamma\epsilon} \eta_{\epsilon\beta} = \delta^\gamma{}_\beta \tag{5.8}$$

And similarly, we get $\Lambda^\beta{}_\alpha \bar{\Lambda}^\alpha{}_\gamma = \delta^\beta{}_\gamma$. In matrix notation, we have $\Lambda\bar{\Lambda} = \bar{\Lambda}\Lambda = \mathbb{I}$ and thus $\bar{\Lambda} = \Lambda^{-1}$. To summarize the transformations of 4-vectors:

- A contravariant vector transforms with Λ
- A covariant vector transforms with $\Lambda^{-1} = \bar{\Lambda}$

The scalar product of two vectors V^α and U^β is defined by

$$V^\alpha U_\alpha = V_\alpha U^\alpha = \eta^{\alpha\beta} V_\alpha U_\beta = \eta_{\alpha\beta} V^\alpha U^\beta \tag{5.9}$$

and is invariant under Lorentz transformations: $V'^\alpha U'_\alpha = \underbrace{\Lambda^\alpha{}_\beta \bar{\Lambda}^\delta{}_\alpha}_{\delta^\delta_\beta} V^\beta U_\delta = V^\beta U_\beta$.

The operator $\frac{\partial}{\partial x^\alpha}$ transforms like a covariant vector: $\frac{\partial}{\partial x'^\alpha} = \frac{\partial x^\beta}{\partial x'^\alpha} \frac{\partial}{\partial x^\beta}$. Since $\frac{\partial x^\beta}{\partial x'^\alpha} = \bar{\Lambda}^\beta{}_\alpha \Rightarrow \frac{\partial}{\partial x'^\alpha} = \bar{\Lambda}^\beta{}_\alpha \frac{\partial}{\partial x^\beta}$. We will now use the notations $\partial_\alpha \equiv \frac{\partial}{\partial x^\alpha}$ (covariant vector) and $\partial^\alpha \equiv \frac{\partial}{\partial x_\alpha}$ (contravariant vector). The D'Alembert operator can be written as $\Box = \partial^\alpha \partial_\alpha = \eta^{\alpha\beta} \partial_\alpha \partial_\beta = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta$ and is a Lorentz scalar.

A quantity is a rank r contravariant tensor if its components transform like the coordinates x^α:

$$T'^{\alpha_1 \ldots \alpha_r} = \Lambda^{\alpha_1}{}_{\beta_1} \ldots \Lambda^{\alpha_r}{}_{\beta_r} T^{\beta_1 \ldots \beta_r} \tag{5.10}$$

Tensors of rank 0 are scalars, tensors of rank 1 are vectors. For "mixed" tensors we have for example:

$$T'^\alpha{}_{\beta\gamma} = \Lambda^\alpha{}_\delta \bar{\Lambda}^\epsilon{}_\beta \bar{\Lambda}^\mu{}_\gamma T^\delta{}_{\epsilon\mu}$$

The following operations can be used to form new tensors:

1. Linear combinations of tensors with the same upper and lower indices: $T^\alpha{}_\beta = aR^\alpha{}_\beta + bS^\alpha{}_\beta$
2. Direct products of tensors: $T^{\alpha\beta} V^\gamma$ (works with mixed indices)
3. Contractions of tensors: $T^{\alpha\beta}{}_\beta$ or $T^{\alpha\beta} V_\beta$ (lowers a tensor by 2 in rank)
4. Differentiation of a tensor field: $\partial_\alpha T^{\beta\alpha}$ (the derivative ∂_α of any tensor is a tensor with one additional lower index α: $\partial_\alpha T^{\beta\gamma} \equiv R_\alpha{}^{\beta\gamma}$)
5. Going from a covariant to a contravariant component of a tensor is defined like in (5.2) and (5.5) (lowering and raising indices with $\eta^{\alpha\beta}$, $\eta_{\alpha\beta}$).

One must be aware that

- the order of the upper and lower indices is important,
- $\Lambda^{\alpha}{}_{\beta}$ is not a tensor.

η can be considered as a tensor: $\eta = \eta^{\alpha\beta} = \eta_{\alpha\beta}$ is the Minkowski tensor.

$$\eta'_{\alpha\beta} = \bar{\Lambda}^{\gamma}{}_{\alpha}\bar{\Lambda}^{\delta}{}_{\beta}\eta_{\gamma\delta} \overset{(3.10)}{=} \bar{\Lambda}^{\gamma}{}_{\alpha}\bar{\Lambda}^{\delta}{}_{\beta}\Lambda^{\mu}{}_{\gamma}\Lambda^{\nu}{}_{\delta}\eta_{\mu\nu} \overset{(5.8)}{=} \eta_{\alpha\beta}$$

η appears in the line element ($\mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}x^{\alpha}\mathrm{d}x^{\beta}$) and is thus the metric tensor in Minkowski space. We also have $\eta^{\alpha}{}_{\beta} = \eta^{\alpha\gamma}\eta_{\gamma\beta} = \delta^{\alpha}{}_{\beta} = \eta_{\beta}{}^{\alpha}$, and thus the Kronecker symbol is also a tensor.

We define the totally antisymmetric tensor or (Levi-Civita tensor) as

$$\epsilon^{\alpha\beta\gamma\delta} = \begin{cases} +1 & (\alpha,\beta,\gamma,\delta) \text{ is an even permutation of } (0,1,2,3) \\ -1 & (\alpha,\beta,\gamma,\delta) \text{ is an odd permutation of } (0,1,2,3) \\ 0 & \text{otherwise} \end{cases} \tag{5.11}$$

Without proof we have: ($det\,(\Lambda) = 1$)

$$\epsilon'^{\alpha\beta\gamma\delta} = \epsilon^{\alpha\beta\gamma\delta},$$

$$\epsilon_{\alpha\beta\gamma\delta} = \eta_{\alpha\alpha'}\eta_{\beta\beta'}\eta_{\gamma\gamma'}\eta_{\delta\delta'}\epsilon^{\alpha'\beta'\gamma'\delta'} = -\epsilon^{\alpha\beta\gamma\delta}.$$

The functions $S(x)$, $V^{\alpha}(x)$, $T^{\alpha\beta}\dots$ with $x = (x^0, x^1, x^2, x^3)$ are a scalar field, a vector field, or a tensor field ... respectively if:

$$S'(x') = S(x)$$

$$V'^{\alpha}(x') = \Lambda^{\alpha}_{\beta}V^{\beta}(x)$$

$$T'^{\alpha\beta}(x') = \Lambda^{\alpha}_{\delta}\Lambda^{\beta}_{\gamma}T^{\delta\gamma}(x)$$

$$\vdots$$

Also the argument has to be transformed, thus x' has to be understood as $x'^{\alpha} = \Lambda^{\alpha}{}_{\beta}x^{\beta}$.

6 Electrodynamics

Maxwell's equations relate the fields $\vec{E}(\vec{r},t)$, $\vec{B}(\vec{r},t)$, the charge density $\rho_e(\vec{r},t)$ and the current density $\vec{\jmath}(\vec{r},t)$:

$$\text{inhomogeneous}\begin{cases}\operatorname{div}\vec{E}=4\pi\rho_c\\ \\ \operatorname{rot}\vec{B}=\dfrac{4\pi}{c}\vec{\jmath}+\dfrac{1}{c}\dfrac{\partial\vec{E}}{\partial t}\end{cases}\qquad\text{homogeneous}\begin{cases}\operatorname{div}\vec{B}=0\\ \\ \operatorname{rot}\vec{E}=-\dfrac{1}{c}\dfrac{\partial\vec{B}}{\partial t}\end{cases}\tag{6.1}$$

The continuity equation

$$\operatorname{div}\vec{\jmath}+\dot{\rho}_c=0\rightarrow\partial_\alpha j^\alpha=0\tag{6.2}$$

with $j^\alpha=(c\rho_e,\vec{\jmath})$ follows from the conservation of charge, which for an isolated system implies $\partial_t\displaystyle\int j^0\,\mathrm{d}^3r=0$. $\partial_\alpha j^\alpha$ is a Lorentz scalar. We can define the *field strength tensor* which is given by the antisymmetric matrix

$$F^{\alpha\beta}=\begin{pmatrix}0&-E_x&-E_y&-E_z\\E_x&0&-B_z&B_y\\E_y&B_z&0&-B_x\\E_z&-B_y&B_x&0\end{pmatrix}.\tag{6.3}$$

Using this tensor we can rewrite the inhomogeneous Maxwell equations

$$\underbrace{\partial_\alpha F^{\alpha\beta}}_{4-\text{vector}}=\frac{4\pi}{c}\underbrace{j^\beta}_{4-\text{vector}},\tag{6.4}$$

and also the homogeneous ones:

$$\epsilon^{\alpha\beta\gamma\delta}\partial_\beta F_{\gamma\delta}=0.\tag{6.5}$$

Both equations are covariant under a Lorentz transformation. Eq. (6.5) allows to represent $F^{\alpha\beta}$ as a "curl" of a 4-vector A^α:

$$F^{\alpha\beta}=\partial^\alpha A^\beta-\partial^\beta A^\alpha.\tag{6.6}$$

We can then reformulate Maxwell's equations for $A^\alpha=(\phi,A^i)$. From (6.6) it follows that the gauge transformation

$$A^\alpha\rightarrow A^\alpha+\partial^\alpha\Theta\tag{6.7}$$

of the 4-vector A^α leaves $F^{\alpha\beta}$ unchanged, where $\Theta(x)$ is an arbitrary scalar field. The Lorenz gauge $\partial_\alpha A^\alpha=0$ leads to the decoupling of the inhomogeneous Maxwell's equation (6.4) to

$$\Box A^\alpha=\frac{4\pi}{c}j^\alpha.\tag{6.8}$$

The generalized equation of motion for a particle with charge q is

$$m\frac{\mathrm{d}u^\alpha}{\mathrm{d}\tau}=\frac{q}{c}F^{\alpha\beta}u_\beta\tag{6.9}$$

The spatial components give the expression of the Lorentz force $\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = q\left(\vec{E} + \frac{\vec{v}}{c} \wedge \vec{B}\right)$ with $\vec{p} = \gamma m \vec{v}$. The energy-momentum tensor for the electromagnetic field is

$$T^{\alpha\beta}_{em} = \frac{1}{4\pi}\left(F^{\alpha}{}_{\gamma}F^{\gamma\beta} + \frac{1}{4}\eta^{\alpha\beta}F_{\gamma\delta}F^{\gamma\delta}\right) \tag{6.10}$$

The 00-component represents the energy density of the field $T^{00}_{em} = u_{em} = \frac{1}{8\pi}\left(\vec{E}^2 + \vec{B}^2\right)$ and the 0i-components the Poynting vector $\vec{S}^i = cT^{0i}_{em} = \frac{c}{4\pi}\left(\vec{E} \wedge \vec{B}\right)^i$. In terms of these tensors, Maxwell's equations are $\partial_\alpha T^{\alpha\beta}_{em} = -\frac{1}{c}F^{\beta\gamma}j_\gamma$. $T^{\alpha\beta}_{em}$ is symmetric and conserved: $\partial_\alpha T^{\alpha\beta}_{em} = 0$. Setting $\beta = 0$ leads to energy conservation whereas $\partial_\alpha T^{\alpha k}_{em} = 0$ leads to conservation of the k^{th} component of the momentum. One should note that $\partial_\alpha T^{\alpha\beta}_{em} = 0$ is valid only if there is no external force, otherwise we can write $\partial_\alpha T^{\alpha\beta}_{em} = f^\beta$, where f^β is the external force density. Such an external force can often be included in the energy-momentum tensor.

7 Accelerated reference systems in special relativity

Non inertial systems can be considered in the context of special relativity. However, then the physical laws no longer have their simple covariant form. In e.g. a rotating coordinate system, additional terms will appear in the equations of motion (centrifugal terms, Coriolis force, etc.).

Let us look at a coordinate system KS' (with coordinates x'^μ) which rotates with constant angular speed with respect to an inertial system IS (x^α):

$$\begin{cases} x = x'\cos(\omega t') - y'\sin(\omega t'), \\ y = x'\sin(\omega t') + y'\cos(\omega t'), \\ z = z', \\ t = t', \end{cases} \tag{7.1}$$

and assume that $\omega^2(x'^2 + y'^2) \ll c^2$. Then we insert (7.1) into the line element $\mathrm{d}s$ (in the known IS form):

$$\begin{aligned} \mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta &= c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \\ &= \left[c^2 - \omega^2(x'^2 + y'^2)\right]\mathrm{d}t'^2 + 2\omega y'\mathrm{d}x'\mathrm{d}t' - 2\omega x'\mathrm{d}y'\mathrm{d}t' - \mathrm{d}x'^2 - \mathrm{d}y'^2 - \mathrm{d}z'^2 \\ &= g_{\mu\nu}\mathrm{d}x'^\mu\mathrm{d}x'^\nu. \end{aligned} \tag{7.2}$$

The resulting line element is more complicated. For arbitrary coordinates x'^μ, $\mathrm{d}s^2$ is a quadratic form of the coordinate differentials $\mathrm{d}x'^\mu$. Consider a general coordinate transformation from x^μ (in IS) to x'^μ (in KS'):

$$x^\alpha \equiv x^\alpha(x') = x^\alpha(x'^0, x'^1, x'^2, x'^3), \tag{7.3}$$

then we get for the line element

$$\mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta = \eta_{\alpha\beta}\frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\mathrm{d}x'^\mu \mathrm{d}x'^\nu = g_{\mu\nu}(x')\mathrm{d}x'^\mu \mathrm{d}x'^\nu, \tag{7.4}$$

with

$$g_{\mu\nu}(x') = \eta_{\alpha\beta}\frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}. \tag{7.5}$$

The quantity $g_{\mu\nu}$ is the metric tensor of the KS' system. It is symmetric ($g_{\mu\nu} = g_{\nu\mu}$) and depends on x'. It is called metric because it defines distances between points in coordinate systems.

In an accelerated reference system we get inertial forces. In the rotating frame we expect to experience the centrifugal force $\vec{Z}$, which can be written in terms of a centrifugal potential ϕ:

$$\phi = -\frac{\omega^2}{2}(x'^2 + y'^2) \quad \text{and} \quad \vec{Z} = -m\vec{\nabla}\phi. \tag{7.6}$$

This enables us to see that g_{00} from (7.2) is

$$g_{00} = 1 + \frac{2\phi}{c^2}. \tag{7.7}$$

The centrifugal potential appears in the metric tensor. We will see later that the first derivatives of the metric tensor are related to the forces in the relativistic equations of motion. To get the meaning of t' in KS' we evaluate (7.2) at a point with $\mathrm{d}x' = \mathrm{d}y' = \mathrm{d}z' = 0$:

$$\mathrm{d}\tau = \frac{\mathrm{d}s_{\text{clock}}}{c} = \sqrt{g_{00}}\,\mathrm{d}t' = \sqrt{1 + \frac{2\phi}{c^2}}\,\mathrm{d}t' = \underbrace{\sqrt{1 - \frac{v^2}{c^2}}}_{\substack{\text{correspond to clocks}\\ \text{time computed in}\\ \text{an inertial system}}}\,\mathrm{d}t \tag{7.8}$$

τ represents the time of a clock at rest in KS'.

In an inertial system we have $g_{\mu\nu} = \eta_{\mu\nu}$ and the clock moves with speed $v = \omega\rho$ ($\rho = \sqrt{x'^2 + y'^2}$). With (7.6) we see that both expressions in (7.8) are the same.

The coefficients of the metric tensor $g_{\mu\nu}(x')$ are functions of the coordinates. Such a dependence will also arise when one uses curved coordinates. Consider for example cylindrical coordinates:

$$x'^0 = ct = x^0, \quad x'^1 = \rho, \quad x'^2 = \theta, \quad x'^3 = z.$$

This results in the line element

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 = c^2\mathrm{d}t^2 - \mathrm{d}\rho^2 - \rho^2\mathrm{d}\theta^2 - \mathrm{d}z^2 = g_{\mu\nu}(x')\mathrm{d}x'^\mu \mathrm{d}x'^\nu. \tag{7.9}$$

Here $g_{\mu\nu}$ is diagonal:

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -\rho^2 & \\ & & & -1 \end{pmatrix}. \tag{7.10}$$

The fact that the metric tensor depends on the coordinates can be either due to the fact that the considered coordinate system is accelerated or that we are using non-cartesian coordinates.

Part III

Towards General Relativity

8 The equivalence principle

The principle of equivalence of gravitation and inertia tells us how an arbitrary physical system responds to an external gravitational field (with the help of tensor analysis). The physical basis of general relativity is the *equivalence principle* as formulated by Einstein:

1. Inertial and gravitational mass are equal
2. Gravitational forces are equivalent to inertial forces
3. In a local inertial frame, we experience the known laws of special relativity without gravitation

8.1 About the masses

The inertial mass m_t is the quantity appearing in Newton's law $\vec{F} = m_t\,\vec{a}$ which acts against acceleration by external forces. In contrast, the gravitational mass m_s is the proportionality constant relating the gravitational force between mass points to each other. For a particle moving in a homogeneous gravitational field, we have the equation $m_t\ddot{z} = -m_s g$, whose solution is

$$z(t) = -\frac{1}{2}\frac{m_s}{m_t}gt^2 \qquad (+v_0t + z_0). \tag{8.1}$$

Galilei stated that "all bodies fall at the same rate in a gravitational field", i.e. $\frac{m_s}{m_t}$ is the same for all bodies. Another experiment is to consider the period T of a pendulum (in the small amplitude approximation): $\left(\frac{T}{2\pi}\right)^2 = \frac{m_s}{m_t}\frac{l}{g}$, where l is the length of the pendulum. Newton verified that this period is independent on the material of the pendulum to a precision of about 10^{-3}. Eötvös (~1890), using torsion balance, got a precision of about 5×10^{-9}. Today's precision is about $10^{-11} \sim 10^{-12}$, this is way we can make the assumption $m_s = m_t$ on safe grounds.

Due to the equivalence between energy and mass ($E = mc^2$), all forms of energy contribute to mass, and due to the first point of the equivalence principle, to the inertial and to the gravitational masses.

8.2 About the forces

As long as gravitational and inertial masses are equal, then gravitational forces are equivalent to inertial forces: going to a well-chosen accelerated reference frame, one can get rid of the gravitational field. As an example take the equation of motion in the homogeneous gravitational field at Earth's surface:

$$m_t\frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = m_s\vec{g} \tag{8.2}$$

This expression is valid for a reference system which is at rest on Earth's surface ($\sim$ to a good approximation an IS). Then we perform the following transformation to an accelerated KS system:

$$\vec{r} = \vec{r}\,' + \frac{1}{2} g t'^2, \qquad t = t', \tag{8.3}$$

and we assume $gt \ll c$. The origin of KS $\vec{r}\,' = 0$ moves in IS with $\vec{r}(t) = \frac{1}{2} g t^2$. Then, inserting (8.3) into (8.2) results in

$$m_t \frac{\mathrm{d}^2}{\mathrm{d}t'^2} \left(\vec{r}\,' + \frac{1}{2} g t'^2 \right) = m_s \vec{g}$$

$$\Rightarrow m_t \frac{\mathrm{d}^2 \vec{r}\,'}{\mathrm{d}t'^2} = (m_s - m_t)\, \vec{g}. \tag{8.4}$$

If $m_s = m_t$, the resulting equation in KS is the one of a free moving particle $\frac{\mathrm{d}^2 \vec{r}\,'}{\mathrm{d}t'^2} = 0$; the gravitational force vanishes. As another example in a "free falling elevator" the "observer" does not feel any gravity.

Einstein generalized this finding postulating that (this is the *Einstein equivalence principle*) "in a free falling accelerated reference system *all* physical processes run as if there is no gravitational field". Notice that the "mechanical" finding is now expanded to *all* types of physical processes (at all times and places). Moreover also non-homogeneous gravitational fields are allowed. The equality of inertial and gravitational mass is also called the *weak equivalence principle* (or *universality of the free fall*).

As an example of a freely falling system, consider a satellite in orbit around Earth (assuming that the laboratory on the satellite is not rotating). Thus the equivalence principle states that in such a system all physical processes run as if there would be no gravitational field. The processes run as in an inertial system: the local IS. However, the local IS is *not* an inertial system, indeed the laboratory on the satellite is accelerated compared to the reference system of the fixed distant stars. The equivalence principle implies that in a local IS the rules of special relativity apply.

- For an observer on the satellite laboratory all physical processes follow special relativity and there are neither gravitational nor inertial forces.
- For an observer on Earth, the laboratory moves in a gravitational field and moreover inertial forces are present, since it is accelerated.

The motion of the satellite laboratory, i.e. its free falling trajectory, is such that the gravitational forces and inertial forces just compensate each other (cf (8.4)). The compensation of the forces is exactly valid only for the center of mass of the satellite laboratory. Thus the equivalence principle applies only to a very small or local satellite laboratory ("how small" depends on the situation).

The equivalence principle can also be formulated as follows:

> "At every space-time point in an arbitrary gravitational field, it is possible to choose a locally inertial coordinate system such that, within a sufficiently small region around the point in question, the laws of nature take the same form as in non-accelerated Cartesian coordinate systems in the absence of gravitation."[2]

[2]Notice the analogy with the axiom Gauss took as a basis of non-Euclidean geometry: he assumed that at any point on a curved surface we may erect a locally Cartesian coordinate system in which distances obey the law of Pythagoras.

The equivalence principle allows us to set up the relativistic laws including gravitation; indeed one can just perform a coordinate transformation to another KS:

$$\left.\begin{matrix}\text{special relativity laws}\\ \text{without}\\ \text{gravitation}\end{matrix}\right\} \xrightarrow[transformation]{coordinate} \left\{\begin{matrix}\text{relativistic laws}\\ \text{with}\\ \text{gravitation}\end{matrix}\right.$$

The coordinate transformation includes the relative acceleration between the laboratory system and KS which corresponds to the gravitational field. Thus from the equivalence principle we can derive the relativistic laws in a gravitational field. However, it does *not* fix the field equations for $g_{\mu\nu}(x)$ since those equations have no analogue in special relativity.

From a geometrical point of view the coordinate dependence of the metric tensor $g_{\mu\nu}(x)$ means that space is curved. In this sense the field equations describe the connection between curvature of space and the sources of the gravitational field in a quantitative way.

8.3 Riemann space

We denote with ξ^α the Minkowski coordinates in the local IS (e.g. the satellite laboratory). From the equivalence principle, the special relativity laws apply. In particular, we have for the line element

$$\mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}\xi^\alpha\mathrm{d}\xi^\beta. \tag{8.5}$$

Going from the local IS to a KS with coordinates x^μ is given by a coordinate transformation $\xi^\alpha = \xi^\alpha(x^0, x^1, x^2, x^3)$. Inserting this into (8.5) results in

$$\mathrm{d}s^2 = \eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial\xi^\beta}{\partial x^\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu = g_{\mu\nu}(x)dx^\mu dx^\nu, \tag{8.6}$$

and thus $g_{\mu\nu}(x) = \eta_{\alpha\beta}\dfrac{\partial\xi^\alpha}{\partial x^\mu}\dfrac{\partial\xi^\beta}{\partial x^\nu}$. A space with such a path element of the form (8.6) is called a *Riemann space.*

The coordinate transformation (expressed via $g_{\mu\nu}$) also describes the relative acceleration between KS and the local IS. Since at two different points of the local IS the accelerations are (in general) different, there is no global transformation in the form (8.6) that can be brought to the Minkowski form (8.5). We shall see that $g_{\mu\nu}$ are the relativistic gravitational potentials, whereas their derivatives determine the gravitational forces.

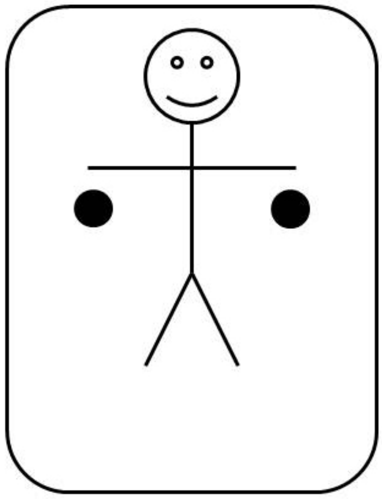


Figure 1: An experimenter and his two stones freely floating somewhere in outer space, i.e. in the absence of forces.

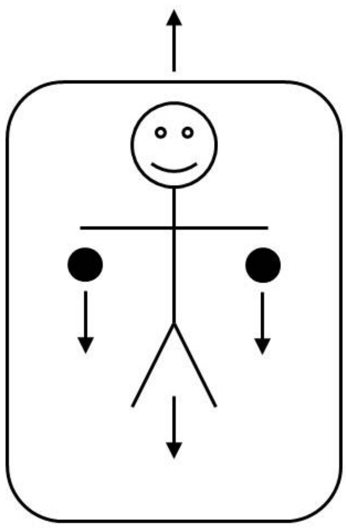


Figure 2: Constant acceleration upwards mimics the effect of a gravitational field: experimenter and stones drop to the bottom of the box.

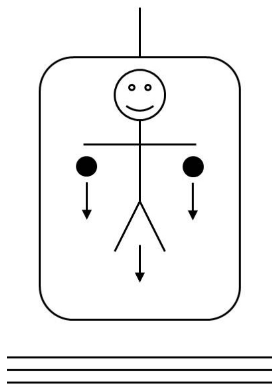


Figure 3: The effect of a constant gravitational field: indistinguishable for our experimenter from that of a constant acceleration as in figure 2.

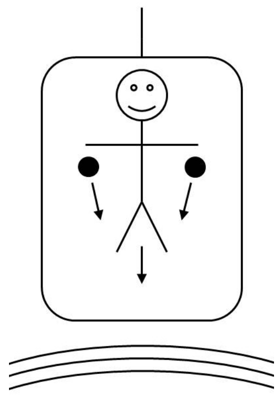


Figure 5: The experimenter and his stones in a non-uniform gravitational field: the stones will approach each other slightly as they fall to the bottom of the elevator.

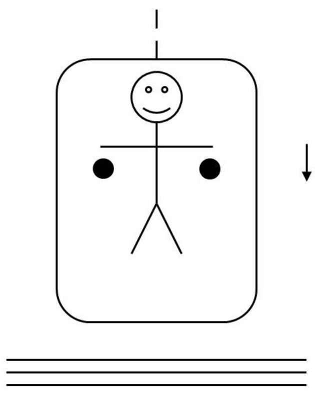


Figure 4: Free fall in a gravitational field has the same effect as no gravitational field (figure 1): experimenter and stones float.

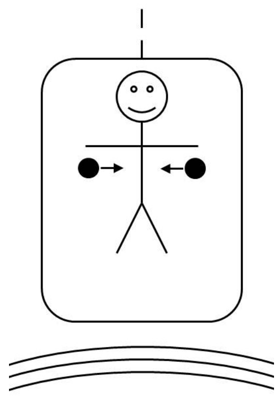


Figure 6: The experimenter and stones freely falling in a non-uniform gravitational field: the experimenter floats, so do the stones, but they move closer together, indicating the presence of some external forces.

9 Physics in a gravitational field

9.1 Equations of motion

According to the equivalence principle, in a local IS the laws of special relativity hold. For a mass point on which no forces act we have

$$\frac{\mathrm{d}^2\xi^\alpha}{\mathrm{d}\tau^2} = 0, \tag{9.1}$$

where the proper time τ is defined through $\mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}\xi^\alpha\mathrm{d}\xi^\beta = c^2\mathrm{d}\tau^2$. We can also define the 4-velocity as $u^\alpha = \frac{\mathrm{d}\xi^\alpha}{\mathrm{d}\tau}$. Solutions of (9.1) are straight lines

$$\xi^\alpha = a^\alpha\tau + b^\alpha. \tag{9.2}$$

Light (or a photon) moves in the local IS on straight lines. However, for photons τ cannot be identified with the proper time since on the light cone $\mathrm{d}s = c\mathrm{d}\tau = 0$. Thus we denote by λ a parameter of the trajectory of photons:

$$\frac{\mathrm{d}^2\xi^\alpha}{\mathrm{d}\lambda^2} = 0. \tag{9.3}$$

Let us now consider a global coordinate system KS with x^μ and metric $g_{\mu\nu}(x)$. At all points x^μ, one can locally bring $\mathrm{d}s^2$ into the form $\mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}\xi^\alpha\mathrm{d}\xi^\beta$. Thus at all points P there exists a transformation $\xi^\alpha(x) = \xi^\alpha(x^0, x^1, x^2, x^3)$ between ξ^α and x^μ. The transformation varies from point to point. Consider a small region around point P. Inserting the coordinate transformation into the line element, we get

$$\mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}\xi^\alpha\mathrm{d}\xi^\beta = \underbrace{\eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial\xi^\beta}{\partial x^\nu}}_{\equiv g_{\mu\nu}(x)\text{ metric tensor}} \mathrm{d}x^\mu\mathrm{d}x^\nu. \tag{9.4}$$

We write (9.1) in the form

$$0 = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\right) = \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2} + \frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau},$$

multiply it by $\frac{\partial x^\kappa}{\partial\xi^\alpha}$ and make use of $\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial x^\kappa}{\partial\xi^\alpha} = \delta^\kappa{}_\mu$. This way we can solve for $\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\tau^2}$ and get the following equation of motion in a gravitational field

$$\frac{\mathrm{d}^2x^\kappa}{\mathrm{d}\tau^2} = -\Gamma^\kappa_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau}, \tag{9.5}$$

with

$$\Gamma^\kappa_{\mu\nu} = \frac{\partial x^\kappa}{\partial\xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\nu}. \tag{9.6}$$

The $\Gamma^\kappa_{\mu\nu}$ are called the *Christoffel symbols* and represent a pseudo force or fictive gravitational field (like centrifugal or Coriolis forces) that arises whenever one describes inertial motion in non-inertial coordinates. Eq. (9.5) is a second order differential equation for the functions $x^\mu(\tau)$ which describe the trajectory of a particle in KS with $g_{\mu\nu}(x)$. Eq. (9.5) can also be written as $m\dfrac{\mathrm{d}u^\alpha}{\mathrm{d}\tau} = f^\alpha$, $u^\alpha = \frac{\mathrm{d}x^\alpha}{\mathrm{d}\tau}$. Comparing with (4.3) one infers that the right hand side of (9.5) describes the gravitational forces. Due to (9.4), the velocity $\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$ has to satisfy the condition

$$c^2 = g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \qquad (\text{for } m \neq 0) \tag{9.7}$$

(assume $\mathrm{d}\tau \neq 0$ and $m \neq 0$). Due to (9.7) only 3 of the 4 components of $\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}$ are independent (this is also the case for the 4-velocity in special relativity). For photons ($m = 0$) one finds instead, using (9.3), a completely analogous equation for the trajectory:

$$\frac{\mathrm{d}^2 x^\kappa}{\mathrm{d}\lambda^2} = -\Gamma^\kappa_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}, \tag{9.8}$$

and since $\mathrm{d}\tau = \mathrm{d}s = 0$, one has instead of (9.7):

$$0 = g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} \qquad (\text{for } m = 0).$$

9.2 Christoffel symbols

The Christoffel symbols can be expressed in terms of the first derivatives of $g_{\mu\nu}$. Consider with (9.4):

$$\begin{aligned}\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\nu} =& \eta_{\alpha\beta}\left[\frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\lambda}\frac{\partial \xi^\beta}{\partial x^\nu} + \underbrace{\cancel{\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial^2\xi^\beta}{\partial x^\nu\partial x^\lambda}}}_{1}\right]\\ &+ \eta_{\alpha\beta}\left[\frac{\partial^2\xi^\alpha}{\partial x^\lambda \partial x^\mu}\frac{\partial \xi^\beta}{\partial x^\nu} + \underbrace{\cancel{\frac{\partial\xi^\alpha}{\partial x^\lambda}\frac{\partial^2\xi^\beta}{\partial x^\nu\partial x^\mu}}}_{2}\right]\\ &- \eta_{\alpha\beta}\left[\underbrace{\cancel{\frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu}\frac{\partial \xi^\beta}{\partial x^\lambda}}}_{2} + \underbrace{\cancel{\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial^2\xi^\beta}{\partial x^\lambda\partial x^\nu}}}_{1}\right].\end{aligned}$$

Using $\eta_{\alpha\beta} = \eta_{\beta\alpha}$ this becomes

$$= 2\eta_{\alpha\beta}\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\lambda}\frac{\partial\xi^\beta}{\partial x^\nu}. \tag{9.9}$$

On the other hand

$$\begin{aligned} g_{\nu\sigma}\Gamma^\sigma_{\mu\lambda} &= \eta_{\alpha\beta}\overbrace{\frac{\partial\xi^\alpha}{\partial x^\nu}\frac{\partial\xi^\beta}{\partial x^\sigma}}^{g_{\nu\sigma}}\overbrace{\frac{\partial x^\sigma}{\partial \xi^\rho}\frac{\partial^2\xi^\rho}{\partial x^\mu\partial x^\lambda}}^{\Gamma^\sigma_{\mu\lambda}} \quad \text{with} \quad \underbrace{\frac{\partial\xi^\beta}{\partial x^\sigma}\frac{\partial x^\sigma}{\partial \xi^\rho}}_{\delta^\beta_\rho}\\ &= \eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\nu}\frac{\partial^2\xi^\beta}{\partial x^\mu\partial x^\lambda}\\ &= \frac{1}{2}\left[\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\nu}\right]. \end{aligned} \tag{9.10}$$

We introduce the inverse matrix $g^{\mu\nu}$ such that $g^{\mu\nu}g_{\nu\sigma} = \delta^\mu{}_\sigma$. Therefore we can solve with respect to the Christoffel symbols:

$$\Gamma^\kappa_{\mu\lambda} = \frac{1}{2}g^{\kappa\nu}\left[\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\nu}\right]. \tag{9.11}$$

Note that the Γ's are symmetric in the lower indices $\Gamma^{\kappa}_{\mu\nu} = \Gamma^{\kappa}_{\nu\mu}$. The gravitational forces on the right hand side of (9.6) are given by derivatives of $g_{\mu\nu}$. Comparing with the equation of motion of a particle in a electromagnetic field shows that the $\Gamma^{\lambda}_{\mu\nu}$ correspond to the field $F^{\alpha\beta}$, whereas the $g_{\mu\nu}$ correspond to the potentials A^{α}.

