General Relativity

Christian G. Böhmer*
Department of Mathematics, University College London,
Gower Street, London, WC1E 6BT, UK

October 2, 2008

Abstract

These lecture notes are based on my handwritten notes used in the academic year 2007/2008. The first chapters mainly follow Mitchell Berger's lecture notes who taught this class for many years in the past. From him I took the idea of introducing the geodesic equations as early as possible and let students work out Christoffel symbol components and solve geodesic equations to find improved coordinate systems. By this I mean writing a flat two dimensional metric in some funny coordinates and solving the geodesic equations which will reveal the geodesics to be straight lines etc. Chapter 4–6 cover the standard material covered by an introductory course on general relativity.

As can be seen in the structure of the course, I am following the "mathematics first – physics second" (i.e. mathematical physics) school when introducing differential geometry and general relativity. However, I do prefer concrete examples and worked out calculations over abstract concepts only.

Note that figures are still missing in these notes. Please inform me about typos, flaws or inaccuracies as this is not yet the final version.

*`c.boehmer@ucl.ac.uk`

Contents

Recommended books

B. F. Schutz
A first course in general relativity
Cambridge University Press

R. Wald
General Relativity
Chicago University Press

S. Carroll
Spacetime and Geometry
Addison Wesley

S. Carroll
Lecture notes on general relativity
http://arxiv.org/abs/gr-qc/9712019

J. Plebanski and A. Krasinski
An Introduction to General Relativity and Cosmology
Cambridge University Press

Misner, Thorne and Wheeler
Gravitation
W. H. Freeman & Co Ltd

S. Weinberg
The general theory of relativity
John Wiley & Sons

Bishop and Goldberg
Tensor analysis on manifold
Dover

1 Manifolds

1.1 Manifolds

A manifold generalises the idea of a surface or a space.

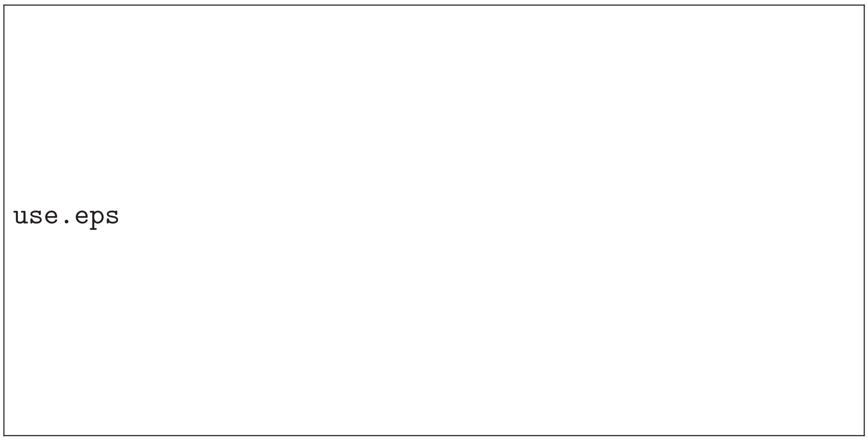
Definition 1.1 A manifold M:

i M is a set of points which can be mapped into $\mathbb{R}^n$, $n \in \mathbb{N}$, where n is called the dimension of the manifold

ii this mapping must be one-to-one

iii if two mappings overlap, one must be a differentiable function of the other

Roughly speaking some neighbourhood of each point admits a coordinate system. Note that this mapping is often called chart in the literature.



$$
\begin{aligned}
\mu &: U \to \mathbb{R}^n \\
\tau &: U \to \mathbb{R}^n \\
\mu &= \mu(X^1, X^2, \ldots, X^n) \\
\tau &= \tau(Y^1, Y^2, \ldots, Y^n)
\end{aligned}
$$

μ and τ are the coordinate maps of the respective neighbourhoods, while $X^1, X^2, \ldots, X^n$ and $Y^1, Y^2, \ldots, Y^n$ are the local coordinates.

Example 1.1 $M = \mathbb{S}^1$ (circle) One coordinate ϕ which can be chosen to go from $-\pi$ to π. However, the point mapped to $\phi = \pi$ is also mapped to $\phi = -\pi$, hence not single valued and therefore not one-to-one. At least two coordinate patches are required to cover the circle.

Example 1.2 $M = \mathbb{S}^2$ (sphere) Similar to the circle, spherical coordinates cannot over the whole sphere. The azimuthal angle ϕ is many valued at the poles and moreover not continuous at $\phi = \pm\pi$ (Globe of the Earth: international date line). Not single valued and not one-to-one. At least two coordinate patches are required to cover the sphere.

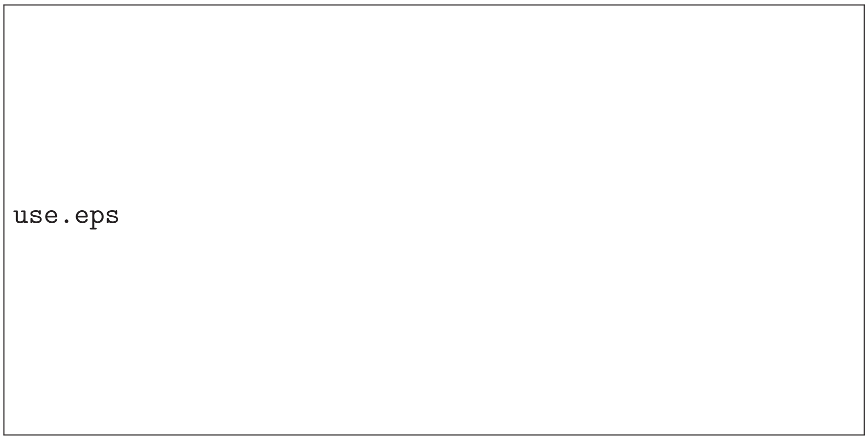
Example 1.3 $M = \mathbb{S}^n$ (n-sphere) The n-sphere is defined by

$$\mathbb{R}^n = \{(X^1 \ldots, X^{n+1}) | (X^1)^2 + \cdots + (X^{n+1})^2 = 1\} \tag{1.1}$$

At least two coordinate patches are required to cover the n-sphere.

Example 1.4 $M = \mathbb{T}^2$ (2-torus) The 2-torus is the surface of an American doughnut. One possible choice for the coordinates consists of the two angles θ and ϕ, representing the and the long way round respectively. These are not continuous at $\pm\pi$.

One can cut the torus twice and open it up. Then $\mathbb{T}^2$ can also be represented as a rectangle.



Note that the two lines $\phi = -\pi$ and $\phi = \pi$ are identified (this means are in fact equal) and likewise for the angle θ.

The torus $\mathbb{T}^2$ is described by two independent circles. This is expressed by the notation

$$\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1 \tag{1.2}$$

and is called a product manifold.

Example 1.5 M = Möbius strip or Möbius band

use.eps

Join the end of a strip of paper with a single half twist.

Have a look at this surface! It has only *o*ne side and only *o*ne boundary component.

One can even visualise a spinor using the Möbius strip as one a rotation by 4π is needed to return to the starting point.

1.2 Coordinate transformations

Many physical situations are symmetric; e.g. the gravitational field of the sun or our Earth are nearly spherically symmetric. In such cases it is useful to work with coordinates adapted to the symmetry of the system.

Let U be open, and let a point $p \in U \subset M$ have coordinates

$$X = (X^1, \dots, X^n), \qquad Y = (Y^1, \dots, Y^n) \tag{1.3}$$

in two respective coordinate systems X and Y. Then the coordinates $Y^1, \dots, Y^n$ must be differentiable functions of the coordinates $X^1, \dots, X^n$, this means

$$Y^1 = Y^1(X^1, \dots, X^n) \tag{1.4}$$
$$Y^2 = Y^2(X^1, \dots, X^n) \tag{1.5}$$
$$\vdots \tag{1.6}$$
$$Y^n = Y^n(X^1, \dots, X^n) \tag{1.7}$$

Recall the Jacobi matrix from vector calculus

$$\frac{\partial Y^a}{\partial X^b} = J^a{}_b \tag{1.8}$$

$$a : \text{labels the rows} \tag{1.9}$$

$$b : \text{labels the columns} \tag{1.10}$$

Recall (iii) in definition 1.1 of a manifold.

1.3 Notation and conventions

In differential geometry and consequently in general relativity the placement of indices is very important. Objects with superscripts are different from those with subscripts.

To simplify the notation of what follows we use the Einstein summation convention. Roughly speaking, sum over twice repeated indices.

Definition 1.2 Einstein summation convention. Given two objects, one indexed with superscripts $A = (A^1, \ldots, A^n)$ and the other with subscripts $B = (B^1, \ldots, B^n)$, one defines

$$A^c B_c = \sum_{c=1}^{n} A^c B_c \tag{1.11}$$

which means that we drop the summation symbol whenever possible.

Definition 1.3 Partial derivative. We abbreviate the notation of partial derivative in the following manner

$$\frac{\partial f}{\partial X^a} = \partial_a f = f_{,a} \tag{1.12}$$

Definition 1.4 Coordinate system and coordinates. Coordinate systems are labelled by either unprimed symbols $X, Y, \ldots$ and primed symbols $X', Y', \ldots$ or other capital letters. Later the actual coordinates will be denoted by non-capital letters.

2 Vectors, tensors and metrics

So far we have defined what is meant by a manifold. We also discussed coordinate systems and changes of coordinate systems. Now we define and introduce objects that live on the manifold.

Advice: We will shortly be defining a vector. For now try to put everything you know about vectors aside. In differential geometry one should not think of a vector as a column of numbers.

2.1 Definitions

Definition 2.1 Scalar fields. Scalar fields are functions that assign numbers to points on the manifold. More precisely, a scalar field is a function f which maps a manifold M to the set of real numbers (scalars do not transform under coordinate transformations)

$$f : M \to \mathbb{R} \tag{2.1}$$

Example 2.1 $M =$ surface of the Earth. $p \in M$, let $f(p)$ be the temperature at p. Watch BBC news for such weather maps.

Definition 2.2 Vector or contravariant vector (this notation is not the modern mathematicians way, however, it is essential to read books and articles on general relativity and gravitation). A vector is an object with one superscript that transforms under coordinate transformations as follows

$$V'^a = \frac{\partial X'^a}{\partial X^b} V^b \tag{2.2}$$

Definition 2.3 1-form or covariant vector. A covariant vector is an object with one subscript that transforms under coordinate transformations as follows

$$W'_b = \frac{\partial X^c}{\partial X'^b} W_c \tag{2.3}$$

Definition 2.4 Tensor. A type $\binom{p}{q}$ tensor is an object with p superscripts and q subscripts. It is said to be of rank $p+q$.

Under coordinate transformations a $\binom{p}{q}$ tensor transforms according to

$$T'^{a_1 \ldots a_p}{}_{b_1 \ldots b_q} = \frac{\partial X'^{a_1}}{\partial X^{c_1}} \cdots \frac{\partial X'^{a_p}}{\partial X^{c_p}} \frac{\partial X^{d_1}}{\partial X'^{b_1}} \cdots \frac{\partial X^{d_q}}{\partial X'^{b_q}} T^{c_1 \ldots c_p}{}_{d_1 \ldots d_q} \tag{2.4}$$

Scalars and vectors are special kinds of tensors, scalars are rank 0 tensors and vectors are rank 1 tensor. A rank 2 tensor can be visualised as a $n \times n$ matrix.

Lemma 2.1 *The transformations (2.2) and (2.3) are inverse.*

Proof Consider the product

$$A^a_c = \frac{\partial X'^a}{\partial X^b}\frac{\partial X^b}{\partial X'^c} \underset{\substack{\uparrow \\ \text{chain rule}}}{=} \frac{\partial X'^a}{\partial X'^c} = \begin{cases} 1 & a = c \\ 0 & a \neq c \end{cases} \tag{2.5}$$

and therefore we find that $A^a_c = \delta^a_c$ which is the identity. $\square$

In the above we defined the Kronecker δ by

$$\delta^a_c = \begin{cases} 1 & a = c \\ 0 & a \neq c \end{cases} \tag{2.6}$$

Note that this Lemma in particular implies that contravariant vectors (vectors) and covariant vectors (1-forms) have different (inn fact inverse) transformation laws. Contravariant and covariant vectors are dual objects (in the Algebra sense), they combine to give a number in $\mathbb{R}$ or $\mathbb{C}$. In elementary matrix algebra row vectors and column vectors multiply to give a number.

In quantum theory a bra vector and a ket vector ($<\psi|$ and $|\phi>$, respectively) give a complex number $<\psi|\phi>$. This is the inner product defined on a Hilbert space.

2.2 Tensor algebra

Definition 2.5 Addition. Two tensors of the same type and the same index structure can be added

$$R^a{}_b{}^c + S^a{}_b{}^c = T^a{}_b{}^c \tag{2.7}$$

but not $A^a + B_a$.

Definition 2.6 Composition. Given a type $\binom{p}{q}$ and another type $\binom{r}{s}$ tensor, these can be combined to a type $\binom{p+r}{q+s}$ tensor.

Example 2.2

$$V^a = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \qquad W_a = \begin{pmatrix} 2 & 3 \end{pmatrix}, \tag{2.8}$$

$$M^a{}_b = V^a W_b = \begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix} \tag{2.9}$$

Definition 2.7 Contraction. Given a type $\binom{p}{q}$ tensor one can sum over one upper and one lower index which results in a type $\binom{p-1}{q-1}$ tensor

$$T^{a_1 \dots d \dots a_p}{}_{b_1 \dots d \dots b_q} = U^{a_1 \dots a_p}{}_{b_1 \dots b_q} \tag{2.10}$$

where $(a_1 \dots a_p)$ contains $p-1$ indices and $(b_1 \dots b_q)$ contains $q-1$ indices.

Example 2.3

$$M^a{}_a = M^1{}_1 + M^2{}_2 = 2 + 6 = 8 \tag{2.11}$$

Definition 2.8 Trace. Let $M^a{}_b$ be a rank 2 tensor of type $\binom{1}{1}$. The trace is defined by

$$\operatorname{tr} M = M^a{}_a \tag{2.12}$$

General advice: Whenever one encounters tensor equations one should check that the index structure of the left-hand side and the right-hand side agree.

