Notes on

General Relativity

v0.0

S.Bernuzzi ©
sebastiano.benuzzi@uni-jena.de
Jena FSU

Last update: April 30, 2021

Contents

0. About

These are semiprivate notes sketching topics and calculations discussed in about 25 lectures of 1.5-2 hours each at Jena FSU. They are not meant to substitute books. Please visit

`http://sbernuzzi.gitpages.tpi.uni-jena.de/gr/`

for an updated list of books, references and other material, including the exercise sheets distributed each week.

I welcome constructive feedbacks. Red text is work in progr...

Conventions. The spacetime and the metric are indicated as $(\mathcal{M}, g_{ab})$ and the notation mostly follows Wald's book: signature convention $(-,+,+,+)$, $a, b, ...$ indexes in abstract notation, $\alpha, \beta, ...$ indexes of tensor components, $i, j, ...$ spatial coordinate indexes, etc. Coordinate basis of the tange vector space $T_p(M)$ are indicated as e_μ; the natural basis of partial derivatives is $e_\mu = \partial_\mu$. The dual basis $e^{*\nu}$ ($e_\mu e^{*\nu} = \delta^\nu_\mu$) is constructed by the gradients of the coordinates is $e^{*\mu} = \mathrm{d}x^\mu$. The exterior derivative of an n-form is indicated with $\mathbf{d}$; applied to scalars it reduces to the gradient (1-form) $\mathbf{d}f = \mathrm{d}f = \mathrm{grad}(f)$ with components $(\mathbf{d}f)_\mu = (\mathrm{d}f)_\mu = \partial_\mu f$. Covariant derivatives (Levi-Civita connection) are indicated with ∇. ∇ applied to scalars reduces to the gradient $\nabla f = \mathrm{d}f$ (components $\nabla_\mu f = (\mathrm{d}f)_\mu = \partial_\mu f$) and it is consistent with the concept of tangent vector $v(f) = v^\mu \nabla_\mu f$. The symbol $:=$ is an assignment, while $\equiv$ an identity.

Units are $c = G = 1$ if not specified.

1. Introduction

②

These introductory lectures briefly summarize the key GR concepts and historical milestones in a very basic and accessible way (BSc level). All the concepts will be developed in detail during the course.

Suggested readings. *There are several popular books, readings and videos on GR, please see the course webpage. I recommend a classic: Kip S. Thorne, "Black holes and time warps: Einstein's outrageous legacy.".*

TODO

Please temporary refer to the handwritten notes at:
`http://sbernuzzi.gitpages.tpi.uni-jena.de/gr/notes/2018/intro_notes.pdf`

2. Special Relativity

③

These lectures summarize the theory of special relativity (SR) from the postulates to the dynamics of particles and fields. I assume the students are already familiar with some SR concepts from the electromagnetism and relativistic physics courses. A definition of tensor is introduced using the transformation rule for its components in SR.

Suggested readings. *Chap. 1 of Landau and Lifschits (1975); Chap. 1 of Schutz (1985); Chap. 1 of Carroll (1997); Chap. 1 of Wald (1984).*

2.1 SR postulates

Definition 2.1.1. Inertial observer (frame) = *reference system in which a free moving body (no forces acting on the body) moves at constant velocity.*

SR postulates [Einstein (1905)]

(i) *Relativity principle*: Laws of nature are the same in all inertial frames.
$\Rightarrow$ Physics laws be invariant with respect to transformation of coordinates connecting inertial frames.

(ii) The *speed of light* in vacuum is the same in all inertial frames, and its value is finite

$$c \simeq 2.99\ldots\ 10^{10}\ \mathrm{cm/s}\ . \tag{2.1}$$

Definition 2.1.2. Spacetime = *continuum, composed of events.*

Definition 2.1.3. Event = *when & where, point of the spacetime labelled by coordinates* $(t, x, y, z) \in \mathbb{R}^4$.

2.2 Spacetime & Causal structure

The SR postulates determine a peculiar causal structure of the spacetime.

Causal structure in pre-relativistic physics. Given an event p, an observer/material body can either (i) move from p to q; or (ii) move from q' to p; and (iii) cannot *be* at p and at p'. Hence one defines

(i) Set of events that can be reached from p, $\{q\}$ = FUTURE of p;
(ii) Set of events that can reach p, $\{q'\}$ = PAST of p;
(iii) Set $\{p'\}$ = SIMULTANEOUS to p.

In other terms,

- there exists an *absolute time* t (observer/frame independent);
- $t = const$ defines a 3D surface of simultaneous events.

This causal structure has an important consequence illustrated in what follows.

Distances in pre-relativistic physics. Consider two distinct events in space, p and q, and two Cartesian frames. The frame $\mathcal{O}' : (t', x', y', z')$ moves a constant velocity $\vec{V} = V\hat{x}$ along the x-axis of frame $\mathcal{O} : (t, x, y, z)$. Another equivalent notation for the coordinate labels is $(t, x, y, z) = (x_0, x_1, x_2, x_3)$ and same with primes.

The Galilean tranformation connecting the coordinates of the two frames is

$$\begin{cases} t' = t \quad \text{(absolute time)} \\ x' = x - Vt \\ y' = y \\ z' = z \end{cases} . \tag{2.2}$$

The formula above can be easily generalized for more complicated orientation of $\mathcal{O}'$; but it sufficient to capture all the concepts described here. The key point is that one always has $t' = t$ in pre-relativistic physics: the time is absolute.

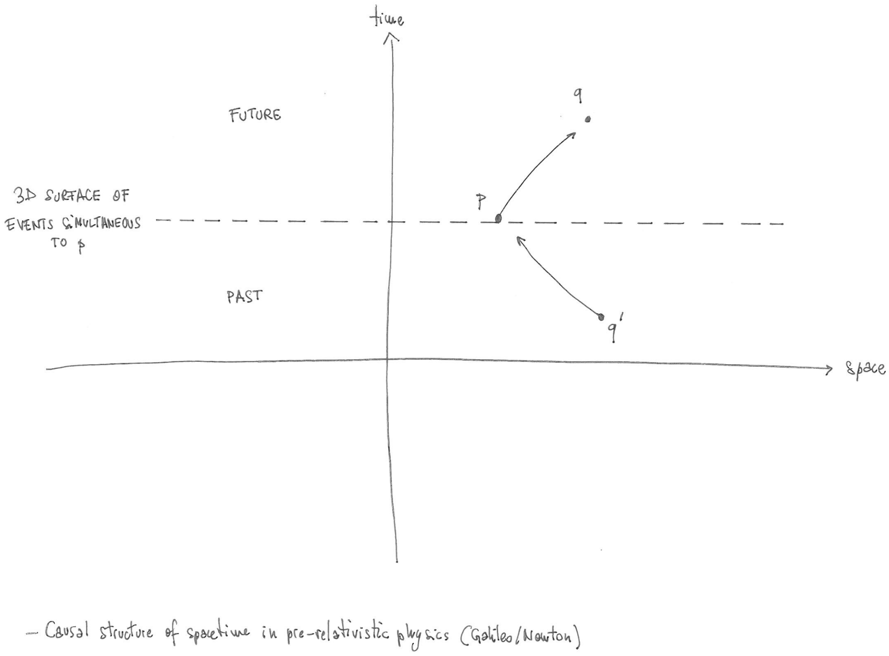

Figure 2.1: Causal structure in pre-relativistic physics.

The (infinitesimal) distance $\bar{pq} = d\ell$ is given by Pythagorean's theorem and can be written in many equivalent ways:

$$d\ell^2 = dx^2 + dy^2 + dz^2 = \sum_{i=1}^{3}(dx_i)^2 \tag{2.3a}$$

$$= \sum_{i=1}^{3}\sum_{j=1}^{3}\delta_{ij}(dx_i)(dx_j) \quad \text{with} \quad \delta_{ij} := \text{diag}(1,1,1) \quad \underline{\text{Euclidean metric}} \tag{2.3b}$$

$$= \delta_{ij}(dx^i)(dx^j) = (dx_i)(dx^i) \ , \tag{2.3c}$$

where in the last line Einstein's sum-convention is introduced.

The square distance ℓ^2 is clearly the square Eucliden distance in $\mathbb{R}^3$ defined by the standard scalar product as a quadratic form in the coordinates. ℓ^2 has the same value in $\mathcal{O}$ and $\mathcal{O}'$; i.e. it is <u>invariant</u> in pre-relativistic physics. This can be immediately verified in the special case by applying Eq. (2.2)

$$\delta\ell^{2'} = (\delta x)^2 = (x'_p - x'_q)^2 = (x_p - Vt - x_q + Vt)^2 = (x_p - x_q)^2 = \delta\ell^2 \ , \tag{2.4}$$

and it is clearly a consequence of the absolute time in Eq. (2.2) that implies one needs to consider only spatial translations and rotations.

Causal structure in SR. Given an event p, material bodies can still move to/from p. Hence, there exists events in the

(i) FUTURE of p;
(ii) PAST of p.

But there exists also events that (iii) can be connected only by light and not by material bodies; (iv) cannot be connected even by light because they would require a speed larger than c. Hence, one defines

(iiia) Set of events that can be reached from p following light rays, $\{i^+\}$ = FUTURE LIGHT CONE of p;
(iiib) Set of events that can reach p following light rays, $\{i^-\}$ = PAST LIGHT CONE of p;
(iv) Set of events causally disconnected from p, $\{p'\}$ = SPACELIKE events.

Consider now an inertial observer $\mathcal{O}$ and draw a *spacetime diagram*, Fig. (2.2). In a spacetime diagram spacetime is often represented with diagram (t, x) where the x axes represent the spatial dimensions in 1D. A point of the diagram is an event, a line represents uniform motion at speed v such that $dt/dx = 1/v$. One uses units with $c = 1$ so that 45 degrees lines represent the motion of light (photons); For example, the lines $t = \pm x$ represent light rays passing trhough the origin. To label events with coordinates the observer must (i) place a rigid frame to measure distances in the 3D space; (ii) place a clock at each point of space; (iii) syncronize the clocks by sending ligh pulses (e.g. each clock starts when the light pulse arrives). The observer $\mathcal{O}$ makes an *observation* when it assigns to the event p the spatial coordinate (x, y, z) and the time t <u>read by the clock at (x, y, z)</u>. Note this is different from a *visual observation* of the

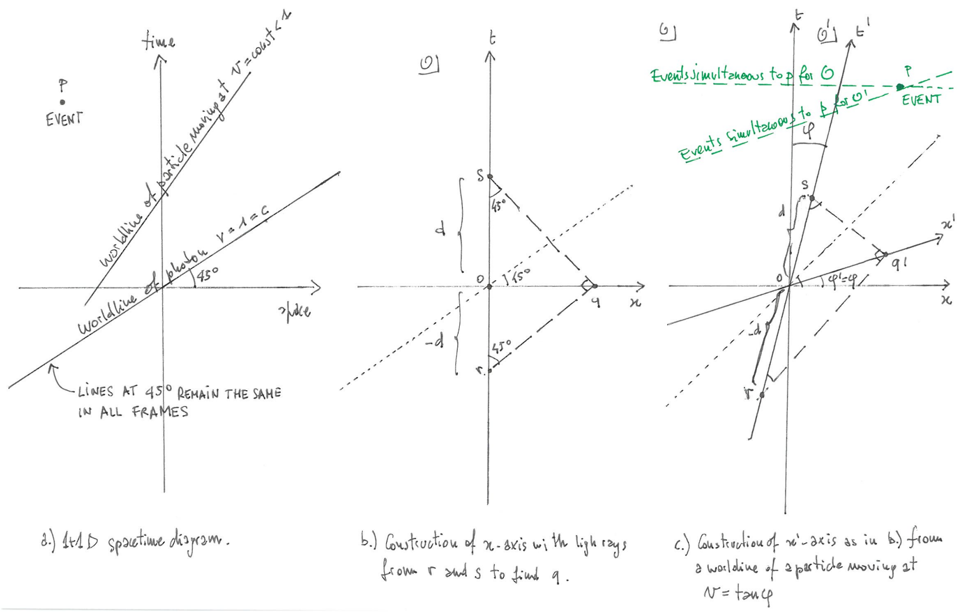

Figure 2.2: Spacetime diagrams and causal structure in SR.

event made by a scientist stitting at, e.g. $(x, y, z) = (0, 0, 0)$. The intertial observers measure as simultaneous all the events happening at the same time as indicate by the clocks in their position ($t = const$); the scientist sitting at the origin sees as simultaneous all the events happening at the same time as indicated by his clock (at his position).

Once a time axis is chosen, the spatial axes (only one in the figure) can be defined by (i) drawing light rays from points r and s chosen such that they are at the same distance from the origin, $\bar{ro} = \bar{so}$; (ii) drawing the line from o to the intersection q of the light rays, Fig. (2.2).

A second inertial observer $\mathcal{O}'$ moving at speed V is set up in an analogous way by first chosing a time axes t' inclined of angle ϕ : $\tan\phi = V$ and then constructing the spatial axes from the light rays. As clear from the picture, the second observer will obviously **not** agree on which events are simultaneous to an event p. For example, the planes $t = const$ and $t' = const$ are different in the figure, and none of the two frames is preferred. The two observers measure the same value of c for light speed, but in general they disagree on the values of dt and $d\ell$.

Summary 2.2.1. *Time in SR is not absolute; time intervals can have different values in different frame of reference. Without notion of simultaneity one cannot define distances (spatial intervals) in a observer-invariant way.*

2.3 Spacetime invariant interval

Question 2.3.1. *Does SR have an "invariant interval" similar to Galileian physics?*

One must consider the 4D spacetime and two inertial observers, $\mathcal{O}$ and $\mathcal{O}'$. Let us label coordinates as x^μ with $\mu = 0, 1, 2, 3$, where as above $(x^0, x^1, x^2, x^3) = (t, x, y, z)$ and events labels are placed as subscripts. Take 2 specific events:

(p) emission of a light pulse at spacetime coordinates x^μ_p for $\mathcal{O}$ and $x^{\mu'}_p$ for $\mathcal{O}'$;

(q) arrival of the light pulse at spacetime coordinates x^μ_q for $\mathcal{O}$ and $x^{\mu'}_q$ for $\mathcal{O}'$.

Because light propagates at c in both frames, one can calculate the spatial distances in the two frames as

$$\delta\ell = \sqrt{(x_q - x_p)^2 + (y_q - y_p)^2 + (z_q - z_p)^2} = c(t_q - t_p) \tag{2.5}$$

$$\delta\ell = \sqrt{(x'_q - x'_p)^2 + (y'_q - y'_p)^2 + (z'_q - z'_p)^2} = c(t'_q - t'_p) \ , \tag{2.6}$$

and observe that the quantity

$$s^2_{qp} = (x_q - x_p)^2 + (y_q - y_p)^2 + (z_q - z_p)^2 - c^2(t_q - t_p)^2 = (x'_q - x'_p)^2 + (y'_q - y'_p)^2 + (z'_q - z'_p)^2 - c^2(t'_q - t'_p)^2 \tag{2.7}$$

is always invariant for events connected by light and has value $s^2_{qp} = 0$ (light-like events).

Observations

- s^2_{qp} is a quadratic form "similar" to the Euclidean distance in $\mathbb{R}^4$ but with a "minus" sign.

- In infinitesimal notation we could write it in many ways:

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 = -c^2dt^2 + \sum_{i=1}^{3}(dx^i)^2 \tag{2.8a}$$

$$= \sum_{\mu=0}^{3}\sum_{\nu=0}^{3}\eta_{\mu\nu}dx^\mu dx^\nu \quad \text{with} \quad \eta_{\mu\nu} := \text{diag}(-c^2, 1, 1, 1) \quad \underline{\text{Flat/Minwowski/Lorentz metric}} \tag{2.8b}$$

$$= \eta_{\mu\nu}dx^\mu dx^\nu = dx_\mu dx^\mu \ , \tag{2.8c}$$

  where in the last line Einstein's sum-convention is used again.
- In general the ds^2 computed by two inertial observers must be infinitesimal of the same order [1], thus proportional to a function a of the relative velocity between the observers,

$$ds^2 = a\ ds^{2'} \quad \text{with} \quad a = a(|\vec{V}|) \ . \tag{2.9}$$

  The fact that the function a is only a function of the relative velocity follows from the basic assumptions of *homogeneity of spacetime* $\Rightarrow$ a cannot depend on x^μ, otherwise different point would not be equivalent; *isotropy of space* $\Rightarrow$ a cannot depend on $\hat{V}$, otherwise there would be a favourite direction.

The last observation allows one to show that

Theorem 2.3.1. *The spacetime interval $ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu$ between two events is invariant for all the inertial observers.*

Proof. *Take 3 inertial observers $\mathcal{O}, \mathcal{O}_1, \mathcal{O}_2$ with relative velocities $\vec{V}_1$ ($\mathcal{O}_1 - \mathcal{O}$), $\vec{V}_2$ ($\mathcal{O}_2 - \mathcal{O}$), $\vec{V}_{12}$ ($\mathcal{O}_1 - \mathcal{O}_2$). Using Eq. (2.10):*

$$\left.\begin{aligned} ds^2 = a(V_1)ds_1^2 = a(V_2)ds_2^2 \\ ds_1^2 = a(V_{12})ds_2^2 \end{aligned}\right\} \quad \Rightarrow \quad a(V_{12}) = \frac{a(V_1)}{a(V_2)} \ . \tag{2.10}$$

The l.h.s. of the last equation depends on the angle between $\vec{V}_1$ and $\vec{V}_2$ because the modulus V_{12} depends on V_1, V_2 and that angle. But the r.h.s. does not depend on the angle, thus $a(V) = const$. The only constant compatible with the equation is however one, thus $a(V) \equiv 1$.