9.3 Newtonian limit

Let us assume that $v^i \ll c$ and the fields are weak and static (i.e. not time dependent). Thus $\frac{\mathrm{d}x^i}{\mathrm{d}\tau} \ll \frac{\mathrm{d}x^0}{\mathrm{d}\tau}$. Inserting this into (9.5) leads to

$$\frac{\mathrm{d}^2 x^\kappa}{\mathrm{d}\tau^2} = -\Gamma^{\kappa}_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \overbrace{\approx}^{\text{small velocity}} -\Gamma^{\kappa}_{00}\left(\frac{\mathrm{d}x^0}{\mathrm{d}\tau}\right)^2. \tag{9.12}$$

For static fields we get from (9.11):

$$\Gamma^{\kappa}_{00} \overbrace{=}^{\text{staticity}} -\frac{g^{\kappa i}}{2}\frac{\partial g_{00}}{\partial x^i} \qquad (i = 1, 2, 3) \tag{9.13}$$

(the other terms contain partial derivative with respect to x^0 which are zero by staticity). We write the metric tensor as $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. For weak fields we have $|h_{\mu\nu}| = |g_{\mu\nu} - \eta_{\mu\nu}| \ll 1$. In this case the coordinates (ct, x^i) are "almost" Minkowski coordinates. Inserting the expansion for $g_{\mu\nu}$ into (9.13) (taking only linear terms in h) gives

$$\Gamma^{\kappa}_{00} = \left(0, \frac{1}{2}\frac{\partial h_{00}}{\partial x^i}\delta^k_i\right). \tag{9.14}$$

Then, let us compute (9.12) for $\kappa = 0$, $\kappa = j$:

$$\frac{\mathrm{d}^2 t}{\mathrm{d}\tau^2} = 0 \Rightarrow \frac{\mathrm{d}t}{\mathrm{d}\tau} = \text{constant} \overbrace{=}^{\text{choice}} 1, \tag{9.15a}$$

$$\frac{\mathrm{d}^2 x^j}{\mathrm{d}\tau^2} = -\frac{c^2}{2}\frac{\partial h_{00}}{\partial x^j}\underbrace{\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2}_{1^2}. \tag{9.15b}$$

Taking $(x^j) = \vec{r}$, we can write

$$\frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = -\frac{c^2}{2}\nabla h_{00}(\vec{r}), \tag{9.16}$$

which is to be compared with the Newtonian case $\dfrac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = -\nabla\phi(\vec{r})$. Therefore:

$$g_{00}(\vec{r}) = 1 + h_{00}(\vec{r}) = 1 + \frac{2\phi(\vec{r})}{c^2}. \tag{9.17}$$

Notice that the Newtonian limit (9.16) gives no clue on the other components of $h_{\mu\nu}$. The quantity $\frac{2\phi}{c^2}$ is a measure of the strength of the gravitational field. Consider a spherically symmetric mass distribution. Then

$$\begin{aligned}
\frac{2\phi(R)}{c^2} &\approx 1.4 \times 10^{-9} && \text{at Earth surface,} \\
&\approx 4 \times 10^{-6} && \text{on the Sun (and similar stars),} \\
&\approx 3 \times 10^{-4} && \text{on a white dwarf,} \\
&\approx 3 \times 10^{-1} && \text{on a neutron star} \rightarrow \text{GR required.}
\end{aligned}$$

10 Time dilation

We study a clock in a static gravitational field and the phenomenon of gravitational redshift.

10.1 Proper time

The proper time τ of the clock is defined through the 4-dimensional line element as

$$\mathrm{d}\tau = \frac{\mathrm{d}s_{\text{clock}}}{c} = \frac{1}{c}\left(\sqrt{g_{\mu\nu}(x)\mathrm{d}x^\mu \mathrm{d}x^\nu}\right)_{\text{clock}}, \tag{10.1}$$

$x = (x^\mu)$ are the coordinates of the clock. The time shown by the clock depends on both the gravitational field and of its motion (the gravitational field being described by $g_{\mu\nu}$).

Special cases:

1. Moving clock in an IS without gravity :

$$\mathrm{d}\tau = \sqrt{1 - \frac{v^2}{c^2}}\mathrm{d}t$$

   ($g_{\mu\nu} = \eta_{\mu\nu}$, $\mathrm{d}x^i = v^i \mathrm{d}t$, $\mathrm{d}x^0 = c\mathrm{d}t$).

2. Clock at rest in a gravitational field ($\mathrm{d}x^i = 0$)

$$\mathrm{d}\tau = \sqrt{g_{00}}\,\mathrm{d}t.$$

   For a weak static field, one has with (9.17):

$$\mathrm{d}\tau = \sqrt{1 + \frac{2\phi(r)}{c^2}}\mathrm{d}t \qquad (|\phi| \ll c^2). \tag{10.2}$$

   The fact that ϕ is negative implies that a clock in a gravitational field goes more slowly than a clock outside the gravitational field.

10.2 Redshift

Let us now consider objects which emit or absorb light with a given frequency. Consider only a static gravitational field ($g_{\mu\nu}$ does not depend on time). A source in $\vec{r}_A$ (at rest) emits a monochromatic electromagnetic wave at a frequency ν_A. An observer at $\vec{r}_B$, also at rest, measures a frequency ν_B.

$$\begin{aligned} &\text{At source:} && \mathrm{d}\tau_A = \sqrt{g_{00}(\vec{r}_A)}\mathrm{d}t_A \\ &\text{At observer:} && \mathrm{d}\tau_B = \sqrt{g_{00}(\vec{r}_B)}\mathrm{d}t_B \end{aligned} \tag{10.3}$$

As a time interval we consider the time between two following peaks departing from A or arriving at B. In this case $\mathrm{d}\tau_A$ and $\mathrm{d}\tau_B$ correspond to the period of the electromagnetic waves at A and B, respectively, and therefore

$$\mathrm{d}\tau_A = \frac{1}{\nu_A}, \qquad \mathrm{d}\tau_B = \frac{1}{\nu_B}. \tag{10.4}$$

Going from A to B needs the same time Δt for the first and the second peak of the electromagnetic wave. Consequently, they will arrive with a time delay which is equal to the one with which they were emitted, thus $\mathrm{d}t_A = \mathrm{d}t_B$. With (10.3) and (10.4) we get:

$$\frac{\nu_A}{\nu_B} = \sqrt{\frac{g_{00}(\vec{r}_B)}{g_{00}(\vec{r}_A)}}, \qquad \text{with} \quad z = \frac{\nu_A}{\nu_B} - 1 = \frac{\lambda_B}{\lambda_A} - 1. \tag{10.5}$$

The quantity z is the *gravitational redshift*:

$$z = \sqrt{\frac{g_{00}(\vec{r}_B)}{g_{00}(\vec{r}_A)}} - 1. \tag{10.6}$$

For weak fields with $g_{00} = 1 + \frac{2\phi}{c^2}$ we have

$$z = \frac{\phi(\vec{r}_B) - \phi(\vec{r}_A)}{c^2} \qquad (|\phi| \ll c^2), \tag{10.7}$$

such a redshift is observed by measuring spectral lines from stars. As an example take solar light with (10.7)

$$z = \frac{\phi(r_B) - \phi(r_A)}{c^2} \approx -\frac{\phi(r_A)}{c^2} = \frac{GM_\odot}{c^2 R_\odot} \approx 2 \times 10^{-6},$$

with $M_\odot \approx 2 \times 10^{30}$ kg and $R_\odot \approx 7 \times 10^8$ m. For a white dwarf we find $z \approx 10^{-4}$ and for a neutron star $z \approx 10^{-1}$. In general there are 3 effects which can lead to a modification in the frequency of spectral lines:

1. Doppler shift due to the motion of the source (or of the observer)
2. Gravitational redshift due to the gravitational field at the source (or at the observer)
3. Cosmological redshift due to the expansion of the Universe (metric tensor is time dependent)

10.3 Photon in a gravitational field

Consider a photon with energy $E_\gamma = \hbar\omega = 2\pi\hbar\nu$, travelling upwards in the homogeneous gravity field of the Earth, covering a distance of $h = h_B - h_A$ (h small). The corresponding redshift is

$$z = \frac{\nu_A}{\nu_B} - 1 = \frac{\phi(r_B) - \phi(r_A)}{c^2} = \frac{g(h_B - h_A)}{c^2} = \frac{gh}{c^2}, \tag{10.8}$$

resulting in a frequency change $\Delta\nu = \nu_B - \nu_A$ ($\nu_A > \nu_B$, $\nu_B = \nu$) and thus

$$\frac{\Delta\nu}{\nu} = -\frac{gh}{c^2}. \tag{10.9}$$

The photon changes its energy by $\Delta E_\gamma = -\frac{E_\gamma}{c^2}gh$ (like a particle with mass $\frac{E_\gamma}{c^2} = m$). This effect has been measured in 1965 (through the Mössbauer effect) as $\dfrac{\Delta\nu_{exp}}{\Delta\nu_{th}} = 1.00 \pm 0.01$ (1% accuracy)[3].

11 Geometrical considerations

In general, the coordinate dependence of $g_{\mu\nu}(x)$ means that spacetime, defined through the line element $\mathrm{d}s^2$, is curved. The trajectories in a gravitational field are the geodesic lines in the corresponding spacetime.

11.1 Curvature of space

The line element in an N-dimensional Riemann space with coordinates $x = (x^1, \ldots, x^N)$ is given as

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu \qquad (\mu, \nu = 1, \ldots, N).$$

Let us just consider a two dimensional space $x = (x^1, x^2)$ with

$$\mathrm{d}s^2 = g_{11}\mathrm{d}x^1\mathrm{d}x^1 + 2g_{12}\mathrm{d}x^1\mathrm{d}x^2 + g_{22}\mathrm{d}x^2\mathrm{d}x^2. \tag{11.1}$$

Examples:

- Plane with Cartesian coordinates $(x^1, x^2) = (x, y)$:

  $$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2, \tag{11.2}$$

  or with polar coordinates $(x^1, x^2) = (\rho, \phi)$:

  $$\mathrm{d}s^2 = \mathrm{d}\rho^2 + \rho^2\mathrm{d}\phi^2 \tag{11.3}$$

- Surface of a sphere with angular coordinates $(x^1, x^2) = (\theta, \phi)$:

  $$\mathrm{d}s^2 = a^2\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2\right) \tag{11.4}$$

The line element (11.2) can, via a coordinate transformation, be brought into the form (11.3). However, there is no coordinate transformation which brings (11.4) into (11.2). Thus:

- The metric tensor determines the properties of the space, among which is also the curvature.
- The form of the metric tensor is not uniquely determined by the space, in other words it depends on the choice of coordinates.

[3]Pound, R. V. and Snider, J. L., *Effect of Gravity on Gamma Radiation*, Physical Review, 140

The curvature of the space is determined via the metric tensor (and it does not depend on the coordinate choice)[4]. If $g_{ik} = \text{const}$ then the space is not curved. In an Euclidian space, one can introduce Cartesian coordinates $g_{ik} = \delta_{ik}$. For a curved space $g_{ik} \neq \text{const}$ (does not always imply that space is curved). For instance by measuring the angles of a triangle and checking if their sum amounts to 180 degrees or differs, one can infer if the space is curved or not (for instance by being on the surface of a sphere).

[4]Beside the curvature discussed here, there is also an exterior curvature. We only consider intrinsic curvatures here.

Part IV

Differential Geometry

12 Differentiable manifolds

A manifold is a topological space that locally looks like the Euclidean $\mathbb{R}^n$ space with its usual topology. A simple example of a curved space is the $\mathcal{S}^2$ sphere: one can setup local coordinates (θ, φ) which map $\mathcal{S}^2$ onto a plane $\mathbb{R}^2$ (a *chart*). Collections of charts are called *atlases*. There is no one-to-one map of $\mathcal{S}^2$ onto $\mathbb{R}^2$; we need several charts to cover $\mathcal{S}^2$.

Definition: Given a (topological) space $\mathcal{M}$, a *chart* on $\mathcal{M}$ is a one-to-one map ϕ from an open subset $\mathcal{U} \subset \mathcal{M}$ to an open subset $\phi(\mathcal{U}) \subset \mathbb{R}^n$, i.e. a map $\phi : \mathcal{M} \to \mathbb{R}^n$. A chart is often called a coordinate system. A set of charts with domain $\mathcal{U}_\alpha$ is called an *atlas* of $\mathcal{M}$, if $\bigcup\limits_\alpha \mathcal{U}_\alpha = \mathcal{M}$, $\{\phi_\alpha | \alpha \in I\}$.

Definition: $\dim \mathcal{M} = n$

Definition: Two charts ϕ_1, ϕ_2 are $\mathcal{C}^\infty$-related if both the maps $\phi_2 \circ \phi_1^{-1} : \phi_1(\mathcal{U}_1 \cap \mathcal{U}_2) \to \phi_2(\mathcal{U}_1 \cap \mathcal{U}_2)$ and its inverse are $\mathcal{C}^\infty$. $\phi_2 \circ \phi_1^{-1}$ is the so-called *transition function* between the two coordinate charts. A collection of $\mathcal{C}^\infty$ related charts such that every point of $\mathcal{M}$ lies in the domain of at least one chart forms an atlas ($\mathcal{C}^\infty$: derivatives of all orders exist and are continuous).

The collection of *all* such $\mathcal{C}^\infty$-related charts forms a *maximal atlas*. If $\mathcal{M}$ is a space and A its maximal atlas, the set $(\mathcal{M}, A)$ is a ($\mathcal{C}^\infty$)-differentiable manifold. (If for each ϕ in the atlas the map $\phi : \mathcal{U} \to \mathbb{R}^n$ has the same n, then the manifold has dimension n.)

Important notions:

- A differentiable function $f : \mathcal{M} \to \mathbb{R}$ belongs to the algebra $\mathcal{F} = \mathcal{C}^\infty(\mathcal{M})$, sum and product of such functions are again in $\mathcal{F} = \mathcal{C}^\infty(\mathcal{M})$.
- $\mathcal{F}_p$ is the algebra of $\mathcal{C}^\infty$-functions defined in any neighbourhood of $p \in \mathcal{M}$ ($f = g$ means $f(q) = g(q)$ in some neighbourhood of p).
- A differentiable curve is a differentiable map $\gamma : \mathbb{R} \to \mathcal{M}$.
- Differentiable maps $F : \mathcal{M} \to \mathcal{M}'$ are differentiable if $\phi_2 \circ F \circ \phi_1^{-1}$ is a differentiable map for all suitable charts ϕ_1 of $\mathcal{M}$ and ϕ_2 of $\mathcal{M}'$.

The notions have to be understood by means of a chart, e.g. $f : \mathcal{M} \to \mathbb{R}$ is differentiable if $x \mapsto f(\underbrace{p(x)}_{\in \mathcal{M}}) \equiv f(x)$ is differentiable. This is independent of the chart representing a neighbourhood of p.

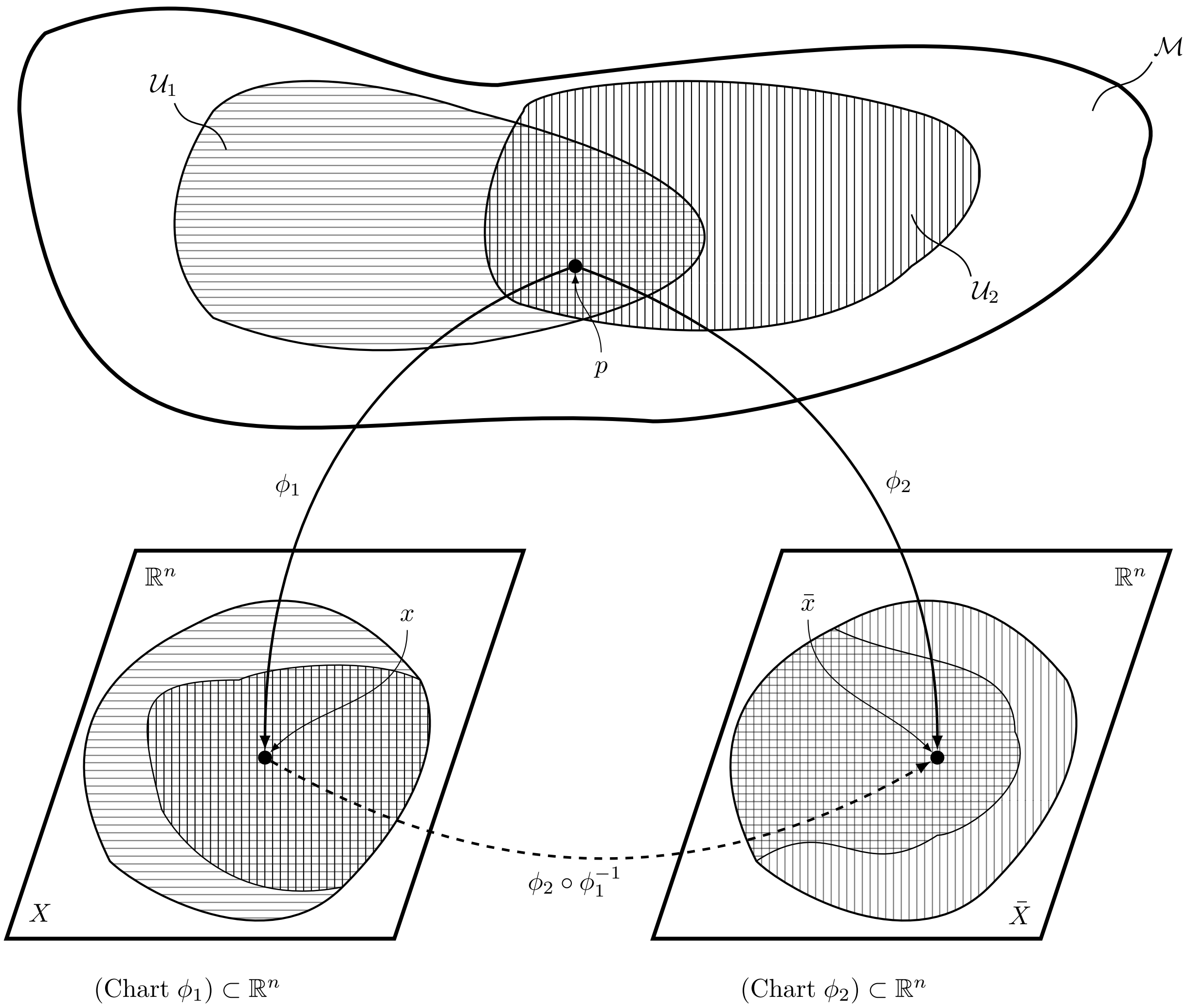


Figure 7: Manifold, charts and transition function.

12.1 Tangent vectors and tangent spaces

At every point p of a differentiable manifold $\mathcal{M}$ one can introduce a linear space, called tangent space $T_p(\mathcal{M})$. A *tensor field* is a (smooth) map which assigns to each point $p \in \mathcal{M}$ a tensor of a given type on $T_p(\mathcal{M})$.

<u>Definition:</u> a $\mathcal{C}^\infty$-curve in a manifold $\mathcal{M}$ is a map γ of the open interval $I = (a, b) \subset \mathbb{R} \to \mathcal{M}$ such that for any chart ϕ, $\phi \circ \gamma : I \to \mathbb{R}^n$ is a $\mathcal{C}^\infty$ map.

Let $f : \mathcal{M} \to \mathbb{R}$ be a smooth function on $\mathcal{M}$. Consider the map $f \circ \gamma : I \to \mathbb{R}$, $t \mapsto f(\gamma(t))$. This has a well-defined derivative: the rate of change of f along the curve. Consider $\underbrace{f \circ \phi^{-1}}_{\substack{\mathbb{R}^n \to \mathbb{R} \\ x^i \to f(x^i) \\ = f(\phi^{-1}(x^i))}} \circ \underbrace{\phi \circ \gamma}_{\substack{I \to \mathbb{R}^n \\ t \to x^i(\gamma(t))}}$ and use the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}t}(f \circ \gamma) = \sum_{i=1}^{n} \left(\frac{\partial}{\partial x^i} f(x^i) \right) \frac{\mathrm{d}x^i(\gamma(t))}{\mathrm{d}t}. \tag{12.1}$$

Thus, given a curve $\gamma(t)$ and a function f, we can obtain a qualitatively new object $\left[\frac{\mathrm{d}}{\mathrm{d}t}(f \circ \gamma) \right] \Big|_{t=t_0}$, the rate of change of f along the curve $\gamma(t)$ at $t = t_0$.

<u>Definition:</u> The *tangent vector* $\dot{\gamma}_p$ to a curve $\gamma(t)$ at a point p is a map from the set of real functions f defined in a neighbourhood of p to $\mathbb{R}$ defined by

$$\dot{\gamma}_p : f \mapsto \left[\frac{\mathrm{d}}{\mathrm{d}t}(f \circ \gamma) \right]_p = (f \circ \gamma)^{\bullet}{}_p = \dot{\gamma}_p(f). \tag{12.2}$$

Given a chart ϕ with coordinates x^i, the components of $\dot{\gamma}_p$ with respect to the chart are

$$(x^i \circ \gamma)^{\bullet}{}_p = \left[\frac{\mathrm{d}}{\mathrm{d}t} x^i(\gamma(t)) \right]_p . \tag{12.3}$$

The set of tangent vectors at p is the *tangent space* $T_p(\mathcal{M})$ at p.

<u>Theorem:</u> If the dimension of $\mathcal{M}$ is n, then $T_p(\mathcal{M})$ is a vector space of dimension n (without proof).

We set $\gamma(0) = p$ $(t = 0)$, $X_p = \dot{\gamma}_p$, and $X_p f = \dot{\gamma}_p(f)$. Eq. (12.3) determines $X_p(x^i)$, the components of X_p with respect to a given basis:

$$\begin{aligned} X_p f &= [f \circ \gamma]^{\bullet}(0) \\ &= \left[f \circ \phi^{-1} \circ \phi \circ \gamma \right]^{\bullet}(0) \\ &= \sum_{i=1}^{n} \frac{\partial}{\partial x^i}(f \circ \phi^{-1}) \frac{\mathrm{d}}{\mathrm{d}t}(x^i \circ \gamma)(0) \\ &= \sum_i \left(\frac{\partial}{\partial x^i} f(x^1, \ldots, x^n) \right) \left(X_p(x^i) \right) . \end{aligned} \tag{12.4}$$

This way we see that

$$X_p = \sum_i \left(X_p(x^i) \right) \left(\frac{\partial}{\partial x^i} \right)_p , \tag{12.5}$$

and so the $\left(\frac{\partial}{\partial x^i} \right)_p$ span $T_p(\mathcal{M})$. From (12.5) we see that $X_p(x^i)$ are the components of X_p with respect to the given basis ($X_p(x^i) = X_p^i$ or X^i).

Suppose that f, g are real functions on $\mathcal{M}$ and $fg : \mathcal{M} \to \mathbb{R}$ is defined as $fg(p) = f(p)g(p)$. If $X_p \in T_p(\mathcal{M})$, then (Leibniz rule)

$$X_p(fg) = (X_p f)g(p) + f(p)(X_p g). \tag{12.6}$$

Notation: $(Xf)(p) = X_p f$, $p \in \mathcal{M}$.

Basis of $T_p(\mathcal{M})$: $T_p = T_P(\mathcal{M})$ has dimension n. In any basis $(e_1, \ldots, e_n)$ we have $X = X^i e_i$. Changes of basis are given by

$$\bar{e}_i = \phi_i{}^k e_k, \qquad \bar{X}^i = \phi^i{}_k X^k. \tag{12.7}$$

The transformations $\phi_i{}^k$ and $\phi^i{}_k$ are inverse transposed of each other. In particular, $e_i = \frac{\partial}{\partial x^i}$ is called *coordinate basis* (with respect to a chart). Upon change of chart $x \mapsto \bar{x}$,

$$\phi_i{}^k = \frac{\partial x^k}{\partial \bar{x}^i}, \qquad \phi^i{}_k = \frac{\partial \bar{x}^i}{\partial x^k}. \tag{12.8}$$

Definition: The *cotangent space* T_p^* (or *dual space* T_p^* of T_p) consists of covectors $\omega \in T_p^*$, which are linear one-forms $\omega : X \mapsto \omega(X) \equiv < \omega, X > \in \mathbb{R}$ ($\omega : T_p \to \mathbb{R}$).

In particular for functions f, $\mathrm{d}f : X \mapsto Xf$ is an element of T_p^*. The elements $\mathrm{d}f = f_{,i}\, \mathrm{d}x^i = \left(\frac{\partial f}{\partial x^i}\right) \mathrm{d}x^i$ form a linear space of dimension n, therefore all of T_p^*.

We can define a *dual basis* $(e^1, \ldots, e^n)$ of T_p^*: $\omega = \omega_i e^i$. In particular the *dual basis* of a basis $(e_1, \ldots, e_n)$ of T_p is given by $< e^i, X >= X^i$ or $< e^i, X^j e_j >= X^j \underbrace{< e^i, e_j >}_{\delta^i{}_j} = X^i$. Thus $\omega_i =< \omega, e_i >$.

Upon changing the basis, the ω_i transform like the e_i and the e^i like the X^i (see (12.7)). In particular we have for the coordinate basis $e_i = \frac{\partial}{\partial x^i}$, $e^i = \mathrm{d}x^i$ ($< e^i, e_j >=< \mathrm{d}x^i, \frac{\partial}{\partial x^j} >= \delta^i{}_j$). The change of basis is:

$$\frac{\partial}{\partial \bar{x}^i} = \frac{\partial x^k}{\partial \bar{x}^i} \frac{\partial}{\partial x^k} = \phi_i{}^k \frac{\partial}{\partial x^k}$$

$$\mathrm{d}\bar{x}^i = \frac{\partial \bar{x}^i}{\partial x^k} \mathrm{d}x^k = \phi^i{}_k \mathrm{d}x^k$$

(Similar to co- and contravariant vectors.)

Tensors on T_p are multilinear forms on T_p^* and T_p, i.e. a tensor T of type $\binom{1}{2}$ (for short $T \in \otimes^1_2 T_p$): $T(\omega, X, Y)$ is a trilinear form on $T_p^* \times T_p \times T_p$. The *tensor product* is defined between tensors of any type, i.e. $T(\omega, X, Y) = R(\omega, X)S(Y)$: $T = R \otimes S$. In components:

$$T(\omega, X, Y) = \underbrace{T(e^i, e_j, e_k)}_{\equiv T^i{}_{jk}} \quad \underbrace{\omega_i X^j Y^k}_{e_i(\omega)e^j(X)e^k(Y)} \quad , \tag{12.9}$$

hence $T = T^i{}_{jk}\, e_i \otimes e^j \otimes e^k$. Any tensor of any type can therefore be obtained as a linear combination of tensor products $X \otimes \omega \otimes \omega'$ with $X \in T_p$, $\omega, \omega' \in T_p^*$. A change of basis can be performed similarly to the ones for vectors and covectors:

$$\bar{T}^i_{jk} = T^\alpha{}_{\beta\gamma}\phi^i{}_\alpha \phi_j{}^\beta \phi_k{}^\gamma \tag{12.10}$$

Trace: any bilinear form $b \in T_p^* \otimes T_p$ determines a linear form $l \in (T_p \otimes T_p^*)^*$ such that $l(X \otimes \omega) = b(X, \omega)$. In particular $\operatorname{tr} T$ is a linear form on tensors T of type $\binom{1}{1}$, defined by $\operatorname{tr}(X \otimes \omega) =< \omega, X >$. In components with respect to a dual pair of bases we have: $\operatorname{tr} T = T^\alpha{}_\alpha$. Similarly $T^i{}_{jk} \mapsto S_k = T^i{}_{ik}$ defines for instance a map from tensors of type $\binom{1}{2}$ to tensors of type $\binom{0}{1}$.

12.2 The tangent map

<u>Definition:</u> Let φ be a differentiable map: $\mathcal{M} \to \bar{\mathcal{M}}$ and let $p \in \mathcal{M}$, $\bar{p} = \varphi(p)$. Then φ induces a linear map (*"push-forward"*):

$$\varphi_* : T_p(\mathcal{M}) \to T_{\bar{p}}(\bar{\mathcal{M}}),$$

which we can describe in two ways:

(a) For any $\bar{f} \in \mathcal{F}_p(\bar{\mathcal{M}})$ ($\mathcal{F}$: space of all smooth functions on $\mathcal{M}$ (or $\bar{\mathcal{M}}$), that is $\mathcal{C}^\infty$ map $f : \mathcal{M} \to \mathbb{R}$):

$$(\varphi_* X)\bar{f} = X(\bar{f} \circ \varphi)$$

(b) Let γ be a representative of X ($X = \dot{\gamma}_p$, see (12.2) and (12.3)). Then let $\bar{\gamma} = \varphi \circ \gamma$ be a representative of $\varphi_* X$. This agrees with (a) since $\frac{\mathrm{d}}{\mathrm{d}t}\bar{f}(\bar{\gamma}(t))\big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}(\bar{f} \circ \varphi)(\gamma(t))\big|_{t=0}$.