Definition 2.9 Let T^{ab} be a rank 2 tensor of type $\binom{2}{0}$. We define symmetrisation as follows

$$T^{(ab)} = \frac{1}{2}(T^{ab} + T^{ba}) \tag{2.13}$$

and anti-symmetrisation by

$$T^{[ab]} = \frac{1}{2}(T^{ab} - T^{ba}) \tag{2.14}$$

Definition 2.10 T^{ab} is called symmetric if

$$T^{ab} = T^{ba} \qquad \Leftrightarrow \qquad T^{ab} = T^{(ab)} \tag{2.15}$$

and a tensor T^{ab} is called anti-symmetric or skew-symmetric if

$$T^{ab} = -T^{ba} \qquad \Leftrightarrow \qquad T^{ab} = T^{[ab]} \tag{2.16}$$

Note that according to definition 2.9 any tensor T^{ab} cab always be written such that

$$T^{ab} = T^{(ab)} + T^{[ab]} \tag{2.17}$$

Definition 2.11 Levi-Civita tensor. The Levi-Civita tensor (actually tensor density) is a completely skew-symmetric tensor of rank n in n dimensions

$$\varepsilon^{ab\dots n} = \begin{cases} 0 & \text{if any two of the indices are equal} \\ +1 & \text{if } (a,b,\dots,c) \text{ is an even permutation} \\ -1 & \text{if } (a,b,\dots,c) \text{ is an odd permutation} \end{cases} \tag{2.18}$$

Example 2.4 In $n = 2$ dimensions we have

$$\varepsilon^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{2.19}$$

Example 2.5 $n = 3$. Let $M = mathbbE^3$ be Euclidean 3-space with Cartesian coordinates x, y, z. The usual cross or vector product can be defined by

$$(\vec{\nabla} \times \vec{A})^i = \varepsilon^{ijk} \partial_j A_k \tag{2.20}$$

Let us for example consider the y-component

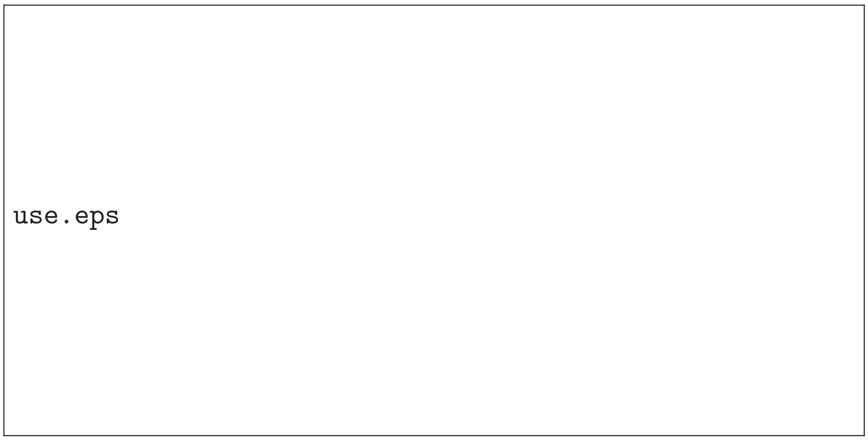
$$\begin{aligned} (\vec{\nabla} \times \vec{A})^y &= \varepsilon^{yjk} \partial_j A_k \\ &= \varepsilon^{yxz} \partial_x A_z + \varepsilon^{yzx} \partial_z A_x \\ &= \partial_z A_x - \partial_x A_z \end{aligned} \tag{2.21}$$

where in the summation only the non-vanishing terms were taken into account. Similarly for the other two components.

Many of the well-known 3-vector identities can easily be proved in index notation.

2.3 Metrics & Geodesics I

The main question we are not going to address is how to measure the distance of the two points in a manifold. Let us recall how to measure the distance the distance of two points in Euclidean 2-space



$$s^2 = x^2 + y^2 \tag{2.22}$$

Next, let us consider small distances $\Delta s, \Delta x, \Delta y$

$$\Delta s^2 = \Delta x^2 + \Delta y^2 \tag{2.23}$$

In the limit of infinitesimal distances we can formally write

$$ds^2 = dx^2 + dy^2 \tag{2.24}$$

Definition 2.12 Metric. Let $(X^1, \ldots, X^n)$ and $(X^1 + dX^1, \ldots, X^n + dX^n)$ be two nearby points. The distance can be defined by introducing the metric tensor g_{ab}. The distance satisfies

$$ds^2 = g_{ab} dX^a dX^b \tag{2.25}$$

In general g_{ab} is an arbitrary function of the coordinates; we also assume it is non-degenerate and therefore its inverse exists $g_{ab}g^{bc} = \delta_a^c$. ds^2 is often called the line element.

Since g_{ab} is a symmetric rank 2 tensor, it has $n(n+1)/2$ independent components in n dimensions.

The metric can be used to lower and raise indices. Therefore, the metric provides us with a one-to-one mapping between contravariant and covariant vector, dual vectors. Since general relativity is formulated on metric manifolds, physicists are happy to drop the distinction between vectors and dual vectors. It is only the index position that matters.

Definition 2.13 Total differential. Let f be a function of several variables $f = f(X^1, \ldots, X^n)$, the total differential is defined by

$$df = f_{,a} dX^a = \sum_{a=1}^{n} \frac{\partial f}{\partial X^a} dX^a \tag{2.26}$$

also known as the differential.

Example 2.6 $M = \mathbb{E}^2$ Euclidean 2-space. In Cartesian coordinates $(X^1, X^2) = (x, y)$ the line element takes the form

$$ds^2 = dx^2 + dy^2 \qquad g_{ab} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathrm{diag}(1, 1) \tag{2.27}$$

Let us introduce spherical coordinates

$$x = r \cos\varphi \tag{2.28}$$

$$dx = \cos\varphi dr + r(-\sin\varphi) d\varphi \tag{2.29}$$

$$dx^2 = \cos^2\varphi dr^2 + r^2 \sin^2\varphi d\varphi^2 - 2r \sin\varphi \cos\varphi dr d\varphi \tag{2.30}$$

$$y = r \sin\varphi \tag{2.31}$$

$$dy = \sin\varphi dr + r \cos\varphi d\varphi \tag{2.32}$$

$$dy^2 = \sin^2\varphi dr^2 + r^2 \cos^2\varphi d\varphi^2 + 2r \sin\varphi \cos\varphi dr d\varphi \tag{2.33}$$

Adding up both equations we find

$$ds'^2 = dr^2 + r^2 d\varphi^2 \qquad g_{ab} = \mathrm{diag}(1, r^2) \tag{2.34}$$

Definition 2.14 Signature of the metric. Using results from linear algebra, one can show that at any point $p \in M$ the metric can be diagonalised. In general one can the diagonal elements ± 1 in a suitable basis. However, the number of $+$ signs and the number of $-$ signs are independent of that choice. Hence, we call the number of $+$ and $-$ signs occurring the signature of the metric. Often their sum is used as the signature.

Example 2.7 $M = \mathbb{E}^3$ Euclidean 3-space with Cartesian coordinates

$$ds^2 = dx^2 + dy^2 + dz^2 \tag{2.35}$$

has signature $(+, +, +)$ or signature 3.

Example 2.8 M = Rindler spacetime

$$ds^2 = -x^2dt^2 + dx^2 \tag{2.36}$$

with coordinate ranges $-\infty < t < \infty$ and $0 < x < \infty$. This metric has signature $(-,+)$ or signature 0.

Definition 2.15 Riemannian metric. Let $V^a \neq 0$ be a contravariant, non-vanishing vector. The metric g_{ab} is called Riemannian if

$$g_{ab}V^aV^b > 0 \qquad \forall\ V^a \neq 0 \tag{2.37}$$

A manifold equipped with a Riemannian metric is called a Riemannian manifold.

Lemma 2.2 *Let M be a n-dimensional manifold of signature n. Then M is Riemannian.*

Proof Let V^a be a non-vanishing vector at some point $p \in M$. Let us choose coordinates at p such that g_{ab} is diagonal. Since g_{ab} has signature n, all diagonal elements are positive and can be made $+1$ locally. Hence

$$g_{ab}V^aV^b = (V^1)^2 + \cdots + (V^n)^2 > 0 \tag{2.38}$$

since V^a is non-vanishing. □

Definition 2.16 Pseudo-Riemannian metric. A metric which in not Riemannian is called pseudo-Riemannian.

Example 2.9 Rindler spacetime (2.36) is a pseudo-Riemannian metric.

Example 2.10 $M = \mathbb{M}^4$ Minkowski spacetime is given by the metric

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2 \tag{2.39}$$

and has signature -2 and is pseudo-Riemannian

Definition 2.17 Lorentzian metrics. Metrics with either of the following signatures $(-,+,\ldots,+)$ or $(+,-,\ldots,-)$ are called Lorentzian metrics (only one sign is different). A manifold with a Lorentzian metric is called Lorentzian manifold.

In the above we defined the notion of a scalar field. It maps the manifold M to the reals $\mathbb{R}$. A curve describes the reverse situation.

Definition 2.18 Curve. A curve is a mapping of the real line (or parts of the real line, or a circle) into the manifold M

$$\gamma : \mathbb{R} \to M \tag{2.40}$$

A smooth curve γ on M is a C^∞ mapping of $\mathbb{R}$ (or parts thereof) into M. A curve is usually parametrised by τ or $lambda$.

Let us assume that γ lies is an open region $U \subset M$. By definition there exists a set of local coordinates μ that map U into $\mathbb{R}^n$. Then the curve provides a set of n coordinate functions of the parameter λ

$$\gamma(\lambda) = \begin{pmatrix} X^1(\lambda) \\ \vdots \\ X^n(\lambda) \end{pmatrix} = X^a(\lambda) \tag{2.41}$$

Definition 2.19 Tangent vector to a curve. Let γ be a smooth curve on M. The tangent to the curve γ is any coordinate basis is given by

$$T^a \frac{dX^a}{d\lambda} \tag{2.42}$$

Let us assume a curve $X^a(\lambda)$ connects two points on a manifold M. Since we know how to measure distances on the manifold, we can also compute the length of the curve

$$\int ds = \int \frac{ds(\lambda)}{d\lambda} d\lambda = \int \sqrt{g_{ab} \dot{X}^a \dot{X}^b} d\lambda \tag{2.43}$$

where we denoted $\dot{X}^a = dX^a/d\lambda$.

The most natural question to ask then: Which curves are the shortest curves to connect two points, or the straightest possible lines?

Definition 2.20 Geodesics. Let $X^a(\lambda)$ be a curve. We define a geodesic to be a curve whose path extremises the functional

$$\int ds = \int \sqrt{g_{ab} \dot{X}^a \dot{X}^b} d\lambda \tag{2.44}$$

Without loss of generality (length is parametrisation independent) we may assume that the curve is parametrised so that

$$g_{ab}\dot{X}^a\dot{X}^b = g_{ab}T^aT^b = 1 \tag{2.45}$$

which is also called the affine parametrisation.

The functional can be extremised using well-known techniques of Lagrangian mechanics. The geodesic equation can also be obtained (in affine parametrisation) from the Lagrangian $L = g_{ab}\dot{X}^a\dot{X}^b$, which simplifies the calculations.

Lemma 2.3 *Geodesic equation. A geodesic satisfies the following equations of motion*

$$\frac{d^2X^a}{d\lambda^2} + \Gamma^a_{bc}\frac{dX^b}{d\lambda}\frac{dX^c}{d\lambda} \tag{2.46}$$

where

$$\Gamma^a_{bc} = \frac{1}{2}g^{ad}(g_{db,c} + g_{cd,b} - g_{bc,d}) \tag{2.47}$$

Proof Consider the Lagrangian

$$L = g_{ab}(X^c)\dot{X}^a\dot{X}^b \tag{2.48}$$

The Euler-Lagrange equations are given by

$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{X}^c} = \frac{\partial L}{\partial X^c} \tag{2.49}$$

$$\frac{\partial L}{\partial X^c} = \frac{\partial g_{ab}}{\partial X^c}\dot{X}^a\dot{X}^b = g_{ab,c}\dot{X}^a\dot{X}^b \tag{2.50}$$

$$\frac{\partial L}{\partial \dot{X}^c} = g_{ab}\dot{X}^a\delta^b_c + g_{ab}\delta^a_c\dot{X}^b = g_{ac}\dot{X}^a + g_{cb}\dot{X}^b = 2g_{ca}\dot{X}^a \tag{2.51}$$

$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{X}^c} = 2g_{ca,b}\dot{X}^b\dot{X}^a + 2g_{ca}\ddot{X}^a \tag{2.52}$$

Hence, we find the following equations of motion

$$g_{ab,c}\dot{X}^a\dot{X}^b = 2g_{ca}\ddot{X}^a + g_{ca,b}\dot{X}^b\dot{X}^a + g_{cb,a}\dot{X}^b\dot{X}^a \tag{2.53}$$

which after sorting the terms leads to

$$2g_{ca}\ddot{X}^a + (g_{ca,b} + g_{cb,a} - g_{ab,c})\dot{X}^a\dot{X}^b = 0 \tag{2.54}$$

Next, we apply g^{cd} to this latter equation and find

$$2\delta^d_a \ddot{X}^a + g^{cd}(g_{ca,b} + g_{bc,a} - g_{ab,c})\dot{X}^a\dot{X}^b = 0 \tag{2.55}$$
$$\ddot{X}^d + g^{dc}(g_{ca,b} + g_{bc,a} - g_{ab,c})\dot{X}^a\dot{X}^b = 0 \tag{2.56}$$

Finally we rename the indices $a \to b$, $b \to c$, $c \to d$, $d \to a$ and arrive at

$$\ddot{X}^a + \Gamma^a_{bc}\dot{X}^b\dot{X}^c = 0 \tag{2.57}$$

□

Γ^a_{bc} is called Christoffel symbol and is of paramount interest in general relativity.

The Christoffel symbol is called symbol because it is NOT a tensor. It does NOT transform like a tensor under general coordinate transformations.

2.4 A glance forward

At the end of the last subsection we derived the geodesic equations of motions. Geometrically speaking these curves extremise the length between its endpoints, they are the 'straightest possible' lines, intuitively speaking.

The geodesic equation determines the movement of a particle in a gravitational field. Let us consider

$$\frac{d^2X^a}{d\lambda^2} = -\Gamma^a_{bc}\dot{X}^b\dot{X}^c \tag{2.58}$$

and compare with the equations of motion of a particle in Newtonian gravity

$$m\ddot{\mathbf{r}} = -m\nabla\Phi(\mathbf{r}) \tag{2.59}$$

where Φ is the gravitational potential and $\mathbf{r}$ is the position vector of the particle.

If for the moment we denote the three components or $\mathbf{r}$ by x^i, the Newtonian equations of motion become

$$\frac{d^2x^i}{dt^2} = -\frac{\partial\Phi}{\partial x^i} \tag{2.60}$$

Since the Christoffel symbol contain first derivatives of the metric tensor, we can suspect that the metric will contain the gravitational potential. Moreover, there should be a well-defined procedure to obtain the Newton's equations from the geometrical equations.

Besides the equations of motion, Newton's theory of gravity is described by the field equation

$$\Delta\Phi(\mathbf{r}) = 4\pi G\rho(\mathbf{r}) \tag{2.61}$$

This is a second order linear partial differential equation. Since we suspect the metric to contain the gravitational potential, we expect to find second order equations in the metric tensor as field equations (first derivatives of the Christoffel symbol).

We will see in Section 4 that the curvature of a manifold contains second derivatives of the metric.

Before going further into differential geometry let us briefly discuss a few aspects of special relativity and Maxwell's theory covariantly formulated.

3 A little special relativity

3.1 Introduction

Special relativity is the study of physics in a universe governed by the Minkowski metric

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2 \tag{3.1}$$

Minkowski spacetime has coordinates

$$(X^0, X^1, X^2, X^3) = (t, x, y, z) \tag{3.2}$$

and therefore the Minkowski metric is given by

$$\eta_{ab} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.3}$$

Conventionally it is denoted by η_{ab} and not g_{ab}.

