General Relativity

HS 2010

Lecture by Prof. Dr. Philippe Jetzer

typeset and revision: Felix Hähl

May 6, 2011

Contents

Part I
Newtonian Gravity and Special Relativity

1 Newton's Theory of Gravitation

The first mathematical theory of gravitation was published by Newton around 1687 in his "Principia". According to Newton, the trajectory of N point masses which attract each other via gravitation is described by the differential equation

$$m_i \frac{d^2\boldsymbol{r}_i}{dt^2} = -G\sum_{j\neq i} \frac{m_i m_j(\boldsymbol{r}_i - \boldsymbol{r}_j)}{|\boldsymbol{r}_i - \boldsymbol{r}_j|^3} \qquad (i = 1, ..., N) \tag{1.1}$$

where $\boldsymbol{r}_i(t)$ describes the position of the i-th mass m_i at time t.
Experiments show that

$$G = 6.6743(7) \cdot 10^{-11}\mathrm{m}^3\mathrm{kg}^{-1}\mathrm{s}^{-2}. \tag{1.2}$$

Note that the gravitational force is *always attractive* and directed along $\boldsymbol{r}_j - \boldsymbol{r}_i$. Furthermore it is proportional to the product of the two masses.
The scalar gravitational potential is

$$\Phi(\boldsymbol{r}) = -G\int d^3r' \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \tag{1.3}$$

with $\rho(\boldsymbol{r}) = \sum_j m_j\delta(\boldsymbol{r} - \boldsymbol{r}_j)$. This reduces to $\Phi(\boldsymbol{r}) = -G\sum_j \frac{m_j}{|\boldsymbol{r}-\boldsymbol{r}_j|}$ for pointlike masses.
So the equation of motion for a particle of mass m reads:

$$m\frac{d^2\boldsymbol{r}}{dt^2} = -m\boldsymbol{\nabla}\Phi(\boldsymbol{r}). \tag{1.4}$$

According to Eq. (1.3) the field $\Phi(\boldsymbol{r})$ is determined by the masses of the other particles. The corresponding field equation is given by

$$\Delta\Phi(\boldsymbol{r}) = 4\pi G\rho(\boldsymbol{r}). \tag{1.5}$$

The source of the field is the *mass density*.
Obviously Eq. (1.5) has the same mathematical structure as the field equation in electrostatics:

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

with charge density ρ_e and electrostatic potential ϕ_e. To this equation we associate the non-relativistic equation of motion for a particle with charge q

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

We want to compare Eqs. (1.4) and (1.7). In Eq. (1.7) the charge q does not depend on the mass m and acts as a coupling constant. In analogy we interpret the m on the LHS of Eq. (1.4) as the same **inertial mass** as on the LHS of Eq. (1.7) and the m on the RHS of Eq. (1.4) as an *a priori different* **gravitational mass** acting as a coupling parameter. Experimentally it is verified that inertial and gravitational mass are the same to an accuracy of $\sim 10^{-13}$. In Newton's theory this (presumable) identity of inertial and gravitational mass is just by chance but in GR it will turn out to be a crucial fact.

For many applications Newtonian gravity is good enough. However it is clear that the above equations cannot always be true since they imply instantaneous action at a distance (a change of the mass distribution changes the field everywhere instantaneously). This problem has already been known to Newton himself but it took more than 300 years until somebody (namely Einstein) could think of *how* to solve this problem correctly and find a more general theory of which Newtonian gravity is a special case.

1.1 Goals of General Relativity

From an experimental point of view, the necessity of a more general theory of gravity is clear. For example the problem of Mercury's perihelion was the first experimental fact that couldn't be explained using only Newtonian gravity and it was the first problem to be solved by Einstein after his discovery of general relativity.
Heuristically one could think of the possibility to make a transition from Newtonian gravity (NG) to general relativity (GR) in an analogous way as electrostatics (ES) is generalized to electrodynamics (ED). The transition from ES to ED is done, in a formal sense, by considering inertial frames in relativistic motion relative to each other. This leads to the conclusion that charge densities ρ_e and current densities $\boldsymbol{j}$ have to be related and have to transform into each other. Formally this corresponds to the following generalization:

$$\rho_e \qquad \longrightarrow \qquad j^\alpha = (\rho_e c, \rho_e \boldsymbol{v}). \tag{1.8}$$

One treats the potentials in an analogous way:

$$\Phi_e \qquad \longrightarrow \qquad A^\alpha = (\Phi_e, \boldsymbol{A}). \tag{1.9}$$

Finally, "completing" Δ to $\Box$,

$$\Delta \qquad \longrightarrow \qquad \Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta, \tag{1.10}$$

yields the transition from the static Poisson equation to a relativistic field equation:

$$\Delta\Phi_e = -4\pi\rho_e \qquad \longrightarrow \qquad \Box A^\alpha = \frac{4\pi}{c} j^\alpha. \tag{1.11}$$

Note that the 0-component of the new equation reduces in the static case to the Poisson equation.
Since NG and ES have the same mathematical structure, one can try to generalize NG in a similar way. The transition $\Delta \to \Box$ is straightforward, but

generalizing the mass density ρ is not so easy: In the case of ES/ED the charge q of a particle is independent of how the particle moves. The mass m, on the other hand, depends on velocity ($m = \gamma m_0$). This is just the fact that q is a Lorentz scalar but m isn't. Since the charge is a Lorentz scalar, the charge density $\rho_e = \frac{\Delta q}{\Delta V}$ transforms as the 0-component of a Lorentz vector j^α (because $\frac{1}{\Delta V}$ transforms into $\gamma \frac{1}{\Delta V}$ under Lorentz transformations). The mass is not a Lorentz scalar, but since the energy $E = mc^2$ transforms as the 0-component of a Lorentz vector, the mass density ρ transforms as the 00-component of a Lorentz tensor: the energy-momentum tensor $T_{\alpha\beta}$. Thus instead of Eq. (1.8) we should generalize in the following way:

$$\rho \qquad \longrightarrow \qquad \begin{pmatrix} \rho c^2 & \rho c \boldsymbol{v}^T \\ \rho c \boldsymbol{v} & \rho \boldsymbol{v} \cdot \boldsymbol{v}^T \end{pmatrix} \sim T_{\alpha\beta}. \tag{1.12}$$

Accordingly we have to generalize the gravitational potential Φ to some tensor depending on two indices, which we shall call the metric tensor $g^{\alpha\beta}$:

$$\Delta\Phi = 4\pi G \rho \qquad \longrightarrow \qquad \Box g^{\alpha\beta} \propto G T^{\alpha\beta}. \tag{1.13}$$

This heuristic approach will turn out not to be completely the right track but Eq. (1.13) will turn out to be a weak-field approximation of GR! From Eq. (1.13) it is also immediately obvious that we will expect to find wave solutions in analogy to ED. But since Eq. (1.13) will turn out to be not exact, gravitational waves will only be approximate solutions in GR, whereas electromagnetic waves are exact solutions in ED.

Note another difference between the transition ES $\to$ ED and NG $\to$ GR: due to the equivalence of mass and energy, the gravitational field, which carries energy, also carries a mass. Therefore the gravitational field is itself a gravitational source. This is the reason for the non-linearities in GR. In ED the electromagnetic waves also carry energy but they do not carry a charge, so they are not a source of new fields.

2 Special Relativity

2.1 Galilei Transformations

In order to formulate the relativity principle we need the notion of a **coordinate system** which is a reference system with a defined choice of coordinates (e.g. cartesian coordinates $x(t), y(t), z(t)$). Important examples are *inertial* reference systems (IS). From a very practical point of view these are systems which move relative to the distant (thus fixed) stars in the sky at a constant velocity. Non-inertial systems are reference systems that are accelerated with respect to an IS. Newton's equations of motion are valid in IS.

The principle of **Galilei invariance** states that all IS are equivalent, i.e. all physical laws are valid in every IS. The laws are **covariant** (i.e. forminvariant) under transformations which lead from IS to IS'. If an event in IS is described by coordinates x^i, t, then the same event is described in IS' by coordinates x'^i, t' and the general Galilei transformation connecting IS and IS' takes the form

$$x'^i = R^i_k x^k + v^i t + b^i, \qquad\qquad t' = t + \tau. \tag{2.1}$$

Obviously $\boldsymbol{b}$ describes a translation, $\boldsymbol{v}$ is the relative velocity between IS and IS'. The relative rotation of the coordinate systems is described by $(R^i_k) \in SO(3)$. Note that $R^i{}_j (R^T)^j{}_k = \delta^i{}_k$ ensures that the line element

$$ds^2 = dx^2 + dy^2 + dz^2 \tag{2.2}$$

remains invariant. Eqs. (2.1) defines a 10 paramter group of transformations, the **Galilei group**.

2.2 Lorentz Transformations

As we know, the laws of mechanics are invariant under Galilei transformations whereas Maxwell's equations are not. The reason for the latter fact is that Maxwell's equations contain the constant speed of light c - a concept which is incompatible with the philosophy of Galilei invariance. If we demand all physical laws (including Maxwell's equations) to be valid in every IS, we are led to Lorentz transformations instead of Galilei transformations. Accordingly the laws of mechanics have to be modified.

In order to describe Lorentz transformations we use Minkowski coordinates $x = (ct, \boldsymbol{r})$. Let an event be at x^α in IS and at x'^α in IS'. The general Poincaré transformation from x^α to x'^α reads

$$x'^\alpha = \Lambda^\alpha{}_\beta x^\beta + b^\alpha. \tag{2.3}$$

where $\Lambda^\alpha{}_\beta$ describes the relative rotation and movement of the inertial frames. Due to homogeneity of space, $\Lambda^\alpha{}_\beta$ has to be linear, i.e. it does not depend on x^α.

As before $\Lambda^\alpha{}_\beta$ has to be such that the square of the line element remains unchanged under the transformation[1]:

$$ds^2 = \eta_{\alpha\beta} dx^\alpha dx^\beta = c^2 dt^2 - d\boldsymbol{r}^2. \tag{2.4}$$

A four-dinemsional space with this line element is called a **Minkowski space**. In the case of light we have

$$d\tau^2 = dt^2 - \frac{dx^2 + dy^2 + dz^2}{c^2} = 0 \qquad \Rightarrow \qquad \left| \frac{d\boldsymbol{r}}{dt} \right|^2 = c^2. \tag{2.5}$$

The fact that $d\tau = 0$ implies $d\tau' = 0$ ensures that the speed of light is the same in all coordinate systems.

Inserting (2.3) into the invariance condition $ds^2 = ds'^2$ we obtain

$$ds'^2 = \eta_{\alpha\beta} dx'^\alpha dx'^\beta = \eta_{\alpha\beta} \Lambda^\alpha{}_\gamma \Lambda^\beta{}_\delta dx^\gamma dx^\delta \stackrel{!}{=} \eta_{\gamma\delta} dx^\gamma dx^\delta \tag{2.6}$$

$$\Rightarrow \qquad \Lambda^\alpha{}_\gamma \Lambda^\beta{}_\delta \eta_{\alpha\beta} = \eta_{\gamma\delta} \tag{2.7}$$

$$\Rightarrow \qquad \Lambda^T \eta \Lambda = \eta. \tag{2.8}$$

Rotations are special cases and are contained in Λ:

$$x'^\alpha = \Lambda^\alpha{}_\beta x^\beta \tag{2.9}$$

[1] We use the convention $\eta_{\alpha\beta} = \mathrm{diag}(1, -1, -1, -1)$

where $\Lambda^i{}_k = R^i{}_k$, $\Lambda^0{}_0 = 1$, $\Lambda^i{}_0 = \Lambda^0{}_i = 0$. The entire group of Lorentz transformations is called **Poincaré group** (10 parameters). The **homogeneous Lorentz group** consists of those Poincaré transformations with $a^\alpha = 0$.
Translations and rotations form a subgroup of the Galilei group as well as the Lorentz group and their properties are the same as before. So the interesting transformations are those which describe relative velocities between IS and IS'.
A Lorentz boost in the x-direction takes the form

$$\Lambda^\alpha{}_\beta = \begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & 0 & 0 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.10}$$

with (according to (2.8))

$$(\Lambda^0{}_0)^2 - (\Lambda^1{}_0)^2 = 1 \tag{2.11}$$
$$(-\Lambda^1{}_1)^2 - (\Lambda^0{}_1)^2 = -1 \tag{2.12}$$
$$\Lambda^0{}_0\Lambda^0{}_1 - \Lambda^1{}_0\Lambda^1{}_1 = 0 \tag{2.13}$$

which can be written as

$$\begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 \\ \Lambda^1{}_0 & \Lambda^1{}_1 \end{pmatrix} = \begin{pmatrix} \cosh\psi & -\sinh\psi \\ -\sinh\psi & \cosh\psi \end{pmatrix}. \tag{2.14}$$

For the origin of IS' we have $x'^1 = 0 = \Lambda^1{}_0 ct + \Lambda^1{}_1 vt$ which gives us

$$\tanh\psi = -\frac{\Lambda^1{}_0}{\Lambda^1{}_1} = \frac{v}{c} =: \beta. \tag{2.15}$$

As a function of velocity this yields

$$\Lambda^0{}_0 = \Lambda^1{}_1 = \gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{2.16}$$
$$\Lambda^0{}_1 = \Lambda^1{}_0 = \frac{-\beta}{\sqrt{1-\beta^2}}. \tag{2.17}$$

Finally this transformation ("Lorentz boost") reads

$$x' = \gamma(x - \beta ct), \qquad y' = y, \qquad z' = z, \qquad ct' = \gamma(ct - \beta x). \tag{2.18}$$

Note that for $v \ll c$ this reduces to a Galilei transformation.
The object $\psi = \mathrm{arctanh}\beta$ is called **rapidity**. We infer that for the addition of parallel velocities we have

$$\psi = \psi_1 + \psi_2 \qquad \rightarrow \qquad \beta = \frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2}. \tag{2.19}$$

2.3 Proper Time

The time coordinate t in IS is the time measured by clocks at rest in IS. The time τ shown by a clock which moves with velocity $\boldsymbol{v}(t)$ is called **proper time**. Consider an IS' at a given moment t_0 which moves at a constant velocity $\boldsymbol{v}_0(t_0)$ with respect to IS such that the clock can be considered at rest in IS' during an interval dt'. We then have

$$d\tau = dt' = \frac{1}{\gamma}dt, \tag{2.20}$$

which follows immediately from Eq. (2.18) by setting $x = v_0 t$.
An infinitesimal time later at time $t+dt$, the clock rests in yet another inertial system IS" which moves with velocity $\boldsymbol{v}_0 = \boldsymbol{v}(t+dt)$ and so on. Summing up all infinitesimal time intervals $d\tau$ yields the total proper time which is measured by a clock moving in IS:

$$\boxed{\tau = \int_{t_1}^{t_2} dt\sqrt{1-\frac{v(t)^2}{c^2}}} \tag{2.21}$$

2.4 Relativistic Mechanics

2.4.1 Equations of Motion

We want to recapitulate the relativistic generalization of Newton's equations of motion. In order to do so, we replace the spatial velocity $v^i = \frac{dx^i}{dt}$ by the 4-velocity

$$u^\alpha = \frac{dx^\alpha}{d\tau}. \tag{2.22}$$

Because $d\tau = \frac{ds}{c}$ is Lorentz-invariant and x'^α transforms like a vector ($x'^\alpha = \Lambda^\alpha{}_\beta x^\beta$) this 4-velocity also shows the correct Lorentz-covariant transformation behaviour:

$$u'^\alpha = \Lambda^\alpha{}_\beta u^\beta. \tag{2.23}$$

All quantities which transform in this way under coordinate transformations $x'^\alpha = \Lambda^\alpha{}_\beta x^\beta$ are called **Lorentz 4-vector**. The relativistic equation of motion reads

$$m\frac{du^\alpha}{d\tau} = f^\alpha. \tag{2.24}$$

Due to Lorentz covariance we can perform a Lorentz transformation and the equation of motion in another inertial system has the same form,

$$m\frac{du'^\alpha}{d\tau} = f'^\alpha. \tag{2.25}$$

As is necessary for consistency, Eq. (2.25) reduces to Newtons law if $v \ll c$. This is because the quantities change according to

$$m\frac{du^\alpha}{d\tau} \quad \xrightarrow{v\ll c} \quad m(0, \frac{d\boldsymbol{v}}{dt}) \tag{2.26}$$

$$f^\alpha = (f^0, \boldsymbol{f}) \quad \xrightarrow{v\ll c} \quad (0, \boldsymbol{K}). \tag{2.27}$$

Note that f'^{α} is determined in any IS through a suitable LT such that $f'^{\alpha} = \Lambda^{\alpha}{}_{\beta} f^{\beta}$.
As an example we find for a boost with velocity v^1 in x-direction that

$$f'^0 = \gamma \frac{v^1 K^1}{c}, \qquad f'^1 = \gamma K^1, \qquad f'^2 = K^2, \qquad f'^3 = K^3. \tag{2.28}$$

For a boost with velocity $\boldsymbol{v}$ in a general direction one can show

$$f'^0 = \gamma \frac{\boldsymbol{v} \cdot \boldsymbol{K}}{c}, \qquad \boldsymbol{f}' = \boldsymbol{K} + (\gamma - 1)\boldsymbol{v} \frac{\boldsymbol{v} \cdot \boldsymbol{K}}{v^2}. \tag{2.29}$$

So if we want to determine the 4-force in a general IS, we simply write it down in a rest frame where it is trivial (Newtonian) and perform a corresponding Lorentz transformation.

2.4.2 Energy and Momentum

Of course, the 4-momentum

$$p^{\alpha} = m u^{\alpha} = m \frac{dx^{\alpha}}{d\tau} = m\gamma \frac{dx^{\alpha}}{dt} = m\gamma(c, \boldsymbol{v}) = (\frac{E}{c}, \boldsymbol{p}) \tag{2.30}$$

is a Lorentz-vector, as well. With Eq. (2.29) the 0-component of Eq. (2.25) in the limit $v \ll c$ becomes

$$\frac{dE}{dt} = \boldsymbol{v} \cdot \boldsymbol{K}. \tag{2.31}$$

Here $\boldsymbol{v} \cdot \boldsymbol{K}$ is just the power given to the particle. This justifies to call the quantity "E" an energy.
From $ds^2 = c^2 d\tau^2 = \eta_{\alpha\beta} dx^{\alpha} dx^{\beta}$ it follows that $\eta_{\alpha\beta} p^{\alpha} p^{\beta} = m^2 c^2$. This is just the energy-momentum relation

$$E^2 = m^2 c^4 + c^2 \boldsymbol{p}^2 \tag{2.32}$$

which again reduces to classical mechanics in the non-relativistic limit:

$$E = \sqrt{m^2 c^4 + c^2 \boldsymbol{p}^2} \simeq \begin{cases} mc^2 + \frac{\boldsymbol{p}^2}{2m} & \text{if } v \ll c, \quad p \ll mc \\ cp & \text{if } v \simeq c, \quad p \gg mc \end{cases} \tag{2.33}$$

For particles with no rest mass (e.g. photons) the second equation $E = cp$ is exact.

2.4.3 Equivalence of Mass and Energy

One can divide the energy into the rest energy and the kinetic energy:

$$\begin{aligned} E_0 &= mc^2 & (2.34) \\ E_{\text{kin}} &= E - E_0 = E - mc^2 = (\gamma - 1)mc^2. & (2.35) \end{aligned}$$

2.5 Tensors in Minkowski space

We want to discuss the transformation properties of physical quantities under Lorentz transformations. As we have seen, a 4-vector transforms as

$$v'^{\alpha} = \Lambda^{\alpha}{}_{\beta} v^{\beta}. \tag{2.36}$$

We call v^{α} (with upper index) a **contravariant** 4-vector. The coordinate system transforms in the same way, according to $x'^{\alpha} = \Lambda^{\alpha}{}_{\beta} x^{\beta}$.
A **covariant** 4-vector is defined through

$$v_{\alpha} = \eta_{\alpha\beta} v^{\beta}. \tag{2.37}$$

The matrix $\eta^{\alpha\beta}$ is defined as the inverse of $\eta_{\alpha\beta}$ such that $\eta^{\alpha\beta}\eta_{\beta\gamma} = \delta^{\alpha}{}_{\gamma}$. Given the transformation behaviour of contravariant vectors we find for covariant vectors

$$v'_{\alpha} = \eta_{\alpha\beta} v'^{\beta} = \eta_{\alpha\beta}\Lambda^{\beta}{}_{\gamma} v^{\gamma} = \eta_{\alpha\beta}\Lambda^{\beta}{}_{\gamma}\eta^{\gamma\delta} v_{\delta} = \overline{\Lambda}^{\delta}{}_{\alpha} v_{\delta} \tag{2.38}$$

with $\overline{\Lambda}^{\delta}{}_{\alpha} = \eta_{\alpha\beta}\Lambda^{\beta}{}_{\gamma}\eta^{\gamma\delta}$. Using Eq. (2.7) we find

$$\overline{\Lambda}^{\gamma}{}_{\alpha}\Lambda^{\alpha}{}_{\beta} = \eta_{\alpha\delta}\eta^{\gamma\varepsilon}\Lambda^{\delta}{}_{\varepsilon}\Lambda^{\alpha}{}_{\beta} = \eta^{\gamma\varepsilon}\eta_{\varepsilon\beta} = \delta^{\gamma}{}_{\beta} \tag{2.39}$$

and similarly $\Lambda^{\beta}{}_{\alpha}\overline{\Lambda}^{\alpha}{}_{\gamma} = \delta^{\beta}{}_{\gamma}$. In matrix notation this means that $\overline{\Lambda} = \Lambda^{-1}$ or with Lorentz indices: $\overline{\Lambda}^{\alpha}{}_{\beta} = (\Lambda^{-1})^{\alpha}{}_{\beta} = \Lambda_{\beta}{}^{\alpha}$. Therefore a contravariant vector transforms with Λ whereas a covariant vector transforms with Λ^{-1}.
The scalar product of two 4-vectors v_{α}, u_{β} is

$$v_{\alpha}u^{\alpha} = v^{\alpha}u_{\alpha} = \eta^{\alpha\beta} v_{\alpha} u_{\beta} \tag{2.40}$$

which is Lorentz-invariant if the vectors are contravariant and covariant, respectiveley:

$$v'^{\alpha}u'_{\alpha} = \underbrace{\Lambda^{\alpha}{}_{\beta}\overline{\Lambda}^{\delta}{}_{\alpha}}_{=\eta^{\delta}_{\beta}} v^{\beta}u_{\delta} = v^{\beta}u_{\beta}. \tag{2.41}$$

Note that $\frac{\partial}{\partial x^{\alpha}}$ transforms like a covariant vector. This is because if the coordinates transform as $x'^{\alpha} = \Lambda^{\alpha}{}_{\beta} x^{\beta}$, then we immediately have

$$x^{\beta} = \Lambda_{\alpha}{}^{\beta} x'^{\alpha} \qquad \Rightarrow \qquad \frac{\partial x^{\beta}}{\partial x'^{\alpha}} = \Lambda_{\alpha}{}^{\beta} \tag{2.42}$$

which immediately gives a covariant transformation behaviour for $\frac{\partial}{\partial x^{\alpha}}$:

$$\frac{\partial}{\partial x'^{\alpha}} = \frac{\partial x^{\beta}}{\partial x'^{\alpha}}\frac{\partial}{\partial x^{\beta}} = \Lambda_{\alpha}{}^{\beta}\frac{\partial}{\partial x^{\beta}}. \tag{2.43}$$

We introduce the notation

$$\underbrace{\partial_{\alpha} \equiv \frac{\partial}{\partial x^{\alpha}}}_{\text{covariant}} \qquad \text{and} \qquad \underbrace{\partial^{\alpha} \equiv \frac{\partial}{\partial x_{\alpha}}}_{\text{contravariant}}. \tag{2.44}$$

The D'Alembert operator reads

$$\Box \equiv \partial^\alpha \partial_\alpha \tag{2.45}$$

which can immediately seen to be a Lorentz scalar, in this notation.
A **contravariant tensor of rank** r is an object whose components transform like the coordinates x^α, i.e.

$$T'^{\alpha_1 \cdots \alpha_r} = \Lambda^{\alpha_1}{}_{\beta_1} \cdots \Lambda^{\alpha_r}{}_{\beta_r} T^{\beta_1 \cdots \beta_r}. \tag{2.46}$$

Tensors of rank 0 are called scalars, such of rank 1 are vectors. In an analogous way we define mixed tensors, for example

$$T'^{\alpha}{}_{\eta\gamma} = \Lambda^\alpha{}_\delta \overline{\Lambda}^\varepsilon{}_\beta \overline{\Lambda}^\nu{}_\gamma T^\delta{}_{\varepsilon\gamma}. \tag{2.47}$$

The following operations can be used in order to form new tensors:

- A linear combination of tensors with the same upper and lower indices, e.g. $T^\alpha{}_\beta = aR^\alpha{}_\beta + bS^\alpha{}_\beta$
- Direct product of tensors, e.g. $T^{\alpha\beta\gamma} = S^{\alpha\beta} R^\gamma$
- Contraction of tensors, e.g. $T^\alpha = S^{\alpha\beta}{}_\beta$
- Differentiation of a tensorfield, e.g. $T_\alpha{}^{\beta\gamma} = \partial_\alpha S^{\beta\gamma}$
  (Caution: If we leave Minkowski space, the derivative of a tensorfield will not necessarily be a tensorfield any more.)

We also notice that two tensors which are equal in one coordinate system are equal in every coordinate system. Because "zero" is a tensor, as well, this implies that the vanishing of a tensor is a Lorentz invariant property.

Transforming a contravariant component of a tensor in a covariant component is defined as in Eq. (2.37) (raising and lowering indices with the metric). Note that the order of upper and lower indices is important.
Some examples for tensors:

- The metric η which is called Minkowski tensor:

  $$\eta'_{\alpha\beta} = \overline{\Lambda}^\gamma{}_\alpha \overline{\Lambda}^\delta{}_\beta \eta_{\gamma\delta} = \overline{\Lambda}^\gamma{}_\alpha \overline{\Lambda}^\delta{}_\beta \Lambda^\mu{}_\gamma \Lambda^\nu{}_\delta \eta_{\mu\nu} = \eta_{\alpha\beta}. \tag{2.48}$$

  In case of the Minkowski metric we have

  $$\eta^\alpha{}_\beta = \eta^{\alpha\gamma}\eta_{\gamma\beta} = \delta^\alpha{}_\beta = \eta_\beta{}^\alpha. \tag{2.49}$$

- Note that $\Lambda^\alpha{}_\beta$ is *not* a tensor.
- $\frac{dx^\mu}{d\tau}$ is a tensor ($\tau' = \tau$ because it refers explicitly to a particular frame), but $\frac{dx^\mu}{dt}$ is not (because t transforms non-trivially under LT).
- The totally antisymmetric tensor $\varepsilon^{\alpha\beta\gamma\delta}$. Under Lorentz transformations ($\det \Lambda = 1$) it behaves like a "pseudotensor", i.e.

  $$\varepsilon'^{\alpha\beta\gamma\delta} = \Lambda^\alpha{}_{\alpha'} \Lambda^\beta{}_{\beta'} \Lambda^\gamma{}_{\gamma'} \Lambda^\delta{}_{\delta'} \varepsilon^{\alpha'\beta'\gamma'\delta'} = \det \Lambda \varepsilon^{\alpha\beta\gamma\delta} \tag{2.50}$$

- The functions $S(x)$, $V^\alpha(x)$, $T^{\alpha\beta}(x)$, ... are scalar fields, vector fields and tensor fields, respectively:

$$S'(x') = S(x) \tag{2.51}$$
$$V'^\alpha(x') = \Lambda^\alpha{}_\beta V^\beta(x) \tag{2.52}$$
$$T'^{\alpha\beta}(x') = \Lambda^\alpha{}_\delta \Lambda^\beta{}_\gamma T^{\delta\gamma}(x). \tag{2.53}$$

In each case the argument has to be transformed, as well: $x'^\alpha = \Lambda^\alpha{}_\beta x^\beta$.