With respect to bases $(e_1, \ldots, e_n)$ of T_p and $(\bar{e}_1, \ldots, \bar{e}_n)$ of $T_{\bar{p}}(\bar{\mathcal{M}})$, this reads $\bar{X} = \varphi_* X$: $\bar{X}^i = (\varphi_*)^i{}_k X^k$ with $(\varphi_*)^i{}_k =< \bar{e}^i, \varphi_* e_k >$ or in case of coordinate bases: $(\varphi_*)^i{}_k = \dfrac{\partial \bar{x}^i}{\partial x^k}$.

<u>Definition:</u> The *adjoint map* (or *"pull-back"*) φ^* of φ_* is defined as $\varphi^* : T^*_{\bar{p}} \to T^*_p$, $\bar{\omega} \mapsto \varphi^* \bar{\omega}$ ($= \omega$ in T_p^*) with $< \varphi^* \bar{\omega}, X >=< \bar{\omega}, \varphi_* X >$. The same result is obtained from the definition

$$\varphi^* : \mathrm{d}\bar{f} \mapsto \mathrm{d}(\bar{f} \circ \varphi), \qquad \bar{f} \in \mathcal{F}(\bar{\mathcal{M}}). \tag{12.10a}$$

In components, $\omega = \varphi^* \bar{\omega}$ reads $\omega_k = \bar{\omega}_i (\varphi_*)^i{}_k$.

Consider (local) diffeomorphisms, i.e. maps φ such that φ^{-1} exists in a neighbourhood of $\bar{p}$. Note that $\dim \mathcal{M} = \dim \bar{\mathcal{M}}$ and $\det \left(\frac{\partial \bar{x}^i}{\partial x^j}\right) \neq 0$. Then φ_* and φ^*, as defined above, are invertible and may be extended to tensors of arbitrary types.

Example: tensor of type $\binom{1}{1}$

$$(\varphi_* T)(\bar{\omega}, \bar{X}) = T(\underbrace{\varphi^* \bar{\omega}}_{\omega}, \underbrace{\varphi_*^{-1} \bar{X}}_{X}),$$

$$(\varphi^* \bar{T})(\omega, X) = \bar{T}(\underbrace{(\varphi^*)^{-1} \omega}_{\bar{\omega}}, \underbrace{\varphi_* X}_{\bar{X}}).$$

Here, φ_* and φ^* are the inverse of each other and we have

$$\begin{aligned} \varphi_*(T \otimes S) &= (\varphi_* T) \otimes (\varphi_* S), \\ \operatorname{tr}(\varphi_* T) &= \varphi_*(\operatorname{tr} T), \end{aligned} \tag{12.11}$$

and similarly for φ^*. In components $\bar{T} = \varphi_* T$ reads

$$\bar{T}^i_k = T^\alpha{}_\beta \frac{\partial \bar{x}^i}{\partial x^\alpha} \frac{\partial x^\beta}{\partial \bar{x}^k} \quad \text{(in a coordinate basis).} \tag{12.12}$$

This is formally the same as for transformation (12.10) when changing basis.

13 Vector and tensor fields

Definition: If to every point p of a differentiable manifold $\mathcal{M}$ a tangent vector $X_p \in T_p(\mathcal{M})$ is assigned, then we call the map $X\colon\, p \mapsto X_p$ a *vector field* on $\mathcal{M}$.

Given a coordinate system x^i and associated basis $\left(\frac{\partial}{\partial x^i}\right)_p$ for each $T_p(\mathcal{M})$, X_p has components X^i_p with $X_p = X^i_p \left(\frac{\partial}{\partial x^i}\right)_p$ and $X^i_p = X_p(x^i)$ (see (12.5)). Eq. (12.8) shows how X^i_p transform under coordinate transformations. The quantity Xf is called the derivative of f with respect to the vector field X. The following rules apply:

$$\begin{aligned} X(f+g) &= Xf + Xg, \\ X(fg) &= (Xf)g + f(Xg) \quad \text{(Leibnitz rule).} \end{aligned} \tag{13.1}$$

The vector fields on $\mathcal{M}$ form a linear space on which the following operations are defined as well:

$$X \mapsto fX \qquad \text{(multiplication by } f \in \mathcal{F}\text{),}$$

$$X, Y \mapsto [X,Y] = XY - YX \qquad \text{(commutator).}$$

$[X,Y]$, unlike XY, satisfies the Leibniz rule (13.1). The components of the commutator of two vector fields X, Y relative to a local coordinate basis can be obtained by its action on x^i. Thus using $X = X^i \frac{\partial}{\partial x^i}$ and $Y = Y^k \frac{\partial}{\partial x^k}$ we get

$$[X,Y]^j = (XY - YX)x^j$$

$$Yx^j = Y^k \frac{\partial x^j}{\partial x^k} = Y^k \delta^j{}_k = Y^j$$

$$XY^j = X^k \frac{\partial}{\partial x^k}(Y^j) = X^k \underbrace{Y^j{}_{,k}}_{\frac{\partial Y^j}{\partial x^k}}$$

$$\Rightarrow XY^j - YX^j = X^k Y^j{}_{,k} - Y^k X^j{}_{,k}$$

In a local coordinate basis, the bracket $[\partial_k, \partial_j]$ clearly vanishes ($X^k = 1$ and $Y^k = 1$, and thus $Y^k_{,j} = 0$). The Jacobi identity holds:

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0. \tag{13.2}$$

Definition: Let $T_p(\mathcal{M})^r_s$ be the set of all tensors of rank (r,s) defined on $T_p(\mathcal{M})$ (contravariant of rank r, covariant of rank s). If we assign to every $p \in \mathcal{M}$ a tensor $t_p \in T_p(\mathcal{M})^r_s$, then the map $t : p \mapsto t_p$ defines a *tensor field* of type $\binom{r}{s}$.

Algebraic operations on tensor fields are defined point-wise; for instance the sum of two tensor fields is defined by $(t+\tilde{s})_p = t_p + \tilde{s}_p$ where $t, \tilde{s} \in T_p(\mathcal{M})^r_s$. Tensor products and contractions of tensor fields are defined analogously. Multiplication by a function $f \in \mathcal{F}(\mathcal{M})$ is given by $(ft)_p = f(p)t_p$. In a neighbourhood $\mathcal{U}$ of p, having coordinates $(x^1, \ldots, x^n)$ a tensor field can be expanded in the form

$$t = \underbrace{t^{i_1 \ldots i_r}{}_{j_1 \ldots j_s}}_{\substack{\text{components of } t \text{ relative}\\ \text{to the coordinate system}\\ (x^1, \ldots, x^n)}} \left(\frac{\partial}{\partial x^{i_1}} \otimes \ldots \otimes \frac{\partial}{\partial x^{i_r}} \right) \otimes \left(\mathrm{d}x^{j_1} \otimes \ldots \otimes \mathrm{d}x^{j_s} \right). \tag{13.3}$$

If the coordinates are transformed to $(\bar{x}^1, \ldots, \bar{x}^n)$ the components of t transform according to

$$\bar{t}^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} \equiv t^{k_1 \ldots k_r}{}_{l_1 \ldots l_s} \frac{\partial \bar{x}^{i_1}}{\partial x^{k_1}} \cdots \frac{\partial \bar{x}^{i_r}}{\partial x^{k_r}} \frac{\partial x^{l_1}}{\partial \bar{x}^{j_1}} \cdots \frac{\partial x^{l_s}}{\partial \bar{x}^{j_s}}. \tag{13.4}$$

(We shall consider C^∞ tensor fields). Covariant tensors of order 1 are also called *one-forms*. The set of tensor fields of type $\binom{r}{s}$ is denoted by $\mathcal{T}^r_s(\mathcal{M})$.

Definition: A *pseudo-Riemannian metric* on a differentiable manifold $\mathcal{M}$ is a tensor field $g \in \mathcal{T}^0_2(\mathcal{M})$ having the properties:

(i) $g(X,Y) = g(Y,X)$ for all X, Y

(ii) For every $p \in \mathcal{M}$, g_p is a non-degenerate ($\neq 0$) bilinear form on $T_p(\mathcal{M})$. This means that $g_p(X,Y) = 0$ for all $X \in T_p(\mathcal{M})$ if and only if $Y = 0$.

The tensor field $g \in \mathcal{T}^0_2(\mathcal{M})$ is a (proper) Riemannian metric if g_p is positive definite at every point p.

Definition: A *(pseudo-)Riemannian* manifold is a differentiable manifold $\mathcal{M}$, together with a (pseudo-) Riemannian metric g.

13.1 Flows and generating vector fields

A *flow* is a 1-parametric group of diffeomorphisms: $\varphi_t : \mathcal{M} \to \mathcal{M}$, $s, t \in \mathbb{R}$ with $\varphi_t \circ \varphi_s = \varphi_{t+s}$. In particular $\varphi_0 = \mathrm{id}$. Moreover, the orbits (or integral curves) of any point $p \in \mathcal{M}$, $t \mapsto \varphi_t(p) \equiv \gamma(t)$ shall be differentiable. A flow determines a vector field X by means of

$$Xf = \frac{\mathrm{d}}{\mathrm{d}t} (f \circ \varphi_t) \Big|_{t=0} \tag{13.5}$$

i.e. $X_p = \frac{\mathrm{d}}{\mathrm{d}t}\gamma(t)\big|_{t=0} = \dot{\gamma}(0)$ (see (12.2) and (12.3)). $\dot{\gamma}(0)$ is the tangent vector to γ at the point $p = \gamma(0)$. At the point $\gamma(t)$ we have then

$$\dot{\gamma}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\varphi_t(p) = \frac{\mathrm{d}}{\mathrm{d}s} (\varphi_s \circ \varphi_t)(p) \Big|_{s=0} = X_{\varphi_t(p)}$$

i.e. $\gamma(t)$ solves the ordinary differential equation:

$$\dot{\gamma}(t) = X_{\gamma(t)}, \quad \gamma(0) = p. \tag{13.6}$$

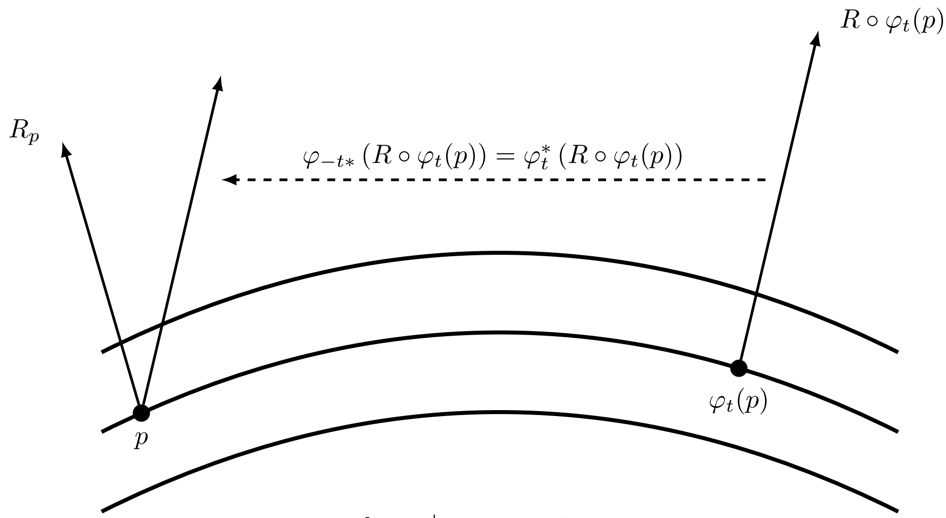
The generating vector field determines the flow uniquely. Not always does (13.6) admit global solutions (i.e. for all $t \in \mathbb{R}$), however for most purposes, "local flows" are good enough.

14 Lie derivative

The derivative of a vector field V rests on the comparison of V_p and $V_{p'}$ at nearby points p, p'. Since $V_p \in T_p$ and $V_{p'} \in T_{p'}$ belong to different spaces their difference can be taken only after $V_{p'}$ has been transported to V_p. This can be achieved by means of the tangent map φ_* (Lie transport). The *Lie derivative* $L_X R$ of a tensor field R in direction of a vector field X is defined by

$$L_X R = \frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^* R\bigg|_{t=0}, \tag{14.1}$$

or more explicitly $(L_X R)_p = \dfrac{\mathrm{d}}{\mathrm{d}t}\varphi_t^* R_{\varphi_t(p)}\bigg|_{t=0}$. Here φ_t is the (local) flow generated by X, where $\varphi_t^* R_{\varphi_t(p)}$ is a tensor on T_p depending on t.



$$L_X R = \frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^* R\bigg|_{t=0} = \lim_{t\to 0}\frac{1}{t}\left(\varphi_t^* R - R\right)$$
$$t \to \varphi_t(p) = \gamma(t); X_p = \tfrac{d}{dt}\gamma(t)|_{t=0} = \dot\gamma(0)$$
(φ^* is the inverse of φ_*)

Figure 8: Illustration of the Lie derivative

In order to express L_X in components we write φ_t in a chart: $\varphi_t : x \mapsto \bar{x}(t)$, and linearize it for small t: $\bar{x}^i = x^i + tX^i(x) + \mathcal{O}(t^2)$, $x^i = \bar{x}^i - tX^i(\bar{x}) + \mathcal{O}(t^2)$, thus $\frac{\partial^2 \bar{x}^i}{\partial x^k \partial t} = -\frac{\partial^2 x^i}{\partial \bar{x}^k \partial t} = X^i{}_{,k}$ at $t = 0$.

As an example, let R be of type $\binom{1}{1}$. By (12.12) we have $(\varphi_t^* R)^i{}_j(x) = R^\alpha{}_\beta(\bar{x})\frac{\partial x^i}{\partial \bar{x}^\alpha}\frac{\partial \bar{x}^\beta}{\partial x^j}$. Taking (according to (14.1)) a derivative with respect to t at $t = 0$ yields

$$(L_X R)^i{}_j = R^i{}_{j,k} X^k - R^\alpha{}_j X^i{}_{,\alpha} + R^i{}_\beta X^\beta{}_{,j} \tag{14.2}$$

(first term: $\underbrace{\dfrac{\partial}{\partial \bar{x}^k} R^\alpha{}_\beta(\bar{x})}_{R^\alpha{}_{\beta,k}(\bar{x})}\underbrace{\dfrac{\partial \bar{x}^k}{\partial t}}_{X^k}\dfrac{\partial x^i}{\partial \bar{x}^\alpha}\dfrac{\partial \bar{x}^\beta}{\partial x^j}\bigg|_{t=0} = R^i{}_{j,k}X^k$).

Properties of L_X:

(a) L_X is a linear map from tensor fields to tensor fields of the same type.

(b) $L_X(\operatorname{tr} T) = \operatorname{tr}(L_X T)$

(c) $L_X(T \otimes S) = (L_X T) \otimes S + T \otimes (L_X S)$

(d) $L_X f = Xf \quad (f \in \mathcal{F}(\mathcal{M}))$

(e) $L_X Y = [X, Y] \quad$ (Y vector field)

(proof: (a) follows from (14.1), (b) and (c) from (12.11), (d) from (13.5), whereas (e) is more involved).

Further properties of L_X: if X,Y are vector fields and $\lambda \in \mathbb{R}$, then

(i) $L_{X+Y} = L_X + L_Y, \quad L_{\lambda X} = \lambda L_X$

(ii) $L_{[X,Y]} = [L_X, L_Y] = L_X \circ L_Y - L_Y \circ L_X$

"Proof" of (ii): Apply it to $f \in \mathcal{F}(\mathcal{M})$,

$$[L_X, L_Y]f = L_X \circ L_Y f - L_Y \circ L_X f = L_X(Yf) - L_Y(Xf) = XYf - YXf = [X,Y]f = L_{[X,Y]}f.$$

Next apply it on a vector field Z:

$$[L_X, L_Y]Z \underset{(e)}{=} [X,[Y,Z]] - [Y,[X,Z]] \underset{\substack{\text{Jacobi}\\\text{identity}}}{=} [[X,Y],Z] = L_{[X,Y]}Z.$$

For higher rank tensors the derivation follows from the use of (c).
If $[X,Y] = 0$ then $L_X L_Y = L_Y L_X$ and for ϕ and ψ, which are the flows generated by X and Y, one finds: $\phi_s \circ \psi_t = \psi_t \circ \phi_s$.

15 Differential forms

Definition: A *p-form* Ω is a totally antisymmetric tensor field of type $\binom{0}{p}$

$$\Omega(X_{\pi(1)},\dots,X_{\pi(p)}) = (\operatorname{sign}\pi)\Omega(X_1,\dots,X_p)$$

for any permutation π of $\{1,\dots,p\}$ ($\pi \in S_p$ (group of permutations)) with $\operatorname{sign}\pi$ being its parity. For $p > \dim\mathcal{M}$, $\Omega \equiv 0$. Any tensor field of type $\binom{0}{p}$ can be antisymmetrized by means of the operation $\mathcal{A}$:

$$(\mathcal{A}T)(X_1,\dots,X_p) = \frac{1}{p!}\sum_{\pi\in S_p}(\operatorname{sign}\pi)T(X_{\pi(1)},\dots,X_{\pi(p)}) \tag{15.1}$$

with $\mathcal{A}^2 = \mathcal{A}$. The *exterior product* of a p_1-form Ω^1 with a p_2-form Ω^2 is the (p_1+p_2)-form:

$$\Omega^1 \wedge \Omega^2 = \frac{(p_1+p_2)!}{p_1!\,p_2!}\mathcal{A}(\Omega^1\otimes\Omega^2) \tag{15.2}$$

Properties:

- $\Omega^1\wedge\Omega^2 = (-1)^{p_1p_2}\,\Omega^2\wedge\Omega^1$
- $\Omega^1\wedge(\Omega^2\wedge\Omega^3) = (\Omega^1\wedge\Omega^2)\wedge\Omega^3 = \dfrac{(p_1+p_2+p_3)!}{p_1!\,p_2!\,p_3!}\mathcal{A}(\Omega^1\otimes\Omega^2\otimes\Omega^3)$

The components in a local basis $(e^1,\dots,e^n)$ of 1-forms are

$$\begin{aligned}
\Omega &= \Omega_{i_1\dots i_p}e^{i_1}\otimes\ldots\otimes e^{i_p} = \mathcal{A}\Omega\\
&= \Omega_{i_1\dots i_p}\mathcal{A}(e^{i_1}\otimes\ldots\otimes e^{i_p})\\
&= \sum_{i_1=1}^{n}\cdots\sum_{i_p=1}^{n}\Omega_{i_1\dots i_p}\frac{1}{p!}e^{i_1}\wedge\ldots\wedge e^{i_p}\\
&= \sum_{1\le i_1<\ldots<i_p\le n}\Omega_{i_1\dots i_p}e^{i_1}\wedge\ldots\wedge e^{i_p}
\end{aligned} \tag{15.3}$$

A covariant tensor of rank p, which is antisymmetric under exchange of any pair of indices (i.e. is a p-form), in n dimensions has $\binom{n}{p} = \frac{n!}{(n-p)!p!}$ independent components.

Examples:

- For 1-forms A, B (vector fields) we have

$$(A\wedge B)_{ik} = A_iB_k - A_kB_i = (-1)(B\wedge A)_{ik}.$$

- For a 2-form A and a 1-form B

$$(A\wedge B)_{ikl} = A_{ik}B_l + A_{kl}B_i + A_{li}B_k, \tag{15.4}$$

since

$$
\begin{aligned}
A \wedge B &= \frac{(1+2)!}{1!\,2!}\mathcal{A}(A \otimes B) \\
&= \frac{3!}{1!\,2!}(A_{ik}B_l)\frac{1}{3!}e^i \wedge e^k \wedge e^l \\
&= \frac{1}{2}(A_{ik}B_l)e^i \wedge e^k \wedge e^l \\
&= \frac{1}{2}\frac{1}{3}(A_{ik}B_l + \text{cyclic permutations})e^i \wedge e^k \wedge e^l \\
&= (A_{ik}B_l + \text{cyclic permutations})\frac{1}{3!}e^i \wedge e^k \wedge e^l .
\end{aligned}
$$

Thus by comparing with (15.3) we get (15.4).

15.1 Exterior derivative of a differential form

The derivative $\mathrm{d}f$ of a 0-form $f \in \mathcal{F}$ is the 1-form $\mathrm{d}f(X) = Xf$: the argument X (vector) acts as a derivation. In a local coordinate basis: $\mathrm{d}f = \frac{\partial f}{\partial x^i}\mathrm{d}x^i$. The exterior derivative is performed by an operator d applied to forms, converting p-forms to $(p+1)$-forms. The derivative $\mathrm{d}\Omega$ of a 1-form Ω is given by

$$\mathrm{d}\Omega(X_1, X_2) = X_1\Omega(X_2) - X_2\Omega(X_1) - \Omega([X_1, X_2]). \tag{15.5}$$

This expression is verified as follows:

$$X_1\Omega(X_2) = X_1 \underbrace{\langle \Omega, X_2 \rangle}_{\text{1-form}} = X_1^i \underbrace{\frac{\partial}{\partial x^i}}_{,i} \left(\Omega_k X_2^k\right) = X_1^i \Omega_{k,i} X_2^k + X_1^i \Omega_k X_{2,i}^k ,$$

$$X_2\Omega(X_1) = X_2^k \Omega_{i,k} X_1^i + X_2^k \Omega_i X_{1,k}^i ,$$

$$\Omega([X_1, X_2]) = \langle \Omega, X_1X_2 - X_2X_1 \rangle = \Omega_i (X_1X_2 - X_2X_1)^i = \Omega_i \left(X_1^k X_{2,k}^i - X_2^k X_{1,k}^i\right) ,$$

then

$$\mathrm{d}\Omega(X_1, X_2) = (\Omega_{k,i} - \Omega_{i,k})X_1^i X_2^k .$$

This is manifestly a 2-form (the coefficient also fits the expectations: $\frac{1}{2!}\frac{(1+1)!}{1!1!} = 1$). One can easily verify that

$$\mathrm{d}\Omega(fX_1, X_2) = f\mathrm{d}\Omega(X_1, X_2). \tag{15.6}$$

For $\Omega \wedge f = f\Omega$ (as f is a 0-form), the product rule

$$\mathrm{d}(\Omega \wedge f) = \mathrm{d}\Omega \wedge f - \Omega \wedge \mathrm{d}f$$

applies, as one can verify

$$\mathrm{d}(\Omega \wedge f)(X_1, X_2) \underset{(15.5)}{=} X_1(f\Omega)(X_2) - X_2(f\Omega)(X_1) - (f\Omega)([X_1, X_2]),$$

and

$$X_1(f\Omega)(X_2) = X_1^i \frac{\partial}{\partial x^i}(f\Omega_k X_2^k) = \underbrace{f X_1^i \frac{\partial}{\partial x^i}\left(\Omega_k X_2^k\right)}_{f X_1 \Omega(X_2)} + \underbrace{X_1^i \frac{\partial f}{\partial x^i} \Omega_k X_2^k}_{\mathrm{d}f(X_1)\Omega(X_2)}.$$

So

$$\mathrm{d}(f \wedge \Omega)(X_1, X_2) = \underbrace{f \mathrm{d}\Omega(X_1, X_2)}_{\mathrm{d}\Omega \wedge f} + \underbrace{\Omega(X_2)\mathrm{d}f(X_1) - \Omega(X_1)\mathrm{d}f(X_2)}_{-\Omega \wedge \mathrm{d}f}. \tag{15.7}$$

Moreover we have $\mathrm{d}^2 f = 0$, since

$$\begin{aligned} \mathrm{d}^2 f(X_1, X_2) &\underset{(15.5)}{=} X_1 \mathrm{d}f(X_2) - X_2 \mathrm{d}f(X_1) - \mathrm{d}f([X_1, X_2]) \\ &= X_1 X_2 f - X_2 X_1 f - [X_1, X_2] f = 0. \end{aligned} \tag{15.8}$$

The generalization of the definition to a p-form Ω gives

$$\begin{aligned} \mathrm{d}\Omega(X_1, .., X_{p+1}) &= \sum_{i=1}^{p+1} (-1)^{i-1} X_i \Omega(X_1, .., \hat{X}_i, .., X_{p+1}) \\ &+ \sum_{i<j}^{p+1} (-1)^{i+j} \Omega([X_i, X_j], X_1, .., \hat{X}_i, .., \hat{X}_j, .., X_{p+1}), \end{aligned} \tag{15.9}$$

where ˆ means omission, e.g. $(X_1, \hat{X}_2, X_3) = (X_1, X_3)$.

One can show that the following properties hold:

(a) d is a linear map from p-forms to $p+1$-forms,

(b) $\mathrm{d}(\Omega^1 \wedge \Omega^2) = \mathrm{d}\Omega^1 \wedge \Omega^2 + (-1)^{p_1} \Omega^1 \wedge \mathrm{d}\Omega^2$,

(c) $\mathrm{d}^2 = 0$, i.e. $\mathrm{d}(\mathrm{d}\Omega) = 0$,

(d) $\mathrm{d}f(X) = Xf \ (f \in \mathcal{F})$,

By means of (a)-(d) we have an alternative definition of d. By eq. (15.3) we have with respect to a coordinate basis

$$\Omega = \frac{1}{p!} \Omega_{i_1 \dots i_p} \mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_p}, \quad \text{and hence} \tag{15.10a}$$

$$\mathrm{d}\Omega \underbrace{=}_{\mathrm{dd}x^{i_p}=0} \frac{1}{p!} \mathrm{d}\Omega_{i_1 \dots i_p} \wedge \mathrm{d}x^{i_1} \wedge \dots \wedge \mathrm{d}x^{i_p} \tag{15.10b}$$

Components:

$$
\begin{aligned}
p!\,\mathrm{d}\Omega &= \Omega_{i_1\dots i_p,i_0}\,\mathrm{d}x^{i_0}\wedge\mathrm{d}x^{i_1}\wedge\ldots\wedge\mathrm{d}x^{i_p}\\
&= -\Omega_{i_0 i_2\dots i_p,i_1}\mathrm{d}x^{i_0}\wedge\ldots\wedge\mathrm{d}x^{i_p}\\
&= (-1)^k\Omega_{i_0\dots\hat{i}_k\dots i_p,i_k}\mathrm{d}x^{i_0}\wedge\ldots\wedge\mathrm{d}x^{i_p} \qquad (k=0,\dots,p)
\end{aligned}
$$

$$
\Rightarrow \mathrm{d}\Omega = \underbrace{\frac{1}{p!}\frac{1}{p+1}}_{\frac{1}{(p+1)!}}\underbrace{\sum_{k=0}^{p}(-1)^k\Omega_{i_0\dots\hat{i}_k\dots i_p,i_k}}_{(\mathrm{d}\Omega)_{i_0\dots i_p}}\mathrm{d}x^{i_0}\wedge\ldots\wedge\mathrm{d}x^{i_p} \tag{15.11}
$$

Examples:

- $p=1$:
$$(\mathrm{d}\Omega)_{ik} = \Omega_{k,i}-\Omega_{i,k} \tag{15.12}$$
- $p=2$:
$$(\mathrm{d}\Omega)_{ikl} = \Omega_{ik,l}+\Omega_{kl,i}+\Omega_{li,k} \tag{15.13}$$

Consider a map $\varphi:\mathcal{M}\to\bar{\mathcal{M}}$ and $\varphi^*:T^*_{\bar{p}}(\bar{\mathcal{M}})\to T^*_p(\mathcal{M})$; then

$$\varphi^*\circ\mathrm{d}=\mathrm{d}\circ\varphi^*. \tag{15.14}$$

A "proof" is found by using (15.10), (12.11) and property (b). It suffices to verify (15.14) on 0-forms and 1-forms. For 0-forms $\bar{f}$, (15.14) is identical to (12.10a). For 1-forms which are differentials $\mathrm{d}\bar{f}$, due to (c) we have

$$
\begin{aligned}
&(\varphi^*\circ\mathrm{d})(\mathrm{d}\bar{f}) = 0 \qquad (\mathrm{d}^2\bar{f}=0),\\
&(\mathrm{d}\circ\varphi^*)(\mathrm{d}\bar{f}) = \mathrm{d}(\varphi^*\circ\mathrm{d}\bar{f}) \underset{\substack{(12.10a)\\(\varphi^*\circ\mathrm{d}\bar{f})\\=\mathrm{d}(\bar{f}\circ\varphi)}}{=} \mathrm{d}(\mathrm{d}(\bar{f}\circ\varphi)) = \mathrm{d}^2(\bar{f}\circ\varphi)=0.
\end{aligned}
$$

Setting $\varphi=\varphi_t$ (the flow generated by X) and forming (14.1) ($L_X R = \frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^* R\big|_{t=0}$), one obtains the infinitesimal version of (15.14):

$$L_X\circ\mathrm{d}=\mathrm{d}\circ L_X. \tag{15.15}$$

Definition: A p-form ω with

- $\omega=\mathrm{d}\eta$ is *exact*
- $\mathrm{d}\omega=0$ is *closed*

An exact p-form is always closed ($\mathrm{d}^2\eta=0$), but the converse is not generally true (Poincaré lemma gives conditions under which the converse is valid).[5] [6]

[5] η is not unique since gauge transformations $\eta\mapsto\eta+\mathrm{d}\rho$, with ρ any $(p-1)$-form, leave $\mathrm{d}\eta$ unchanged.

[6] This is a generalization of the results of three-dimensional vector analysis: $\operatorname{rot}\operatorname{grad}f=0$ and $\operatorname{div}\operatorname{rot}\vec{k}=0$.

The integral of an n-form:

$\mathcal{M}$ is orientable within an atlas of "positively oriented" charts, if $\det\left(\frac{\partial \bar{x}^i}{\partial x^j}\right) > 0$ for any change of coordinates. For an n-form ω ($n = \dim \mathcal{M}$):

$$\omega = \omega_{i_1 \ldots i_n} \frac{1}{n!} \mathrm{d}x^{i_1} \wedge \ldots \wedge \mathrm{d}x^{i_n} = \underbrace{\omega_{1 \ldots n}}_{\omega(x)} \mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n \tag{15.16}$$

is determined by the single component $\omega(x)$; under a change of coordinates $\omega(x)$ transforms like

$$\bar{\omega}(\bar{x}) = \bar{\omega}_{1 \ldots n} = \underbrace{\omega_{i_1 \ldots i_n}}_{\substack{\text{totally} \\ \text{antisymmetric}}} \frac{\partial x^{i_1}}{\partial \bar{x}^1} \cdots \frac{\partial x^{i_n}}{\partial \bar{x}^n} = \omega(x) \det\left(\frac{\partial x^i}{\partial \bar{x}^j}\right). \tag{15.17}$$

The integral of a n-form is defined as follows:

$$\int_{\mathcal{M}} \omega = \int_{\mathcal{U}} \mathrm{d}x^1 \ldots \mathrm{d}x^n \, \omega(x^1, \ldots, x^n) \qquad \text{(if the support of } \omega \text{ is contained in a chart } \mathcal{U}\text{).}$$

This integral is independent of the choice of coordinates, since in different coordinates $\int \mathrm{d}x^1 \ldots \mathrm{d}x^n \, \omega(x) = \int \mathrm{d}\bar{x}^1 \ldots \mathrm{d}\bar{x}^n \, \omega(x) \left|\det\left(\frac{\partial x^i}{\partial \bar{x}^j}\right)\right|$ and (15.17) applies. [7] [8]

15.2 Stokes theorem

Let D be a region in a n dimensional differentiable manifold $\mathcal{M}$. The boundary ∂D consists of those $p \in D$ whose image x in some chart satisfies e.g. $x^1 = 0$. One can show that ∂D is a closed $(n-1)$ dimension submanifold of $\mathcal{M}$. If $\mathcal{M}$ is orientable then ∂D is also orientable. D shall have a smooth boundary and be such that $\bar{D}$ is compact. Then for every $(n-1)$-form ω we have

$$\int_D \mathrm{d}\omega = \int_{\partial D} \omega \tag{15.18}$$

15.3 The inner product of a p-form

Definition: Let X be a vector field on $\mathcal{M}$. For any p-form Ω we define the *inner product* as

$$(i_X \Omega)(X_1, \ldots, X_{p-1}) \equiv \Omega(X, X_1, \ldots, X_{p-1}) \tag{15.19}$$

(and zero if $p = 0$).