A few notes:

i since space and time should be measured using the same units we have either $X^0 = ct$ or $X^1 = x/c$, $X^2 = y/c$ and $X^3 = z/c$. In order to avoid factors of c we set $c = 1$

ii there is another convention for the Minkowski metric, namely

$$\eta_{ab} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.4}$$

both conventions are commonly used

iii the time coordinate is either called $X^0 = t$ but sometime, especially in the older literature, one finds $X^4 = t$ in which case the metric is often presented as

$$ds^2 = dx^2 + dy^2 + dz^2 - dt^2 \tag{3.5}$$

Special relativity is based on the principle of the constancy of the speed of light.

The purely spatial part of the metic is Euclidean 3-space. Hence this part is invariant under translations and spatial rotations. However, since we consider the four (space + time) dimensional manifold, in principle there are rotations involving both space and time coordinates.

Definition 3.1 Boost. A boost is a transformation to a coordinate system moving at constant relative velocity with respect to the original one.

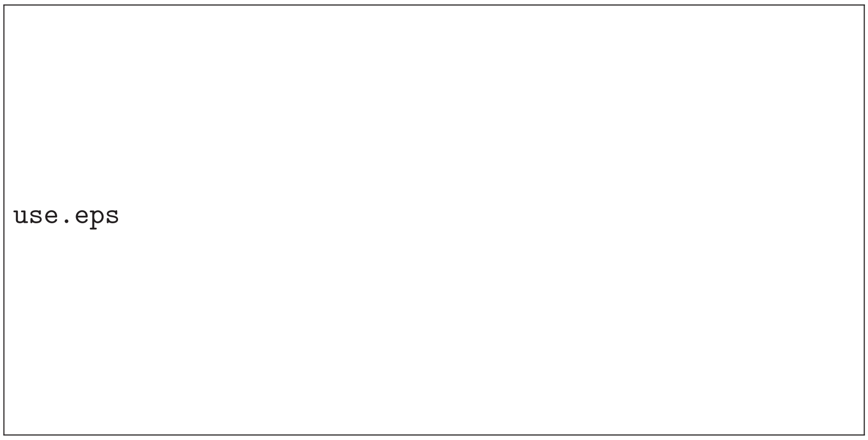
Definition 3.2 Inertial reference frame. An inertial reference frame is a coordinate system with Cartesian coordinates, and where there exists no inertial (fictitious) forces.

Definition 3.3 Einstein's axioms of special relativity:

i The laws of physics are invariant to translations, rotations and boosts.

ii The speed of light is equal in all inertial reference frames.

3.2 Spacetime diagrams



Definition 3.4 World-line. The world-line of an object is the path it traces in spacetime.

Consider one space dimension, the velocity of an object is dx/dt. So the slope of the world-line is the velocity. Since photon move with the speed of light $c = 1$, light in spacetime diagrams moves at an angle of $\pi/4$.

Definition 3.5 Proper time. Choose a point p on an object's world-line to be at $\tau = 0$. Let τ be the arc-length away from p

$$\tau = \int \sqrt{g_{ab} dX^a dX^b} \tag{3.6}$$

The arc-length τ along a world-line is called proper time.

Definition 3.6 Event. A spacetime event p is a point in spacetime.

Suppose at an event p a light signal is sent. This signal corresponds to an expanding sphere of light in space. In a spacetime diagram this looks as follows:

use.eps

In a spacetime diagram the expanding sphere traces out a cone, the light cone.

Definition 3.7 Let p and q be two events. The interval pq is called light-like if $\Delta\tau^2 = 0$, q is on the past light cone of p. The interval pq is space-like is $\Delta\tau^2 < 0$ and time-like if $\Delta\tau^2 > 0$.

Since all massive objects move slower than c, they travel within the future light cone. They can only reach time-like parts.

A particle with 3-velocity $\mathbf{v} = (dx/dt, dy/dt, dz/dt)$ satisfies

$$|\mathbf{v}|^2 dt^2 = dx^2 + dy^2 + dz^2 \tag{3.7}$$

For a photon, on the other hand, we have $|\mathbf{v}| = 1$ and hence

$$dt^2 = dx^2 + dy^2 + dz^2 \quad \Leftrightarrow \quad dt^2 - dx^2 - dy^2 - dz^2 = 0 \tag{3.8}$$

If we define an interval between two points by $d\tau$, we have

$$d\tau^2 = dt^2 - dx^2 - dy^2 - dz^2 = 0 \tag{3.9}$$

Note that $d\tau^2$ is precisely the line element.

Let $X^a(\tau)$ be a world-line and τ proper time. Let $T^a = dX^a/d\tau$ be the tangent vector to the world-line. Then we find

$$\eta_{ab} X^a X^b = \left(\frac{dt}{d\tau}\right)^2 - \left(\frac{dx}{d\tau}\right)^2 - \left(\frac{dy}{d\tau}\right)^2 - \left(\frac{dz}{d\tau}\right)^2 \tag{3.10}$$

$$= \frac{dt^2 - dx^2 - dy^2 - dz^2}{d\tau^2} = \frac{ds^2}{d\tau^2} = 1 \tag{3.11}$$

3.3 Lorentz transformations

Consider Minkowski spacetime $\mathbb{M}^4$

$$ds^2 = \eta_{ab} dX^a dX^b \tag{3.12}$$

There exit coordinate transformation that leave the line element invariant

$$X'^a = L^a{}_b X^b + A^a \tag{3.13}$$

These are called Poincare transformations or inhomogeneous Lorentz transformations. The vector A^a changes the origin, translations, while $L^a{}_b$ leaves the origin unchanged, rotations. When $A^a = 0$ these are called the homogeneous Lorentz transformations.

The set of all η_{ab} preserving transformations is called the Poincare group. Likewise, the group of η_{ab} preserving transformations which leave the origin fixed is called Lorentz group. If we consider

$$\eta_{ab} = L^c{}_a L^d{}_b \eta_{cd} \tag{3.14}$$

and take the determinant of both sides, we get

$$(\det L)^2 = 1 \tag{3.15}$$

and therefore $\det L = \pm 1$.

Definition 3.8 Proper and improper Lorentz transformations. The proper Lorentz transformations satisfy $\det L = +1$, while the improper ones are defined by $\det L = -1$.

3.4 Lorentz boosts

Suppose a spaceship moves with velocity v in the x-direction with respect to the Earth. We have the ship's rest frame S and Earth's rest frame E. We assume their origins coincide at an event p. What is the form of the transformation

$$\begin{pmatrix} t \\ x \end{pmatrix}_S = \begin{pmatrix} \gamma & \delta \\ \mu & \nu \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix}_E \tag{3.16}$$

(i) speed of light is an absolute constant

use.eps

the right moving photon passes through $(t, x) = (t, t)$ while the left moving photon passes through $(t, x) = (t, -t)$.

Suppose in the Earth frome a photon passes through the event

$$\begin{pmatrix} t \\ x \end{pmatrix}_E = \begin{pmatrix} t_0 \\ t_0 \end{pmatrix}_E \tag{3.17}$$

In the coordinates of the ship

$$\begin{pmatrix} t \\ x \end{pmatrix}_S = \begin{pmatrix} t \\ t \end{pmatrix} \begin{pmatrix} \gamma & \delta \\ \mu & \nu \end{pmatrix} \begin{pmatrix} t_0 \\ t_0 \end{pmatrix}_E \tag{3.18}$$

$$\Rightarrow \gamma + \delta = \mu + \nu \tag{3.19}$$

This must also hold for the left moving photons and hence

$$\begin{pmatrix} t \\ x \end{pmatrix}_S = \begin{pmatrix} t \\ -t \end{pmatrix} \begin{pmatrix} \gamma & \delta \\ \mu & \nu \end{pmatrix} \begin{pmatrix} t_0 \\ -t_0 \end{pmatrix}_E \tag{3.20}$$

$$\Rightarrow \gamma - \delta = \nu - \mu \tag{3.21}$$

Both results combine to $\gamma = \nu$ and $\delta = \mu$.

(ii) Follow the spatial origin in the ship's coordinates. On the ship $(t, x)_S = (t, 0)_S$. However, from Earth we see it move with velocity v and so $(t, vt)_E$, hence

$$\begin{pmatrix} t \\ 0 \end{pmatrix}_S = \begin{pmatrix} \gamma & \delta \\ \delta & \gamma \end{pmatrix} \begin{pmatrix} t \\ vt \end{pmatrix}_E \tag{3.22}$$

$$\Rightarrow \delta = -\gamma v \tag{3.23}$$

and so the transformation matrix takes the form

$$\begin{pmatrix} \gamma & -\gamma v \\ -\gamma v & \gamma \end{pmatrix} \tag{3.24}$$

(iii) Finally we assume proper Lorentz transformations

$$\det \begin{pmatrix} \gamma & -\gamma v \\ -\gamma v & \gamma \end{pmatrix} = \gamma^2 - \gamma^2 v^2 = \gamma^2(1 - v^2) = 1 \tag{3.25}$$

which results in the famous γ factor

$$\gamma = \frac{1}{\sqrt{1 - v^2}} \tag{3.26}$$

Note that we are working with units where $c = 1$.

3.5 Relativistic dynamics

The 3-momentum of a particle is defined as $\mathbf{p} = m\mathbf{v}$. For an object travelling at speed $\mathbf{v}$ a Lorentz transformation from the rest frame of the object gives

$$u^a = \begin{pmatrix} \gamma \\ \gamma \mathbf{v} \end{pmatrix} \tag{3.27}$$

One verifies that $u^a u_a = 1$. In classical mechanics the 3-momentum is conserved, and the energy is conserved.

Definition 3.9 4-momentum. The 4-momentum p_a (its covariant form) is defined by

$$p_a = m u_a = (E, -\mathbf{p}) \tag{3.28}$$

and therefore

$$p^a = \begin{pmatrix} \gamma m \\ \gamma m \mathbf{v} \end{pmatrix} \tag{3.29}$$

Let us consider the energy E for velocities satisfying $v \ll 1$ ($v \ll c$) which we call non-relativistic

$$E = p^0 = \gamma m = m \frac{1}{\sqrt{1 - v^2}} = m + \frac{1}{2} m v^2 + O(v^4) \tag{3.30}$$

which encodes Einstein's famous equation

$$E = mc^2 \tag{3.31}$$

For massive particles one finds $E^2 = |\mathbf{p}|^2 + m^2$.

Definition 3.10 Newton's force law. In special relativity Newton's force law becomes

$$f^b = ma^b \tag{3.32}$$

where a^b is the 4-acceleration defined by

$$a^b = \frac{dT^b}{d\tau} = \frac{d^2 X^b}{d\tau^2} \tag{3.33}$$

Lemma 3.1

$$a_b T^b = 0 \tag{3.34}$$

Proof

$$a_b T^b = \eta_{bc} a^b T^c = \eta_{bc} \frac{dT^b}{d\tau} T^c = \eta_{bc} \frac{1}{2} \frac{d}{d\tau} (T^b T^c) = \frac{1}{2} \frac{d}{d\tau} (\eta_{bc} T^b T^c) = \frac{1}{2} \frac{d}{d\tau} (1) = 0 \tag{3.35}$$

$\square$

In the limit $v \to 1$ the γ factor diverges. However, since photons are massless one still has a well defined 4-momentum. Let $v \to 1$ while $m \to 0$, keeping $E = \gamma m$ constant

$$p_a = (\gamma m, -\gamma m \mathbf{v}) = (E, -E\mathbf{v}) \tag{3.36}$$

Note that for massless particles $|\mathbf{p}|^2 = m^2$. The 3-vector $\mathbf{v}$ is now a unit vector determining the direction of the travelling photon.

Important: The world line of a photon cannot be parametrised by proper time τ since proper time down not exits for a photon. Along the path of a photon $d\tau = 0$. However, other parameters can be used, for example the coordinate time t is some reference frame.

3.6 Maxwell's equations

The internal structure equations are given by

$$\nabla \cdot \mathbf{B} = 0 \tag{3.37}$$
$$\nabla \times \mathbf{E} + \partial_t \mathbf{B} = 0 \tag{3.38}$$

while the source equations read

$$\nabla \cdot \mathbf{E} = 4\pi\rho \tag{3.39}$$

$$\nabla \times \mathbf{B} - \partial_t \mathbf{E} = \mathbf{J} \tag{3.40}$$

The Lorentz force equation reads

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{3.41}$$

Recall the electric potential ϕ and the vector potential $\mathbf{A}$ can be used to define the electric and magnetic fields respectively

$$\mathbf{E} = -\nabla\phi - \partial_t \mathbf{A} \tag{3.42}$$

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{3.43}$$

In order to define Maxwell's equations in tensor form, we firstly define the Faraday tensor

$$F_{ab} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{pmatrix} \tag{3.44}$$

The Faraday tensor is a skew-symmetric tensor of type $\binom{0}{2}$.

A skew-symmetric tensor of type $\binom{0}{2}$ is often called a 2-form. Likewise, a totally skew-symmetric tensor of type $\binom{0}{p}$ is called p-form, more precisely the components of a p-form. We will not use this terminology throughout the lectures.

In tensor form, or covariant form, the structure equations are given by

$$\partial_a F_{bc} + \partial_b F_{ca} + \partial_c F_{ab} = 0 \tag{3.45}$$

while the source equations can be written

$$\partial_b F^{ab} = j^a \tag{3.46}$$

where

$$j^a = \begin{pmatrix} \rho \\ \mathbf{J} \end{pmatrix} \tag{3.47}$$

and $F^{ab} = \eta^{ac}\eta^{bd}F_{cd}$, in special relativity. On a generic manifold M we raise and lower indices with the metric g_{ab}.

The charge conservation equation immediately follows from (3.46)

$$\partial_a \partial_b F^{ab} = 0 = \partial_a j^a \tag{3.48}$$

where the first terms vanished because a symmetric and a skew-symmetric tensor come together.

In tensor notation the Lorentz force law becomes

$$f^a = q u_b F^{ba} \tag{3.49}$$

As above, it easily follows that force and velocity are orthogonal

$$u_a f^a = q u_a u_b T^{ba} = 0 \tag{3.50}$$

By defining the 4-potential

$$A^b = \begin{pmatrix} \phi \\ \mathbf{M} \end{pmatrix} \quad \text{or} \quad A_b = \begin{pmatrix} \phi \\ -\mathbf{M} \end{pmatrix} \tag{3.51}$$

the Faraday tensor simply becomes

$$F_{ab} = A_{a,b} - A_{b,a} \tag{3.52}$$

Note that there are two different conventions, the other being

$$F_{ab} = \partial_a A_b - \partial_b A_a \tag{3.53}$$

Maxwell's equations are invariant under gauge transformations

$$A^b \mapsto A^b + \partial^b \psi \tag{3.54}$$

3.7 Stress-energy-momentum tensors

Continuous matter distributions in spacial relativity are described by a symmetric tensor T_{ab} called the stress-energy-momentum tensor, or energy-momentum tensor or stress-energy tensor.

Definition 3.11 The stress-energy tensor of the electromagnetic field is given by

$$T_{ab} = \frac{1}{4\pi} \left(F_{ac} F_b{}^c - \frac{1}{4} \eta_{ab} F_{mn} F^{mn} \right) \tag{3.55}$$

Note that $\partial^a T_{ab} = 0$ by virtue of Maxwell's equations.