2.6 Electrodynamics

Maxwell's equations read

$$\mathrm{div}\boldsymbol{E} = 4\pi\rho_e, \qquad \mathrm{rot}\boldsymbol{B} = \frac{4\pi}{c}\boldsymbol{j} + \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} \tag{2.54}$$
$$\mathrm{rot}\boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}, \qquad \mathrm{div}\boldsymbol{B} = 0 \tag{2.55}$$

and the continuity equation is

$$\mathrm{div}\boldsymbol{j} + \frac{\partial \rho_e}{\partial t} = 0 \qquad \Leftrightarrow \qquad \partial_\alpha j^\alpha = 0 \tag{2.56}$$

with $j^\alpha = (c\rho_e, \boldsymbol{j})$. The continuity equation follows from the conservation of charge in isolated systems:

$$\partial_t \underbrace{\int d^3r\, j^0}_{=q} = 0. \tag{2.57}$$

We introduce the (antisymmetric) field strength tensor

$$F^{\alpha\beta} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}. \tag{2.58}$$

In terms of the field strength tensor, Maxwell's equations read

$$\partial_\alpha F^{\alpha\beta} = \frac{4\pi}{c} j^\beta \qquad \text{(inhomogeneous)} \tag{2.59}$$
$$\varepsilon^{\alpha\beta\gamma\delta}\partial_\beta F_{\gamma\delta} = 0 \qquad \text{(homogeneous)}. \tag{2.60}$$

Maxwell's equations are now expressed in a *manifestly* Lorentz-covariant notation.
Eq. (2.60) allows us to represent $F^{\alpha\beta}$ as a "curl" of a 4-vector A^α:

$$F^{\alpha\beta} = \partial^\alpha A^\beta - \partial^\beta A^\alpha \tag{2.61}$$

which automatically solves the homogeneous equation (2.60).
Using $A^\alpha = (\phi, A^i)$, we can reformulate Maxwell's equations. From Eq. (2.61) it follows that the gauge transformation

$$A^\alpha \qquad \longrightarrow \qquad A^\alpha + \partial^\alpha \chi \tag{2.62}$$

leaves $F^{\alpha\beta}$ unchanged ($\chi(x)$ is an arbitrary, differentiable scalar field). We can use this freedom in order to demand a certain gauge condition. The Lorenz gauge condition

$$\partial_\alpha A^\alpha = 0 \tag{2.63}$$

leads to the decoupling of the inhomogeneous Maxwell equations:

$$\Box A^\alpha = \frac{4\pi}{c} j^\alpha. \tag{2.64}$$

The equation of motion for a particle with charge q reads

$$m\frac{du^\alpha}{d\tau} = \frac{q}{c} F^{\alpha\beta} u_\beta. \tag{2.65}$$

One can also define the energy-momentum tensor of the electromagnetic field

$$T^{\alpha\beta}_{\text{e.m.}} = \frac{1}{4\pi}\left(F^\alpha{}_\gamma F^{\gamma\beta} - \frac{1}{4}\eta^{\alpha\beta} F_{\gamma\delta} F^{\gamma\delta}\right). \tag{2.66}$$

The 00-component is just the energy density of the field and the $0i$-components give rise to the Poynting vector:

$$T^{00}_{\text{e.m.}} = u_{\text{e.m.}} = \frac{1}{8\pi}(\boldsymbol{E}^2 + \boldsymbol{B}^2) \tag{2.67}$$

$$T^{0i}_{\text{e.m.}} = \frac{1}{c}\boldsymbol{S} = \frac{1}{4\pi}[\boldsymbol{E} \wedge \boldsymbol{B}]^i. \tag{2.68}$$

Maxwell's equations can also be written in terms of the energy-momentum tensor:

$$\partial_\alpha T^{\alpha\beta}_{\text{e.m.}} = -\frac{1}{c} F^{\beta\gamma} j_\gamma. \tag{2.69}$$

An important property of this tensor is that it is conserved ($\partial_\alpha T^{\alpha\beta}_{\text{e.m.}} = 0$) if there is no external force present. The 0-component is just energy conservation and for $\alpha = 1, 2, 3$ this expresses the conservation of momentum. If we have an external force density f^α, then

$$\partial_\alpha T^{\alpha\beta}_{\text{e.m.}} = f^\beta. \tag{2.70}$$

2.7 Accelerated Reference Systems in Special Relativity

Non-inertial systems can be considered in special relativity in a consistent way. But then the physical laws do no longer have the simple covariant form.
For example, consider a rotating coordinate system. We expect additional terms in the equation of motion (e.g. Coriolis force, centrifugal force). Let the coordinate system KS' (with coordinates x'^α) rotate with constant velocity relative to the inertial system IS (with coordinates x^α):

$$\begin{aligned} &t = t', \qquad z = z' \\ &x = x'\cos(\omega t') - y'\sin(\omega t') \\ &y = x'\sin(\omega t') + y'\cos(\omega t') \end{aligned} \tag{2.71}$$

with $\omega^2(x'^2+y'^2) \ll c^2$. Insert Eq. (2.71) into the line element ds:

$$\begin{aligned}
ds^2 &= \eta_{\alpha\beta}dx^\alpha dx^\beta \\
&= c^2dt^2 - dx^2 - dy^2 - dz^2 \\
&= \left[c^2 - \omega^2(x'^2+y'^2)\right]dt'^2 + 2\omega y'dx'dt' - 2\omega x'dy'dt' - dx'^2 - dy'^2 - dz'^2 \\
&= g_{\alpha\beta}dx'^\alpha dx'^\beta \\
&= ds'^2.
\end{aligned} \tag{2.72}$$

For arbitrary coordinates x'^α, ds^2 is a quadratic form in the coordinate differentials dx'^α. To this end, consider a general coordinate transformation

$$x^\alpha = x^\alpha(x') = x^\alpha(x'^0, x'^1, x'^2, x'^3). \tag{2.73}$$

We find

$$\begin{aligned}
ds^2 &= \eta_{\alpha\beta}dx^\alpha dx^\beta \\
&= \underbrace{\eta_{\alpha\beta}\frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{dx'^\nu}}_{=g_{\mu\nu}(x')}dx'^\mu dx'^\nu
\end{aligned} \tag{2.74}$$

where $g_{\mu\nu}$ is defined as the metric tensor of the KS' system. It is symmetric in $\mu \leftrightarrow \nu$ and it is not constant in spacetime.
As in classical mechanics where accelerated reference systems feel inertial forces, we find a centrifugal force $\boldsymbol{Z}$ in a rotating frame. In terms of the centrifugal potential ϕ we have

$$\phi = -\frac{\omega^2}{2}(x'^2+y'^2) \qquad \text{and} \qquad \boldsymbol{Z} = -m\boldsymbol{\nabla}\phi. \tag{2.75}$$

We see from Eqs. (2.72) that

$$g_{00} = 1 + \frac{2\phi}{c^2}. \tag{2.76}$$

Thus the centrifugal potential appears in the metric tensor. First derivatives of the metric tensor are related to the forces in the relativistic equations of motion.
In order to interpret the meaning of t' in KS' we evaluate Eq. (2.72) at a point with $dx' = dy' = dz' = 0$:

$$d\tau \equiv \frac{ds_{\text{clock}}}{c} = \sqrt{g_{00}}dt' = \sqrt{1+\frac{2\phi}{c^2}}dt' \stackrel{!}{=} \sqrt{1-\frac{v^2}{c^2}}dt \tag{2.77}$$

because the proper time τ measured by a clock resting in KS' is a Lorentz scalar. In case of an inertial system we have $g_{\mu\nu} = \eta_{\mu\nu}$ and the clock moves with speed $v = \omega\rho = \omega\sqrt{x'^2+y'^2}$. With Eq. (2.75) we see that both expressions in (2.77) are indeed equal.

As we have seen, the coefficients of the metric tensor $g_{\mu\nu}(x')$ are functions of the coordinates if the coordinate system is not inertial. Such a dependence will also arise if we use curved coordinates instead of flat $\eta_{\alpha\beta}$. In order to describe our example of a rotating reference system, we could also use cylindrical coordinates

$$x'^0 = ct = x^0, \qquad x'^1 = \rho, \qquad x'^2 = \varphi, \qquad x'^3 = z \tag{2.78}$$

with the line element

$$\begin{aligned} ds^2 &= c^2 dt^2 - d\rho^2 - \rho^2 d\varphi^2 - dz^2 \\ &= g_{\mu\nu}(x') dx'^\mu dx'^\nu . \end{aligned} \tag{2.79}$$

Hence $g_{\mu\nu} = \mathrm{diag}(1, -1, -\rho^2, -1)$ is diagonal. These coordinates describe flat Minkowski space from the point of view of an inertial system. Nevertheless the metric tensor depends on the coordinates because which is now due to the fact that we don't use Cartesian coordinates. The fact that the metric tensor depends on the coordinates can either be interpreted as an effect of the acceleration of the coordinate system, or it can be interpreted as a consequence of the fact that we don't use cartesian coordinates.

Part II
Basic Ideas about Gravity and Curvature

3 The Equivalence Principle

The equivalence principle of gravitation and inertia tells us how an arbitrary physical system responds to an external gravitational field.
The physical basis of general relativity is the equivalence principle as formulated by Einstein. At first we want to discuss different forms of the equivalence principle:

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

The first point (**"weak equivalence principle"**) is concerned with the inertial mass m_i which appears in Newton's law $F = m_i a$, and the gravitational mass m_g which the gravitational force is proportional to. A particle in a homogeneous gravitational field satisfies the classical equation

$$m_i \ddot{z} = -m_g g \tag{3.1}$$

with the solution

$$z(t) = -\frac{1}{2}\frac{m_g}{m_i} g t^2 \quad (+v_0 t + z_0). \tag{3.2}$$

Galilei's statement ("All bodies fall at the same rate in a gravitational field.") means that $\frac{m_g}{m_i}$ is the same for all bodies.
A first experimental confirmation is due to Newton who considered the period T of a pendulum (small amplitude):

$$\left(\frac{T}{2\pi}\right)^2 = \frac{m_i}{m_g}\frac{l}{g} \qquad l: \text{ length of pendulum.} \tag{3.3}$$

Using this method, Newton verified $m_i = m_g$ to a precision of $\sim 10^{-3}$. In 1890 Eötvös could verify this statement to a precision of $\sim 10^{-9}$ using torsion balance. The most recent experiments confirm this equality to a precision of $\sim 10^{-12}$.
Due to the equivalence of energy and mass, all forms of energy contribute to a body's inertial as well as to its gravitational mass.

The second postulate reflects the fact that as long as gravitational and inertial mass are the same, also gravitational forces are equivalent to inertial

[2] This postulate is also called the "universality of free fall".

forces. This means that going to an accelerated reference frame, one can get rid of the gravitational field. For example, in the (approximately) homogeneous gravitational field at the surface of the earth we have

$$m_i \ddot{\boldsymbol{r}} = m_g \boldsymbol{g}. \tag{3.4}$$

This is valid for a reference system resting at the earth's surface which is an inertial system for us (to a good approximation). We can perform a transformation to an accelerated coordinate system KS via

$$\boldsymbol{r} = \boldsymbol{r}' + \frac{1}{2}\boldsymbol{g}t^2, \qquad t = t' \tag{3.5}$$

where we assume $gt \ll c$. The origin of KS, $\boldsymbol{r}' = 0$, moves in IS with $\boldsymbol{r}(t) = \frac{1}{2}\boldsymbol{g}t^2$. Inserting (3.5) into (3.4), it follows that

$$m_i \frac{d^2}{dt^2}(\boldsymbol{r}' + \frac{1}{2}gt'^2) = m_g \boldsymbol{g} \tag{3.6}$$

$$\Rightarrow m_i \frac{d^2\boldsymbol{r}'}{dt^2} = (m_g - m_i)\boldsymbol{g} \tag{3.7}$$

which reduces to the equation of motion of a freely moving particle in KS if $m_i = m_g$. Hence the gravitational force vanishes in the freely falling system KS.
Note that something similar is not possible in electromagnetism because charge is certainly not the same as inertial mass.

The weak equivalence principle can also be stated in the following form: an observer who is sealed in some box is not able to tell only by observing freely falling test-particles if his box is accelerating in free space or if there is a gravitational field present. In both cases a freely falling particle behaves exactly the same. The only restriction we have to impose is that he should only be allowed to perform experiments in (infinitesimally) small regions of spacetime. If he were, for example, to compare two freely falling particles which are separated by some large distance, then the particles could behave differently in case of a gravitational field being present. If the field was inhomogeneous (e.g. the gravitational field of the Earth), then the particles would move closer together as they fall down. This effect would not be observed if, instead of the gravitational field of the Earth, the observer and his test-particles would be uniformly accelerated.

Einstein generalizes this postulate. He assumes (**Einstein equivalence principle**) that *in a freely falling reference system,* all *physical processes run exactly as if there was no gravitational field.* This extends the previously purely *mechanical* statement ("weak" equivalence principle) to all types of physical processes and also non-homogeneous gravitational fields. In other words: the observer in the sealed box cannot distinguish a gravitational field from pure acceleration by means of *any* physical process.

As an example of a freely falling system consider a satellite in a stable orbit around the earth and assume that the laboratory on the satellite is not rotating. Then the equivalence principle tells us that in such a system all physical processes run as if there was no gravitational field. So the processes in this laboratory run as in an inertial system. But it is only a *local* IS because the satellite is on a circular orbit and thus it is not resting with respect to the stars. This is the content of the third formulation: in a *local IS* (which is not the same as a general IS) the processes run exactly as in an IS, i.e. the laws of special relativity hold true. The observer in the satellite's laboratory will observe that all physical processes follow the laws of special relativity and that there are neither gravitational nor inertial forces. On the other hand, the observer at earth will say that the satellite moves in a gravitational field and that inertial forces are present since it is accelerated. For him, the satellite follows the path of a freely falling object for which the gravitational and inertial forces just compensate each other.

Note that the compensation of the forces is *exactly* valid only for the center of mass of the satellite system. Thus the equivalence principle applies only to *small* (infinitesimal) local systems.[3]

The equivalence principle can also be formulated in the following way:

> 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 of the point in question, the laws of nature take the same form as in unaccelerated cartesian coordinate systems in the absence of gravitation.[4]

For example, a laboratory bound to the earth's surface can equivalently be seen as in free space without gravity but accelerating "upwards".

The equivalence principle allows to set up the relativistic laws including gravitation. Indeed one can just perform a coordinate transformation to another system: a physical law expressed in terms of special relativity without gravitation can be translated to a relativistic law with gravitation by a coordinate transformation which includes the relative acceleration between the systems (for example, the satellite system and another coordinate system). This relative acceleration corresponds to the gravitational field. Therefore we can derive the relativistic laws in a gravitational field from the equivalence principle. However this does not fix the field equations for $g_{\mu\nu}(x)$ because these equations do not have any analogue in special relativity.

From a geometrical point of view, the coordinate-dependence of the metric tensor means that space is curved. In this sense the field equations describe the

[3] We note that sometimes the following argument is used to show that the equivalence principle is actually wrong: one can in principle construct arbitrarily small apparatuses which are able to detect the inhomogeneity in an inhomogeneous gravitational field which cannot be due to pure acceleration. Therefore, it is argued, free acceleration and gravity can be distinguished in this case even in arbitrarily small regions. The misconception is that we should not talk about arbitrarily small regions in space but in space-*time*. And because such "accelerometers" are certainly not able to measure the inhomogeneity during an arbitrarily short time, the argument fails.

[4] Notice the analogy with the axiom of Gauss taken 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.

connection between the curvature of space and the sources of the gravitational field in a quantitative way. The deep (philosophical) meaning of the equivalence principle is not that gravitational mass as a coupling parameter has accidentally the same numerical value as inertial mass. Einstein rather stated that inertial and gravitational mass are *intrinsically* the same, meaning that gravity is not some external force acting on bodies within spacetime but that rather gravity is a property *inherent* to spacetime. This makes the geometrical interpretation of gravity possible: gravity *is* the manifestation of the intrinsic geometrical shape of spacetime.

This also means that gravity is actually no longer a "force". What has been the uniformly moving inertial observer in Newtonian theory is now the freely falling body. So if we wanted to measure the "gravitational force" by some experiment, we would run into serious problems because we have no "force-free" reference frame! Every body is falling in exactly the same way under the influence of gravity because gravity is no external force depending on some characteristics of the body but rather the shape of spacetime which determines the geodesic path for any free motion. If we want to measure a magnetic field, we can, for instance, compare the motion of a charged particle in the field to ourselves sitting in an inertial laboratory system not being influenced by the magnetic field. According to the equivalence principle something similar is impossible with gravity because the "inertial" laboratory system is that which is freely falling in the fabric of spacetime, so it is moving in exactly the same way as the particle that has to be observed. There is no "background" observer who can be considered as the distinguished inertial frame observing the particle's motion as due to the gravitational "force".

Consequently, we redefine what we mean by "acceleration". In the context of GR we define an unaccelerated observer as one who is freely falling in whatever gravitational field is present. Acceleration is never due to gravity but only due to other forces.

3.1 Riemann Space

Let ξ^α be the Minkowski coordinates in the local IS (e.g. satellite in orbit around earth). The equivalence principle tells us that special relativity should apply. In particular we have

$$ds^2 = \eta_{\alpha\beta} d\xi^\alpha d\xi^\beta. \tag{3.8}$$

The transition from the local IS to another coordinate system KS with coordinates x^μ be given as

$$\xi^\alpha = \xi^\alpha(x^0, x^1, x^2, x^3) \tag{3.9}$$

which yields

$$ds^2 = \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\nu} dx^\mu dx^\nu = g_{\mu\nu}(x) dx^\mu dx^\nu \tag{3.10}$$

$$\Rightarrow g_{\mu\nu}(x) = \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\nu}. \tag{3.11}$$

A space with a path element of the form (3.10) is called a **Riemann Space**. The coordinate transformation (described by $g_{\mu\nu}$) describes also the relative acceleration between KS and the local IS. Since the accelerations are (in general) different at two different points of the local IS, there is no global transformation that brings (3.10) to a Minkowski form. We shall see that $g_{\mu\nu}$ reflects the relativistic gravitational potentials.

4 Physics in a Gravitational Field

Without having really developed the mathematical language of general relativity yet, we can nevertheless draw very interesting conclusions solely from the equivalence principle. At first we want to derive the equation of motion for particles in the presence of a gravitational field (i.e. in a curved space with non-trivial metric). From the equation of motion we can immediately infer the bending of light as it travels past large accumulations of mass. As a second application we will see that a strong gravitational field causes a dilation of time similar to that for lage velocities in special relativity. As a consequence light emitted from a massive star is redshifted when analyzed far away from the star.

4.1 Equation of Motion

The equivalence principle states that in a local IS the laws of special relativity hold. Consider a particle on which no forces but gravity act. Due to the equivalence principle there is a reference frame with coordinates ξ^α in which our freely falling particle moves straight, i.e. the equations of motion assume the form

$$\frac{d^2\xi^\alpha}{d\tau^2} = 0 \qquad\qquad (u^\alpha = \frac{d\xi^\alpha}{d\tau}) \tag{4.1}$$

where τ is the proper time defined by $ds^2 = \eta_{\alpha\beta} d\xi^\alpha d\xi^\beta$. The solution of this equation of motion is just a straight line:

$$\xi^\alpha = a^\alpha \tau + b^\alpha. \tag{4.2}$$

Note that for photons we need to write the equation of motion in the form

$$\frac{d^2\xi^\alpha}{d\lambda^2} = 0 \qquad \text{(photons)} \tag{4.3}$$

where λ is some parameter parametrizing the trajectory (proper time does not make sense because for photons $ds = cd\tau = 0$).

Now consider a global coordinate system KS with coordinates x^μ and metric $g_{\mu\nu}$. At all points x^μ one can locally bring ds^2 to the form $ds^2 = \eta_{\alpha\beta} d\xi^\alpha d\xi^\beta$. Thus at every point P there exists a transformation $\xi^\alpha(x) = \xi^\alpha(x^0, x^1, x^2, x^3)$ from ξ^α to x^μ. Note that this transformation varies from point to point. Consider a small neighbourhood of P and apply the coordinate transformation:

$$ds^2 = \eta_{\alpha\beta} d\xi^\alpha d\xi^\beta = \underbrace{\eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\nu}}_{\equiv g_{\mu\nu(x)}} dx^\mu dx^\nu. \tag{4.4}$$

We write (4.1) in the form

$$0 = \frac{d}{d\tau}\left(\frac{\partial \xi^\alpha}{\partial x^\mu}\frac{dx^\mu}{d\tau}\right)$$
$$= \frac{\partial \xi^\alpha}{\partial x^\mu}\frac{d^2 x^\mu}{d\tau^2} + \frac{\partial^2 \xi^\alpha}{\partial x^\mu x^\nu}\frac{dx^\nu}{d\tau}\frac{dx^\mu}{d\tau}. \tag{4.5}$$

If we multiply this with $\frac{\partial x^\kappa}{\partial \xi^\alpha}$ and use $\frac{\partial \xi^\alpha}{\partial x^\mu}\frac{\partial x^\kappa}{\partial \xi^\alpha} = \delta^\kappa_\mu$, we find the following equation of motion:

$$\boxed{\frac{d^2 x^\kappa}{d\tau^2} = -\Gamma^\kappa_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}} \tag{4.6}$$

$$\text{with the Christoffel symbols } \Gamma^\kappa_{\mu\nu} = \frac{\partial x^\kappa}{\partial \xi^\alpha}\frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu}. \tag{4.7}$$

We see that the world line of a freely falling particle as described by x^μ is a *geodesic* in the Riemannian manifold with metric $g_{\mu\nu}$ (these expressions will be explained later). In special relativity the metric is constant and thus the Christoffel symbols vanish and the equations of motion are those of a particle on which no forces are acting. The Christoffel symbols represent a pseudo-force or fictitious gravitational force which is of the kind of a Coriolis force. It arises whenever inertial motion is described from the point of view of a non-inertial system. Eq. (4.6) is a second order differential equation for the functions $x^\mu(\tau)$ which describe the trajectory of a particle in KS with the metric $g_{\mu\nu}(x)$. We can rewrite this equation as

$$m\frac{du^\alpha}{d\tau} = f^\alpha \qquad \text{with } u^\alpha = \frac{dx^\alpha}{d\tau}. \tag{4.8}$$

We thus interpret the right hand side of (4.6) as describing gravitational forces. Due to Eq. (4.4), the 4-velocity $\frac{dx^\mu}{d\tau}$ has to satisfy the condition

$$c^2 = g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} \qquad (\text{if } m \neq 0). \tag{4.9}$$

Because of this condition, only three of the four components of $\frac{dx^\mu}{d\tau}$ are independent (the same holds true in the case of 4-velocities in special relativity).
Using Eq. (**??**), we find a completely analogous geodesic equations for photons ($m = 0$):

$$\frac{d^2 x^\kappa}{d\lambda^2} = -\Gamma^\kappa_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}. \tag{4.10}$$

Since $d\tau = ds = 0$ we find instead of (4.9)

$$0 = g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} \qquad (\text{if} m = 0). \tag{4.11}$$

We come to the conclusion that only by means of the equivalence principle, light rays can seen to be bent by gravity.

Note that neither the right hand side nor the left hand side of Eq. (4.6) is a tensor. But one can easily show that the whole equation is still true in a tensorial sense because the additional terms which appear under coordinate transformations cancel.

4.1.1 Christoffel Symbols

We want to show that the Christoffel symbols can be expressed in terms of $g_{\mu\nu}$ and its first derivatives only. Considering Eq. (4.4) we note that

$$
\begin{aligned}
\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\nu} &= \eta_{\alpha\beta}\left(\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\lambda}\frac{\partial\xi^\beta}{\partial x^\nu} + \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial^2\xi^\beta}{\partial x^\nu\partial x^\lambda}\right) \\
&\quad + \eta_{\alpha\beta}\left(\frac{\partial^2\xi^\alpha}{\partial x^\lambda\partial x^\mu}\frac{\partial\xi^\beta}{\partial x^\nu} + \frac{\partial\xi^\alpha}{\partial x^\lambda}\frac{\partial^2\xi^\beta}{\partial x^\nu\partial x^\mu}\right) \\
&\quad - \eta_{\alpha\beta}\left(\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\nu}\frac{\partial\xi^\beta}{\partial x^\lambda} + \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial^2\xi^\beta}{\partial x^\lambda\partial x^\nu}\right) \\
&= 2\eta_{\alpha\beta}\frac{\partial^2\xi^\alpha}{\partial x^\mu\partial x^\lambda}\frac{\partial\xi^\beta}{\partial x^\nu}.
\end{aligned}
\tag{4.12}
$$

On the other hand we calculate

$$
\begin{aligned}
g_{\nu\sigma}\Gamma^\sigma_{\mu\lambda} &= \eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\nu}\underbrace{\frac{\partial\xi^\beta}{\partial x^\sigma}\cdot\frac{\partial x^\sigma}{\partial\xi^\kappa}}_{=\delta^\beta_\kappa}\frac{\partial^2\xi^\kappa}{\partial x^\mu\partial x^\lambda} \\
&= \eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\nu}\frac{\partial^2\xi^\beta}{\partial x^\mu\partial x^\lambda}.
\end{aligned}
\tag{4.13}
$$

Comparison of (4.12) and (4.13) yields the result

$$
\boxed{\Gamma^\kappa_{\mu\lambda} = \frac{1}{2}g^{\kappa\nu}\left(\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\nu}\right)}
\tag{4.14}
$$

with the inverse metric $g^{\mu\nu}$,

$$
g^{\kappa\nu}g_{\nu\sigma} = \delta^\kappa_\sigma.
\tag{4.15}
$$

We immediately see that the Christoffel symbols are symmetric in the lower indices:

$$
\Gamma^\kappa_{\mu\lambda} = \Gamma^\kappa_{\lambda\mu}.
\tag{4.16}
$$

Comparing this with the equations of motion of a particle in an electromagnetic field we note a certain analogy. The $\Gamma^\lambda_{\mu\nu}$ correspond to the fields $F^{\alpha\beta}$, whereas $g_{\mu\nu}$ corresponds to A^α.

4.1.2 Newtonian Limit

Assume small velocities ($v^i \ll c$) as well as weak and static (time-independent) fields. Thus

$$
\frac{dx^i}{d\tau} \ll \frac{dx^0}{d\tau}.
\tag{4.17}
$$

We expect to recover the Newtonian theory in this limit. Inserting (4.17) into Eq. (4.6), we find

$$\frac{d^2x^\kappa}{d\tau^2} = -\Gamma^\kappa_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} \simeq -\Gamma^\kappa_{00}\left(\frac{dx^0}{d\tau}\right)^2. \tag{4.18}$$

In the case of static fields ($g_{\mu\nu,0} \simeq 0$), Eq. (4.14) yields

$$\Gamma^\kappa_{00} = -\frac{1}{2}g^{\kappa i}\frac{\partial g_{00}}{\partial x^i}. \tag{4.19}$$

For weak fields, set

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \qquad \text{with } |h_{\mu\nu}| \ll 1. \tag{4.20}$$

This means that the coordinates (ct, x^i) are "almost" Minkowski coordinates. Inserting this into (4.19) we find in a linear order in h that

$$\Gamma^\kappa_{00} = \left(0, \frac{1}{2}\frac{\partial h_{00}}{\partial x^i}\right). \tag{4.21}$$

We evaluate (4.18) for $\kappa = 0$, $\kappa = j$:

$$\frac{d^2t}{d\tau^2} = 0 \qquad \left(\Rightarrow \frac{dt}{d\tau} = \text{const.}\right) \tag{4.22}$$

$$\frac{d^2x^j}{d\tau^2} = -\frac{c^2}{2}h_{00,j}\left(\frac{dt}{d\tau}\right)^2. \tag{4.23}$$

If we choose coordinates in which $\frac{dt}{d\tau} = 1$, the second equation becomes

$$\frac{d^2x^j}{d\tau^2} = -\frac{c^2}{2}h_{00,j}. \tag{4.24}$$

Hence

$$\frac{d^2\boldsymbol{r}}{dt^2} = -\frac{c^2}{2}\boldsymbol{\nabla}h_{00}(\boldsymbol{r}) \tag{4.25}$$

with $(x^j) = \boldsymbol{r}$. The corresponding Newtonian equation is

$$\frac{d^2\boldsymbol{r}}{dt^2} = -\boldsymbol{\nabla}\phi(\boldsymbol{r}). \tag{4.26}$$

We conclude that

$$g_{00}(\boldsymbol{r}) = 1 + h_{00}(\boldsymbol{r}) = 1 + \frac{2\phi(\boldsymbol{r})}{c^2}. \tag{4.27}$$

But note that the Newtonian limit (4.26) contains no information about the other components of $h_{\mu\nu}$. The term $\frac{2\phi(\boldsymbol{r})}{c^2}$ serves as a measure for the strength of the gravitational fields.

For example, consider a spherically symmetric mass distribution. At the surface of the earth we have a general relativistic correction of the order

$$\frac{2\phi(\boldsymbol{r})}{c^2} \simeq 1.4 \cdot 10^{-9}. \tag{4.28}$$

Compare this to the data of other astrophysical objects:

$$\frac{2\phi(\boldsymbol{r})}{c^2} \simeq \begin{cases} 4 \cdot 10^{-6} & \text{(sun)} \\ 3 \cdot 10^{-4} & \text{(white dwarf)} \\ 3 \cdot 10^{-1} & \text{(neutron star)} \end{cases} \tag{4.29}$$

5 Time Dilation

Consider a clock in a static gravitational field and the effect of gravitational redshift.

5.0.3 Proper Time

The proper time τ of the clock with coordinates $x = (x^\mu)$ in an arbitrary reference system is

$$d\tau = \frac{ds_{\text{clock}}}{c} = \frac{1}{c}\left(\sqrt{g_{\mu\nu}(x)dx^\mu dx^\nu}\right)_{\text{clock}}. \tag{5.1}$$

(According to the equivalence principle, if the coordinate system which we use to describe the motion is locally inertial, the laws of special relativity hold, i.e. Eq. (5.1) holds with $\eta_{\mu\nu}$ instead of $g_{\mu\nu}$).
The time shown by the clock depends on both the gravitational field and on the motion of the clock. We consider the following special cases:

- A clock moving in an IS without gravity:

$$d\tau = \sqrt{1 - \frac{v^2}{c^2}}dt. \qquad (g_{\mu\nu} = \eta_{\mu\nu}, \quad dx = (cdt, \boldsymbol{v}dt)) \tag{5.2}$$

- Clock at rest in a gravitational field has spatial coordinates $dx^i = 0$. Hence we have

$$d\tau = \sqrt{g_{00}}dt. \tag{5.3}$$

  If the field is static, we have, using Eq. (4.27),

$$d\tau = \sqrt{1 + \frac{2\phi(r)}{c^2}}dt \qquad (|\phi| \ll c^2). \tag{5.4}$$

  Because of ϕ being negative, a clock in a gravitational field goes slower than a free clock without gravitational field.