Properties:

[7] Actually, it is often impossible to cover the whole manifold with a single set of coordinates. In the general case it is necessary to introduce different sets of coordinates in different overlapping patches of the manifold, with the constraint that in the overlap between the patch with coordinate x^i and another patch with coordinate $\bar{x}^i$, the x^i can be expressed in a smooth one-to-one way as functions of $\bar{x}^i$ and vice-versa (orientable manifold).

[8] The integral over a p-form over the overlap between two patches (x^i and $\bar{x}^i$) can be evaluated using either coordinate system, provided $\det\left(\frac{\partial x^i}{\partial \bar{x}^j}\right) > 0$.

(a) i_X is a linear map from p-forms to $(p-1)$-forms,

(b) $i_X(\Omega^1 \wedge \Omega^2) = (i_X\Omega^1) \wedge \Omega^2 + (-1)^{p_1}\Omega^1 \wedge (i_X\Omega^2)$,

(c) ${i_X}^2 = 0$,

(d) $i_X \mathrm{d}f = Xf = \langle \mathrm{d}f, X\rangle$ with $f \in \mathcal{F}(\mathcal{M})$,

(e) $L_X = i_X \circ \mathrm{d} + \mathrm{d} \circ i_X$.

Proof of (e): for 0-forms f we have

$$L_X f = Xf,$$

$$i_X \circ \mathrm{d}f + \underbrace{\cancel{\mathrm{d} \circ i_X f}}_{=0} = i_X df = Xf,$$

and for 1-forms $\mathrm{d}f$

$$L_X \mathrm{d}f \underset{\substack{(15.15)\\ L_X \circ \mathrm{d} = \mathrm{d} \circ L_X}}{=} \mathrm{d}(L_X f) = \mathrm{d}(Xf),$$

$$i_X \circ \underbrace{\cancel{\mathrm{d}\mathrm{d}f}}_{=0} + \mathrm{d} \circ i_X \mathrm{d}f = \mathrm{d}(Xf).$$

Application: Gauss theorem

Let X be a vector field. Then $\mathrm{d}(i_X\eta)$ is an n-form with $dim\ M = n$. η is an n-form, and if $\eta_p \neq 0\ \forall p \in \mathcal{M}$, then η is a "volume form". A function $\operatorname{div}_\eta X \in \mathcal{F}$ is defined through

$$(\operatorname{div}_\eta X)\eta = \mathrm{d}(i_X\eta) = L_X\eta.^9 \tag{15.20}$$

We can apply Stokes theorem since $\mathrm{d}(i_X\eta)$ is an n-form and thus $i_X\eta$ an $(n-1)$-form:

$$\int_D \mathrm{d}(i_X\eta) = \int_D (\operatorname{div}_\eta X)\eta = \int_{\partial D} i_X\eta. \tag{15.21}$$

The standard volume form η is given by $\eta = \sqrt{|g|}\,\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n$.

[9] $\mathrm{d}\eta = 0$, thus $L_X = i_X \circ \mathrm{d} + \mathrm{d} \circ i_X$ applied on η gives $L_X\eta = i_X \circ \cancel{\mathrm{d}\eta} + \mathrm{d}(i_X\eta)$.

Expression for $\mathrm{div}_\eta X$ in *local* coordinates:

Let $\eta = a(x)\,\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n$, $X = X^i \frac{\partial}{\partial x^i}$. Then since $(\mathrm{div}_\eta X)\eta = L_X \eta$, we have (using property (c) of the Lie derivative):

$$L_X \eta = (Xa)\,\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n + a \sum_{i=1}^{n} \mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}(Xx^i) \wedge \ldots \wedge \mathrm{d}x^n.$$

Since $\mathrm{d}(Xx^i) = \mathrm{d}(X^k \underbrace{\frac{\partial}{\partial x^k} x^i}_{\delta^i_k}) = \mathrm{d}X^i(x) = X^i{}_{,j}\mathrm{d}x^j$, but $\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^j \wedge \ldots \wedge \mathrm{d}x^n \neq 0$ only if $j = i$ (otherwise we have two identical $\mathrm{d}x^i$) we find

$$\begin{aligned} L_X \eta &= \underbrace{Xa}_{X^i \frac{\partial a}{\partial x^i}}\,\mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n + a \sum_{i=1}^{n} X^i{}_{,i} \mathrm{d}x^1 \wedge \ldots \wedge \mathrm{d}x^n \\ &= (X^i a_{,i} + a X^i{}_{,i}) \frac{1}{a} \eta \end{aligned}$$

$$\Rightarrow \mathrm{div}_\eta X = \frac{1}{a}(aX^i)_{,i} = \frac{1}{\sqrt{|g|}} \left(\sqrt{|g|} X^i\right)_{,i} \qquad \text{for the “standard” } \eta. \tag{15.22}$$

16 Affine connections: Covariant derivative of a vector field

Definition: An *affine* (linear) *connection* or *covariant differentiation* on a manifold $\mathcal{M}$ is a mapping ∇ which assigns to every pair X, Y of $\mathcal{C}^\infty$ vector fields on $\mathcal{M}$ another $\mathcal{C}^\infty$ vector field $\nabla_X Y$ with the following properties:

$$\begin{aligned} &(i)\ \nabla_X Y \text{ is bilinear in } X \text{ and } Y, \\ &(ii)\ \text{if } f \in \mathcal{F}(\mathcal{M}), \text{ then} \\ &\qquad \nabla_{fX} Y = f \nabla_X Y, \\ &\qquad \nabla_X (fY) = f \nabla_X Y + X(f) Y. \end{aligned} \tag{16.1}$$

Lemma: Let X and Y be vector fields. If X vanishes at the point p on $\mathcal{M}$, then $\nabla_X Y$ also vanishes at p.

Proof: Let $\mathcal{U}$ be a coordinate neighbourhood of p. On $\mathcal{U}$ we have the representation $X = \xi^i \frac{\partial}{\partial x^i}$, $\xi^i \in \mathcal{F}(\mathcal{U})$ with $\xi^i(p) = 0$. Then $(\nabla_X Y)_p = \nabla_{\xi^i \frac{\partial}{\partial x^i}} Y = \underbrace{\xi^i(p)}_{=0} [\nabla_{\frac{\partial}{\partial x^i}} Y]_p = 0$.

Since $\nabla_X Y$ produces again a vector field, the result of the covariant differentiation can only be a linear combination of again the basis in the current chart. This leads us to the following statement:

Definition: One sets, relative to a chart $(X^1, \dots, X^n)$ for $\mathcal{U} \subset \mathcal{M}$:

$$\boxed{\nabla_{\frac{\partial}{\partial x^i}} \left(\frac{\partial}{\partial x^j} \right) = \Gamma^k_{ij} \frac{\partial}{\partial x^k}} \tag{16.2}$$

The n^3 functions $\Gamma^k_{ij} \in \mathcal{F}(\mathcal{U})$ are called *Christoffel symbols* (or *connection coefficients*) of the connection ∇ in a given chart.[10]

The Christoffel symbols are not tensors:

$$\nabla_{\frac{\partial}{\partial \bar{x}^a}} \left(\frac{\partial}{\partial \bar{x}^b} \right) = \bar{\Gamma}^c_{ab} \frac{\partial}{\partial \bar{x}^c} = \bar{\Gamma}^c_{ab} \frac{\partial x^k}{\partial \bar{x}^c} \frac{\partial}{\partial x^k}. \tag{16.3}$$

If we use (16.1):

$$\begin{aligned} \nabla_{\frac{\partial}{\partial \bar{x}^a}} \left(\frac{\partial}{\partial \bar{x}^b} \right) &= \nabla_{\left(\frac{\partial x^i}{\partial \bar{x}^a} \frac{\partial}{\partial x^i} \right)} \left(\frac{\partial x^j}{\partial \bar{x}^b} \frac{\partial}{\partial x^j} \right) = \frac{\partial x^i}{\partial \bar{x}^a} \left[\frac{\partial x^j}{\partial \bar{x}^b} \Gamma^k_{ij} \frac{\partial}{\partial x^k} + \frac{\partial}{\partial x^i} \left(\frac{\partial x^j}{\partial \bar{x}^b} \right) \frac{\partial}{\partial x^j} \right] \\ &= \frac{\partial x^i}{\partial \bar{x}^a} \frac{\partial x^j}{\partial \bar{x}^b} \Gamma^k_{ij} \frac{\partial}{\partial x^k} + \frac{\partial^2 x^j}{\partial \bar{x}^a \partial \bar{x}^b} \frac{\partial}{\partial x^j}. \end{aligned}$$

Comparison with 16.3:

$$\begin{aligned} \frac{\partial x^k}{\partial \bar{x}^c} \bar{\Gamma}^c_{ab} &= \frac{\partial x^i}{\partial \bar{x}^a} \frac{\partial x^j}{\partial \bar{x}^b} \Gamma^k_{ij} + \frac{\partial^2 x^k}{\partial \bar{x}^a \partial \bar{x}^b} \\ \Rightarrow \bar{\Gamma}^c_{ab} &= \frac{\partial x^i}{\partial \bar{x}^a} \frac{\partial x^j}{\partial \bar{x}^b} \frac{\partial \bar{x}^c}{\partial x^k} \Gamma^k_{ij} + \frac{\partial \bar{x}^c}{\partial x^k} \frac{\partial^2 x^k}{\partial \bar{x}^a \partial \bar{x}^b} \end{aligned} \tag{16.4}$$

The second term is not compatible with being a tensor.

If for every chart there exist n^3 functions Γ^k_{ij} which transform according to (16.4) under a change of coordinates, then one can show that there exists a *unique* affine connection ∇ on $\mathcal{M}$ which satisfies (16.3).

Definition: for every vector field X we can introduce the tensor $\nabla X \in T^1_1(\mathcal{M})$ defined by

$$\nabla X(Y, \omega) \equiv \langle \omega, \nabla_Y X \rangle, \tag{16.5}$$

where ω is a one-form. ∇X is called the *covariant derivative* of X.

In a chart $(x_1, \dots, x_n)$, let $X = \xi^i \partial_i$ and $\nabla X = \xi^i{}_{;j} \mathrm{d}x^j \otimes \partial_i$ ($< dx^i, \partial_i >= \delta^i_k$):

$$\xi^i{}_{;j} = \nabla X(\partial_j, \mathrm{d}x^i) = \left\langle \mathrm{d}x^i, \nabla_{\partial_j} X \right\rangle = \left\langle \mathrm{d}x^i, \xi^k{}_{,j} \partial_k + \xi^k \Gamma^\delta_{jk} \partial_\delta \right\rangle = \xi^i{}_{,j} + \Gamma^i_{jk} \xi^k \quad [11] \tag{16.6}$$

[10]For a pseudo-Riemannian manifold, the corresponding connection coefficients are given by (9.6) or (9.11).

[11]semicolon shall denote the covariant derivative ("normal derivative" + additional terms, that vanish in (cartesian) Euclidean or Minkowski space)

16.1 Parallel transport along a curve

Definition: let $\gamma : I \to \mathcal{M}$ be a curve in $\mathcal{M}$ with velocity field $\dot{\gamma}(t)$, and let X be a vector field on some open neighbourhood of $\gamma(I)$. X is said to be *autoparallel along* γ if

$$\nabla_{\dot{\gamma}} X = 0. \tag{16.7}$$

The vector $\nabla_{\dot{\gamma}} X$ is sometimes denoted as $\frac{DX}{\mathrm{d}t}$ or $\frac{\nabla X}{\mathrm{d}t}$ (covariant derivative along γ). In terms of coordinates, we have $X = \xi^i \partial_i$, $\dot{\gamma} = \frac{\mathrm{d}x^i}{\mathrm{d}t}\partial_i$ (see (12.3)). With (16.1) and (16.2) we get

$$\begin{aligned}
\nabla_{\dot{\gamma}} X &= \nabla_{\frac{\mathrm{d}x^i}{\mathrm{d}t}\partial_i}(\xi^k \partial_k) \\
&= \frac{\mathrm{d}x^i}{\mathrm{d}t} \nabla_{\partial_i}(\xi^k \partial_k) \\
&= \frac{\mathrm{d}x^i}{\mathrm{d}t} \left[\xi^k \Gamma^j_{ik} \partial_j + \partial_i \xi^k \partial_k\right] \\
&= \frac{\mathrm{d}x^i}{\mathrm{d}t} \left[\xi^j \Gamma^k_{ij} \partial_k + \partial_i \xi^k \partial_k\right] \\
&= \left[\frac{\mathrm{d}\xi^k}{\mathrm{d}t} + \Gamma^k_{ij} \frac{\mathrm{d}x^i}{\mathrm{d}t} \xi^j\right] \partial_k,
\end{aligned} \tag{16.8}$$

where we used $\dfrac{\mathrm{d}x^i}{\mathrm{d}t}\dfrac{\partial \xi^k}{\partial x^i} = \dfrac{\mathrm{d}\xi^k}{\mathrm{d}t}$. This shows that $\nabla_{\dot{\gamma}} X$ only depends on the values of X along γ. In terms of coordinates we get for (16.7)

$$\frac{\mathrm{d}\xi^k}{\mathrm{d}t} + \Gamma^k_{ij} \frac{\mathrm{d}x^i}{\mathrm{d}t} \xi^j = 0. \tag{16.9}$$

For a curve γ and any two point $\gamma(s)$ and $\gamma(t)$ consider the mapping

$$\tau_{t,s} : T_{\gamma(s)}(\mathcal{M}) \to T_{\gamma(t)}(\mathcal{M}),$$

which transforms a vector $v(s)$ at $\gamma(s)$ into the parallel transported vector $v(t)$ at $\gamma(t)$. The mapping $\tau_{t,s}$ is the *parallel transport* along γ from $\gamma(s)$ to $\gamma(t)$. We have $\tau_{s,s} = \mathbb{1}$ and $\tau_{r,s} \circ \tau_{s,t} = \tau_{r,t}$.

We can now give a geometrical interpretation of the covariant derivative that will be generalized to tensors. Let X be a vector field along γ, then

$$\nabla_{\dot{\gamma}} X(\gamma(t)) = \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} \tau_{t,s} X(\gamma(s)), \tag{16.10}$$

Proof: Let's work in a given chart. By construction, $v(t) = \tau_{t,s} v(s)$ with $v(s) \in T_{\gamma(s)}(\mathcal{M})$ and due to (16.8) it satisifies: $\dot{v}^i + \Gamma^i_{kj} \dot{x}^k v^j = 0$. If we write $(\tau_{t,s} v(s))^i = (\tau_{t,s})^i{}_j v^j(s) = v^i(t)$ (with $\tau_{t,s} = (\tau_{s,t})^{-1}$

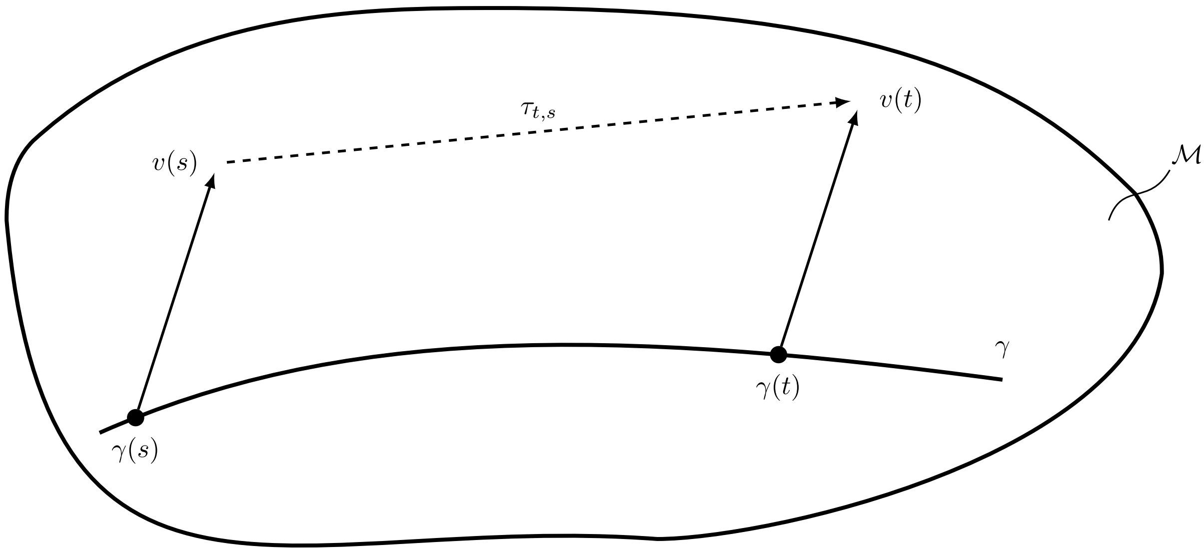


Figure 9: Illustration of parallel transport.

and $\tau_{s,s} = \mathbb{1}$), we get

$$\begin{aligned}\dot{v}^i(s) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=s} v^i(t) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=s} \left[(\tau_{t,s})^i{}_j v^j(s)\right] \\ &= \left(\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=s} (\tau_{t,s})^i{}_j\right) v^j(s) \\ &= -\Gamma^i_{kj}\dot{x}^k v^j(s).\end{aligned}$$

$$\Rightarrow \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=s} (\tau_{t,s})^i{}_j = -\Gamma^i_{kj}\dot{x}^k \tag{16.11}$$

Since $\tau_{t,s} = (\tau_{s,t})^{-1}$, $\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} (\tau_{t,s})^i{}_j = -\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=s} (\tau_{t,s})^i{}_j = \Gamma^i_{kj}\dot{x}^k$. Then

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} [\tau_{t,s}X(\gamma(s))]^i &= \left(\left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} \tau_{t,s}\right)^i{}_j X^j + \left.\frac{\mathrm{d}}{\mathrm{d}s}\right|_{s=t} X^i(\gamma(s)) \\ &= \Gamma^i_{kj}\dot{x}^k X^j + X^i{}_{,j} \left.\frac{\mathrm{d}x^j(\gamma(s))}{\mathrm{d}s}\right|_{s=t},\end{aligned}$$

which is again (16.8) ($X = \xi^i\partial_i$ and the second term gives $\frac{\mathrm{d}\xi^i}{\mathrm{d}t}$).

Definition: If $\nabla_X Y = 0$, then Y is said to be *parallel transported* with respect to X.

Geometrical interpretation of parallel transport: Consider the differential $\mathrm{d}A^i = A^i{}_{,j}\mathrm{d}x^j = A^i(x + \mathrm{d}x) - A^i(x)$. In order that the difference of two vectors be a vector, we have to consider them at the same position. The transport has to be chosen such that for cartesian coordinates there is no change in transporting it. The covariant derivative exactly achieves this.

Definition: Let X be a vector field such that $\nabla_X X = 0$. Then the integral curves of X are called *geodesics.*

In local coordinates x^i the curve is given by (using (12.3) and (13.6)) the requirement $\frac{\mathrm{d}}{\mathrm{d}t}x^i(t) = X^i(x(t))$. Inserting this into (16.8) and using $\frac{\mathrm{d}^2x^i}{\mathrm{d}t^2} = \frac{\mathrm{d}X^i}{\mathrm{d}t}$, we get

$$\ddot{x}^k + \Gamma^k_{ij}\dot{x}^i\dot{x}^j = 0. \tag{16.12}$$

For a vector parallel transported along a geodesic, its length and angle with the geodesic does not change.

16.2 Round trips by parallel transport

Consider (16.8) and denote $\xi^i = v^i$, thus

$$\dot{v}^i = -\Gamma^i_{kj}\dot{x}^k v^j. \tag{16.13}$$

Let $\gamma : [0, 1] \to \mathcal{M}$ be a closed path, wih $\gamma(0) = p = \gamma(1)$. Displace a vector $v_0 \in T_p(\mathcal{M})$ parallel along γ and obtain the field $v(t) = \tau_{t,0}v_0 \in T_{\gamma(t)}(\mathcal{M})$. We assume that the closed path is sufficiently small (such that we can work in the image of some chart), thus we can expand $\Gamma^i_{kj}(x)$ around the point $x(0) = x_0$ on the curve:

$$\Gamma^i_{kj}(x) \simeq \Gamma^i_{kj}(x_0) + (x^\rho - x^\rho_0)\,\frac{\partial}{\partial x^\rho}\Gamma^i_{kj}(x)\bigg|_{x=x_0} + \cdots \tag{16.14}$$

Thus (16.13) is to first order in $(x^k - x^k_0)$:

$$\int_0^t \dot{v}^i\,\mathrm{d}t = v^i(t) - v^i_0 = -\int_0^t \Gamma^i_{kj}\underbrace{v^j}_{\simeq v^j_0}\dot{x}^k\,\mathrm{d}t \approx -\Gamma^i_{kj}(x_0)v^j_0 \underbrace{\int_0^t \dot{x}^k\,\mathrm{d}t}_{x^k(t)-x^k_0},$$

taking only the first term in the expansion of Γ. And hence,

$$v^i(t) = v^i_0 - \Gamma^i_{kj}(x_0)(x^k(t) - x^k_0)v^j_0 + \cdots \tag{16.15}$$

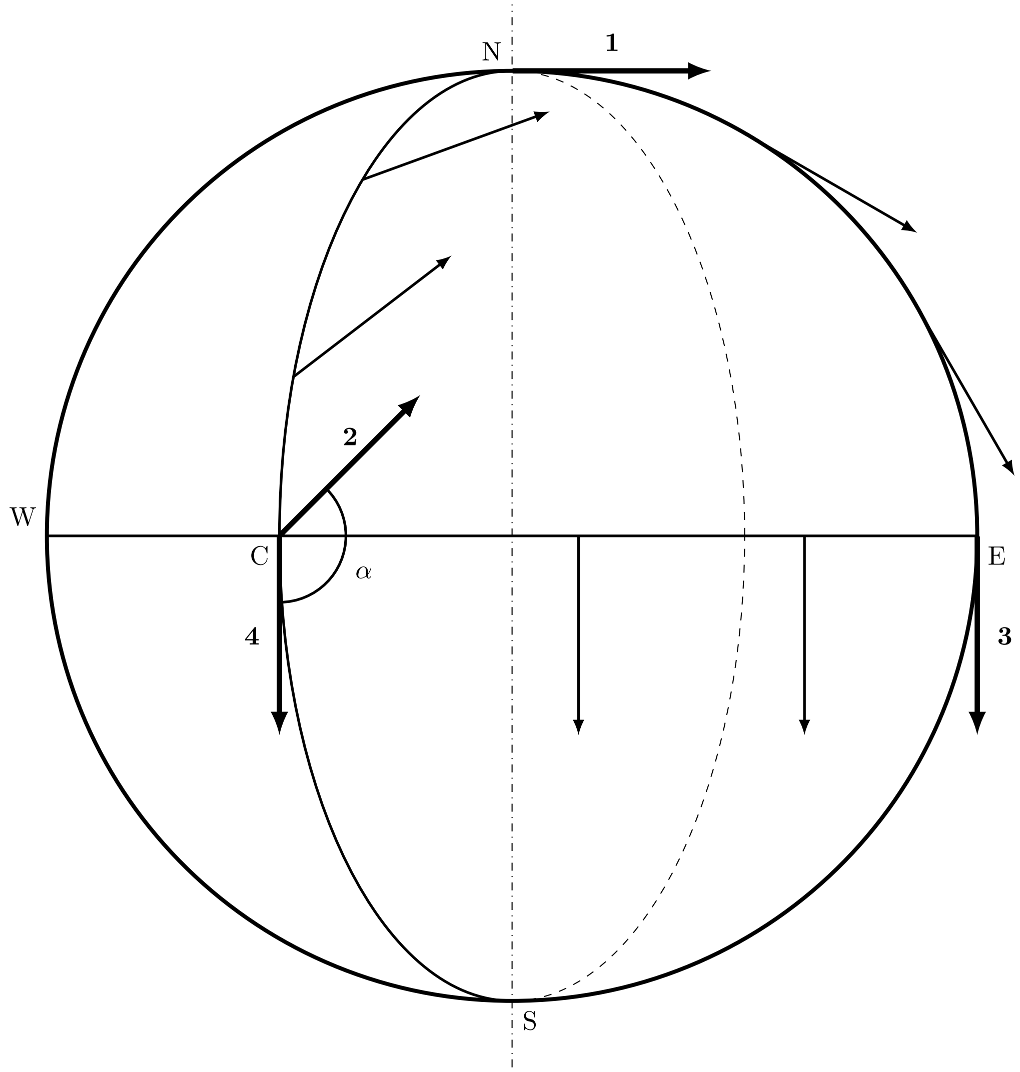


Figure 10: Illustration of the path dependence of parallel transport on a curved space: vector 1 at N can be parallel transported along the geodesic N-S to C, giving rise to vector 2. Alternatively, it can be first transported along the geodesic N-S to E (vector 3) and then along E-W to C to give the vector 4. Clearly these two are different. The angle α between them reflects the curvature of the two-sphere.

By plugging (16.14) and (16.15) into (16.13), we obtain an equation valid to second order:

$$\int_0^1 \dot{v}^i \,\mathrm{d}t = -\int_0^1 \Gamma^i_{kj}\dot{x}^k v^j dt \tag{16.16}$$

$$v^i(1) - v^i_0 \simeq -\int_0^1 \left(\Gamma^i_{kj}(x_0) + (x^\rho - x^\rho_0)\frac{\partial}{\partial x^\rho}\Gamma^i_{kj}(x_0) + \cdots\right) \times$$

$$\left(v^j_0 - \Gamma^j_{\tilde{k}\tilde{j}}(x_0)(x^{\tilde{k}}(t) - x^{\tilde{k}}_0)v^{\tilde{j}}_0 + \cdots\right)\dot{x}^k \,\mathrm{d}t.$$

Multiplying out and discarding terms of third order or higher in $x^k - x_0^k$, we get:

$$v^i(1) \simeq v_0^i - \Gamma^i_{kj}(x_0)v_0^j \underbrace{\int_0^1 \dot{x}^k \,\mathrm{d}t}_{\substack{x^k(1)-x^k(0)\\=0}} - \left[\frac{\partial}{\partial x^\rho}\Gamma^i_{kj}(x_0) - \Gamma^i_{k\tilde{j}}(x_0)\Gamma^{\tilde{j}}_{\rho j}(x_0)\right] v_0^j \int_0^1 (x^\rho - x_0^\rho)\dot{x}^k \,\mathrm{d}t.$$

Since we are considering a closed path ($\int_0^1 \dot{x}\, dt = x^k(1) - x^k(0) = 0$),

$$\Delta v^i = v^i(1) - v^i(0) = -\left[\frac{\partial}{\partial x^\rho}\Gamma^i_{kj}(x_0) - \Gamma^i_{kl}(x_0)\Gamma^l_{\rho j}(x_0)\right] v_0^j \int_0^1 x^\rho \dot{x}^k \,\mathrm{d}t,$$

with

$$\oint_0^1 x^\rho \dot{x}^k \,\mathrm{d}t = \underbrace{\oint_0^1 \frac{\mathrm{d}}{\mathrm{d}t}(x^\rho x^k)\,\mathrm{d}t}_{=0} - \oint_0^1 \dot{x}^\rho x^k \,\mathrm{d}t = -\oint_0^1 \dot{x}^\rho x^k \,\mathrm{d}t,$$

antisymmetric in (ρ, k). Then

$$\Delta v^i = -\frac{1}{2}\underbrace{\left[\frac{\partial}{\partial x^\rho}\Gamma^i_{kj} - \Gamma^i_{kl}\Gamma^l_{\rho j} - \frac{\partial}{\partial x^k}\Gamma^i_{\rho j} + \Gamma^i_{\rho l}\Gamma^l_{kj}\right](x_0)}_{-R^i_{jk\rho}} v_0^j \int_0^1 x^\rho \dot{x}^k \,\mathrm{d}t,$$

$$\Delta v^i = \frac{1}{2}R^i_{jk\rho}(x_0)v_0^j \int_0^1 x^\rho \dot{x}^k \,\mathrm{d}t. \tag{16.17}$$

We shall see that $R^i_{jk\rho}$ is the curvature tensor.

$$R^i_{jk\rho} = \frac{\partial}{\partial x^k}\Gamma^i_{\rho j} - \frac{\partial}{\partial x^\rho}\Gamma^i_{kj} + \Gamma^l_{\rho j}\Gamma^i_{kl} - \Gamma^l_{kj}\Gamma^i_{\rho l} \tag{16.18}$$

Thus an arbitrary vector v^i will not change when parallel transported around an arbitrary small closed curve at x_0 if and only if $R^i_{jk\rho}$ vanishes at x_0.

16.3 Covariant derivatives of tensor fields

The parallel transport is extended to tensors by means of the requirements:

$$\tau_{s,t}(T \otimes S) = (\tau_{s,t}T) \otimes (\tau_{s,t}S),$$

$$\tau_{s,t}\,\mathrm{tr}(T) = \mathrm{tr}(\tau_{s,t}T),$$

$$\tau_{s,t}\, c = c \qquad (c \in \mathbb{R}).$$

For e.g. a covariant vector ω, $\langle\tau_{s,t}\,\omega, \tau_{s,t}\,X\rangle_{\gamma(s)} = \langle\omega, X\rangle_{\gamma(t)}$ and for a tensor of type $\binom{1}{1}$: $\tau_{s,t}\,T(\tau_{s,t}\,\omega, \tau_{s,t}\,X) = T(\omega, X)$. In components:

$$(\tau_{s,t}T)^{i}{}_{k} = T^{\alpha}{}_{\beta}(\tau_{s,t})^{i}{}_{\alpha}(\tau_{s,t})_{k}{}^{\beta} \tag{16.19}$$

($\tau_i{}^k$ is inverse transpose of $\tau^i{}_k$). The covariant derivative ∇_X (X vector field, T tensor field) associated to τ is

$$(\nabla_X T)_p = \frac{\mathrm{d}}{\mathrm{d}t}\tau_{0,t}T_{\gamma(t)}\bigg|_{t=0}, \tag{16.20}$$

with $\gamma(t)$ any curve with $\gamma(0) = p$ and $\dot\gamma(0) = X_p$ (generalization of (16.10)).

Properties of the covariant derivative:

(a) ∇_X is a linear map from tensor fields to tensor fields of the same type $\binom{r}{r}$,

(b) $\nabla_X f = Xf$,

(c) $\nabla_X(\mathrm{tr}\,T) = \mathrm{tr}(\nabla_X T)$,

(d) $\nabla_X(T \otimes S) = (\nabla_X T) \otimes S + T \otimes (\nabla_X S)$.