Let an observer have 4-velocity V^a, the quantity

$$T_{ab}V^aV^b \tag{3.56}$$

is interpreted as the energy density. For normal matter this quantity will be non-negative

$$T_{ab}V^aV^b \geq 0 \tag{3.57}$$

If W^a is orthogonal to V^a, the quantity

$$-T_{ab}V^aW^b \tag{3.58}$$

is interpreted as the momentum density of the matter in the W^b direction.

Definition 3.12 Perfect fluid. A perfect fluid is defined to be a continuous distribution of matter with energy-momentum tensor

$$T_{ab} = \rho u_a u_b - p(\eta_{ab} - u_a u_b) \tag{3.59}$$

where u_a is a unit time-like vector representing the 4-velocity of the fluid. ρ is the energy density of the fluid and p is its pressure as measured in its rest frame. It satisfies the equations of motion

$$\partial^a T_{ab} = 0 \tag{3.60}$$

Although no scalar field has been observed so far in nature, it is often instructive to consider a scalar field ϕ satisfying the Klein-Gordon equations

$$\partial^a \partial_a \phi + m^2 \phi = 0 \tag{3.61}$$

Definition 3.13 The energy-momentum tensor of a scalar field is given by

$$T_{ab} = -\partial_a\phi\partial_b\phi + \eta_{ab}\frac{1}{2}\left(\eta^{cd}\partial_c\phi\partial_d\phi - m^2\phi^2\right) \tag{3.62}$$

To finish this section, we will work out explicitely the equations of motion of the perfect fluid and consider its non-relativistic limit.

Let us write out equations (3.60) explicitely

$$\begin{aligned}\partial^a \rho u_a u_b + \rho u_a \partial^a u_b + \rho u_b \partial^a u_a& \\ - \partial^a p(\eta_{ab} - u_a u_b) + p u_a \partial^a u_b + p u_b \partial^a u_a &= 0\end{aligned} \tag{3.63}$$

We multiply this equation by u^b which yields

$$u^a \partial_a \rho + (\rho + p)\partial^a u_a = 0 \tag{3.64}$$

The terms that vanish after multiplication are orthogonal to the other terms and hence must vanish separately

$$(\rho + p)u_a \partial^a u_b - (\eta_{ab} - u_a u_b)\partial^a p = 0 \tag{3.65}$$

Now we consider the non-relativistic limit

$$v \ll 1 \quad p \ll \rho \quad u^a = \begin{pmatrix} 1 \\ \mathbf{v} \end{pmatrix} \quad |\mathbf{v}|\frac{dp}{dt} \ll |\nabla p| \tag{3.66}$$

Then equation (3.64) becomes

$$u^a \partial_a \rho + \rho \partial^a u_a = \partial_t \rho + \mathbf{v} \cdot \nabla \rho + 0 + \rho \nabla \cdot \mathbf{v} = \partial_t \rho + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{3.67}$$

which is the hydrodynamical conservation equation of mass, well known in the mechanics of fluids.

The second equation (3.65) in the non-relativistic limit is

$$\rho u^a \partial_a u_b - (\delta^a_b - u_b u^a)\partial_a p = 0 \tag{3.68}$$

The time component is identically satisfied, so we consider the spatial components

$$\rho \partial_t(-\mathbf{v}) + \rho \mathbf{v} \cdot \nabla(-\mathbf{v}) - \nabla p + (-\mathbf{v})[\partial_t p + \mathbf{v} \cdot \nabla p] = 0 \tag{3.69}$$

Since the last two terms are of higher order, we are left with

$$\rho\left(\partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla)\mathbf{v}\right) + \nabla p = 0 \tag{3.70}$$

which is Euler's equation of hydrodynamic.

4 Curvature

4.1 Covariant derivative ad parallel transport

Definition 4.1 Covariant derivative. A covariant derivative ∇_a (sometimes called derivative operator) on a manifold M is a map which takes a type $\binom{p}{q}$ tensor to a type $\binom{p}{q+1}$ tensor. It satisfies the following properties:

i linearity: for all $\alpha, \beta \in \mathbb{R}$

$$\nabla_c(\alpha A^{a_1 \dots a_m}{}_{b1 \dots b_n} + \beta B^{a_1 \dots a_m}{}_{b1 \dots b_n}) = \alpha \nabla_c A^{a_1 \dots a_m}{}_{b1 \dots b_n} + \beta \nabla_c B^{a_1 \dots a_m}{}_{b1 \dots b_n} \tag{4.1}$$

ii Leibnitz rule:

$$\nabla_c(A^{\cdots}{}_{\dots} B^{\cdots}{}_{\dots}) = B^{\cdots}{}_{\dots} \nabla_c A^{\cdots}{}_{\dots} + A^{\cdots}{}_{\dots} \nabla_c B^{\cdots}{}_{\dots} \tag{4.2}$$

iii commutativity with contraction:

$$\nabla_c(\alpha A^{a_1 \dots k \dots a_m}{}_{b1 \dots k \dots b_n}) = \nabla_c \alpha A^{a_1 \dots k \dots a_m}{}_{b1 \dots k \dots b_n} \tag{4.3}$$

iv torsion free: for all $f \in C^\infty(M)$

$$\nabla_a \nabla_b f = \nabla_b \nabla_a f \tag{4.4}$$

In general relativity the covariant derivative is always assumed to be torsion free. In Einstein-Cartan theory for example this assumption is dropped.

In Euclidean 3-space with Cartesian coordinates, the covariant derivative should correspond to the familiar partial derivative. Moreover, for any smooth function f, the covariant derivative should coincide with the partial derivative

$$\nabla_a f = \partial_a f = f_{,a} \tag{4.5}$$

One easily verifies that $\partial_a f$ transforms like a tensor.

Given a covariant derivative ∇_a, its action on a vector A^a (or on any tensor $A^{a_1 \dots a_m}{}_{b_1 \dots b_n}$) should depend only on the value of quantities at some point p. Let us consider

$$\nabla_a A^b \tag{4.6}$$

We know that $\partial_a A^b$ does not transform as a tensor. Together with the above locality assumption, the difference between $\nabla_a A^b$ and $\partial_a A^b$ must be expressible in terms of A^a, therefore let us write

$$\nabla_a A^b - \partial_a A^b = C^b_{ac} A^c \tag{4.7}$$

Property (iv) of the covariant derivative implies that C^b_{ac} is symmetric in the lower pair of indices

$$C^b_{ac} = C^b_{ca} \tag{4.8}$$

Moreover, since $\nabla_a A^b$ is a tensor by definition, however $\partial_a A^b$ does not transform like a tensor, we conclude that C^b_{ac} cannot be a tensor.

Rewriting (4.7) as follows

$$\nabla_a A^b = \partial_a A^b + C_b^{ac} A^c \tag{4.9}$$

implies that the right-hand side must transform like a tensor. The inhomogeneous parts of the individual transformations must cancel each other.

Since ∇_a when acting on scalars equals the partial derivative, let us consider the scalar $A_b A^b$

$$\begin{aligned}\nabla_a(A_b A^b) &= A_b \nabla_a A^b + \nabla_a A_b A^b \\ &= A_b(\partial_a A^b + C_b^{ac} A^c) + \nabla_a A_b A^b = \partial_a(A_b A^b) = A_b \partial_a A^b + \partial_a A_b A^b\end{aligned} \tag{4.10}$$

Now we rewrite this equation

$$C^b_{ac} A^c A_b + A^b \nabla_a A_b = A^b \partial_a A_b \tag{4.11}$$

$$A^b \nabla_a A_b = A^b \partial_a A_b - C^b_{ac} A^c A_b \tag{4.12}$$

$$A^b \nabla_a A_b = A^b \partial_a A_b - C^c_{ab} A^b A_c \tag{4.13}$$

and therefore we arrive at

$$\nabla_a A_b = \partial_a A_b - C^c_{ab} A_c \tag{4.14}$$

Since we now know how ∇_a acts on a contravariant and on a covariant index, we can compute the covariant derivative on any type $\binom{p}{q}$ tensor

$$\begin{aligned}\nabla_c A^{a_1\ldots a_p}{}_{b_1\ldots b_q} = \partial_c A^{a_1\ldots a_p}{}_{b_1\ldots b_q} \\ + \Gamma^{a_1}_{ck} A^{k\ldots a_p}{}_{b_1\ldots b_q} + \cdots + \Gamma^{a_p}_{ck} A^{a_1\ldots k}{}_{b_1\ldots b_q} \\ - \Gamma^k_{cb_1} A^{a_1\ldots a_p}{}_{k\ldots b_q} - \cdots - \Gamma^k_{cb_q} A^{a_1\ldots a_p}{}_{b_1\ldots k}\end{aligned} \tag{4.15}$$

Theorem 4.1 *Let M be a manifold and g_{ab} be a metric. Then there exists a unique covariant derivative operator satisfying*

$$\nabla_a g_{bc} = 0 \tag{4.16}$$

which is called the metricity condition.

Proof Idea of the proof. Use the action of ∇_a on g_{bc} and try to solve the resulting equation for C^a_{bc}.

$$\nabla_a g_{bc} = g_{bc,a} - C^d_{ab} g_{dc} - C^d_{ac} g_{bd} = 0 \tag{4.17}$$

$$\nabla_c g_{ab} = g_{ab,c} - C^d_{ca} g_{db} - C^d_{cb} g_{ad} = 0 \tag{4.18}$$

$$\nabla_b g_{ca} = g_{ca,b} - C^d_{bc} g_{da} - C^d_{ba} g_{cd} = 0 \tag{4.19}$$

Let us consider the following combination (4.17) + (4.17) − (4.17) of the metricity conditions

$$g_{bc,a} + g_{ab,c} - g_{ca,b} - 2C^d_{ca} g_{db} = 0 \tag{4.20}$$

Apply g^{bm} to this equation and we find

$$\delta^m_d C^d_{ca} = \frac{1}{2} g^{mb}(g_{bc,a} + g_{ab,c} - g_{ca,b}) \tag{4.21}$$

Finally we rename indices $c \to b, a \to c, m \to a, b \to d$ and arrive at

$$C^a_{bc} = \Gamma^a_{bc} = \frac{1}{2} g^{ad}(g_{db,c} + g_{cd,b} - g_{bc,d}) \tag{4.22}$$

Hence, C^a_{bc} is uniquely fixed to be the Christoffel symbol and therefore ∇_a is unique. Note that Γ^a_{bc} is often called connection. □

Definition 4.2 Parallel transport. Let ∇_a be a covariant derivative and γ be a curve with tangent vector T^a. A vector V^a at each point on the curve is said to be parallelly transported along γ if

$$T^a \nabla_a V^b = 0 \tag{4.23}$$

is satisfied along the curve. Parallel transport of an arbitrary type $\binom{p}{q}$ tensor is defined by

$$T^a \nabla_a A^{a_1 \ldots a_p}{}_{b_1 \ldots b_q} = 0 \tag{4.24}$$

Using the definition of the covariant derivative we find explicitely

$$T^a \partial_a V^b + T^a \Gamma^b_{ac} V^c = 0 \tag{4.25}$$

If the curve γ is parametrised by τ so that we have $X^a = X^a(\tau)$, then the tangent vector is given by $T^a = dX^a/d\tau$. Hence, we can write

$$T^a \partial_a V^b = \frac{dX^a}{d\tau} \frac{dV^b}{dX^a} = \frac{dV^a}{d\tau} \tag{4.26}$$

and therefore

$$\frac{dV^b}{d\tau} + T^a \Gamma^b_{ac} V^c = 0 \tag{4.27}$$

This is a first order ordinary differential equation and has a unique solution for given initial value of V^a. Thus, a vector at one point on γ uniquely defines the parallelly transported vector everywhere else on the curve.

Lemma 4.1 *Let V^a and W^a be two vectors parallelly transported along a curve γ. Then the scalar $V_a W^a$ remains unchanged if parallelly transported along γ.*

Proof To show that $V_a W^a$ is constant along the curve, let us consider the quantity

$$T^a \nabla_a (V_b W^b) \tag{4.28}$$

We can rewrite this as follows

$$T^a \nabla_a (g_{bc} V^b W^c) = T^a V^b W^c \nabla_a g_{bc} + T^a g_{bc} W^c \nabla_a V^b + T^a g_{bc} V^b \nabla_a W^c \tag{4.29}$$

We order the terms such that

$$T^a V^b W^c \nabla_a g_{bc} + g_{bc} W^c T^a \nabla_a V^b + g_{bc} V^b T^a \nabla_a W^c \overset{!}{=} 0 \tag{4.30}$$

The first term vanishes in view of Theorem 4.1, the second term and the third term both vanish since we assume that V^a and W^a are parallelly transported. Hence $V_b W^b$ is constant along the curve. □

4.2 Riemann curvature tensor

From the definition of the covariant derivative it follows that $\nabla_a \nabla_b$ commutes when acting on scalars

$$(\nabla_a \nabla_b - \nabla_b \nabla_a) f = 0 \tag{4.31}$$

However, it does not commute, when acting on vectors. Let us work out this commutator

$$\begin{aligned}\nabla_a \nabla_b A^c = \nabla_a (\partial_b A^c + \Gamma^c_{bd} A^d) \\ &= \partial_{ab} A^c - \Gamma^d_{ab} \partial_d A^c + \Gamma^c_{ad} \partial_b A^d \\ &\quad + \partial_a (\Gamma^c_{bd} A^d) - \Gamma^e_{ab} \Gamma^c_{ed} A^d + \Gamma^c_{ae} \Gamma^e_{bd} A^d \end{aligned} \tag{4.32}$$

Exchange the indices a and b

$$\begin{aligned}\nabla_b \nabla_a A^c = \partial_{ba} A^c - \Gamma^d_{ba} \partial_d A^c + \Gamma^c_{bd} \partial_a A^d \\ + \partial_b (\Gamma^c_{ad} A^d) - \Gamma^e_{ba} \Gamma^c_{ed} A^d + \Gamma^c_{be} \Gamma^e_{ad} A^d\end{aligned} \tag{4.33}$$

Since we wish to compute $(\nabla_a \nabla_b - \nabla_b \nabla_a) A^c$ we note that terms 1, 2 and 5 cancel each other. Let us collect the remaining terms and expand the partial derivatives

$$\begin{aligned}(\nabla_a \nabla_b - \nabla_b \nabla_a) A^c = \Gamma^c_{ad} \partial_b A^d + \partial_a \Gamma^c_{bd} A^d + \Gamma^c_{bd} \partial_a A^d \\ - \Gamma^c_{bd} \partial_a A^d - \partial_b \Gamma^c_{ad} A^d + \Gamma^c_{ad} \partial_b A^d \\ + \Gamma^c_{ae} \Gamma^e_{bd} A^d - \Gamma^c_{be} \Gamma^e_{ad} A^d\end{aligned} \tag{4.34}$$

Terms 1 and 6 cancel each other and so do terms 3 and 4, hence we have

$$\begin{aligned}(\nabla_a \nabla_b - \nabla_b \nabla_a) A^c &= (\partial_a \Gamma^c_{bd} - \partial_b \Gamma^c_{ad} + \Gamma^c_{ae} \Gamma^e_{bd} - \Gamma^c_{be} \Gamma^e_{ad}) A^d \\ &= (\Gamma^c_{bd,a} - \Gamma^c_{ad,b} + \Gamma^e_{bd} \Gamma^c_{ea} - \Gamma^e_{ad} \Gamma^c_{eb}) A^d\end{aligned} \tag{4.35}$$

and we define

$$(\nabla_a \nabla_b - \nabla_b \nabla_a) A^c = R_{bad}{}^c A^d \tag{4.36}$$

$$R_{bad}{}^c := \Gamma^c_{bd,a} - \Gamma^c_{ad,b} + \Gamma^e_{bd} \Gamma^c_{ea} - \Gamma^e_{ad} \Gamma^c_{eb} \tag{4.37}$$

where $R_{bad}{}^c$ is called the Riemann curvature tensor.