5.0.4 Redshift

We consider atoms (e.g. on stars) emitting and absorbing light with a given frequency as "clocks". Furthermore we consider only a static gravitational field, i.e. $g_{\mu\nu}$ does not depend on time. Let the source be resting at $\boldsymbol{r}_A$ and let it emit a monochromatic electromagnetic wave with frequency ν_A. An observer at $\boldsymbol{r}_B$ will measure a frequency ν_B. In general, A and B will measure different proper times:

$$\text{Proper time of A:} \qquad d\tau_A = \sqrt{g_{00}(\boldsymbol{r}_A)}dt_A, \tag{5.5}$$

$$\text{Proper time of B:} \qquad d\tau_B = \sqrt{g_{00}(\boldsymbol{r}_B)}dt_B. \tag{5.6}$$

As a time interval we consider the time between two subsequent peaks which are emitted by A or received by B, respectively. In this case $d\tau_A$ and $d\tau_B$ correspond to the period of the waves in A and B, i.e.

$$d\tau_A = \frac{1}{\nu_A}, \qquad d\tau_B = \frac{1}{\nu_B}. \tag{5.7}$$

Going from A to B needs the same time Δt for the peaks of both signals. Therefore they will arrive with a time delay which is equal to the one with which they were emitted:

$$dt_A = dt_B. \tag{5.8}$$

Using Eqs. (5.6) and (5.7) we find

$$\frac{\nu_A}{\nu_B} = \sqrt{\frac{g_{00}(\boldsymbol{r}_B)}{g_{00}(\boldsymbol{r}_A)}}. \tag{5.9}$$

We introduce the **redshift**

$$z = \frac{\nu_A}{\nu_B} - 1 = \frac{\lambda_B}{\lambda_A} - 1. \tag{5.10}$$

From Eq. (5.9) we conclude that the gravitational redshift is

$$z = \sqrt{\frac{g_{00}(\boldsymbol{r}_B)}{g_{00}(\boldsymbol{r}_A)}} - 1. \tag{5.11}$$

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

$$z = \frac{\phi(\boldsymbol{r}_B) - \phi(\boldsymbol{r}_A)}{c^2} \qquad (|\phi| \ll c^2). \tag{5.12}$$

To summarize, we expect to measure a smaller frequency $\nu_B < \nu_A$ if the light has to "escape" a gravitational field to some extend when propagating from A to B.

For example, consider the light of the sun. With (5.12) we find

$$z \simeq \frac{-\phi(r_A)}{c^2} = \frac{GM_\odot}{c^2 R_\odot} \simeq 2 \cdot 10^{-6}. \tag{5.13}$$

Compare this to $z \simeq 10^{-4}$ on white dwarfs or even $z \simeq 10^{-1}$ in the case of neutron stars.

In general there are three effects which can lead to a change of 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.
3. **Cosmological redshift:** Due to the expansion of the universe. In this case the metric tensor depends on time.

5.0.5 Photons in a Gravitational Field

Consider a photon with energy $E_\gamma = \hbar\omega = 2\pi\hbar\nu$ propagating in the homogeneous gravitational field of the Earth. We assume that it moves a distance $h = h_B - h_A$ "upwards" where h is so small that the field can assumed to be constant. The redshift is

$$z = \frac{\nu_A}{\nu_B} - 1 \approx \frac{\phi(r_B) - \phi(r_A)}{c^2} = \frac{g(h_B - h_A)}{c^2} = \frac{gh}{c^2}. \tag{5.14}$$

This leads to a change in the photon's frequency which amounts to

$$\Delta\nu = \nu_B - \nu_A \qquad \Rightarrow \qquad \frac{\Delta\nu}{\nu} \approx -\frac{gh}{c^2} \tag{5.15}$$

with $\nu_A > \nu_B \equiv \nu$. Thus the photon's energy decreases by $\Delta E_\gamma = -\frac{E_\gamma}{c^2} gh$ which is exactly the change in potential energy of a particle with mass $\frac{E_\gamma}{c^2}$ when lifted by the same distance h.
Note that from a Newtonian point of view, this thought experiment leads to contradictions because the photon would gain potential energy without losing any other kind of energy.

This effect has been measured first in 1965 through the Mössbauer effect. It has been observed that

$$\left(\frac{\Delta\nu_{\text{exp.}}}{\Delta\nu_{\text{theoretical}}}\right) \simeq 1.00 \pm 0.01. \tag{5.16}$$

Part III

The Mathematical Language of General Relativity

6 Geometrical Considerations

In general the coordinate dependence of $g_{\mu\nu}(x)$ can be interpreted in the sense that space, defined by the line element ds^2, is curved. The trajectories in a gravitational field are the geodesics in the corresponding space.

6.1 Curvature of Space

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

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\nu \qquad (\mu, \nu = 1, ..., N). \tag{6.1}$$

Let's consider a two dimensional space, for example the Euclidean plane with cartesian coordinates $(x^1, x^2) = (x, y)$ and $ds^2 = dx^2 + dy^2$. It can also be described in terms of polar coordinates $(x^1, x^2) = (\rho, \varphi)$ with $ds^2 = d\rho^2 + \rho^2 d\varphi^2$. The two line elements can be transformed into each other via a coordinate transformation. Another example is the surface of a sphere with angular coordinates $(x^1, x^2) = (\theta, \varphi)$ and $ds^2 = a^2(d\theta^2 + \sin^2\theta d\varphi^2)$, a being the radius of the sphere. Note that there is no coordinate transformation which brings the line element of the sphere to the form of the line element of a plane. Thus the metric tensor determines the properties of the space (e.g. the curvature).
However, the form of the metric tensor is not uniquely determined by the space, because it depends on the choice of coordinates.
Note that besides the curvature discussed here, there is also an *exterior* curvature. We are only interested in intrinsic curvature.
If g_{ik} =const., then the space is not curved. In a Euclidean space one can then introduce cartesian coordinates such that $g'_{ik} = \delta_{ik}$. For curved spaces $g_{ik} \neq$const.. But obviously the fact that $g_{ik} \neq$const. does not imply that the space is curved (see, for example, the above examples).
For instance, by measuring the angles of a triangle, one can infer if the space is curved or not. It is such properties which can be inferred without external reference, which we call **intrinsic**.

7 Differentiable Manifolds

A manifold is a topological space which locally looks like Euclidean $\mathbf{R}^n$. A simple example is the sphere S^2: On can set up local coordinates (θ, ϕ) which map S^2 locally to the plane. The local coordinates are called a **chart** and a collection of charts is an **atlas**. There is no 1-1 map from S^2 onto $\mathbf{R}^2$ - we need several charts to cover S^2. In general we have the

Definition. Given a topological space M, a **chart** on M is a 1-1 map ϕ from an open subset $U \subset M$ to an open subset $\phi(U) \subset \mathbf{R}^n$. A chart is often called a coordinate system. A set of charts with domain U_α is called an atlas of M s.t. $\bigcup_\alpha U_\alpha = M$ with a family of maps $\{\phi_\alpha | \alpha \in I\}$. The dimension of the manifold is $n :=\dim(M)$.
We will always assume that the maps ϕ are C^∞-diffeomorphisms.

Definition. If (ϕ_1, U_1) and (ϕ_2, U_2) are two charts, then ϕ_1 and ϕ_2 are **C^∞-related** if both the **transition function**

$$\phi_2 \circ \phi_1^{-1} : \qquad \phi_1(U_1 \cap U_2) \subset \mathbf{R}^n \longrightarrow \phi_2(U_1 \cap U_2) \subset \mathbf{R}^n \tag{7.1}$$

and its inverse are C^∞, i.e. it is a diffeomorphism. A collection of C^∞-related charts such that every point of M lies in the domain of at least one chart form an atlas. The collection of all such C^∞-related charts forms a **maximal atlas**. If M is a topological space and A its maximal atlas, then the pair (M, A) is called a **C^∞-differentiable manifold**.
(If for each ϕ in the atlas the map $\phi : U \longrightarrow \mathbf{R}^n$ has the same n, then the manifold has dimension n.)
Some notions:

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

7.1 Tangent Vectors and Tangent Spaces

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

Definition. A C^∞-curve in a manifold M is a map $\gamma : I = (a, b) \longrightarrow M$ such that for any chart ϕ, $\phi \circ \gamma : I \longrightarrow \mathbf{R}^n$ is a C^∞-map.
Let $f : M \longrightarrow \mathbf{R}$ be a smooth function on M. Consider the map $f \circ \gamma : I \longrightarrow \mathbf{R}$, $t \mapsto f(\gamma(t))$. This map has a well-defined derivative which is the rate of change of f along the curve γ. Consider

$$\underbrace{f \circ \phi^{-1}}_{\mathbf{R}^n \to \mathbf{R}} \circ \underbrace{\phi \circ \gamma}_{I \to \mathbf{R}^n} \tag{7.2}$$

which allows for a derivative (chain rule!):

$$\frac{d}{dt}(f \circ \gamma) = \sum_{i=1}^{N} \left(\frac{\partial}{\partial x^i} f(\phi^{-1}(x^i)) \right) \frac{dx^i(\gamma(t))}{dt}. \tag{7.3}$$

Thus given a curve $\gamma(t)$ and a function f we can obtain the rate of change of f along $\gamma(t)$ at $t = t_0$:

$$\left[\frac{d}{dt}(f \circ \gamma) \right]_{t=t_0}. \tag{7.4}$$

Definition. The tangent vector $\dot{\gamma}_p = \left(\frac{d\gamma}{dt}\right)_p$ to a curve $\gamma(t)$ at a point p is a map from the set of real functions $f : U \longrightarrow \mathbf{R}$ from a neighbourhood U of p to $\mathbf{R}$, defined by

$$\dot{\gamma}_p : \quad f \mapsto \left[\frac{d}{dt}(f \circ \gamma) \right]_p = (f \circ \gamma)^{\cdot}_p = \dot{\gamma}_p(f). \tag{7.5}$$

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

$$(x^i \circ \gamma)^{\cdot}_p = \left[\frac{d}{dt} x^i(\gamma(t)) \right]_t. \tag{7.6}$$

The set of all tangent vectors at p span the **tangent space**.

Theorem. If the dimension of M is n then T_pM is a vector space of dimension n.
We set $\gamma(0) = p$, $X_p = \dot{\gamma}_p$ and $X_p f = \dot{\gamma}_p(f)$. i.e.

$$\begin{aligned} X_p f &= (f \circ \gamma)^{\cdot}(0) \\ &= \left[f \circ \phi^{-1} \circ \phi \circ \gamma \right]^{\cdot}(0) \\ &= \sum_{i=1}^{N} \frac{\partial}{\partial x^i}(f \circ \phi^{-1})(x^i \circ \gamma)^{\cdot}(0) \\ &= \sum_{i=1}^{N} \left(\frac{\partial}{\partial x^i} f(x) \right) (X_p(x^i)) \\ &= \sum_{i=1}^{N} \left(\frac{\partial}{\partial x^i} f(x^i) \right) (X_p(x^i)) \end{aligned} \tag{7.7}$$

where $X_p(x^i)$ are the components of X_p with respect to the given basis. This way we find

$$X_p = \sum_{i=1}^{N} (X_p(x^i)) \left(\frac{\partial}{\partial x^i} \right)_p \tag{7.8}$$

and so the $\left(\frac{\partial}{\partial x^i}\right)$ span T_pM.

Suppose now that f, g are real functions on M and $fg : M \longrightarrow \mathbf{R}$ is defined by $fg(p) = f(p)g(p)$. If $X_p \in T_pM$, then the Leibniz rule holds:

$$X_p(fg) = (X_pf)g(p) + f(p)(X_pg). \tag{7.9}$$

We introduce the notation

$$(Xf)(p) \equiv X_pf. \tag{7.10}$$

We can now introduce a basis of T_pM. T_pM has dimension n. In any basis $(e_1, ..., e_n)$ we have

$$X = X^i e_i. \tag{7.11}$$

A change of basis is given by

$$\bar{e}_i = \phi_i{}^k e_k, \qquad \bar{x}^i = \phi^i{}_k x^k. \tag{7.12}$$

In particular, $e_i = \frac{\partial}{\partial x^i}$ is called the coordinate basis (with respect to a chart). Upon change of chart $x \to \bar{x}$ we have

$$\phi_i{}^k = \frac{\partial x^k}{\partial \bar{x}^i}, \qquad \phi^i{}_k = \frac{\partial \bar{x}^i}{\partial x^k}. \tag{7.13}$$

7.2 The Cotangent Space T_p^*M (or dual space of T_pM)

The dual space of T_pM is the space of all covectors $\omega \in T_p^*M$ which are linear forms $\omega : X \mapsto \omega(X) \equiv \langle \omega, X \rangle \in \mathbf{R}$.
In particular, we have that the differential $df : X \to Xf$ of functions f on M is an element of T_p^*. The elements $df = f_{,i}\, dx^i = \left(\frac{\partial}{\partial x^i} f\right) dx^i$ span all of T_p^*.

Definition. Consider a basis $(e^1, ..., e^n)$ of T_p^*. It is the called the **dual basis** of a basis $(e_1, ...e_n)$ of T_p if

$$\langle e^i, X \rangle = x^i \tag{7.14}$$

$$\Leftrightarrow \quad \langle e^i, X^j e_j \rangle = X^j \underbrace{\langle e^i, e_j \rangle}_{=\delta^i_j} = X^i. \tag{7.15}$$

For an arbitrary element $\omega = \omega_i e^i$ of T_p^*M this means that $\omega_i = \langle \omega, e_i \rangle$. The components ω_i transform as the contravariant basis e_i and the e^i transform like the X^i. In particular, we have for the coordinate basis $e_i = \frac{\partial}{\partial x^i}$, $e^i = dx^i$ the following usual contravariant and covariant transformation behaviour:

$$\frac{\partial}{\partial \bar{x}^i} = \frac{\partial x^k}{\partial \bar{x}^i} \frac{\partial}{\partial x^k} = \phi_i{}^k \frac{\partial}{\partial x^k} \tag{7.16}$$

$$d\bar{x}^i = \frac{\partial \bar{x}^i}{\partial x^k} dx^k = \phi^i{}_k dx^k. \tag{7.17}$$

7.3 Tensors on T_pM

A tensor on T_p is a multilinear form on T_p^*M and T_pM. For example, a tensor T of type $(1,2)$ is a trilinear form $T : T_p^*M \times T_pM \times T_pM \to \mathbb{R}$. We introduce the notation $T \in \otimes_2^1 T_pM$.
The tensor product is defined between tensors of any type:

$$T = R \otimes S \Rightarrow T(\omega, X, Y) = R(\omega, X) \cdot S(Y). \tag{7.18}$$

Written in components, this reads

$$\underbrace{T(e^i, e_j, e_k)}_{=:T^i{}_{jk}} \quad \underbrace{\omega_i X^j Y^k}_{e_i(\omega)e^j(X)e^k(Y)} \quad , \tag{7.19}$$

hence

$$T = T^i{}_{jk} e_i \otimes e^j \otimes e^k. \tag{7.20}$$

Any tensor of this type can therefore be obtained as a linear combination of tensor products $X \otimes \omega \otimes \omega'$ with $X \in T_pM$, ω, $\omega' \in T_p^*M$.
The generalization to tensors of type (p,q) is obvious.

If we perform a change of bases, our $(1,2)$-tensor transforms as

$$\bar{T}^i{}_{jk} = T^\alpha{}_{\beta\gamma} \phi^i{}_\alpha \phi_j{}^\beta \phi_k{}^\gamma. \tag{7.21}$$

Note that any bilinear form $b \in T_p^*M \otimes T_pM$ determines a linear form $l \in (T_pM \otimes T_p^*M)^*$ via

$$l(X \otimes \omega) = b(X, \omega). \tag{7.22}$$

In particular, the **trace** $\mathrm{tr}T$ is a linear form on tensors T of type $(1,1)$, defined by

$$\mathrm{tr}(X \otimes \omega) = \langle \omega, X \rangle \tag{7.23}$$

$$\Leftrightarrow \qquad \mathrm{tr}T = T^\alpha{}_\alpha \quad \text{(w.r.t. dual pair of bases).} \tag{7.24}$$

Similarly, given a (p,q)-tensor, one can obtain a $(p-1, q-1)$-tensor by means of **contraction**. For example,

$$T^i{}_{jk} \quad \longrightarrow \quad S_k = T^i{}_{ik}. \tag{7.25}$$

To summarize the last two sections, we have the following transformation behaviours for coordinate basis, dual basis, vector components and covector components:[5]

$$\frac{\partial}{\partial \bar{x}^i} = \frac{\partial x^k}{\partial \bar{x}^i}\frac{\partial}{\partial x^k}, \tag{7.27}$$

$$d\bar{x}^i = \frac{\partial \bar{x}^i}{\partial x^k} dx^k,$$

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

$$\bar{\omega}_i = \frac{\partial x^k}{\partial \bar{x}^i}\omega_k. \tag{7.28}$$

7.3.1 The Tangent Map

Definition. Let $\varphi : M \to \bar{M}$ be a differentiable map and let $p \in M$, $\bar{p} = \varphi(p)$. Then φ induces a linear map („**pushforward**")

$$\varphi_* : T_pM \to T_{\bar{p}}\bar{M}. \tag{7.29}$$

We denote by $\mathcal{F}(\bar{M})$ the space of all smooth functions $M \to \mathbf{R}$. The pushforward can by characterized in two (equivalent) ways:

- $\forall \bar{f} \in \mathcal{F}_p(\bar{M}) : \qquad (\varphi_* X)\bar{f} = X(\bar{f} \circ \varphi)$
- Choose a representative γ of X, i.e. $X = \dot{\gamma}_p$. Then $\bar{\gamma} = \varphi \circ \gamma$ is a representative of $\varphi_* X$. This agrees with the first characterization since

$$\frac{d}{dt}\bar{f}(\bar{\gamma}(t))\Big|_{t=0} = \frac{d}{dt}(\bar{f} \circ \varphi)(\gamma(t))\Big|_{t=0}. \tag{7.30}$$

With respect to bases $(e_1, ..., e_n)$ of T_pM, $(\bar{e}_1, ..., \bar{e}_n)$ of $T_{\bar{p}}\bar{M}$ the condition reads

$$\bar{X} = \varphi_* X \qquad \text{with} \quad \bar{X}^i = (\varphi_*)^i{}_k X^k = \langle \bar{e}^i, \varphi_* e_k \rangle X^k \tag{7.31}$$

or in case of coordinate bases:

$$(\varphi_*)^i{}_k = \frac{\partial \bar{x}^i}{\partial x^k}. \tag{7.32}$$

Definition. The **adjoint map** (”**pullback**") φ^* of φ_* is defined as

$$\varphi^* : \quad T^*_{\bar{p}}M \longrightarrow T^*_pM, \tag{7.33}$$

$$\bar{\omega} \mapsto \varphi^* \bar{\omega} \qquad (= \omega \text{ in } T^*_pM) \tag{7.34}$$

$$\text{with} \quad \langle \varphi^* \bar{\omega}, X \rangle = \langle \bar{\omega}, \varphi_* X \rangle. \tag{7.35}$$

[5]Note how this transformation behaviour makes the independence on coordinates explicit: let, for example, $X = X^i \partial_i$ be a vector field. Then

$$\bar{X} = \bar{X}^i \frac{\partial}{\partial \bar{x}^i} = \frac{\partial \bar{x}^i}{\partial x^k} X^k \frac{\partial x^j}{\partial \bar{x}^i}\frac{\partial}{\partial x^j} = \delta^j_k X^k \frac{\partial}{\partial x^j} = X^k \frac{\partial}{\partial x^k} = X. \tag{7.26}$$

The same result is obtained from the definition

$$\varphi^* : \quad d\bar{f} \mapsto d(\bar{f} \circ \varphi) \qquad \text{for } \bar{f} \in \mathcal{F}(\bar{M}). \tag{7.36}$$

In components, $\omega = \varphi^* \bar{\omega}$ reads

$$\omega_k = \bar{\omega}_i (\varphi_*)^i{}_k. \tag{7.37}$$

Consider (local) diffeomorphisms, i.e. differentiable maps φ such that φ^{-1} exists in a neighbourhood of $\bar{p}$ and demand that dim$M$ =dim$\bar{M}$ 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 type. Again, we consider the example of a $(1,1)$-tensor where the generalization is straightforward. Then

$$(\varphi_* T)(\bar{\omega}, \bar{X}) = T\big(\underbrace{\varphi^* \bar{\omega}}_{\omega}, \underbrace{\varphi_*^{-1} \bar{X}}_{X}\big) \tag{7.38}$$

$$(\varphi^* \bar{T})(\omega, X) = \bar{T}\big(\underbrace{(\varphi^*)^{-1} \omega}_{\bar{\omega}}, \underbrace{\varphi_* X}_{\bar{X}}\big). \tag{7.39}$$

Hence φ_* and φ^* are each others inverse and we have

$$\varphi_*(T \otimes S) = (\varphi_* T) \otimes (\varphi_* S) \tag{7.40}$$

$$\mathrm{tr}(\varphi_* T) = \varphi_*(\mathrm{tr} T) \tag{7.41}$$

and similar for φ^*. In components, $\bar{T} = \varphi_* T$ reads in a coordinate basis

$$\bar{T}^i{}_k = T^\alpha{}_\beta \frac{\partial \bar{x}^i}{\partial x^\alpha} \frac{\partial \bar{x}^\beta}{\partial x^k}. \tag{7.42}$$

8 Vector and Tensor Fields

Definition. If to every point p of a differentiable manifold M a tangent vector $X_p \in T_pM$ is assigned, then we call the map $p \mapsto X_p$ a **vector field** on M. Given a coordinate system x^i and an assigned basis $\left(\frac{\partial}{\partial x^i}\right)_p$ for every tangent space T_pM, X has components X_p^i where

$$X_p = X_p^i \left(\frac{\partial}{\partial x^i}\right)_p \qquad \text{and} \qquad X_p^i = X_p(x^i). \tag{8.1}$$

Eq. (7.13) shows how the components X_p^i transform under coordinate transformations. Xf is called the *derivative of f with respect to the vector field X*. The following rules apply:

- $X(f+g) = Xf + Xg$
- $X(f \cdot g) = (Xf)g + f(Xg)$ (Leibniz rule)

The vector fields on M form a linear space on which the following operations are well defined:

- $X \mapsto fX$ (multiplication by $f \in \mathcal{F}$)
- $X, Y \mapsto [X, Y] = XY - YX$ (commutator). Note that $[X, Y]$, in contrast to XY, satisfies the Leibniz rule.
  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:

  $$[X, Y]^j = (XY - YX)x^j. \tag{8.2}$$
  $$\tag{8.3}$$

  To evaluate this, we use that for $X = X^i \frac{\partial}{\partial x^i}$, $Y = Y^k \frac{\partial}{\partial x^k}$, we have $Yx^j = Y^k \frac{\partial x^j}{\partial x^k} = Y^k \delta^j_k = Y^j$ and $XY^j = X^k \frac{\partial}{\partial x^k}(Y^j) = X^k Y^j{}_{,k}$ which implies

  $$XY^j - YX^j = X^k Y^j{}_{,k} - Y^k X^j{}_{,k}\,. \tag{8.4}$$

  Furthermore, in a local coordinate basis, the bracket $[\partial_k, \partial_j]$ vanishes.
  We note that this commutator satisfies the Jacobi identity:

  $$[X, [Y, Z]] + [Y, [X, Z]] + [Z, [X, Y]] = 0. \tag{8.5}$$

Definition. Let $T_pM^r_s$ of tensorrs of rank (r, s) defined on T_pM. A **tensor field** of rank (r, s) is a map which assigns a tensor t_p of rank (r, s) on T_pM to every point $p \in M$.
Algebraic operations on tensor fields are defined pointwise. For instance, the sum of two tensor fields is defined by

$$(t + s)_p = t_p + s_p \qquad \text{for } t, s \in T_pM^r_s. \tag{8.6}$$

Tensor products and contractions of tensors are defined analogously. Tensor fields can be multiplied with functions in a natural way: let $f \in \mathcal{F}(M)$, $t \in T_pM^r_s$, then

$$(ft)_p = f(p)t_p. \tag{8.7}$$

In a coordinate neighbourhood U with coordinates $(x^1, ..., x^n)$ a tensor field can be expanded in the form

$$t = t^{i_1 \cdots i_r}{}_{j_1 \cdots j_s} \left(\frac{\partial}{\partial x^{i_1}} \otimes \cdots \otimes \frac{\partial}{\partial x^{i_r}} \right) \otimes \left(dx^{j_1} \otimes \cdots \otimes dx^{j_s} \right). \tag{8.8}$$

If we perform a coordinate transformation to $(\bar{x}^1, ..., \bar{x}^n)$, the components of t transform according to

$$\bar{t}^{i_1 \cdots i_r}{}_{j_1 \cdots j_s} = t^{k_1 \cdots k_r}{}_{l_1 \cdots l_s} \frac{\partial \bar{x}^{i_1}}{\partial x^{k_1}} \cdots \frac{\partial \bar{x}^{i_r}}{\partial x^{k_r}} \frac{\partial x^{l_1}}{\partial \bar{x}^{j_1}} \cdots \frac{\partial x^{l_s}}{\partial \bar{x}^{j_s}}. \tag{8.9}$$

A **one-form** is a covariant tensor of rank 1. So if ω is a 1-form, this means that at every point $p \in M$, ω_p is a mapping $T_pM \to \mathbb{R}$. Moreover to every basis $\{\boldsymbol{e}_i\}$ of T_pM, there is a dual basis $\{\boldsymbol{e}^j\}$ consisting of 1-forms, such that $\boldsymbol{e}^j(\boldsymbol{e}_i) = \delta^j_i$. The canonical example is, of course, $\boldsymbol{e}_i = \partial_i$, $\boldsymbol{e}^j = dx^j$.
We denote by $\mathcal{T}^r_s(M)$ the set of all tensor fields of rank (r, s).

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

$$(i) \quad g(X,Y) = g(Y,X) \quad \forall X, Y \tag{8.10}$$

$$(ii) \quad \forall p \in M, \quad g_p \text{ is a non-degenerate bilinear form on } T_pM. \tag{8.11}$$

$$\text{i.e. } g_p(X,Y) = 0 \quad \forall X \in T_pM \quad \Leftrightarrow \quad Y = 0. \tag{8.12}$$

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

Definition. A **(pseudo-)Riemannian manifold** is a differentiable manifold M together with a (pseudo-)Riemannian metric.

8.1 Flows and Generating Vector Fields

A **flow** is a one-parameter group of diffeomorphisms, i.e.

$$\varphi_t : M \longrightarrow M \text{ such that } \varphi_t \circ \varphi_s = \varphi_{t+s}. \tag{8.13}$$

In particular φ_0 =id. Moreover, the orbits (or **integral curves**) of any point $p \in M$, given by

$$t \mapsto \varphi_t(p) \equiv \gamma(t), \tag{8.14}$$

shall be differentiable.
A flow determines a vector field X via

$$Xf = \frac{d}{dt}(f \circ \varphi_t)\big|_{t=0}. \tag{8.15}$$

$$\text{i.e.} \quad X_p = \frac{d}{dt}\gamma(t)\big|_{t=0} = \dot{\gamma}(0). \tag{8.16}$$

At the point $\gamma(t)$ we have

$$\dot{\gamma}(t) = \frac{d}{dt}\varphi_t(p) = \frac{d}{ds}(\varphi_s \circ \varphi_t)(p)\big|_{s=0} = X_{\varphi_t(p)}, \tag{8.17}$$

hence $\gamma(t)$ solves the ordinary differential equation

$$\dot{\gamma}(t) = X_{\gamma(t)}, \quad \gamma(0) = p. \tag{8.18}$$

The generating vector field determines the flow uniquely due to the Picard-Lindelöf-theorem.

9 The Lie Derivative

In order to define the derivative of a vector field V, we need to have a method to compare V_p and $V_{p'}$ at nearby points $p, p' \in M$. Since V_p and $V_{p'}$ live in different spaces, their difference can not be taken unless $V_{p'}$ has been transported to T_pM in a consistent way. 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 the vector field X is defined by

$$L_X R = \frac{d}{dt}\varphi_t^* R\big|_{t=0} \tag{9.1}$$

or more explicitly:

$$(L_X R)_p = \frac{d}{dt}\varphi_t^* R_{\varphi_t(p)}\big|_{t=0}. \tag{9.2}$$

Here, φ_t is the (local) flow generated by X, hence $\varphi_t^* R_{\varphi_t(p)}$ is a tensor on T_p depending on t. The Lie derivative provides us with a generalization of the directional derivative: it compares the tensor R at a point with the same tensor at an infinitesimally close point in the direction of X; so it is a measure of how R changes if one moves in the direction of the given vector field X.[6]
In order to express L_X in components we write φ_t in a chart, $\varphi_t : x \mapsto \bar{x}(t)$ and linearize for small t:

$$\bar{x}^i = x^i + tX^i(x) + \mathcal{O}(t^2), \qquad x^i = \bar{x}^i - tX^i(\bar{x}) + \mathcal{O}(t^2). \tag{9.4}$$

Therefore

$$\frac{\partial^2 \bar{x}^i}{\partial x^k \partial t} = -\frac{\partial^2 x^i}{\partial \bar{x}^k \partial t} = X^i{}_{,k} \tag{9.5}$$

at $t = 0$. As an example, let R be of type $(1, 1)$. According to Eq. (7.42) we have

$$(\varphi_t^* R)^i{}_j(x) = R^\alpha{}_\beta(\bar{x}) \frac{\partial x^i}{\partial \bar{x}^\alpha} \frac{\partial \bar{x}^\beta}{\partial x^j}. \tag{9.6}$$

Taking a derivative with respect to t and $t = 0$ yields

$$(L_X R)^i{}_j = R^i{}_{j,k} X^k - R^\alpha{}_j X^i{}_{,\alpha} + R^i{}_\beta X^\beta{}_{,j}\,. \tag{9.7}$$

The first term can be computed using

$$\frac{\partial}{\partial \bar{x}^k} R^\alpha{}_\beta(\bar{x}) \frac{\partial \bar{x}^k}{\partial t} \frac{\partial x^i}{\partial \bar{x}^\alpha} \frac{\partial \bar{x}^\beta}{\partial x^j}\big|_{t=0}. \tag{9.8}$$

The Lie derivative enjoys the following properties:

- L_X maps tensor fields to tensor fields of the same type linearly.