This follows from the properties of $\tau_{s,t}$. For a 1-form ω we have:

$$\begin{aligned}(\nabla_X\omega)(Y) &= \mathrm{tr}(\nabla_X\omega \otimes Y)\\ &= \mathrm{tr}(\nabla_X(\omega \otimes Y)) - \mathrm{tr}(\omega \otimes \nabla_X Y)\\ &= \nabla_X\,\mathrm{tr}(\omega \otimes Y) - \omega(\nabla_X Y)\\ &= X\omega(Y) - \omega(\nabla_X Y).\end{aligned} \tag{16.21}$$

General differentiation rule for a tensor field of type $\binom{1}{1}$:

$$\boxed{(\nabla_X T)(\omega, Y) = XT(\omega, Y) - T(\nabla_X\omega, Y) - T(\omega, \nabla_X Y)} \tag{16.22}$$

Due to (a)-(d), the operation ∇_X is completely determined by its action on vector fields Y, which are the affine connections (see (16.1) and (16.2)).

16.4 Local coordinate expressions for covariant derivative

Let $T \in T^q_p(\mathcal{U})$ be a tensor of rank (p, q) with local coordinates $(x^1, \dots, x^n)$ valid in a region $\mathcal{U}$. We have $T^{i_1\dots i_p}{}_{j_1\dots j_q}\,\partial_{i_1} \otimes \dots \otimes \partial_{i_p} \otimes \mathrm{d}x^{j_1} \otimes \dots \otimes \mathrm{d}x^{j_q}$ and $X = X^k\partial_k$. Let us use

$$XT^{i_1\dots i_p}{}_{j_1\dots j_q} = X^k T^{i_1\dots i_p}{}_{j_1\dots j_q,k} \tag{16.23}$$

and write (16.2):

$$\nabla_X(\partial_i) = X^k\nabla_{\partial_k}\partial_i = X^k\Gamma^l_{ki}\partial_l\ . \tag{16.24}$$

Moreover,

$$(\nabla_X \mathrm{d}x^j)(\partial_i) \overset{(16.21)}{=} \underbrace{X \underbrace{\langle \mathrm{d}x^j, \partial_i \rangle}_{\delta^j{}_i}}_{0} - \langle \mathrm{d}x^j, \nabla_X \partial_i \rangle$$

$$= -X^k \Gamma^j_{ki} ,$$

$$\text{or} \quad \nabla_X \mathrm{d}x^j = -X^k \Gamma^j_{ki} \mathrm{d}x^i. \tag{16.25}$$

Using (16.23), (16.24) and (16.25) for $\omega^j = \mathrm{d}x^j$, $Y_i = \partial_i$ we obtain the following expression for $\nabla_X T$:

$$\boxed{\begin{aligned} T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q ; k} = {} & T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q , k} + \Gamma^{i_1}_{kl} T^{l i_2 \ldots i_p}{}_{j_1 \ldots j_q} + \ldots + \Gamma^{i_p}_{kl} T^{i_1 \ldots i_{p-1} l}{}_{j_1 \ldots j_q} \\ & - \Gamma^l_{k j_1} T^{i_1 \ldots i_p}{}_{l j_2 \ldots j_q} - \ldots - \Gamma^l_{k j_q} T^{i_1 \ldots i_p}{}_{j_1 \ldots j_{q-1} l}. \end{aligned}} \tag{16.26}$$

Examples:

- Contravariant and covariant vector fields:

$$\xi^i{}_{;k} = \xi^i{}_{,k} + \Gamma^i_{kl} \xi^l,$$

$$\eta_{i;k} = \eta_{i,k} - \Gamma^l_{ki} \eta_l,$$

- Kronecker tensor:

$$\delta^i{}_{j;k} = 0,$$

- Tensor $\binom{1}{1}$:

$$T^i{}_{k;r} = T^i{}_{k,r} + \Gamma^i_{rl} T^l{}_k - \Gamma^l_{rk} T^i{}_l.$$

The covariant derivative of a tensor is again a tensor. Consider the covariant derivative of the metric $g_{\mu\nu}$:

$$g_{\mu\nu;\lambda} = \frac{\partial g_{\mu\nu}}{\partial x^\lambda} - \Gamma^\rho_{\lambda\mu} g_{\rho\nu} - \Gamma^\rho_{\lambda\nu} g_{\rho\mu}. \tag{16.27}$$

Inserting into this the expressions of $\Gamma^\rho_{\lambda\mu}$ given by (9.11) leads us to

$$\boxed{g_{\mu\nu;\lambda} = 0.} \tag{16.28}$$

This is not surprising since $g_{\mu\nu;\lambda}$ vanishes in locally inertial coordinates and being a tensor it is then zero in all systems.

Covariance principle: Write the appropriate special relativistic equations that hold in the absence of gravitation, replace $\eta_{\alpha\beta}$ by $g_{\alpha\beta}$, and replace all derivatives with covariant derivatives $(, \to ;)$. The resulting equations will be generally covariant and true in the presence of gravitational fields.

17 Curvature and torsion of an affine connection, Bianchi identities

Let an affine connection be given on $\mathcal{M}$, let X, Y, Z be vector fields.

Definition:

$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] \tag{17.1}$$

$$R(X,Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]} \tag{17.2}$$

$T(X,Y)$ is antisymmetric and f-linear in X, Y and then defines a tensor of type $\binom{1}{2}$ through: $(\omega, X, Y) \to \langle \omega, T(X,Y) \rangle$ is thus a $\binom{1}{2}$ tensor field called the *torsion tensor*.

f-linearity:

$$T(fX, gY) = fgT(X,Y) \qquad f, g, \in \mathcal{F}(\mathcal{M}).$$

In local coordinates, the components of the torsion tensor are given by:

$$\begin{aligned} T^k{}_{ij} = \left\langle \mathrm{d}x^k, T(\partial_i, \partial_j) \right\rangle &= \left\langle \mathrm{d}x^k, \underbrace{\nabla_{\partial_i}\partial_j}_{=\Gamma^l_{ij}\partial_l} - \nabla_{\partial_j}\partial_i - \underbrace{[\partial_i, \partial_j]}_{=0} \right\rangle \\ &= \Gamma^k_{ij} - \Gamma^k_{ji} \end{aligned} \tag{17.3}$$

(using that $\left\langle \mathrm{d}x^k, \partial_l \right\rangle = \delta^k{}_l$). In particular, we have $T^k{}_{ij} = 0 \Leftrightarrow \Gamma^k_{ij} = \Gamma^k_{ji}$.

$R(X,Y) = -R(Y,X)$ is antisymmetric in X,Y. The vector field $R(X,Y)Z$ is f-linear in X, Y, Z: $(R(fX, gY)\, hZ = fgh\, R(X,Y)Z;\ f, g, h \in \mathcal{F}(\mathcal{M}))$. R determines a tensor of type $\binom{1}{3}$: the *Riemann tensor* or *curvature tensor*.

$$(\omega, Z, X, Y) \to \langle \omega, R(X,Y)Z \rangle \equiv R^i{}_{jkl}\omega_i Z^j X^k Y^l$$

In components with respect to local coordinates:

$$\boxed{\begin{aligned} R^i{}_{jkl} &= \left\langle \mathrm{d}x^i, R(\partial_k, \partial_l)\partial_j \right\rangle = \left\langle \mathrm{d}x^i, (\nabla_{\partial_k}\nabla_{\partial_l} - \nabla_{\partial_l}\nabla_{\partial_k})\partial_j \right\rangle^{12} \\ &= \left\langle \mathrm{d}x^i, \nabla_{\partial_k}(\Gamma^s_{lj}\partial_s) - \nabla_{\partial_l}(\Gamma^s_{kj}\partial_s) \right\rangle \\ &= \Gamma^i_{lj,k} - \Gamma^i_{kj,l} + \Gamma^s_{lj}\Gamma^i_{ks} - \Gamma^s_{kj}\Gamma^i_{ls}. \end{aligned} \qquad (17.4)}$$

[12] Notice that $\nabla_{\underbrace{[\partial_k, \partial_l]}_{=0}}\partial_j = 0$.

Eq. (17.4) is exactly the the same as defined in (16.18). It is antisymmetric in the last two indices: $R^i{}_{jkl} = -R^i{}_{jlk}$.

Definition: The *Ricci tensor* is the following contraction of the curvature tensor:

$$\boxed{R_{jl} \equiv R^i{}_{jil} = \Gamma^i_{lj,i} - \Gamma^i_{ij,l} + \Gamma^s_{lj}\Gamma^i_{is} - \Gamma^s_{ij}\Gamma^i_{ls}} \tag{17.5}$$

The *scalar curvature* is the trace of the Ricci tensor:

$$\boxed{\mathcal{R} \equiv g^{lj}R_{jl} = R^l{}_l} \tag{17.6}$$

Example: For a pseudo-Riemannian manifold the connection coefficients are given by (9.11). Consider a two-sphere (which is a pseudo Riemannian manifold) with the metric $\mathrm{d}s^2 = a^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2)$, then:

$$g_{\theta\phi} = a^2 \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix}, \qquad g^{\theta\phi} = \frac{1}{a^2}\begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\sin^2\theta} \end{pmatrix}.$$

The non-zero Γ are:

$$\Gamma^\theta_{\phi\phi} = -\sin\theta\cos\theta,$$

$$\Gamma^\phi_{\theta\phi} = \Gamma^\phi_{\phi\theta} = \cot\theta.$$

The Riemann tensor is given by

$$\begin{aligned} R^\theta_{\phi\theta\phi} &= \partial_\theta\Gamma^\theta_{\phi\phi} - \partial_\phi\Gamma^\theta_{\theta\phi} + \Gamma^\theta_{\theta\lambda}\Gamma^\lambda_{\phi\theta} - \Gamma^\theta_{\phi\lambda}\Gamma^\lambda_{\theta\phi} \\ &= (\sin^2\theta - \cos^2\theta) - 0 + 0 - (-\sin\theta\cos\theta)\cot\theta \\ &= \sin^2\theta. \end{aligned}$$

The Ricci tensor has the following components:

$$\begin{aligned} R_{\phi\phi} &= R^\theta_{\phi\theta\phi} + \underbrace{R^\phi_{\phi\phi\phi}}_{=0} = \sin^2\theta, \\ R_{\theta\theta} &= 1, \\ R_{\theta\phi} &= R_{\phi\theta} = 0. \end{aligned}$$

The Ricci scalar is

$$\begin{aligned} \mathcal{R} &= \underbrace{g^{\theta\theta}}_{\frac{1}{a^2}}\underbrace{R_{\theta\theta}}_{1} + \underbrace{g^{\phi\phi}}_{\frac{1}{a^2\sin^2\theta}}\underbrace{R_{\phi\phi}}_{\sin^2\theta} + g^{\theta\phi}\underbrace{R_{\theta\phi}}_{0} + g^{\phi\theta}\underbrace{R_{\phi\theta}}_{0} \\ &= \frac{1}{a^2} + \frac{1}{a^2\sin^2\theta}\sin^2\theta \\ &= \frac{2}{a^2}. \end{aligned}$$

The Ricci scalar is constant over this two-sphere and positive, thus the sphere is “positively curved”. [13] [14] [15]

17.1 Bianchi identities for the special case of zero torsion

X, Y and Z are vector fields, then

$$R(X,Y)Z + \text{cyclic} = 0 \qquad \text{(1st Bianchi identity)}, \tag{17.7}$$

$$(\nabla_X R)(Y,Z) + \text{cyclic} = 0 \qquad \text{(2nd Bianchi identity)}. \tag{17.8}$$

Proof of the 1st identity: Torsion $= 0 \Rightarrow \nabla_X Y - \nabla_Y X = [X,Y]$. Then

$$\begin{aligned}
&(\nabla_X\nabla_Y - \nabla_Y\nabla_X)Z + (\nabla_Z\nabla_X - \nabla_X\nabla_Z)Y \\
&\qquad + (\nabla_Y\nabla_Z - \nabla_Z\nabla_Y)X - \nabla_{[X,Y]}Z - \nabla_{[Z,X]}Y - \nabla_{[Y,Z]}X \\
&= \nabla_X(\nabla_Y Z - \nabla_Z Y) - \nabla_{[Y,Z]}X + \text{cyclic} \\
&= [X,[Y,Z]] + \text{cyclic} \\
&= 0 \qquad \text{due to the Jacobi identity (13.2).}
\end{aligned}$$

(See textbooks for proof of the 2nd Bianchi identity.)

[13] For a position independent metric (e.g. Cartesian coordinates) the Riemann tensor (and thus the scalar curvature) vanishes as the Γ vanish.

[14] For a plane with polar coordinates we get a position dependent metric $\begin{pmatrix}1 & 0\\ 0 & r^2\end{pmatrix}$: $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$ and thus the Γ do not vanish. However the curvature vanishes.

[15] The curvature does not depend on the choice of coordinates.

18 Riemannian connections

Metric: Let $\mathcal{M}$ be equipped with a pseudo-Riemannian metric: a symmetric, non-degenerate tensor field: $g(X,Y) = \langle X, Y \rangle$ of type $\binom{0}{2}$.

- Non-degenerate means that, for any point $p \in \mathcal{M}$, $X, Y \in T_p$, one has $g_p(X,Y) = 0 \ \forall Y \in T_p \Rightarrow X = 0$.
- In components, $\langle X, Y \rangle = g_{ik} X^i Y^k$, $g_{ik} = g_{ki}$ (symmetric) and $\det g_{ik} \neq 0$.
- With the metric we can lower and raise indices:

$$\tilde{X}_i = g_{ik} X^k, \qquad \tilde{\omega}^i = g^{ik} \omega_k,$$

  where g^{ik} denotes the inverse of g_{ik}. It also works for tensor fields of different types: $T^i{}_k = T_{lk} g^{il} = T^{il} g_{lk}$.
- Given a basis $(e_1, \dots, e_n)$ of T_p, the covectors of the dual basis $(e^1, \dots, e^n)$ become themselves vectors; indeed $e_i = g_{ij} e^j$.

Riemann connection: The metric tensor g at a point p in $\mathcal{M}$ is a symmetric $\binom{0}{2}$ tensor. It assigns a magnitude $\sqrt{|g(X,X)|}$ to each vector X on $T_p(\mathcal{M})$, denoted by $\mathrm{d}(X)$ and defines the angle between any two vectors X, Y ($\neq 0$) on $T_p(\mathcal{M})$ via

$$a(X,Y) = \arccos\left(\frac{g(X,Y)}{\mathrm{d}(X)\mathrm{d}(Y)}\right). \tag{18.1}$$

If $a(X,Y) = \frac{\pi}{2}$ then X and Y are orthogonal. Further observations:

- The length of a curve with tangent vector X between t_1 and t_2 is $L(t_1, t_2) = \int\limits_{t_1}^{t_2} \mathrm{d}(X)\, \mathrm{d}t$.
- If (e_a) is a basis of $T_p(\mathcal{M})$, the components of g with respect to this basis are $g_{ab} = g(e_a, e_b)$.
- Like in special relativity we classify vectors at a point as timelike ($g(X,X) > 0$), null ($g(X,X) = 0$) and space like ($g(X,X) < 0$).

Definition: let $(\mathcal{M}, g)$ be a pseudo-Riemannian manifold. An affine connection is a *metric connection* if parallel transport along any smooth curve γ on $\mathcal{M}$ preserve the inner product: for autoparallel fields $X(t)$, $Y(t)$ (see (16.7)), $g_{\gamma(t)}(X(t), Y(t))$ is independent of t along γ.

Theorem: an affine connection ∇ is metric if and only if (no proof)

$$\nabla g = 0. \tag{18.2}$$

Eq. (18.2) is equivalent for $(g(Y,Z))$ to

$$\nabla_X g = 0 = X g(Y,Z) - g(\nabla_X Y, Z) - g(Y, \nabla_X Z)$$

or

$$Xg(Y,Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z). \tag{18.3}$$

Theorem: For every pseudo-Riemannian manifold $(\mathcal{M}, g)$, there exists a unique affine connection such that

(a) ∇ has vanishing torsion (∇ is symmetric),

(b) ∇ is metric.

Proof: $T = 0$ (vanishing torsion) means $\nabla_X Y = \nabla_Y X + [X,Y]$. Inserting this into (18.3) (and the linearity of g) gives

$$Xg(Y,Z) = g(\nabla_Y X, Z) + g([X,Y], Z) + g(Y, \nabla_X Z). \tag{18.4}$$

By cyclic permutations one obtains as well

$$Yg(Z,X) = g(\nabla_Z Y, X) + g([Y,Z], X) + g(Z, \nabla_Y X), \tag{18.5}$$

$$Zg(X,Y) = g(\nabla_X Z, Y) + g([Z,X], Y) + g(X, \nabla_Z Y). \tag{18.6}$$

Taking the linear combination (18.5) + (18.6) - (18.4), we get (*Koszul formula*):

$$\begin{aligned} 2g(\nabla_Z Y, X) &= -Xg(Y,Z) + Yg(Z,X) + Zg(X,Y) \\ &\quad -g([Z,X],Y) - g([Y,Z],X) + g([X,Y],Z). \end{aligned} \tag{18.7}$$

The right hand side is independent of ∇. Since g is non-degenerate, the uniqueness of ∇ follows from (18.7).

Definition: the unique connection on $(\mathcal{M}, g)$ from the above theorem is called the *Riemannian* or *Levi-Civita connection.*

We determine the Christoffel symbols for the Riemannian connection in a given chart $(\mathcal{U}, x^1, \ldots, x^n)$. For this purpose we take $X = \partial_k$, $Y = \partial_j$, $Z = \partial_i$ in (18.7) and we use $[\partial_i, \partial_j] = 0$ as well as $\langle \partial_i, \partial_j \rangle = g_{ij}$. The result is

$$\langle \nabla_{\partial_i} \partial_j, \partial_k \rangle = \Gamma^l_{ij} \underbrace{\langle \partial_l, \partial_k \rangle}_{g_{lk}},$$

$$2\Gamma^l_{ij} g_{lk} = -\partial_k \underbrace{\langle \partial_i, \partial_j \rangle}_{g_{ij}} + \partial_j \underbrace{\langle \partial_i, \partial_k \rangle}_{g_{ik}} + \partial_i \underbrace{\langle \partial_k, \partial_j \rangle}_{g_{kj}},$$

or

$$g_{lk}\Gamma^l_{ij} = \frac{1}{2}(g_{kj,i} + g_{ik,j} - g_{ji,k}). \tag{18.8}$$

g^{ij} denoting the inverse matrix of g_{ij}, we obtain

$$\Gamma^l_{ij} = \frac{1}{2}g^{lk}(g_{kj,i} + g_{ik,j} - g_{ji,k}), \tag{18.9}$$

which is exactly equation (9.11).

Properties of the Riemannian connection:

(i) The inner product of any two vectors remains constant upon parallel transporting them along any curve γ $(g(X,Y)_{\gamma(t)} = g(X,Y)_{\gamma(0)})$.

(ii) The covariant derivative commutes with raising or lowering indices, e.g. $T^i_{k;l} = (g_{km}T^{im})_{;l} = g_{km}T^{im}{}_{;l}$, because $g_{km;l} = 0$.

<u>Riemann tensor:</u> the curvature tensor of a Riemannian connection has the following additional symmetry properties (without proof):

$$\langle R(X,Y)Z,U\rangle = -\langle R(X,Y)U,Z\rangle\,, \tag{18.10}$$

$$\langle R(X,Y)Z,U\rangle = \langle R(Z,U)X,Y\rangle\,. \tag{18.11}$$

In coordinate expression the Riemann tensor satisfies the following symmetries:

$$R^i{}_{jkl} = -R^i{}_{jlk} \qquad \text{is always the case,} \tag{18.12}$$

$$\sum_{(jkl)} R^i{}_{jkl} = 0 \qquad \text{1st Bianchi identity,} \tag{18.13}$$

$$\sum_{(klm)} R^i{}_{jkl;m} = 0 \qquad \text{2nd Bianchi identity.} \tag{18.14}$$

Eqs. (18.13) and (18.14) are valid for vanishing torsion. Here $\sum_{(jkl)}$ denotes the cyclic sum. Additionally,

$$R_{ijkl} = -R_{jikl}, \tag{18.15}$$

$$R_{ijkl} = R_{klij}, \tag{18.16}$$

for the Riemannian connection with $R_{ijkl} = g_{is}R^s{}_{jkl}$.

<u>Ricci and Einstein tensor</u>

$$R_{ik} = R^j{}_{ijk} \qquad \text{Ricci tensor} \qquad (18.17)$$

$$\mathcal{R} = R^i{}_i \qquad \text{scalar curvature} \qquad (18.18)$$

$$G_{ik} = R_{ik} - \frac{1}{2}\mathcal{R}g_{ik} \qquad \text{Einstein tensor} \qquad (18.19)$$

By symmetry, $R_{ik} = R_{ki}$, $G_{ik} = G_{ki}$ and

$$R^k{}_{i;k} = \frac{1}{2}\mathcal{R}_{;i}, \tag{18.20}$$

$$G^k{}_{i;k} = 0, \tag{18.21}$$

which are the contracted 2nd Bianchi identity.

<u>Proof:</u> $R_{ik} = g^{jl}R_{lijk} = g^{jl}R_{jkli}$, 2nd Bianchi identity gives:

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

Then we take the (ik)-trace:

$$R_{jl;m} + \underbrace{R^i{}_{jlm;i}}_{-g^{ik}R_{jklm;i}} - R_{jm;l} = 0,$$

$$R^j{}_{l;m} - g^{ik}R^j{}_{klm;i} - R^j{}_{m;l} = 0,$$

(jm)-trace:

$$\underbrace{R^j{}_{l;j} + g^{ik}R_{kl;i}}_{2R^j{}_{l;j}} - R_{;l} = 0.$$

$$\Rightarrow (18.20)$$

For (18.21):

$$G^k{}_i = R^k{}_i - \frac{1}{2}g^k{}_i\mathcal{R} = R^k{}_i - \frac{1}{2}\delta^k{}_i\mathcal{R}$$

$$G^k{}_{i;k} = R^k{}_{i;k} - \frac{1}{2}(\delta^k_i\mathcal{R})_{;k} = \underbrace{\frac{1}{2}\mathcal{R}_{;i}}_{(18.20)} - \frac{1}{2}\mathcal{R}_{;i} = 0$$

Without proof in n dimensions, the Riemann tensor has $c_n = \frac{n^2(n^2-1)}{12}$ independent components ($c_1 = 0$, $c_2 = 1$, $c_3 = 6$, $c_4 = 20$).

Part V

General Relativity

19 Physical laws with gravitation

19.1 Mechanics

The physical laws are relations among tensors (scalars and vectors being tensors of rank 0 and 1 respectively). Thus the physical laws read the same in all coordinate systems (provided the physical quantities are transformed suitably) and satisfy *general covariance* (same form). Practically, this means that from the special relativity laws that hold in absence of gravitation, we have to replace $\eta_{\alpha\beta}$ by $g_{\alpha\beta}$ and replace derivation by covariant derivation.

In an inertial system, we have the equation of motion (see (4.3))

$$m\frac{\mathrm{d}u^\alpha}{\mathrm{d}\tau} = f^\alpha. \tag{19.1}$$

According to the equivalence principle, (19.1) holds in a local IS. f^α does not contain gravitational forces as they would vanish in a local IS. We transform it to general KS (coordinate system), then the Lorentz vector f^α gets transformed to $f^\mu = \frac{\partial x^\mu}{\partial \xi^\alpha} f^\alpha$ (ξ^α is in local IS, x^μ is in KS). Equation (19.1) holds in a local IS reads then

$$m \underbrace{\frac{\mathrm{D}u^\mu}{\mathrm{d}\tau}}_{\substack{\text{covariant derivative}\\ \text{given in (16.8)}}} = f^\mu, \tag{19.2}$$

($\nabla_X \to \frac{\mathrm{D}}{\mathrm{d}\tau}$) with $\xi^\mu \to u^\mu$ ($\mathrm{d}t \to \mathrm{d}\tau$ and $\frac{\mathrm{d}x^i}{\mathrm{d}\tau} = u^i$) and thus

$$\frac{\mathrm{D}u^\mu}{\mathrm{d}\tau} = \frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} + \Gamma^\mu_{\nu\lambda} u^\nu u^\lambda.$$

Then equation (19.2) reads

$$m\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = f^\mu - m\Gamma^\mu_{\nu\lambda} u^\nu u^\lambda. \tag{19.3}$$

We see that on the right hand side there are now gravitational forces appearing explicitly (via $\Gamma^\mu_{\nu\lambda}$). Equation (19.3) (or (19.2)) is covariant (it has the same form in all coordinate systems) and reduces for $g_{\mu\nu} \to \eta_{\alpha\beta}$ (thus $\Gamma^\mu_{\nu\lambda} = 0$) to equation (19.1) (in a local IS). The components of u^μ are not independent but satisfy the condition $g_{\mu\nu} u^\mu u^\nu = c^2$.

19.2 Electrodynamics

According to the equivalence principle, Maxwell's equations (see (6.4) and (6.5))

$$\partial_\alpha F^{\alpha\beta} = \frac{4\pi}{c} j^\beta \qquad \text{and} \qquad \epsilon^{\alpha\beta\gamma\delta} \partial_\beta F_{\gamma\delta} = 0$$

are valid in a local IS. Applying the covariance principle, they become as follows in a general KS:

$$F^{\mu\nu}{}_{;\nu} = \frac{4\pi}{c} j^\mu \qquad \text{and} \qquad \epsilon^{\mu\nu\lambda\kappa} F_{\lambda\kappa;\nu} = 0, \tag{19.4}$$

provided that going from coordinates ξ^α in a local IS to the KS coordinates x^μ we have

$$j^\alpha \to j^\mu = \frac{\partial x^\mu}{\partial \xi^\alpha} j^\alpha \qquad \text{and} \qquad F^{\mu\nu} = \frac{\partial x^\mu}{\partial \xi^\alpha}\frac{\partial x^\nu}{\partial \xi^\beta} F^{\alpha\beta}.$$

Gravity enters via the $\Gamma^\mu_{\nu\lambda}$ in the covariant derivative. The continuity equation $\partial_\alpha j^\alpha = 0$ translates to $j^\mu{}_{;\mu} = 0$. It can be shown that in the homogeneous equation the terms with Γ vanish. Thus the covariant derivative reduces to the ordinary derivative (,) [16].

19.3 Energy-momentum tensor

For an ideal fluid, given by (in a local IS)

$$T^{\mu\nu} = \left(\rho + \frac{p}{c^2}\right) u^\mu u^\nu - \eta^{\mu\nu} p, \tag{19.8}$$

with

- u^μ: four-velocity,
- ρ: proper energy density,
- p: pressure of the fluid.

In a KS this becomes

$$T^{\mu\nu} = \left(\rho + \frac{p}{c^2}\right) u^\mu u^\nu - g^{\mu\nu} p. \tag{19.9}$$

In the IS the conservation law implies $T^{\mu\nu}{}_{,\nu} = 0$ and in the KS $T^{\mu\nu}{}_{;\nu} = 0$ (explicitely, $T^{\mu\nu}{}_{;\nu} = T^{\mu\nu}{}_{,\nu} + \Gamma^\mu_{\nu\lambda} T^{\nu\lambda} + \Gamma^\nu_{\nu\lambda} T^{\mu\lambda} = 0$). With (19.5), $\Gamma^\nu_{\nu\lambda} = \frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}}{\partial x^\lambda}$, we get instead

$$T^{\mu\nu}{}_{;\nu} = \frac{1}{\sqrt{g}}\frac{\partial \sqrt{g} T^{\mu\nu}}{\partial x^\nu} + \Gamma^\mu_{\nu\lambda} T^{\nu\lambda} = 0. \tag{19.10}$$

This is no longer a conservation law, as we cannot form any constant of motion from (19.10). This should also not be expected, since the system under consideration can exchange energy and momentum with the gravitational field.

[16] $g = \det(g_{ik}) = \epsilon^{i_1\dots i_n} g_{1i_1}\dots g_{ni_n}$. Consider $\frac{\partial g}{\partial x^l} = \sum_{k=1}^n \epsilon^{i_1\dots i_n} g_{1i_1}\dots \frac{\partial g_{ki_k}}{\partial x^l}\dots g_{ni_n}$ and use $\frac{\partial g_{ki_k}}{\partial x^l} = \frac{\partial g_{km}}{\partial x^l}\delta^m{}_{i_k} = \frac{\partial g_{km}}{\partial x^l} g^{mr} g_{ri_k}$. Due to the antisymmetry of ϵ, only the term $r = k$ survives. Thus

$$\frac{\partial g}{\partial x^l} = \frac{\partial g_{km}}{\partial x^l} g^{mk} g.$$

Plugging this into the definition of Γ^k_{kl} (one contraction):

$$\Gamma^k_{kl} = \frac{g^{km}}{2}\left(\frac{\partial g_{mk}}{\partial x^l} + \underbrace{\frac{\partial g_{ml}}{\partial x^k} - \frac{\partial g_{kl}}{\partial x^m}}_{\substack{\text{vanish by interchanging}\\ (m\leftrightarrow k)}}\right) = \frac{g^{km}}{2}\frac{\partial g_{mk}}{\partial x^l} = \frac{\partial \ln\sqrt{g}}{\partial x^l} = \frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}}{\partial x^l}. \tag{19.5}$$

With (19.5) one can show that the inhomogeneous Maxwell equation in KS can be written as

$$\frac{1}{\sqrt{g}}\frac{\partial(\sqrt{g}F^{\mu\nu})}{\partial x^\nu} = \frac{4\pi}{c^2} j^\mu, \tag{19.6}$$

and the continuity equation: $j^\mu{}_{;\mu} = 0$ becomes

$$\frac{\partial(\sqrt{g} j^\mu)}{\partial x^\mu} = 0. \tag{19.7}$$

20 Einstein's field equations

The field equations cannot be derived by using the covariance principle, since there is no equivalent equation in a local IS. We have to make some requirements/assumptions.

Requirements:

- The Newtonian limit is well confirmed through all observations: $\Delta\phi = 4\pi G\rho$.
- From the Newtonian limit of the equation of motion for a particle we derived (equation (9.17)) $g_{00} \approx 1 + 2\frac{\phi}{c^2}$.
- The non-relativistic limit should then be

$$\Delta g_{00} = \frac{8\pi G}{c^4} T_{00}, \tag{20.1}$$

with $T_{00} \approx \rho c^2$ (other T_{ij} are small).

Thus a generalization should lead to something of type $G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}$ where $G_{\mu\nu}$ has to satisfy the following requirements:

(1) $G_{\mu\nu}$ is a tensor ($T_{\mu\nu}$ *is* tensor).

(2) $G_{\mu\nu}$ has the "dimension" of a second derivative. If we assume that no new dimensional constant enter in $G_{\mu\nu}$ then it has to be a linear combination of terms which are either second derivatives of the metric $g_{\mu\nu}$ or quadratic in the first derivative of $g_{\mu\nu}$.

(3) Since $T_{\mu\nu}$ is symmetric, $G_{\mu\nu}$ also has to be symmetric and due to the fact that $T_{\mu\nu}$ is covariantly conserved, i.e. $T^{\mu\nu}{}_{;\nu} = 0$, it follows that $G_{\mu\nu}$ must satisfy $G_{\mu\nu} = G_{\nu\mu}$ and $G_{\mu\nu}{}^{;\nu} = 0$.

(4) For a weak stationary field we shall get (20.1), thus $G_{00} \simeq \Delta g_{00}$.