Remark 2.3.1. *Despite its notation, the spacetime invariant is not positive definite. Two events are precisely characterized by the value of ds^2 in a observer independent (absolute) way. The spacetime interval between them can be*

$$ds^2 \begin{cases} = 0 \ : & \textit{lightlike or null} \\ < 0 \ : & \textit{timelike} \\ > 0 \ : & \textit{spacelike} \ . \end{cases} \tag{2.11}$$

Definition 2.3.1. Proper time = *the time interval measured by an observer at rest (carrying his clock). Since for an observer at rest $ds^2 = -c^2dt^2$, proper time is an invariant. Equivalently, one can define the proper time as the interval $dt^2 = -ds^2/c^2$ for a timelike worldline; proper time is thus the time elapsed between two events as measured by an observer moving on a straight path between the two events.*

Referrring again to the spacetime diagram of Fig. (2.2) it is easy to idenfy which event are null/timelike/spacelike w.r.t. to o.

- Region I:

$$\begin{cases} t > 0 & \Rightarrow \text{all event occur after } o \\ t > x & \Rightarrow \ -t^2 + x^2 < 0 \ \Rightarrow \ \text{timelike events} \end{cases} \tag{2.12}$$

  Region I is the future of event o. No events are simultaneous to o
- Region III:

$$\begin{cases} t < 0 & \Rightarrow \text{all event occur before } o \\ t > x & \Rightarrow \ -t^2 + x^2 < 0 \ \Rightarrow \ \text{timelike events} \end{cases} \tag{2.13}$$

  Region III is the past of event o. No events are simultaneous to o
- Region II and IV. Events are connected to o by spacelike intervals: they are causally disconnected from o. For any event in these region there exist a frame such that the event occur before/after/simultaneously to o [exercise].

2.4 Lorentz transformations

Inertial frames must be connected by transformations that leave the spacetime interval ds^2 invariant (Relativity principle). Let us find these trasformations.

[1] This just follows from the fact that $ds^2 \to 0 \Rightarrow ds^{2'} \to 0$ for arbitrary coordinates. See Schutz (1985) for a rigorous derivation.

Translations. The coordinate transformation

$$x^\mu \mapsto x^{\mu'} = x^\mu + a^\mu \quad \text{with} \quad a^\mu \in \mathbb{R}^4 \ , \tag{2.14}$$

leaves ds^2 invariant because $dx^\mu = dx^{\mu'}$. [Note: as in the equation above, the prime for the new coordinates $'$ is sometimes indicated with the index, instead of being "attached" to the x'. This is a slight abuse of notation as indexes at l.h.s. and r.h.s. "do not match", but it is often used.].

Lorentz Transformations. Consider the linear coordinate trasformation

$$x^\mu \mapsto x^{\mu'} = \Lambda^{\mu'}{}_\nu x^\nu \quad \text{with} \quad \Lambda^{\mu'}{}_\nu \ 4\times 4 \text{ matrix} \ , \tag{2.15}$$

or in matrix notation [2]

$$x' = \Lambda x \ . \tag{2.16}$$

Combining the above equation with the distance between two events in the two frames one obtains an equation for the Λ matrices

$$\begin{cases} ds^2 &= (dx)^{\mathrm{T}}\eta(dx) = (dx')^{\mathrm{T}}\eta(dx') \\ dx' &= \Lambda dx \end{cases} \quad \Rightarrow \quad ds^2 = (dx)^{\mathrm{T}}\eta(dx) = (\Lambda dx)^{\mathrm{T}}\eta(\Lambda dx) \quad \Rightarrow \quad \eta = \Lambda^{\mathrm{T}}\eta\Lambda \tag{2.17}$$

or in components

$$\eta_{\mu\nu} = \Lambda^{\mu'}{}_\mu \Lambda^{\nu'}{}_\nu \eta_{\mu'\nu'} \quad \underline{\text{Lorentz transformation}} \ . \tag{2.18}$$

Every coordinate transformation where Λ satisfies Eq. (2.18) preserves ds^2 by construction.

Eq. (2.18) reminds to the expression for **rotation** matrices R in 3D:

$$\delta_{ij} = R_i^{i'} R_j^{j'} \delta_{i'j'} \quad \text{or} \quad I = R^{\mathrm{T}} I R \ , \tag{2.19}$$

where $I := \mathrm{diag}(1,1,1)$ is the identity matrix. In fact, rotations are a special case of Lorentz transformations restricted to the spatial sector. If the time coordinate does not change, the first column and row in the l.h.s. of Eq. (2.18) are trivially satisfied and one needs only to deal with the "spatial" sub-block which is precisely Eq. (2.19). Schematically,

$$\eta = \begin{bmatrix} -1 & 0 \\ 0 & I \end{bmatrix} = \Lambda^{\mathrm{T}}\eta\Lambda = \begin{bmatrix} 1 & 0 \\ 0 & R^{\mathrm{T}} \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix} \ . \tag{2.20}$$

The Λ above matrix that describes a rotation of an angle $\theta \in (0, 2\pi]$ around the $\hat{z}$ axis is for example,

$$\Lambda^{\mu'}{}_\nu = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \ , \tag{2.21}$$

and one can verify immediately that Eq. (2.18) holds. A generic rotation is a combination of rotations around each axis, thus it is described by 3 angles.

In 4D one can additionally consider "rotations involving the time coordinate". For example, the matrix

$$\Lambda^{\mu'}{}_\nu = \begin{bmatrix} \cosh\phi & -\sinh\phi & 0 & 0 \\ -\sinh\phi & \cosh\phi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \ , \tag{2.22}$$

leads to the coordinate transformation

$$\begin{cases} t' &= t\cosh\phi - x\sinh\phi \\ x' &= -t\sinh\phi + x\cosh\phi \\ y' &= y \\ z' &= z \ , \end{cases} \tag{2.23}$$

which is similar to the rotation except that it is expressed with hyperbolic sine and cosine and $\phi \in (-\infty, \infty)$. The hyperbolic functions are clearly needed to fullfill Eq. (2.18) (now we are changing the time coordinate!) and one can easily prove by using

$$\cosh^2(x) - \sinh^2(x) = 1 \tag{2.24}$$

that the first two lines of Eq. (2.23) are the most general transformation preserving $ds^2 = -dt^2 + dx^2$ [Exercise].

What does these "time rotations" describe? Hint: inertial observers are moving with respect to each other by constant velocity ... the transformation must connect an inertial observer with another one in motion at constant

[2] Notation: the Λ matrix element $(\mu\nu)$ corresponds to $\Lambda^\mu{}_\nu$; the Λ^{T} matrix element $(\mu\nu)$ corresponds to $\Lambda^\nu{}_\mu$.

velocity. It is simple to see this explicitely. Consider the origin in $\mathcal{O}'$, i.e. point $x' = 0$; using Eq. (2.23) it is immediate to see that it moves at $\tanh\phi$ with respect to $\mathcal{O}$:

$$x' = 0 = -t\sinh\phi + x\cosh\phi \quad \Rightarrow \quad \frac{x}{t} = \frac{\sinh\phi}{\cosh\phi} = \tanh\phi =: V \ . \tag{2.25}$$

Using Eq. (2.24) to invert [exercise]

$$\begin{cases} \cosh\phi & = \left[1 - \left(\frac{V}{c}\right)^2\right]^{-1/2} =: \gamma \ , \quad \underline{\text{Lorentz factor}} \ (1 \leq \gamma < \infty) \\ \sinh\phi & = V\gamma \ , \end{cases} \tag{2.26}$$

one obtains the transformation in the most usual form (factors c are restored for clarity)

$$\begin{cases} t' & = \gamma(t - \frac{V}{c^2}x) \\ x' & = \gamma(x - Vt) \end{cases} \ . \tag{2.27}$$

These transformations are called **boosts**. A generic boost is a combination of boosts along each axis, thus it is described by 3 boost parameters (velocities).

Observations

- The *inverse* of a Lorentz transformation from the unprimed to the primed coordinates is also a Lorentz transformation, this time from the primed to the unprimed systems. Schematically

$$x^\mu \xleftrightarrow[(\Lambda^{-1})^{\nu'}{}_\mu]{\Lambda^{\nu'}{}_\mu} x^{\nu'} \quad \textit{i.e.} \quad x' = \Lambda x \quad \text{and} \quad x = \Lambda^{-1}x' \ . \tag{2.28}$$

  It is convenient to define the notation

$$\Lambda_{\nu'}^{\ \mu} := (\Lambda^{-1})^{\nu'}{}_\mu \tag{2.29}$$

  such that

$$x^{\nu'} = \Lambda^{\nu'}{}_\mu x^\mu \quad \text{and} \quad x^\mu = \Lambda_{\nu'}^{\ \mu} x^{\nu'} \ . \tag{2.30}$$

  and

$$\Lambda_{\nu'}^{\ \mu}\Lambda^{\rho'}{}_\mu = \delta^{\rho'}_{\nu'} \quad \text{and} \quad \Lambda_{\nu'}^{\ \mu}\Lambda^{\nu'}{}_\rho = \delta^\mu_\rho \tag{2.31}$$

  where $\delta^\mu_\mu = \mathrm{diag}(1,1,1,1)$.
- Lorentz transformations contain Galileian ones for sufficiently small velocities

$$V \ll c \text{ or } c \to \infty \quad \Rightarrow \quad \gamma \to 1 \text{ and } \frac{V}{c^2}\gamma \to 0 \quad \Rightarrow \begin{cases} t' & = t \\ x' & = x - Vt \end{cases} \tag{2.32}$$

- Boosts are undefined for $V > c$ as transformations in $\mathbb{R}^4$ because $(V/c)^2 > 1$ and γ becomes imaginary.
- No material body can move at $V = c$ because $\gamma \to \infty$.
- Boosts transformations generate time dilation and length contractions.

Remark 2.4.1. *Lorentz transformations form a six-parameter (3 rotations, 3 boosts) group called* Lorentz group*. The Lorentz group is nonabelian since the transformations do not commute. The set of translations and Lorentz transformation form a ten-parameter nonabelian group called the* Poincare' group*. The Poincare group encodes the isometries of SR.*

Example 2.4.1. Time dilation and proper time. *Consider two events o and p occurring at the same point in $\mathcal{O}$; the interval between the events is $dt = t_p - t_o$. Take $\mathcal{O}'$ in relative motion with relative velocity V; the time interval between the events is*

$$\left.\begin{aligned} t'_o &= \gamma(t_o - Vx_o) = \gamma t_o \\ t'_p &= \gamma(t_p - Vx_p) = \gamma t_p \end{aligned}\right\} \quad \Rightarrow \quad dt' = t'_p - t'_o = \gamma dt = (1 - V^2)^{-1/2} dt > dt \ , \tag{2.33}$$

This illustrates that proper time is the minimal *time interval measured by inertial frames for a given worldline between two events, i.e. time runs slower for moving clocks.*

However, consider now the same two events connected by two different worldlines: a straight line and a nonstraight line returning at the same spatial location (Cf. Twin paradox). In general, one worldline describes uniform motion (constant speed) from o to p but the second worldline must accelerate at some point, the simplest case being composed of moving away at constant speed, turning at halfway and returning at constant speed. In this case the Lorentz transformations do not apply because the second observer is noninertial. The proper time is given by the integral *of $-ds^2$ along the path. If δt is the time interval for the straight path, for the second simplest path it is trivial to calculate that each half travel at constant speed v takes $\sqrt{(\delta t^2/2)-(\delta x)^2}$ with $\delta x = \delta t\, v/2$. Hence, one obtains that the accelerated path has the* shorter *proper time,*

$$\delta t' = 2\times\sqrt{(\delta t^2/2)-(\delta x)^2} = 2\times\sqrt{(\delta t^2/2)-(v\delta t/2)^2} = \sqrt{1-v^2}\delta t < \delta t \ . \tag{2.34}$$

In Euclidean space straight lines are the shorter distance between two points; in Mikowski spacetime straight lines are the longest proper time interval between two events.

2.5 4-vectors and tensors in SR

A first definition of vectors, one-forms and tensors in terms of the transformation of the components.

Definition 2.5.1. 4-vector = *an object made of 4 (contravariant) components v^μ that under a Lorentz transformation change as*

$$v^{\mu'} = \Lambda^{\mu'}{}_{\mu} v^\mu \ . \tag{2.35}$$

From the definition one sees immediately that the scalar quantity given by the contraction of the 4-vector with the flat metric,

$$\eta_{\mu'\nu'}v^{\mu'}v^{\nu'} = \eta_{\mu'\nu'}\Lambda^{\mu'}{}_{\mu}\Lambda^{\nu'}{}_{\nu}v^\mu v^\nu = \eta_{\mu\nu}v^\mu v^\nu \ , \tag{2.36}$$

is invariant under Lorentz transformation. A vector is thus classified accordingly

$$v_\mu v^\mu \begin{cases} = 0 \ : & \text{null vector} \\ < 0 \ : & \text{timelike vector} \\ > 0 \ : & \text{spacelike vector} \ . \end{cases} \tag{2.37}$$

Remark 2.5.1. *If $x^\mu(\lambda)$ is a curve in the spacetime parametrized by λ, the components of the tanget vector at a given point are*

$$v^\mu = \frac{dx^\mu}{d\lambda} \ , \tag{2.38}$$

and Eq. (2.35) follows from the transformation of the coordinates, from the fact that λ is left unaltered by the Lorentz transformation, and from the linearity of the derivative. Vectors at a given point form a vector space called tangent space T_p.

Definition 2.5.2. covector/dual vector/1-form = *an object made of 4 (covariant) components w_μ that under a Lorentz transformation change as*

$$w_{\mu'} = \Lambda_{\mu'}{}^{\mu} w_\mu \ . \tag{2.39}$$

Remark 2.5.2. *Covectors form a tangent space T_p^* that is dual [3] to T_p:*

$$w(v) = w_\mu v^\mu \in \mathbb{R} \ . \tag{2.40}$$

One can interpret vectors as linear maps on dual vectors by defining

$$v(w) := w(v) = w_\mu v^\mu \in \mathbb{R} \ , \tag{2.41}$$

which shows dual space of the dual vector space is the original space $T_p^{**} = T_p$.

The invariant $v^\mu v_\mu$ can be considered as the contraction of the vector components v^μ with those of the associated covector computed as

$$v_\mu = \eta_{\mu\nu}v^\nu \ . \tag{2.42}$$

Note here $v^\mu v_\mu$ is analogous to the vector's norm given by the Euclidean scalar product but it can be zero or negative. It should also remind to the *bra* $\langle\Psi|$ and *ket* $|\varphi\rangle$ of quantum mechanics. Similarly, given covector components w_μ one obtains the components of the associated vector using

$$w^\mu = (\eta^{-1})^{\mu\nu}w_\nu = \eta^{\mu\nu}w_\nu \ , \tag{2.43}$$

where in the last passage we introduce a short notation for the inverse metric: $\eta_{\mu\rho}\eta^{\rho\nu} = \delta^\nu_\mu$. Note the inverse metric $\eta^{\mu\nu}$ has exactly the same components as $\eta_{\mu\nu}$ in flat space and in Cartesian coordinates.

In order to generalize the notion of vectors, it is useful to observe that the relation

$$\frac{\partial x^{\mu'}}{\partial x^\mu} = \frac{\partial}{\partial x^\mu}(\Lambda^{\mu'}{}_{\nu}x^\nu) = \Lambda^{\mu'}{}_{\nu}\frac{\partial x^\nu}{\partial x^\mu} = \Lambda^{\mu'}{}_{\nu}\delta^\nu_\mu = \Lambda^{\mu'}{}_{\mu} \ , \tag{2.44}$$

allows on to use the following notation for the transformation of the ctor/covector components

$$v^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu}v^\mu \ \text{ and } \ w_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}}w_\mu \ ; \tag{2.45}$$

Cf. notation introduced in Eq. (2.29).

[3] Recall the dual space is the space of all linear maps from the original vector to real numbers.

Definition 2.5.3. (k,l)-tensor = *an object with arbitrary components* $T^{\mu_1\cdots\mu_k}_{\nu_1\ldots\nu_l}$ *that under coordinate transformation change as*

$$T^{\mu'_1\cdots\mu'_k}_{\nu'_1\cdots\nu'_l} = \frac{\partial x^{\mu'_1}}{\partial x^{\mu_1}}\cdots\frac{\partial x^{\mu'_k}}{\partial x^{\mu_k}}\frac{\partial x^{\nu_1}}{\partial x^{\nu'_1}}\cdots\frac{\partial x^{\nu_k}}{\partial x^{\nu'_k}}T^{\mu_1\cdots\mu_k}_{\nu_1\cdots\nu_l} \tag{2.46}$$

Example: $(0,2)$ **tensors**

- $T_{\mu\nu}$ is a tensor of type $(0,2)$, the component transform as $T_{\mu\nu} = \frac{\partial x^{\mu'}}{\partial x^{\mu}}\frac{\partial x^{\nu'}}{\partial x^{\nu}}T_{\mu\nu}$.
- $T_{\mu\nu}$ is a *symmetric* $(0,2)$ tensor iff $T_{\mu\nu} = T_{\nu\mu}$. A generic $(0,2)$ tensor can be symmetrized $T_{(\mu\nu)} := (T_{\mu\nu}+T_{\nu\mu})/2$.
- $T_{\mu\nu}$ is a *antisymmetric* $(0,2)$ tensor iff $T_{\mu\nu} = -T_{\nu\mu}$. A generic $(0,2)$ tensor can be antisymmetrized $T_{[\mu\nu]} := (T_{\mu\nu} - T_{\nu\mu})/2$.
- In 4D $T_{\mu\nu}$ has $4\times 4 = 16$ independent components; $T_{(\mu\nu)}$ has $4+6 = 10$ indep. comp. (diagonal+lower tringular part); $T_{[\mu\nu]}$ has 6 indep. comp. (the diagonal is zero!).
- $\eta_{\mu\nu}$ is an example of $(0,2)$ tensor. In this case however the components remain unchanged in any Cartesian coordinate system in flat spacetime.