[6] Remember the directional derivative in a flat, Euclidean geometry: denote by $\gamma(\lambda)$ a path in $\mathbb{R}^n$. Then the directional derivative of a scalar function $f(\boldsymbol{x})$ in the direction of γ at the point $p = \gamma(0)$ is

$$\nabla_\gamma f_p = \frac{d}{d\lambda} f(\gamma(\lambda))\Big|_{\lambda=0} = \dot{\gamma}(0) \cdot \boldsymbol{\nabla} f(p). \tag{9.3}$$

The Lie derivative is obviously a generalization of this. We only need to pay attention that (in contrast to the Euclidean case) tangent spaces at different points cannot be compared, so that vectors in different tangent spaces have to all be transported to the same tangent space.

- $L_X(trT) = tr(L_X T)$.
- $L_X(T \otimes S) = (L_X T) \otimes S + T \otimes (L_X S)$.
- $L_X f = Xf$ for $f \in \mathcal{F}(M)$.
- $L_X Y = [X, Y]$ for any vectorfield Y.

The first point follows from (9.1), the second and third property are due to (7.41) and the fourth property follows from (8.15). The last point is a bit more involved.

The following properties of L_X are also important: If X, Y are vector fields and $\lambda \in \mathbf{R}$, then

- $L_{X+Y} = L_X + L_Y$ and $L_{\lambda X} = \lambda L_X$.
- $L_{[X,Y]} = [L_X, L_Y] = L_X \circ L_Y - L_Y \circ L_X$ which can easily be verified by applying the expression to a function $f \in \mathcal{F}(M)$:

$$[L_X, L_Y]f = XYf - YXf = [X, Y]f = L_{[X,Y]}f. \tag{9.9}$$

which is due to

$$L_X(Y \otimes f) = (L_X Y)f + Y L_X f = [X, Y]f + YXf = XYf = L_X(Yf) \tag{9.10}$$

(note that $(L_X Y)f \neq L_X(Yf)$). Now we can apply the commutator to vector fields using the Jacobi identity:

$$[L_X, L_Y]Z = [X, [Y, Z]] - [Y, [X, Z]] = [[X, Y], Z] \tag{9.11}$$

which proves the property.

If $[X, Y] = 0]$, then $L_X L_Y = L_Y L_X$ and for the two flows ϕ, ψ generated by X and Y we have

$$\phi_s \circ \psi_t = \psi_t \circ \phi_s. \tag{9.12}$$

10 Differential Forms

Definition. A **p-form** Ω is a totally antisymmetric tensor field of type $(0, p)$ with the property that

$$\Omega(X_{\pi(1)}, ..., X_{\pi(p)}) = (\text{sign}(\pi))\Omega(X_1, ..., X_p) \tag{10.1}$$

for any permutation π of $\{1, ..., p\}$. Note that a p-form with $p >$dim(M) vanishes identically.
Any tensor field of type $(0, p)$ can be antisymmetrized by means of the operator $\mathcal{A}$:

$$(\mathcal{A}T)(X_1, ..., X_p) = \frac{1}{p!} \sum_{\pi \in S_p} T(X_{\pi(1)}, ..., X_{\pi(p)}). \tag{10.2}$$

Obviously, we have the property that $\mathcal{A}^2 = \mathcal{A}$.

The **exterior product** of a p_1-form Ω^1 with a p_2-form Ω^2 is the (p_1+p_2)-form

$$\Omega^1 \wedge \Omega^2 = \frac{(p_1+p_2)!}{p_1!p_2!}\mathcal{A}(\Omega^1 \otimes \Omega^2). \tag{10.3}$$

As an example consider the exterior product of two 1-forms ω_1 and ω_2:

$$\omega_1 \wedge \omega_2 = \omega_1 \otimes \omega_2 - \omega_2 \otimes \omega_1. \tag{10.4}$$

If we have n 1-forms, the corresponding antisymmetrization reads

$$\omega_1 \wedge ... \wedge \omega_n = \sum_{\pi \in S_n} \mathrm{sgn}(\pi)\ \omega_{\pi(1)} \otimes ... \otimes \omega_{\pi(n)}. \tag{10.5}$$

The exterior product has the following properties:

- $\Omega^1 \wedge \Omega^2 = (-1)^{p_1 p_2} \Omega^2 \wedge \Omega^1$
- $\Omega^1 \wedge (\Omega^2 \wedge \Omega^3) = (\Omega^1 \wedge \Omega^2) \wedge \Omega^3 = \frac{(p_1+p_2+p_3)!}{p_1!p_2!p_3!}\mathcal{A}(\Omega^1 \otimes \Omega^2 \otimes \Omega^3)$
- We can write the components of a p-form in a (local) basis of 1-forms $(e^1, ..., e^n)$:

$$\begin{aligned}\Omega &= \Omega_{i_1 \cdots i_p} e^{i_1} \otimes \cdots \otimes e^{i_p} = \mathcal{A}\Omega \\ &= \Omega_{i_1 \cdots i_p} \mathcal{A}(e^{i_1} \otimes \cdots \otimes e^{i_p}) \\ &= \Omega_{i_1 \cdots i_p} \frac{1}{p!} e^{i_1} \wedge \ldots \wedge e^{i_p} \end{aligned} \tag{10.6}$$

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

- Vector fields A, B are 1-forms. We have

$$(A \wedge B)_{ik} = A_i B_k - A_k B_i = (-1)(B \wedge A)_{ik} \tag{10.7}$$

- Consider the following example, involving a 2-form A and a 1-form B:

$$(A \wedge B)_{ikl} = A_{ik} B_l + A_{kl} B_i + A_{li} B_k \tag{10.8}$$

  which is due to

$$\begin{aligned} A \wedge B &= \frac{(1+2)!}{1!2!}\mathcal{A}(A \otimes B) \\ &= 3(A_{ik}B_l)\frac{1}{3!}e^i \wedge e^k \wedge e^l \\ &= \frac{1}{2}(A_{jk}B_l)e^i \wedge e^k \wedge e^l \\ &= \frac{1}{2}\frac{1}{3}(A_{ik}B_l + \text{cyclic permutations})e^i \wedge e^k \wedge e^l \\ &= (A_{ik}B_l + \text{cyclic permutations})\frac{1}{3!}e^i \wedge e^k \wedge e^l. \end{aligned} \tag{10.9}$$

10.1 Exterior Derivative of a Differential Form

The derivative df of a 0-form $f \in \mathcal{F}$ is the 1-form $df(X) = Xf$: the argument X (X being a vector) acts as a derivative. In local coordinates, we can write this as

$$df = \frac{\partial f}{\partial x^i} dx^i. \tag{10.10}$$

Exterior differentiation is effected by an operator d applied to forms. It converts a p-form into a $(p+1)$-form (if d acts as a derivative, it is clear, that this has to be so: it should somehow tell how the p-form changes in the direction of some vector, so the resulting object depends on $p+1$ vectos). The derivative $d\Omega$ of a 1-form Ω is given by

$$d\Omega(X_1, X_2) := X_1\Omega(X_2) - X_2\Omega(X_1) - \Omega([X_1, X_2]) \tag{10.11}$$

which is obviously antisymmetric. We can verify that this is indeed a 2-form since

$$\begin{aligned}
X_1\Omega(X_2) &= X_1\underbrace{\langle \Omega, X_2\rangle}_{\text{a 1-form}} \\
&= X_1^i \frac{\partial}{\partial x^i}(\Omega_k X_2^k) \\
&= X_1^i \Omega_{k,1} X_2^k + X_1^i \Omega_k X_{2,i}^k && (10.12) \\
-X_2\Omega(X_1) &= -X_2^k \Omega_{i,k} X_1^i - X_2^k \Omega_i X_{1,k}^i && (10.13) \\
\Omega([X_1, X_2]) &= \langle \Omega, X_1X_2 - X_2 - X_1\rangle \\
&= \Omega_i (X_1X_2 - X_2X_1)^i \\
&= \Omega_i (X_1^k X_{2,k}^i - X_2^k X_{1,k}^i) && (10.14)
\end{aligned}$$

and therefore

$$d\Omega(X_1, X_2) = (\Omega_{k,i} - \Omega_{i,k}) X_1^i X_2^k. \tag{10.15}$$

This is manifestly a 2-form!
One can easily verify that

$$d\Omega(fX_1, X_2) = f d\Omega(X_1, X_2). \tag{10.16}$$

Furthermore,

$$\Omega \wedge f = f\Omega \tag{10.17}$$

since f is a 0-form.
We have a product rule for the exterior derivative:

$$d(\Omega \wedge f) = d\Omega \wedge f - \Omega \wedge df \tag{10.18}$$

which can be verified easily using Eq. (10.2):

$$\begin{aligned} d(f\Omega)(X_1,X_2) &= d(\Omega\wedge f)(X_1,X_2) = X_1(f\Omega)(X_2) \\ &= X_1(f\Omega)(X_2) - X_2(f\Omega)(X_1) - (f\Omega)([X_1,X_2]). \end{aligned} \tag{10.19}$$

$$\begin{aligned} X_1(f\Omega)(X_2) &= X_1^i \frac{\partial}{\partial x^i}(f\Omega_k X_2^k) \\ &= \underbrace{f X_1^i \frac{\partial}{\partial x^i}(\Omega_k X_2^k)}_{=fX_1\Omega(X_2)} + \underbrace{X_1^i \frac{\partial f}{\partial x^i}\Omega_k X^k}_{=df(X_1)\Omega(X_2)} \\ &= \underbrace{f d\Omega(X_1,X_2)}_{=d\Omega\wedge f} + \underbrace{\Omega(X_2)df(X_1) - \Omega(X_1)df(X_2)}_{-\Omega\wedge df} \end{aligned} \tag{10.20}$$

Moreover, we have that $d^2 f = 0$ since

$$\begin{aligned} d^2 f(X_1,X_2) &= X_1 df(X_2) - X_2 df(X_1) - df([X_1,X_2]) \\ &= X_1 X_2 f - X_2 X_1 f - [X_1,X_2]f \\ &= 0. \end{aligned} \tag{10.21}$$

We can generalize the exterior derivative to the case of p-forms Ω in the following way:

$$\begin{aligned} d\Omega(X_1,...,X_{p+1}) = &\sum_{i=1}^{p+1}(-1)^{i+j}\Omega(X_1,...,\hat{X}_i,...,X_{p+1}) \\ &+ \sum_{i<j}(-1)^{i+j}\Omega([X_i,X_j],,X_1,...,\hat{X}_i,...,\hat{X}_j,...,X_{p+1}). \end{aligned} \tag{10.22}$$

One can show that the following properties hold:

- d is a linear map which maps p-forms to $(p+1)$-forms.
- $d(\Omega^1\wedge\Omega^2) = d\Omega^1\wedge\Omega^2 + (-1)^{p_1}\Omega^1\wedge d\Omega^2$.
- $d^2 = 0$, i.e. $d(d\Omega) = 0$.
- $df(X) = Xf$ for all $f \in \mathcal{F}$.

Note that these properties can also be used as the *definition* of the exterior derivative.

By means of Eq. (10.6), we have with respect to a coordinate basis

$$\Omega = \frac{1}{p!}\Omega_{i_1\cdots i_p} dx^{i_1}\wedge\ldots\wedge dx^{i_p}. \tag{10.23}$$

Thus, using that $ddx^{i_p} = 0$, we find

$$
\begin{aligned}
& d\Omega = \frac{1}{p!} d\Omega_{i_1\cdots i_p} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_n} && (10.24)\\
\Rightarrow \quad & p!\, d\Omega = \Omega_{i_1\cdots i_p,i_0} dx^{i_0} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_p} \\
& \qquad = -\Omega_{i_0,i_2,\ldots,i_p,i_1} dx^{i_0} \wedge \ldots \wedge dx^{i_p} \\
& \qquad = (-1)^k \Omega_{i_0\cdots\hat{i}_k\cdots i_p,i_k} dx^{i_0} \wedge \ldots \wedge dx^{i_p} \qquad (k = 0, \ldots, p) \\
& \qquad = \frac{1}{p+1} \sum_{k=0}^{p} (-1)^k \Omega_{i_0\cdots\hat{i}_k\cdots i_p,i_k} dx^{i_0} \wedge \ldots \wedge dx^{i_p} && (10.25)\\
\Rightarrow \quad & d\Omega = \frac{1}{(p+1)!} (d\Omega)_{i_0\cdots i_p}. && (10.26)
\end{aligned}
$$

For example, consider a 1-form Ω. Then

$$(d\Omega)_{ik} = \Omega_{k,i} - \Omega_{i,k}. \tag{10.27}$$

If, on the other hand, Ω is a 2-form, we have that

$$(d\Omega)_{ikl} = \Omega_{ik,l} + \Omega_{kl,i} + \Omega_{li,k}. \tag{10.28}$$

Consider a map $\varphi :\ M \to \bar{M}$ and $\varphi^* :\ T^*_{\bar{p}}\bar{M} \to T^*_p M$. Then one can prove that

$$\varphi^* \circ d = d \circ \varphi^*. \tag{10.29}$$

The proof rests on the observation that according to (10.23), (10.26) it suffices to show the equality for 0-forms and 1-forms. For 0-forms, (10.29) is equivalent to (7.36). For 1-forms, i.e. differentials $d\bar{f}$, we have

$$
\begin{aligned}
(\varphi^* \circ d)(d\bar{f}) &= 0 \\
(d \circ \varphi^*)(d\bar{f}) = d(\varphi^* \circ d\bar{f}) = d(d(\bar{f} \circ \varphi)) = d^2(\bar{f} \circ f) &= 0.
\end{aligned}
$$

Setting $\varphi = \varphi_t$ (the flow generated by X) and taking the time derivative of (9.1) at $t = 0$, one obtains the infinitesimal version of (10.29):

$$L_X \circ d = d \circ L_X \tag{10.30}$$

(because $L_X R = \frac{d}{dt}\varphi^*_t R\big|_{t=0}$.)

Definition. A p-form ω is called

$$
\begin{aligned}
&\textbf{exact} \text{ if } \omega = d\eta \\
&\textbf{closed} \text{ if } d\omega = 0.
\end{aligned}
$$

Note that every exact p-form is closed (since $d^2\eta = 0$). Under the assumptions of the *Poincare lemma*, the converse is true, as well.
Note that η is not unique: A "gauge transformation" $\eta \to \eta + d\rho$ leaves $d\eta$ invariant for any $(p-2)$-form ρ.
In three dimensional vector analysis, these results take the familiar form $\mathrm{div}(\mathrm{grad}(f)) = 0$ and $\mathrm{div}(\mathrm{rot}(\boldsymbol{K})) = 0$.

10.2 The Integral of an n-Form

Assume that M is orientable, i.e. there exists an atlas of "positively oriented" charts with $\det\left(\frac{\partial\bar{x}^i}{\partial x^j}\right) > 0$ for any change of coordinates. An n-form

$$\begin{aligned}\omega &= \omega_{i_1\cdots i_n}\frac{1}{n!}dx^{i_1}\wedge\ldots\wedge dx^{i_n}\\ &= \underbrace{\omega_{1\cdots n}}_{\equiv\omega(x)}dx^1\wedge\ldots\wedge dx^n \qquad (10.31)\end{aligned}$$

transforms under a change of coordinates as

$$\begin{aligned}\bar{\omega}(\bar{x}) &= \bar{\omega}_{1\cdots n}\\ &= \omega_{i_1\cdots i_n}\frac{\partial x^{i_1}}{\partial\bar{x}^1}\cdots\frac{\partial x^{i_n}}{\partial\bar{x}^n}\\ &= \omega(x)\det\left(\frac{\partial x^i}{\partial\bar{x}^j}\right). \qquad (10.32)\end{aligned}$$

The integral of ω over M is defined as

$$\boxed{\int_M \omega = \int_U dx^1\cdots dx^n\omega(x^1,...,x^n)} \qquad (10.33)$$

for some chart U which contains the support of ω. This definition is independent of the choice of coordinates since

$$\int d\bar{x}^1\cdots d\bar{x}^n\bar{\omega}(\bar{x}) = \int dx^1\cdots dx^n\omega(x)\left|\det\left(\frac{\partial x^i}{\partial\bar{x}^j}\right)\right| \qquad (10.34)$$

If there is no single chart U which covers all of M, one has to introduce different charts which cover M in such a way that in overlapping patches the coordinates of one chart can be expressed in a smooth one-to-one way as functions of the coordinates of the other patch. This requires that M is orientable. The integral over a p-form over the overlap region of two patches can then be evaluated using either coordinate system (because of the coordinate-independence of the integral).

10.2.1 Stoke's Theorem

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

$$\int_D d\omega = \int_{\partial D}\omega. \qquad (10.35)$$

10.3 The Inner Product of a p-Form

Let X be a vector field on M. For any p-form Ω let

$$(i_X\Omega)(X_1,...,X_{p-1}) := \Omega(X, X_1,...,X_{p-1}) \tag{10.36}$$

which vanishes if $p = 0$. The map i_X has the following properties:

- i_X maps p-forms to $(p-1)$-forms.
- $i_X(\Omega^1 \wedge \Omega^2) = i_X(\Omega^1) \wedge \Omega^2 + (-1)^{p_1}\Omega^1 \wedge (i_X\Omega^2)$.
- $i_X^2 = 0$.
- $i_X df = Xf$ for all $f \in \mathcal{F}(M)$.
- $L_X = i_X \circ d + d \circ i_X$
(proof: in case of 0-forms, we have $L_X f = Xf$ and $i_X \circ df + d \circ i_X f = i_X df = Xf$ which is the same. In case of 1-forms df, we find $i_X \circ ddf + d \circ i_X df = d(Xf)$ and $L_X df = d(L_X f) = d(Xf)$ which is also the same.)

10.3.1 Application: Gauss' Theorem

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

$$(\mathrm{div}_\eta X)\eta = d(i_X\eta). \tag{10.37}$$

Note that $d\eta = 0$, thus $L_X = i_X \circ d + d \circ i_X$ applied to η yields

$$L_X\eta = i_X \circ d\eta + d(i_X\eta) = d(i_X\eta). \tag{10.38}$$

We can apply Stoke's theorem: $d(i_X\eta)$ is an n-form, hence $i_X\eta$ is an $(n-1)$-form and we find **Gauss Theorem**:

$$\int_D d(i_X\eta) = \int_{\partial D} i_X\eta = \int (\mathrm{div}_\eta X)\eta. \tag{10.39}$$

The **standard volume form** η is given by[7]

$$\eta = \sqrt{|g|}dx^1 \wedge \ldots \wedge dx^n. \tag{10.41}$$

[7] We can convince ourselves that $\sqrt{|g|}$ is the correct factor in front of $dx^1 \wedge \ldots \wedge dx^n$: because $dx^1 \wedge \ldots \wedge dx^n$ is the only n-form up to a multiplicative factor, we need to find the correct prefactor, so that we obtain a tensorial quantity. Under coordinate transformations, $\Omega = e^1 \wedge ... \wedge e^n$ transforms as follows:

$$\begin{aligned}\bar{\Omega} &= \bar{e}^1 \wedge ... \wedge \bar{e}^n \\ &= \phi^1{}_{k_1} \cdots \phi^n{}_{k_n}\ e^{k_1} \wedge ... \wedge e^{k_n} \\ &= \phi^1{}_{k_1} \cdots \phi^n{}_{k_n}\ \mathrm{sgn}(\det(\phi))\varepsilon^{k_1 \cdots k_n}\ e^1 \wedge ... \wedge e^n \\ &= \det(\phi)\Omega.\end{aligned} \tag{10.40}$$

If we express $\mathrm{div}_\eta X$ in local coordinates,

$$\eta = a(x) dx^1 \wedge \ldots \wedge dx^n, \tag{10.42}$$

we find

$$L_X \eta = (Xa) dx^1 \wedge \ldots \wedge dx^n + a \sum_{i=1}^{n} dx^1 \wedge \ldots \wedge d(Xx^i) \wedge \ldots \wedge dx^n \tag{10.43}$$

since $(\mathrm{div}_\eta X)\eta = L_X \eta$. Note that

$$\begin{aligned} d(Xx^i) &= d\left(X^k \frac{\partial x^i}{\partial x^k} \right) \\ &= dX^i(x) \\ &= X^i{}_{,j} dx^j \end{aligned} \tag{10.44}$$

but $dx^1 \wedge \ldots \wedge dx^j \wedge \ldots \wedge dx^k \neq 0$ only if $j = i$. Therefore

$$\begin{aligned} L_X \eta &= X a dx^1 \wedge \ldots \wedge dx^n + a \sum_{i=1}^{n} X^i{}_{,i} dx^1 \wedge \ldots \wedge dx^n \\ &= \left(X^i a_{,i} + a X^i_{,i} \right) \cdot \frac{\eta}{a(x)} \\ &= \mathrm{div}_\eta X \cdot \eta = \frac{1}{a} \left(a X^i \right)_{,i} \eta \end{aligned} \tag{10.45}$$

which is a reminiscent of $\frac{1}{\sqrt{|g|}} \left(\sqrt{|g|} \frac{\partial}{\partial x^i} \right)_{,i}$.

11 Affine Connections

11.1 The Covariant Derivative of a Vector Field

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

- $\nabla_X Y$ is bilinear in X and Y
- If $f \in \mathcal{F}(M)$, then

$$\begin{aligned} & \nabla_{fX} Y = f \nabla_X Y \\ \text{and} \quad & \nabla_X (fY) = f \nabla_X Y + X(f) Y \end{aligned} \tag{11.1}$$

We expect that this covariant derivative provides us with a measure of how the vector field Y changes when we move in the direction of X.[8]

[8] At first sight the asymmetry in the C^∞-linearity in the two arguments may seem strange. But actually it is clear, that if X is rescaled by a function f, then this kind of "directional derivative" has to rescale as a whole because it actually depends only on the direction, not on the magnitude of X and also it depends only on X_p. The vector Y, on the other hand, is the one whose rate of change we are interested in. So if we multiply Y by a function f, then the derivative of f has to play a role, of course.

Lemma. Let X, Y be two vector fields. If X vanishes at p, then $\nabla_X Y$ vanishes, as well.

Proof. Let U be a coordinate neighbourhood of p. On U, we have the representation $X = \xi^i \frac{\partial}{\partial x^i}$ with $\xi^i(p) = 0$. Thus

$$(\nabla_X Y)_p = \nabla_{\xi^i \frac{\partial}{\partial x^i}} Y = \xi^i(p) \left[\nabla_{\frac{\partial}{\partial x^i}} Y\right]_p = 0. \tag{11.2}$$

□

Definition. Let $(x^1, ..., x^n)$ be a chart for $U \subset M$. Set

$$\boxed{\nabla_{\frac{\partial}{\partial x^i}} \left(\frac{\partial}{\partial x^j}\right) := \Gamma^k{}_{ij} \frac{\partial}{\partial x^k}} \tag{11.3}$$

with n^3 functions $\Gamma^k{}_{ij} \in \mathcal{F}(M)$ which are called the **Christoffel symbols** or **Connection coefficients** of the connection ∇ in a given chart.

Note that for a pseudo-Riemannian manifold the corresponding connection coefficients are given by (**??**) or (4.14).
The Christoffel symbols are not tensors:

$$\nabla_{\frac{\partial}{\partial \bar{x}^a}} \left(\frac{\partial}{\partial \bar{x}^b}\right) = \bar{\Gamma}^c{}_{ab} \frac{\partial x^k}{\partial \bar{x}^c} \frac{\partial}{\partial x^k} \tag{11.4}$$

where (by means of Eq. (11.1))

$$\begin{aligned}
\nabla_{\frac{\partial}{\partial \bar{x}^a}} \left(\frac{\partial}{\partial \bar{x}^b}\right) &= \nabla_{\left(\frac{\partial x^i}{\partial \bar{x}^a} \frac{\partial}{\partial x^i}\right)} \left(\frac{\partial x^j}{\partial \bar{x}^b} \frac{\partial}{\partial x^j}\right) \\
&= \frac{\partial x^i}{\partial \bar{x}^a} \left[\frac{\partial x^j}{\partial \bar{x}^b} \Gamma^k{}_{ij} \frac{\partial}{\partial x^k} + \frac{\partial}{\partial x^i} \left(\frac{\partial x^j}{\partial \bar{x}^b}\right) \frac{\partial}{\partial x^j}\right] \\
&= \frac{\partial x^i}{\partial \bar{x}^a} \frac{\partial x^j}{\partial \bar{x}^b} \Gamma^k{}_{ij} \frac{\partial}{\partial x^k} + \frac{\partial^2 x^j}{\partial \bar{x}^a \partial \bar{x}^b} \frac{\partial}{\partial x^j} \qquad (11.5) \\
& \qquad (11.6)
\end{aligned}$$

hence

$$\bar{\Gamma}^c{}_{ab} = \frac{\partial x^i}{\partial \bar{x}^a} \frac{\partial x^j}{\partial \bar{x}^b} \frac{\partial \bar{x}^c}{\partial x^k} \Gamma^k{}_{ij} + \frac{\partial \bar{x}^c}{\partial x^k} \frac{\partial^2 x^k}{\partial \bar{x}^a \partial \bar{x}^b}. \tag{11.7}$$

If for every chart there exist n^3 functions $\Gamma^k{}_{ij}$ which transform according to (11.7) under a change of coordinates, then one can show that there exists a unique affine connection ∇ on M which satisfies (11.4).

Definition. For every vector field X we can introduce the $(1,1)$-tensor ∇X defined by

$$\nabla X(Y, \omega) := \langle \omega, \nabla_Y X \rangle. \tag{11.8}$$

We call ∇X the **covariant derivative** of X. This is motivated by the fact, that a covariantly differentiated vector field still transforms as a tensor.
In a specific chart $(x^1, ..., x^n)$ let $X = \xi^i \partial_i$ and $\nabla X = \xi^i{}_{;j} dx^j \otimes \partial_i$ with

$$\begin{aligned}
\xi^i{}_{;j} &= \nabla X(\partial_j, dx^i) \\
&= \langle dx^i, \nabla_{\partial_j}(\xi^i \partial_i)\rangle \\
&= \langle dx^i, \xi^k{}_{,j}\, \partial_k + \xi^k \Gamma^\delta{}_{jk} \partial_\delta \rangle \\
&= \xi^i{}_{,j} + \Gamma^i{}_{jk}\xi^k.
\end{aligned} \tag{11.9}$$

The semicolon in expressions like $\xi^i{}_{;j}$ denotes the covariant derivative which consits of the "usual" derivative plus additional terms that vanish in Euclidean or Minkowski space. The first term in Eq. (11.9) describes the usual derivative in the coordinate chart. But since the coordinate chart is flat whereas the underlying manifold may be curved, there is the second term. It's meaning can be understood as follows: if a vector is parallel transported along an infinitesimal path in the manifold, the projection of this transport to the coordinate chart does not have to look very "parallel". One has to add some correction to the Euclidean parallel transport in order to get the vector which really corresponds to the transported one. This correction is just the term proportional to $\Gamma^i{}_{jk}$. If the manifold itself is flat, this term vanishes.

11.2 Parallel Transport along a Curve

Definition. Let $\gamma : I \to M$ be a curve in 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

$$\nabla_{\dot\gamma} X = 0 \tag{11.10}$$

along γ. The vector $\nabla_{\dot\gamma} X$ is sometimes denoted as $\frac{DX}{dt}$ or $\frac{\nabla X}{dt}$.
In terms of coordinates, we have $X = \xi^i \partial_i$ and $\dot\gamma = \frac{dx^i}{dt}\partial_i$. With Eqs. (11.1) and (11.3), we find

$$\begin{aligned}
\nabla_{\dot\gamma} X &= \nabla_{\left(\frac{dx^i}{dt}\partial_i\right)}(\xi^k \partial_k) \\
&= \frac{dx^i}{dt} \nabla_{\partial_i}(\xi^k \partial_k) \\
&= \frac{dx^i}{dt}\left[\xi^k \Gamma^j{}_{ik}\partial_j + \partial_i \xi^k \partial_k\right] \\
&= \frac{dx^i}{dt}\left[\xi^j \Gamma^k{}_{ij}\partial_k + \partial_i \xi^k \partial_k\right] \\
&= \left[\frac{d\xi^k}{dt} + \Gamma^k{}_{ij}\frac{dx^i}{dt}\xi^j\right]\partial_k
\end{aligned} \tag{11.11}$$

where we used $\frac{dx^i}{dt}\frac{\partial \xi^k}{\partial x^i} = \frac{d\xi^k}{dt}$.
This shows that $\nabla_{\dot\gamma} X$ only depends on the values of X along γ.