Conditions (1)-(4) determine $G_{\mu\nu}$ uniquely. (1) and (2) imply that $G_{\mu\nu}$ has to be a linear combination

$$G_{\mu\nu} = aR_{\mu\nu} + b\mathcal{R}g_{\mu\nu} \tag{20.2}$$

of $R_{\mu\nu}$, the Ricci tensor, and $\mathcal{R}$, the Ricci scalar[17]. The symmetry of $G_{\mu\nu}$ is automatically satisfied. The contracted Bianchi identity (18.20), (18.21) suggests that $G_{\mu\nu}{}^{;\nu} = 0$ on the Einstein tensor, what implies $b = -\frac{a}{2}$. Thus we find

$$G_{\mu\nu} = a(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu}\mathcal{R}) = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{20.3}$$

The constant a has to be determined by performing the Newtonian limit. Consider weak fields: $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, $|h_{\mu\nu}| \ll 1$ (non relativistic velocities: $v^i \ll c$), then $|T_{ik}| \ll |T_{00}| \Rightarrow |G_{ik}| \ll |G_{00}|$. Compute the trace of $G_{\mu\nu}$:

$$g^{\mu\nu} G_{\mu\nu} \begin{cases} = a(\mathcal{R} - 2\mathcal{R}) = -a\mathcal{R} & \text{from (20.3)} \\ \approx G_{00} = a\left(R_{00} - \frac{\mathcal{R}}{2} \underbrace{g_{00}}_{\substack{\approx\eta_{00} \\ =1}}\right) = a(R_{00} - \mathcal{R}/2) & \end{cases} . \tag{20.4}$$

[17] It can be shown that indeed the Ricci tensor is the only tensor made of the metric tensor and first and second deivatives of it, and which is linear in the second derivative.

Comparing the two results gives $\mathcal{R} \approx -2R_{00}$, thus

$$G_{00} \simeq a\left(R_{00} - \frac{\mathcal{R}}{2}\right) \simeq 2aR_{00}. \tag{20.5}$$

For weak fields all terms quadratic in $h_{\mu\nu}$ can be neglected in the Riemann tensor; we get to leading order:

$$R_{\mu\nu} = R^{\rho}{}_{\mu\rho\nu} \simeq \frac{\partial \Gamma^{\rho}_{\mu\nu}}{\partial x^{\rho}} - \frac{\partial \Gamma^{\rho}_{\rho\mu}}{\partial x^{\nu}} \qquad (|h_{\mu\nu}| \ll 1).$$

For weak stationary fields we find:

$$R_{00} = \frac{\partial \Gamma^{i}_{00}}{\partial x^{i}} \quad \text{with} \quad \Gamma^{i}_{00} = \frac{1}{2}\frac{\partial g_{00}}{\partial x^{i}}.$$

Thus $G_{00} \approx 2a\frac{\partial \Gamma^{i}_{00}}{\partial x^{i}} = a\Delta g_{00} \stackrel{!}{=} \Delta g_{00}$, therefore $a = 1$. Einstein's field equations are [18] (found 1915 by Albert Einstein):

$$\boxed{R_{\mu\nu} - \frac{\mathcal{R}}{2}g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}} \tag{20.6}$$

Together with the geodesic equation ((16.12) or (19.3)), these are the fundamental equations of general relativity. By contraction of (20.6), we find also

$$R^{\mu}{}_{\mu} - \frac{\mathcal{R}}{2}\underbrace{\delta^{\mu}{}_{\mu}}_{=4} = -\mathcal{R} = \frac{8\pi G}{c^4}T. \tag{20.7}$$

$\mathcal{R}$ can be expressed in (20.6) in terms of T , and we get:

$$\boxed{R_{\mu\nu} = \frac{8\pi G}{c^4}\left(T_{\mu\nu} - \frac{T}{2}g_{\mu\nu}\right)} \tag{20.8}$$

an equivalent version of the field equations. For the vacuum case where $T_{\mu\nu} = 0$ we have

$$R_{\mu\nu} = 0. \tag{20.9}$$

Significance of the Bianchi identity

Einstein's equation constitutes a set of non-linear coupled partial differential equations whose general solution is not known. Usually one makes some assumptions, for instance spherical symmetry. Because the Ricci tensor is symmetric, the Einstein equations constitute a set of 10 algebraically independent second order differential equations for $g_{\mu\nu}$.

The Einstein equations are generally covariant, so that they can at best determine the metric up to coordinate transformation ($\rightarrow$ 4 functions). Therefore we expect only 6 independent generally covariant equations for the metric. Indeed the (contracted) Bianchi identities tell us that (equation (18.21)) $G^{\nu}{}_{\mu;\nu} = 0$ and hence there are 4 differential relations among the Einstein's equations. Bianchi identities can also be understood as a consequence of the general covariance of the Einstein equations.

[18]Depending on the convention used for the Riemann tensor, one could also encounter a minus in front of the energy-momentum tensor, as for example in Weinberg.

20.1 The cosmological constant

As a generalization, one can relax condition (2) and have a linear term in $g_{\mu\nu}$ [19]. The field equations become

$$R_{\mu\nu} - \frac{\mathcal{R}}{2} g_{\mu\nu} - \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{20.10}$$

where Λ is a constant: the cosmological constant ($[\Lambda] = L^{-2}$). For point (4) the Newtonian limit of (20.10) leads to

$$\Delta\phi = 4\pi\rho G - c^2\Lambda. \tag{20.11}$$

The right-hand side can also be written as $4\pi G(\rho - \rho_{\text{vacuum}})$, with

$$c^2\rho_{\text{vac}} = \frac{c^4}{4\pi G}\Lambda. \tag{20.12}$$

Λ corresponds to the (constant) energy density of empty space (vacuum). $\Lambda^{-1/2}$ (distance) has to be much larger than the dimension of the solar system.

21 The Einstein-Hilbert action

The field equations (20.6) can be obtained from a covariant variational principle. The action for the metric g is

$$S_{\mathcal{D}}[g] = \int\limits_{\mathcal{D}} \mathcal{R}(g)\, \mathrm{d}v, \tag{21.1}$$

where $\mathcal{D} \subset \mathcal{M}$ is a compact region space-time, $\mathcal{R}$ is a scalar curvature and $\mathrm{d}v$ a volume element:

$$\mathrm{d}v = \sqrt{|g|}\mathrm{d}^4x \tag{21.2}$$

($g = \det g_{ik}$, d^4x in 4 dimensions). The Euler-Lagrange equations are the field equations in vacuum:

$$\delta S_{\mathcal{D}}[g] = 0.$$

We have

$$\delta \int\limits_{\mathcal{D}} \mathcal{R}(g)\, \mathrm{d}v = \int\limits_{\mathcal{D}} \delta(g^{\mu\nu} R_{\mu\nu}\sqrt{-g})\, \mathrm{d}^4x = \int\limits_{\mathcal{D}} (\delta R_{\mu\nu}) g^{\mu\nu}\sqrt{-g}\, \mathrm{d}^4x + \int\limits_{\mathcal{D}} R_{\mu\nu}\delta(g^{\mu\nu}\sqrt{-g})\, \mathrm{d}^4x. \tag{21.3}$$

Consider first $\delta R_{\mu\nu}$:

$$R_{\mu\nu} = \partial_\alpha \Gamma^\alpha_{\mu\nu} - \partial_\nu \Gamma^\alpha_{\mu\alpha} + \Gamma^\rho_{\mu\nu}\Gamma^\alpha_{\rho\alpha} - \Gamma^\rho_{\nu\alpha}\Gamma^\alpha_{\rho\mu}. \tag{21.4}$$

Let us compute the variation of $R_{\mu\nu}$ at any point p in normal coordinates, whose center is in p itself ($x(p) = 0$, then $\Gamma^\alpha_{\beta\gamma}(0) = 0$). Thus $\delta R_{\mu\nu}$ reduces (at any such p) to

$$\delta R_{\mu\nu} = \left(\delta\Gamma^\alpha_{\mu\nu}\right)_{,\alpha} - \left(\delta\Gamma^\alpha_{\mu\alpha}\right)_{,\nu}. \quad [20] \tag{21.5}$$

Without proof one finds that $\delta\Gamma^\alpha_{\mu\nu}$ is indeed a tensor although $\Gamma^\alpha_{\mu\nu}$ is not a tensor. (21.5) is thus a tensor equation, it holds in every coordinate system and we can also take the covariant derivative:

$$\delta R_{\mu\nu} = \left(\delta\Gamma^\alpha_{\mu\nu}\right)_{;\alpha} - \left(\delta\Gamma^\alpha_{\mu\alpha}\right)_{;\nu} \tag{21.6}$$

[19] Note that $g_{\mu\nu;\sigma} = 0$.

[20] variation (with respect to g) δ and normal derivative commute

(*Palatini identity*). Since $g_{\mu\nu;\sigma} = 0$ we can write (21.6) as

$$
\begin{aligned}
g^{\mu\nu}\delta R_{\mu\nu} &= \left(g^{\mu\nu}\delta\Gamma^{\alpha}_{\mu\nu}\right)_{;\alpha} - \left(g^{\mu\nu}\delta\Gamma^{\alpha}_{\mu\alpha}\right)_{;\nu} &(21.7)\\
&= \omega^{\alpha}{}_{;\alpha}\\
&= \omega^{\alpha}{}_{,\alpha} + \underbrace{\Gamma^{\alpha}_{\alpha\mu}}_{\frac{1}{\sqrt{-g}}\frac{\partial\sqrt{-g}}{\partial x^{\mu}}}\omega^{\mu}\\
&= \frac{1}{\sqrt{-g}}\frac{\partial(\sqrt{-g}\,\omega^{\mu})}{\partial x^{\mu}}. &(21.8)
\end{aligned}
$$

Inserting this into the integral (21.3) and applying (15.21) (Gauss theorem), we get

$$\int\limits_{\mathcal{D}} (\operatorname{div}_g \omega)\eta = \int\limits_{\partial\mathcal{D}} i_{\omega}\eta,$$

where $\operatorname{div}_g \omega = \omega^{\alpha}{}_{;\alpha}$ and thus

$$\int\limits_{\mathcal{D}} (\delta R_{\mu\nu}) g^{\mu\nu}\sqrt{-g}\,\mathrm{d}^4x = \int\limits_{\partial\mathcal{D}} \omega^{\alpha}\sqrt{-g}\,\mathrm{d}0_{\alpha}.$$

$\mathrm{d}0_{\alpha}$ is the coordinate normal to $\partial\mathcal{D}$ and

$$\omega^{\alpha} = g^{\mu\nu}\delta\Gamma^{\alpha}_{\mu\nu} - g^{\mu\alpha}\delta\Gamma^{\nu}_{\mu\nu} \tag{21.9}$$

is a vector field. If the variations of $\delta g^{\mu\nu}$ vanish outside a region contained in $\mathcal{D}$, then the boundary term vanishes as well.

As for the second term in (21.3) ($\int\limits_{\mathcal{D}} R_{\mu\nu}\delta(g^{\mu\nu}\sqrt{-g})\,\mathrm{d}^4x$), we recall that for an $n\times n$ matrix $A(\lambda)$ we have (see linear algebra):

i) $\frac{\mathrm{d}}{\mathrm{d}\lambda}\det A = \det A \operatorname{tr}\left(A^{-1}\frac{\mathrm{d}A}{\mathrm{d}\lambda}\right)$,

ii) $\frac{\mathrm{d}}{\mathrm{d}\lambda}(A^{-1})A = -A^{-1}\frac{\mathrm{d}A}{\mathrm{d}\lambda}$.

Thus $(\delta g^{\mu\nu})g_{\nu\sigma} = -g^{\mu\nu}\delta g_{\nu\sigma}$ comes from ii) and $\delta g = g g^{\mu\nu}\delta g_{\nu\mu}$ comes from i) with $A^{-1} = g^{\mu\nu}$. Hence we find the desired expressions

$$
\begin{aligned}
\delta\sqrt{-g} &= \frac{1}{2}\sqrt{-g}\,g^{\mu\nu}\delta g_{\nu\mu} = -\frac{1}{2}\sqrt{-g}\,g_{\alpha\beta}\,\delta g^{\alpha\beta},\\
\delta(g^{\mu\nu}\sqrt{-g}) &= \sqrt{-g}\,\delta g^{\mu\nu} - \frac{1}{2}\sqrt{-g}\,g^{\mu\nu}\,g_{\alpha\beta}\,\delta g^{\alpha\beta}. &(21.10)
\end{aligned}
$$

And thus

$$
\begin{aligned}
0 = \int\limits_{\mathcal{D}} R_{\mu\nu}\delta(g^{\mu\nu}\sqrt{-g})\,\mathrm{d}^4x &= \int\limits_{\mathcal{D}} \underbrace{\sqrt{-g}\,\mathrm{d}^4x}_{\mathrm{d}v}\left(R_{\mu\nu}\delta g^{\mu\nu} - \frac{1}{2}\underbrace{R_{\mu\nu}g^{\mu\nu}}_{\mathcal{R}}\underbrace{g_{\alpha\beta}\delta g^{\alpha\beta}}_{\substack{\alpha\to\mu\\ \beta\to\nu}}\right)\\
&= \int\limits_{\mathcal{D}} \mathrm{d}v\,\underbrace{\left(R_{\mu\nu} - \frac{1}{2}\mathcal{R}g_{\mu\nu}\right)}_{=G_{\mu\nu}=0}\delta g^{\mu\nu}.
\end{aligned}
$$

Therefore $\delta S_{\mathcal{D}}[g_{\mu\nu}] = 0 \Rightarrow G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}\mathcal{R} g_{\mu\nu} = 0$. Since $\delta \int\limits_{\mathcal{D}} \sqrt{-g}\,\mathrm{d}^4x = \int\limits_{\mathcal{D}} \sqrt{-g}\,\mathrm{d}^4x \frac{1}{2}\, g^{\mu\nu}\delta g_{\mu\nu} = -\frac{1}{2}\int\limits_{\mathcal{D}} \sqrt{-g}\,\mathrm{d}^4x\, g_{\mu\nu}\,\delta g^{\mu\nu}$, it follows that if we have a cosmological constant , the Einstein's vacuum equations are obtained from the action principle applied on

$$S_{\mathcal{D}}[g] = \int\limits_{\mathcal{D}} (\mathcal{R} - 2\Lambda)\sqrt{-g}\,\mathrm{d}^4x. \tag{21.11}$$

The variational principle extends to matter described by any field $\psi = (\psi_A)$ $(A = 1, \dots, N)$, (we include also the electromagnetic field among the ψ_A) transforming as a tensor under change of coordinates. Consider an action of the form

$$S_{\mathcal{D}}[\psi] = \int\limits_{\mathcal{D}} \mathcal{L}(\psi, \nabla_g\psi)\sqrt{-g}\,\mathrm{d}^4x, \tag{21.12}$$

where ∇_g is the Riemannian connection of the metric g. If we know $\mathcal{L}$ in flat space, the equivalence principle prescribes to replace $\eta_{\alpha\beta}$ by $g_{\alpha\beta}$ and replace ordinary derivatives by covariant ones.

<u>Example:</u> electromagnetic field

$$\mathcal{L} = -\frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu} = -\frac{1}{16\pi}F_{\mu\nu}\,F_{\sigma\rho}\,g^{\sigma\mu}\,g^{\rho\nu},$$

and the Euler-Lagrange equations in this case ($F_{\mu\nu} = A_{\nu;\mu} - A_{\mu;\nu} = A_{\nu,\mu} - A_{\mu,\nu}$) for the basic 4-potential A_μ field read:

$$\frac{\partial\mathcal{L}}{\partial A_\nu} - \nabla_\mu \frac{\partial\mathcal{L}}{\partial\nabla_\mu A_\nu} = 0, \qquad \text{with } \nabla_\mu A_\nu = A_{\nu;\mu};$$

in this case $\frac{\partial\mathcal{L}}{\partial A_\nu} = 0$, and $\frac{\partial\mathcal{L}}{\partial\nabla_\mu A_\nu} = -\frac{1}{4\pi}F^{\mu\nu}$. The Euler-Lagrange equations are then ${F^{\mu\nu}}_{;\nu} = 0$, which are the Maxwell equations for vanishing current j^μ (${F^{\mu\nu}}_{;\nu} = \frac{4\pi}{c}j^\mu$ and $\mathcal{L} = -\frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu} - \frac{1}{c}j^\mu A_\mu$ with $j^\mu A_\mu = g^{\mu\nu}j_\nu A_\mu$).

Variations in (21.12) with respect to the fields ψ_A lead to the Euler-Lagrange equations, whereas variations with respect to the metric (which is also a function and is determined by solving Einstein's equations) gives (without proof)

$$\delta_g \int\limits_{\mathcal{D}} \mathcal{L}(\psi, \nabla_g\psi)\sqrt{-g}\,\mathrm{d}^4x = -\frac{1}{2}\int\limits_{\mathcal{D}} T^{\mu\nu}\delta g_{\mu\nu}\sqrt{-g}\,\mathrm{d}^4x. \tag{21.13}$$

This term has to be added to the one proportional to $\delta g_{\mu\nu}$ in Einstein's action:

$$\int\limits_{\mathcal{D}} \sqrt{-g}\,\mathrm{d}^4x \underbrace{\left(G_{\mu\nu}\frac{c^4}{16\pi G} - \frac{1}{2}T_{\mu\nu}\right)}_{=0}\delta g^{\mu\nu}$$

and thus $G_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}$. For electrodynamics: $T_{\mu\nu} = \frac{1}{4\pi}\left(F_{\mu\sigma}F^\sigma_\nu - \frac{1}{4}F_{\sigma\rho}F^{\sigma\rho}g_{\mu\nu}\right)$ (or $T^{\alpha\beta} = -F^\alpha_\mu F^{\mu\beta} - \mathcal{L}g^{\alpha\beta}$). And similarly for other "matter" fields.

22 Static isotropic metric

22.1 Form of the metric

For the gravity field of Earth and Sun we assume a spherically symmetric distribution of the matter (rotation velocities $v^i \ll c$). Thus we need a spherically symmetric and static solution for the metric $g_{\mu\nu}(x)$. We first give the general form of such a metric (static and isotropic) which we then use as an ansatz to solve the field equations. For $r \to \infty$, the Newtonian gravitational potential $\varphi = -\frac{GM}{r}$ goes to zero. Thus, asymptotically, the metric should be Minkowskian: $\mathrm{d}s^2 \underset{r\to\infty}{=} c^2\mathrm{d}t^2 - \mathrm{d}r^2 - r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)$, in spherical coordinates r, θ, ϕ and t. Thus,

$$\mathrm{d}s^2 = B(r)c^2\mathrm{d}t^2 - A(r)\mathrm{d}r^2 - C(r)r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2). \tag{22.1}$$

Due to isotropy and time independence, A, B and C cannot depend on θ, ϕ and t (and no linear terms in $\mathrm{d}\theta$ and $\mathrm{d}\phi$). Freedom in the choice of coordinates allows to introduce a new radial coordinate in (22.1): $C(r)r^2 \to r^2$, thus $C(r)$ can be absorbed into r. We get the standard form:

$$\mathrm{d}s^2 = B(r)c^2\mathrm{d}t^2 - A(r)\mathrm{d}r^2 - r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2) \tag{22.2}$$

(θ and ϕ have the same significance as in Minkowski coordinates). Due to our asymptotic requirements ($r \to \infty$) we can assume that $B(r) \to 1$ and $A(r) \to 1$.

22.2 Robertson expansion

Even without knowing the solution to the field equations, we can give an expansion of the metric for weak fields outside the mass distribution. The metric can only depend on the total mass of the considered object (Earth or Sun for instance), on the distance from it and on the constants G, c. Since A and B are dimensionless, they can only depend on a combination of the dimensionless quantity $\frac{GM}{c^2r}$. For $\frac{GM}{c^2r} \ll 1$ we can then have the following expansion:

$$\begin{aligned} B(r) &= 1 - 2\frac{GM}{c^2r} + 2(\beta-\gamma)\left(\frac{GM}{c^2r}\right)^2 + \dots \\ A(r) &= 1 + 2\gamma\frac{GM}{c^2r} + \dots \end{aligned} \tag{22.3}$$

which is the *Robertson expansion.* The linear term in $B(r)$ has no free parameter since it is constrained by the Newtonian limit: $g_{00} \simeq 1 + 2\frac{\phi}{c^2}$, $\phi = -\frac{GM}{r}$ (Newtonian potential), therefore $B \to g_{00}$. The coefficient $2(\beta-\gamma)$ comes from historical reasons, β and γ are independent coefficients. In the solar system, $\frac{GM}{c^2r} \leq \frac{GM}{c^2R_\odot} \simeq 2\times10^{-6}$, then only linear terms in γ and β play a role. For general relativity: $\gamma = \beta = 1$ (Newtonian gravity: $\gamma = \beta = 0$).

22.3 Christoffel symbols and Ricci tensor for the standard form

The metric tensor $g_{\mu\nu}$ is diagonal.

$$g_{00} = B(r) \qquad g_{11} = -A(r) \qquad g_{22} = -r^2 \qquad g_{33} = -r^2 \sin^2\theta \tag{22.4}$$

$$g^{00} = \frac{1}{B(r)} \qquad g^{11} = -\frac{1}{A(r)} \qquad g^{22} = -\frac{1}{r^2} \qquad g^{33} = -\frac{1}{r^2 \sin^2\theta} \tag{22.5}$$

The non-vanishing components of $\Gamma^{\sigma}_{\lambda\mu} = \frac{g^{\sigma\nu}}{2}\left(\frac{\partial g_{\mu\nu}}{\partial x^{\lambda}} + \frac{\partial g_{\lambda\nu}}{\partial x^{\mu}} - \frac{\partial g_{\mu\lambda}}{\partial x^{\nu}}\right)$ are

$$\begin{aligned}
&\Gamma^0_{01} = \Gamma^0_{10} = \frac{B'}{2B} && \Gamma^1_{00} = \frac{B'}{2A} && \Gamma^1_{11} = \frac{A'}{2A} \\
&\Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r} && \Gamma^1_{22} = -\frac{r}{A} && \Gamma^1_{33} = -\frac{r\sin^2\theta}{A} \\
&\Gamma^3_{13} = \Gamma^3_{31} = \frac{1}{r} && \Gamma^3_{23} = \Gamma^3_{32} = \cot\theta && \Gamma^2_{33} = -\sin\theta\cos\theta
\end{aligned} \tag{22.6}$$

where $'$ stands for $\frac{\partial}{\partial r}$. With

$$-g = r^4 AB \sin^2\theta \tag{22.7}$$

we get

$$(\Gamma^{\rho}_{\mu\rho}) = \left(\frac{\partial \ln\sqrt{-g}}{\partial x^{\mu}}\right) = \left(0, \frac{2}{r} + \frac{A'}{2A} + \frac{B'}{2B}, \cot\theta, 0\right). \tag{22.8}$$

The Ricci tensor can then be calculated as

$$R_{\mu\nu} = \frac{\partial \Gamma^{\rho}_{\mu\nu}}{\partial x^{\rho}} - \frac{\partial \Gamma^{\rho}_{\rho\mu}}{\partial x^{\nu}} + \Gamma^{\sigma}_{\mu\nu}\Gamma^{\rho}_{\rho\sigma} - \Gamma^{\sigma}_{\rho\mu}\Gamma^{\rho}_{\nu\sigma}, \tag{22.9}$$

and we get as a result

$$\begin{aligned}
R_{00} &= \frac{B''}{2A} - \frac{A'B'}{2A^2} - \frac{B'^2}{2AB} + \frac{B'}{2A}\left(\frac{2}{r} + \frac{A'}{2A} + \frac{B'}{2B}\right), \\
&= \frac{B''}{2A} - \frac{B'}{4A}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{B'}{rA},
\end{aligned} \tag{22.10}$$

$$R_{11} = -\frac{B''}{2B} + \frac{B'}{4B}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{A'}{rA}, \tag{22.11}$$

$$R_{22} = 1 + \frac{r}{2A}\left(\frac{A'}{A} - \frac{B'}{B}\right) - \frac{1}{A}, \tag{22.12}$$

$$R_{33} = R_{22}\sin^2\theta, \tag{22.13}$$

The non-diagonal components $R_{\mu\nu}$ with $\mu \neq \nu$ vanish.

22.4 Schwarzschild metric

We assume a static, spherically symmetric, mass distribution with finite extension:

$$\rho(r) \begin{cases} \neq 0 & r \leq r_0 \\ = 0 & r > r_0 \end{cases} \tag{22.14}$$

Similarly, the pressure $P(r)$ is thought to vanish for $r > r_0$. The four velocity vector within the mass distribution in the static case is $u^\mu = (u^0 = \text{constant}, 0, 0, 0)$. This way, the energy-momentum tensor (describing matter) does not depend on time. We then adopt the ansatz for the metric elaborated in (22.2): $g_{\mu\nu} = \text{diag}(B(r), -A(r), -r^2, -r^2 \sin^2\theta)$. Outside the mass distribution ($r \geq r_0$), the Ricci tensor vanishes: $R_{\mu\nu} = 0$. We have already calculated the coefficients $R_{\mu\nu}$ in equations (22.10) – (22.13). For $\mu \neq \nu$, $R_{\mu\nu} = 0$ is trivially satisfied while the diagonal components should be set to zero: $R_{00} = R_{11} = R_{22} = R_{33} = 0$ ($r \geq r_0$).

Consider $\dfrac{R_{00}}{B} + \dfrac{R_{11}}{A} = -\dfrac{1}{rA}\left(\dfrac{B'}{B} + \dfrac{A'}{A}\right) = 0$ and thus $\dfrac{\mathrm{d}}{\mathrm{d}r}(\ln AB) = 0$ (since $rA \neq 0$) or $AB = \text{constant}$ (or $\ln AB = \text{constant}$). For $r \to \infty$ we require $A = B = 1$, therefore $AB = 1 \Rightarrow A(r) = \frac{1}{B(r)}$. Introducing this into R_{22} (22.12) and R_{11} (22.11) leads to

$$R_{22} = 1 - rB' - B = 0, \tag{22.15}$$

$$R_{11} = -\frac{B''}{2B} - \frac{B'}{rB} = -\frac{rB'' + 2B'}{2rB} = \frac{1}{2rB}\frac{\mathrm{d}R_{22}}{\mathrm{d}r} = 0. \tag{22.16}$$

With (22.15), (22.16) is automatically satisified (since $R_{22} = 0$ also its derivative vanishes). We write (22.15) as

$$\frac{\mathrm{d}}{\mathrm{d}r}(rB) = 1. \tag{22.17}$$

We integrate it and get $rB = r + \underbrace{\text{constant}}_{-2a} = r - 2a$. Then

$$\begin{aligned} B(r) &= 1 - \frac{2a}{r}, \\ A(r) &= \frac{1}{1 - \frac{2a}{r}}, \end{aligned} \tag{22.18}$$

for $r \geq r_0$. This solution for the vacuum Einstein' equations was found in 1916 by Schwarzschild. The *Schwarzschild solution* is

$$\boxed{\mathrm{d}s^2 = \left(1 - \frac{2a}{r}\right)c^2\mathrm{d}t^2 - \frac{\mathrm{d}r^2}{1 - \frac{2a}{r}} - r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2)} \tag{22.19}$$

The constant can be determined by considering the Newtonian limit:

$$g_{00} = B(r) \xrightarrow[r\to\infty]{} 1 + 2\frac{\phi}{c^2} = 1 - 2\frac{GM}{c^2 r} = 1 - \frac{2a}{r}.$$

Thus one introduces the so called *Schwarzschild radius*:

$$\boxed{r_S = 2a = \frac{2GM}{c^2}}$$

The Schwarzschild radius of the Sun is $r_{s,\odot} = \frac{2GM_\odot}{c^2} \simeq 3$ km ($M_\odot \simeq 2 \times 10^{30}$ kg, $R_\odot = 7 \times 10^5$ km) so $\frac{r_{S,\odot}}{R_\odot} = \frac{2GM_\odot}{c^2 R_\odot} \simeq 4 \times 10^{-6}$ [21]. A clock at rest in r has the proper time $\mathrm{d}\tau = \sqrt{B}\,\mathrm{d}t$, thus $\frac{\mathrm{d}t}{\mathrm{d}\tau}$ diverges at $r \to r_S$. This implies that a photon emitted at $r = r_S$ will be infinitely redshifted (t is not a good coordinate either for events taking place at $r \leq r_S$). A star, whose radius r_{star} is smaller than r_S, is a black hole since photons emitted at its surface cannot reach regions with $r > r_S$.

Expanding the Schwarzschild metric in power of $\frac{r_S}{r}$ and comparing it with the Robertson expansion (22.3), one finds $\beta = \gamma = 1$ for general relativity.

23 General equations of motion

We now consider the motion of a freely falling material particle or photon in a static isotropic gravitational field (e.g. motion of planets around the Sun). For the relativistic orbit $x^k(\lambda)$ of a particle in a gravitational field we have:

$$\frac{\mathrm{d}^2 x^k}{\mathrm{d}\lambda^2} = -\Gamma^k_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} \tag{23.1}$$

and

$$g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} = \left(\frac{\mathrm{d}s}{\mathrm{d}\lambda}\right)^2 = c^2 \left(\frac{\mathrm{d}\tau}{\mathrm{d}\lambda}\right)^2 = \begin{cases} c^2 & m \neq 0, \quad \lambda = \tau \\ 0 & m = 0 \end{cases} . \tag{23.2}$$

For a massive particle we can take the proper time as a parameter for the trajectory or orbit ($\mathrm{d}\lambda = \mathrm{d}\tau$). For massless particles one has to choose another parameter. For the spherically symmetric gravitational field, we use the metric ($r > r_\odot$, radius of the star)

$$\mathrm{d}s^2 = B(r)c^2 \mathrm{d}t^2 - \mathrm{d}r^2 A(r) - r^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2), \tag{23.3}$$

with the coordinates $(x^0, x^1, x^2, x^3) = (ct, r, \theta, \phi)$. Equations (23.1) – (23.3) define the relativistic Kepler problem. Using the Christoffel symbols given in (22.6), we get for (23.1):

$$\frac{\mathrm{d}^2 x^0}{\mathrm{d}\lambda^2} = -\frac{B'}{B} \frac{\mathrm{d}x^0}{\mathrm{d}\lambda} \frac{\mathrm{d}r}{\mathrm{d}\lambda}, \tag{23.4}$$

$$\frac{\mathrm{d}^2 r}{\mathrm{d}\lambda^2} = -\frac{B'}{2A} \left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda}\right)^2 - \frac{A'}{2A} \left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 + \frac{r}{A} \left(\frac{\mathrm{d}\theta}{\mathrm{d}\lambda}\right)^2 + \frac{r\sin^2\theta}{A} \left(\frac{\mathrm{d}\phi}{\mathrm{d}\lambda}\right)^2, \tag{23.5}$$

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}\lambda^2} = -\frac{2}{r} \frac{\mathrm{d}\theta}{\mathrm{d}\lambda} \frac{\mathrm{d}r}{\mathrm{d}\lambda} + \sin\theta\cos\theta \left(\frac{\mathrm{d}\phi}{\mathrm{d}\lambda}\right)^2, \tag{23.6}$$

$$\frac{\mathrm{d}^2\phi}{\mathrm{d}\lambda^2} = -\frac{2}{r} \frac{\mathrm{d}\phi}{\mathrm{d}\lambda} \frac{\mathrm{d}r}{\mathrm{d}\lambda} - 2\cot\theta \frac{\mathrm{d}\theta}{\mathrm{d}\lambda} \frac{\mathrm{d}\phi}{\mathrm{d}\lambda}. \tag{23.7}$$

Equation (23.6) can be solved by

$$\theta = \frac{\pi}{2} = \text{constant}. \tag{23.8}$$

Without loss of generality we can choose the coordinate system such that $\theta = \frac{\pi}{2}$, this way the trajectory lies on the plane with $\theta = \frac{\pi}{2}$. $\frac{\mathrm{d}^2\theta}{\mathrm{d}\lambda^2} = 0$ corresponds to angular momentum conservation. With (23.8)

[21] Apparently it seems that the Schwarzschild metric is singular for $r = r_s$, but this is not the case. It is only an artefact of the coordinate choice. To be discussed later.

we get for (23.7) :

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(r^2\frac{\mathrm{d}\phi}{\mathrm{d}\lambda}\right) = 0, \tag{23.9}$$

which leads to

$$r^2\frac{\mathrm{d}\phi}{\mathrm{d}\lambda} = \text{constant} = l. \tag{23.10}$$

l can be interpreted as the (orbital) angular momentum (per unit mass). Equations (23.8) and (23.10) follow from angular momentum conservation, which is a consequence of spherical symmetry (rotation invariance).