Remark 2.5.3. *The quantities $\eta_{\mu\nu}$, $\eta^{\mu\nu}$, δ^{μ}_{ν} are tensor (components) that, even though they all transform according to the tensor transformation law Eq. (2.46), their components remain unchanged in any Cartesian coordinate system in flat spacetime. In more general coordinate systems the components can change, except for the Kronecker delta that has exactly the same components in any coordinate system in any spacetime. This follows from the fact that tensors are, in their abstract definition, linear maps and the Kronecker tensor, among them, is the identity map.*

2.6 Kinematics

Concepts of 4-velocity and acceleration.

Definition 2.6.1. Worldline = *set of spacetime events corresponding to the motion of a particle (material body or photon). In general it is indicated as $x^\mu(\lambda)$, where λ is the curve parameter.*

Definition 2.6.2. *Given a worldline $x^\mu(\lambda)$, the* 4-velocity *is*

$$u^\mu := \frac{dx^\mu}{d\lambda} = \dot{x}^\mu \ . \tag{2.47}$$

While λ is a generic parameter, for timelike wordlines one can use the proper time calculated as

$$\tau = \int dt = \int \sqrt{-ds^2} = \int \sqrt{-\eta_{\mu\nu}dx^\mu dx^\nu} = \int \sqrt{-\eta_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}d\lambda \ , \tag{2.48}$$

and reparametrize the worldline with $\tau(\lambda)$. In this case the 4-velocity is normalized:

$$u^\mu = \frac{dx^\mu}{d\tau} \ \Rightarrow\ u^\mu u_\mu = \eta_{\mu\nu}u^\mu u^\nu = \eta_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = \frac{\eta_{\mu\nu}dx^\mu dx^\nu}{d\tau^2} = \frac{ds^2}{d\tau^2} = -1. \tag{2.49}$$

The above calculation can be repeated with the generic parametrization in λ leading to $u^\mu u_\mu = ds^2/d\lambda^2$: while one obviously loses the normalization, the sign of $u^\mu u_\mu$ is still determined by the sign of ds^2. This way the 4-velocity explains the link between the vector classification and the classification of spacetime intervals given above: for two events sufficiently close and laying on the same curve $x^\mu(\lambda)$ are separated by a null/timelike/spacelike interval if the norm of the 4-velocity (tangent vector to the curve) is null/timelike/spacelike. If along the curve the tangent vector remain (for all λ) the curve is called null/timelike/spacelike.

More properties.

- u^μ is defined dimensionless
- Components in the rest frame $u^\mu = (1,0,0,0)$.
- Components in a generic frame $u^\mu = (\gamma,\gamma v^i)$. This can be readily seen by boosting the rest frame component along e.g. x-direction: $u^{\mu'} = \Lambda^{\mu'}_{\mu}u^\mu = (\gamma,\gamma V,0,0)$ using Eq. (2.22).

Definition 2.6.3. *Given a worldline $x^\mu(\lambda)$, the* 4-acceleration *is*

$$a^\mu := \frac{d^2x^\mu}{d\lambda^2} = \dot{u}^\mu = \ddot{x}^\mu \ . \tag{2.50}$$

The acceleration is orthogonal to the velocity:

$$0 = \frac{d}{d\tau}(-1) = \frac{d}{d\tau}(u^\mu u_\mu) = \frac{d}{d\tau}(\eta_{\mu\nu}u^\mu u^\nu) = 2\eta_{\mu\nu}\dot{u^\mu}u^\nu = 2a^\mu u_\mu \ . \tag{2.51}$$

2.7 Dynamics of particles

The motion of material bodies is described by a Lagrangian L and a action $S = \int L dt$. Let us find an action for free bodies.

Since the Lagrangian has dimensions of energy $[L] = E$, the action has dimension $[S] = ET = ML^2T^{-1}$. Requiring that the action is Lorentz invariant, the simplest form for particles (timelike worldlines) is

$$S = K\int ds = K\int \sqrt{-\eta_{\mu\nu}u^\mu u^\nu} d\lambda = Kc\int d\tau = \int \underbrace{Kc\gamma^{-1}}_{=L} dt \ , \tag{2.52}$$

where c is included for clarity and K is a constant with dimensions $[K] = MLT^{-1}$. The latter can be determined by forcing the Lagrangian to the Newtonian limit

$$L_{\rm Newt} = \frac{1}{2}mv^2 \ . \tag{2.53}$$

Expanding for small velocities $v/c \ll 1$ and up to an overall constant (see below) one obtains [4]

$$L = K(-\eta_{\mu\nu}\dot{x}^\mu\dot{x}^\nu)^{1/2} = Kc\gamma^{-1} = Kc\sqrt{1-\frac{v^2}{c^2}} \approx Kc\left(1 - \frac{1}{2}\frac{v^2}{c^2} + \mathcal{O}(\frac{v^4}{c^2})\right) \quad \Rightarrow \quad K = -mc \ . \tag{2.54}$$

Note the action is invariant with respect changes of the parametrization of the worldline, since λ appears as mute variable in the integral.

Given the Lagrangian one can calculate the conjugate 3-momentum $\vec{p} = (p^i)$ and the Hamiltonian

$$p_i = \frac{\partial L}{\partial v^i} = -mc^2\frac{1}{2}\frac{-2}{\sqrt{1-\frac{v^2}{c^2}}}\frac{v_i}{c^2} = \gamma m v_i \tag{2.55}$$

$$H = \vec{p}\cdot\vec{v} - L = \gamma m v^2 - mc^2\sqrt{1-\frac{v^2}{c^2}} = \gamma m v^2 + mc^2\gamma^{-1} = m(\gamma v^2 + \gamma^{-1}c^2) = m\frac{v^2 + c^2(1-v^2/c^2)}{\sqrt{1-v^2/c^2}} = m\gamma c^2 \tag{2.56}$$

Observations

- Newtonian limit $v \ll c$: $\vec{p} \approx m\vec{v}$ and $H \approx mc^2 + mv^2/2$
- $v \to c \Rightarrow \vec{p} \to \infty$.
- For a particle at rest $\gamma = 1$ and $H = mc^2 \Rightarrow$ the energy of the particle has a contribution from the mass.

Equations of motion (EOM). Considering in general $L = L(x^\mu, \dot{x}^\mu)$ the Euler equations follows from the stationarity of the action. Denoting with δ the variation,

$$\delta S = 0 \quad \Rightarrow \quad \frac{\partial L}{\partial x^\mu} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^\mu} = 0 \ . \tag{2.57}$$

Applying the equations above to $L = K(-\eta_{\mu\nu}\dot{x}^\mu\dot{x}^\nu)^{1/2}$ one obtains:

$$0 = \underbrace{\frac{\partial L}{\partial x^\alpha}}_{=0} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{x}^\alpha} = -\frac{d}{d\tau}\frac{\partial}{\partial u^\alpha}(-\eta_{\alpha\beta}u^\alpha u^\beta)^{1/2} = +\frac{d}{d\tau}\left[\frac{1}{2}\left(-\eta_{\alpha\beta}u^\alpha u^\beta\right)^{-1/2} 2\eta_{\mu\nu}u^\nu\right] \quad \Rightarrow \quad \frac{d}{d\tau}u^\mu = a^\mu = 0 \ . \tag{2.58}$$

The last equation is trivial as there is always a multiplicative term $\dot{u}^\mu$ when taking the time derivative of the term in braket. The last expression of the first equation deserves instead a closer look as this is a calculation that will often return:

$$\frac{\partial}{\partial u^\mu}(-\eta_{\alpha\beta}u^\alpha u^\beta)^{1/2} = \frac{1}{2}\left(-\eta_{\alpha\beta}u^\alpha u^\beta\right)^{-1/2}\frac{\partial}{\partial u^\mu}(\eta_{\alpha\beta}u^\alpha u^\beta) \ ; \tag{2.59}$$

note the term under square root is a scalar and remember the Einstein's sum-convention, the interesting part is

$$\frac{\partial}{\partial u^\mu}(\eta_{\alpha\beta}u^\alpha u^\beta) = \frac{\partial}{\partial u^\mu}(\eta_{00}u^0u^0 + \eta_{11}u^1u^1 + \eta_{22}u^2u^2 + \eta_{33}u^3u^3) \tag{2.60a}$$

$$= \eta_{00}\underbrace{\frac{\partial u^0}{\partial u^\mu}}_{=\delta^0_\mu}u^0 + \eta_{00}u^0\frac{\partial u^0}{\partial u^\mu} + \ldots = 2\eta_{\mu 0}u^0 + \ldots \tag{2.60b}$$

$$= 2(\eta_{\mu 0}u^0 + \eta_{\mu 1}u^1 + \eta_{\mu 2}u^2 + \eta_{\mu 3}u^3) = 2\eta_{\mu\nu}u^\nu \ . \tag{2.60c}$$

[4] Recall the expansion $\sqrt{1-x^2} \approx 1 - x^2/2 + \mathcal{O}(x^4)$ for $x \ll 1$.

The result for the EOM is what one expects: a free body has no acceleration. In case presence of a force f^μ the EOM are

$$a^\mu = f^\mu \ . \tag{2.61}$$

For example, but without giving a full derivation, the EOM of a particle with charge q is

$$m\frac{du^\mu}{d\tau} = \frac{q}{c}F^{\mu\nu}u_\nu \ , \tag{2.62}$$

where $F^{\mu\nu}$ is the Faraday tensor (see below). The above equation can be derived from the action

$$S = -mc\int ds + \frac{q}{c}\int A_\mu dx^\mu \ , \tag{2.63}$$

where A_μ is the vector potential (see below for its definition).

Definition 2.7.1. *The* 4-momentum *is* $p^\mu := mcu^\mu = mc(\gamma, \gamma v^i) = (E/c, p^i)$.

Observations

- From the definition, $p_\mu p^\mu = m^2c^2u^\mu u_\mu = -m^2c^2$ and $p_\mu p^\mu = -E^2/c^2 + p^2$. Together these two equations implies the expression for the relativistic Hamiltonian with its appropriate Newtonian limit,

$$E = c\sqrt{p^2 + m^2c^2} \approx mc^2 + \frac{p^2}{2m} \ . \tag{2.64}$$

- The definition is consistent with the definition $p_\mu = -\partial S/\partial x^\mu$.

2.8 Dynamics of fields

The equations of motion for a classical field $\phi(x^\mu)$ (or a set of such fields) are derived from the action

$$S = \int L dt = \int dt \int d^3x \mathcal{L}[\phi; \partial_\mu\phi] \ , \tag{2.65}$$

written in terms of the Lagrangian density $\mathcal{L}$, Lorentz invariant. Schematically the EOM are obtained by

- Varying the fields and their derivatives

$$\phi \mapsto \phi + \delta\phi \ , \quad \partial_\mu\phi \mapsto \partial_\mu\phi + \delta(\partial_\mu\phi) = \partial_\mu\phi + \partial_\mu(\delta\phi) \ ; \tag{2.66}$$

- Varying the Lagrangian

$$\mathcal{L}[\phi + \delta\phi; \partial_\mu\phi + \partial_\mu(\delta\phi)] \approx \mathcal{L}[\phi; \partial_\mu\phi] + \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\partial_\mu\delta\phi \ ; \tag{2.67}$$

- Extremizing the action

$$0 = \delta S = \int d^4x \left[\frac{\partial\mathcal{L}}{\partial\phi}\delta\phi + \underbrace{\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\partial_\mu\delta\phi}_{b.p.}\right] = \int d^4x \frac{\partial\mathcal{L}}{\partial\phi}\delta\phi - \int d^4x \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\right)\delta\phi + \underbrace{\int d^4x \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\delta\phi\right)}_{\text{total derivative}} \tag{2.68a}$$

$$\Rightarrow \quad \frac{\partial\mathcal{L}}{\partial\phi} - \partial_\mu\left(\frac{\partial\mathcal{L}}{\partial\partial_\mu\phi}\right) = 0 \tag{2.68b}$$

where the second term is integrated by parts and the total derivative give a boundary term which is set to zero assuming $\delta\phi|_{\text{boundary}} = 0$. The EOM follows from the remaning terms since the field variation is generic and non zero.

Example 2.8.1. *The Lagrangian density of the scalar field φ with potential V is*

$$\mathcal{L} = -\frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - V(\varphi) \ , \tag{2.69}$$

and it is manifestly Lorentz invariant. The EOM

$$\frac{\partial\mathcal{L}}{\partial\varphi} = -\frac{dV}{d\varphi} \ , \quad \frac{\partial\mathcal{L}}{\partial\partial_\mu\varphi} = -\eta^{\alpha\mu}\partial_\alpha\varphi \quad \Rightarrow \quad 0 = \partial_\mu\left(\eta^{\mu\nu}\partial_\nu\varphi\right) - V' = \Box\varphi - V' \ . \tag{2.70}$$

Note that above we have defined $\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu = -\partial_t^2 + \sum_{i=1}^3 \partial_i^2$. It is left as exercise to find the dimentions of $\mathcal{L}$ in units $c = G = \hbar = 1$, and specify the equation to the Klein-Gordon potential $V = m^2\varphi^2/2$.

2.9 Maxwell equations

Maxwell equations can be written in an explicitely Lorentz invariant form using 4D tensor quantities.

Let us start from

$$\nabla \times \vec{B} - \partial_t \vec{E} = 0 = \epsilon^{ijk}\partial_j B_k - \partial_t E^i \tag{2.71a}$$
$$\nabla \cdot \vec{E} = 4\pi\rho = \partial_i E^i \tag{2.71b}$$
$$\nabla \times \vec{E} - \partial_t \vec{B} = 0 = \epsilon^{ijk}\partial_j E_k - \partial_t B^i \tag{2.71c}$$
$$\nabla \cdot \vec{B} = 0 = \partial_i B^i \ , \tag{2.71d}$$

where the blue equations are written down in terms of components rather than in vector notation. Note that spatial indexes $i, j, k, ... = 1, 2, 3$ can be put up or down arbitrarily because the spatial part of the Lorentz/Minkowski metric is the identiy, $\eta_{ij} = \eta^{ij} = \delta_{ij}$. The quantity ϵ^{ijk} is the usual Levi-Civita symbol; its values are $+1$ for even permutation of the index, -1 for odd, 0 if two indexes are the same (for 3 indexes even permutation coincides with cyclc permutations, odd permutations with anticyclic). Note the definition generalizes to higher dimensions/multiple indexes.

Since the divergence of the magnetic field is zero, the magnetic field can be expressed in terms of a vector potential. Substituing into the "curl equation" for the electric field, the latter can be expressed as combination of the vector and a scalar potential (See Jackson (1975) for the derivation of some of the equations of his paragraph),

$$\vec{B} = \nabla \times \vec{A} \ , \quad \vec{E} = \nabla\phi - \partial_t \vec{A} \ . \tag{2.72}$$

In terms of the potentials Maxwell equations read

$$\nabla\phi - \partial_t \nabla \cdot \vec{A} = \rho \ , \quad \Box\vec{A} + \nabla(\nabla \cdot \vec{A} - \partial_t \phi) = 4\pi \vec{J} \ . \tag{2.73}$$

The potentials definition in Eq. (2.72) is up to scalar function,

$$\phi \mapsto \phi - \partial_t \chi \ , \quad \vec{A} \mapsto \vec{A} + \nabla\chi \ , \tag{2.74}$$

that determines the gauge freedomn of electromagnetism. Choosing the Lorentz gauge

$$\nabla \cdot \vec{A} - \partial_t \phi = 0 \ , \tag{2.75}$$

leads to wave equations for the potentials

$$\Box\phi = \rho \ , \quad \Box\vec{A} = 4\pi\vec{J} \ . \tag{2.76}$$

Towards a manifestly covariant set of equation one defines the 4-vectors

Definition 2.9.1. *The* 4-potential $A^\mu = (\phi, A^i)$ *and the* 4-current $J^\mu = (\rho, J^i)$. *(Note $A_\mu = \eta_{\mu\nu}A^\nu = (-\phi, A_i)$.)*

Maxwell equation in terms of the 4-potential becomes the single equation

$$\Box A^\mu = 4\pi J^\mu \quad \text{with} \quad \partial_\mu A^\mu = 0 \quad \text{(Lorentz gauge)} \ , \tag{2.77}$$

which is manifestly Lorentz invariant.

Definition 2.9.2. Faraday/Maxwell tensor $F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu$.

Properties

- Antisymmetric by construction $F_{\mu\nu} = -F_{\nu\mu}$.
- Electric field: $F_{0i} = \partial_0 A_i - \partial_i A_0 = \partial_t A_i - \partial_i \phi = -E_i$.
- Magnetic field: $F_{ij} = \partial_i A_j - \partial_j A_i = \epsilon_{ijk}B^k$.
- Matrix form

$$F_{\mu\nu} = \begin{bmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & B_3 & -B_2 \\ E_2 & -B_3 & 0 & B_1 \\ E-3 & B_2 & -B_1 & 0 \end{bmatrix} . \tag{2.78}$$

Inverse $F^{\alpha\beta} = \eta^{\mu\alpha}\eta^{\nu\beta}F_{\mu\nu}$

$$F^{\mu\nu} = \begin{bmatrix} 0 & +E_1 & +E_2 & +E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E-3 & B_2 & -B_1 & 0 \end{bmatrix} , \tag{2.79}$$

Note it changes only the first row and columns; the nonzero components are

$$F^{0i} = E^i \ , \quad F^{ij} = \epsilon^{ijk}B_k \ . \tag{2.80}$$

- Lorentz transformation (verify it is a tensor) $F_{\mu'\nu'} = \Lambda^{\mu}{}_{\mu'}\Lambda^{\nu}{}_{\nu'}F_{\mu\nu}$. For example a boost in x-direction gives

$$\begin{cases} E_1' &= E_1 \\ E_2' &= \gamma(E_2 - VB_3) \\ E_3' &= \gamma(E_3 - VB_2) \end{cases} \quad \text{and} \quad \begin{cases} B_1' &= B_1 \\ B_2' &= \gamma(B_2 - VE_3) \\ B_3' &= \gamma(B_3 - VE_2) \end{cases} \tag{2.81}$$

that are the trasformations for the electric and magnetic fields derived in Jackson (1975).