In coordinates, we can write (11.10) as

$$\frac{d\xi^k}{dt} + \Gamma^k{}_{ij}\frac{dx^i}{dt}\xi^j = 0. \tag{11.12}$$

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

$$\tau_{t,s} : T_{\gamma(s)}M \longrightarrow T_{\gamma(t)}M \tag{11.13}$$

which transforms a vector v at $\gamma(s)$ into the parallel transported vector $v(t)$ at $\gamma(t)$. The mapping $\tau_{s,t}$ is the **parallel transport** along γ from $\gamma(s)$ to $\gamma(t)$. The parallel transported vector field $v(t)$ is defined by the property

$$\nabla_{\dot\gamma(t)}v(t) = \nabla_{\dot\gamma(t)}(\tau_{t,s}v) = 0. \tag{11.14}$$

We have $\tau_{s,s} = \mathbf{1}$ and $\tau_{t,s}\circ\tau_{s,r} = \tau_{t,r}$.
We can now give a geometrical interpretation of the covariant derivative that will be generalized to tensors. To this end, let X be a vector field along γ. We then have

$$\nabla_{\dot\gamma}X(\gamma(t)) = \frac{d}{ds}\Big|_{s=t}\tau_{s,t}X(\gamma(s)). \tag{11.15}$$

Proof. We work in a chart. By construction, $v(t) = \tau_{s,t}v(s)$ with $v(s) \in T_{\gamma(s)}M$ and, due to (11.11)

$$\dot v^i + \Gamma^i{}_{kj}\dot x^k v^j = 0. \tag{11.16}$$

If we write $(\tau_{s,t}v(s))^i = (\tau_{s,t})^i_j v^j(s) = v^i(t)$ with $\tau_{s,t} = (\tau_{t,s})^{-1}$ and $\tau_{s,s} = \mathbf{1}$ we find

$$\begin{aligned}
\dot v^i(s) &= \frac{d}{dt}\Big|_{t=s} v^i(t)\\
&= \frac{d}{dt}\Big|_{t=s}\Big[(\tau_{s,t})^i_j v^j(s)\Big]\\
&= \left(\frac{d}{dt}\Big|_{t=s}(\tau_{s,t})^i_j\right)v^j(s)\\
&\overset{!}{=} -\Gamma^i{}_{kj}\dot x^k v^j(s)
\end{aligned} \tag{11.17}$$

and thus

$$\frac{d}{dt}\Big|_{t=s}(\tau_{s,t})^i_j = -\Gamma^i{}_{kj}\dot x^k. \tag{11.18}$$

Since $\tau_{s,t} = (\tau_{t,s})^{-1}$, we have

$$\begin{aligned}
\frac{d}{ds}\Big|_{s=t}(\tau_{s,t})^i_j &= -\frac{d}{dt}\Big|_{s=t}(\tau_{s,t})^i_j\\
&= \Gamma^i{}_{kj}\dot x^k.
\end{aligned} \tag{11.19}$$

This yields

$$\left.\frac{d}{ds}\right|_{s=t}[\tau_{s,t}X(\gamma(s))]^i = \left(\left.\frac{d}{ds}\right|_{s=t}\tau_{s,t}\right)^i_{\ j} X^j + \left.\frac{d}{ds}\right|_{s=t} X^i(\gamma(s))$$
$$= \Gamma^i{}_{kj}\dot{x}^k X^j + X^i{}_{,j}\left.\frac{dx^j(\gamma(s))}{ds}\right|_{s=t} \tag{11.20}$$

which is again (11.11). □

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

We can now give a geometrical interpretation to the previous considerations. Consider the differential $dA^i = A^i{}_{,j}dx^j = A^i(x+dx) - A^i(x)$. In general, we cannot compare A^i at the two different points $x+dx$ and x because $A^i(x+dx)$ and $A^i(x)$ live in different spaces. In order that the difference of two vectors be again a vector, we have to compare them in the same tangent space. The transport has to be chosen such that for cartesian coordinates the vector doesn't change when it is transported. 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 thus meant to satisfy

$$\frac{d}{dt}x^i(t) = X^i(x(t)) \tag{11.21}$$

where we made use of Eqs. (7.6) and (8.18). Inserting this into (11.11) and using $\frac{d^2x^i}{dt^2} = \frac{dX^i}{dt}$, we infer that

$$\boxed{\ddot{x}^k + \Gamma^k{}_{ij}\dot{x}^i\dot{x}^j = 0.} \tag{11.22}$$

11.2.1 Roundtrip by Parallel Transport

Considering Eq. (11.11) and denoting $\xi^i = v^i$, we find

$$\dot{v}^i = -\Gamma^i{}_{kj}\dot{x}^k v^j. \tag{11.23}$$

Let $\gamma:\ [0,1] \to M$ be a closed path with $\gamma(0) = \gamma(1) = p$. Displace a vector $v_0 \in T_pM$ parallel along γ. This yields the parallel transported vector field

$$v(t) = \tau_{t,0}v_0 \in T_{\gamma(t)}M. \tag{11.24}$$

We assume the closed path to be sufficiently small such that the relevant region can be covered by a single chart. We can then expand $\Gamma^i{}_{jk}(x)$ around the point $x(0) = x_0$ on the curve:

$$\Gamma^i{}_{kj}(x) \simeq \Gamma^i{}_{kj}(x_0) + (x^\rho - x_0^\rho)\frac{\partial}{\partial x^\rho}\Gamma^i{}_{kj}(x)\big|_{x=x_0} + ... \tag{11.25}$$

To first order in $(x^\rho - x_0^\rho)$ we have (due to (11.23))

$$\begin{aligned}
\int_0^t \dot{v}^i dt' &= v^i(t) - v_0^i \\
&= -\int_0^t \Gamma^i{}_{kj}(x_0) \underbrace{v^j}_{\simeq v_0^j} \dot{x}^k dt' \\
&= -\Gamma^i{}_{kj}(x_0) v_0^j \int_0^t \dot{x}^k dt' \\
&= -\Gamma^i{}_{kj}(x_0) v_0^j (x^k(t) - x_0^k) \qquad (11.26)
\end{aligned}$$

$$\Rightarrow \qquad v^i(t) \simeq v_0^i - \Gamma^i{}_{kj}(x_0)(x^k(t) - x_0^k) v_0^j + \ldots \qquad (11.27)$$

Using Eqs. (11.25) and (11.27) and plugging them into (11.23), we obtain an equation valid to second order:

$$\begin{aligned}
\underbrace{\int_0^1 \dot{v}^i dt}_{=v^i(1)-v_0^i} &= -\int_0^1 \Gamma^i{}_{kj}(x) \dot{x}^k v^j dt \\
&\simeq -\int_0^1 \left(\Gamma^i{}_{kj}(x_0) + (x^\rho - x_0^\rho)\frac{\partial}{\partial x^\rho}\Gamma^i{}_{kj} + \ldots\right)\left(v_0^j - \Gamma^j_{\tilde{k}\tilde{j}}(x_0)(x^{\tilde{k}} - x_0^{\tilde{k}}) v_0^{\tilde{j}} + \ldots\right) \dot{x}^k dt.
\end{aligned} \qquad (11.28)$$

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

$$v^i(1) \simeq v_0^i - \Gamma^i{}_{kj}(x_0) v_0^j \int_0^1 \dot{x}^k dt - \left(\frac{\partial}{\partial x^\rho}\Gamma^i{}_{kj} - \Gamma^i{}_{k\tilde{j}}(x_0)\Gamma^{\tilde{j}}{}_{\rho j}(x_0)\right) v_0^j \int_0^1 (x^\rho - x_0^\rho)\dot{x}^k dt. \qquad (11.29)$$

Since the path is closed, the second term vanishes due to $\int_0^1 \dot{x}^k dt = 0$. We conclude that

$$\begin{aligned}
\Delta v^i &= v^i(1) - v_0^i \\
&= -\left(\frac{\partial}{\partial x^\rho}\Gamma^i{}_{kj}(x_0) - \Gamma^i{}_{kl}(x_0)\Gamma^l{}_{\rho j}(x_0)\right) v_0^j \int_0^1 x^\rho \dot{x}^k dt \qquad (11.30)
\end{aligned}$$

where

$$\oint_0^1 x^\rho \dot{x}^k dt = \oint_0^1 \frac{d}{dt}(x^\rho x^k) dt - \oint_0^1 \dot{x}^\rho x^k dt = -\oint_0^1 x^k \dot{x}^\rho dt \qquad (11.31)$$

is antisymmetric in the indices ρ and j which implies that

$$\begin{aligned}
\Delta v^i &= -\frac{1}{2}\left(\frac{\partial}{\partial x^\rho}\Gamma^i{}_{kj}(x_0) - \Gamma^i{}_{kl}\Gamma^l{}_{\rho j} - \frac{\partial}{\partial x^k}\Gamma^i{}_{\rho j} + \Gamma^i{}_{\rho l}\Gamma^l{}_{kj}\right) v_0^i \oint_0^1 x^\rho \dot{x}^k dt \\
&=: \frac{1}{2} R^i{}_{jk\rho}(x_0) v_0^j \oint_0^1 x^\rho \dot{x}^k dt \qquad (11.32)
\end{aligned}$$

with the **curvature tensor**

$$\boxed{R^i{}_{jk\rho}(x_0) = \frac{\partial}{\partial x^k}\Gamma^i{}_{\rho j} - \frac{\partial}{\partial x^\rho}\Gamma^i{}_{kj} + \Gamma^l{}_{\rho j}\Gamma^i{}_{kl} - \Gamma^l{}_{kj}\Gamma^i{}_{\rho l}.} \tag{11.33}$$

Therefore an arbitrary vector v^i will not change when parallel transported around an arbitrarily small closed curve at x_0 if and only if $R^i{}_{jk\rho} = 0$.

11.3 Covariant Derivative of Tensor Fields

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

- $\tau_{t,s}(T \otimes S) = (\tau_{t,s}T) \otimes (\tau_{t,s}S)$,
- $\tau_{t,s}\ \mathrm{tr}(T) = \mathrm{tr}(\tau_{t,s}T)$,
- $\tau_{t,s}c = c$ for $c \in \mathbb{R}$.

For example, for a covector ω and a vector X, we have

$$\langle \tau_{t,s}\omega, \tau_{t,s}X\rangle = \langle \omega, X\rangle_{\gamma(t)} \tag{11.34}$$

and for a tensor of type (1,1):

$$\tau_{t,s}T(\tau_{s,t}\omega, \tau_{s,t}X) = T(\omega, X) \tag{11.35}$$

which reads as follows in components:

$$(\tau_{t,s}T)^i{}_k = T^\alpha{}_\beta(\tau_{t,s})^i{}_\alpha(\tau_{t,s})^\beta{}_k \tag{11.36}$$

(note that $\tau^k{}_i$ is the transposed inverse of $\tau^i{}_k$).

Definition. The **covariant derivative** $\nabla_X T$ of a tensor field T with respect to a vector field X associated to τ is

$$(\nabla_X T)_p = \frac{d}{dt}\tau_{0,t}T_{\gamma(t)}\big|_{t=0} \tag{11.37}$$

where $\gamma(t)$ is any curve satisfying $\gamma(0) = p$ and $\dot\gamma(0) = X_p$.
Some properties of the covariant derivative:

- ∇_X is a linear map from tensor fields to tensor fields of the same type (r, s),
- $\nabla_X f = Xf$,
- $\nabla_X(\mathrm{tr}T) = \mathrm{tr}(\nabla_X T)$,
- $\nabla_X(T \otimes S) = (\nabla_X T) \otimes S + T \otimes \nabla_X S$.

We conclude that for a 1-form ω we have

$$\begin{aligned}(\nabla_X\omega)(Y) &= \mathrm{tr}\,(\nabla_X\omega\otimes Y)\\ &= \mathrm{tr}\nabla_X\,(\omega\otimes Y) - \mathrm{tr}(\omega\otimes\nabla_X Y)\\ &= \nabla_X\mathrm{tr}(\omega\otimes Y) - \omega(\nabla_X Y)\\ &= X\omega(Y) - \omega(\nabla_X Y). \end{aligned} \tag{11.38}$$

The general rule for differentiation for a tensor field of type $(1,1)$ reads

$$(\nabla_X T)(\omega, Y) = XT(\omega, Y) - T(\nabla_X\omega, Y) - T(\omega, \nabla_X Y). \tag{11.39}$$

Due to the above properties of ∇_X, the covariant derivative is completely determined by its action on vector fields Y which are the affine connections (c.f. Eqs. (11.1) and (11.3)).

11.3.1 Local Coordinate Expressions for the Covariant Derivative

Let $T\in\mathcal{T}^q_p(U)$ where U is a chart described by local coordinates $(x^1,...,x^n)$, i.e.

$$T^{i_1\cdots i_p}{}_{j_1\cdots j_q}\partial_{i_1}\otimes\cdots\otimes\partial_{i_p}\otimes dx^{j_1}\otimes\cdots\otimes dx^{j_q}. \tag{11.40}$$

We can therefore compute

$$XT^{i_1\cdots i_p}{}_{j_1\cdots j_q} = X^k T^{i_1\cdots i_p}{}_{j_1\cdots j_q,k}. \tag{11.41}$$

Using Eq. (11.3), we find

$$\begin{aligned}\nabla_X(\partial_i) &= X^k\nabla_{\partial_k}\partial_i\\ &= X^k\Gamma^l{}_{ki}\partial_l. \end{aligned} \tag{11.42}$$

$$\begin{aligned}(\nabla_X dx^j)(\partial_i) &= X\underbrace{\langle dx^j,\partial_i\rangle}_{=0} - \langle dx^j, \nabla_X\partial_i\rangle\\ &= -X^k\Gamma^j{}_{ki} \end{aligned} \tag{11.43}$$

where we used (11.38) to obtain the second expression. We can write the second expression as

$$\boxed{\nabla_X dx^j = -X^k\Gamma^j{}_{ki}dx^i.} \tag{11.44}$$

Using Eqs. (11.41), (11.42), (11.44) for $\omega_j = dx^j$, $Y_i = \partial_i$, we obtain

$$\begin{aligned}T^{i_1\cdots i_p}{}_{j_1\cdots j_q;k} = T^{i_1\cdots i_p}{}_{j_1\cdots j_q,k} &+ \Gamma^{i_1}{}_{kl}T^{li_2\cdots i_p}{}_{j_1\cdots j_q} + \ldots + \Gamma^{i_p}{}_{kl}T^{i_1\cdots i_{p-1}l}{}_{j_1\cdots j_q}\\ &- \Gamma^l{}_{kj_1}T^{i_1\cdots i_p}{}_{lj_2\cdots j_k} - \ldots - \Gamma^l{}_{kj_q}T^{i_1\cdots i_p}{}_{j_1\cdots j_{q-1}l}. \end{aligned} \tag{11.45}$$

For example, we find for contravariant and covariant vector fields:

$$\xi^i{}_{;k} = \xi^i{}_{,k} + \Gamma^i{}_{kl}\xi^l \tag{11.46}$$

$$\eta_{i;k} = \eta_{i,k} - \Gamma^l{}_{ki}\eta_l. \tag{11.47}$$

For the Kronecker tensor, we find

$$\delta^i{}_{j;k} = 0. \tag{11.48}$$

Similarly for a $(1,1)$ tensor:

$$T^i{}_{k;r} = T^i{}_{k,r} + \Gamma^i{}_{rl} T^l{}_k - \Gamma^l{}_{rk} T^i{}_l. \tag{11.49}$$

Consider the covariant derivative of the metric $g_{\mu\nu}$:

$$\boxed{g_{\mu\nu;\lambda} = \frac{\partial g_{\mu\nu}}{\partial x^\lambda} - \Gamma^\rho{}_{\lambda\mu} g_{\rho\nu} - \Gamma^\rho{}_{\lambda\nu} g_{\rho\mu} = 0} \tag{11.50}$$

where the last step can easily be verified by inserting the expressions given by (4.14) for $\Gamma^\rho{}_{\lambda\mu}$.

So our "receipt" to make special-relativistic equations that hold in the absence of gravitation valid in curved spacetime is to replace $\eta_{\alpha\beta}$ by $g_{\alpha\beta}$ and replace all derivatives by covariant derivatives $(, \to ;)$. The resulting equations will be generally covariant and true in the presence of gravitational fields.

12 Curvature and Torsion of an Affine Connection - Bianchi Identities

Let an affine connection be given on M and let X, Y, Z be vector fields.

Definition. We define the **torsion tensor**

$$T(X,Y) := \nabla_X Y - \nabla_Y X - [X,Y]. \tag{12.1}$$

The torsion tensor $T(X,Y)$ is antisymmetric and f-linear in X, Y:

$$T(fX, gY) = fgT(X,Y) \qquad \text{for } f,g \in \mathcal{F}(M). \tag{12.2}$$

It thus defines a tensor of type $(1,2)$ through

$$(\omega, X, Y) \longrightarrow \langle \omega, T(X,Y) \rangle. \tag{12.3}$$

In local cordinates the components of the torsion tensor are given by

$$\begin{aligned} T^k{}_{ij} &= \langle dx^k, T(\partial_i, \partial_j) \rangle \\ &= \langle dx^k, \nabla_{\partial_i} \partial_j - \nabla_{\partial_j} \partial_i - \underbrace{[\partial_i, \partial_j]}_{=0} \rangle \\ &= \Gamma^k{}_{ij} - \Gamma^k{}_{ji}. \end{aligned} \tag{12.4}$$

In particular

$$T^k{}_{ij} = 0 \quad \Leftrightarrow \Gamma^k{}_{ij} = \Gamma^k{}_{ji}. \tag{12.5}$$

Definition. We define

$$R(X,Y) := \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}. \tag{12.6}$$

This is also linear in the sense that

$$R(fX, gY)hZ = fghR(X,Y)Z \quad \text{for } f, g, h \in \mathcal{F}(M) \tag{12.7}$$

and antisymmetric in X, Y. R determines a type-$(1,3)$ tensor which is called the **Riemann tensor** or **curvature tensor** given by

$$(\omega, Z, X, Y) \longrightarrow \langle \omega, R(X,Y)Z \rangle \equiv R^i{}_{jkl}\omega_i Z^j X^k Y^l. \tag{12.8}$$

The components with respect to local coordinates are

$$\begin{aligned} R^i{}_{jkl} &= \langle dx^i, R(\partial_k, \partial_l)\partial_j \rangle \\ &= \langle dx^i, (\nabla_{\partial_k}\nabla_{\partial_l} - \nabla_{\partial_l}\nabla_{\partial_k})\partial_j \rangle \\ &= \langle dx^i, \nabla_{\partial_k}(\Gamma^s{}_{lj}\partial_s) - \nabla_{\partial_l}(\Gamma^s{}_{kj}\partial_s) \rangle \\ &= \Gamma^i{}_{lj,k} - \Gamma^i{}_{kj,l} + \Gamma^s{}_{lj}\Gamma^i{}_{ks} - \Gamma^s{}_{kj}\Gamma^i{}_{ls}. \end{aligned} \tag{12.9}$$

Note that Eq. (12.9) is exactly the same as defined in (11.33). It is antisymmetric in the last two indices:

$$R^i{}_{jkl} = -R^i{}_{jlk}. \tag{12.10}$$

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

$$\begin{aligned} R_{jl} &= R^i{}_{jil} \\ &= \Gamma^i{}_{lj,i} - \Gamma^i{}_{ij,l} + \Gamma^s{}_{lj}\Gamma^i{}_{is} - \Gamma^s{}_{ij}\Gamma^i{}_{ls}. \end{aligned} \tag{12.11}$$

Furthermore, we define the **scalar curvature**

$$R := g^{lj} R_{jl} = R^l{}_l \tag{12.12}$$

As an example, consider the connection coefficients for a pseudo-Riemannian manifold given by (4.14). One can easily compute $\Gamma^i{}_{jk}$ using (4.14):

- Consider the two-sphere S^2 which is a pseudo-Riemannian manifold with the metric

$$ds^2 = a^2(d\theta^2 + \sin^2\theta d\phi^2) \quad \Rightarrow \quad g_{\theta\phi} = a^2 \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix} \tag{12.13}$$

  The non-zero components are

$$\begin{aligned} \Gamma^\theta{}_{\phi\phi} &= -\sin\theta\cos\theta \\ \Gamma^\phi{}_{\theta\phi} = \Gamma^\phi{}_{\phi\theta} &= \cot\theta \end{aligned}$$

so the Riemann tensor reads

$$\begin{aligned}R^{\theta}{}_{\phi\theta\phi} &= \partial_\theta \Gamma^{\theta}{}_{\phi\phi} - \partial_\phi \Gamma^{\theta}{}_{\theta\phi} + \Gamma^{\theta}{}_{\theta\lambda}\Gamma^{\lambda}{}_{\phi\theta} - \Gamma^{\theta}{}_{\phi\lambda}\Gamma^{\lambda}\theta\phi \\ &= (\sin^{\theta} - \cos^2\theta) - 0 + 0 - (-\sin\theta\cos\theta)\mathrm{cot}\theta \\ &= \sin^2\theta. \end{aligned} \tag{12.14}$$

The Ricci tensor has the following components:

$$\begin{aligned} R_{\phi\phi} &= \sin^2\theta \\ R_{\theta\theta} &= 1 \\ R_{\phi\theta} &= R_{\theta\phi} = 0. \end{aligned} \tag{12.15}$$

The Ricci scalar takes the form

$$\begin{aligned} R &= g^{\theta\theta}R_{\theta\theta} + g^{\phi\phi}R_{\phi\phi} + g^{\theta\phi}R_{\phi\theta} + g^{\phi\theta}R_{\theta\phi} \\ &= \frac{1}{a^2} + \frac{1}{a^2\sin^2\theta}\sin^2\theta \\ &= \frac{2}{a^2}. \end{aligned} \tag{12.16}$$

The Ricci scalar is constant over S^2. It is also strictly positive which means that S^2 has "positive curvature".

Note that for a position independent metric (e.g. cartesian coordinates) the Riemann tensor (and thus the scalar curvature) vanishes.
For a plane with polar coordinates, we get a position dependent metric $\begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$, so the Christoffel symbols do *not* vanish! However, the curvature vanishes. The curvature does not depend on the choice of coordinates.

12.0.2 Bianch Identities for the Special Case of Vanishing Torsion

Let X, Y, Z be vector fields. Then

$$R(X,Y)Z + \text{cyclic} = 0 \quad \text{(first Bianchi identity)} \tag{12.17}$$

$$(\nabla_X R)(Y,Z) + \text{cyclic} = 0 \quad \text{(second Bianchi identity).} \tag{12.18}$$

Proof. We prove the first Bianchi identity:
Due to vanishing torsion, $\nabla_X Y - \nabla_Y X = [X,Y]$, we have

$$\begin{aligned} &(\nabla_X\nabla_Y - \nabla_Y\nabla_X)Z + (\nabla_Y\nabla_Z - \nabla_Z\nabla_Y)X + (\nabla_Z\nabla_X - \nabla_X\nabla_Z)Y \\ &\qquad\qquad - \nabla_{[X,Y]}Z - \nabla_{[Y,Z]}X - \nabla_{[Z,X]}Y \\ &= \nabla_X \underbrace{(\nabla_Y Z - \nabla_Z Y)}_{=[Y,Z]=T} - \nabla_{\underbrace{[Y,Z]}_{=T}} X + \text{cyclic} \\ &= [X,[Y,Z]] + \text{cyclic} \\ &= 0 \end{aligned} \tag{12.19}$$

where we used the Jacobi identity (8.5). □

13 Riemannian Connections

Definition. Let M be equipped with a **pseudo-Riemannian metric**, i.e. a symmetric, non-degenerate type $(0,2)$ tensor field $g(X,Y) \equiv (X,Y)$. Non-degeneracy means that for any $p \in M$ and $X, Y \in T_pM$ we have

$$g_p(X,Y) = 0 \;\forall Y \in T_pM \quad \Rightarrow X = 0. \tag{13.1}$$

In components the metric reads

$$(X,Y) = g_{ik}X^iY^k \quad \text{with } g_{ik} = g_{ki} \text{ and } \det(g_{ik}) \neq 0. \tag{13.2}$$

We can use the metric to lower and raise indices[9]:

$$\widetilde{X}_i = g_{ik}X^k \;\text{ and }\; \widetilde{\omega}^i = g^{ik}\omega_k. \tag{13.7}$$

This also works for tensor fields of different type, of course. For example,

$$T^i{}_k = T_{lk}g^{il} = T^{il}g_{lk}. \tag{13.8}$$

Given a basis $(e_1, ..., e_n)$ of T_pM, the covectors of the dual basis $(e^1, ..., e^n)$ become themselves vectors:

$$e_i = g_{ij}e^j. \tag{13.9}$$

Definition. A metric tensor g at a point $p \in M$ is a symmetric $(0,2)$-tensor. It assigns a length

$$d(X) := \sqrt{|g(X,X)|} \tag{13.10}$$

to each vector $X \in T_pM$. It also defines the angle between any two vectors X, $Y (\neq 0)$ via

$$a(X,Y) = \arccos\left(\frac{g(X,Y)}{d(X)d(Y)}\right). \tag{13.11}$$

[9] In order to see this, we introduce the mapping

$$\flat : TM_p \to T^*M_p, \quad X_p \mapsto X_p^\flat \tag{13.3}$$
$$\text{where } X_p^\flat(Y_p) = g_p(X_p, Y_p). \tag{13.4}$$

One can show that $\flat$ provides an isomorphism $T_pM \to T_p^*M$ with well-defined inverse. In components we have on the one hand

$$X_p^\flat(Y_p) = g(X^i\partial_i, Y^j\partial_j) = X^iY^jg(\partial_i,\partial_j) = g_{ij}X^iY^j. \tag{13.5}$$

On the other hand

$$X_p^\flat(Y_p) = X_i dx^i(Y^j\partial_j) = X_iY^j\delta^i_j = X_iY^i \tag{13.6}$$

where $X^\flat = X_i dx^i$ is the one-form associated to X. Comparison of these two results yields $X_i = g_{ij}X^j$. Similarly one can show that $X^i = g^{ij}X_j$.

If $g(X,Y)=0$, then X and Y are said to be orthogonal.
The length of a curve with tangent vector X between t_1 and t_2 is

$$L(t_1,t_2)=\int_{t_1}^{t_2} d(X)dt. \tag{13.12}$$

If (e_a) is a basis of T_pM, the components of g with respect to this basis are $g_{ab}=g(e_a,e_b)$. Like in special relativity, we can classify vectors at a point as

- timelike: $g(X,X)>0$,
- null: $g(X,X)=0$,
- spacelike: $g(X,X)<0$.

Definition. Let (M,g) be a pseudo-Riemannian manifold. An affine connection is a **metric connection** if parallel transport along any smooth curve γ in M preserves the inner product. For autoparallel fields $X(t)$, $Y(t)$ (c.f. Eq. (11.10)) along γ, $g_{\gamma(t)}(X(t),Y(t))$ is independent of t.

Theorem. An affine connection ∇ is metric if and only if

$$\nabla g=0. \tag{13.13}$$

(without proof)
Eq. (13.13) is equivalent to

$$0=\nabla_X g = Xg(Y,Z)-g(\nabla_X Y,Z)-g(Y,\nabla_X Z) \tag{13.14}$$

$$\Leftrightarrow \quad Xg(Y,Z)=g(\nabla_X Y,Z)+g(Y,\nabla_X Z). \tag{13.15}$$

Theorem. For every pseudo-Riemannian manifold (M,g) there exists a unique affine connection such that

1. ∇ has vanishing torsion (∇ is *symmetric*) and
2. ∇ is metric.

This particular connection is called **Riemannian** or **Levi-Civita connection**.

Proof. Vanishing torsion means that $\nabla_X Y=\nabla_Y X+[X,Y]$. Inserting this condition into (13.15) we obtain

$$Xg(Y,Z)=g(\nabla_Y X,Z)+g([X,Y],Z)+g(Y,\nabla_X Z) \tag{13.16}$$

and the same equation twice more with X, Y, Z cyclically permuted:

$$Yg(Z,X)=g(\nabla_Z Y,X)+g([Y,Z],X)+g(Z,\nabla_Y X) \tag{13.17}$$

$$Zg(X,Y)=g(\nabla_X Z,Y)+g([Z,X],Y)+g(X,\nabla_Z Y). \tag{13.18}$$

We take the linear combination (13.17)+(13.18)−(13.16) and obtain the *Koszul formula*:

$$\begin{aligned}2g(\nabla_Z Y,X)=&-Xg(Y,Z)+Yg(Z,X)+Zg(X,Y)\\ &-g([Z,X],Y)-g([Y,Z],X)+g([X,Y],Z).\end{aligned} \tag{13.19}$$

The right hand side is independent of ∇. Since g is non-degenerate, uniqueness of ∇ follows from (13.19). □

We want to determine the Christoffel symbols for the Levi-Civita connection in a chart $(x^1, ..., x^n)$. For this purpose we take $X = \partial_k$, $Y = \partial_j$, $Z = \partial_i$ in Eq. (13.19) and use $[\partial_i, \partial_j] = 0$ as well as $\langle \partial_i, \partial_j \rangle = g_{ij}$. The result is

$$\begin{aligned} 2\langle \nabla_{\partial_i}\partial_j, \partial_k \rangle &= 2\Gamma^l{}_{ij} \langle \partial_l, \partial_k \rangle \\ &= 2\Gamma^l{}_{ij} g_{lk} \\ &= -\partial_k \underbrace{\langle \partial_j, \partial_i \rangle}_{=g_{ji}} + \partial_j \underbrace{\langle \partial_i, \partial_k \rangle}_{=g_{ik}} + \partial_i \underbrace{\langle \partial_k, \partial_j \rangle}_{=g_{kj}} \end{aligned} \tag{13.20}$$

which is equivalent to

$$g_{lk}\Gamma^l{}_{ij} = \frac{1}{2}(g_{kj,i} + g_{ik,j} - g_{ji,k}). \tag{13.21}$$

Denoting by g^{kj} the inverse metric, we obtain

$$\Gamma^l{}_{ij} = \frac{1}{2} g^{lk}(g_{kj,i} + g_{ik,j} - g_{ji,k}) \tag{13.22}$$

which is exactly the same as Eq. (4.14).