Equation (23.4) can be written as ($B = B(r(\lambda))$)

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\ln\left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda}\right) + \ln B\right) = 0, \tag{23.11}$$

which can be integrated as $\ln\left[\left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda}\right)B\right] = \text{constant}$ or

$$B\frac{\mathrm{d}x^0}{\mathrm{d}\lambda} = \text{constant} = F. \tag{23.12}$$

In (23.5) we use (23.8), (23.10) and (23.12) and get:

$$\frac{\mathrm{d}^2 r}{\mathrm{d}\lambda^2} + \frac{F^2 B'}{2AB^2} + \frac{A'}{2A}\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 - \frac{l^2}{Ar^3} = 0. \tag{23.13}$$

We multiply it with $2A\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)$ and get

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left[A\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 + \frac{l^2}{r^2} - \frac{F^2}{B}\right] = 0. \quad ^{22} \tag{23.14}$$

Integration gives

$$A\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 + \frac{l^2}{r^2} - \frac{F^2}{B} = -\epsilon = \text{constant}. \tag{23.15}$$

Integrating it once more we get $r = r(\lambda)$. Inserting then this result into (23.10) and (23.12), we obtain with one more integration $\phi = \phi(\lambda)$ and $t = t(\lambda)$. Next we eliminate λ and get $r = r(t)$ and $\phi = \phi(t)$. Together with $\theta = \frac{\pi}{2}$, this is then a complete solution (generally it has to be done numerically).

Equation (23.2) becomes

$$g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} = B\left(\frac{\mathrm{d}x^0}{\mathrm{d}\lambda}\right)^2 - A\left(\frac{\mathrm{d}r}{\mathrm{d}\lambda}\right)^2 - r^2\left(\frac{\mathrm{d}\theta}{\mathrm{d}\lambda}\right)^2 - r^2\sin^2\theta\left(\frac{\mathrm{d}\phi}{\mathrm{d}\lambda}\right)^2 = \epsilon, \tag{23.16}$$

using (23.8), (23.10), (23.12) and (23.15). On the other hand

$$\epsilon = \begin{cases} c^2 & (m \neq 0) \\ 0 & (m = 0) \end{cases}.$$

We are left with two integration constants, F and l.

[22]Notice: $\frac{\mathrm{d}A}{\mathrm{d}\lambda} = A'\frac{\mathrm{d}r}{\mathrm{d}\lambda}$

23.1 Trajectory

From (23.15) we get

$$\frac{\mathrm{d}r}{\mathrm{d}\lambda} = \sqrt{\frac{\frac{F^2}{B} - \frac{l^2}{r^2} - \epsilon}{A}}, \tag{23.17}$$

and with (23.10),

$$\frac{\mathrm{d}\phi}{\mathrm{d}r} = \frac{\mathrm{d}\phi}{\mathrm{d}\lambda}\frac{\mathrm{d}\lambda}{\mathrm{d}r} = \frac{l}{r^2}\sqrt{\frac{A}{\frac{F^2}{B} - \frac{l^2}{r^2} - \epsilon}}. \tag{23.18}$$

Thus,

$$\phi(r) = \int \frac{\mathrm{d}r}{r^2}\sqrt{\frac{A(r)}{\frac{F^2}{B(r)l^2} - \frac{1}{r^2} - \frac{\epsilon}{l^2}}}. \tag{23.19}$$

With this we can find the trajectory $\phi = \phi(r)$ in the orbital plane. (Massive particles: 2 integration constants $\frac{F^2}{l^2}$ and $\frac{\epsilon}{l^2}$, massless particles: only $\frac{F^2}{l^2}$).

Trajectory in Schwarzschild metric:

Insert Schwarzschild metric: $B(r) = A^{-1}(r) = 1 - \frac{r_S}{r} = 1 - \frac{2a}{r}$ and write:

$$\dot{t} = \frac{\mathrm{d}t}{\mathrm{d}\lambda}, \quad \dot{r} = \frac{\mathrm{d}r}{\mathrm{d}\lambda}, \quad \dot{\phi} = \frac{\mathrm{d}\phi}{\mathrm{d}\lambda}.$$

Then with (23.8), (23.10), (23.12) and (23.15) we get

$$\theta = \frac{\pi}{2}, \quad c\dot{t}\left(1 - \frac{2a}{r}\right) = F, \quad r^2\dot{\phi} = l. \tag{23.20}$$

Multiplying (23.15) with B and using $AB = 1$, we have

$$\frac{\dot{r}^2}{2} - \frac{a\epsilon}{r} + \frac{l^2}{2r^2} - \frac{al^2}{r^3} = \frac{F^2 - \epsilon}{2} = \text{constant}. \tag{23.21}$$

The radial component can be written as

$$\frac{\dot{r}^2}{2} + V_{\text{eff}}(r) = \text{constant}, \tag{23.22}$$

with the effective potential ($2a = \frac{2GM}{c^2}$, $\epsilon = \{c^2, 0\}$)

$$V_{\text{eff}}(r) = \begin{cases} -\frac{GM}{r} + \frac{l^2}{2r^2} - \frac{GMl^2}{c^2r^3} & (m \neq 0) \\ \frac{l^2}{2r^2} - \frac{GMl^2}{c^2r^3} & (m = 0) \end{cases}. \tag{23.23}$$

A formal solution $r = r(\lambda)$ of (23.22) is given through the following integral :

$$\lambda = \pm \int \frac{\mathrm{d}r}{\sqrt{2(\text{constant} - V_{eff}(r))}}. \tag{23.24}$$

Due to the $\frac{1}{r^3}$ term (relativistic), this is an elliptical integral which has to be solved numerically.

For small values of r, centrifugal potential term dominates (as long as l is not too small), then for even smaller values of r the attractive relativistic term takes over: [23] [24]

[23] $l \sim r \times v \rightarrow \frac{l^2}{r^2} \sim v^2$

[24] $\frac{v}{c} \sim 10^{-4} \rightarrow \frac{v^2}{c^2} \sim 10^{-8}$

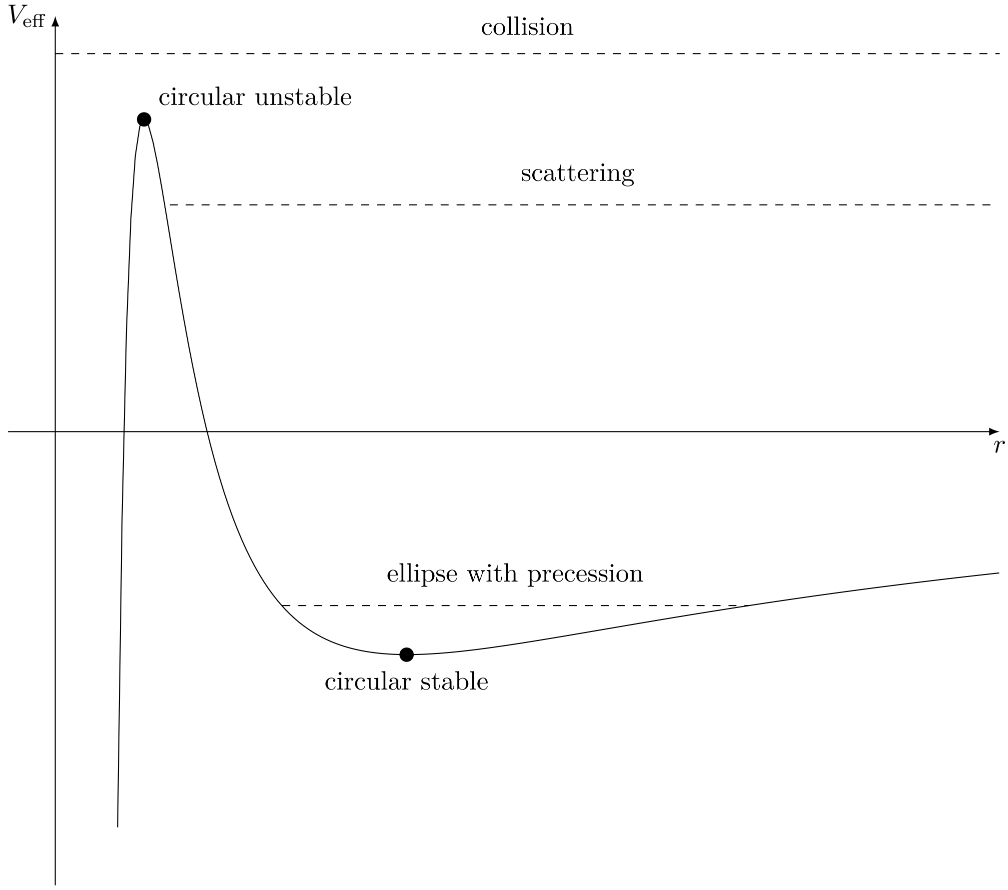


Figure 11: Effective potential for massive particles in Schwarzschild metric

$$-\frac{GM}{r}\frac{l^2}{c^2r^2} \simeq -\frac{GM}{r}\frac{v^2}{c^2} \tag{23.25}$$

Eq. (23.22) differs from the non-relativistic case by an additional $\frac{1}{r^3}$ term and $\dot{r} = \frac{\mathrm{d}r}{\mathrm{d}\tau}$ differs from $\frac{\mathrm{d}r}{\mathrm{d}t}$ by terms of order $\frac{v^2}{c^2}$.

Observations:

- Where V_{eff} has a minimum there are bounded solutions, however due to the relativistic effects there will be small deviations from the elliptical orbits (precession of the perihelion). As a special case, with $\dot{r} = 0$, the circular orbit is a possible solution (in which case the constant in (23.22) is equal to the value of V_{eff} at its minimum).
- The solution at the maximum of V_{eff} is an unstable circular orbit.
- If the constant is positive one gets non-bounded trajectories (corresponding to hyperbolic solutions in the non-relativistic case).

- If the constant is larger than the maximum value of the potential, the particle falls into the center.
- At minimum and maximum we have $\frac{\mathrm{d}V_{eff}}{\mathrm{d}r} = 0$. For $m \neq 0$ we get

$$\frac{c^2}{l^2}r^2 - 2\frac{r}{r_S} + 3 = 0. \tag{23.26}$$

In order to have two real solutions, we need $\frac{3c^2}{l^2} < \frac{1}{r_S^2}$. That means

$$l \geq l_{crit} = \sqrt{3}\, r_S c. \tag{23.27}$$

For $l \to l_{crit}$, the angular momentum barrier gets smaller and smaller until the maximum and minimum fall together for $l = l_{crit}$. For $l < l_{crit}$, the potential decreases monotonically for $r \to 0$.

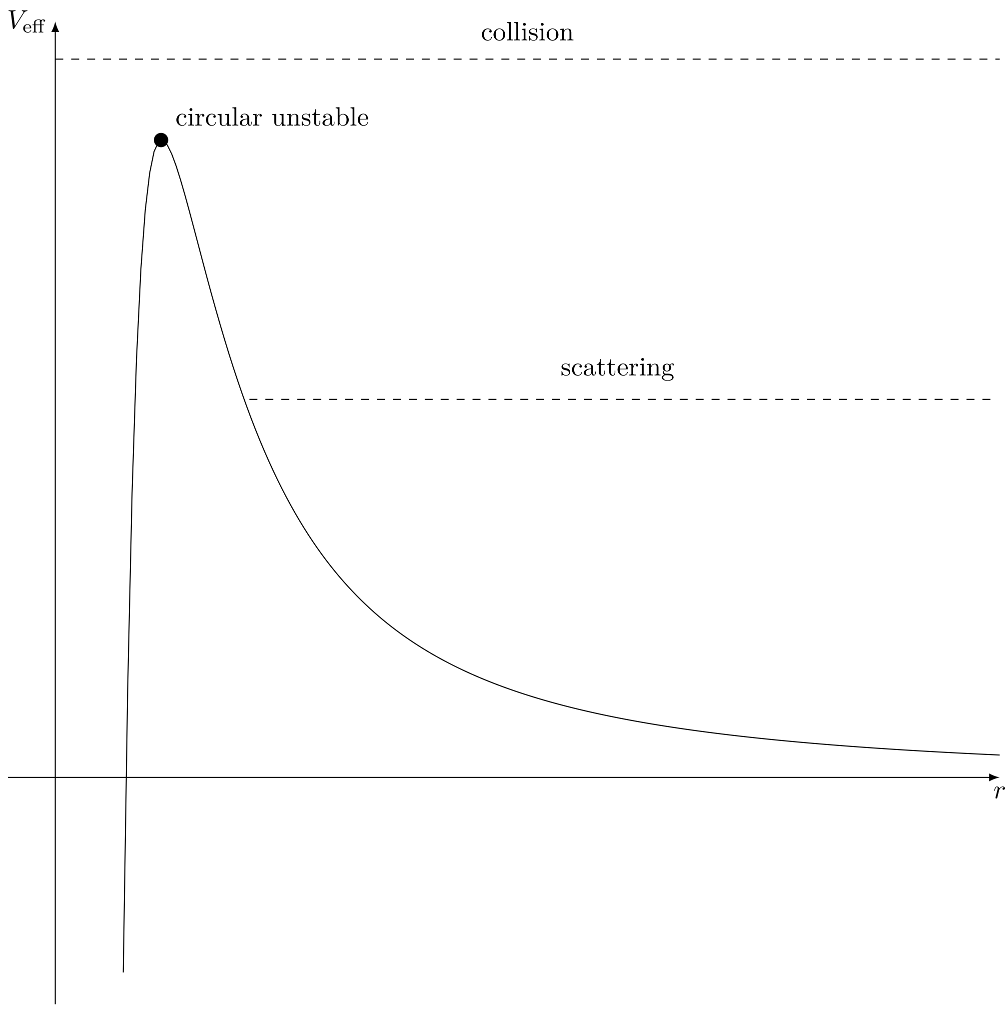


Figure 12: Effective potential for massless particles in Schwarzschild metric

Here both terms are proportional to l^2, thus the shape of V_{eff} does not depend on l. At $r_{max} = \frac{3}{2}r_S$ the potential has a maximum. At r_{max} the photons can move on a circular orbit, which is unstable. If the constant in (23.22) is smaller than $V_{\text{eff}}(r_{max})$ then the incoming photon will be scattered, whereas if the constant is bigger the photon will be absorbed at the center. [25]

[25] for $r \leq r_S$ the Schwarzschild solution is not applicable

Part VI

Applications of General Relativity

24 Light deflection

The trajectory $r = r(\phi)$ of a photon in the gravitational field is given by (23.19) ($\epsilon = 0$):

$$\phi(r) = \phi(r_0) + \int_{r_0}^{r} \frac{\mathrm{d}\tilde{r}}{\tilde{r}^2} \sqrt{\frac{A(\tilde{r})}{\frac{F^2}{B(\tilde{r})l^2} - \frac{1}{\tilde{r}^2}}}. \tag{24.1}$$

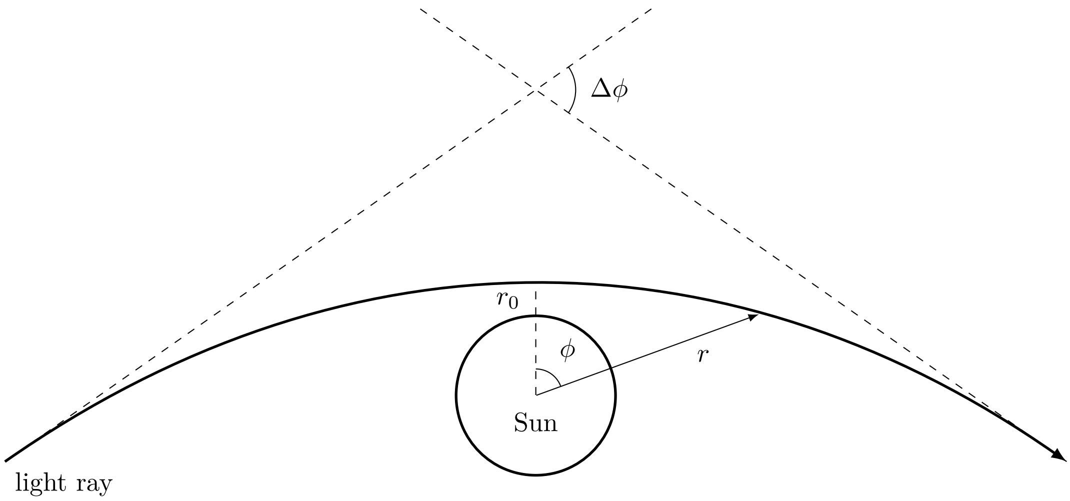


Figure 13: Deflection of light by the Sun

We will now show that light is deflected by a massive body, carrying through calculations for the Sun. In fig. 13, the following quantities are defined: light is deflected by $\Delta\phi$ and r_0 is the minimal distance (or impact parameter) from the Sun. For simplification we assume also $r_0 \gg r_S$.

As starting point of the integration we choose the minimum distance r_0, where we set $\phi(r_0) = 0$. Going from r_0 till r_∞ the angle changes by $\phi(\infty)$. Along the drawn trajectory the radial vector turns by $2\phi(\infty)$. If the trajectory would be a straight line, then $2\phi(\infty) = \pi$.
Thus $\Delta\phi = \pi - \pi = 0$ for a straight line and in general ($\phi(r_0) = 0$):

$$\Delta\phi = 2\phi(\infty) - \pi. \tag{24.2}$$

At r_0, $r(\phi)$ is a minimum, thus

$$\left(\frac{\mathrm{d}r}{\mathrm{d}\phi}\right)\bigg|_{r_0} = 0. \tag{24.3}$$

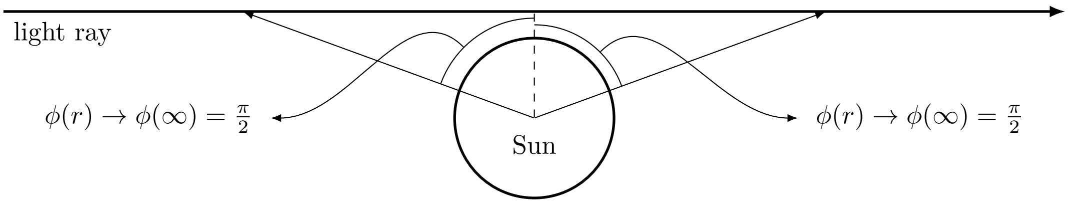


Figure 14: Non-deflected ray of light

From (24.3) we get with (23.17) and (23.18) the condition

$$\frac{F^2}{l^2} = \frac{B(r_0)}{r_0^2}. \tag{24.4}$$

This way we can eliminate the constants F and l in terms of r_0 with (24.1):

$$\phi(\infty) = \int_{r_0}^{\infty} \frac{\mathrm{d}r}{r} \sqrt{\frac{A(r)}{\frac{B(r_0)}{B(r)}\frac{r^2}{r_0^2} - 1}}. \tag{24.5}$$

Let us compute the integral by inserting the Robertson expansion $A(r) = 1 + \gamma\frac{2a}{r}$, $B(r) = 1 - \frac{2a}{r}$ (with $a = \frac{r_s}{2} = \frac{GM}{c^2}$). We keep terms up to $\frac{a}{r}$ with

$$\begin{aligned}\frac{B(r_0)}{B(r)}\frac{r^2}{r_0^2} - 1 &\simeq \frac{r^2}{r_0^2}\left[1 + 2a\left(\frac{1}{r} - \frac{1}{r_0}\right)\right] - 1 \\ &= \left[\frac{r^2}{r_0^2} - 1\right]\left[1 - \frac{2ar}{r_0(r + r_0)}\right].\end{aligned}$$

We get using $\sqrt{1+x} = 1 + \frac{x}{2}$,

$$\begin{aligned}\phi(\infty) &\simeq \int_{r_0}^{\infty} \frac{\mathrm{d}r}{\sqrt{r^2 - r_0^2}} \frac{r_0}{r}\left(1 + \gamma\frac{a}{r} + \frac{ar}{r_0(r + r_0)}\right) \\ &= \left[\arccos\left(\frac{r_0}{r}\right) + \gamma\frac{a}{r_0}\frac{\sqrt{r^2 - r_0^2}}{r} + \frac{a}{r_0}\sqrt{\frac{r - r_0}{r + r_0}}\right]_{r_0}^{\infty} \\ &= \frac{\pi}{2} + \gamma\frac{a}{r_0} + \frac{a}{r_0}.\end{aligned} \tag{24.6}$$

With (24.2) we get

$$\Delta\phi = \frac{4a}{r_0}\left(\frac{1+\gamma}{2}\right) = \frac{2r_S}{r_0}\left(\frac{1+\gamma}{2}\right). \tag{24.7}$$

For general relativity, $\gamma = 1$, $r_S = \frac{2GM}{c^2}$ and thus

$$\boxed{\Delta\phi = \frac{2r_S}{r_0}}$$

For a light ray which just grazes the surface of the Sun ($r_0 = R_\odot = 7\times10^5$ km) we get ($\pi = 180\times3600''$):

$$\Delta\phi = 1.75'' \left(\frac{1+\gamma}{2}\right). \quad [26] \tag{24.8}$$

On May 29, 1919, an eclipse allowed experimental confirm of this result.

Figure 15: Gravitational lensing in the Abel 2218 galaxy cluster

[26]"cheating" with Newton's theory gives half this result that is $0.84''$

25 Perihelion precession

Consider the elliptical orbit of a planet around the Sun:

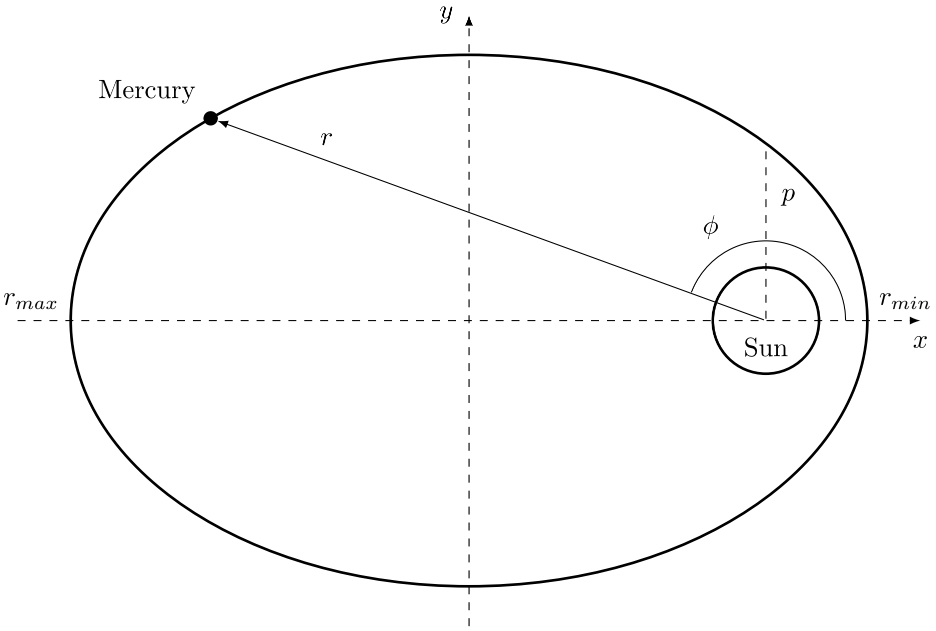


Figure 16: Non-relativistic elliptical orbit of Mercury around the Sun

We will use the following notations: minimum distance $r_- = r_{min}$, maximum distance $r_+ = r_{max}$, $\phi_\pm = \phi(r_\pm)$, $A_\pm = A(r_\pm)$, $B_\pm = B(r_\pm)$. The relativistic orbit follows from equation (23.19) for $r = r(\phi)$ with $\epsilon = c^2$. The integral gives for the change in angle between r_- and r_+:

$$\phi_+ - \phi_- = \int_{r_-}^{r_+} \frac{\mathrm{d}r}{r^2} \sqrt{\frac{A(r)}{\frac{F^2}{B(r)l^2} - \frac{1}{r^2} - \frac{c^2}{l^2}}} = \int_{r_-}^{r_+} \frac{\mathrm{d}r}{r^2} \sqrt{\frac{A(r)}{K(r)}}. \tag{25.1}$$

For a full orbit the angle is 2π, i.e. twice the integral (25.1). The shift of the perihelion (per complete orbit) is given by

$$\Delta\phi = 2(\phi_+ - \phi_-) - 2\pi. \tag{25.2}$$

The integrand in (25.1) is equal to $\frac{\mathrm{d}\phi}{\mathrm{d}r}$. For $r = r_\pm$ due to $\frac{\mathrm{d}r}{\mathrm{d}\phi} = 0$, $\sqrt{K(r)}\,r^2$ has to vanish, thus $K(r_\pm) = 0$:

$$\frac{F^2}{B_\pm l^2} = \frac{1}{r_\pm^2} + \frac{c^2}{l^2}. \tag{25.3}$$

This way we can express F and l through $r_\pm$:

$$\frac{F^2}{l^2} = \frac{\frac{1}{r_+^2} - \frac{1}{r_-^2}}{\frac{1}{B_+} - \frac{1}{B_-}} = \frac{r_-^2 - r_+^2}{r_+^2 r_-^2 \left(\frac{1}{B_+} - \frac{1}{B_-}\right)}, \tag{25.4}$$

$$\frac{c^2}{l^2} = -\frac{\frac{B_+}{r_+^2} - \frac{B_-}{r_-^2}}{B_+ - B_-} = -\frac{\frac{r_+^2}{B_+} - \frac{r_-^2}{B_-}}{r_+^2 r_-^2 \left(\frac{1}{B_+} - \frac{1}{B_-}\right)}. \tag{25.5}$$

This leads us to an expression for $K(r)$:

$$K(r) = \frac{r_-^2\left(\frac{1}{B(r)} - \frac{1}{B_-}\right) - r_+^2\left(\frac{1}{B(r)} - \frac{1}{B_+}\right)}{r_+^2 r_-^2\left(\frac{1}{B_+} - \frac{1}{B_-}\right)} - \frac{1}{r^2}. \tag{25.6}$$

For A and B insert the Robertson expansion

$$A(r) = 1 + \gamma\frac{2a}{r} + \ldots, \tag{25.7}$$

$$B(r) = 1 - \frac{2a}{r} + 2(\beta - \gamma)\left(\frac{a}{r}\right)^2 + \ldots, \tag{25.8}$$

$$\frac{1}{B(r)} = 1 + \frac{2a}{r} + 2(2 - \beta + \gamma)\left(\frac{a}{r}\right)^2 + \ldots . \quad [27] \tag{25.9}$$

With eq. (25.9), $K(r)$ becomes a quadratic form in $\frac{1}{r}$. Since $\frac{\mathrm{d}\phi}{\mathrm{d}r} = \infty$ for $r = r_\pm$, $K_+ = K_- = 0$. This determines $K(r)$ up to a constant $\tilde{c}$:

$$K(r) = \tilde{c}\left(\frac{1}{r_-} - \frac{1}{r}\right)\left(\frac{1}{r} - \frac{1}{r_+}\right). \tag{25.10}$$

$\tilde{c}$ can be obtained by comparing with (25.6) for $r \to \infty$. With (25.9) one gets

$$\tilde{c} = 1 - (2 - \beta + \gamma)\left(\frac{a}{r_+} + \frac{a}{r_-}\right). \tag{25.11}$$

We get thus the following integral:

$$\phi_+ - \phi_- = \frac{1}{\sqrt{\tilde{c}}}\int_{r_-}^{r_+}\frac{\mathrm{d}r}{r^2}\underbrace{\left(1 + \gamma\frac{a}{r}\right)}_{\text{from}\sqrt{A}\approx 1+\gamma\frac{a}{r}}\left[\left(\frac{1}{r_-} - \frac{1}{r}\right)\left(\frac{1}{r} - \frac{1}{r_+}\right)\right]^{-\frac{1}{2}}. \tag{25.12}$$

We perform the following substitution:

$$\frac{1}{r} = \frac{1}{2}\left(\frac{1}{r_+} + \frac{1}{r_-}\right) + \frac{1}{2}\left(\frac{1}{r_+} - \frac{1}{r_-}\right)\sin\psi; \tag{25.13}$$

[27] $\frac{v^2}{c^2} \approx \frac{a}{r}$; terms $g_{00}u^0u^0 \approx Bc^2$ and $g_{11}u^1u^1 \approx Av^2 \approx Ac^2\frac{a}{r}$ show up both and have thus to be expanded to the same order in $\frac{a}{r}$. Therefore B has to be expanded one order in $\frac{a}{r}$ more than A.

r_+ and r_- correspond to $\psi = \frac{\pi}{2}$ and $\psi = -\frac{\pi}{2}$, respectively. With

$$\mathrm{d}\left(\frac{1}{r}\right) = -\frac{1}{r^2}\mathrm{d}r = \frac{1}{2}\left(\frac{1}{r_+} - \frac{1}{r_-}\right)\cos\psi\,\mathrm{d}\psi, \tag{25.14}$$

$$\frac{1}{r_-} - \frac{1}{r} = \frac{1}{2}\left(\frac{1}{r_-} - \frac{1}{r_+}\right)(1 + \sin\psi), \tag{25.15}$$

$$\frac{1}{r} - \frac{1}{r_+} = \frac{1}{2}\left(\frac{1}{r_-} - \frac{1}{r_+}\right)(1 - \sin\psi), \tag{25.16}$$

we get for the integral

$$\phi_+ - \phi_- = \frac{1}{\sqrt{\tilde{c}}}\int\limits_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \mathrm{d}\psi\left[1 + \gamma\frac{a}{2}\left(\frac{1}{r_-} + \frac{1}{r_+}\right) + \gamma\frac{a}{2}\left(\frac{1}{r_+} - \frac{1}{r_-}\right)\sin\psi\right]. \tag{25.17}$$

Now introduce the parameter p of the ellipse (see figure 16):

$$\frac{2}{p} = \frac{1}{r_+} + \frac{1}{r_-}. \tag{25.18}$$

Integration of eq. (25.17) leads to

$$\phi_+ - \phi_- = \frac{\pi}{\sqrt{\tilde{c}}}\left[1 + \gamma\frac{a}{p}\right] = \pi\left[1 + (2 - \beta + \gamma)\frac{a}{p}\right]\left[1 + \gamma\frac{a}{p}\right] = \pi\left[1 + (2 - \beta + 2\gamma)\frac{a}{p}\right]. \tag{25.19}$$