To write the Maxwell equations in terms of the Faraday tensor let us compute the derivatives of $F^{\mu\nu}$

$$\partial_i F^{0i} = \partial_i(\eta^{00}\eta^{ii}F_{0i}) = \partial_i E^i \tag{2.82a}$$

$$\partial_\mu F^{i\mu} = \partial_0 \underbrace{F^{i0}}_{=-E^i} + \partial_j \underbrace{F^{ij}}_{=\epsilon^{ijk}B_k} = -\partial_t E^i + \epsilon^{ijk}\partial_j B_k \ . \tag{2.82b}$$

Comparing to Eq. (2.71), Maxwell equations are rewritten as

$$\partial_\mu F^{\nu\mu} = 4\pi J^\mu \tag{2.83a}$$

$$\partial_{[\mu}F_{\nu\lambda]} = \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \ , \tag{2.83b}$$

where the first equation corresponds to Eq. (2.71b) and Eq. (2.71a), and the second equation to the other two (derivation left as exercise.)

Remark 2.9.1. Charge conservation *follows from antisymmetry by taking a derivative of the first equation:*

$$0 = \underbrace{\partial_\mu\partial_\nu}_{sym} \underbrace{F^{\mu\nu}}_{asym} = 4\pi\partial_\sigma J^\sigma \ . \tag{2.84}$$

The above equation is the 4-divergence of the 4-current; taking the integral of a 3D volume Σ *closed by a 2D surface* $\partial\Sigma$ *and using Gauss theorem, one gets the usual expression*

$$\frac{dQ}{dt} = \int_\Sigma \partial_t\phi = \int_\Sigma \nabla\cdot\vec{J} = \int_{\partial\Sigma} \vec{J}\cdot\hat{n} \ , \tag{2.85}$$

stating that the charge variation follows from the current (flux) through $\partial\Sigma$.

Remark 2.9.2. *The maxwell equations can be derived from the following* Lagrangian density

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{c}J_\mu A^\mu \ . \tag{2.86}$$

3. Manifolds & Tensors

④

These lectures in differential geometry introduce the definitions of manifolds, tensors and other objects to arrive to the presentation of Stokes theorem. The concepts of metric and stress energy tensor are introduced here.

Suggested readings. *Chap. 2 of Wald (1984); Chap. 2 of Carroll (1997); Chap. 2-4,6 of Schutz (1985); O'Neill (1983) book.*

3.1 Differential geometry language for GR

An event in spacetime is characterized by four numbers x^μ. In pre-relativity and SR spacetime is globally in a one-to-one correspondence with $\mathbb{R}^4$. In GR, instead, the spacetime is not "fixed" but determined by the matter content via Einstein equations. Thus, we need to introduce and use some more general mathematical formalism that allows us to describe non-Euclidean, arbitrary geometries, and write differential equations in those geometries. This is what differential geometry provides us with.

Example 3.1.1. *Measure distances and describe waves on the* 2-sphere S^2. *Imagine to live in this 2D world: locally a sphere behaves like* $\mathbb{R}^2$ *and one can compute small distances from point A to point B as* $\bar{AB} = \sqrt{(x_B - x_A)^2 - (y_B - y_A)^2}$, *but measuring global distances need something else. Note there is no one-to-one correspondence between all the points on the sphere and the two coordinates (an subset of* $\mathbb{R}^2$*): using the usual angles* (ϕ, θ)*, the polar angle* ϕ *is undefined at the poles of* S^2*. A way to approach this problem is to* embed S^2 *in* $\mathbb{R}^3$*, i.e. consider the sphere as immersed in a Euclidean space of higher dimension (3D). This approach allows one to calculate meridians or parallel as path of minimal lenght between 2 points on the sphere. It also allows to write wave equations on* S^2 *using the standard transformation from Cartesian to spherical coordinates (and "throwing away terms" i.e. reducing the equation from 3D to 2D). However, this is not the aproach we can take in GR because (i) we do not experience higher dimensions than 4, and (ii) we do not know in advance the geometry of spacetime. Another possibility to approach the* S^2 *problem could be to employ multiple (at least 2) coordinate patches that are slighly overlapping (think of wrapping up a ball with small pieces of a paper sheet) and for which the points on the sphere covered by the patch are in one-to-one correspondence with the two coordinates of the patch. This goes in the right direction ...*

3.2 Manifold

The concept of *manifold* is introduced in order to map a generic set into $\mathbb{R}^n$. Such a map is not always possible for the whole set (global map), but it is often possible *locally.* The example is the atlas that covers Earth with many charts. A manifold is such a "structure" used to map a generic geometry into a Eucliden space, eventually piece-by-piece.

Definition 3.2.1. Map between two sets $\phi : M \mapsto N$ *such that is assign to each element* $p \in M$ *one element* $q \in N$.

A map can be
- *injective* : $\forall$ elements of N $\exists$ at most one element of M;
- *surjective* : $\forall$ elements of N $\exists$ at least one element of M;
- *bijective* (invertible) : injective and surjective, i.e. $\exists \phi^{-1}$ such that the composition $\phi^{-1} \circ \phi(p) = p$.
- None of the above.

Definition 3.2.2. Open ball *in* $\mathbb{R}^n$ *of radius* r *around* $y \in \mathbb{R}^n$ *is the set of points:* $B_r = \{x \in \mathbb{R}^n \; : \; |x - y| = \left(\sum_{i=1}^n (x_i - y_i)^2\right)^{1/2} < r\}$.

Definition 3.2.3. Open set *in* $\mathbb{R}^n$ = *any set of* $\mathbb{R}^n$ *that can be expressed as a union of balls.*

Definition 3.2.4. Manifold *is a set* M *together with a set of subsets* $\{O_\alpha \subset M\}$ *such that*
1. $\{O_\alpha\}$ *cover* M*, i.e. each* $p \in M$ *is contained in a* at least *one of the* O_α.

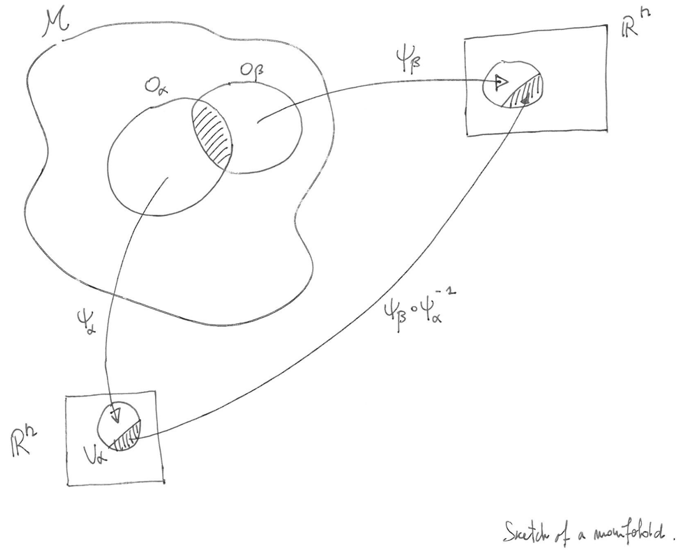

Figure 3.1: Illustration of a manifold.

2. $\forall\ \alpha\ \exists$ a bijective map to an <u>*open*</u> *subset of $\mathbb{R}^n$, $\psi_\alpha : O_\alpha \mapsto U_\alpha \subset \mathbb{R}^n$.*

3. If any two sets overlap $O_\alpha \cap O_\beta$, then the function $\psi_\beta \circ \psi_\alpha^{-1} : O_\alpha \cap O_\beta \subset \mathbb{R}^n \mapsto \mathbb{R}^n$ is $\mathcal{C}^\infty$.

The picture is shown in Fig. (3.1). ψ_α is called a chart *or a* coordinate system. *$\{\psi_\alpha\}$ is called an atlas. The manifold $(M, \{O_\alpha\}, \{\psi_\alpha\})$ is usually shortly indicated with $\mathcal{M}$. The dimension of $\mathcal{M}$ is n.*

Examples

- $\mathcal{M} = \mathbb{R}^n$ is a manifold: $O = \mathbb{R}^n$ (one chart), $\psi =$ *identity*.
- *Unit circle* S^1 The natural coordinate system for the unit circle is the angle $\theta : S^2 \mapsto \mathbb{R}$. Does that single chart define a manifold? Check properties 1)-3) of the Def. 3.2.4.
  1) $\theta \in [0, 2\pi)$ or $\theta \in (0, 2\pi]$ cover S^1. Ok.
  2) Including $\theta = 0$ (or $\theta = 2\pi$) gives a <u>close</u> interval in $\mathbb{R}$. Excluding both $\theta = 0, 2\pi$ does not cover the full circle.
  $\Rightarrow$ the circle cannot be covered with one chart !

  Let us try with two charts given by two angles each spanning only a part of the circle and overlapping on two arcs segments. Say, $\theta_1 \in (0, 5/4\pi)$ ad $\theta_2 \in (\pi, \pi/4)$. By construction the two sets are now open, and cover the unit circle. A point on the unit circle is identified with the value of one of these two angles. Hence S^1 with these two charts is a manifold. In general, one can use an atlas with more than two of similarly constructed charts and still make a manifold.

  Another option is to use stereographic projections. Consider the circle embedded in standard Euclidean coordinates in $\mathbb{R}^2$. One can define two charts $P_{N,S}$ by taking the north or south pole of the circle, finding any other point on the circle and projecting the line segment onto the x-axis. This provides the mapping from a point of S^1 to $\mathbb{R}^1$. Clearly each of the maps excludes the other pole. They cover the circle (property 1.) and maps to open sets of $\mathbb{R}$ (2.). Note the local coordinates for the charts are different: the same point on the circle mapped via the two charts do not map to the same point in $\mathbb{R}$. Explicitely, the map is [1]

$$u_p = P_N(p) = P_N(x_p, y_p) = \frac{x_p}{1 - y_p} \ , \quad u'_p = P_S(p) = P_S(x_p, y_p) = \frac{x_p}{1 + y_p} \ . \tag{3.1}$$

  The inverse maps are $P^{-1}_{N/S}(u_p) = (x_p, y_p)$ such that

$$x_p = \frac{2u_p}{u_p^2 + 1} \ , \quad y_p = \pm \frac{u_p^2 - 1}{u_p^2 + 1} \ , \tag{3.2}$$

  and the composition is

$$u'_p = P_S \circ P_N^{-1}(u_p) = P_S(P_N^{-1}(u_p)) = P_S\left(\frac{2u_p}{u_p^2 + 1}, \frac{u_p^2 - 1}{u_p^2 + 1}\right) = \frac{1}{u_p} \ . \tag{3.3}$$

[1] Recall the circle has unit radius and the triangles formed by the (pole-point on x-axis-point on S^1) are similar.

The latter is $\mathcal{C}^\infty$ in the overlap region, i.e. for all points of the circles except the poles. This verifies property 3.

- An example of a set which is not a manifold is a plane with a line ending on it. It is left as an exercise to discuss why (no calculations required, just think of an argument).

Remark 3.2.1. *A manifold is a space that is locally like $\mathbb{R}^n$ but globally different (generic). This allows one to import all the analysis tools from Euclidean spaces to generic spaces! For example, a map $f : \mathcal{M} \mapsto \mathcal{M}'$ connecting the points of the two manifolds $p' = f(p)$ can be identified, and should be always understood as, a "regular" function*

$$F : \psi_\alpha \circ f \circ \psi'_\alpha : \mathbb{R}^n \mapsto \mathbb{R}^{n'} \ . \tag{3.4}$$

We abandon immediately the notation f and F and always consider "functions on manifolds" in the above sense. In particular one talks of smoothness of f thinking of $F \in C^\infty$. A smooth and invertible function is called diffeomorphism, *an important concept for GR.*

3.3 Tangent vector space

Introduce the concept of tanget vector space at a point of the manifold by starting with the Euclean analogy.

Euclidean geometry. Here one uses vectors to describe "displacements". In $\mathbb{R}^n$ one has that

- Vectors form a vector space;
- Vectors are defined globally in $\mathbb{R}^n$. There is a intuitive notion of "rigidly transporting" vectors at different points, sum and subtract them;
- A vector introduces a <u>direction</u>, and can be naturally associated to the <u>derivative</u> of a function in that direction

$$\text{vector} \quad \underbrace{\vec{v} = (v_1, ..., v_n)}_{direction} \quad \leftrightarrow \quad \text{directional derivative} \qquad \underbrace{\sum_{\mu=0}^{n-1} v^\mu \partial_\mu}_{derivative\ in\ direction\ \vec{v}} \qquad . \tag{3.5}$$

In particular, given a curve $\gamma : \mathbb{R} \mapsto \mathbb{R}^n$ with $p = x^\mu(\lambda) \in \gamma \subset \mathbb{R}^n$, the *tangent vector* to γ at point p is made of components

$$v^\mu(p) = \frac{dx^\mu}{d\lambda} \ . \tag{3.6}$$

Any function derivative long γ is

$$\frac{df}{d\lambda} = \sum_{\mu=0}^{n-1} \frac{dx^\mu}{d\lambda} \frac{\partial_\mu f}{\partial x^\mu} = v^\mu \partial_\mu f \ . \tag{3.7}$$

Recall that the properties of derivatives are <u>linearity</u> and <u>Leibnitz rule</u>.

Arbitrary geometry. The concept of tangent vector space is globally lost and it is not obvious how to transport vectors "rigidly" from point to point. Rather, tangent vectors are defined only at a point. The idea can be visualized by embedding the manifold in a larger $\mathbb{R}^n$ space. If one thinks for example to S^2 in $\mathbb{R}^3$, a vector on the surface at point p "goes out" of the sphere in a neighbourg of p and has thus meaning only at point p; vectors on the sphere are meaningful only to describe <u>infinitesimal displacements</u>.

In order to generalize the concept of vectors to generic geometries, and to give an intrinsic way of characterizing them without the need of embedding the space, one uses theirs identification with direction derivatives. Consider the set of smooth functions $\mathcal{F} = \{f : M \mapsto \mathbb{R} \ : f \in \mathcal{C}^\infty\}$.

Definition 3.3.1. Tangent vector space *at point $p \in \mathcal{M}$ is $T_p\mathcal{M} = \{v : \mathcal{F} \mapsto \mathbb{R}\}$ such that the map $v \in T_p\mathcal{M}$ is*

1. *linear, $v(af + bg) = av(f) + bv(g)$ with $a, b \in \mathbb{R}$ and $f, g \in \mathcal{F}$;*
2. *Leibniz rule, $v(fg) = fv(g) + gv(f)$;*

and $T_p\mathcal{M}$ is a vector space,

3. *$(v_a + v_2)(f) = v_1(f) + v_2(f)$;*
4. *$(av)(f) = av(f)$.*

Note that if $\{e_\mu \in T_p\mathcal{M}\}$ is a set of n independent vectors, then any vector of the vector space can be written as $v = \sum_\mu v^\mu e_\mu$ such that its action on a any function is $v(f) = \sum_\mu v^\mu e_\mu(f)$.

Properties

1. $f(p) = const \equiv K \Rightarrow v(f) = 0$.
   Proof.

$$v(f^2) = \begin{cases} v(Kf) = Kv(f) \text{ linearity} \\ f(p)v(f) + f(p)v(f) = 2Kv(f) \text{ Leibnitz} \end{cases} \Rightarrow v(f) = 0 \ . \tag{3.8}$$

2. Dimension and basis are determined by the following

Theorem 3.3.1. *The dimension of $T_p\mathcal{M}$ is n and $v \in T_p\mathcal{M}$ is such that $v = \sum_\mu v^\mu \partial_\mu$ with $\partial_\mu := \partial/\partial x^\mu$ are the partial derivatives in a coordinate basis.*
Proof. *Introduce a basis of $T_p\mathcal{M}$ composed of the n vectors $\{e_\mu\}$ and show that the choice $e_\mu = \partial_\mu$ corresponds to (i) n independent vectors; (ii) spans $T_p\mathcal{M}$. $\{\partial_\mu\}$ is called* natural basis *or* coordinate basis.
(i) Take local coordinates $\psi(p) = x^\mu \in \mathbb{R}^n$. The coordinate representation of f is $f \circ \psi^{-1}$, that is a function from $\mathbb{R}^n$ to $\mathbb{R}$. The set $\{\partial_\mu := \partial/\partial x^\mu\}$ is composed of n tangent vectors acting on f that are linearly independent (each is a derivative in one of the directions of the coordinate axes). The action of each basis vector is

$$e_\mu(f) = \partial_\mu(f) = \frac{\partial}{\partial x^\mu}(f \circ \psi^{-1})|_{\psi(p)} \ . \tag{3.9}$$

(ii) A fundamental calculus theorem for the function $F : O \subset \mathbb{R}^n \mapsto \mathbb{R}$, where O the unit radius ball centered at the origin, says that

$$F(x) - F(0) = \int_0^1 \frac{d}{dt} F(tx^1, ..., tx^n)dt = \sum_\mu x^\mu \underbrace{\int_0^1 \frac{\partial}{\partial x^\mu} F(tx^1, ..., tx^n)dt}_{=:H_\mu(x^\nu)} = \sum_\mu x^\mu H_\mu(x^\nu) \ . \tag{3.10}$$

Note that $H_\mu(0) = \partial F/\partial x^\mu|_0$. For a function on manifolds one takes $F = f \circ \psi^{-1}$, $\psi(p) = 0$, and the result can be re-written for any $q \in O$ close to p as

$$f(q) = f(p) + \sum_\mu \left(x^\mu \circ \psi(q) - \underbrace{x^\mu \circ \psi(p)}_{=0} \right) H_\mu \circ \psi(q) \ . \tag{3.11}$$

Stress: this is just rewriting the equation above in the way it should be interpreted for functions on manifolds. Apply the vector (directional derivative at p) to the generic $f = f(q)$:

$$v\,(f(q)) = \underbrace{v\,(f(p))}_{=0} + v\left(\sum_\mu (x^\mu \circ \psi(q))\, H_\mu \circ \psi(q)\right) = \sum_\mu v\,(\,(x^\mu \circ \psi(q))\, H_\mu \circ \psi(q)\,) \tag{3.12a}$$

$$= \sum_\mu \underbrace{(x^\mu \circ \psi(p))}_{=0} v\,(H_\mu \circ \psi) + \sum_\mu v\,(x^\mu \circ \psi)\, H_\mu \circ \psi(p) \tag{3.12b}$$

$$= \sum_\mu \underbrace{v\,(x^\mu \circ \psi)}_{=:v^\mu} \partial_\mu f = \sum_\mu v^\mu \partial_\mu f \quad \forall f \ . \tag{3.12c}$$

The first line uses linearity and the results above that $v(K) = 0$. The second line uses Leibnitz. The third lines follows from the fact that $H_\mu \circ \psi(p) = H_\mu(\psi(p)) = \partial_\mu f$ and the definition of the vector component as the values of v applied to the function $x^\mu \circ \psi$. Hence, any vector at point p is generated by the coordinate basis.