Likewise, when acting on a covariant vector we find

$$(\nabla_a \nabla_b - \nabla_b \nabla_a) V_c = R_{abc}{}^d A_d \tag{4.38}$$

To calculate $R_{abc}{}^d$ starting from the metric g_{ab}, we first calculate the Christoffel symbol components Γ^c_{ab}, and from there one computed the the Riemann curvature tensor. Before further discussing the Riemann curvature tensor, we show its geometrical significance.

Let us consider a parallelly displaced vector Y^c along a curve γ

use.eps

For infinitesimal displacements we find $Y^c(X^a)$ at the point p, while at the point q we have

$$\begin{aligned} Y^c(X^a + dX^a) &= Y^c(X^a) + dX^a \frac{\partial Y^c}{\partial X^a}(X^a) + O(dX^a)^2 \\ &= Y^c(X^a) + dX^a \nabla_a Y^c - dX^a \Gamma^c_{ab} Y^b + O(dX^a)^2 \end{aligned} \tag{4.39}$$

The second term vanished because of our parallel transport assumption and hence

$$\delta Y^c = Y^c_{(q)} - Y^c_{(p)} = -dX^a \Gamma^c_{ab} Y^b \tag{4.40}$$

in lowest order approximation.

Let us now consider parallelly transporting a vector A^a around a small closed loop.

use.eps

Let us firstly transport the vector A^b from p_0 to p_1. At p_1 the displaced vector is given by

$$A^b + \delta A^b = A^b_{p_0} - dX^a \Gamma^b_{ac} A^c_{p_0} \tag{4.41}$$

Next, parallel transport of the transported vector to $\bar{p}$, along the direction $d\bar{X}^a$

$$\begin{aligned} \delta \bar{A}^b &= -d\bar{X}^a \Gamma^b_{ac\,(p_1)} (A^c + \delta A^c) \simeq -d\bar{X}^a (\Gamma^b_{ac} + \Gamma^b_{ac,d} dX^d)(A^c + \delta A^c) \\ &\simeq -d\bar{X}^a \Gamma^b_{ac} A^c - d\bar{X}^a \Gamma^b_{ac,d} dX^d A^c + d\bar{X}^a dX^d \Gamma^b_{ac} \Gamma^c_{de} A^e \end{aligned} \tag{4.42}$$

where we kept terms to second order only. If we now consider the other way to reach the point $\bar{p}$ we obtain

$$\delta \tilde{A}^b \simeq -dX^a \Gamma^b_{ac} A^c - dX^a \Gamma^b_{ac,d} d\bar{X}^d A^c + dX^a d\bar{X}^d \Gamma^b_{ac} \Gamma^c_{de} A^e \tag{4.43}$$

Finally, we are interested in the difference between the vector transported the one or the other way

$$\begin{aligned} \Delta A^b = (A^b + \delta \bar{A}^b) - (A^b + \delta \tilde{A}^b) &= (dX^a - d\bar{X}^a) \Gamma^b_{ac} A^c \\ &-d\bar{X}^a dX^d A^e \left[\Gamma^b_{ae,d} - \Gamma^b_{de,a} + \Gamma^b_{ac} \Gamma^c_{de} - \Gamma^b_{dc} \Gamma^c_{ae} \right] \end{aligned} \tag{4.44}$$

The term in the square bracket is again the Riemann curvature tensor and therefore we can write

$$\Delta A^b = (dX^a - d\bar{X}^a) \Gamma^b_{ac} A^c - d\bar{X}^a dX^d R_{ade}{}^b A^e \tag{4.45}$$

Assume we transport dX^a along $d\bar{X}^a$ and vice versa. Then for the term we find

$$dX^a \Gamma^b_{ac} d\bar{X}^c - d\bar{X}^a \Gamma^b_{ac} dX^c = dX^a d\bar{X}^c (\Gamma^b_{ac} - \Gamma^b_{ca}) \stackrel{!}{=} 0 \tag{4.46}$$

because the connection is assumed to be symmetric. This means that infinitesimal parallelograms always close! On manifolds with torsion this no longer holds.

Therefore we have obtained that the Riemann curvature tensor measures the path dependence of parallel transport. This path dependence allows to define an intrinsic notion of curvature. The failure of a parallel transported vector around a closed loop coinciding with the original vector measures the curvature (the failure of pointing in the same direction after transport).

The definition of the Riemann curvature tensor implies symmetry properties which are if great importance. In particular the twice contracted

Bianchi identities yield the left-hand side (geometrical side) of the field equations of general relativity.

In n dimensions the Riemann curvature tensor has

$$\frac{1}{12}n^2(n^2-1) \tag{4.47}$$

independent components.

Lemma 4.2 *The Riemann curvature tensor has the following properties:*

i $R_{abcd} = -R_{bacd}$,

ii $R_{abcd} + R_{cabd} + R_{bcab} = 0$,

iii $R_{abcd} = -R_{abdc}$,

iv $\nabla_e R_{abcd} + \nabla_d R_{abec} + \nabla_c R_{abde} = 0$. *This is the famous Bianchi identity.*

Proof These identities can be proved more or less straightforwardly

i trivial by definition

ii To prove this identity we consider the permutations of $\nabla_a \nabla_b W_c$ and write

$$\nabla_a \nabla_b W_c - \nabla_b \nabla_a W_c = R_{abc}{}^d W_d \tag{4.48}$$

$$\nabla_c \nabla_a W_b - \nabla_a \nabla_c W_b = R_{cab}{}^d W_d \tag{4.49}$$

$$\nabla_b \nabla_c W_a - \nabla_c \nabla_b W_a = R_{bca}{}^d W_d \tag{4.50}$$

and add these three equations up. Firstly, observe that

$$\nabla_a \nabla_b W_c - \nabla_a \nabla_c W_b = \nabla_a (\partial_b W_c - \partial_c W_b) \tag{4.51}$$

and let us denote $T_{bc} = \partial_b W_c - \partial_c W_b$. Then the left-hand side of the added up equations becomes

$$\nabla_a T_{bc} + \nabla_b T_{ca} + \nabla_c T_{ab} \tag{4.52}$$

Due to the skew-symmetry of T_{ab} the Christoffel symbols drop out the the covariant derivative become partial derivative. As these commute, we find that the left-hand side vanished and so the identity follows.

iii We can use the fact that $\nabla_a g_{bc} = 0$ and write

$$\begin{aligned} 0 &= (\nabla_a \nabla_b - \nabla_b \nabla_a) g_{cd} = R_{abc}{}^e g_{ed} + R_{abd}{}^e g_{ce} \\ &= R_{abcd} + R_{abdc} \end{aligned} \tag{4.53}$$

iv no proof for now

□

Properties (i)–(iii) imply another symmetry, namely

$$R_{abcd} = R_{cdab} \tag{4.54}$$

4.3 Ricci, Weyl and Einstein tensor

The Riemann tensor can be decomposed in a 'trace part' and a 'trace-free part.' Since the Riemann tensor is skew-symmetric in the first and second pair of indices, we can trace over the first and third (or second and fourth) index. This defined the

Definition 4.3 Ricci tensor

$$R_{ab} = R_{acb}{}^c \tag{4.55}$$

Note that the symmetry properties of the Riemann tensor imply that the Ricci tensor is symmetric.

Definition 4.4 Scalar curvature. The trace of the Ricci tensor is the scalar curvature or the Ricci scalar

$$R = R_a{}^a \tag{4.56}$$

Definition 4.5 Weyl tensor. The 'trace-free part' of the Riemann tensor if the so called Weyl tensor. For manifold of dimension $n \geq 3$ it is defined by

$$C_{abcd} = R_{abcd} - \frac{2}{n-2}(g_{a[c}R_{d]b} - g_{b[c}R_{d]a}) + \frac{2}{(n-1)(n-2)} R g_{a[c}g_{d]b} \tag{4.57}$$

The Weyl tensor is also called conformal tensor because it is invariant under conformal transformations of the metric $g_{ab} \mapsto \Omega^2(X^a)\, g_{ab}$

Let us consider the contracted (apply the metric once) Bianchi identity

$$\nabla_a R_{bcd}{}^a + \nabla_b R_{cd} - \nabla_c R_{bd} = 0 \tag{4.58}$$

and apply g^{bd} to that equation

$$\nabla_a R_c{}^a + \nabla_b R_c{}^b - \nabla_c R = 0 \tag{4.59}$$
$$2\nabla_a R_c{}^a - \nabla_c R = 0 \tag{4.60}$$
$$\nabla^a R_{ca} - \frac{1}{2}\nabla^b g_{bc} R = 0 \tag{4.61}$$
$$\nabla^a (R_{ca} - \frac{1}{2} R g_{ca}) = 0 \tag{4.62}$$

Definition 4.6 Einstein tensor. The tensor G_{ab} defined by

$$G_{ab} = R_{ab} - \frac{1}{2} g_{ab} R \tag{4.63}$$

is called the Einstein tensor. Its covariant derivative vanishes.

Using the definition of the Riemann and Ricci tensors, the Ricci tensor can be calculated directly from the Christoffel symbols

$$R_{mr} = \Gamma^n_{mr,n} - \Gamma^n_{nr,m} + \Gamma^a_{mr}\Gamma^n_{an} - \Gamma^a_{nr}\Gamma^n_{am} \tag{4.64}$$

4.4 Geodesics II

As already discussed in Subsection 2.3, geodesics are the 'straightest possible' line, they 'curve as little as possible.' Using the notion of parallel transport, we can give a very geometrical definition of geodesics.

Lemma 4.3 *Let ∇_a be a covariant derivative. A geodesic is a curve whose tangent vector is parallelly transported along itself, this means the tangent vector T^a satisfies*

$$T^a \nabla_a T^b = 0 \tag{4.65}$$

Proof Let γ be a curve with affine parametrisation $X^a(\lambda)$. The tangent vector to γ is given by $T^a = dX^a/d\lambda$. Moreover, $\nabla_a T^b$ can be written as follows

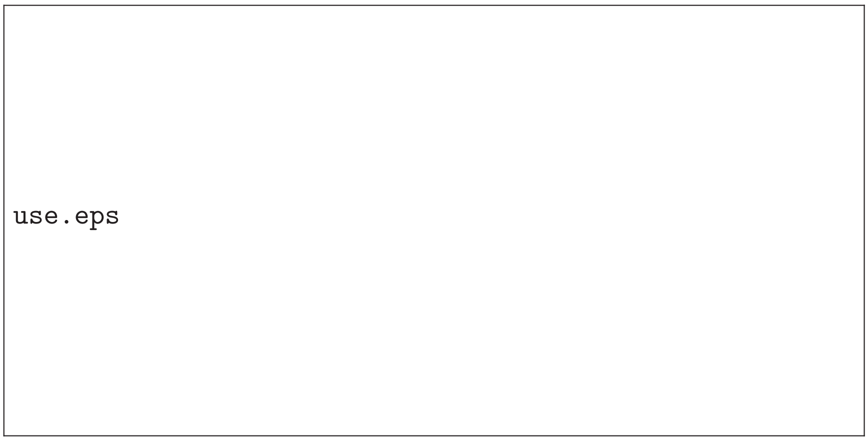
$$\nabla_a T^b = \partial_a T^b + \Gamma^b_{ac} T^c \tag{4.66}$$
$$T^a \partial_a T^b = T^a \frac{\partial T^b}{\partial X^a} = \frac{\partial X^a}{\partial \lambda}\frac{\partial T^b}{\partial X^a} = \frac{\partial T^b}{\partial \lambda} = \frac{\partial^2 X^b}{\partial \lambda^2} \tag{4.67}$$

and therefore the equation of parallel transport becomes

$$\frac{\partial^2 X^b}{\partial \lambda^2} + \Gamma^b_{ac} T^a T^c = \frac{\partial^2 X^b}{\partial \lambda^2} + \Gamma^b_{ac} \frac{\partial X^a}{\partial \lambda} \frac{\partial X^c}{\partial \lambda} = 0 \tag{4.68}$$

which is the geodesic equation. □

Another way to interpret the Riemann tensor is by considering a family of geodesics which span a surface $X^a = X^a(\tau, \lambda)$



The $\lambda =$ const. curves are geodesics parametrised by the affine parameter τ. Let $U^a = dX^a/d\tau$ be the tangent vector to the geodesic and let $N^a = dX^a/d\lambda$ be the displacement vector to an infinitesimally nearby geodesic. We can choose $U_a U^a = 1$, N^a and U^a are orthogonal $g_{ab} U^a N^b = 0$.

Since partial derivatives commute, we have

$$\frac{\partial^2 X^a}{\partial \tau \partial \lambda} = \frac{\partial^2 X^a}{\partial \lambda \partial \tau} \tag{4.69}$$

and hence we find

$$U^a N^b{}_{,a} = N^a U^b{}_a \tag{4.70}$$

Since the connection is symmetric this can re-written such that

$$U^a \nabla_a N^b = N^a \nabla_a U^b \tag{4.71}$$

The quantity $v^a = U^b \nabla_b N^a$ gives the rate of change along a geodesic of the displacement to a nearby geodesic. One can regards v^a as the relative velocity. Similarly

$$a^a = U^c \nabla_c u^a \tag{4.72}$$

can be interpreted as the relative acceleration of an infinitesimal nearby geodesic.

$$
\begin{aligned}
a^a &= U^c\nabla_c(U^b\nabla_b N^a) = U^c\nabla_c(N^b\nabla_b U^a) && (4.73)\\
&= U^c\nabla_c N^b\nabla_b U^a + U^c N^b\nabla_c\nabla_b U^a && (4.74)\\
&= N^c\nabla_c U^b\nabla_b U^a + N^b U^c\nabla_b\nabla_c U^a - R_{cbd}{}^a N^b U^c U^d && (4.75)\\
&= N^c\nabla_c(U^b\nabla_b U^a) - R_{cbd}{}^a N^b U^c U^d && (4.76)
\end{aligned}
$$

The first term vanishes as we are dealing with geodesics and hence

$$a^a = -R_{cbd}{}^a N^b U^c U^d \tag{4.77}$$

which is the geodesic deviation equation.

The acceleration vanishes for all families of geodesics if and only if $R_{abcd} = 0$. Some geodesics will accelerate toward or away from each other if and only if $R_{abcd} \neq 0$.