The Levi-Civita connection enjoys the following properties:

- The inner product of any two vectors stays constant upon parallel transport along any curve, i.e.

$$g(X, Y)_{\gamma(t)} = g(X, Y)_{\gamma(0)}. \tag{13.23}$$

- The covariant derivative commutes with raising and lowering indices. e.g.

$$T^i{}_{k;l} = \left(g_{km} T^{im}\right)_{;l} = g_{km} T^{im}{}_{;l}. \tag{13.24}$$

  because of $g_{km;l} = 0$.

13.0.3 The Riemann Tensor

The curvature tensor of a Riemannian connection has the following additional symmetry properties (without proof):

$$(i) \quad \langle R(X,Y)Z, U \rangle = -\langle R(X,Y)U, Z \rangle \quad (\Leftrightarrow R_{ijkl} = -R_{jikl}), \tag{13.25}$$

$$(ii) \quad \langle R(X,Y)Z, U \rangle = \langle R(Z,U)X, Y \rangle \quad (\Leftrightarrow R_{ijkl} = R_{klij}). \tag{13.26}$$

In coordinates the Riemann tensor satisfies the following symmetries:

$$(i) \quad R^i{}_{jkl} = -R^i{}_{jlk} \tag{13.27}$$

$$(ii) \quad R^i{}_{[jkl]} \equiv \sum_{(jkl)} R^i{}_{jkl} = 0 \quad \text{(1. Bianchi identity)}, \tag{13.28}$$

$$(iii) \quad R^i{}_{j[kl;m]} \equiv \sum_{(klm)} R^i{}_{jkl;m} = 0 \quad \text{(2. Bianchi identity)} \tag{13.29}$$

where $\sum_{(klm)}$ denotes the cyclic sum. Note that the Bianchi identities only hold if the connection is torsion free.

13.0.4 The Ricci and Einstein-Tensor

The Ricci tensor is defined as

$$R_{ik} = R^j{}_{ijk}. \tag{13.30}$$

The scalar curvature is

$$R = R^i{}_i. \tag{13.31}$$

The **Einstein tensor** G_{ik} is defined as follows:

$$G_{ik} = R_{ik} - \frac{1}{2}Rg_{ik}. \tag{13.32}$$

Apart from the obvious symmetries $R_{ik} = R_{ki}$ and $G_{ik} = G_{ki}$. we have the following identities:

$$R_i{}^k{}_{;k} = \frac{1}{2}R_{;i} \tag{13.33}$$

$$G_i{}^k{}_{;k} = 0. \tag{13.34}$$

Proof. The second Bianchi identity reads

$$R^i{}_{jkl;m} + R^i{}_{jlm;k} + R^i{}_{jmk;l} = 0. \tag{13.35}$$

Contracting the indices i and k, we obtain

$$R_{jl;m} + \underbrace{R^i{}_{jlm;i}}_{=-g^{ik}R_{jklm;i}} - R_{jm;l} = 0 \tag{13.36}$$

$$\Rightarrow \quad R^j{}_{l;m} - g^{ik}R^j{}_{lkm;i} - R^j{}_{m;l} = 0. \tag{13.37}$$

Taking the trace in the indices j and m, this implies

$$\underbrace{R^j{}_{l;j} - g^{ik}R_{kl;i}}_{=2R^j{}_{l;j}} - R_{;l} = 0 \tag{13.38}$$

which is equivalent to Eq. (13.33)

In order to prove the identity (13.34), we observe that

$$G_i{}^k = R_i{}^k - \frac{1}{2}g_i{}^k R = R_i{}^k - \frac{1}{2}\delta_i{}^k R \tag{13.39}$$

$$\Rightarrow \quad G_i{}^k{}_{;k} = R_i{}^k{}_{;k} - \frac{1}{2}(\delta_i{}^k R_{;k})$$

$$= \frac{1}{2}R_{;i} - \frac{1}{2}R_{;i} = 0 \tag{13.40}$$

which implies (13.34). □

Note that 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 IV
General Relativity

14 Physical Laws with Gravitation

The physical laws are relations between tensors (scalars, vectors, ...). Thus the physical laws read the same in every coordinate system (provided the physical quantities are transformed suitably), i.e. the equations assume the same *form* in every system. This property is called **general covariance**.
Practically, we get general covariant equations which are valid in the presence of gravity by replacing $\eta_{\alpha\beta}$ by $g_{\alpha\beta}$ and derivatives by covariant derivatives in the special relativistic laws which hold in the absence of gravitation.
Examples:

14.0.5 Mechanics

In an inertial system, we have the equation of motion

$$m\frac{du^\alpha}{d\tau} = f^\alpha. \tag{14.1}$$

According to the equivalence principle, this equation holds in a local IS. The force f^α does not contain gravitational forces as they would vanish in a local IS.
We can now easily transform this equation to a general coordinate system KS. The Lorentz vector f^α becomes

$$f^\mu(x) = \frac{\partial x^\mu}{\partial \xi^\alpha} f^\alpha(\xi(x)) \tag{14.2}$$

where ξ^α are the coordinates in the local IS and x^μ is the general KS. Eq. (14.1) reads then

$$m\frac{Du^\mu}{d\tau} = f^\mu \tag{14.3}$$

where $\frac{D}{d\tau}$ denotes the covariant derivative as in (11.11), i.e.

$$\frac{Du^\mu}{d\tau} = \frac{du^\mu}{d\tau} + \Gamma^\mu{}_{\nu\lambda} u^\nu u^\lambda. \tag{14.4}$$

We can thus write Eq. (14.3) as

$$m\frac{du^\mu}{d\tau} = f^\mu - m\Gamma^\mu{}_{\nu\lambda} u^\nu u^\lambda. \tag{14.5}$$

We see that on the right-hand side there appear now gravitational forces explicitly (via $\Gamma^\mu{}_{\nu\lambda}$).
Eq. (14.5) is satisfies general covariance, i.e. it has the same form in all coordinate systems. If we replace $g_{\mu\nu}$ by $\eta_{\mu\nu}$ (i.e. we are in a local inertial system), Eq. (14.5) reduces to (14.1).

Note that the components of u^μ are not independent but satisfy the condition $g_{\mu\nu} u^\mu u^\nu = c^2$.

14.0.6 Electrodynamics

According to the equivalence principle Maxwell's equations (2.59), (2.60),

$$\partial_\alpha F^{\alpha\beta} = \frac{4\pi}{c^2} j^\beta \quad \text{and} \quad \varepsilon^{\alpha\beta\gamma\delta}\partial_\beta F_{\gamma\delta} = 0 \tag{14.6}$$

are valid in a local inertial system.
Using the covariance principle, these equations take the following form in a general coordinate system KS:

$$F^{\mu\nu}{}_{;\mu} = \frac{4\pi}{c} j^\nu \quad \text{and} \quad \varepsilon^{\mu\nu\lambda\kappa} F_{\lambda\kappa;\nu} = 0 \tag{14.7}$$

provided that the transformation from coordinates ξ^α in a local IS to the general coordinates x^μ is given by

$$j^\alpha \longrightarrow j^\mu = \frac{\partial x^\mu}{\partial \xi^\alpha} j^\alpha \tag{14.8}$$

$$\Rightarrow \quad F^{\alpha\beta} \longrightarrow F^{\mu\nu} = \frac{\partial x^\mu}{\partial \xi^\alpha}\frac{\partial x^\nu}{\partial \xi^\beta} F^{\alpha\beta}. \tag{14.9}$$

Again, gravity enters via the Christoffel symbols in the covariant derivative.
The continuity equation $\partial_\alpha j^\alpha$ becomes

$$j^\mu{}_{;\mu} = 0. \tag{14.10}$$

It can be shown that due to the antisymmetry of $F_{\mu\nu}$ the terms involving Christoffel symbols cancel. Thus the covariant derivative reduces to the ordinary derivative in case of the homogeneous Maxwell equation in (14.7).

Note that

$$g = \det(g_{ik}) = \varepsilon^{i_1\cdots i_n} g_{1i_1}\cdots g_{ni_n}. \tag{14.11}$$

We can use this to consider

$$\begin{aligned} \frac{\partial g}{\partial x^l} &= \sum_{k=1}^{n} \varepsilon^{i_1\cdots i_n} g_{1i_1}\cdots \frac{\partial g_{ki_k}}{\partial x^l}\cdots g_{ni_n} \\ &= \frac{\partial g_{km}}{\partial x^l} g^{mk} g \end{aligned} \tag{14.12}$$

where we used the antisymmetry of $\varepsilon^{i_1\cdots i_n}$ and

$$\frac{\partial g_{ki_k}}{\partial x^l} = \frac{\partial g_{km}}{\partial x^l}\delta^m{}_{i_k} = \frac{\partial g_{km}}{\partial x^l} g^{mr} g_{ri_k}. \tag{14.13}$$

Plugging (14.12) into the definition of $\Gamma^k{}_{kl}$, we find

$$\begin{aligned}
\Gamma^k{}_{kl} &= \frac{\partial g^{km}}{2}(g_{mk,l} + \underbrace{g_{ml,k} - g_{kl,m}}_{=0\ (m\leftrightarrow k)}) \\
&= \frac{g^{km}}{2}\frac{\partial g_{mk}}{\partial x^l} \\
&= \frac{\partial \ln\sqrt{g}}{\partial x^l} \\
&= \frac{1}{\sqrt{g}}\frac{\partial\sqrt{g}}{\partial x^l}. \qquad (14.14)
\end{aligned}$$

With this relation, one can show that the inhomogeneous Maxwell equation and the continuity equation in KS can be written as

$$\frac{1}{\sqrt{g}}\frac{\partial(\sqrt{g}F^{\mu\nu})}{\partial x^\nu} = \frac{4\pi}{c^2}j^\mu \qquad (14.15)$$

$$\frac{\partial(\sqrt{g}j^\mu)}{\partial x^\mu} = 0 \qquad (14.16)$$

14.0.7 The Energy-Momentum Tensor

For an ideal fluid, the energy-momentum tensor in a local IS is given by

$$T^{\mu\nu} = \left(\rho + \frac{p}{c^2}\right)u^\mu u^\nu - \eta^{\mu\nu}p \qquad (14.17)$$

where u^μ is the four-velocity, ρ the proper energy density and p the pressure of the fluid.
In a general coordinate system KS, the energy-momentum tensor reads

$$T^{\mu\nu} = \left(\rho + \frac{p}{c^2}\right)u^\mu u^\nu - g^{\mu\nu}p. \qquad (14.18)$$

In the IS the conservation law implies $T^{\mu\nu}{}_{,\nu} = 0$ which reads as follows in KS:

$$T^{\mu\nu}{}_{;\nu} = T^{\mu\nu}{}_{,\nu} + \Gamma^\mu{}_{\nu\lambda}T^{\nu\lambda} + \Gamma^\nu{}_{\nu\lambda}T^{\mu\lambda} = 0. \qquad (14.19)$$

With Eq. (14.14) we get instead

$$T^{\mu\nu}{}_{;\nu} = \frac{1}{\sqrt{g}}\frac{\partial(\sqrt{g}T^{\mu\nu})}{\partial x^\nu} + \Gamma^\mu{}_{\nu\lambda}T^{\nu\lambda} = 0. \qquad (14.20)$$

This is no longer a conservation law because we cannot form any constant of motion from (14.20). The physical reason is that the system under consideration can exchange energy and momentum with the gravitational field.

15 Einstein's Field Equations

The field equations cannot be derived by the covariance principle because there is no equivalent equation in a local IS. So we cannot apply the simple procedure

from the previous section. We have to make some assumptions: Newtonian gravity as a limiting case should be well confirmed through all observations: $\Delta\phi = 4\pi G\rho$.
From the Newtonian limit of the equations of motion for a particle, we derived (c.f. Eq. (4.27)): $g_{00} = 1 + \frac{2\phi}{c^2}$. So the non-relativistic limit should read

$$\Delta g_{00} = \frac{2}{c^2} 4\pi G\rho = \frac{8\pi G}{c^4} T_{00} \tag{15.1}$$

with $T_{00} \simeq \rho c^2$ (the other components T_{ij} are small). Then ρ transforms as the 00-component of a tensor which is necessary because of $\rho \simeq \frac{\Delta m}{\Delta V}$.[10]
We conclude that a generalization of (15.1) 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 (because $T_{\mu\nu}$ is a tensor, as well).

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

3. Since $T_{\mu\nu}$ is symmetric, $G_{\mu\nu}$ should also be symmetric. Furthermore, we expect ${G_{\mu\nu}}^{;\nu} = 0$ since $T_{\mu\nu}$ is covariantly conserved (${T^{\mu\nu}}_{;\nu} = 0$).

4. For a weak stationary field we should get Eq. (15.1) and thus $G_{00} \simeq \Delta g_{00}$.

It turns out that these four conditions suffice to determine $G_{\mu\nu}$ uniquely: The first two points imply that $G_{\mu\nu}$ has to be a linear combination of the Ricci tensor and the the Ricci scalar[11]:

$$G_{\mu\nu} = aR_{\mu\nu} + bRg_{\mu\nu}. \tag{15.2}$$

The symmetry of $G_{\mu\nu}$ is obviously satisfied. Using the contracted Bianchi identity (Eq. (13.33)) and the fact that ${g_{\mu\nu}}^{;\nu} = 0$, we infer that also ${G_{\mu\nu}}^{;\nu} = 0$ is satisfied if $b = -\frac{a}{2}$. Thus we find

$$G_{\mu\nu} = a\left(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R\right) = \frac{8\pi G}{c^2} T_{\mu\nu}. \tag{15.3}$$

The constant a can be determined by considering the Newtonian limit and demanding property 4. from above. In order to do this, we consider the weak field limit, i.e. $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ with $|h_{\mu\nu}| \ll 1$. In this limit, it follows

$$|T_{il}| \ll T_{00} \quad \text{and} \quad |G_{ik}| \ll |G_{00}|. \tag{15.4}$$

[10] Note that Δm is not a Lorentz scalar! The electromagnetic charge density $\rho_e \simeq \frac{\Delta q}{\Delta V}$ transforms as the 0-component of a 4-vector because charge is a Lorentz scalar.
[11] It can be shown that indeed the Ricci tensor is the only tensor made of the metric and first and second derivatives of it which is linear in the second derivative terms.

Taking the trace of (15.3), we find

$$g^{\mu\nu}G_{\mu\nu}\begin{cases} = a(R-2R) = -aR & \text{(from (15.3))} \\ \approx G_{00} = a\left(R_{00} - \dfrac{R}{2}g_{00}\right) = a\left(R_{00} - \dfrac{R}{2}\right) & \end{cases} \tag{15.5}$$

Comparing these two results, we infer that

$$R \simeq -2R_{00} \quad \Rightarrow \quad G_{00} \simeq a\left(R_{00} - \frac{R}{2}\right) \simeq 2aR_{00}. \tag{15.6}$$

For weak fields all terms quadratic in $h_{\mu\nu}$ can be neglected in the Riemann tensor. In leading order this implies

$$R_{\mu\nu} = R^{\rho}{}_{\mu\rho\nu} \simeq \frac{\partial \Gamma^{\rho}{}_{\mu\rho}}{\partial x^{\nu}} - \frac{\partial \Gamma^{\rho}{}_{\mu\nu}}{\partial x^{\rho}} \qquad (|h_{\mu\nu}| \ll 1). \tag{15.7}$$

For weak stationary fields we get

$$R_{00} = -\frac{\partial \Gamma^{i}{}_{00}}{\partial x^{i}} \qquad \text{with } \Gamma^{i}{}_{00} = \frac{1}{2}\frac{\partial g_{00}}{\partial x^{i}}. \tag{15.8}$$

(note that we use another sign convention as in Eq. (12.6). We use the Riemann curvature tensor with another sign now.) Therefore

$$G_{00} \simeq -2a\frac{\partial \Gamma^{i}{}_{00}}{\partial x^{i}} = -a\Delta g_{00} \stackrel{!}{=} \Delta g_{00} \tag{15.9}$$

which implies $a = -1$. This leads to the final **Einstein field equations** (Einstein, 1915)

$$\boxed{R_{\mu\nu} - \frac{R}{2}g_{\mu\nu} = -\frac{8\pi G}{c^4}T_{\mu\nu}.} \tag{15.10}$$

Together with the geodesic equation (11.22 or 14.5), these equations are the foundation of the theory of general relativity.
By means of contracting Eq. (15.10), we get also

$$R^{\mu}{}_{\mu} - \frac{R}{2}\underbrace{\delta^{\mu}{}_{\mu}}_{=4} = -R = -\frac{8\pi G}{c^4}T \tag{15.11}$$

where $T := T^{\mu}{}_{\mu}$. We can thus express R in terms of T and rewrite Einstein's field equations (15.10) in the form

$$\boxed{R_{\mu\nu} = -\frac{8\pi G}{c^4}\left(T_{\mu\nu} - \frac{T}{2}g_{\mu\nu}\right).} \tag{15.12}$$

Note that in vacuum (i.e. $T_{\mu\nu} = 0$) we have $R_{\mu\nu} = 0$.

The interesting thing in Eq. (15.10) is that on the left hand side, there appears a purely *geometrical* quantity describing the intrinsic geometrical properties (in particular curvature) of the manifold, whereas the object on the right hand side is purely "physical" in the sense that it descibes velocities and pressure. So Einstein's equations tell us that curvature and physical field content are essentially the same.

15.0.8 Significance of the Bianchi Identities

Einsteins's field equations constitute a set of non-linear coupled differential equations whose general solution is not known. Usually one makes some simplifying assumptions, for example spherical symmetry.
Because of the symmetry of the Ricci tensor, the Einstein field equations constitute a set of 10 algebraically independent second order differential equations for $g_{\mu\nu}$.
Due to the fact that the equations are generally covariant, the metric can at most be determined up to coordinate transformations. Therefore we expect only 6 independent generally covariant equations for the metric (4 degrees of freedom are canceled by the freedom to choose coordinates). Indeed the (contracted) Bianchi identities tell us that

$$G_{\mu}{}^{\nu}{}_{;\nu} = 0 \tag{15.13}$$

giving 4 more equations. So the Bianchi identities can also be understood as a consequence of general covariance and they ensure the conservation of $T_{\mu\nu}$ in curved spacetime.

15.0.9 The Cosmological Constant

As a generalization of the field equations one can relax the second condition and allow a term which is linear in $g_{\mu\nu}$ (note that $g_{\mu\nu}{}^{;\nu} = 0$). The field equations become

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} + \Lambda g_{\mu\nu} = -\frac{8\pi G}{c^4} T_{\mu\nu} \tag{15.14}$$

with the **cosmological constant** Λ ($[\Lambda] =$length^{-2}). The Newtonian limit of (15.14) reads

$$\Delta\phi = 4\pi\rho G + \frac{c^2}{2}\Lambda. \tag{15.15}$$

The right-hand side can also be written as

$$4\pi G(\rho + \rho_{\text{vacuum}}) \qquad \text{with} \qquad c^2 \rho_{\text{vacuum}} = \frac{c^4}{8\pi G}\Lambda. \tag{15.16}$$

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

16 The Einstein-Hilbert Action

The Einstein field equations (15.10) can be obtained from a covariant variation principle. We claim that the correct action for the metric g is given by

$$S_D[g] = \int_D R(g) dv \tag{16.1}$$

where $D \subset M$ is a compact region in the spacetime manifold, R is the scalar curvature and dv is the volume element

$$dv = \sqrt{|g|}d^4x. \tag{16.2}$$

We have to show that the Euler-Lagrange equations which are derived from this action lead to the Einstein field equations. The complete calculation is quite lengthy, so we only sketch the ideas. The Euler-Lagrange equations are the field equations in vacuum:

$$\begin{aligned} 0 = \delta S_D[g] &= \int_D \delta \left(g^{\mu\nu} R_{\mu\nu} \sqrt{-g}\right) d^4x \\ &= \int_D (\delta R_{\mu\nu})\, g^{\mu\nu} \sqrt{-g} d^4x + \int_D R_{\mu\nu} \delta \left(g^{\mu\nu} \sqrt{-g}\right) d^4x. \end{aligned} \tag{16.3}$$

We start with the first of these integrals. We will show that it vanishes. We start with $\delta R_{\mu\nu}$ where

$$R_{\mu\nu} = \partial_\alpha \Gamma^\alpha{}_{\mu\nu} - \partial_\nu \Gamma^\alpha{}_{\mu\alpha} + \Gamma^\rho{}_{\mu\nu}\Gamma^\alpha{}_{\rho\alpha} - \Gamma^\rho{}_{\mu\alpha}\Gamma^\alpha{}_{\rho\nu}. \tag{16.4}$$

We first compute the variation of $R_{\mu\nu}$ locally at any point p in normal coordinates centered at p (i.e. $p \hat{=} x = 0$ and $\Gamma^\alpha{}_{\beta\gamma}(0) = 0$)[12]:

$$\delta R_{\mu\nu} = \left(\delta \Gamma^\alpha{}_{\mu\nu}\right)_{,\alpha} - \left(\delta \Gamma^\alpha{}_{\mu\alpha}\right)_{,\nu} \tag{16.5}$$

(note that the variation with respect to g commutes with ordinary derivatives). One can show that $\delta\Gamma^\alpha{}_{\mu\nu}$ is indeed a tensor although $\Gamma^\alpha{}_{\mu\nu}$ is not. So (16.5) is a tensor equation which is valid in every coordinate system and we can also use covariant derivatives:

$$\delta R_{\mu\nu} = \left(\delta \Gamma^\alpha{}_{\mu\nu}\right)_{;\alpha} - \left(\delta \Gamma^\alpha{}_{\mu\alpha}\right)_{;\nu} \tag{16.6}$$

(so called Palatini identity). Using $g_{\mu\nu;\sigma} = 0$, we can write (16.6) as

$$\begin{aligned} g^{\mu\nu}\delta R_{\mu\nu} &= \left(g^{\mu\nu}\delta \Gamma^\alpha{}_{\mu\nu}\right)_{;\alpha} - \left(g^{\mu\nu}\delta \Gamma^\alpha{}_{\mu\alpha}\right)_{;\nu} \\ &= w^\alpha{}_{;\alpha} \\ &= w^\alpha{}_{,\alpha} + \Gamma^\alpha{}_{\alpha\mu} w^\mu \\ &= w^\alpha{}_{,\alpha} + \frac{1}{\sqrt{-g}} \frac{\partial \sqrt{-g}}{\partial x^\mu} w^\mu \\ &= \frac{1}{\sqrt{-g}} \frac{\partial (w^\mu \sqrt{-g})}{\partial x^\mu}. \end{aligned} \tag{16.7}$$

Inserting this into the integral (16.3) and remembering Gauss theorem (10.39),

$$\int_D (\mathrm{div}_g w)\, \eta = \int_{\partial D} i_w \eta \tag{16.8}$$

[12]It can be shown that one can always choose a local coordinate system on T_pM such that p corresponds to $x = 0$ and $g_{ij}(0) = \eta_{ij}$ and $\Gamma^k{}_{ij}(0) = 0$. Such coordinates are called *normal coordinates*.

with $\mathrm{div}_g w = w^\alpha{}_{;\alpha}$, we find

$$\int_D \delta R_{\mu\nu} g^{\mu\nu} \sqrt{-g} d^4x = \int_{\partial D} w^\alpha \sqrt{-g} do_\alpha \tag{16.9}$$

where do_α is the surface element for coordinates normal to ∂D and

$$w^\alpha = g^{\mu\nu} \delta\Gamma^\alpha{}_{\mu\nu} - g^{\mu\alpha} \delta\Gamma^\nu{}_{\mu\nu} \tag{16.10}$$

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

In order to compute the second term in (16.3), we recall from linear algebra that for any $n \times n$ matrix $A(\lambda)$, we have

$$\frac{d}{d\lambda} \det A = \det A \cdot \mathrm{tr}\left(A^{-1} \frac{dA}{d\lambda}\right) \tag{16.11}$$

$$\text{and} \qquad \frac{d}{d\lambda}(A^{-1}) \cdot A = -A^{-1} \frac{dA}{d\lambda}. \tag{16.12}$$

Thus

$$\delta g = g g^{\mu\nu} \delta g_{\nu\mu} \tag{16.13}$$

$$\text{and} \qquad (\delta g^{\mu\nu}) g_{\nu\sigma} = -g^{\mu\nu} \delta g_{\nu\sigma}. \tag{16.14}$$

Using these equations, we infer that

$$\begin{aligned} \delta\sqrt{-g} &= \frac{1}{2}\sqrt{-g} g^{\mu\nu} \delta g_{\nu\mu} = -\frac{1}{2}\sqrt{-g} g_{\alpha\beta} \delta g^{\alpha\beta} \\ \Rightarrow \quad \delta(g^{\mu\nu}\sqrt{-g}) &= \sqrt{-g}\delta g^{\mu\nu} - \frac{1}{2}\sqrt{-g} g^{\mu\nu} g_{\alpha\beta} \delta g^{\alpha\beta}. \end{aligned} \tag{16.15}$$

We can therefore compute the second term in (16.3):

$$\begin{aligned} 0 \overset{!}{=} \int_D R_{\mu\nu} \delta(g^{\mu\nu}\sqrt{-g}) d^4x &= \int_D \underbrace{\sqrt{-g} d^4x}_{=dv} \left(R_{\mu\nu} \delta g^{\mu\nu} - \frac{1}{2} \underbrace{R_{\mu\nu} g^{\mu\nu}}_{=R} g_{\alpha\beta} \delta g^{\alpha\beta} \right) \\ &= \int_D dv \left(R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right) \delta g^{\mu\nu} \end{aligned} \tag{16.16}$$

from which we conclude immediately the Einstein equations:

$$\delta S_D[g_{\mu\nu}] = 0 \quad \Rightarrow \quad G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 0. \tag{16.17}$$

This proves that variation of the action (16.1) gives the correct field equations.

Note that from

$$\delta \int_D \sqrt{-g} d^4x = \int_D \sqrt{-g} d^4x \frac{1}{2} g^{\mu\nu} \delta g_{\mu\nu} = \int_D \sqrt{-g} d^4x \frac{1}{2} g_{\mu\nu} \delta g^{\mu\nu} \tag{16.18}$$

it follows that if we have a non-vanishing cosmological constant, the Einstein equations in vacuum are obtained from the action principle applied to the action

$$S_D[g] = \int_D (R - 2\Lambda)\sqrt{-g}d^4x. \tag{16.19}$$

The variation principle extends to matter described by any field $\psi = (\psi_A)$ $(A = 1, ..., N)$ transforming as a tensor under change of coordinates (we include also the electromagnetic field among the ψ_A). Consider any action of the form

$$S_D[\psi] = \int_D \mathcal{L}(\psi, \nabla_g \psi)\sqrt{-g}d^4x \tag{16.20}$$

where ∇_g is the Riemannian connection of the metric g. If we know $\mathcal{L}$ in flat space, the equivalence principle tells us that we just need to replace $\eta_{\alpha\beta}$ by $g_{\alpha\beta}$ and replace ordinary derivatives by covariant derivatives in order to obtain a $\mathcal{L}$ in curved spacetime.
As an example, consider the electromagnetic field described by the Lagrangian

$$\mathcal{L} = -\frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu} = -\frac{1}{16\pi}F_{\mu\nu}F_{\sigma\rho}g^{\sigma\mu}g^{\rho\nu}. \tag{16.21}$$

In electrodynamics, we have the special feature that $F_{\mu\nu} = A_{\nu;\mu} - A_{\mu;\nu} = A_{\nu,\mu} - A_{\mu,\nu}$. With $\nabla_\mu A_\nu \equiv A_{\nu;\mu}$, the Euler-Lagrange equations read

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial A_\nu} - \nabla_\mu \frac{\partial \mathcal{L}}{\partial(\nabla_\mu A_\nu)} &= 0 \\ \Leftrightarrow \quad -\nabla_\mu\left(-\frac{1}{4}F_{\sigma\rho}g^{\mu\sigma}g^{\nu\sigma}\cdot 4\right) &= 0 \\ \Leftrightarrow \quad F^{\mu\nu}{}_{;\mu} &= 0. \end{aligned} \tag{16.22}$$

(these are just the inhomogeneous Maxwell equations for vanishing current j^ν. Remember also the general inhomogeneous Maxwell equations $F^{\mu\nu}{}_{;\mu} = \frac{1}{c}j^\nu$ which can easily be derived from the Lagrangian $\mathcal{L} = -\frac{1}{16\pi}F_{\mu\nu}F^{\mu\nu} - \frac{1}{c}j^\mu A_\mu$.