3. Under a change of coordinate $x^\mu \mapsto x^{\mu'}$, the basis changes as

$$\partial_\mu \mapsto \partial_{\mu'} = \sum_\mu \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial}{\partial x^\mu} = \sum_\mu \frac{\partial x^\mu}{\partial x^{\mu'}} \partial_\mu \tag{3.13}$$

where $\frac{\partial x^\mu}{\partial x^{\mu'}}$ is the Jacobian of the transformation. Since the value of $v(f)$ cannot depend on the basis choice, the components must change as

$$v^{\mu'} = \sum_\mu \frac{\partial x^{\mu'}}{\partial x^\mu} v^\mu = \frac{\partial x^{\mu'}}{\partial x^\mu} v^\mu \ , \tag{3.14}$$

where in the last expression the Einstein sum-convention is used for the first time in this chapter. The transformation law above is the same derived int he context of SR, but it is not generic and valid for every spacetime/geometry.

Definition 3.3.2. Tangent vector field = *assignment of $T_p\mathcal{M}$ for each point of the manifold.*

Observations.

1. If $p \neq q$, then $T_p\mathcal{M} \neq T_q\mathcal{M}$. Tangent spaces at different points are different !
2. $\forall p$, the action of the vector $v(f) \in T_p\mathcal{M}$ is a function $\mathcal{M} \mapsto \mathbb{R}$. The vector field at point p is said *smooth* iff $\forall f \in \mathcal{C}^\infty \Rightarrow v(f) \in \mathcal{C}^\infty$. Note that if ∂_μ and v are smooth, then the components v^μ are smooth functions in $\mathbb{R}^n \mapsto \mathbb{R}$.
3. From now on we will **not** distinguish between tangent vectors and tangent vector fields and their relative spaces. The context should be sufficient to clarify which object one is using. The extension from tensors to tensors field is analogous and we won't bother too much also in that case. In most of GR applications we will deal with tensor fields.

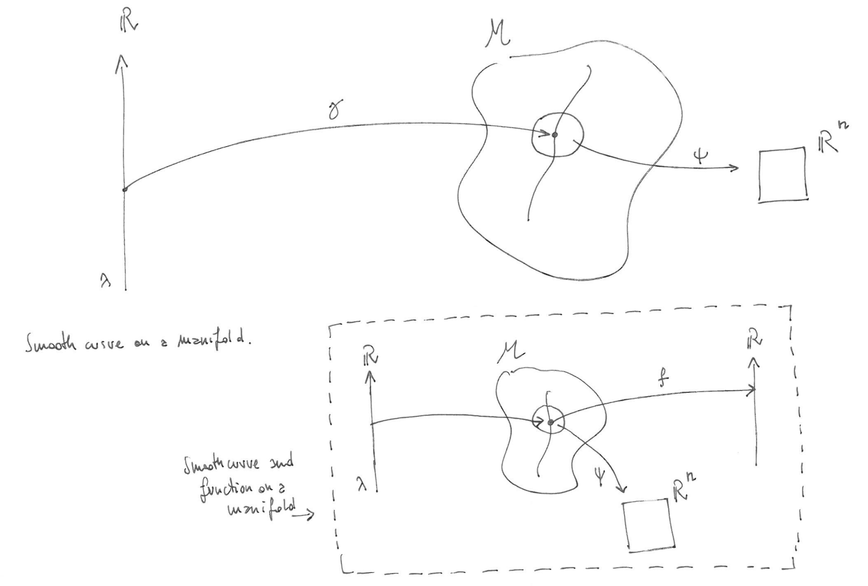

Figure 3.2: Illustration of a curve on manifold. In the inset a function on the manifold is also illustrated.

3.4 Smooth curve on $\mathcal{M}$

The definition of vector given above is compatible with what one expects from vector tangent to a curve.

Definition 3.4.1. *A* smooth curve $\mathcal{M}$ *is a* $\mathcal{C}^\infty$ *function* $\gamma : \mathbb{R} \mapsto \mathcal{M}$*, associating a real parameter to points of the manifold,* $\gamma(\lambda) = p$*.*

For each point of the curve one can define the *tangent vector* $v_\gamma : \mathcal{F} \mapsto \mathbb{R}$ such that

$$v_\gamma(f) \equiv \frac{d}{d\lambda}(f \circ \gamma) = \frac{d}{d\lambda}(f \circ \psi^{-1} \circ \psi \circ \gamma) = \frac{d}{d\lambda}(\underbrace{f \circ \psi^{-1}}_{\mathbb{R} \mapsto \mathbb{R}^n} \circ \underbrace{\psi \circ \gamma}_{x^\mu(\lambda)}) = \underbrace{\frac{dx^\mu(\lambda)}{d\lambda}}_{=v^\mu_\gamma} \underbrace{\frac{\partial}{\partial x^\mu}}_{\text{basis}}(f \circ \psi^{-1}) \ . \tag{3.15}$$

This shows that the components of the tangent vector are the partial derivatives of the coordinates, as expected.

Observations

- Given a curve in local coordinates $x^\mu(\lambda)$, one can find the tangent vector at all points from its components $v^\mu = dx^\mu/d\lambda$.
- Given a vector at one point, one can construct the curce γ by solving the ODE $dx^\mu/d\lambda = v^\mu$ with the given initial condition. This is a ODE system of 1st order for which local existance and uniqueness of the solution is guaranteed.

Remark 3.4.1. Vector fields as generators of diffeomorphisms. *The tangent vector can be interpreted as infinitesimal displacement. Consider the 1-parameter family of diffeomorphisms*

$$\phi_t : \mathbb{R} \times \mathcal{M} \mapsto \mathcal{M} \ \ , t \in \mathbb{R} \ \ \textit{such that} \ \ \phi_t \circ \phi_s = \phi_{t+s} \ \ \forall\, t, s \in \mathbb{R} \ . \tag{3.16}$$

The map describes a smooth transformation between points of the manifold and ϕ_0 *is the identity map. Given* $p \in \mathcal{M}$ *one can define the curve through* p *as*

$$\gamma_p := \phi_t : \mathbb{R} \mapsto \mathcal{M} \ , \textit{such that} \ \phi_0(p) = p \ . \tag{3.17}$$

This curve is now associated with a tangent vector at each point which is the infinitesimal generator of the transformations on the manifold.

Summary 3.4.1. *The definition of tangent vector is compatible and extend the definition of 4-vectors given in SR. The abstract definition given here is "intrinsic" and apply to arbitrary geometries, dimensions, and coordinate systems. Key steps to the definition: (i) intuitive idea of tangent vectors as infinitesimal displacements; (ii) vectors as direction derivatives; (iii) compatibility with the tangent to a curve at given point.*

3.5 Dual or cotangent vector space

Definition 3.5.1. *The* dual space *of $T_p\mathcal{M}$ is the set of linear maps $T_p^*\mathcal{M} = \{\omega : T_p\mathcal{M} \mapsto \mathbb{R},\ linear\}$.*

Observations

- $T_p^*\mathcal{M}$ is a vector space and ω is called *dual vector* or *1-form.*
- *Basis of $T_p\mathcal{M}$.* Given a basis $\{e_\mu\}$ of $T_p\mathcal{M}$ one defines the set $\{e^{*\mu} \in T_p^*\mathcal{M}\}$ such that $e^{*\mu}(e_\nu) = \delta^\mu_\nu$.
  (i) $\{e^{*\mu}\}$ is a basis of $T_p^*\mathcal{M}$ and $\omega = \omega_\mu e^{*\mu}$.
  (ii) The dimension of $T_p^*\mathcal{M}$ is the same as the dimension of $T_p\mathcal{M}$.
  (iii) The action of a 1-form ω is simply given by the action of the basis:
$$\omega(v) = \sum_\mu \omega_\mu e^{*\mu}(v) = \sum_\mu \omega_\mu e^{*\mu}(\sum_\nu v^\nu e_\nu) = \sum_{\mu,\nu} \omega_\mu v^\nu \underbrace{e^{*\mu}(e_\nu)}_{=\delta^\mu_\nu} = \sum_\mu \omega_\mu v^\mu = \omega_\mu v^\mu \ , \tag{3.18}$$
  where $\omega_\mu v^\mu$ is a function at p, i.e. a number.
- For a given basis there is a correspondence between vector and covectors: one can think of vectors as linear maps on duals
$$v(\omega) \equiv \omega(v) = \omega_\mu v^\nu \ . \tag{3.19}$$
  So in this sense $T_p^{**}\mathcal{M} = T_p^*\mathcal{M}$.
- The *natural basis* of $T_p^*\mathcal{M}$ is the gradient of the local coordinates
$$\mathrm{d}x^\mu(\partial_\nu) = \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu \quad \Rightarrow \quad \omega = \omega_\mu \mathrm{d}x^\mu. \tag{3.20}$$
  Under a coordinate transformation the basis and the components transform as
$$\mathrm{d}x^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu}\mathrm{d}x^\mu \quad \Rightarrow \quad \omega_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}}\omega_\mu \ . \tag{3.21}$$
  One thus find again 1-form are the generalization of 4-covectors.

Example 3.5.1. Gradient of a scalar field. *Consider a worldline in Mikowski spacetime, i.e. a curve $x^\mu(\lambda)$ in $\mathbb{R}^4$, and a scalar field $\varphi(x^\mu)$. The tangent vector to the curve has components $v^\mu = dx^\mu/d\lambda = \dot{x}^\mu$; the values of the scalar field along the curve is $\varphi(x^\mu(\lambda))$. The "rate of change" of the scalar field along the curve is*
$$\frac{d\varphi}{d\lambda} = \frac{\partial\varphi}{\partial x^\mu}\frac{dx^\mu}{d\lambda} = v^\mu \frac{\partial\varphi}{\partial x^\mu} \ . \tag{3.22}$$
The equation above can be viewed as a map $v \mapsto d\varphi/d\lambda \in \mathbb{R}$, i.e. it is a 1-form! The components of the 1-form are precisely the components of the gradient:
$$\omega(v) = \omega_\mu v^\mu = v^\mu \frac{\partial\varphi}{\partial x^\mu} \quad \Rightarrow \quad \omega_\mu = \frac{\partial\varphi}{\partial x^\mu} = (\frac{\partial\varphi}{\partial t}, \frac{\partial\varphi}{\partial x^1}, \frac{\partial\varphi}{\partial x^2}, \frac{\partial\varphi}{\partial x^3}) \ . \tag{3.23}$$

Example 3.5.2. *The simplest 1-form is the* gradient $\mathrm{d}f$ *of a function on $\mathcal{M}$. Consider the vector $v = d/d\lambda$ (you can think of it as the vector tangent to a curve, but it can be also a general vector), the 1-form defined as the gradient of the function f acts on this vector and results in the directional derivative of the function:*
$$\omega(v) = \mathrm{d}f(v) = \mathrm{d}f(\frac{d}{d\lambda}) = \frac{df}{d\lambda} \ . \tag{3.24}$$
The gradient is the simplest choice that allows one to obtain the derivative at point p of the function.

Example 3.5.3. Visualization of 1-forms. *A vector v is interpreted as an abstract displacement and it is usually illustrated/imagined as an "arrow". How can one imagine a 1-form?*

Consider a topographical map in which countours of constant elevation are equisurfaces of a scalar field, e.g. each contour represent the same altitude h, the closer the lines are, the largest the gradient $\mathrm{d}h$ is. In order to know the elevation from point a point A to a point B one draws the vector $\vec{AB}$ and count the lines that the vector crosses. In other terms, the values of the gradient is the number of surfaces crossed by the vector, Fig. (3.3).

A 1-form can be thus thought as a series of surfaces and abstractly visualized as straight lines. Surfaces are parallel because the 1-form is defined at one point in the manifold. If one draws a vector, the value of $\omega(v)$ is the number of lines the vectors crosses.

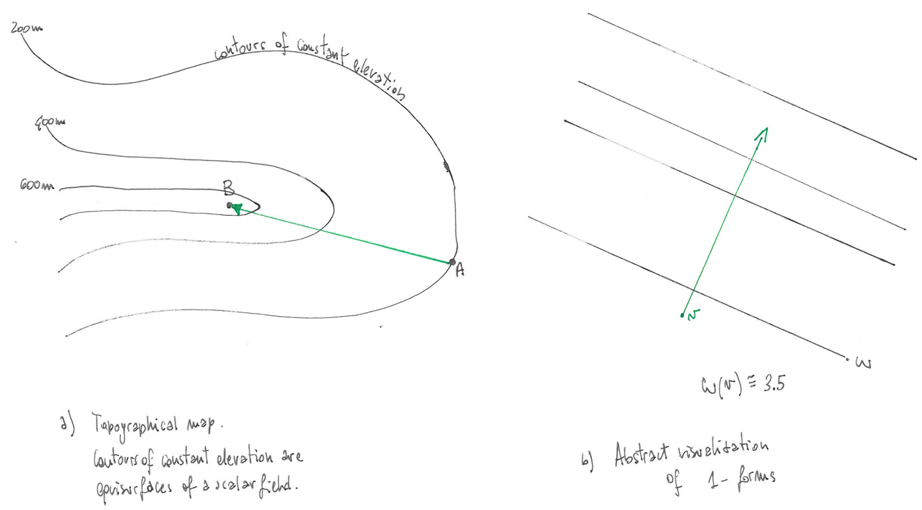

Figure 3.3: Abstract visualization of 1-form as series of surfaces Schutz (1985).

3.6 Tensors

Tangent vectors and covectors generalize to tensors.

Definition 3.6.1. *A* tensor *of type (rank)* (k,l) *is a* <u>multilinear</u> *map taking* k *1-forms and* l *vectors and giving a number,* $T : \underbrace{T_p^*\mathcal{M} \times ... \times T_p^*\mathcal{M}}_{k} \times \underbrace{T_p\mathcal{M} \times ... \times T_p\mathcal{M}}_{l} \mapsto \mathbb{R}$.

Multilinear means linear in each of the arguments.

Examples (math).
- $(0,0)$ tensor = scalar;
- $(0,1)$ tensor = dual vector;
- $(1,0)$ tensor = vector.

Examples (physics).
- The Farady/Maxwell trensor is a $(0,2)$ tensor field;
- The stress-energy tensor for particles, matter fields, etc is a $(0,2)$ tensor.