4.5 Einstein's field equations

In Newtonian gravity, the gravitational field may be represented by the potential Φ. The total acceleration of two nearby particles if given by

$$-(\mathbf{x}\cdot\nabla)\nabla\Phi \tag{4.78}$$

where $\mathbf{x}$ is the separation vector. Comparison with the geodesic deviation equation suggests

$$R_{cbd}{}^a U^c U^d \leftrightarrow \partial_b\partial^a\Phi \tag{4.79}$$

However, Poisson's equation read

$$\Delta\Phi = 4\pi\rho \tag{4.80}$$

Recalling the discussion of the energy-momentum tensors, we defined the energy density to be

$$T_{ab}U^a U^b = \rho \tag{4.81}$$

which seems to suggest

$$R_{cad}{}^a U^c U^d = 4\pi T_{cd}U^c U^d \tag{4.82}$$

and would indicate $R_{cd} = 4\pi T_{cd}$ as field equations. This was indeed originally suggested by Einstein, however is not divergence free. The energy-momentum tensor must satisfy $\nabla^a T_{ab} = 0$. However, the contracted Bianchi identities yielded such a tensor, namely the Einstein tensor. Hence, the field equations of general relativity are given by

$$G_{ab} := R_{ab} - \frac{1}{2} R g_{ab} = 8\pi T_{ab} \tag{4.83}$$

> *Matter tells spacetime how to curve and*
> *spacetime tells matter how to move.*

Since the covariant divergence of the Einstein tensor vanishes identically, these equations imply

$$\nabla^a T_{ab} = 0 \tag{4.84}$$

This is equivalent to saying that world liner of test bodies are geodesics. It is also a direct consequence of the field equations.

From now on we study the field equations.

Taking the trace of the Einstein equations (apply g^{ab}) yields

$$g^{ab}(R_{ab} - \frac{1}{2} R g_{ab}) = R - \frac{1}{2} \cdot 4R = -R \tag{4.85}$$

$$g^{ab} T_{ab} = T \tag{4.86}$$

and therefore we find

$$R_{ab} - \frac{1}{2} R g_{ab} = 8\pi T_{ab} \tag{4.87}$$

$$R_{ab} = 8\pi (T_{ab} - \frac{1}{2} T g_{ab}) \tag{4.88}$$

In the absence of matter $T_{ab} = 0$ and the vacuum field equations reduce to

$$R_{ab} = 0 \tag{4.89}$$

Manifolds satisfying the vacuum field equations are often called Ricci flat. This is very different from Riemann flat, where the full Riemann tensor vanishes. In general, solutions of the vacuum field equations are not Riemann flat.

As stated above, the field equations imply that test bodies follows geodesics

$$\frac{d^2 X^a}{d\lambda^2} + \Gamma^a_{bc} \frac{dX^b}{d\lambda} \frac{dX^c}{d\lambda} = 0 \tag{4.90}$$

Recall that these equations were obtained from the Lagrangian

$$L = g_{ab}\frac{dX^a}{d\lambda}\frac{dX^b}{d\lambda} \tag{4.91}$$

Note that for massive particles the affine parametrisation means we can choose

$$L = g_{ab}\frac{dX^a}{d\lambda}\frac{dX^b}{d\lambda} = \pm 1 \tag{4.92}$$

For massless particles moving with the speed of light we have $L = 0$ (null geodesics or null curves). In summary we have

$$L = \begin{cases} \pm 1 & m \neq 0 \\ 0 & m = 0 \end{cases} \tag{4.93}$$

± 1: the sign depends on the signature of the metric one works with

$$\text{signature } (-,+,+,+) \qquad -1 \tag{4.94}$$

$$\text{signature } (+,-,-,-) \qquad +1 \tag{4.95}$$

5 The Schwarzschild solutions

In order to test general relativity we are interested to find solutions of the vacuum field equations which describe the exterior gravitational field of a static and spherically symmetric body, like for example the Earth or our Sun.

The Schwarzschild solutions is the most important known exact solutions of the field equations.

5.1 Metric ansatz and Christoffel symbols

We want to find all 4-dimensional metrics with Lorentzian signature whose Ricci tensor vanishes and which are static and spherically symmetric. The most general static and spherically symmetric metric has the form

$$ds^2 = -e^{\nu(r)}dt^2 + e^{\lambda(r)}dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \tag{5.1}$$

which contains two unknown quantities ν and λ that are functions of the radial coordinate r.

The most efficient way to compute Christoffel symbol components is via the Lagrangian and the geodesic equation. We have $X^a = (t, r, \theta, \phi)$ with the Lagrangian given by $L = g_{ab}\dot{X}^a\dot{X}^b$, which yields

$$L = -e^{\nu}\dot{t}^2 + e^{\lambda}\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2 \tag{5.2}$$

Recall that the dot denotes differentiation with respect to an affine parameter and NOT differentiation with respect to time t. Also note that the four coordinates are functions of this affine parameter. The Euler-Lagrange equations read

$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q} \tag{5.3}$$

$$q = t: \qquad \frac{\partial L}{\partial t} = 0 \qquad \frac{\partial L}{\partial \dot{t}} = -2e^{\nu}\dot{t} \tag{5.4}$$

$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{t}} = -2\nu' e^{\nu}\dot{r}\dot{t} - 2e^{\nu}\ddot{t} \tag{5.5}$$

$$\Rightarrow \qquad \ddot{t} + \nu'\dot{r}\dot{t} = 0 \tag{5.6}$$

$$\Rightarrow \qquad \Gamma^t_{tr} = \frac{1}{2}\nu' \tag{5.7}$$

and all other components of the form Γ^t_{ab} vanish. The prime denote differentiation of the function with respect to its argument, so $\nu' = d\nu/dr$.

$$q = \phi: \qquad \frac{\partial L}{\partial \phi} = 0 \qquad \frac{\partial L}{\partial \dot{\phi}} = 2r^2\sin^2\theta\dot{\phi} \tag{5.8}$$

$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{t}} = 4r\dot{r}\sin^2\theta\dot{\phi} + 2r^2 2\sin\theta\cos\theta\dot{\theta}\dot{\phi} + 2r^2\sin^2\theta\ddot{\phi} \tag{5.9}$$

$$\Rightarrow \qquad \ddot{\phi} + 2\cot\theta\dot{\theta}\dot{\phi} + \frac{2}{r}\dot{r}\dot{\phi} = 0 \tag{5.10}$$

$$\Rightarrow \qquad \Gamma^{\phi}_{\phi r} = \frac{1}{r} \qquad \Gamma^{\phi}_{\theta\phi} = \cot\theta \tag{5.11}$$

all other components of the form Γ^{ϕ}_{ab} vanish.

$$q = \theta: \qquad \frac{\partial L}{\partial \phi} = r^2 2\sin\theta\cos\theta\dot{\phi}^2 \qquad \frac{\partial L}{\partial \dot{\phi}} = 2r^2\dot{\theta} \tag{5.12}$$

$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{\theta}} = 4r\dot{r}\dot{\theta} + 2r^2\ddot{\theta} \tag{5.13}$$

$$\Rightarrow \qquad \ddot{\theta} + \frac{2}{r}\dot{r}\dot{\theta} - \sin\theta\cos\theta\dot{\phi}^2 = 0 \tag{5.14}$$

$$\Rightarrow \qquad \Gamma^{\theta}_{\theta r} = \frac{1}{r} \qquad \Gamma^{\theta}_{\phi\phi} = -\sin\theta\cos\theta \tag{5.15}$$

all other components of the form Γ^θ_{ab} vanish.

Finally we consider $q = r$

$$q = r: \qquad \frac{\partial L}{\partial r} = -\nu' e^\nu \dot{t}^2 + \lambda' e^\lambda \dot{r}^2 + 2r\dot{\theta}^2 + 2r\sin^2\theta\dot{\phi}^2 \tag{5.16}$$

$$\frac{\partial L}{\partial \dot{r}} = 2e^\lambda \dot{r} \tag{5.17}$$

$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{r}} = 2\lambda' e^\lambda \dot{r}^2 + 2e^\lambda \ddot{r} \tag{5.18}$$

$$\Rightarrow \qquad \ddot{r} + \frac{1}{2}\lambda'\dot{r}^2 + \frac{1}{2}\nu' e^{\nu-\lambda}\dot{t}^2 - e^{-\lambda} r\dot{\theta}^2 - r\sin^2\theta e^{-\lambda}\dot{\phi}^2 = 0 \tag{5.19}$$

$$\Rightarrow \qquad \Gamma^r_{rr} = \frac{1}{2}\lambda' \qquad \Gamma^r_{tt} = \frac{1}{2}\nu' e^{\nu-\lambda} \tag{5.20}$$

$$\Gamma^r_{\theta\theta} = -re^{-\lambda} \qquad \Gamma^r_{\phi\phi} = -r\sin^2\theta e^{-\lambda} \tag{5.21}$$

all other components of the form Γ^r_{ab} vanish.

In order to compute the Ricci tensor we also need trace terms of the Christoffel symbol Γ^c_{ac}. Using the non-vanishing components we find

$$\Gamma^c_{tc} = 0 \tag{5.22}$$

$$\Gamma^c_{rc} = \Gamma^t_{rt} + \Gamma^r_{rr} + \Gamma^\theta_{r\theta} + \Gamma^\phi_{r\phi} = \frac{2}{r} + \frac{1}{2}(\nu' + \lambda') \tag{5.23}$$

$$\Gamma^c_{\theta c} = \cot\theta \tag{5.24}$$

$$\Gamma^c_{\phi c} = 0 \tag{5.25}$$

5.2 Ricci tensor components

The Ricci tensor is defined by

$$R_{ab} = \Gamma^n_{ba,n} - \Gamma^n_{bn,a} + \Gamma^m_{ba}\Gamma^n_{nm} - \Gamma^m_{bn}\Gamma^n_{am} \tag{5.26}$$

and for its components we find

$$R_{tt} = \Gamma^n_{tt,n} - \Gamma^n_{tn,t} + \Gamma^m_{tt}\Gamma^n_{nm} - \Gamma^m_{tn}\Gamma^n_{tm} \tag{5.27}$$

$$\Gamma^n_{tt,n} = \Gamma^r_{tt,r} = (\nu' e^{\nu-\lambda}/2)_{,r} = \frac{1}{2}\nu'' e^{\nu-\lambda} + \frac{1}{2}\nu'(\nu' - \lambda')e^{\nu-\lambda} \tag{5.28}$$

$$\Gamma^n_{nt,t} = 0 \tag{5.29}$$

$$\Gamma^m_{tt}\Gamma^n_{mn} = \Gamma^r_{tt}\Gamma^n_{rn} = \frac{1}{2}\nu' e^{\nu-\lambda}\left(\frac{2}{r} + (\nu' + \lambda')/2\right) \tag{5.30}$$

$$\Gamma^m_{tn}\Gamma^n_{mt} = \Gamma^t_{tn}\Gamma^n_{tt} + \Gamma^r_{tn}\Gamma^n_{rt} = \Gamma^t_{tr}\Gamma^r_{tt} + \Gamma^r_{tt}\Gamma^t_{rt} = \frac{1}{2}(\nu')^2 e^{\nu-\lambda} \tag{5.31}$$

All of the above terms are proportional to $e^{\nu-\lambda}$, putting them together we get

$$
\begin{aligned}
R_{tt} &= e^{\nu-\lambda}\left[\frac{1}{2}\nu'' + \frac{1}{2}(\nu')^2 - \frac{1}{2}\nu'\lambda' + \frac{1}{r}\nu' + \frac{1}{4}(\nu')^2 + +\frac{1}{4}\nu'\lambda' - \frac{1}{2}(\nu')^2\right] \\
&= e^{\nu-\lambda}\left[\frac{1}{2}\nu'' + \frac{1}{4}(\nu')^2 + \frac{1}{r}\nu' - \frac{1}{4}\nu'\lambda'\right] \qquad (5.32)
\end{aligned}
$$

The other components of the Ricci tensor can be calculated following the same routine, and are given by

$$R_{rr} = -\frac{1}{2}\nu'' - \frac{1}{4}(\nu')^2 + \frac{1}{4}\nu'\lambda' + \frac{1}{r}\lambda' \qquad (5.33)$$

$$R_{\theta\theta} = 1 - e^{-\lambda} + \frac{1}{2}r\lambda' e^{-\lambda} - \frac{1}{2}r\nu' e^{-\lambda} \qquad (5.34)$$

$$R_{\phi\phi} = \sin^2\theta R_{\theta\theta} \qquad (5.35)$$

5.3 The Schwarzschild solution

The vacuum field equations are

$$R_{ab} = 0 \qquad (5.36)$$

Since we have two unknown functions and three equations one should check that only two equations are independent. Let us firstly consider

$$(tt): \qquad \frac{1}{2}\nu'' + \frac{1}{4}(\nu')^2 + \frac{1}{r}\nu' - \frac{1}{4}\nu'\lambda' = 0 \qquad (5.37)$$

$$(rr): \qquad -\frac{1}{2}\nu'' - \frac{1}{4}(\nu')^2 + \frac{1}{4}\nu'\lambda' + \frac{1}{r}\lambda' = 0 \qquad (5.38)$$

Adding both equations yields

$$\frac{1}{r}(\nu' + \lambda') = 0 \qquad (5.39)$$

which is easily integrate and to give $\nu + \lambda = \tilde{C}$ where $\tilde{C}$ is a constant of integration. Exponentiating both sides leads to

$$e^{\nu} = Ce^{-\lambda} \qquad (5.40)$$

where C is another constant of integration $C = \exp\tilde{C}$. The constant of integration can be set to one by rescaling the time coordinate $t \mapsto \sqrt{C}t$ which results in

$$e^{\nu} = e^{-\lambda} \qquad (5.41)$$

Let us put this into the $(\theta\theta)$ equation

$$1 - e^{\nu} - \frac{1}{2}r\nu' e^{\nu} - \frac{1}{2}r\nu' e^{\nu} = 1 - e^{\nu} - r\nu' e^{\nu} = 0 \tag{5.42}$$

Note that

$$\frac{d}{dr}(re^{\nu}) = e^{\nu} + r\nu' e^{\nu} \tag{5.43}$$

and therefore we have to solve

$$\frac{d}{dr}(r - re^{\nu}) = 0 \tag{5.44}$$

which is easily integrated and results in $r - re^{\nu} = \mathcal{C}$ which leads to

$$e^{\nu} = 1 - \frac{\mathcal{C}}{r} \tag{5.45}$$

This solves the vacuum field equations and describes the exterior of a static and spherically symmetric object. This solution was first discovered by Karl Schwarzschild in 1916. The complete line element now reads

$$ds^2 = -\left(1 - \frac{\mathcal{C}}{r}\right)dt^2 + \left(1 - \frac{\mathcal{C}}{r}\right)^{-1}dr^2 + r^2 d\Omega^2 \tag{5.46}$$

where we denoted $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$.

In the limit $r \to \infty$ the metric component of the Schwarzschild solution approach those of Minkowski spacetime in spherical coordinates. This supports the interpretation of the Schwarzschild metric as the exterior gravitational field of an isolated body.

In physical units, this means reinserting the speed of light c and the gravitational constant G, we firstly observe that $1-\mathcal{C}/r$ should be dimensionless. If we write $G\mathcal{C}/c^2$ for $\mathcal{C}$, then the constant $\mathcal{C}$ should have dimensions of mass.