Variations in (16.20) with respect to the fields ψ_A lead to the Euler-Lagrange equations, whereas variations with respect to the metric give (without proof)

$$\delta_g \int_D \mathcal{L}(\psi, \nabla_g \psi)\sqrt{-g}d^4x = -\frac{1}{2}\int_D T^{\mu\nu}\delta g_{\mu\nu}\sqrt{-g}d^4x. \tag{16.23}$$

This term has to be added to the term which is proportional to $\delta g_{\mu\nu}$ in Einstein's action:

$$\begin{aligned} 0 &= \int_D \sqrt{-g}d^4x \underbrace{\left(G_{\mu\nu}\frac{c^4}{16\pi G} - \frac{1}{2}T_{\mu\nu}\right)}_{\stackrel{!}{=}0}\delta g^{\mu\nu} \\ \Rightarrow \qquad G_{\mu\nu} &= -\frac{8\pi G}{c^4}T_{\mu\nu}. \end{aligned} \tag{16.24}$$

In case of electrodynamics we have

$$T_{\mu\nu} = \frac{1}{4\pi}\left(F_{\mu\sigma}F_\nu{}^\sigma - \frac{1}{4}F_{\sigma\rho}F^{\sigma\rho}g_{\mu\nu}\right). \tag{16.25}$$

17 Static Isotropic Metric

17.1 The Form of the Metric

For the gravity field of the Earth or the Sun, we assume a spherically symmetric distribution of matter (rotation velocity $v^i \ll c$). Our goal is to find a spherically symmetric and static solution $g_{\mu\nu}(x)$ to the Einstein field equations.
For $r \to \infty$, we expect to recover the Newtonian gravitational potential $\phi = -\frac{GM}{r} \to 0$, i.e. in this limit, the metric becomes just the Minkowski metric $\eta_{\mu\nu}$:

$$ds^2 = c^2dt^2 - dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2) \quad \text{as } r \to \infty. \tag{17.1}$$

This motivates the following ansatz:

$$ds^2 = B(r)c^2dt^2 - A(r)dr^2 - C(r)r^2(d\theta^2 + \sin^2\theta d\varphi^2) \tag{17.2}$$

where A, B, C cannot depend on θ, φ or t due to spherical symmetry. Due to isotropy there cannot be any terms which are linear in $d\theta$ or $d\varphi$. The freedom in the choice of coordinates allows us to introduce a new radial coordinate in (17.2):

$$C(r) \;\longrightarrow\; r^2. \tag{17.3}$$

So we can absorb $C(r)$ into r. We get the standard form

$$\begin{aligned} ds^2 &= B(r)c^2dt^2 - A(r)dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2) \\ &\qquad\text{with } B(r),\ A(r) \to 1 \text{ as } r \to \infty. \end{aligned} \tag{17.4}$$

17.1.1 Robertson Expansion

Although we don't know the solution of 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 (e.g. the Sun), on the distance to that object 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 have the following expansion[13]:

$$B(r) = 1 - 2\frac{GM}{c^2r} + 2(\beta - \gamma)\left(\frac{GM}{c^2r}\right)^2 + \ldots \tag{17.5}$$

$$A(r) = 1 + 2\gamma\frac{GM}{c^2r} + \ldots \tag{17.6}$$

(**Robertson Expansion**).
The field equations contain the Newtonian limit[14]

$$g_{00} \simeq 1 + \frac{2\phi}{c^2} \quad \text{with } \phi = -\frac{GM}{r}. \tag{17.7}$$

[13] Note that $\frac{v^2}{c^2} \sim \frac{a}{r}$; terms like $g_{00}u^0u^0 \sim Bc^2$ and $g_{11}u^1u^1 \sim Av^2 \sim Ac^2\frac{a}{r}$ show up and have to be expanded to the same order in $\frac{a}{r}$. Thus B has to be expanded to one order higher than A.

[14] The factor $2(\beta - \gamma) \equiv G_{\text{eff.}}$ is for historical reasons. The parameters β and γ are independent.

In the solar system, for instance, we have $\frac{GM}{c^2 r} \leq \frac{GM}{c^2 R_\odot} \simeq 2 \cdot 10^{-6}$. So we can forget about terms which are of higher than linear order in β and γ.
In general relativity we have $\gamma = \beta = 1$ (in Newtonian gravity: $\beta = \gamma = 0$).

We calculate the Christoffel symbols for the standard form as given in Eq. (17.4). First of all, $g_{\mu\nu}$ is diagonal:

$$g_{00} = B(r), \quad g_{11} = -A(r), \quad g_{22} = -r^2, \quad g_{33} = -r^2 \sin^2\theta \tag{17.8}$$

$$g^{\mu\mu} = \frac{1}{g_{\mu\mu}}. \tag{17.9}$$

The non-vanishing Christoffel symbols are

$$\begin{aligned}
&\Gamma^0{}_{01} = \Gamma^0{}_{10} = \frac{B'(r)}{2B(r)} \\
&\Gamma^1{}_{00} = \frac{B'}{2A}, \quad \Gamma^1{}_{11} = \frac{A'}{2A}, \quad \Gamma^1{}_{22} = -\frac{r}{A}, \quad \Gamma^1{}_{33} = -\frac{r^2 \sin^2\theta}{A} \\
&\Gamma^2{}_{12} = \Gamma^2{}_{21} = \frac{1}{r}, \quad \Gamma^2{}_{33} = -\sin\theta\cos\theta \\
&\Gamma^3{}_{31} = \Gamma^3{}_{13} = \frac{1}{r}, \quad \Gamma^3{}_{23} = \Gamma^3{}_{32} = \mathrm{ctg}\theta.
\end{aligned} \tag{17.10}$$

Writing

$$-g = r^4 AB \sin^2\theta, \tag{17.11}$$

we find

$$\Gamma^\rho{}_{\mu\rho} = \left(\frac{\partial \ln\sqrt{-g}}{\partial x^\mu}\right) = \left(0, \frac{2}{r} + \frac{A'}{2A} + \frac{B'}{2B}, \mathrm{ctg}\theta, 0\right) \tag{17.12}$$

and thus the Ricci tensor

$$R_{\mu\nu} = \frac{\partial \Gamma^\rho{}_{\mu\rho}}{\partial x^\nu} - \frac{\partial \Gamma^\rho{}_{\mu\nu}}{\partial x^\rho} + \Gamma^\sigma{}_{\mu\rho}\Gamma^\rho{}_{\sigma\nu} - \Gamma^\sigma{}_{\mu\nu}\Gamma^\rho{}_{\sigma\rho} \tag{17.13}$$

has the following non-vanishing components:

$$\begin{aligned}
R_{00} &= -\frac{B''}{2A} + \frac{A'B'}{2A^2} + \frac{B'^2}{2AB} - \frac{B'}{2A}\left(\frac{2}{r} + \frac{A'}{2A} + \frac{B'}{2B}\right) \\
&= -\frac{B''}{2A} + \frac{B'}{4A}\left(\frac{A'}{A} + \frac{B'}{B}\right) - \frac{B'}{rA},
\end{aligned} \tag{17.14}$$

$$R_{11} = \frac{B''}{2B} - \frac{B'}{4B}\left(\frac{A'}{A} + \frac{B'}{B}\right) - \frac{A'}{rA} \tag{17.15}$$

$$R_{22} = -1 - \frac{r}{2A}\left(\frac{A'}{A} - \frac{B'}{B}\right) + \frac{1}{A} \tag{17.16}$$

$$R_{33} = R_{22} \cdot \sin^2\theta. \tag{17.17}$$

17.2 The Schwarzschild Metric

We assume a static, spherically symmetric mass distribution of finite extension:

$$\rho(r) \quad \begin{cases} \neq 0 & r \leq r_0 \\ = 0 & r > r_0 \end{cases} \tag{17.18}$$

and similarly for the pressure $p(r) \neq 0$ for $r \leq r_0$. In the static case, the four-velocity vector within the mass distribution is $u^\mu = (u^0 = \text{const.},\ 0,\ 0,\ 0)$. So the energy-momentum tensor which describes the properties of the matter does not depend on time. We adopt the following ansatz for $g^{\mu\nu}$ (cf. Eq. (17.4)):

$$g_{\mu\nu} = \begin{pmatrix} B(r) & & & \\ & -A(r) & & \\ & & -r^2 & \\ & & & -r^2\sin^2\theta \end{pmatrix}. \tag{17.19}$$

Outside the mass distribution ($r \geq r_0$) we need to have $R_{\mu\nu} = 0$ (this is because of $p = \rho = 0$ for $r \geq r_0$ which always implies $R_{\mu\nu} = 0$ as we remember from Eq. (15.12)). On the other hand, the components $R_{\mu\nu}$ are given by Eqs. (17.14)-(17.17). For $\mu \neq \nu$, the condition $R_{\mu\nu} = 0$ is trivially satisfied. For $r \geq r_0$, we demand $R_{\mu\mu} = 0$. It obviously holds true that

$$\begin{aligned} 0 = \frac{R_{00}}{B} + \frac{R_{11}}{A} &= -\frac{1}{rA}\left(\frac{B'}{B} + \frac{A'}{A}\right) \\ \Rightarrow \qquad \frac{d}{dr}(\ln(AB)) &= 0. \end{aligned} \tag{17.20}$$

Accordingly, $AB = \text{const.}$. Since the metric should be Minkowski in the limit $r \to \infty$, we require $A(r) = A(r \to \infty) = 1 = B(r \to \infty) = B(r)$. Plugging this into into R_{22} from Eq. (17.16) and into R_{11} from Eq. (17.15), we see that

$$R_{22} = -1 + rB' + B = 0 \tag{17.21}$$

$$R_{11} = \frac{B''}{2B} + \frac{B'}{rB} = \frac{rB'' + 2B'}{2rB} = \frac{1}{2rB}\frac{dR_{22}}{dr} \overset{(17.21)}{=} 0. \tag{17.22}$$

We write (17.21) as

$$\frac{d}{dr}(rB) = 1 \tag{17.23}$$

and integrate this equation:

$$rB = r + \text{const.} \equiv r - 2a. \tag{17.24}$$

This implies

$$B(r) = 1 - \frac{2a}{r} \qquad \text{and} \qquad A(r) = \frac{1}{1 - \frac{2a}{r}} \qquad \text{for } r \geq r_0. \tag{17.25}$$

This solution of the Einstein equations in vacuum was found in 1916 by Schwarzschild. The **Schwarzschild solution** is[15]

$$\boxed{ds^2 = \left(1 - \frac{2a}{r}\right) c^2 dt^2 - \frac{1}{\left(1 - \frac{2a}{r}\right)} dr^2 - r^2 \left(d\theta^2 + \sin^2 d\varphi^2\right).} \tag{17.26}$$

The constant a can be determined by considering the Newtonian limit

$$g_{00} = B(r) \overset{r\to\infty}{\longrightarrow} 1 + \frac{2\phi}{c^2} = 1 - \frac{2GM}{c^2 r} = 1 - \frac{2a}{r} \tag{17.27}$$

We introduce the so called **Schwarzschild radius**

$$r_S = 2a = \frac{2GM}{c^2}. \tag{17.28}$$

For the Sun, $r_{S,\odot} \simeq 3$ km. The significance of the Schwarzschild radius is that the Schwarzschild metric (17.26) seems to be singular. But this is actually not the case. As we will see later, this singularity is only an artefact of the choice of coordinates: r is not a good coordinate in the region where $r \leq r_S$.
A clock which rests at radius r has proper time $d\tau = \sqrt{B}dt$, so $\frac{dt}{d\tau}$ is divergent as $r \to r_S$. This implies that a photon emitted at $r = r_S$ will be infinitely redshifted.
A star with radius $r_{\text{star}} < r_S$ is a black hole since photons which are emitted at its surface cannot reach regions with $r > r_S$.

Expanding the Schwarzschild metric in powers of $\frac{r}{r_S}$ and comparing it to the Robertson expantion (17.6), one finds $\beta = \gamma = 1$ for GR.

18 General Equations of Motion

We now consider the motion of a freely falling material particle or a photon in a static, isotropic gravitational field (for example, the motion of a planet around the Sun). This is slightly different from the Kepler problem because the Kepler problem can be solved exactly whereas now we neglect the influence of the test particle on the gravitational field.
For the relativistic orbit $x^k(\lambda)$ of a particle in a gravitational field, we have the geodesic equation

$$\frac{d^2 x^k}{d\lambda^2} = -\Gamma^k{}_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \tag{18.1}$$

where the Christoffel symbols can be computed from (17.26). Furthermore, we have

$$g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = \left(\frac{ds}{d\lambda}\right)^2 = c^2 \left(\frac{d\tau}{d\lambda}\right)^2 = \begin{cases} c^2 & (m \neq 0,\ \lambda = t) \\ 0 & (m = 0) \end{cases} \tag{18.2}$$

[15] The fundamental reason why we are able to find this solution so easily, is the Bianchi identities!

For a massive particle we can use the proper time to parametrize the trajectory (or orbit): $d\lambda = d\tau$. In case of massless particles one has to choose another parameter λ.
For the spherically symmetric gravitational field of the Sun, we can use the Schwarzschild metric for $r \geq r_\odot$:

$$ds^2 = B(r)c^2dt^2 - A(r)dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2). \tag{18.3}$$

The equations (18.1)-(18.3) describe the relativistic Kepler problem. Using the Christoffel symbols as given in Eq. (17.10), we get for (18.1)

$$\frac{d^2x^0}{d\lambda^2} = -\frac{B'}{B}\frac{dx^0}{d\lambda}\frac{dr}{d\lambda} \tag{18.4}$$

$$\frac{d^2r}{d\lambda^2} = -\frac{B'}{2A}\left(\frac{dx^0}{d\lambda}\right)^2 - \frac{A'}{2A}\left(\frac{dr}{d\lambda}\right)^2 + \frac{r}{A}\left(\frac{d\theta}{d\lambda}\right)^2 + \frac{r\sin^2\theta}{A}\left(\frac{d\varphi}{d\lambda}\right)^2 \tag{18.5}$$

$$\frac{d^2\theta}{d\lambda^2} = -\frac{2}{r}\frac{d\theta}{d\lambda}\frac{dr}{d\lambda} + \sin\theta\cos\theta\left(\frac{d\varphi}{d\lambda}\right)^2 \tag{18.6}$$

$$\frac{d^2\varphi}{d\lambda^2} = -\frac{2}{r}\frac{d\varphi}{d\lambda}\frac{dr}{d\lambda} - 2\mathrm{ctg}\theta\frac{d\theta}{d\lambda}\frac{d\varphi}{d\lambda}. \tag{18.7}$$

Eq. (18.6) can be solved by $\theta = \frac{\pi}{2} = $ const.. Without loss of generality we can always choose the coordinate system such that $\theta = \frac{\pi}{2}$ is satisfied (i.e. the trajectory lies on the plane with $\theta = \frac{\pi}{2}$. The fact that $\frac{d^2\theta}{d\lambda^2} = 0$ corresponds to angular momentum conservation.
With $\theta = \frac{\pi}{2}$, Eq. (18.7) reads

$$\frac{1}{r^2}\frac{d}{d\lambda}\left(r^2\frac{d\varphi}{d\lambda}\right) = 0. \tag{18.8}$$

It follows immediately that

$$r^2\frac{d\varphi}{d\lambda} = \text{const.} = l. \tag{18.9}$$

The constant l can be interpreted as the (orbital) angular momentum per mass (this is, of course, motivated by the classical formula for the angular momentum: $L = mr^2\dot{\phi}$). It has units of $\frac{\text{length}^2}{\text{time}}$. Eq. (18.8) and $\theta = \frac{\pi}{2}$ follow from angular momentum conservation which is a consequence of the spherical symmetry (rotational invariance).
Eq. (18.4) can be written as

$$\frac{d}{d\lambda}\left(\log\left(\frac{1}{c}\frac{dx^0}{d\lambda}\right) + \log B\right) = 0 \tag{18.10}$$

which can easily be integrated:

$$\log\left(\left(\frac{dx^0}{d\lambda}\right)\cdot\frac{B}{c}\right) = \text{const.'} \quad \Rightarrow \quad B\frac{dx^0}{d\lambda} = \text{const.} = F. \tag{18.11}$$

In Eq. (18.5) we use $\theta = \frac{\pi}{2}$ and Eqs. (18.9) and (18.11) and get

$$\frac{d^2r}{d\lambda^2} + \frac{F^2B'}{2AB^2} + \frac{A'}{2A}\left(\frac{dr}{d\lambda}\right)^2 - \frac{l^2}{Ar^3} = 0. \tag{18.12}$$

If we multiply this with $2A\frac{dr}{d\lambda}$, we obtain

$$\frac{d}{d\lambda}\left(A\left(\frac{dr}{d\lambda}\right)^2 + \frac{l^2}{r^2} - \frac{F^2}{B}\right) = 0. \tag{18.13}$$

Again, integration yields

$$A\left(\frac{dr}{d\lambda}\right)^2 + \frac{l^2}{r^2} - \frac{F^2}{B} = \text{const.} = -\varepsilon. \tag{18.14}$$

The next steps have to be done numerically except in special simple cases. In principle, one proceeds as follows: We can integrate this equation once more in order to get $r = r(\lambda)$. Inserting this result into (18.9) and (18.11), one can obtain $\varphi = \varphi(\lambda)$ and $t = t(\lambda)$. Next, one eliminates λ and finds $r = r(t)$ and $\varphi = \varphi(t)$. Together with $\theta = \frac{\pi}{2}$ this is then a complete solution.
Using $\theta = \frac{\pi}{2}$, (18.9), (18.11) and (18.14), one can see that Eq. (18.2) takes a simple form:

$$\begin{aligned} g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} &= B\left(\frac{dx^0}{d\lambda}\right)^2 - A\left(\frac{dr}{d\lambda}\right)^2 - r^2\left(\frac{d\theta}{d\lambda}\right)^2 - r^2\sin^2\theta\left(\frac{d\varphi}{d\lambda}\right)^2 \\ &= \varepsilon. \end{aligned} \tag{18.15}$$

Because the result has to coincide with Eq. (18.2), we infer

$$\varepsilon = \begin{cases} c^2 & (m \neq 0) \\ 0 & (m = 0) \end{cases} \tag{18.16}$$

so we are left with the two constants of integration F and l.

We want to examine the trajectory closer. From (18.14) we get

$$\frac{dr}{d\lambda} = \sqrt{\frac{\frac{F^2}{B} - \frac{l^2}{r^2} - \varepsilon}{A}}. \tag{18.17}$$

On the other hand, Eq. (18.9) yields

$$\frac{d\varphi}{dr} = \frac{d\varphi}{d\lambda}\frac{d\lambda}{dr} = \frac{l}{r^2}\sqrt{\frac{A}{\frac{F^2}{B} - \frac{l^2}{r^2} - \varepsilon}} \tag{18.18}$$

which implies

$$\boxed{\varphi(r) = \int \frac{dr}{r^2}\frac{\sqrt{A(r)}}{\sqrt{\frac{F^2}{B(r)l^2} - \frac{1}{r^2} - \frac{\varepsilon}{l^2}}}.} \tag{18.19}$$

This confines the trajectory $\varphi = \varphi(r)$ to the orbital plane (characterized by $\theta = \frac{\pi}{2}$). For massive particles, we have two integration constants $\frac{F^2}{l^2}$ and $\frac{\varepsilon}{l^2}$. For massless particles, ε vanishes, so there is actually only one integration constant, namely $\frac{F^2}{l^2}$.

We can specialze the trajectory to the case of the Schwarzschild metric, i.e.

$$B(r) = \frac{1}{A(r)} = 1 - \frac{r_S}{r} = 1 - \frac{2a}{r} \tag{18.20}$$

and write

$$\dot{t} = \frac{dt}{d\lambda}, \quad \dot{r} = \frac{dr}{d\lambda}, \quad \dot{\varphi} = \frac{d\varphi}{d\lambda}. \tag{18.21}$$

Using Eqs. (18.9) and (18.11) and $\theta = \frac{\pi}{2}$, we find

$$c\dot{t}\left(1 - \frac{2a}{r}\right) = F \qquad \text{and} \qquad r^2\dot{\varphi} = l. \tag{18.22}$$

Inserting $B(r) = 1 - \frac{2a}{r}$, the radial equation (18.11) takes the form

$$\frac{\dot{r}^2}{2} - \frac{a\varepsilon}{r} + \frac{l^2}{2r^2} - \frac{al^2}{r^3} = \frac{F^2 - \varepsilon}{2} = \text{const.} \tag{18.23}$$

which can be written as

$$\frac{\dot{r}^2}{2} + V_{\text{eff}}(r) = \text{const.} = \frac{F^2 - \varepsilon}{2} \tag{18.24}$$

with the **effective potential**

$$\boxed{V_{\text{eff}}(r) = \begin{cases} -\frac{GM}{r} + \frac{l^2}{2r^2} - \frac{GMl^2}{c^2r^3} & (m \neq 0) \\ \frac{l^2}{2r^2} - \frac{GMl^2}{c^2r^3} & (m = 0) \end{cases}} \tag{18.25}$$

This effective potential puts us in the situation that we are able to describe the particle's motion with general relativistic effects in the same way as we would describe a particle in flat spacetime moving in the effective potential (18.25) which differs from the corresponding Newtonian effective potential by the term $\propto \frac{1}{r^3}$.

A formal solution $r = r(\lambda)$ of (18.24) is given implicitly by the following integral:

$$\lambda = \pm \int \frac{dr}{\sqrt{2(\text{const.} - V_{\text{eff}}(r))}}. \tag{18.26}$$

Due to the $\frac{1}{r^3}$ term (which is the relativistic correction), this is an elliptical integral which has to be solved numerically. This term makes that the behaviour of the classical effective potential changes for small r:

- **If** $m \neq 0$:
For small radii, the effective potential does not go to $+\infty$ as in the classical case but to $-\infty$ instead: the effective potential becomes *attractive* for very small r which is in contrast to the classical Kepler problem. The attractive relativistic term becomes (for small r)

$$-\frac{GM}{r}\frac{l^2}{c^2r^2} \simeq -\frac{GM}{r}\frac{v^2}{c^2}. \tag{18.27}$$

Eq. (18.24) differs from the non-relativistic case not only due to the term which is proportional to $\frac{1}{r^3}$ but also due to the fact that $\dot{r} = \frac{dr}{d\tau}$ differs from $\frac{dr}{dt}$ by terms of order $\frac{v^2}{c^2}$.
Note that where V_{eff} has a minimum, there are bound solutions. However, there will be small deviations from the elliptical orbits due to relativistic corrections (precession of the perihelion). As a special case, the circular orbit is a possible solution because it is characterized by $\dot{r} = 0$.
Note that if the constant is positive, one gets unbound trajectories corresponding to the hyperbolic solutions in the non-relativistic case. If the constant is bigger than the maximum value of the potential, then the particle falls into the center.
The effective potential has a stable minimum and an unstable maximum where $\frac{dV_{\text{eff}}}{dr} = 0$. For $m \neq 0$ this is equivalent to the condition

$$\frac{c^2}{l^2}r^2 - 2\frac{r}{r_S} + 3 = 0. \tag{18.28}$$

In order for this equation to have two real solutions (a minimum and a maximum), we need to have

$$\frac{3c^2}{l^2} < \frac{1}{r_S^2} \tag{18.29}$$

$$\Leftrightarrow l \geq l_{\text{crit}} = \sqrt{3}r_S c. \tag{18.30}$$

For $l \to l_{\text{crit.}}$, the angular momentum barrier gets smaller and smaller until the minimum and the maximum coincide. If $l < l_{\text{crit.}}$, the effective potential decreases monotonically as $r \to 0$.

- **If** $m = 0$**:**
Both terms in (18.25) are proportional to l^2, so the shape of $V_{\text{eff.}}$ does not depend on l^2. The potential has a maximum at $r_{\text{max.}} = \frac{3}{2}r_S$ where photons could in principle move on a circular orbit. This orbit is unstable, of course.
If the constant in (18.24) is smaller than $V_{\text{eff.}}(r_{\text{max.}})$, then the incoming photon will be scattered, whereas it will be absorbed if the constant is bigger than $r_{\text{max.}}$. Note that for $r \leq r_S$, the Schwarzschild solution is no longer applicable.

19 Deflection of Light

The trajectory $r = r(\varphi)$ in the gravitational field is given by (18.19) with $\varepsilon = 0$:

$$\varphi(r) = \varphi(r_0) + \int_{r_0}^{r} \frac{dr'}{r'^2} \frac{\sqrt{A(r')}}{\sqrt{\frac{F^2}{B(r')l^2} - \frac{1}{r'^2}}} \tag{19.1}$$

where r_0 is the impact parameter (smallest distance to the gravitating mass). We define φ in such a way that $\varphi(r = r_0) = 0$. Sufficiently far away from the mass, the light ray is straight in a Minkowski metric. The angle between the incoming and outgoing asymptotic lightrays is $\Delta\varphi$. In total, the light is deflected by this angle $\Delta\varphi$. Going from r_0 to r_∞, the angle changes by $\varphi(\infty)$. Therefore, going on the light ray from $r_{-\infty}$ to r_∞, the radial vector rotates by $2\varphi(\infty)$ (if the light ray was a straight line, then $2\varphi(\infty) = \pi$). In general we have

$$\Delta\varphi = 2\varphi(\infty) - \pi. \tag{19.2}$$

Because of $r(\varphi)$ being minimal at $r_0 = r(0)$, we have

$$\left.\frac{dr}{d\varphi}\right|_{r_0} = 0 \tag{19.3}$$

which implies (using (18.17) and (18.18))

$$\frac{F^2}{l^2} = \frac{B(r_0)}{r_0^2}. \tag{19.4}$$

This allows us to eliminate the constants F and l in terms of r_0:

$$\phi(\infty) = \int_{r_0}^{\infty} \frac{dr}{r} \frac{\sqrt{A(r)}}{\sqrt{\frac{B(r_0)}{B(r)}\frac{r^2}{r_0^2} - 1}}. \tag{19.5}$$

To compute tis integral we insert 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 only keep terms of order $\mathcal{O}\left(\frac{a}{r}\right)$. This approximation is applicable if $r_0 \ll r_S$. With

$$\begin{aligned} \frac{r^2}{r_0^2}\frac{B(r_0)}{B(r)} - 1 &\simeq \frac{r^2}{r_0^2}\left[1 + 2a\left(\frac{1}{r} - \frac{1}{r_0}\right)\right] - 1 \\ &= \left(\frac{r^2}{r_0^2} - 1\right)\left(1 - \frac{2ar}{r_0(r + r_0)}\right) \end{aligned} \tag{19.6}$$

we find (using $\sqrt{1+x} \simeq 1 + \frac{x}{2}$)

$$\begin{aligned} \varphi(\infty) &\simeq \int_{r_0}^{\infty} \frac{dr}{\sqrt{r^2 - r_0^2}} \frac{r_0}{r}\left(1 + \gamma\frac{a}{r} + \frac{ar}{r_0(r + r_0)}\right) \\ &= \left[\cos^{-1}\left(\frac{r_0}{r}\right) + \gamma\frac{a}{r_0}\frac{\sqrt{r^2 - r_0^2}}{r} + \frac{a}{r_0}\sqrt{\frac{r - r_0}{r + r_0}}\right]\Bigg|_{r_0}^{\infty} \end{aligned} \tag{19.7}$$

$$\varphi(\infty) = \frac{\pi}{2} + \gamma\frac{a}{r_0} + \frac{a}{r_0}. \tag{19.8}$$

With (19.2) this yields

$$\begin{aligned}\Delta\varphi &= \frac{4a}{r_0}\left(\frac{1+\gamma}{2}\right) \\ &= \frac{2r_S}{r_0}\left(\frac{1+\gamma}{2}\right).\end{aligned} \tag{19.9}$$

In the case of GR, $\gamma = 1$, so

$$\Delta\varphi = \frac{2r_S}{r_0}. \tag{19.10}$$

For a light ray which just grazes the surface of the Sun ($r_0 = R_\odot \simeq 7\cdot 10^5$ km), we find

$$\Delta\varphi_\odot^{\text{max.}} = 1''.75\left(\frac{1+\gamma}{2}\right) \tag{19.11}$$

($\pi = 180 \times 3600''$).
Note that a similar calculation can be done for massive particles with a Newtonian potential. One finds that the formula for $\Delta\varphi$ does not depend on the mass of the particle. So one could argue that light has an infinitesimal mass μ which is set to zero at the end of the calculation and draw the conclusion that light is deflected also in Newtonian gravity. But note that the angle found in this way is too small by a factor of 2 (so it coincides with Eq. (19.11) for $\gamma = 0$).
Based on this idea, Einstein initiated real measurements of $\Delta\varphi$ (before he discovered GR). These measurements had to be done during a solar eclipse. The first attempt had to be canceled due to World War I. The experiment could not be done before 1919 when Einstein had already found the general relativistic equation.
Today, gravitational deflection of light is confirmed by the observation of gravitational lensing. Since $\Delta\varphi \propto r_S \propto M$, one can infer the mass of the gravitating object (e.g. galaxy clusters). It turns out that the observed mass is far smaller than the mass necessary to account for the lensing effect. This fact provides clear evidence for dark matter.
Gravitational lensing can also be used to detect stars and planets in our galaxy and nearby (microlensing): light that is deflected by stars or planets is magnified. So differences in the magnitude of background light is an evidence for masses between the light source and the observer.

20 Perihelion Precession

Even before the confirmation of GR by means of the observation of light deflection, Einstein could explain the longstanding problem of the precession of Mercury's perihelion.

Consider the elliptical orbit of a planet around the sun. We denote by $r_- = r_{\text{min}}$ the minimal distance and by $r_+ = r_{\text{max}}$ the maximum distance to the Sun and use the notation $X_\pm = X(r = r_\pm)$ for $X = A, B, \varphi, \ldots$.