Properties.
- Tensors on a n-dimensional manifold form a vector space $\tau(k,l)$ of dimension n^{k+l}.
- *Tensor product.* Given two tensor $T_1 \in \tau(k,l)$ and $T_2 \in \tau(k',l')$, the tensor product is defined as

$$\otimes : \tau(k,l) \times \tau(k',l') \mapsto \tau(k+k,l+l') \tag{3.25a}$$
$$T_1 \otimes T_2(\omega_1,...,\omega_{k+k'},v_1,...,v_{l+l'}) := T_1(\omega_1,...,\omega_k,v_1,...,v_l)T_2(\omega_{k+1},...,\omega_{k+k'},v_{l+1},...,v_{l+l'}) \tag{3.25b}$$

  Its action is simply the product of the action of the two tensors on the appropriate arguments. Note that $T_1 \otimes T_2 \neq T_2 \otimes T_1$ and that the subscripts of ω and v are not the components but label the vectors and 1-forms.
- *Tensor basis.* Given $\{e_\mu\}$ basis of $T_p\mathcal{M}$ and $\{e^{*\mu}\}$ basis of $T_p^*\mathcal{M}$ let us construct a basis for the elements of $\tau(k,l)$ by considering the action of the tensor on 1-forms and vectors :

$$T(\omega_1,...,\omega_k,v_1,...,v_l) = T(...,\underbrace{\omega_{\mu_j}e^{*\mu_j}}_{j\text{th argument}},...,\underbrace{v^{\nu_i}e_{\nu_j}}_{i\text{th argument}},...) \tag{3.26a}$$
$$= \sum_{\mu_1}...\sum_{\nu_1}...\,\omega_{\mu_1}...\omega_{\mu_k}v^{\nu_1}...v^{\nu_l}\underbrace{T(...,e^{*\mu_j},...,e_{\nu_j},...)}_{\text{tensor components } =:T^{\mu_1...\mu_k}_{\nu_1...\nu_l}} \tag{3.26b}$$
$$= \omega_{\mu_1}...\omega_{\mu_k}v^{\nu_1}...v^{\nu_l}T^{\mu_1...\mu_k}_{\nu_1...\nu_l}\ . \tag{3.26c}$$

The first line writes the argument in terms of components; note that in component notation the label of the vector/covectors gets attached to the index. The second line uses linearity and defines the tensor components $T^{\mu_1...\mu_k}_{\nu_1...\nu_l}$. The last line uses the sum-convention. The claim here is that a basis for the tensor is given by the tensor product (note the indexes):

$$e_{\mu_1} \otimes ... \otimes e_{\mu_k} \otimes e^{*\nu_1} \otimes ... \otimes e^{*\nu_l} \ . \tag{3.27}$$

Direct verification:

$$\begin{aligned}
T(\omega_1, ..., \omega_k, v_1, ..., v_l) &= T^{\mu_1...\mu_k}_{\nu_1...\nu_l} e_{\mu_1} \otimes ... \otimes e_{\mu_k} \otimes e^{*\nu_1} \otimes ... e^{*\nu_l} (\omega_1, ..., \omega_k, v_1, ..., v_l) && (3.28a)\\
&= T^{\mu_1...\mu_k}_{\nu_1...\nu_l} e_{\mu_1} \otimes ... \otimes e_{\mu_k} \otimes e^{*\nu_1} \otimes ... e^{*\nu_l} (..., \omega_{\alpha_j} e^{*\alpha_j}, ..., v^{\beta_i} e_{\beta_i}, ...) && (3.28b)\\
&= T^{\mu_1...\mu_k}_{\nu_1...\nu_l} \omega_{\alpha_1} ... \omega_{\alpha_k} v^{\beta_1} ... v^{\beta_l} e_{\mu_1} \otimes ... \otimes e_{\mu_k} \otimes e^{*\nu_1} \otimes ... e^{*\nu_l} (..., e^{*\alpha_i}, ..., e_{\beta_j}, ...) && (3.28c)\\
&= T^{\mu_1...\mu_k}_{\nu_1...\nu_l} \omega_{\alpha_1} ... \omega_{\alpha_k} v^{\beta_1} ... v^{\beta_l} e_{\mu_1}(e^{*\mu_1}) ... e_{\mu_k}(e^{*\alpha_k}) e^{*\nu_1}(e_{\beta_1}) ... e^{*\nu_l}(e_{\beta_l}) && (3.28d)\\
&= T^{\mu_1...\mu_k}_{\nu_1...\nu_l} \omega_{\alpha_1} ... \omega_{\alpha_k} v^{\beta_1} ... v^{\beta_l} \delta^{*\mu_1}_{\mu_1} ... \delta^{*\alpha_k}_{\mu_k} \delta^{*\nu_1}_{\beta_1} ... \delta^{*\nu_l}_{\beta_l} && (3.28e)\\
&= \omega_{\mu_1} ... \omega_{\mu_k} v^{\nu_1} ... v^{\nu_l} T^{\mu_1...\mu_k}_{\nu_1...\nu_l} \ . && (3.28f)
\end{aligned}$$

First line write simply the tensor in terms of "components×basis". The second line writes the arguments in terms of the components; third line uses linearity; fourth line uses the definition of tensor product; fifth line uses the definition of basis of the dual space. Note how useful is here the Einstein-sum convention. Note also that the tensor product is terms of components is simply the product of the components

$$(T_1 \otimes T_2)^{\mu_1...\mu_{k+k'}}_{\nu_1...\nu_{l+l'}} = T_1{}^{\mu_1...\mu_k}_{\nu_1...\nu_l} T_2{}^{\mu_{k+1}...\mu_{k+k'}}_{\nu_{l+1}...\nu_{l+l'}} \ . \tag{3.29}$$

- *Tensor contraction.* The operation

$$\begin{aligned}
&C_{(ij)} : \tau(k, l) \mapsto \tau(k-1, l-1) && (3.30a)\\
&C_{(ij)} T := \sum_\sigma T(..., \underbrace{e^{*\sigma}}_{j\text{th argument}}, ..., \underbrace{e_\sigma}_{i\text{th argument}}, ...) \ , && (3.30b)
\end{aligned}$$

  is called a contraction. In terms of components $(C_{(ij)}T)^{\mu_1...\mu_{k-1}}_{\nu_1...\nu_{l-1}} = T^{\mu_1...\sigma...\mu_k}_{\nu_1...\sigma...\nu_l}$, and the notation $C_{(ij)}$ is usually omitted, e.g. $T^\alpha_\beta = T^{\alpha\sigma}_{\beta\sigma}$ is the contraction $C_{(22)}$ (see below Remark 3.6.1).
- *Change of coordinates.* The natural basis made of $e_\mu = \partial_\mu$ and $e^{*\mu} = \mathrm{d}x^\mu$ change under coordinate transformation as

$$\partial_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \partial_\mu \ , \quad \mathrm{d}x^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu} \mathrm{d}x^\mu \ . \tag{3.31}$$

  Substituting in the tensor basis $e_{\mu_1} \otimes ... \otimes e_{\mu_k} \otimes e^{*\nu_1} \otimes ... e^{*\nu_l}$ and using linearity one immediately sees that the components change as

$$T^{\mu'_1...\mu'_k}_{\nu'_1...\nu'_l} = \frac{\partial x^{\mu'_1}}{\partial x^{\mu_1}} ... \frac{\partial x^{\mu'_k}}{\partial x^{\mu_k}} \frac{\partial x^{\nu_1}}{\partial x^{\nu'_1}} ... \frac{\partial x^{\nu_k}}{\partial x^{\nu'_k}} T^{\mu_1...\mu_k}_{\nu_1...\nu_l} \ , \tag{3.32}$$

  which again generalize the SR result, Eq. (2.46), to arbitrary geometries.

Remark 3.6.1. Abstract notation. *Tensors v, ω, T are often indicated by their components (functions) $v^\mu, \omega_\mu, T^{\mu_1...\mu_k}_{\nu_1...\nu_l}$ with greek indexes $\alpha, \beta, ..., \mu, \nu, ...$ employed by convention. The component notation simplify many calculations and but has the drawback that the equations in terms components might not be tensor equations. In fact, a particular basis might simplify the equations for the components (for example because it exploits coordinates adapted to the symmetries of the problem) but such equations might be not valid for every basis.*

Wald (1984) proposes a third notation $v^a, \omega_a, T^{a_1...a_k}_{b_1...b_l}$ with latin indexes $a, b, c, ...$ where the symbols indicate the tensor (not its components). In this notation one writes tensor equations valid in every basis. The notation has the advantage of indicating what type of tensor is, "where" the component indexes must go, and keeping the formulas compact (e.g. avoid to introduce other symbols for contrations). Examples: T^{abc}_{de} indicate an element of $\tau(3, 2)$; T^{abc}_{be} indicate an element of $\tau(2, 1)$ obtained by the contration $C_{(24)}$; $T^{abc}_{de} S^f_g$ is an element of $\tau(4, 3)$ obtainted by the tensor product $(T^{abc}_{de}) \otimes (S^f_g)$.

Symmetries and antisymmetries .

- $S \in \tau(0, 2)$ is *symmetric* iff $S(v_1, v_2) = s(v_2, v_1)$ $\forall v_{1,2} \in T_p\mathcal{M}$ (symmetry is defined according to the exchange of arguments). In abstract notation one has $S_{ab} v^a v^b = S_{ba} v^a v^b$, which implies for components $S_{\mu\nu} = S_{\nu\mu}$. A $(0, 2)$ tensor T can be symmetrized by defining $S(v_1, v_2) = (T(v_1.v_2) + T(v_2, v_1)))/2$; in abstract notation $S_{(ab)} = (T_{ab} + T_{ba})/2$.
- $A \in \tau(0, 2)$ is *antisymmetric* iff $A(v_1, v_2) = -A(v_2, v_1)$ $\forall v_{1,2} \in T_p\mathcal{M}$. In abstract notation one has $A_{ab} = -A_{ba}$. A $(0, 2)$ tensor T can be antisymmetrized by defining $A_{[ab]} = (T_{ab} - T_{ba})/2$.

- Generic totally symmetric/antisymmetric tensors are given by

$$T_{(a_1 \dots a_n)} = \frac{1}{n!}\sum_\pi T_{a_{\pi(1)\dots\pi(n)}} \quad , \quad T_{[a_1 \dots a_n]} = \frac{1}{n!}\sum_\pi \sigma_\pi T_{a_{\pi(1)\dots\pi(n)}} \quad , \tag{3.33}$$

where π are the permutations and $\sigma_\pi = \pm 1$ for even/odd permutation. It is possible to symmetrize also upper indexes in the same way and group of indexes. Examples

$$T_{[abc]} = \frac{1}{3!}(T_{abc} + T_{cab} + T_{bca} - T_{bac} - T_{acb} - T_{cba}) \tag{3.34a}$$

$$T^{(ab)c}_{de} = \frac{1}{2}(T^{abc}_{de} + T^{bac}_{de}) \tag{3.34b}$$

$$T^{(ab)c}_{de} = \frac{1}{2}(T^{abc}_{de} + T^{bac}_{de}) \tag{3.34c}$$

$$T^{(ab)c}_{[de]} = \frac{1}{2}(T^{abc}_{[de]} + T^{bac}_{[de]}) = \frac{1}{4}(T^{abc}_{de} - T^{abc}_{ed} + T^{bac}_{de} - T^{bac}_{ed}) \tag{3.34d}$$

Example 3.6.1. *The physical meaning of a tensor object is clarified by the* stress-energy tensor *of matter. The latter is an object that, given the 4-velocity of the particles worldline, output the energy density. This is exactly the action of a tensor. The energy and momenum of matter fields are thus described by $T \in \tau(0,2)$ symmetric tensor. Specifically, given the timelike vector field u representing the 4-velocity tangent to the worldline of an observer $\mathcal{O}$* [2]*, the* stress-energy tensor *is defined as that bilinear map such that $T(u,u)$ = energy density measured by $\mathcal{O}$ [Complete definition is given below].*

3.7 Metric

Given a point $p \in \mathcal{M}$, we introduce here an object that gives the "infinitesimal squared interval" (or *line element*) associated to an infinitesimal displacement. The idea is clearly to generalize the intervals in $\mathbb{R}^4$

$$d\ell^2 = \delta_{ij} dx^i dx^j \quad \text{Euclidean spacetime} \tag{3.35a}$$

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \quad \text{Mikowski spacetime} \ , \tag{3.35b}$$

to arbitrary manifolds. Since an infinitesimal displacement is represented by a vector $v \in T_p\mathcal{M}$ and ds^2 must be a quadratic form, the object we are looking for must take two vectors and give a number, i.e. it must be a symmetric $(0,2)$ tensor.

Definition 3.7.1. Metric $= g \in \tau(0,2)$, *symmetric $g(v_1, v_2) = g(v_1, v_2)$ $(g_{ab} = g_{ba})$, and nondegenerate $g(v, v_1) = 0$ $\forall v \Rightarrow v_1 = 0$.*

In a coordinate basis the metric tensor writes

$$g = g_{\mu\nu} \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu \ , \tag{3.36}$$

where one should note that $\mathrm{d}x^\mu$ are **not** differentials but 1-forms. The action on two vectors is

$$g(u,v) = g_{\mu\nu}\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu(u,v) = u^\alpha v^\beta g_{\mu\nu}\mathrm{d}x^\mu \otimes \mathrm{d}x^\nu(\partial_\alpha, \partial_\beta) = u^\alpha v^\beta g_{\mu\nu} \underbrace{\mathrm{d}x^\mu(\partial_\alpha)}_{\delta^\mu_\alpha} \underbrace{\mathrm{d}x^\nu(\partial_\beta)}_{\delta^\nu_\beta} = u^\mu v^\nu g_{\mu\nu} \ . \tag{3.37}$$

If one takes the vector $v = dx^\mu \partial_\mu$ connecting point p to an infiniteismally close point $p + \delta p$, the line element is

$$ds^2 = g(v,v) = dx^\mu dx^\nu g_{\mu\nu}\mathrm{d}x^\mu(\partial_\alpha)\mathrm{d}x^\nu(\partial_\beta) = dx^\mu dx^\nu g_{\mu\nu} \ . \tag{3.38}$$

Properties.

1. Since the metric is nondegenrate, the determinant of its components $\det g \neq 0$. Once can thus define the *inverse metric* as the tensor g^{ab} such that $g^{ac}g_{bc} = \delta^a_c$. The inverse metric is clearly symmetric. Differently from SR the components of the inverse metric are in general different from those of the metric.
2. It is always possible to introduce special coordinates called *normal coordinates* such that the metric at a point is diagonal with elements ± 1 (Schutz, 1985; Carroll, 1997; O'Neill, 1983),

$$g(e_\mu, e_\nu) = g_{\mu\nu} = \mathrm{diag}(-1, -1, ..., +1, +1, ...) \ . \tag{3.39}$$

Physically these coordinates must exists in GR because the Einstein's equivalence principle (EEP) states that locally (at point p) one must be able to reduce to SR ("remove gravity").
Mathematically the number of basis vectors giving "-1" and those giving "+1" is independent on the specific choice of the basis. One can thus define unambiguously

[2] A timelike vector field was defined in SR as a vector with negative norm; the concept generalizes in GR where the norm is given by the metric (see below). For the moment one can think of the SR case.

Definition 3.7.2. Signature *of* g = sign(g) *the set of "-1" and "+1" in normal coordinates.*

3. A metric g is called
   - *Euclidean* or Riemannian iff sign(g) has all "+1" (metric is positive definite)
   - *Lorentzian* iff sign(g) has one "-1" (metric is locally Mikowski)
4. The metric g establishes a natural correspondence between $T_p\mathcal{M}$ and $T_p^*\mathcal{M}$. Given a 1-form $\omega \in T_p^*\mathcal{M}$ acting on vectors $v \in T_p\mathcal{M}$, one can associate a vector $\bar{\omega} \in T_p\mathcal{M}$

$$\omega \to g(.,\bar{\omega}) \ , \tag{3.40}$$

such that

$$\omega(v) = g(v,\bar{\omega}) = \underbrace{g_{\mu\nu}\bar{\omega}^\nu}_{\omega_\mu} v^\mu = \omega_\nu v^\mu \ . \tag{3.41}$$

The metric defines the isomorphism between $T_p\mathcal{M}$ and $T_p^*\mathcal{M}$. This property allows on to "raise and lower" indexes with the metric. Note this is what happens in SR where the metric is $g = \eta$.
5. In 4D the number of independent components of the metric tensor is 10: diagonal + lower (or upper) triangular part. In dimension n the number of independent components is $n(n+1)/2$ because there are n components on the diagonal and the lower triangular part has $(n^2-n)/2 = n(n-1)/2$.

Remark 3.7.1. Raising and lowering indexes of tensor components using the metric *(and its inverse) corresponds to the well defined operations of tensor product and contraction. Applying these operations result in building other tensors, different from the originals. For example by lowering the index of a vector one is defining the 1-form associated by the metric tensor. In practical calculation: (a) raising and lowering does not change the position of an index relative to other indices; (b) free indices (not summed over) must bet he same on both sides of an equation; (c) mute indices (summed over) only appear on one side.*

Example 3.7.1. *The metric of the* 2-sphere S^2 *in terms of the usual* (θ,ϕ) *coordinates is* $g_{\mu\nu} = \mathrm{diag}(1,\sin^2\theta)$*. The line element is* $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$*.*

Example 3.7.2. *In* Euclidean $\mathbb{R}^3$ *the metric components in Cartesian and spherical coordinates are given by*

$$g_{\mu\nu} = \mathrm{diag}(1,1,1) \quad g_{\mu\nu} = \mathrm{diag}(1,r^2,r^2\sin^2\theta) \ . \tag{3.42}$$

While the components are different, the line element is the same

$$d\ell^2 = g_{\mu\nu}dx^\mu dx^\nu = dx^2 + dy^2 + dz^2 = dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2 \ . \tag{3.43}$$

Example 3.7.3. *In* Mikowski spacetime $\mathbb{R}^4$ *the metric components in Cartesian coordinates are* $g_{\mu\nu} = \eta_{\mu\nu} = \mathrm{diag}(-1,1,1,1)$.