Now, let consider the metric

$$ds^2 = -\left(1 - \frac{G\mathcal{C}}{c^2 r}\right)dt^2 + \left(1 - \frac{G\mathcal{C}}{c^2 r}\right)^{-1}dr^2 + r^2 d\Omega^2 \tag{5.47}$$

and compute the non-vanishing Christoffel symbol components. Next, consider the limit $c^2 \to \infty$ or $1/c^2 \to 0$. Only one of the non-vanishing Christoffel symbol components contains the constant $\mathcal{C}$, namely

$$\lim_{c^2 \to \infty} \Gamma^r_{tt} = \frac{G\mathcal{C}}{2r^2} \tag{5.48}$$

Compare this with Newton's law per unit mass

$$\nabla\Phi = \frac{GM}{r^2} \tag{5.49}$$

Therefore, we interpret $\mathcal{C}/2$ as the total mass of the body (or of the Schwarzschild field). We finally write the the Schwarzschild metric in the following form

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{5.50}$$

5.4 Newtonian limit, weak field limit and equations of motion

The Newtonian equations of motions of a particle in a gravitational field are

$$m\ddot{\mathbf{r}} = -m\nabla\Phi \tag{5.51}$$

which in index notation reads

$$\frac{d^2X^i}{dt^2} = -\frac{\partial\Phi}{\partial X^i} \tag{5.52}$$

Let us consider the metric given by

$$g_{ab} = \eta_{ab} - h_{ab} \tag{5.53}$$

where eta_{ab} is the Minkowski metric. By weak fields we mean

$$|h_{ab}| \ll 1 \tag{5.54}$$

Note that in Lorentzian manifolds there is no positive norm and therefore there is no natural 'smallness'. We assume that all components of h_{ab} are smaller than one.

In first order approximation the inverse metric is given by

$$g^{ab} = \eta^{ab} + h^{ab} \tag{5.55}$$

To see this:

$$\begin{aligned}
\delta_a^c = g_{ab}g^{bc} &= (\eta_{ab} - h_{ab})(\eta^{bc} + h^{bc}) \\
&= \delta_a^c - h_{ab}\eta^{bc} + \eta_{ab}h^{bc} - h_{ab}h^{bc} \\
&= \delta_a^c - h_a{}^c + h_a{}^c + O(h^2) = \delta_a^c
\end{aligned} \tag{5.56}$$

where the indices of h_{ab} are raised and lowered by η_{ab}. □

Let us consider coordinates $X^a = (ct, x, y, z)$. If we consider the movement of a particle described by a geodesic parametrised by λ, its equations of motion are

$$\frac{d^2X^a}{d\lambda^2} + \Gamma^a_{bc}\frac{dX^b}{d\lambda}\frac{dX^c}{d\lambda} \tag{5.57}$$

The tangent vector to this curve , or its 4-velocity is $U^a = dX^a/d\lambda$. Let us assume small velocities

$$\frac{dX^a}{d\lambda} \ll \frac{dX^t}{d\lambda} \qquad i = 1, 2, 3 \tag{5.58}$$

Now the geodesic equation becomes

$$\frac{d^2X^a}{d\lambda^2} = -\Gamma^a_{bc}\frac{dX^b}{d\lambda}\frac{dX^c}{d\lambda} \simeq \Gamma^a_{tt}\frac{dt}{d\lambda}\frac{dt}{d\lambda} \tag{5.59}$$

By definition of the Christoffel symbol we have

$$\Gamma^a_{tt} = \frac{1}{2}g^{ad}(g_{dt,t} + g_{td,t} - g_{tt,d}) \tag{5.60}$$

which if we consider weak, static gravitational fields leads to

$$\Gamma^a_{tt} = -\frac{1}{2}g^{ad}g_{tt,d} \tag{5.61}$$

$$\Gamma^t_{tt} \simeq -\frac{1}{2}(\eta^{td} + h^{td})(\eta_{tt,d} - h_{tt,d}) = 0 \tag{5.62}$$

$$\Gamma^t_{tt} \simeq -\frac{1}{2}(\eta^{id} + h^{id})(\eta_{tt,d} - h_{tt,d}) \simeq \frac{1}{2}\eta^{ii}h_{tt,i} = \frac{1}{2}\frac{\partial h_{tt}}{\partial X^i} \tag{5.63}$$

Hence, the time component satisfies

$$\frac{d^2t}{d\lambda^2} = 0 \quad \Rightarrow \quad dt = Cd\lambda \tag{5.64}$$

where C is some constant of integration. Next, ket us consider the spatial part of the geodesic equations

$$\frac{d^2X^i}{d\lambda^2} \simeq -\frac{1}{2}\frac{\partial h_{tt}}{\partial X^i}\frac{cdt}{d\lambda}\frac{cdt}{d\lambda} \tag{5.65}$$

which can be rewritten using Eq. (5.64) and leads to

$$\frac{d^2X^i}{dt^2} = -c^2\frac{1}{2}\frac{\partial h_{tt}}{\partial X^i} \tag{5.66}$$

Comparison with the Newtonian equations yields

$$-c^2 \frac{1}{2} \frac{\partial h_{tt}}{\partial X^i} = -\frac{\partial \Phi}{\partial X^i} \tag{5.67}$$

and therefore $h_{tt} = \frac{2\Phi}{c^2}$ which in turn gives

$$g_{tt} = \eta_{tt} - h_{tt} = -1 - \frac{2\Phi}{c^2} = -\left(1 + \frac{2\Phi}{c^2}\right) \tag{5.68}$$

For a spherically symmetric mass distribution in Newtonian gravity we have

$$\Phi(r) = -\frac{GM}{r} \tag{5.69}$$

which then corresponds to

$$g_{tt} = -\left(1 - \frac{2GM}{c^2 r}\right) \tag{5.70}$$

Using $g^{ab} = \eta^{ab} + h^{ab}$ and its inverse one can in principle compute the linearised Ricci and Einstein tensor and consequently one can study the linearised field equations. This is for instance important when gravitational waves are discussed.

5.5 Significance of the Schwarzschild solution

The Schwarzschild solution

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega^2 \tag{5.71}$$

is the most important known exact solution of the vacuum field equations.

An important feature of the components of the metric is that they become singular at $r = 2m$ and also at $r = 0$. This indicates that either we used bad coordinates or there is a true geometrical singularity. It turns out that the $r = 2m$ hypersurface is well defined geometrically and one can construct improved coordinates regular at $r = 2m$. On the other hand, $r = 0$ is a physical singularity where the notion of spacetime breaks down.

Let us consider the numerical value of the $r = 2m$ surface's radius in physical units. One finds

$$r_s = \frac{2GM}{c^2} \approx 3 \left(\frac{M}{M_\odot}\right) \text{km} \tag{5.72}$$

where $M_\odot \approx 2 \times 10^{30}$kg is the mass of the Sun. Hence, for 'normal' bodies of astrophysical interest like the Sun, the Earth or a neutron star, the Schwarzschild radius r_s is well inside the radius of the body where the vacuum solution is no longer valid.

However, if a body undergoes complete gravitational collapse, its surface will eventually disappear and be inside its Schwarzschild radius. Such a situation is then described by the Schwarzschild solutions and the $r = 2M$ hypersurface will be of physical importance. Such an astrophysical object is called a black and has many interesting properties. The $r = 2M$ hypersurface of a black hole of mass M is called the event horizon.

Since the Schwarzschild solution describes the exterior of a spherical mass distribution it is also the basis to predict deviations from Newtonian gravity around the Sun or the Earth.

The three classical tests of general relativity are the perihelion precession of Mercury, the deflection or bending of light by the Sun and gravitational redshift of light. These three tests are based on solving or approximating the geodesics of the Schwarzschild spacetime. Geodesics of the Schwarzschild spacetime will be discussed in Section 6.

5.6 The Schwarzschild interior solution

As already discussed, the Schwarzschild solution describes the exterior spacetime of a spherical body. Next we are interested in modelling a general relativistic star, an interior solution.

5.6.1 Field equations and conservation equation

Let us assume the star's interior is described by a perfect fluid with energy-momentum tensor

$$T_{ab} = \rho u_a u_b + p(g_{ab} + u_a u_b) \tag{5.73}$$

where ρ and p are functions of r only. Note the sign difference to Section 3.7 because we now work with signature $(-, +, +, +)$.

The static and spherically symmetric metric reads

$$ds^2 = -e^{\nu(r)}dt^2 + e^{\lambda(r)}dr^2 + r^2 d\theta^2 + r^2 d\Omega^2 \tag{5.74}$$

The fluid's 4-velocity is given by

$$u^t = e^{-\nu/2} \tag{5.75}$$

and all others vanish, which yields $g_{ab}u^a u^b = -1$.

The components of the energy-momentum tensor are therefore

$$T_{tt} = \rho(-e^{\nu/2})^2 + p(-e^{\nu} + (-e^{\nu/2})^2) = \rho e^{\nu} \tag{5.76}$$

$$T_{ii} = p g_{ii} \qquad i = r, \theta, \phi \tag{5.77}$$

and therefore

$$T_{ab} = \mathrm{diag}(\rho e^{\nu}, p e^{\lambda}, p r^2, p r^2 \sin^2\theta) \tag{5.78}$$

By raising one index of the energy-momentum tensor, the metric functions will disappear and we find

$$T^a_b = \mathrm{diag}(-\rho, p, p, p) \tag{5.79}$$

Let us also compute the trace of the energy-momentum tensor

$$T = g^{ab}T_{ab} = T^a_a = -\rho + p + p + p = -\rho + 3p \tag{5.80}$$

Since we already computed the Ricci tensor components, it is useful to use the field equations in the form

$$R_{ab} = 8\pi(T_{ab} - \frac{1}{2}T g_{ab}) \tag{5.81}$$

$$e^{\nu-\lambda}\left[\frac{1}{2}\nu'' + \frac{1}{4}(\nu')^2 + \frac{1}{r} - \frac{1}{4}\nu'\lambda'\right] = 8\pi\left[\rho e^{\nu} + \frac{1}{2}e^{\nu}(-\rho + 3p)\right] \tag{5.82}$$

$$-\frac{1}{2}\nu'' - \frac{1}{4}(\nu')^2 + \frac{1}{4}\nu'\lambda' + \frac{1}{r}\lambda' = 8\pi\left[p e^{\lambda} - \frac{1}{2}e^{\lambda}(-\rho + 3p)\right] \tag{5.83}$$

$$1 - e^{-\lambda} + \frac{1}{2}r\lambda' e^{-\lambda} - \frac{1}{2}r\nu' e^{-\lambda} = 8\pi\left[\rho r^2 + \frac{1}{2}r^2(-\rho + 3p)\right] \tag{5.84}$$

which are the (tt), (rr) and $(\theta\theta)$ components of the field equations, respectively.

Before solving these equations, recall that the twice contracted Bianchi identities imply the conservation of the energy-momentum tensor $\nabla_a T^{ab}$. In our case this yields

$$\partial_a T^{ab} + \Gamma^a_{ac} T^{cb} + \Gamma^b_{ac} T^{ac} = 0 \tag{5.85}$$

Of these four equations, only the $b = r$ equation gives a non-trivial result

$$\partial_a T^{ar} + \Gamma^a_{ac} T^{cr} + \Gamma^r_{ac} T^{ac} = 0 \tag{5.86}$$

The first two terms are quickly computed

$$\partial_a T^{ar} = \partial_r T^{rr} = \partial_r(pe^{-\lambda}) = p'e^{-\lambda} - p\lambda' e^{-\lambda} \tag{5.87}$$

$$\Gamma^a_{ac} T^{cr} = \Gamma^a_{ar} T^{rr} = \left(\frac{2}{r^2} + \frac{1}{2}(\nu' + \lambda')\right) pe^{-\lambda} \tag{5.88}$$

while the last term is

$$\begin{aligned} \Gamma^r_{ac} T^{ac} &= \Gamma^r_{tt} T^{tt} + \Gamma^r_{rr} T^{rr} + \Gamma^r_{\theta\theta} T^{\theta\theta} + \Gamma^r_{\phi\phi} T^{\phi\phi} \\ &= \frac{1}{2}\nu' e^{\nu-\lambda}\rho e^{-\nu} + \frac{1}{2}\lambda' P e^{-\lambda} - re^{-\lambda} p\frac{1}{r^2} - r\sin^2\theta e^{-\lambda} p \frac{1}{r^2}\frac{1}{\sin^2\theta} \end{aligned} \tag{5.89}$$

Adding up these three terms lead to the conservation equation of a static and spherically perfect fluid

$$p' + \frac{1}{2}\nu'(\rho + p) = 0 \tag{5.90}$$

5.6.2 Tolman-Oppenheimer-Volkoff equation

Let us start solving the field equations by considering the combination (5.82) + (5.83) + 2 × (5.84) which gives

$$e^{-\lambda}\left[\frac{1}{r}\nu' + \frac{1}{r}\lambda' + (2e^{\lambda} - 2)/r^2 + \frac{1}{r}\lambda' - \frac{1}{r}\nu'\right] = 8\pi[4\rho]\frac{1}{2} \tag{5.91}$$

Simplify this expression and multiply by r^2 yields

$$e^{-\lambda}\left[\lambda' r + e^{\lambda} -\right] = 8\pi\rho r^2 \tag{5.92}$$

$$1 - e^{-\lambda} + r\lambda' e^{-\lambda} = 8\pi\rho r^2 \tag{5.93}$$

$$\frac{d}{dr}(r - re^{-\lambda}) = 8\pi\rho r^2 \tag{5.94}$$

Let us define the mass up to r by

$$m(r) = \int_0^r 4\pi\rho(r') r'^2 dr' \tag{5.95}$$

and hence Eq. (5.94) becomes

$$r - re^{-\lambda} = 2m(r) + C \tag{5.96}$$

where C is a constant of integration. Assuming a regular centre of the perfect fluid sphere we set $C = 0$ and find

$$e^{-\lambda} = 1 - \frac{2m(r)}{r} \tag{5.97}$$

The combination (5.82) + (5.83) leads to

$$e^{-\lambda}\left[\frac{1}{r}(\nu' + \lambda')\right] = 8\pi(\rho + p) \tag{5.98}$$

Let us use

$$\lambda' e^{-\lambda} = \frac{2m'r - 2m}{r^2} = 8\pi\rho r - \frac{2m}{r^2} \tag{5.99}$$

which yields

$$e^{-\lambda}\nu' = 8\pi(\rho + p)r + \frac{2m}{r^2} - 8\pi rhor \tag{5.100}$$

and finally we can solve by ν' and obtain

$$\frac{\nu'}{2} = \frac{1}{r^2}\frac{m + 4\pi pr^3}{1 - 2m/r} \tag{5.101}$$

Recall the conservation equation

$$\frac{\nu'}{2} = -\frac{p'}{\rho + p} \tag{5.102}$$

Combining both results yields a differential equation for the star's pressure

$$p' = -\frac{1}{r^2}\frac{(4\pi pr^3 + m)(\rho + p)}{1 - \frac{2m}{r}} \tag{5.103}$$

Collecting results:

$$p' = -\frac{1}{r^2}\frac{(4\pi pr^3 + m)(\rho + p)}{1 - \frac{2m}{r}} \tag{5.104}$$

$$m' = 4\pi\rho r^2 \tag{5.105}$$

$$e^{-\lambda} = 1 - \frac{2m}{r} \tag{5.106}$$

$$\nu' = -\frac{2p'}{\rho + p} \tag{5.107}$$

The differential equation for the pressure is known as the Tolman-Oppenheimer-Volkoff (TOV) equation of hydrostatic equilibrium. Note that there are four unknown functions, namely ν ,λ, ρ and p but only three independent equations. Likewise Eqs. (5.104) and (5.105) determine $p(r)$ and $m(r)$ but contain three unknowns m, ρ and p. One further condition needs to be imposed to close the system of equations.