The relativistic limit follows from Eq. (18.19) for $r = r(\varphi)$ with $\varepsilon = c^2$. The integral gives for the change of the angle between r_- and r_+

$$\begin{aligned}\varphi_+ - \varphi_- &= \int_{r_-}^{r_+} \frac{dr}{r^2} \frac{\sqrt{A(r)}}{\sqrt{\frac{F^2}{B(r)l^2} - \frac{1}{r^2} - \frac{c^2}{l^2}}} \\ &\equiv \int_{r_-}^{r_+} \frac{dr}{r^2} \frac{\sqrt{A(r)}}{\sqrt{K(r)}}. \end{aligned} \tag{20.1}$$

This corresponds to half of an orbit. For the shift due to a full orbit we would have to take twice the integral (20.1). The shift of the perihelion per full orbit is given by

$$\Delta\varphi = 2(\varphi_+ - \varphi_-) - 2\pi. \tag{20.2}$$

The integrand in (20.1) is equal to $\frac{d\varphi}{dr}$. At $r = r_\pm$ we have $\frac{d\varphi}{dr} = \infty$, so $r^2\sqrt{K(r)}$ has to vanish, i.e.

$$K(r_\pm) = 0 \tag{20.3}$$

$$\Rightarrow \quad \frac{F^2}{l^2 B_\pm} = \frac{1}{r_\pm^2} + \frac{c^2}{l^2}. \tag{20.4}$$

This allows us to express F and l in terms of $r_\pm$:

$$\begin{aligned}\frac{F^2}{l^2} &= \frac{\frac{1}{r_+^2} - \frac{1}{r_-^2}}{\frac{1}{B_+} - \frac{1}{B_-}} \\ &= \frac{r_-^2 - r_+^2}{r_+^2 r_-^2 \left(\frac{1}{B_+} - \frac{1}{B_-}\right)}, \end{aligned} \tag{20.5}$$

$$\begin{aligned}\frac{c^2}{l^2} &= \frac{\frac{B_+}{r_+^2} - \frac{B_-}{r_-^2}}{B_+ - B_-} \\ &= \frac{\frac{r_+^2}{B_+} - \frac{r_-^2}{B_-}}{r_+^2 r_-^2 \left(\frac{1}{B_+} - \frac{1}{B_-}\right)}. \end{aligned} \tag{20.6}$$

This way we get for $K(r)$

$$K(r) = \frac{r_-^2\left(\frac{1}{B(r)} - \frac{1}{B_-}\right) - r_+^2\left(\frac{1}{B(r)} - \frac{1}{B_+}\right)}{r_+^2 r_-^2 \left(\frac{1}{B_+} - \frac{1}{B_-}\right)} - \frac{1}{r^2}. \tag{20.7}$$

We insert the Robertson expansion for A and B:

$$A(r) = 1 + \gamma\frac{2a}{r} + \ldots \tag{20.8}$$

$$B(r) = 1 - \frac{2a}{r} + 2(\beta - \gamma)\frac{a^2}{r^2} + \ldots \tag{20.9}$$

$$\frac{1}{B(r)} = 1 + \frac{2a}{r} + 2(2 - \beta + \gamma)\frac{a^2}{r^2} + \ldots \tag{20.10}$$

With (20.10), $K(r)$ becomes a quadratic form in $\frac{1}{r}$. Since $\frac{d\phi}{dr} = \infty$, we have $K_+ = K_- = 0$. This determines $K(r)$ up to a constant $\widetilde{c}$:

$$K(r) = \widetilde{c}\left(\frac{1}{r_-} - \frac{1}{r}\right)\left(\frac{1}{r} - \frac{1}{r_+}\right). \tag{20.11}$$

The constant $\widetilde{c}$ can be determined by comparing with (20.7) for $r \to \infty$. Using Eq. (20.10), one finds

$$\widetilde{c} = 1 - (2 - \beta + \gamma)\left(\frac{a}{r_+} + \frac{a}{r_-}\right). \tag{20.12}$$

We are left with the following integral:

$$\varphi_+ - \varphi_- = \frac{1}{\sqrt{\widetilde{c}}}\int_{r_-}^{r_+} \frac{dr}{r^2}\left(1 + \gamma\frac{a}{r}\right)\left[\left(\frac{1}{r_-} - \frac{1}{r}\right)\left(\frac{1}{r} - \frac{1}{r_+}\right)\right]^{-1/2} \tag{20.13}$$

where the factor $\left(1 + \gamma\frac{a}{r}\right)$ comes from the expansion $\sqrt{A} \simeq 1 + \gamma\frac{a}{r}$. In order to evaluate this integral, we perform the following substitution:

$$\frac{1}{r} = \frac{1}{2}\left(\frac{1}{r_+} + \frac{1}{r_-}\right) + \frac{1}{2}\left(\frac{1}{r_+} - \frac{1}{r_-}\right)\sin\psi \tag{20.14}$$

where r_+ and r_- correspond to $\psi = \pm\frac{\pi}{2}$, respectively. With

$$d\left(\frac{1}{r}\right) = -\frac{1}{r^2}dr = \frac{1}{2}\left(\frac{1}{r_+} - \frac{1}{r_-}\right)\cos\psi d\psi \tag{20.15}$$

$$\frac{1}{r_-} - \frac{1}{r} = \frac{1}{2}\left(\frac{1}{r_-} - \frac{1}{r_+}\right)(1 + \sin\psi) \tag{20.16}$$

$$\frac{1}{r} - \frac{1}{r_+} = \frac{1}{2}\left(\frac{1}{r_-} - \frac{1}{r_+}\right)(1 - \sin\psi) \tag{20.17}$$

the integral (20.1) becomes

$$\begin{aligned}
\varphi_+ - \varphi_- &= \frac{1}{\sqrt{\widetilde{c}}}\int_{-\pi/2}^{\pi/2} d\psi\left[1 + \gamma\frac{a}{r}\left(\frac{1}{r_+} + \frac{1}{r_-}\right) + \gamma\frac{a}{2}\left(\frac{1}{r_+} - \frac{1}{r_-}\right)\sin\psi\right] \\
&= \frac{\pi}{\sqrt{\widetilde{c}}}\left[1 + \gamma\frac{a}{p}\right] \\
&= \pi\left[1 + \frac{1}{2}(2 - \beta + \gamma)\frac{2a}{p} + \mathcal{O}\left(\left(\frac{a}{p}\right)^2\right)\right]\cdot\left[1 + \gamma\frac{a}{p}\right] \\
&= \pi\left[1 + (2 - \beta + 2\gamma)\frac{a}{p}\right] + \mathcal{O}\left(\left(\frac{a}{p}\right)^2\right)
\end{aligned} \tag{20.18}$$

where we introduced the parameter p of the the ellipse which is defined by

$$\frac{2}{p} = \frac{1}{r_+} + \frac{1}{r_-}. \tag{20.19}$$

The precession of the perihelion during one orbit is thus given by

$$\begin{aligned}\Delta\varphi &= 2(\varphi_+ - \varphi_-) - 2\pi \\ &= \frac{6\pi a}{p}\frac{2-\beta+2\gamma}{3}.\end{aligned} \tag{20.20}$$

In GR ($\beta = \gamma = 1$) this reduces to

$$\Delta\varphi = \frac{6\pi a}{p}. \tag{20.21}$$

[1ex]
In case of Mercury ($p \simeq 55 \cdot 10^6$ km, $2a_\odot \simeq 3$ km, $\pi = 180° \cdot 3600''$) we find (with $\beta = \gamma = 1$):

$$\Delta\varphi_{\text{Mercury}} \simeq 0.104'' \tag{20.22}$$

per full orbit. In 100 years, Mercury orbits the Sun 415 times and the total perihelion shift is $\Delta\varphi \simeq 43''$.

For more distant planets (Venus, ...), $\Delta\varphi$ is at most $\sim 5''$ per century. Already in 1882, Newcomb found a perihelion precession of $43''$ per century for Mercury as due to the influence of the Sun. The full perihelion precession amounts to $575''$ per century, $532''$ of which are due to the influence of other planets (due to Newtonian gravity).

Modern measurements suggest that

$$\frac{2-\beta+2\gamma}{3} = 1.003 \pm 0.005 \tag{20.23}$$

in good agreement with GR. More recent experiments (radon echoes delay) with the Cassini spacecraft even yield

$$|\gamma - 1| = (2.1 \pm 2.3) \cdot 10^{-5}. \tag{20.24}$$

20.1 Quadrupolmoment of the Sun

A quadrupolmoment of the Sun could also induce a perihelion precession of Mercury. The mass quadrupolmoment of the Sun due to its rotation is

$$Q = J_2 M_\odot R_\odot^2 \qquad \text{with } J_2 = \frac{2}{5}\frac{R_{||} - R_\perp}{R_\odot} \tag{20.25}$$

where $R_{||}$ is the radius as measured to the poles of the Sun and $R_\perp$ is the radius in the equatorial plane. The induced gravitational potential in the orbital plane of the planets (which is the equatorial plane of the Sun) is

$$\phi(r) = -\frac{GM}{r} - \frac{GQ}{2r^3}. \tag{20.26}$$

The additional term has the same r-dependence as the additional term which is due to GR in V_{eff}:

$$V_{\text{eff}} - \frac{l^2}{2r^2} = -\frac{GM}{r} - \frac{GMl^2}{c^2r^3}. \tag{20.27}$$

In order to quantify the significance of the quadrupol-term (20.26) as compared to the general relativistic correction $\frac{GMl^2}{c^2r^3}$, we calculate the relative strength (using $l \sim pv$ and $v^2 \sim \frac{GM}{p}$):

$$\frac{GQ}{GM\frac{l^2}{c^2}} \sim \frac{J_2R_\odot^2}{p^2\frac{v^2}{c^2}} \sim \frac{J_2R_\odot^2}{p\frac{GM}{c^2}} = \frac{J_2R_\odot^2}{pa}. \tag{20.28}$$

We see that the full perihelion precession is given by

$$\Delta\varphi = \frac{6\pi a}{p}\left(\underbrace{\frac{2-\beta+2\gamma}{3}}_{\text{due to GR}} + \underbrace{\frac{J_2R_\odot^2}{2ap}}_{\text{due to rotation}}\right). \tag{20.29}$$

From observations, one finds $J_2 \sim (1-1.7)\cdot 10^{-7}$, so the additional term (in case of Mercury) is

$$\frac{J_2R_\odot^2}{2ap} \simeq 5\cdot 10^{-4}. \tag{20.30}$$

Since we know that the error in $\frac{2-\beta+2\gamma}{3}$ is of the order $5\cdot 10^{-3}$, at most $\frac{1}{10}$ of the error in (20.23) can be due to a quadrupol moment of the Sun.

21 The Lie Derivative of the Metric and Killing Vectors

Consider the Lie derivative of the metric tensor $g_{\mu\nu}$ in the direction of the vectorfield K. According to Eq. (9.7), we get

$$\begin{aligned} L_K g_{\mu\nu} &= g_{\mu\nu,k}K^k + g_{\mu k}K^k{}_{,\nu} + g_{k\nu}K^k{}_{,\mu} \\ &= \frac{\partial K_\nu}{\partial x^\mu} + \frac{\partial K_\mu}{\partial x^\nu} + K^k\left[\frac{\partial g_{\mu\nu}}{\partial x^k} - \frac{\partial g_{k\nu}}{\partial x^\mu} - \frac{\partial g_{\mu k}}{\partial x^\nu}\right] \\ &= \frac{\partial K_\nu}{\partial x^\mu} + \frac{\partial K_\mu}{\partial x^\nu} - 2K_k\Gamma^k{}_{\mu\nu} \qquad (21.1) \\ &= K_{\nu;\mu} + K_{\mu;\nu} \qquad (21.2) \end{aligned}$$

where we used

$$\begin{aligned} K^k{}_{,\nu}g_{\mu k} &= \frac{\partial K^k g_{\mu k}}{\partial x^\nu} - K^k\frac{\partial g_{\mu k}}{\partial x^\nu} \\ &= \frac{\partial K_\mu}{\partial x^\nu} - K^k\frac{\partial g_{\mu k}}{\partial x^\nu} \qquad (21.3) \end{aligned}$$

in the second step. An infinitesimal coordinate transformation is a **symmetry of the metric** if $L_K g_{\mu\nu} = 0$, i.e.

$$\boxed{K_{\mu;\nu} + K_{\nu;\mu} = 0}. \tag{21.4}$$

Any 4-vector which satisfies this equation is called a **Killing vector**. The idea is that Killing vector fields preserve the metric, i.e. flows of Killing vector fields generate symmetries (continuous isometries) of the spacetime manifold: moving every point of the manifold by the same amount in the direction of the flow of the Killing field does not change the distance between any two points.

As an example consider a stationary gravitational field for which there exists a coordinate system $\{x^\mu\}$ in which $g_{\mu\nu}$ does not depend on $x^0 = ct$. As a vector field take $K = \delta_0^\mu \partial_\mu$. Then K is obviously a Killing vector field as can be seen by inserting it into Eq. (21.4).
The interpretation of this example is simple: If the metric doesn't depend on time, then the spacetime manifold has to have a Killing vector field which is associated to time translations.

We notice that if K_1 and K_2 are Killing vector fields, then $[K_1, K_2]$ is a Killing vector field, as well. This is because from $L_{K_i} g_{\mu\nu}$, it follows that

$$0 = [L_{K_1} g_{\mu\nu}, L_{K_2} g_{\mu\nu}] = L_{[K_1,K_2]} g_{\mu\nu}. \tag{21.5}$$

We are used to the fact that symmetries of physical laws lead to conserved quantities. In the present case, we mean by symmetries of the gravitational field simply symmetries of the metric. The existence of Killing vectors will therefore lead to conserved quantities. More precisely:
If K^μ is a Killing vector and $x^\mu(\tau)$ is a geodesic, then $K_\mu \dot{x}^\mu$ is constant along the geodesic. This can easily be verified:

$$\begin{aligned}
\frac{d}{d\tau} K_\mu \dot{x}^\mu &= (K_{\mu;\nu} \dot{x}^\nu)\, \dot{x}^\mu + K_\mu \underbrace{\dot{x}^\mu{}_{;\nu}}_{=0 \text{ (geodesic)}} \dot{x}^\nu \\
&= \frac{1}{2} (K_{\mu;\nu} + K_{\nu;\mu})\, \dot{x}^\mu \dot{x}^\nu \\
&= 0.
\end{aligned} \tag{21.6}$$

Furthermore we note that if $T^{\mu\nu}$ denotes the covariantly conserved energy-momentum tensor ($T^{\mu\nu}{}_{;\mu} = 0$), then $J^\mu = T^{\mu\nu} K_\nu$ is a covariantly conserved current:

$$\begin{aligned}
J^\mu{}_{;\mu} &= \underbrace{T^{\mu\nu}{}_{;\mu}}_{=0} K_\nu + T^{\mu\nu} K_{\nu;\mu} \\
&= \frac{1}{2} T^{\mu\nu} (K_{\mu;\nu} + K_{\nu;\mu}) \\
&= 0
\end{aligned} \tag{21.7}$$

where we used the symmetry of $T^{\mu\nu}$.

22 Maximally Symmetric Space

A **maximally symmetric space** is a space which allows for the maximum number of Killing vectors (which turns out to be $\frac{n(n+1)}{2}$ for an n-dimensional space).
A suitable description of "space" in a cosmological spacetime in the context of the cosmological principle is given by *homogeneous* ("the same at every point") and *isotropic* ("the same in every direction") spaces.

We know that

$$\left([\nabla_X, \nabla_Y] - \nabla_{[X,Y]}\right) V^\lambda = R^\lambda{}_{\sigma\mu\nu} X^\mu Y^\nu V^\sigma \tag{22.1}$$

where $X = X^\mu \partial_\mu$, $Y = Y^\mu \partial_\mu$. Using

$$\nabla_{Y^\nu \partial_\nu} V^\lambda = Y^\nu \nabla_{\partial_\nu} V^\lambda = Y^\nu V^\lambda{}_{;\nu} \tag{22.2}$$

$$\nabla_X \nabla_Y V^\lambda = X^\mu Y^\nu{}_{;\mu} V^\lambda{}_{;\nu} + X^\mu Y^\nu V^\lambda{}_{;\nu;\mu} \tag{22.3}$$

we thus find

$$[\nabla_\mu, \nabla_\nu] V^\lambda = R^\lambda{}_{\sigma\mu\nu} V^\sigma. \tag{22.4}$$

Using the first Bianchi identity, one can show (no proof here) that for a Killing vector K_μ, the following holds true:

$$\nabla_\lambda \nabla_\mu K_\nu(x) = R^\rho{}_{\lambda\mu\nu} K_\rho(x) \tag{22.5}$$

for $x = x_0$. Thus a Killing vector 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.

A set of Killing vectors $\{K_\mu^{(i)}(x)\}$ is said to be **independent** if any linear relation of the form

$$\sum_i c_i K_\mu^{(i)}(x) = 0 \tag{22.6}$$

with constant coefficients c_i implies $c_i = 0$. Obviously there can at most be n independent vectors $K_\mu^{(i)}(x_0)$ at a spacetime point x_0 if the space is n-dimensional. Furthermore there can at most be $\frac{n(n-1)}{2}$ independent antisymmetric matrices $\nabla_\mu K_\nu^{(i)}(x_0)$. We draw the conclusion that in an n-dimensional spacetime there can exist at most

$$n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2} \tag{22.7}$$

independent Killing vectors. We distinguish the following cases:

- **Homogeneous space:** The n-dimensional spacetime admits n translational Killing vectors. This means that there exist infinitesimal isometries in any arbitrary direction on the spacetime manifold.

- **Isotropic space:** $\nabla_\mu K_\nu(x_0)$ is an arbitrary antisymmetric matrix. We can choose a set of $\frac{n(n-1)}{2}$ independent Killing vectors.
- **Maximally symmetric space:** A space with a metric and the maximum number of $\frac{n(n+1)}{2}$ Killing vectors.

The Riemann curvature tensor looks simpler in maximally symmetric spaces. One can show (without proof) that in maximally symmetric spaces

$$R_{ijkl} = k(g_{ik}g_{jl} - g_{il}g_{jk}) \tag{22.8}$$

for some constant k. The Ricci tensor and scalar read then

$$R_{ij} = (n-1)kg_{ij}, \tag{22.9}$$
$$R = n(n-1)k. \tag{22.10}$$

The Einstein tensor becomes

$$G_{ik} = R_{ik} - \frac{1}{2}Rg_{ik} = k(n-1)\left(1 - \frac{n}{2}\right)g_{ik}. \tag{22.11}$$

From the Bianchi identity it follows that k is a constant.

In the following we consider spacetimes in which the metric is spherically symmetric and homogeneous on each hypersurface of constant time. This assumption of the spatial subspace of the spacetime manifold being maximally symmetric, i.e. homogeneous and isotropic, is called *cosmological principle*. In our case $n = 4$ and the maximally symmetric subspace is 3-dimensional. First, we consider the metric of this 3-dimensional subspace:

$$d\sigma^2 = A(r)dr^2 + r^2 d\Omega^2 \tag{22.12}$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta d\varphi^2$. For the Christoffel symbols, we can use our previous results from the Schwarzschild solution with $B(r) = 0$:

$$R_{rr} = R_{11} = \frac{A'}{rA} \tag{22.13}$$
$$R_{\theta\theta} = R_{22} = -\frac{1}{A} + 1 + \frac{rA'}{2A^2}. \tag{22.14}$$

On the other hand, we can compute the same quantities using Eq. (22.9):

$$R_{rr} = 2kA \tag{22.15}$$
$$R_{\theta\theta} = 2kg_{\theta\theta} = 2kr^2. \tag{22.16}$$

Comparing these results, we infer

$$2kA = \frac{A'}{rA} \quad \Rightarrow \quad A' = 2krA^2 \tag{22.17}$$

and the second equation yields

$$
\begin{aligned}
2kr^2 &= -\frac{1}{A} + 1 + \frac{rA'}{2A^2} \\
&= -\frac{1}{A} + 1 + \frac{2kr^2A^2}{2A^2} \\
&= -\frac{1}{A} + 1 + kr^2 && (22.18) \\
\Rightarrow \qquad kr^2 &= -\frac{1}{A} + 1 \quad \text{or} \quad A = \frac{1}{1-kr^2}. && (22.19)
\end{aligned}
$$

This solves also Eq. (22.17). The metric on the 3-dimensional subspace (maximally symmetric) is

$$
d\sigma^2 = \frac{dr^2}{1-kr^2} + r^2 \underbrace{d\Omega^2}_{d\theta^2+\sin^2\theta d\varphi^2} . \tag{22.20}
$$

It can be shown that k can have the following values:

$$
k = \begin{cases} +1 & \text{(sphere with positive curvature)} \\ -1 & \text{(hyperbolic space with negative curvature)} \\ 0 & \text{(plane with zero curvature).} \end{cases} \tag{22.21}
$$

The full metric (including time) takes the form

$$
\boxed{ds^2 = c^2dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\varphi^2)\right]} \tag{22.22}
$$

where $a(t)$ is the cosmic scale factor which will be determined by solving the Einstein equations with the matter content of the universe.
This metric (First discovered by Friedmann-Lemaître-Robertson-Walker) is a reasonable ansatz for describing the universe. There is good evidence that the universe - on large scales - is surprisingly homogeneous and isotropic. Experimental evidence for this is provided by by measurements of redshifts of galaxies and cosmic microwave background radiation.

Note that we didn't use the field equations so far. We derived the metric solely from some assumptions concerning the symmetry properties of the underlying spacetime.

23 Friedmann Equations

We write the metric (22.22) as follows:

$$
ds^2 = c^2dt^2 - a^2(t)\widetilde{g}_{ij}dx^idx^j \tag{23.1}
$$

where the tilde refers to 3-dimensional quantities. The Christoffel symbols are given by (notice that ${\Gamma^\mu}_{00} = 0$)

$$
{\Gamma^i}_{jk} = {\widetilde{\Gamma}^i}_{jk} \tag{23.2}
$$

$$
{\Gamma^i}_{j0} = \frac{\dot{a}}{a}\delta^i_j \tag{23.3}
$$

$$
{\Gamma^0}_{ij} = \dot{a}a\widetilde{g}_{ij}. \tag{23.4}
$$

Therefore the relevant components of the Riemann tensor are

$$R^i{}_{0j0} = -\frac{\ddot{a}}{a}\delta^i_j \tag{23.5}$$

$$R^0{}_{i0j} = a\ddot{a}\widetilde{g}_{ij} \tag{23.6}$$

$$R^k{}_{ikj} = \widetilde{R}_{ij} + 2\dot{a}^2\widetilde{g}_{ij}. \tag{23.7}$$

We can now use $\widetilde{R}_{ij} = 2k\widetilde{g}_{ij}$ (maximal symmetry of 3-dimensional subspace) to compute $R_{\mu\nu}$. The nonzero components are

$$R_{00} = -3\frac{\ddot{a}}{a} \tag{23.8}$$

$$R_{ij} = (a\ddot{a} + 2\dot{a}^2 + 2k)\widetilde{g}_{ij} = \left(\frac{\ddot{a}}{a} + 2\frac{\dot{a}^2}{a^2} + \frac{2k}{a^2}\right) g_{ij} \tag{23.9}$$

where $g_{ij} = a^2\widetilde{g}_{ij}$. The Ricci scalar is

$$R = \frac{6}{a^2}(a\ddot{a} + \dot{a}^2 + k) \tag{23.10}$$

and the Einstein tensor takes the form

$$G_{00} = 3\left(\frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) \tag{23.11}$$

$$G_{0i} = 0 \tag{23.12}$$

$$G_{ij} = -\left(2\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right) g_{ij}. \tag{23.13}$$

Now we have to specify the matter content which should be in accordance with our assumption of homogeneity and isotropy. We treat the universe as a perfect fluid consisting or a collection of non-interacting particles (galaxies). A perfect fluid has the following energy-momentum tensor (cf. Eq. (14.17)):

$$T_{\mu\nu} = \left(\frac{p}{c^2} + \rho\right) u_\mu u_\nu - g_{\mu\nu}p \tag{23.14}$$

where p is the pressure, ρ is the energy density and u_μ is the velocity field of the fluid (for example, in the comoving coordinate system, we have $u_\mu = (1, 0, 0, 0)$). We need some equation of state $p = p(\rho)$, for example

$$p = \omega\rho \tag{23.15}$$

where ω is the equation of state parameter. An example would be the case of non-interacting particles where $p = \omega = 0$ such that matter is referred to as **dust**.
The trace of the energy-momentum tensor is (with $c = 1$)

$$T^\mu{}_\mu = \rho - 3p. \tag{23.16}$$

In case of radiation, the energy-momentum tensor (as in Maxwell's theory) is traceless so that radiation is described by the equation of state

$$p = \frac{1}{3}\rho, \tag{23.17}$$

so $\omega = \frac{1}{3}$.
As we will see, a cosmological constant Λ corresponds to a "matter"-contribution with $\omega = -1$.
Note that we have the following conservation law: From $T^{\nu\mu}{}_{;\mu} = 0$ we immediately infer $T^{0\mu}{}_{;\mu} = 0$ or

$$\partial_\mu T^{\mu 0} + \Gamma^\mu{}_{\mu\nu} T^{\nu 0} + \Gamma^0{}_{\mu\nu} T^{\mu\nu} = 0. \tag{23.18}$$

For a perfect fluid, this equation reads

$$\partial_t \rho(t) + \Gamma^\mu{}_{\mu 0}\rho + \Gamma^0{}_{00}\rho + \Gamma^0{}_{ij} T^{ij} = 0 \qquad (i, j = 1, 2, 3). \tag{23.19}$$

Inserting the expressions for the Christoffel symbols (23.4), we get

$$\dot\rho = -3(\rho + p)\frac{\dot a}{a}. \tag{23.20}$$

For dust, this equation becomes

$$\frac{\dot\rho}{\rho} = -3\frac{\dot a}{a}. \tag{23.21}$$

Integration yields

$$\rho a^3 = \text{const.} \qquad \text{or } \rho \sim a^{-3}. \tag{23.22}$$

For a radiation dominated universe, we have $p = \frac{\rho}{3}$ and thus

$$\frac{\dot\rho}{\rho} = -4\frac{\dot a}{a} \tag{23.23}$$

which leads to

$$\rho a^4 = \text{const.} \qquad \text{or } \rho \sim a^{-4}. \tag{23.24}$$

More generally, one gets for (23.15)

$$\rho a(t)^{3(1+\omega)} = \text{const.}, \tag{23.25}$$

so ρ is constant for $\omega = -1$.

Recall the Einstein equations with cosmological constant Λ:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = -\frac{8\pi G}{c^4} T_{\mu\nu}. \tag{23.26}$$

These equations finally read

$$-3\frac{\ddot a}{a} = 4\pi G(\rho + 3p) - \Lambda \tag{23.27}$$

$$\left(\frac{\ddot a}{a} + 2\frac{\dot a^2}{a} + \frac{2k}{a^2}\right) = 4\pi G(\rho - p) + \Lambda. \tag{23.28}$$

Additionally, we have Eq. (23.20) as a conservation law.
Using the first equation to eliminate $\ddot{a}$ from the second equation, one obtains the **Friedmann equations**:

$$\frac{\dot{a}^2}{a^2} + \frac{k}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} \tag{23.29}$$

$$-3\frac{\ddot{a}}{a} = 4\pi G(\rho + 3p) - \Lambda \tag{23.30}$$

$$\dot{\rho} = -3(\rho + p)\frac{\dot{a}}{a}. \tag{23.31}$$

We introduce the **Hubble parameter** $H(t) = \frac{\dot{a}(t)}{a(t)}$ and the **deceleration parameter** $q(t) = -\frac{a(t)\ddot{a}(t)}{\dot{a}^2(t)}$. Present day values are denoted by $H_0 \equiv H(t_0)$, $q_0 \equiv q(t_0)$ where t_0 is the recent age of the universe. In terms of these parameters, the Friedmann equations read

$$\boxed{\begin{aligned} H^2 &= \frac{8\pi G}{3}\rho - \frac{k}{a^2} + \frac{\Lambda}{3} \\ q &= \frac{1}{3H^2}(4\pi G(\rho + 3p) - \Lambda) \\ \frac{d}{dt}(\rho a^3) &= -3Hpa^3. \end{aligned}} \tag{23.32}$$

For vanishing cosmological constant ($\Lambda = 0$), we define the **critical density** $\rho_{\text{crit}} = \frac{3H^2}{8\pi G}$. The **density parameter** is defined as $\Omega = \frac{\rho}{\rho_{\text{crit}}}$ and we have

$$\begin{array}{lcll} \rho < \rho_{\text{crit}} & \Leftrightarrow & k = -1 & \text{(open universe)} \\ \rho = \rho_{\text{crit}} & \Leftrightarrow & k = 0 & \text{(flat universe)} \\ \rho > \rho_{\text{crit}} & \Leftrightarrow & k = +1 & \text{(closed universe)} \end{array}$$

(a "closed" universe is one which collapses after a finite time). The density receives contributions from ordinary baryonic matter ($\Omega_{\text{baryons}} \sim 0.04$), dark matter ($\Omega_{\text{DM}} \sim 0.24$) and dark energy from the cosmological constant $\Omega_\Lambda = \frac{\rho_\Lambda}{\rho_{\text{crit}}} \sim 0.72$ with $\rho_\Lambda = \frac{\Lambda}{8\pi G}$. In total, $\Omega = \Omega_{\text{baryons}} + \Omega_{\text{DM}} + \Omega_\Lambda \sim 1$.

We note that H_0^{-1} is related to the age of the universe. As far as we know, $H_0^{-1} \sim 70\frac{\text{km}}{\text{sec/Mpc}}$, giving an age of ~ 13.7 billion years.

Literature

- T. Fliessbach, *Allgemeine Relativitätstheorie*, Spektrum Verlag, 1995
- 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
- 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