Example 3.7.4. Surfaces *are naturally represented by 1-forms. Consider Mikowski spacetime with metric* η *and a surface* $\varphi = const$ *determined by the scalar field* $\varphi(x^\mu)$*. The gradient 1-form of* φ *has components*

$$(\mathrm{d}\varphi)_\mu = (\frac{\partial\varphi}{\partial t},\frac{\partial\varphi}{\partial x},\frac{\partial\varphi}{\partial y},\frac{\partial\varphi}{\partial z}) \ , \tag{3.44}$$

and uniquely identifies the surface $\varphi = const$*. The normal vector to the surface is determined by demanding that it is the vector orthogonal to all the vectors laying on the surface* $\varphi = const$*. It is the vector associated to the 1-form throughout the metric tensor,* $n^\mu = \eta^{\mu\nu}(\mathrm{d}\varphi)_\nu$*, the components are thus*

$$n^\mu = (-\frac{\partial\varphi}{\partial t},\frac{\partial\varphi}{\partial x},\frac{\partial\varphi}{\partial y},\frac{\partial\varphi}{\partial z}) \ . \tag{3.45}$$

Note that, while the normal vector requires the metric η *and the concept of orthogonality, using* $\mathrm{d}\varphi$ *allows for a metric-independent characterization of the surface.*

Example 3.7.5. Lorentzian geometry of an expanding universe. *Consider a 2D universe with metric*

$$g = -\mathrm{d}t\otimes\mathrm{d}t + a^2(t)\mathrm{d}x\otimes\mathrm{d}x = -\mathrm{d}t^2 + a^2(t)\mathrm{d}x^2 \ , \tag{3.46}$$

with coordinates $x^\mu = (t,x)$ *(*$t \in (0,\infty)$ *and* $x \in (-\infty,+\infty)$*) and scale factor* $a(t) := t^q$ *with* $0 < q < 1$*. Note the second expression of* g *above is often used: it has the same meaning as the first but it is slightly less rigourous and should not be confused with the differential. The metric Eq. (3.46) is a special case of the Roberts-Walker metric in lower dimensions and describes a universe that at fixed t-coordinate is a flat 1D Eucliden space whose volume expands in time according to the scale factor* $a(t)$*.*

We want to compute the causal structure of this universe, i.e. to compute the "light cones" for this spacetime. Specifically, one looks for curves $x^\mu(\lambda)$ *with null tangent vectors* $v = \dot{x}^\mu\partial_\mu$ *defined by analogy to SR:*

$$v \ \text{ is null vector iff } \ g(v,v) = 0 \ . \tag{3.47}$$

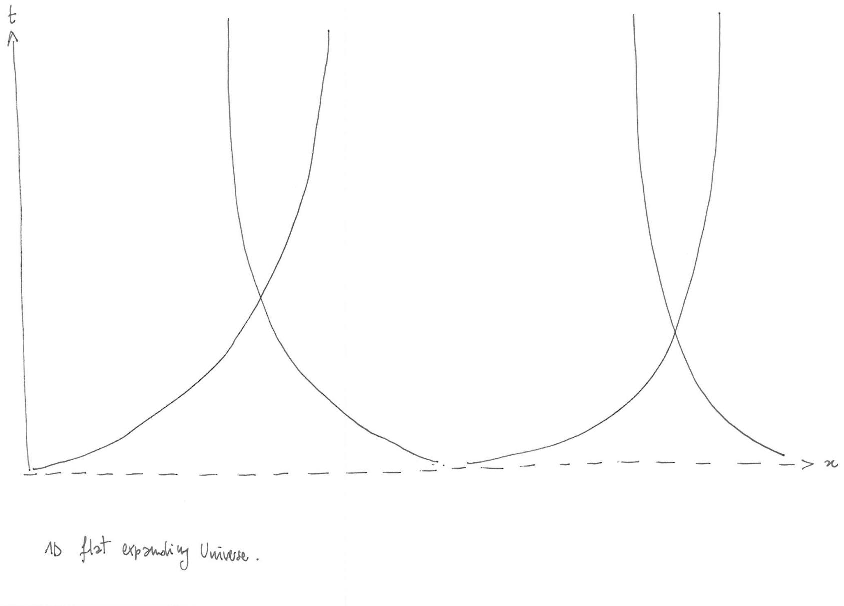

Figure 3.4: 2D expanding universe.

The computation is easy if one remembers that e.g. $\mathrm{d}t(\partial_\mu)$ *is the gradient of* t,

$$0 = g(v,v) = -\mathrm{d}t \otimes \mathrm{d}t(v,v) + a^2 \mathrm{d}x \otimes \mathrm{d}x(v,v) = -\mathrm{d}t(v)\mathrm{d}t(v) + a^2 \mathrm{d}x(v)\mathrm{d}x(v) \tag{3.48a}$$

$$= -\mathrm{d}t(\dot{x}^\mu \partial_\mu)\mathrm{d}t(\dot{x}^\nu \partial_\nu) + a^2 \mathrm{d}x(\dot{x}^\alpha \partial_\alpha)\mathrm{d}x(\dot{x}^\beta \partial_\beta) \tag{3.48b}$$

$$= -\left(\dot{x}^\mu \mathrm{d}t(\partial_\mu)\right)\left(\dot{x}^\nu \mathrm{d}t(\partial_\nu)\right) + a^2 \left(\dot{x}^\alpha \mathrm{d}x(\partial_\alpha)\right)\left(\dot{x}^\beta \mathrm{d}x(\partial_\beta)\right) \tag{3.48c}$$

$$= -\left(\dot{x}^\mu \frac{\partial t}{\partial x^\mu}\right)^2 + a^2 \left(\dot{x}^\alpha \frac{\partial x}{\partial x^\alpha}\right)^2 = -\left(\frac{dt}{d\lambda}\right)^2 + a^2\left(\frac{dx}{d\lambda}\right)^2 \tag{3.48d}$$

where in the last expression now differential do appear! Thus the light curves are

$$\left(\frac{dt}{d\lambda}\right)^2 = a^2\left(\frac{dx}{dt}\frac{dt}{d\lambda}\right)^2 \quad \Rightarrow \quad \frac{dx}{dt} = (a^2)^{-1/2} = \pm t^{-q} \quad \Rightarrow \quad t = \left((1-q)(\pm x - x_0)\right)^{1/(1-q)} \ . \tag{3.49}$$

Note that one would have obtained the same result by incorrectly interpreting the equation for the metric with the differentials instead of 1-forms, and simply dividing by dx^2,

$$0 = -dt^2 + a^2 dx^2 \quad \Rightarrow \quad 0 = -\frac{dt^2}{dx^2} + a^2 \ . \tag{3.50}$$

A plot of the light curves is shown in Fig. (3.4). The light "cones" are asymptotically tangent to the x*-axis at* $t = 0$*, which is not a singular point for the metric/scale factor since* $t = 0$ *is excluded. Thus, the light cones of two events do not necessairly intersect (in SR they always do!); there exist worldlines that are causally disconnected. This example suggests how curvature (see next chapter) can generate "horizons", i.e. surfaces that divide set of events causally disconnected from each other.*

3.8 Stress-energy tensor

Let us consider a timelike vector field $u \in T\mathcal{M}$ tangent to the worldline of an observer $\mathcal{O}$, and a set of three vectors fields e_i (we work in 4D in this section) that are orthogonal to the worldline, $g(u, e_i) = 0$, and form a basis of the 3D space perpendicular to the worldline ($g(e_i, e_j) = \delta_{ij}$). One can think of Mikowski spacetime, although the spacetime can be general e.g. the one determined by Einstein equations.

Definition 3.8.1. *The* stress-energy tensor *is a* <u>symmetric</u> $(0,2)$ *tensor field physically defined by the following conditions*

- $T(u,u) =: E$ *is the energy density measured by* $\mathcal{O}$;
- $-\frac{1}{c}T(e_i,u) =: P^i$ *is the impulse density measured by* $\mathcal{O}$ *(3 numbers);*
- $-cT(u,e_i) =: \varphi^i$ *is the energy flux measured by* $\mathcal{O}$ *(3 numbers);*
- $T(e_i,e_j) =: S_{ij}$ *is the force exerted by the matter in direction* e_i *on the unit surface identified by* e_j.

Observations

- $P = P^i e_i$ and $\varphi = \varphi^i e_i$ are, respectively, the impulse vector of the matter and the energy per unit of time measured trhough a surface element perpendicular to u.
- By symmetry [3] one has
$$\varphi^i = c^2 P^i \quad \text{i.e. energy flux} = c^2 \times \text{impulse density} , \tag{3.51}$$
which generalizes the "$E = mc^2$" of SR.

Definition 3.8.2. *Matter obeys the*
Weak energy condition (WEC) *iff* $E = T(u,u) > 0$ $\forall u$ *timelike.*
Dominant energy condition (DEC) *iff* $E^2 \geq c^2 P^2$, $\forall P$ *not spacelike* $P^i P_i \leq 0$.

Note that DEC⇒WEC. Standard matter forms (particles, fluids, electromagnetic fields) obey the DEC.

Example 3.8.1. *The* perfect fluid *model is a fluid (continuum medium characterized by a 4-velocity field) characterized by isotropic pressure in the fluid's rest-frame. The stress-energy tensor in a generic spacetime is*
$$T = (\rho c^2 + p) u \otimes u + p g , \tag{3.52}$$
where g is the spacetime metric, u is the 1-form associated to the fluid's 4-velocity $u_\mu = g_{\mu\nu} u^\nu$ (or $u = g(., \bar{u})$), ρ, p are the energy density and pressure (scalars) in the fluid's rest frame (or comoving rest frame, $u = 0$). In abstract notation,
$$T_{ab} = (\rho c^2 + p) u_a u_b + p g_{ab} . \tag{3.53}$$
The expression above can be justified from the general definition. Consider an observer $\mathcal{O}$ with velocity V with respect to the fluid. The energy measured by $\mathcal{O}$ is
$$E = T(V,V) = (\rho c^2 + p) u \otimes u(V,V) + p \underbrace{g(V,V)}_{g_{ab}V^aV^b = V_aV^a = -1} \tag{3.54a}$$
$$= (\rho c^2 + p) u(V) u(V) - p = (\rho c^2 + p) \underbrace{g(V,u)}_{=u_aV^a=-W} \underbrace{g(V,u)}_{=u_bV^b=-W} - p \tag{3.54b}$$
$$= (\rho c^2 + p) W^2 - p , \tag{3.54c}$$
where W is the Lorentz factor between $\mathcal{O}$ and the fluid's frame [4]. If $\bar{u} = V$, then $W = 1$ and ρc^2 is the density energy of the fluid. Compare the expression for E to SR's "$E = Wmc^2$": the square W^2 appears because here we denote as "E" the energy density (E = energy/volume) and an additional W comes from the length contraction along the direction of movement (volume $= W \cdot$ proper volume) that reduces the volume and increases the energy density.

Similarly, the momentum density is
$$P^i = -T(V, e_i) = (\rho c^2 + p) \underbrace{g(V,u)}_{=u_aV^a=-W} g(e_i, u) + p g(V, e_i) \tag{3.55a}$$
$$= W(\rho c^2 + p) \underbrace{g(e_i, u)}_{=u_a e_i^a =: WV^i/c} + p g(V, e_i) \tag{3.55b}$$
$$= W^2 (\rho + \frac{p}{c^2}) V^i + p \underbrace{g(V, e_i)}_{=0} \tag{3.55c}$$
$$= (\rho + \frac{p}{c^2}) W^2 V^i , \tag{3.55d}$$
where $V^i = cW^{-1} g_{ab} u^a e_i^b$ is the relative velocity between $\mathcal{O}$ and the fluid's frame in direction "i". Finally,
$$S_{ij} = T(e_i, e_j) = (\rho c^2 + p) g(e_i, u) g(e_j, u) + p g(e_i, e_j) = (\rho c^2 + p)(u_a e_i^a)(u_b e_j^b) + p\delta_{ij} \tag{3.56a}$$
$$= W^2 (\rho + \frac{p}{c^2}) V^i V^j + p \delta_{ij} \tag{3.56b}$$
If $V = \bar{u}$, then $W = 1$ and $V^i = 0$ and $S_{ij} = p\delta_{ij}$ is the isotropic pressure in the fluid's frame. The WEC ($E \geq 0$) and DEC ($P^i P_i \leq E$) imply that $\rho \geq 0$ and $\rho c^2 + p \geq 0$, and $\rho c^2 \geq \sqrt{P^i P_i}$ respectively.

3.9 Differential forms or p-forms

Definition 3.9.1. p-form = *totally antisymmetric* $(0,p)$ *tensor.*

The vector space of p-forms is indicated as $\Lambda_p \mathcal{M}$; $\omega \in \Lambda_p \mathcal{M}$ is indicated as $\omega_{[a_1 \ldots a_p]}$. An example of 2-form is the Faraday/Maxwell tensor.

[3] For the moment symmetry is assumed from physical considerations, mathematical justification will be given in the context of GR and Einstein equations.

[4] If not obvious, one can think of SR and take $V = (1, 0^i)$ (rest frame) and $u^a = (W, Wu^i)$ (generic expression): $u_a V^a = \eta_{ab} u^a V^b = -W$. Note the change of notation with respect to SR's γ.

Properties

- Dimension of $\mathcal{M}$ is $n \Rightarrow$ dimension of $\Lambda_p\mathcal{M}$ is $n!/p!(n-p)!$.
- There are no p-forms for $p > n$. For example, for $n = 2$ one has 0-forms (scalars with 1 independent component), 1-forms (dual vectors with 2 components), and 2-forms (2×2 antisymmetric matrices with 1 independent components):

$n=2$	p	0	1	2
	$\dim\Lambda_p\mathcal{M}$	1	2	1

For $n = 3$ one has scalars, dual vectors with 3 components, antisymmetric matrices with 3 independent components and 3-forms with 1 component:

$n=3$	p	0	1	2	3
	$\dim\Lambda_p\mathcal{M}$	1	3	3	1

For $n = 4$ one has scalars, dual vectors with 4 components, antisymmetric matrices with 6 independent components (e.g. Faraday tensor), 3-forms with 4 components and 4-forms with one component:

$n=4$	p	0	1	2	3	4
	$\dim\Lambda_p\mathcal{M}$	1	4	6	4	1

etc. etc.

- Given basis $\{e_\mu\}$ and $\{e^{*\mu}\}$ for $T_p\mathcal{M}$ and $T^*_p\mathcal{M}$, a basis for a $(0,p)$ tensor is given by

$$e^{*\mu_1} \otimes ... \otimes e^{*\mu_p} \ . \tag{3.57}$$

If the tensor is also antisymmetric several components in this basis would be redundant. In the following we define a "better basis" for p-forms.

Definition 3.9.2. *The* wedge product *of two 1-forms is*

$$\wedge : \Lambda_1\mathcal{M} \times \Lambda_1\mathcal{M} \mapsto \Lambda_2\mathcal{M} \tag{3.58a}$$
$$(\omega, \eta) \in \Lambda_1\mathcal{M} \mapsto \omega \wedge \eta := \omega \otimes \eta - \eta \otimes \omega \tag{3.58b}$$

Check that it is a 2-form, $\forall\ u, v \in T_p\mathcal{M}$:

$$\omega \wedge \eta(u,v) = \omega \otimes \eta(u,v) - \eta \otimes \omega(u,v) = \omega(u)\eta(v) - \eta(u)\omega(v) \tag{3.59a}$$
$$= \eta(v)\omega(u) - \omega(v)\eta(u) \quad \text{(product does commute!)} \tag{3.59b}$$
$$= -\left(\omega(v)\eta(u) - \eta(v)\omega(u)\right) \tag{3.59c}$$
$$= -\left(\omega \otimes \eta(v,u) - \eta \otimes \omega(v,u)\right) = -\omega \wedge \eta(v,u) \ . \tag{3.59d}$$

The wedge product is <u>associative</u> $(a \wedge b) \wedge c = a \wedge b \wedge c$ and can be extended to arbitrary p-forms:

$$\wedge : \Lambda_p\mathcal{M} \times \Lambda_q\mathcal{M} \mapsto \Lambda_{p+q}\mathcal{M} \tag{3.60a}$$
$$(\omega, \eta) \mapsto \omega \wedge \eta := \text{antisymmetric combination of} \omega \otimes \eta \tag{3.60b}$$
$$(\omega \wedge \eta)_{a_1 ... a_p a_{p+1} ... a_{p+q}} = \frac{(p+q)!}{p!q!} \omega_{[a_1 ... a_p} \eta_{a_{p+1} ... . a_{p+q}]} \ . \tag{3.60c}$$

Note that $\omega \wedge \eta = (-1)^{pq}(\eta \wedge \omega)$.

Basis of $\Lambda_p\mathcal{M}$. Let us find the basis of $\Lambda_p\mathcal{M}$.

- A basis for $\Lambda_2\mathcal{M}$ is given by the set $\{e^{*\mu} \wedge e^{*\nu}\}$, i.e. $\alpha \in \Lambda_2\mathcal{M}$ is such that $\alpha = \alpha_{\mu\nu} e^{*\mu} \wedge e^{*\nu}$ with components $\alpha_{\mu\nu} = \alpha_{[\mu\nu]}$ given by

$$\alpha(e_\mu, e_\nu) = \alpha_{\rho\sigma} e^{*\rho} \otimes e^{*\sigma}(e_\mu, e_\nu) = \alpha_{\rho\sigma}\, e^{*\rho}(e_\mu) e^{*\sigma}(e_\nu) = \alpha_{\rho\sigma}\, \delta^\rho_\mu \delta^\sigma_\nu = \alpha_{\mu\nu} \ . \tag{3.61}$$

Verify first that the components of the 2-form are antisymmetric: it is immediate from the expression above to obtain

$$\alpha(e_\mu, e_\nu) = -\alpha(e_\nu, e_\mu) \quad \Rightarrow \quad \alpha_{\mu\nu} = -\alpha_{\nu\mu} \ . \tag{3.62}$$

Now verfy that $\{e^{*\mu} \wedge e^{*\nu}\}$ is a basis:

$$e^{*\rho} \wedge e^{*\sigma}(e_\mu, e_\nu) = e^{*\rho}(e_\mu) e^{*\sigma}(e_\nu) - e^{*\sigma}(e_\mu) e^{*\rho}(e_\nu) = \delta^\rho_\mu \delta^\sigma_\nu - \delta^\sigma_\mu \delta^\rho_\nu \tag{3.63a}$$
$$e^{*\rho} \wedge e^{*\sigma}(e_\nu, e_\mu) = \delta^\rho_\nu \delta^\sigma_\mu - \delta^\sigma_\nu \delta^\rho_\mu \ , \tag{3.63b}$$

hence, taking linear combinations of $\{e^{*\mu} \wedge e^{*\nu}\}$ span all the 2-forms:

$$\alpha_{\rho\sigma} e^{*\rho} \wedge e^{*\sigma}(e_\mu, e_\nu) = \alpha_{\nu\mu} - \alpha_{\mu\nu} \tag{3.64a}$$
$$\alpha_{\rho\sigma} e^{*\rho} \wedge e^{*\sigma}(e_\nu, e_\mu) = \alpha_{\mu\nu} - \alpha_{\nu\mu} = -\left(\alpha_{\nu\mu} - \alpha_{\mu\nu}\right) \ . \tag{3.64b}$$

- A basis of $\Lambda_p\mathcal{M}$ is by the set $\{e^{*\mu_1} \wedge ... \wedge e^{*\mu_p}\}$, a p-form is written as

$$\omega = \omega_{\mu_1 ... \mu_p} e^{*\mu_1} \wedge ... \wedge e^{*\mu_p} \ . \tag{3.65}$$

If $e^{*\mu} = \mathrm{d}x^\mu$ is the dual basis associated to the natural basis $e_\mu = \partial_\mu$, then a basis for the p-forms is

$$\mathrm{d}x^{\mu_1} \wedge ... \wedge \mathrm{d}x^{\mu_p} \tag{3.66}$$

n-form on a n-manifold. A special case of p-form is when $p = n$. In this case the dimension of $\Lambda_n\mathcal{M}$ is $= 1$, and an element $\omega \in \Lambda_n\mathcal{M}$ can be written as

$$\omega = a(x^\mu)\mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^n \ , \tag{3.67}$$

where a is a function, the single component of the n-form (note there is no summation in the baove expression). Under a coordinate transformation, the component transforms as

$$x^\mu \mapsto x^{\mu'} \quad \Rightarrow \quad a(x^\mu) \mapsto a(x^{\mu'})|\frac{\partial x^\mu}{\partial x^{\mu'}}| \ , \tag{3.68}$$

where $|...|$ is the <u>determinant</u> of the coordinate transformation.