From an astrophysical point of view it were natural to prescribe the matter distribution $\rho = \rho(r)$. This, however, often leads to a divergent pressure near the centre. The most physical is to specify an equation of state $\rho = \rho(p)$ or $p = p(\rho)$. For realistic equations of state, these equations can in general not be integrated analytically.

5.6.3 Constant density stars

Let us assume the special case $\rho(r) = \rho_0 =$ const. which approximates a very dense object like a neutron star or a white dwarf

$$m(r) = int_0^r 4\pi\rho_0 r'^2 dr' = \frac{4\pi}{3}\rho_0 r^3 \tag{5.108}$$

$$e^{-\lambda} = 1 - \frac{2m(r)}{r} = 1 - \frac{8\pi}{3}\rho_0 r^2 \tag{5.109}$$

Next, consider the TOV equation

$$p' = -r\frac{4\pi p + (\frac{4\pi}{3}\rho_0)(\rho_0 + p)}{1 - \frac{8\pi}{3}\rho_0 r^2} \tag{5.110}$$

Separation of variables yields

$$\frac{dp}{(p + \rho_0/3)(p + \rho_0)} = -\frac{4\pi r dr}{1 - \frac{8\pi}{3}\rho_0 r^2} \tag{5.111}$$

The left-hand side can be integrated using partial fractions

$$\begin{aligned} \int \frac{dp}{(p + \rho_0/3)(p + \rho_0)} &= \frac{9}{2\rho_0}\int \frac{dp}{\rho_0 + 3p} - \frac{3}{2\rho_0}\int \frac{dp}{\rho_0 + p} \\ &= \frac{3}{2\rho_0}\log(2\rho_0(\rho_0 + p)) - \frac{3}{2\rho_0}\log(\rho_0 + p) \end{aligned} \tag{5.112}$$

while the right-hand side gives

$$-\int \frac{4\pi r dr}{1 - \frac{8\pi}{3}\rho_0 r^2} = \frac{3}{4\rho_0}\log\left(3(1 - \frac{8\pi}{3}\rho_0 r^2)\right) \tag{5.113}$$

Putting both terms together and adding a constant of integration gives

$$\frac{3}{2\rho_0}\log\left(\frac{2\rho_0(\rho_0+p)}{\rho_0+p}\right)=\frac{3}{4\rho_0}\log\left(3(1-\frac{8\pi}{3}\rho_0 r^2)\right)+C \tag{5.114}$$

$$\log\left(\frac{2\rho_0(\rho_0+p)}{\rho_0+p}\right)=\log\sqrt{3(1-\frac{8\pi}{3}\rho_0 r^2)}+\tilde{C} \tag{5.115}$$

$$\frac{2\rho_0(\rho_0+p)}{\rho_0+p}=\bar{C}\sqrt{3(1-\frac{8\pi}{3}\rho_0 r^2)} \tag{5.116}$$

Let us fix the constant of integration by

$$p(r=0)=p_c \tag{5.117}$$

where p_c denotes the central pressure and we get

$$\frac{2\rho_0(\rho_0+p_c)}{\rho_0+p_c}=\bar{C}\sqrt{3} \tag{5.118}$$

Now insert the value of $\bar{C}$ into Eq. (5.116) yields

$$\frac{2\rho_0(\rho_0+p)}{\rho_0+p}=\sqrt{1-\frac{8\pi}{3}\rho_0 r^2}\,\frac{2\rho_0(\rho_0+p_c)}{\rho_0+p_c} \tag{5.119}$$

This equation we now solve for the pressure

$$p=\rho_0\frac{(\rho_0+3p_c)\sqrt{1-\frac{8\pi}{3}\rho_0 r^2}-(\rho_0+p_c)}{3(\rho+p_c)-(\rho_0+3p_c)\sqrt{1-\frac{8\pi}{3}\rho_0 r^2}} \tag{5.120}$$

which is the star's pressure as a function of the radius.

We define the surface of the star to be the vanishing pressure surface $p(R)=0$

$$(\rho_0+3p_c)\sqrt{1-\frac{8\pi}{3}\rho_0 R^2}-(\rho_0+p_c)=0 \tag{5.121}$$

$$\sqrt{1-\frac{8\pi}{3}\rho_0 R^2}=\frac{\rho_0+p_c}{\rho_0+3p_c} \tag{5.122}$$

which can be solved for R.

An important relation follows if we express the central pressure p_c in terms of the radius R, the energy density ρ_0 and the mass of to R, the total mass M

$$M=m(R)=\int_0^R 4\pi r'^2 dr'=\frac{4\pi}{3}\rho_0 R^3 \tag{5.123}$$

From Eq. (5.122) we find

$$\sqrt{1-\frac{2M}{R}}=\frac{\rho_0+p_c}{\rho_0+3p_c} \tag{5.124}$$

which we solve for the central pressure

$$p_c=\rho_0\frac{1-\sqrt{1-\frac{2M}{R}}}{3\sqrt{1-\frac{2M}{R}}-1} \tag{5.125}$$

Since we are interested in astrophysical objects with finite central pressure (regular centre), we find

$$p_c<\infty \quad \Rightarrow \quad \sqrt{1-\frac{2M}{R}}>\frac{1}{3} \quad \Leftrightarrow \quad \frac{2M}{R}<\frac{8}{9}\overset{!}{<}1 \tag{5.126}$$

A uniform density star with $M>4R/9$ cannot exist in general relativity. This result holds independently of the pressure p throughout the star. It in particular implies that the radius of a static perfect fluid sphere exceed the corresponding Schwarzschild radius.

We defined the mass up to r by

$$m(r)=\int_0^r 4\pi\rho(r')r'^2dr' \tag{5.127}$$

which is formally related to the mass in Newton's theory of gravity. This formal analogy must be read with care since the proper volume element is

$$\sqrt{^3g}d^3x=e^{\lambda/2}r^2\sin\theta drd\theta d\phi \tag{5.128}$$

where ^3g denotes the determinant of the spatial part of the metric.

Hence, the proper mass up to r is

$$m_p(r)=\int_0^r 4\pi\rho(r')r'^2\left(1-\frac{2m(r')}{r'}\right)^{-1/2}dr' \tag{5.129}$$

The difference between m and m_p of the total masses (integration up to the surface) can be interpreted as the gravitational binding energy. If R is the surface of the star ($p(R)=0$), then

$$E_b=m_p(R)-m(R) \tag{5.130}$$

which is strictly positive since $m_p>m$.

Finally, let us consider the Newtonian limit of the TOV equation $p \ll \rho$ and $m(r) \ll r$. Then we find

$$p' = -\frac{\rho(r)m(r)}{r^2} \tag{5.131}$$
$$m' = 4\pi\rho(r)r^2 \tag{5.132}$$

which are the structure equations of Newtonian astrophysics. Their solutions describe Newtonian stars.

6 Geodesics of the Schwarzschild solutions

6.1 Geodesic equations

Without loss of generality we can assume motion in the equatorial plane $\theta = \pi/2$, $\dot{\theta} = 0$. The motion of massive and massless particles follows from the Lagrangian

$$L = -\left(1 - \frac{2M}{r}\right)\dot{t}^2 + \left(1 - \frac{2M}{r}\right)^{-1}\dot{r}^2 + r^2\dot{\phi}^2 \tag{6.1}$$

with

$$L = \begin{cases} -1 \quad m \neq 0 \quad \text{time-like geodesics} \\ 0 \quad m = 0 \quad \text{null geodesics} \end{cases} \tag{6.2}$$

Since the Lagrangian is independent of t and ϕ, there will be two constants of motion, namely E (energy) and ℓ (angular momentum)

$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{t}} = \frac{\partial L}{\partial t} = 0 \quad \Rightarrow \quad \frac{\partial L}{\partial \dot{t}} = \text{const.} \tag{6.3}$$

and we write

$$-2\left(1 - \frac{2M}{r}\right)\dot{t} = -2E \tag{6.4}$$
$$E = \left(1 - \frac{2M}{r}\right)\dot{t} \tag{6.5}$$

Likewise

$$\frac{d}{d\lambda}\frac{\partial L}{\partial \dot{\phi}} = \frac{\partial L}{\partial \phi} = 0 \quad \Rightarrow \quad \frac{\partial L}{\partial \dot{\phi}} = \text{const.} \tag{6.6}$$

$$2r^2\dot{\phi} = 2\ell \tag{6.7}$$

$$\ell = r^2\dot{\phi} \tag{6.8}$$

Substituting both constants of motion back into the Lagrangian we find

$$L = -\left(1 - \frac{2M}{r}\right)\frac{E^2}{\left(1-\frac{2M}{r}\right)^2} + \frac{\dot{r}^2}{\left(1-\frac{2M}{r}\right)} + r^2\frac{\ell^2}{r^4} \tag{6.9}$$

which can be rewritten in the following way

$$\frac{1}{2}\dot{r}^2 + \frac{1}{2}\left(1 - \frac{2M}{r}\right)\left(\frac{\ell^2}{r^2} - L\right) = \frac{1}{2}E^2 \tag{6.10}$$

This equation shows that the radial motion of a geodesic is analogous to the equations of motion of a test particle with unit mass and energy $E^2/2$. The motion is determined by the effective potential

$$V_{\text{eff}} = \frac{1}{2}\left(1 - \frac{2M}{r}\right)\left(\frac{\ell^2}{r^2} - L\right) \tag{6.11}$$

One can think of ordinary 1-dimensional non-relativistic mechanics

$$V_{\text{eff}} = \frac{\ell^2}{2r^2} - \frac{M\ell^2}{r^3} + L\frac{M}{r} - L\frac{1}{2} \tag{6.12}$$

$$\frac{\ell^2}{2r^2} \quad \text{centrifugal barrier term} \tag{6.13}$$

$$-\frac{M\ell^2}{r^3} \quad \text{new term: dominates over barrier term for small } r \tag{6.14}$$

$$L\frac{M}{r} \quad \text{normal Newtonian term } -M/r \tag{6.15}$$

Another useful quantity to consider is the spatial orbit parametrised by the radius r. This means the curve $\phi = \phi(r)$

$$\frac{d\phi}{dr} = \frac{d\phi}{d\lambda}\frac{d\lambda}{dr} = \frac{\dot{\phi}}{\dot{r}} = \frac{\ell}{r^2}\frac{1}{\dot{r}} \tag{6.16}$$

From Eq. (6.10) we have

$$\dot{r}^2 = E^2 - \left(1 - \frac{2M}{r}\right)\left(\frac{\ell^2}{r^2} - L\right) \tag{6.17}$$

and therefore

$$\frac{d\phi}{dr} = \frac{\ell}{r^2}\left[E^2 - \left(1 - \frac{2M}{r}\right)\left(\frac{\ell^2}{r^2} - L\right)\right]^{-1/2} \tag{6.18}$$

6.2 Light deflection

use.eps

The deflection angle is given by

$$\Delta\phi = \phi_{+\infty} - \phi_{-\infty} - \pi = 2\phi(\infty) - \pi \tag{6.19}$$

For light rays, massless particles, we have $L = 0$ and we find

$$\frac{d\phi}{dr} = \frac{\ell}{r^2}\left[E^2 - \left(1 - \frac{2M}{r}\right)\frac{\ell^2}{r^2}\right]^{-1/2} \tag{6.20}$$

Since $r(\phi)$ is minimal at r_0 we have

$$\frac{dr}{d\phi}(r = r_0) = 0 \tag{6.21}$$

$$E^2 = \left(1 - \frac{2M}{r_0}\right)\frac{\ell^2}{r_0^2} \tag{6.22}$$

Therefore

$$\phi(r) = \int_{r_0}^{r} \frac{\ell}{r'^2}\left[\left(1 - \frac{2M}{r_0}\right)\frac{\ell^2}{r_0^2} - \left(1 - \frac{2M}{r'}\right)\frac{\ell^2}{r'^2}\right]^{-1/2} dr' \tag{6.23}$$

Note that all ℓ cancel. To find the deflection angle, we need to find $\phi(\infty)$, and therefore we have to evaluate

$$\phi(\infty) = \int_{r_0}^{\infty} \frac{dr'}{\sqrt{\left(1 - \frac{2M}{r_0}\right)\frac{r'^4}{r_0^2} - (r'^2 - 2Mr')}} \tag{6.24}$$

Let us perform a change of variables $u = 1/r'$, $du = -1/r'^2\, dr$

$$\phi(0) = \int_0^{1/r_0} \frac{du}{\sqrt{\left(1 - \frac{2M}{r_0}\right)\frac{1}{r_0^2} - u^2 - 2Mu^3}} \tag{6.25}$$

To first order in the mass M one can use the following trick to evaluate this integral (like a Taylor expansion)

$$\phi(0) = \phi(0)[M=0] + \frac{\partial \phi(0)}{\partial M}[M=0] \cdot M + O(M^2) \tag{6.26}$$

$$\phi(0)[M=0] = \int_0^{1/r_0} \frac{du}{\sqrt{1/r_0^2 - u^2}} = \frac{\pi}{2} \tag{6.27}$$

$$\begin{aligned}\frac{\partial \phi(0)}{\partial M}[M=0] &= \int_0^{1/r_0} \frac{1/r_0^3 - u^3}{1/r_0^2 - u^2} du \\ &= -\frac{2 + r_0 u}{1 + r_0 u}\sqrt{1/r_0 - u^2}\Big|_0^{1/r_0} = \frac{2}{r_0}\end{aligned} \tag{6.28}$$

which leads to

$$\Delta\phi = 2\phi(\infty) - \pi = \pi + \frac{4M}{r_0} - \pi + O(M^2) \simeq \frac{4M}{r_0} \tag{6.29}$$

For a light ray passing nearby the Sun we have

$$r_0 \simeq R_\odot \simeq 7 \times 10^5\,\mathrm{km} \tag{6.30}$$

$$M = \frac{GM_\odot}{c^2} \simeq 1.5\,\mathrm{km} \tag{6.31}$$

Using that $\pi = 180 \cdot 3600''$ we find

$$\Delta\phi 1.75'' \tag{6.32}$$

The deflection or bending of light has been observed during solar eclipses beginning with the 1919 expedition of Eddington. This confirmed an important prediction of general relativity.

Newtonian theory predicts an angle of

$$\Delta\phi_{\mathrm{N}} = \frac{2M}{r_0} = \frac{1}{2}\Delta\phi_{\mathrm{GR}} \tag{6.33}$$

and hence experiments rule out Newtonian gravity.

Modern experiments analyse radio waves emitted by quasars; there is no need to wait for solar eclipses using radio signals.