Precession per orbit for the perihelion is:

$$\Delta\phi = 2(\phi_+ - \phi_-) - 2\pi = \frac{6\pi a}{p}\left(\frac{2 - \beta + 2\gamma}{3}\right). \tag{25.20}$$

In general relativity $\gamma = \beta = 1$ and so $\frac{2-\beta+2\gamma}{3} = 1$. Thus,

$$\boxed{\Delta\phi = \frac{6\pi a}{p}}$$

Consider Mercury: $p = 55 \times 10^6$ km, $2a_\odot \approx 3$ km, $\pi = 180° \times 3600''$ which give $\Delta\phi = \frac{6\pi a}{p} = 0.104''$ (per full orbit). In 100 years Mercury fulfills 415 orbits around the Sun, this way we get $\Delta\phi = 43''$ (per century). For more distant planets (Venus, Earth,...) $\Delta\phi$ is at most $\sim 5''$ per century. Already in 1882, Newcomb found a perihelion precession of $43''$ per century for Mercury. Full perihelion precession amounts to $575''$ per century of which $532''$ are due to the influence of other planets (this within Newtonian theory). One finds

$$\frac{2 - \beta + 2\gamma}{3} = 1.003 \pm 0.005, \tag{25.21}$$

in good agreement with general relativity. So far, the parameters of the Robertson expansion are constrained to $|\gamma - 1| < 3 \times 10^{-4}$ and $|\beta - 1| < 3 \times 10^{-3}$. More recently (radar echoes delay from Cassini spacecraft): $|\gamma - 1| = (2.1 \pm 2.3) \times 10^{-5}$.[28]

[28] B. Bertotti et al. Nature 425, 374 (2003)

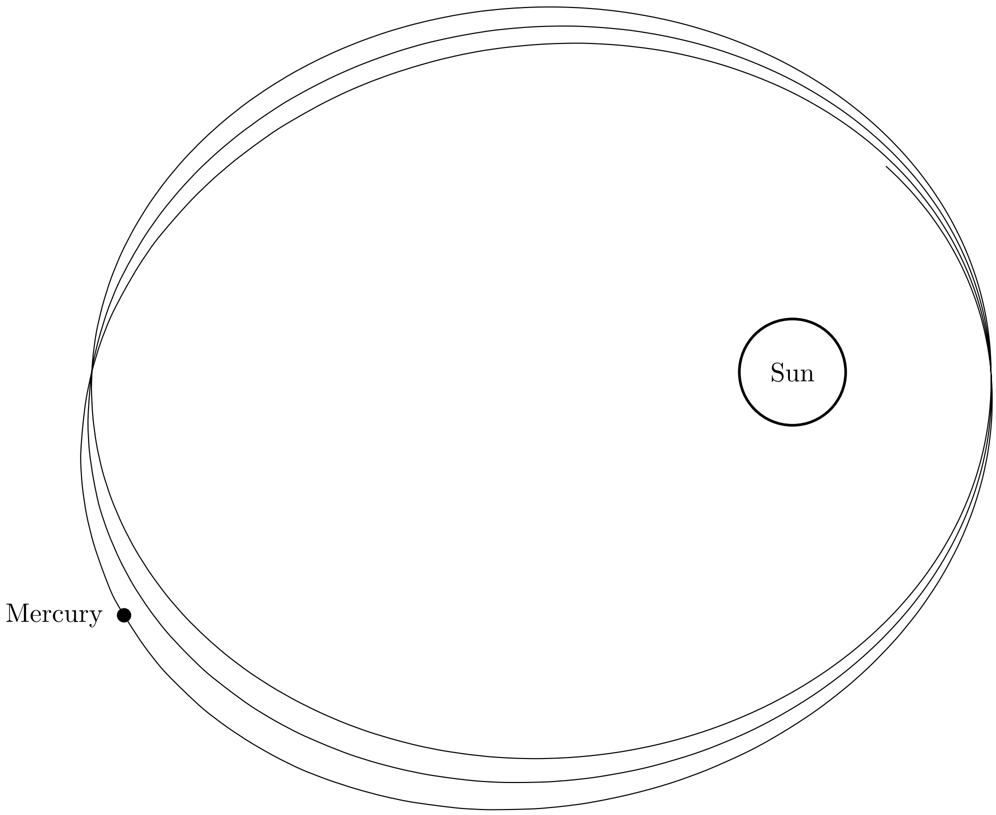


Figure 17: Illustration of the perihelion precession of Mercury (effect strongly exagerated)

25.1 Quadrupole moment of the Sun

A quadrupole moment of the Sun could also influence a perihelion precession of Mercury, that is why one has to study it. The mass quadrupole moment of the Sun (due to its rotation) is

$$Q = \mathcal{J}_2 M_\odot R_\odot^2 \quad \text{with} \quad \mathcal{J}_2 = \frac{2}{5}\frac{R_\parallel - R_\perp}{R_\odot}, \tag{25.22}$$

and

- $R_\parallel$: orthogonal (to $R_\perp$) radius,
- $R_\perp$: radius orthogonal to the plane containing the planet orbits and parallel to the rotation axis of the Sun.

The induced gravitational potential in the planet's orbital plane (which is also the equatorial plane of the Sun) is

$$\phi(r) = -\frac{GM}{r} - \frac{GQ}{2r^3}. \tag{25.23}$$

The additional term has the same r dependence than the additional relativistic term:

$$V_{\text{eff}} = -\frac{GM}{r} + \frac{l^2}{2r^2} - \frac{GMl^2}{c^2r^3}. \tag{25.24}$$

With $l \sim pv$ and $v^2 \sim \frac{GM}{p}$ we can compare the two terms (their relative strength)

$$\frac{GQ}{GM\frac{l^2}{c^2}} \sim \frac{\mathcal{J}_2 R_\odot^2}{p^2\frac{v^2}{c^2}} \sim \frac{\mathcal{J}_2 R_\odot^2}{p\frac{GM}{c^2}} \sim \frac{\mathcal{J}_2 R_\odot^2}{pa}.$$

We see that the full expression for the perihelion precession is given by

$$\Delta\phi = \frac{6\pi a}{p}\left(\frac{2-\beta+2\gamma}{3} + \frac{\mathcal{J}_2 R_\odot^2}{2ap}\right). \tag{25.25}$$

From observations one finds $\mathcal{J}_2 \sim (1-1.7)\times 10^{-7}$; thus the additional term is $\frac{\mathcal{J}_2 R_\odot^2}{2ap} \approx 5\times 10^{-4}$, accounting for at most 1/10 of the error given in (25.21) and is thus negligible.

26 Lie derivative of the metric and Killing vectors

Consider the Lie derivative of the metric tensor $g_{\mu\nu}$ in the direction of the vector K. According to equation (14.2) we get:

$$\mathcal{L}_K g_{\mu\nu} = g_{\mu\nu,\kappa}K^\kappa + g_{\mu\kappa}K^\kappa{}_{,\nu} + g_{\kappa\nu}K^\kappa{}_{,\mu}. \tag{26.1}$$

To rewrite this expression we observe the identities

$$K_\sigma = g_{\sigma\mu}K^\mu,$$

$$K^\kappa{}_{,\nu}g_{\mu\kappa} = \frac{\partial K^\kappa}{\partial x^\nu}g_{\mu\kappa} = \frac{\partial(K^\kappa g_{\mu\kappa})}{\partial x^\nu} - K^\kappa\frac{\partial g_{\mu\kappa}}{\partial x^\nu} = \frac{\partial K_\mu}{\partial x^\nu} - K^\kappa\frac{\partial g_{\mu\kappa}}{\partial x^\nu}.$$

Hence eq. (26.1) can also be written as

$$\begin{aligned}\mathcal{L}_K g_{\sigma\rho} &= \frac{\partial K_\sigma}{\partial x^\rho} + \frac{\partial K_\rho}{\partial x^\sigma} + K^\mu\left[\frac{\partial g_{\rho\sigma}}{\partial x^\mu} - \frac{\partial g_{\mu\sigma}}{\partial x^\rho} - \frac{\partial g_{\rho\mu}}{\partial x^\sigma}\right]\\ &= \frac{\partial K_\sigma}{\partial x^\rho} + \frac{\partial K_\rho}{\partial x^\sigma} - 2K_\mu\Gamma^\mu_{\rho\sigma} \qquad (26.2)\\ &= K_{\sigma;\rho} + K_{\rho;\sigma}.\end{aligned}$$

An infinitesimal coordinate transformation is a *symmetry of the metric* if $\mathcal{L}_K g_{\mu\nu} = 0$, thus if

$$\boxed{K_{\sigma;\rho} + K_{\rho;\sigma} = 0} \tag{26.3}$$

Any 4-vector $K_\sigma(x)$ satisfying this equation will be said to form a *Killing vector.* [29]

Example: Consider a stationary gravitational field, for which there exists coordinates $\{x^\mu\}$ such that the components of $g_{\mu\nu}$ do not depend on $ct = x^0$ (for instance Schwarzschild metric). Let $K^\mu = \delta^\mu{}_0$ with the corresponding vector field $\delta^\mu{}_0\partial_\mu$ ($\to \partial_0$). Inserting K^μ into (26.1) one gets $\mathcal{L}_K g_{\mu\nu} = g_{\mu\nu,0} + 0 + 0 = \frac{\partial}{\partial x^0}g_{\mu\nu} = 0$ (since $g_{\mu\nu}$ does not depend on x^0). K is a Killing vector or *Killing field* or an *infinitesimal isometry.*

[29] Named after 19$^{\text{th}}$ century mathematician Wilhelm Killing

Notice that, due to the properties of the Lie derivative, if K_1 and K_2 are Killing vectors, $\mathcal{L}_{K_1} g_{\mu\nu} = 0$, $\mathcal{L}_{K_2} g_{\mu\nu} = 0$ then $[K_1, K_2]$ is also a Killing vector since

$$[\mathcal{L}_{K_1}, \mathcal{L}_{K_2}] g_{\mu\nu} = \mathcal{L}_{[K_1,K_2]} g_{\mu\nu} = 0. \tag{26.4}$$

We are used to the fact that symmetries lead to conserved quantities: in classical mechanics the angular momentum of a particle moving in a rotationally symmetric field is conserved. In the present context, the concept of "symmetries of a gravitational field" is replaced by "symmetries of the metric" and we therefore expect conserved quantities to be associated with the presence of Killing vectors.

Let K^μ be a Killing vector and $x^\mu(\tau)$ be a geodesic. Then the quantity $K_\mu \dot{x}^\mu$ is constant along the geodesic. Indeed,

$$\begin{aligned}\frac{D}{d\tau}(K_\mu \dot{x}^\mu) &= (\nabla_\nu K_\mu \dot{x}^\nu)\dot{x}^\mu + K_\mu \underbrace{(\nabla_\nu \dot{x}^\mu)}_{=0 \text{ geodesic}} \dot{x}^\nu \\ &= \frac{1}{2} \underbrace{(\nabla_\nu K_\mu + \nabla_\mu K_\nu)}_{=0\ (26.3)} \dot{x}^\mu \dot{x}^\nu = 0.\end{aligned} \tag{26.5}$$

$T^{\mu\nu}$ is the covariantly conserved symmetric energy-momentum tensor with $\nabla_\mu T^{\mu\nu} = 0$. Then $J^\mu = T^{\mu\nu} K_\nu$ is a covariantly conserved current:

$$\nabla_\mu J^\mu = \underbrace{(\nabla_\mu T^{\mu\nu})}_{=0} K_\nu + T^{\mu\nu} \nabla_\mu K_\nu = \frac{1}{2} T^{\mu\nu} \underbrace{(\nabla_\mu K_\nu + \nabla_\nu K_\mu)}_{=0\ (26.3)} = 0,$$

to which we can associate a conserved charge.

27 Maximally symmetric spaces

Maximally symmetric spaces are spaces that admit the maximal number of Killing vectors (which below will turn out to be $\frac{n(n+1)}{2}$ for an n-dimensional space). In the context of the cosmological principle such spaces, which are simultaneously homogeneous ("the same at every point") and isotropic ("the same in every direction"), provide a description of space in a cosmological space-time.

From equation (17.2) we had (from definition of Riemann tensor and covariant derivative)

$$([\nabla_X, \nabla_Y] - \nabla_{[X,Y]}) V^\lambda = R^\lambda{}_{\sigma\mu\nu} X^\mu Y^\nu V^\sigma, \tag{27.1}$$

along with $X = X^\mu \partial_\mu$, $Y = Y^\nu \partial_\nu$ and $\nabla_{Y^\nu \partial_\nu} V^\lambda = Y^\nu \nabla_{\partial_\nu} V^\lambda = Y^\nu \nabla_\nu V^\lambda$, $\nabla_X \nabla_Y V^\lambda = X^\mu \nabla_\mu (Y^\nu \nabla_\nu V^\lambda) = X^\mu (\nabla_\mu Y^\nu) \nabla_\nu V^\lambda + X^\mu Y^\nu \nabla_\mu \nabla_\nu V^\lambda$ etc, we get

$$[\nabla_\mu, \nabla_\nu] V^\lambda = R^\lambda{}_{\sigma\mu\nu} V^\sigma. \tag{27.2}$$

Taking into account the first Bianchi identity, it is possible to find that for a Killing vector K_μ, one has (no proof here)

$$\nabla_\lambda \nabla_\mu K_\nu(x) = R^\rho_{\lambda\mu\nu} K_\rho(x) \tag{27.3}$$

for $x = x_0$. Thus a Killing vector $K^\mu(x)$ is completely determined everywhere by the values of $K^\mu(x_0)$ and $\nabla_\mu K_\nu(x_0)$ at a single point x_0 (think of Taylor expansion). A set of Killing vectors $\{K_\mu^{(i)}(x)\}$ is said to be *independent* if any linear relation of the form

$$\sum_i c_i K_\mu^{(i)}(x) = 0, \tag{27.4}$$

with constant coefficients c_i; implying $c_i = 0$. Since in an n-dimensional space-time there can be at most n linearly independent vectors $K_\mu^{(i)}(x_0)$ at a point, and at most $\frac{n(n-1)}{2}$ independent anti-symmetric matrices $(\nabla_\mu K_\nu(x_0))$, we reach the conclusion that an n-dimensional space-time can have at most

$$n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2} \tag{27.5}$$

independent Killing vectors.

- Homogeneous space is meaning that the n-dimensional space(-time) admits n-translational Killing vectors.
- Isotropic space: $\nabla_\mu K_\nu(x_0)$ is an arbitrary anti-symmetric matrix ($\rightarrow$ rotation). We can choose a set of $\frac{n(n-1)}{2}$ Killing vectors.
- We define a *maximally symmetric space* to be a space with a metric with a maximal number of $\frac{n(n+1)}{2}$ Killing vectors.

The Riemann curvature tensor of a maximally symmetric space becomes simpler. One can show (no proof) that it becomes

$$R_{ijkl} = k(g_{ik}g_{jl} - g_{il}g_{jk}) \tag{27.6}$$

for some constant k. The Ricci tensor then becomes

$$R_{ij}(x) = (n-1)kg_{ij}. \tag{27.7}$$

The Ricci scalar can be obtained to be

$$\mathcal{R}(x) = n(n-1)k, \tag{27.8}$$

and the Einstein tensor

$$G_{ik} = R_{ik} - \frac{1}{2}\mathcal{R}g_{ik} = k\,(n-1)\left(1 - \frac{n}{2}\right) g_{ik}.$$

The Bianchi identity implies that k is a constant in order for ${G_{ik}}^{;k}$ to vanish. We shall deal with space-times in which the metric is spherically symmetric and homogenous on each "plane" of constant time. In our case $n = 4$ and the maximally symmetric subspace has 3 dimensions. Consider first the metric on the 3-dimensional subspace

$$\mathrm{d}\sigma^2 = A(r)\mathrm{d}r^2 + r^2 \underbrace{\mathrm{d}\Omega^2}_{\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2}$$

For the Christoffel symbols, we use the ones for the general form of a static isotropic metric (22.6) with $B(r) = 0$. Since the Christoffel symbols are invariant under an overall sign change of the metric, also the Ricci tensor is and thus one can apply (22.10)-(22.13) with $B(r) = 0$ for this three-dimensional space, without caring for the sign in front of A. Hence we get for equation (27.7)

$$\begin{aligned} R_{rr} &= R_{11} = \frac{A'}{rA}, \\ R_{\theta\theta} &= R_{22} = -\frac{1}{A} + 1 + \frac{rA'}{2A^2}. \end{aligned} \tag{27.9}$$

From eq. (27.7), we have $R_{rr} = 2kA$, and $R_{\theta\theta} = 2kg_{\theta\theta} = 2kr^2$. Thus from equating the two first equation leads us to

$$2kA = \frac{A'}{rA} \Rightarrow A' = 2krA^2, \tag{27.10}$$

while we get for the second one

$$2kr^2 = -\frac{1}{A} + 1 + \frac{rA'}{2A^2} = -\frac{1}{A} + 1 + \frac{2kr^2A^2}{2A^2} = -\frac{1}{A} + 1 + kr^2 \Rightarrow kr^2 = -\frac{1}{A} + 1,$$

which leads to

$$A = \frac{1}{1 - kr^2}, \tag{27.11}$$

and solves also (27.10). Then the metric on the 3-dimensional subspace (maximally symmetric) is

$$\mathrm{d}\sigma^2 = \frac{\mathrm{d}r^2}{1 - kr^2} + r^2 \mathrm{d}\Omega^2. \tag{27.12}$$

It can be shown that k can have the following values: 0, ± 1.

$$k = \begin{cases} +1 & \text{sphere, positive curvature} \\ -1 & \text{hyperbola, negative curvature} \\ 0 & \text{plane, zero curvature} \end{cases}$$

The full metric (with time coordinate) has then the form:

$$\boxed{\mathrm{d}s^2 = c^2 \mathrm{d}t^2 - a^2(t) \left\{ \frac{\mathrm{d}r^2}{1 - kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\, \mathrm{d}\phi^2) \right\}} \tag{27.13}$$

where $a(t)$ is the cosmic scale factor, which has to be determined by solving the Einstein's equations via the matter content of the universe. This metric (first discovered by Friedmann, Lemaître, Roberston and Walker) is a reasonable ansatz for describing the universe. There is good evidence that the universe (on large scales) is surprisingly homogeneous and isotropic (from redshift surveys of galaxies and cosmic microwave background radiation).

28 Friedmann equations

We write the metric (27.13) as follows:

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\tilde{g}_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j, \tag{28.1}$$

where tildes denote 3-dimensional quantities calculated with the metric $\tilde{g}_{ij}$. The Christoffel symbols are given by (notice $\Gamma^\mu_{00} = 0$):

$$\Gamma^i_{jk} = \tilde{\Gamma}^i_{jk}, \qquad \Gamma^i_{j0} = \frac{\dot{a}}{a}\,\delta^i{}_j, \qquad \Gamma^0_{ij} = \dot{a}\,a\,\tilde{g}_{ij}, \tag{28.2}$$

where dot denotes derivation with respect to t. The relevant components of the Riemann tensor are:

$$R^i{}_{0j0} = -\frac{\ddot{a}}{a}\delta^i{}_j, \qquad R^0{}_{i0j} = a\,\ddot{a}\,\tilde{g}_{ij}, \qquad R^k{}_{ikj} = \tilde{R}_{ij} + 2\dot{a}^2\,\tilde{g}_{ij} \tag{28.3}$$

We can make use of $\tilde{R}_{ij} = 2k\tilde{g}_{ij}$ (maximal symmetry of the 3-dimensional subspace) to compute $R_{\mu\nu}$. The non-zero components are then

$$R_{00} = -3\frac{\ddot{a}}{a}, \qquad R_{ij} = (a\ddot{a} + 2\dot{a}^2 + 2k)\tilde{g}_{ij} = -\left(\frac{\ddot{a}}{a} + 2\frac{\dot{a}^2}{a^2} + 2\frac{k}{a^2}\right)g_{ij}, \tag{28.4}$$

where $g_{ij} = -a^2\tilde{g}_{ij}$. The Ricci scalar becomes $\mathcal{R} = -\frac{6}{a^2}(a\ddot{a} + \dot{a}^2 + k)$ and the non-zero components of the Einstein tensor are

$$G_{00} = 3\left(\frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right), \qquad G_{0i} = 0, \qquad G_{ij} = \left(2\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right)g_{ij}. \tag{28.5}$$

Next we have to specify the matter content. We treat here the universe as non-interacting particles or a *perfect fluid*. A perfect fluid has energy-momentum tensor (19.8)

$$T^{\mu\nu} = \left(\frac{p}{c^2} + \rho\right)u^\mu u^\nu - g^{\mu\nu}p, \tag{28.6}$$

where p is the pressure, ρ the energy density and u^μ the velocity field of the fluid ($u^\mu = (c, 0, 0, 0)$ in a comoving coordinate system). The trace of the energy-momentum tensor is then

$$T^\mu{}_\mu = \rho c^2 - 3p. \tag{28.7}$$

The *equation of state* is $p = p(\rho)$ and in particular one assumes

$$p = w\rho, \tag{28.8}$$

where w is the equation of state parameter.

Examples:

- For non-interacting particles we have $p = 0$, $w = 0$. Such matter is referred to as *dust*. The energy-momentum tensor is $T^{\mu\nu} = \rho\,u^\mu u^\nu$.
- For radiation the energy-momentum tensor is (like in Maxwell's theory) traceless and hence *radiation* has the equation of state

  $$p = \frac{1}{3}\rho, \tag{28.9}$$

  thus $w = \frac{1}{3}$.

- As we will see, a cosmological constant Λ corresponds to a "matter" contribution with $w = -1$.

The conservation law $T^{\nu\mu}{}_{;\mu} = 0$ implies $T^{0\mu}{}_{;\mu} = 0$ or $\partial_\mu T^{\mu 0} + \Gamma^\mu_{\mu\nu} T^{\nu 0} + \Gamma^0_{\mu\nu} T^{\mu\nu} = 0$. For a perfect fluid: $\partial_t \rho(t) + \Gamma^\mu_{\mu 0}\rho + \Gamma^0_{00}\rho + \Gamma^0_{ij} T^{ij} = 0$ (with $i, j = 1, 2, 3$). Inserting the expressions for the Christoffel symbols (28.2) we get:

$$\dot{\rho} = -3(\rho + p)\frac{\dot{a}}{a}. \tag{28.10}$$

For dust ($p = 0$):

$$\frac{\dot{\rho}}{\rho} = -3\frac{\dot{a}}{a}. \tag{28.11}$$

Integration gives $\rho a^3 =$ constant or $\rho \propto a^{-3}$. For a radiation dominated universe we get

$$p = \frac{\rho}{3} \Rightarrow \frac{\dot{\rho}}{\rho} = -4\frac{\dot{a}}{a}. \tag{28.12}$$

Integration gives $\rho a^4 =$ constant or $\rho \propto a^{-4}$. More generally for (28.8) one gets:

$$\rho a(t)^{3(1+w)} = \text{constant}. \tag{28.13}$$

The Einstein equations with Λ (equation (20.10)) are

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} + \Lambda g_{\mu\nu}.$$

Using (19.8) and that $u^\mu = (c, 0, 0, 0)$ in a comoving coordinate system, let us write down the 00-component and the ij-component of this equation:

$$3\left(\frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) = 8\pi G\rho + \Lambda, \tag{28.14}$$

$$\left(2\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) g_{ij} = \left(-8\pi G p + \Lambda\right) g_{ij}. \tag{28.15}$$

One has in addition equation (28.10) from the conservation law. Using the first equation to eliminate $\frac{\dot{a}^2}{a^2} + \frac{k}{a^2}$ from the second one, one obtains the *Friedmann equations*:

$$\frac{\dot{a}^2}{a^2} + \frac{k}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3}, \tag{28.16}$$

$$-3\frac{\ddot{a}}{a} = 4\pi G(\rho + 3p) - \Lambda, \tag{28.17}$$

$$\dot{\rho} = -3(\rho + p)\frac{\dot{a}}{a}. \tag{28.18}$$

Notice that one could also use the form (20.8) of the Einstein equations to derive the above equations. Introducing the *Hubble parameter*: $H(t) = \frac{\dot{a}(t)}{a(t)}$ and the *deceleration parameter*: $q(t) = -\frac{a(t)\ddot{a}(t)}{\dot{a}^2(t)}$, with their present day values denoted by $H_0 = H(t_0)$ and $q_0 = q(t_0)$, where t_0 is the age of the universe;

we get instead

$$\boxed{\begin{aligned} H^2 &= \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3} \\ q &= \frac{1}{3H^2}(4\pi G(\rho + 3p) - \Lambda) \\ \frac{\mathrm{d}}{\mathrm{d}t}(\rho a^3) &= -3Hpa^3 \end{aligned}} \tag{28.19}$$

In the case of $\Lambda = 0$, we define a *critical density* $\rho_{\text{crit}} = \frac{3H^2}{8\pi G}$ and a *density parameter* $\Omega = \frac{\rho}{\rho_{\text{crit}}}$. Then

$$\begin{cases} \rho < \rho_{crit} \Leftrightarrow k = -1 & \text{open universe} \\ \rho = \rho_{crit} \Leftrightarrow k = 0 & \text{flat universe} \\ \rho > \rho_{crit} \Leftrightarrow k = +1 & \text{close universe} \end{cases}$$

Let us now assume that the density is a combination of dust (that we shall simply denote as "matter") and radiation: $\rho = \rho_m + \rho_r$. Moreover, we assume that $\rho_m \sim a^{-3}$ and $\rho_r \sim a^{-4}$. This is valid if radiation and matter are decoupled, or if one density is much bigger that the other one (notice that in today's universe $\rho_m \gg \rho_r$). Let us introduce the constants $K_m = \frac{8\pi G}{3}\rho_m a^3$ and $K_r = \frac{8\pi G}{3}\rho_r a^4$. Inserting them into equation (28.16) leads to

$$\dot{a}^2 - \frac{K_r}{a^2} - \frac{K_m}{a} - \frac{1}{3}\Lambda a^2 = -k. \tag{28.20}$$

This equation reads as

$$\dot{a}^2 + V(a) = -k,$$

where

$$V(a) = -\frac{K_r}{a^2} - \frac{K_m}{a} - \frac{1}{3}\Lambda a^2 \tag{28.21}$$

plays the role of an effective potential, see figure 18.

Consider the solution for $a \to 0$: in that case the terms $\frac{K_r}{a^2}$ and $\frac{K_m}{a}$ dominate and the behavior does not depend neither on k nor on Λ.

$$\dot{a}^2 \approx \frac{K_r}{a^2} \quad \to \quad a(t) \sim \sqrt{t} \tag{28.22}$$

$$\dot{a}^2 \approx \frac{K_m}{a} \quad \to \quad a(t) \sim t^{\frac{2}{3}} \tag{28.23}$$

For $a \to 0$, $\dot{a}$ goes to ∞. If $K_r \neq 0$, then for $a \to 0$ $a \sim \sqrt{t}$. From figure 18 we can discriminate different types of solutions:

1. For $\Lambda < 0$, there is for all k-values a maximal a_{max}. Thus there will be a periodic solution going between $a = 0$ and a_{max}.

2. $\Lambda = 0$ (*Einstein-de Sitter universe*): for $k = 1$ there is a bounded solution as in the case 1. For $k = 0$ the expansion velocity goes towards 0, instead for $k = -1$ it goes towards a constant value.

3. For $\Lambda > 0$ there are several cases:

a) $\Lambda = \Lambda_{crit}$ and $k = 1$. The value of Λ_{crit} is obtained for the horizontal line $-k$ (figure 18) just going through the maximum of the potential. Assuming $K_r \approx 0$ (as in today's universe) we get with $\dot{a} = 0$, $V = -1 \rightarrow \frac{dV}{da} = 0$:

$$\Lambda = \Lambda_{crit} = \frac{4}{9K_m^2} \quad and \quad a = a_{stat} = \frac{3K_m}{2}. \tag{28.24}$$

This corresponds to the static *Einstein solution*. Einstein introduced a cosmological constant in order to get such a static solution, which is however unstable. Small perturbations lead either to a contraction or to an exponential growth.

b) $\Lambda < \Lambda_{crit}$ and $k = 1$. The horizontal line $-k$ intercepts the curve $V(a)$ in two points a_1 and a_2. We obtain either a periodic solution between 0 and a_1 or an unbounded solution with $a > a_2$.

c) $\Lambda = \Lambda_{crit}(1 + \epsilon)$ and $k = 1$. For $0 < \epsilon \ll 1$ the horizontal line $-k$ lies just above the maximum value of $V(a)$. Thus the expansion velocity $\dot{a}$ will be very low in this region (*Lemaître universe*).

d) $\Lambda > \Lambda_{crit}$ and $k = 1$. The line $-k$ does not intercept the $V(a)$ curve. Around the maximum of $V(a)$ the expansion is lowered.

e) $\Lambda > 0$ and $k = -1, 0$: as in the previous case, but there may be less deceleration in the region of the maximum of $V(a)$.

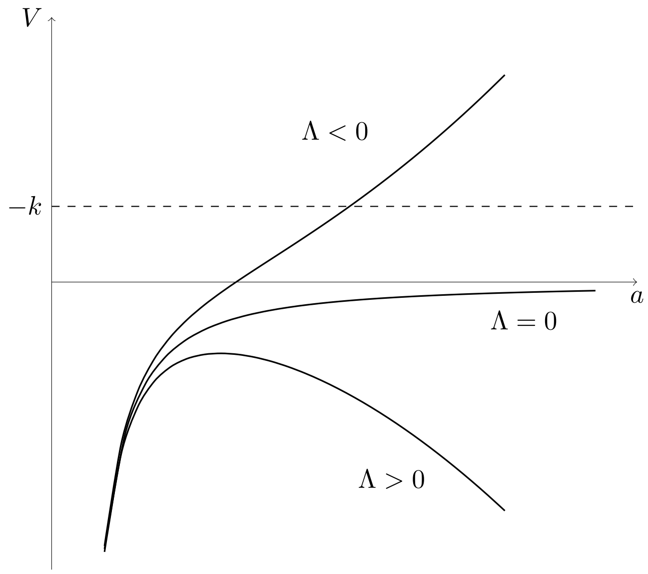


Figure 18: Sketch of the "effective potential" $V(a)$ for different values of Λ.

For $\Lambda > 0$, all solutions are unbounded. In the limit $a \rightarrow \infty$, $V(a)$ is dominated by the Λ-term:

$$\dot{a}^2 \approx \frac{\Lambda}{3} a^2,$$

thus

$$a(t) \sim \exp\left(\sqrt{\frac{\Lambda}{3}} t\right). \tag{28.25}$$

The expansion is exponentially accelerated.

According to the currently most accepted model (so-called Λ*CDM-model*), the main contributions to the density are

- ordinary matter (baryons) $\Omega_{\text{baryons}} \sim 0.05$
- dark matter $\Omega_{\text{DM}} \sim 0.27$
- cosmological constant $\Omega_\Lambda = \frac{\rho_\Lambda}{\rho_{\text{crit}}} \sim 0.68$ with $\rho_\Lambda = \frac{\Lambda}{8\pi G}$ (also called dark energy).

Moreover, the universe seems to be almost flat: $k \approx 0$. These cosmological parameters would thus correspond to the case 3e) of the previous discussion.

Finally, H_0^{-1} is related to the age of the universe. $H_0 \sim 67.80 \pm 0.77 \, \frac{\text{km}}{\text{sec/Mpc}}$ gives an age of ~ 13.8 billion years.[30]

[30] Planck 2013 results. XVI. Cosmological parameters - http://arxiv.org/abs/1303.5076