Let us verify explicitely the above formula for $n = 2$, considering a transformation $(x^1, x^2) \mapsto (y^1, y^2)$ with $\mathrm{d}y^\mu = \partial y^\mu/\partial x^1 \mathrm{d}x^1 + \partial y^\mu/\partial x^2 \mathrm{d}x^2$:

$$\omega = a(x^1,x^2)\mathrm{d}x^1 \wedge \mathrm{d}x^2 = a(x^1,x^2)\left(\mathrm{d}x^1 \otimes \mathrm{d}x^2 - \mathrm{d}x^2 \otimes \mathrm{d}x^2\right) \tag{3.69a}$$

$$= a(y^1,y^2)\mathrm{d}y^1 \wedge \mathrm{d}y^2 = a(y^1,y^2)\left(\mathrm{d}y^1 \otimes \mathrm{d}y^2 - \mathrm{d}y^2 \otimes \mathrm{d}y^2\right) = \tag{3.69b}$$

$$= a(y^1,y^2)\left[\left(\frac{\partial y^1}{\partial x^1}\mathrm{d}x^1 + \frac{\partial y^1}{\partial x^2}\mathrm{d}x^2\right) \otimes \left(\frac{\partial y^2}{\partial x^1}\mathrm{d}x^1 + \frac{\partial y^2}{\partial x^2}\mathrm{d}x^2\right) - \left(\frac{\partial y^2}{\partial x^1}\mathrm{d}x^1 + \frac{\partial y^2}{\partial x^2}\mathrm{d}x^2\right) \otimes \left(\frac{\partial y^1}{\partial x^1}\mathrm{d}x^1 + \frac{\partial y^1}{\partial x^2}\mathrm{d}x^2\right)\right] \tag{3.69c}$$

$$= a(y^1,y^2)\left[\left(\underline{\underline{\frac{\partial y^1}{\partial x^1}\frac{\partial y^2}{\partial x^1}\mathrm{d}x^1 \otimes \mathrm{d}x^1}} + \frac{\partial y^1}{\partial x^2}\frac{\partial y^2}{\partial x^2}\mathrm{d}x^2 \otimes \mathrm{d}x^1 + \frac{\partial y^1}{\partial x^1}\frac{\partial y^2}{\partial x^2}\mathrm{d}x^1 \otimes \mathrm{d}x^2 + \underline{\frac{\partial y^2}{\partial x^2}\frac{\partial y^2}{\partial x^2}\mathrm{d}x^2 \otimes \mathrm{d}x^2}\right)\right. \tag{3.69d}$$

$$\left. - \left(\underline{\underline{\frac{\partial y^2}{\partial x^1}\frac{\partial y^1}{\partial x^1}\mathrm{d}x^1 \otimes \mathrm{d}x^1}} + \frac{\partial y^2}{\partial x^1}\frac{\partial y^2}{\partial x^2}\mathrm{d}x^1 \otimes \mathrm{d}x^2 + \frac{\partial y^2}{\partial x^2}\frac{\partial y^1}{\partial x^1}\mathrm{d}x^2 \otimes \mathrm{d}x^1 + \underline{\frac{\partial y^2}{\partial x^2}\frac{\partial y^2}{\partial x^2}\mathrm{d}x^2 \otimes \mathrm{d}x^2}\right)\right]$$

$$= a(y^1,y^2)\left[\underbrace{\left(\frac{\partial y^1}{\partial x^1}\frac{\partial y^2}{\partial x^2} - \frac{\partial y^2}{\partial x^1}\frac{\partial y^1}{\partial x^2}\right)}_{=\det\frac{\partial y}{\partial x}}\mathrm{d}x^1 \otimes \mathrm{d}x^2 - \underbrace{\left(\frac{\partial y^1}{\partial x^1}\frac{\partial y^2}{\partial x^2} - \frac{\partial y^2}{\partial x^1}\frac{\partial y^1}{\partial x^2}\right)}_{=\det\frac{\partial y}{\partial x}}\mathrm{d}x^2 \otimes \mathrm{d}x^1\right] \tag{3.69e}$$

$$= a\det\frac{\partial y}{\partial x}(\mathrm{d}x^1 \otimes \mathrm{d}x^2 - \mathrm{d}x^2 \otimes \mathrm{d}x^1) = a\det\left|\frac{\partial y}{\partial x}\right|\mathrm{d}x^1 \wedge \mathrm{d}x^2 \ . \tag{3.69f}$$

Above, underlined terms cancel each other.

3.10 Integration on $\mathcal{M}$

Integrals on a manifold of dimension n are defined using n-forms $\omega \in \Lambda_n\mathcal{M}$.

Euristic argument. An integral of a function

$$\int f(x)\mathrm{d}\mu \ , \tag{3.70}$$

is given by specifying a *measure*

$$\mathrm{d}\mu : \text{infinitesimal region} \ \mapsto \ \text{infinitesimal volume (number)} \ . \tag{3.71}$$

An infinitesimal region is identified by vectors. For example in $n = 3$, three orthogonal vectors make a small cube, and $\mathrm{d}\mu(v_1, v_2, v_3)$ gives the volume of such small cube. In other terms the measure is an object that

- maps vectors to numbers (n vectors if the dimension of $\mathcal{M}$ is n);
- it is linear in the vectors and $\mathrm{d}\mu(av_1, bv_2, cv_3) = abc\mathrm{d}\mu(v_1, v_2, v_3)$ with $a, b, c \in \mathbb{R}$;
- it is antisymmetric because if one changes orientation to one of the vectors the measure has to change sign.

The properties above should clarify that the measure is a n-form.

Integration of a n-form on a n-manifold. Given a $\omega \in \Lambda_n\mathcal{M}$ on a manifold $\mathcal{M}$ of dimension n and introducing a coordinate systems such that

$$\omega = a(x^\mu)\mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^n \ , \tag{3.72}$$

the integral of the n-form on the local set $O \subset \mathcal{M}$ is the defined by the number given by the integral on $\mathbb{R}^n$ of the function a:

$$\int_O \omega := \int_{\psi(O)} a(x^\mu)dx^1...dx^n \ . \tag{3.73}$$

The key observation here is that because of the transformation properties of the n-forms, the definition does **not** depend on the particular coordinate system employed. This can be verified immediately considering a transformation

$x^\mu \mapsto x^{\mu'}$ such that $a \mapsto a' = a|\partial x/\partial x'|$: the integral does not change because the transformation is equivalent to the standard "change of variable" in $\mathbb{R}^n$

$$\int_{\psi'(O)} a(x^{\mu'})dx^{1'}...dx^{n'} = \int_{\psi'(O)} a(x^{\mu'})dx^{1'}...dx^{n'} = \int_{\psi'(O)} a(x^\mu)a\frac{\partial x}{\partial x'}dx^{1'}...dx^{n'} = \int_{\psi(O)} a(x^\mu)dx^1...dx^n \ . \tag{3.74}$$

Caveats and extensions:

- The relation above holds for fixed O. If one changes O than the relation hold only in the intersection $O \cap O'$. One can think of "smoothly contracting" O until it is contained in O'. In general this is possible in *simply connected manifolds*, in which curves can be contracted to a point.
- The definition has a sign ambiguity. The sign can be fixed by introducing an *orientation* in the manifold, i.e. a continuous nonvanishing n-form (Wald, 1984). Simply connected manifolds admit an orientation. An example of nonorientable manifold is the Moebius strip.
- The extension of the integral to the entire manifold is obtained by "summing the integrals" on the subsets, that is possible under some general conditions (Wald, 1984).

3.11 Integration of functions on a n-manifold.

Given a function $f : \mathcal{M} \mapsto \mathbb{R}$ the integral on a n-dimensional manifold is defined as

$$\int f := \int f\varepsilon \ , \tag{3.75}$$

where ε is a n-form providing us with the measure.

Let us find ε. Because the measure is an n-form it must be

$$\varepsilon = a\,\mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^n \ ; \tag{3.76}$$

with a yet unspecified function a. The metric is the object that allows one to caculate lengths, so it should be used to determined a. However, the metric is a tensor and one needs a function such that under coordinate transformation

$$a = a(\text{metric}) \ : \ a \ \mapsto \ a' = a\left|\frac{\partial x'}{\partial x}\right| \ . \tag{3.77}$$

Noting that a similar transformation is given by the determinant of the metric

$$\det(g) \mapsto \det(g)\left|\frac{\partial x'}{\partial x}\right|^2 \ , \tag{3.78}$$

one can take $a := \sqrt{\det(g)} = \sqrt{|g|}$ and obtain the measure to be used in the definition of the integral:

$$\varepsilon = \sqrt{|g|}\mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^n \ . \tag{3.79}$$

The definition of integral of a n-form extends to functions:

$$\int_O f := \int_O f\sqrt{|g|}\mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^n = \int_{\psi(O)} f\sqrt{|g|}dx^1...dx^n \ . \tag{3.80}$$

3.12 Exterior derivative

Forms can also be derived.

Definition 3.12.1. Exterior derivative $\mathbf{d} : \Lambda_p\mathcal{M} \mapsto \Lambda_{p+1}\mathcal{M}$ *such that* $\forall \alpha, \beta \in \Lambda_p\mathcal{M}$

1. *Linear* $\mathbf{d}(\alpha + \beta) = \mathbf{d}\alpha + \mathbf{d}\beta$*;*
2. *Leibnitz antiderivative property:* $\mathbf{d}(\alpha \wedge \beta) = \mathbf{d}\alpha \wedge \beta + (-1)^p\alpha \wedge \mathbf{d}\beta$*;*
3. $\mathbf{d}(\mathbf{d}\alpha) = \mathbf{d}^2\alpha = 0$*.*

Observations.

- For 0-forms (functions) the 1-form given by the gradient is an external derivative, $\mathbf{d} = \mathrm{d}$.
- Property 3. guarantees the exterior derivative is an antisymmetric object. Consider for example the exterior derivative of the 1-form given by the grandient

$$\mathbf{d}(\mathbf{d}f) = \mathbf{d}(\mathrm{d}f) = \mathbf{d}(\frac{\partial f}{\partial x^\mu}\mathrm{d}x^\mu) \approx \frac{\partial^2}{\partial x^\mu \partial x^\nu} \ , \tag{3.81}$$

because 2nd partial derivatives are symmetric, one must require $\mathbf{d}^2\alpha = 0$ to have an antisymmetric object.

- Properties 1.-3. define an object formed by an antisymmetric combination of derivatives. In the natural basis of $\Lambda_{p+1}\mathcal{M}$, the components of the exterior derivative must be then written in terms of the the components of the p-form in argument as
$$(\mathbf{d}\omega)_{\mu_1\dots\mu_p\mu_{p+1}} = (p+1)\partial_{[\mu_1}\omega_{\mu_2\dots\mu_{p+1}]} \ . \tag{3.82}$$
The above expression can be alternatively taken as a definition of $\mathbf{d}\omega$, and one verifies that these components transform as a tensor and properties 1.-3. [exercise]. Note again property 3. simply follows from the antisymmetric combination of second derivatives.

Definition 3.12.2. *A p-form $\omega \in \Lambda_p\mathcal{M}$ is called* closed *iff* $\mathbf{d}\omega = 0$.
A p-form $\eta \in \Lambda_p\mathcal{M}$ is called exact *iff $\eta = \mathbf{d}\omega$, for some $\omega \in \Lambda_{p-1}$.*

Note that exact $\Rightarrow$ closed.

Example 3.12.1. Exterior derivatives *in $\mathbb{R}^3$ ($n=3$). The exterior derivative of a 0-form (function) is the gradient $\mathbf{d}f = \mathrm{d}f$. The exterior derivative of a 1-form $\omega = \omega_i \mathrm{d}x^i$ is the 2-form of dimension 3*
$$\mathbf{d}\omega = \mathbf{d}(\omega_i \mathrm{d}x^i) = \partial_j\omega_i \mathrm{d}x^j \wedge \mathrm{d}x^i \tag{3.83a}$$
$$= (\partial_1\omega_2 - \partial_2\omega_1)\mathrm{d}x^1\wedge\mathrm{d}x^2 + (\partial_2\omega_3 - \partial_3\omega_2)\mathrm{d}x^2\wedge\mathrm{d}x^3 + (\partial_3\omega_1 - \partial_1\omega_3)\mathrm{d}x^3\wedge\mathrm{d}x^1 \ . \tag{3.83b}$$
If one identifies the components of the 1-form with those of a 3-vector $\vec{\omega}$ ($\omega^i = \delta^{ij}\omega_j$), then the 2-form $\mathbf{d}\omega$ can be interpreted as the curl $\nabla\times\vec{\omega}$ (with components $\epsilon^{ijk}\partial_j\omega_k$.)

The exterior derivative of a 2-form $\eta = \eta_{12}\mathrm{d}x^1\wedge\mathrm{d}x^2 + \eta_{23}\mathrm{d}x^2\wedge\mathrm{d}x^3 + \eta_{31}\mathrm{d}x^3\wedge\mathrm{d}x^1$ is a 3-form of dimension 1, i.e. can be written as
$$\mathbf{d}\eta = \textit{(a function)}\,\mathrm{d}x^1\wedge\mathrm{d}x^2\wedge\mathrm{d}x^3 =: a\,\mathrm{d}x^1\wedge\mathrm{d}x^2\wedge\mathrm{d}x^3 \ . \tag{3.84}$$
If one identifies the 3 components of the 2-form with those of a 3-vector $\vec{v} = (\eta_{13}, \eta_{31}, \eta_{12})$, then the component of the 3-form $\mathbf{d}\eta$ is identified with the divergence $a = \nabla\cdot\vec{v}$. The latter is the only scalar that can be constructed from the derivatives of the vector $\vec{v}$.

Note that the property $\mathbf{d}^2 = 0$ implies the usual calculus identities
$$\begin{cases} f\in\Lambda_0 & \mathbf{d}(\mathbf{d}f) = 0 \quad\Rightarrow\quad \nabla\times(\nabla f) = 0 \\ \omega\in\Lambda_1 & \mathbf{d}(\mathbf{d}\omega) = 0 \quad\Rightarrow\quad \nabla\cdot(\nabla\times\vec{\omega}) = 0 \ . \end{cases} \tag{3.85}$$

Exterior derivative of an (n-1)-form in a n-dimensional manifold: divergence of a vector. The Example 3.12.1 suggests the existance of a relation between the exterior derivative of a (n-1)-form in a n-dimensional manifold and the divergence of a vector. This relation becomes manifest introducing the Hodge operator [5] and anticipating the use of the covariant derivative ∇ (If uncomfortable, skip this paragraph and return here later. If comfortable, can think of ∇ as the "appropriate" tensorial derivative generalizing ∂, or simply think of Mikowski and identify $\nabla = \partial$.)

Definition 3.12.3. *The* Hodge *operator $* : \Lambda_1\mathcal{M} \mapsto \Lambda_{n-1}\mathcal{M}$ in a n-dimensional manifold transforms a 1-form η into a (n-1)-form $\omega = *\eta$ such that $\omega_{a_1\dots a_{n-1}} = (*\eta)_{a_1\dots a_{n-1}} := \varepsilon^{b}_{a_1\dots a_n}\eta_b$, where ε is the Levi-Civita tensor antisymmetric tensor.*

Note that the Hodged (n-1)-form can be written in term of the vector η^b
$$(*\eta)_{a_1\dots a_{n-1}} = \varepsilon^{b}_{a_1\dots a_n}\eta_b = \varepsilon_{ba_1\dots a_n}\eta^b \ , \tag{3.88}$$
and that its exterior derivative is the n-form given by the total antisymmetric derivative (all the index but the already contracted one!):
$$(\mathbf{d}(*\eta))_{ca_1\dots a_{n-1}} = n\nabla_{[c|}\varepsilon_{b|a_1\dots a_n]}\eta^b = n\varepsilon_{b[a_1\dots a_n}\nabla_{c]}\eta^b \ . \tag{3.89}$$
We anticipate here the expression in terms of the covariant derivative ∇ and in the last passage we used metric compatibility (Alt. if the metric is flat and $\nabla = \partial$, the expression above holds in Cartesian coordinates). But the above n-form must be $\mathbf{d}(*\eta) = (function)\times\varepsilon$ since it has dimension one. The function/component can be found applying the Hodge again (see footnote and Carroll (1997) Appendix E):
$$(*\mathbf{d}(*\eta)) = \frac{n}{n!}\varepsilon^{ca_1\dots a_n}\varepsilon_{b[a_1\dots a_n}\nabla_{c]}\eta^b = (-1)^s\delta^c_b\nabla_c\eta^b = (-1)^s\nabla_b\eta^b \quad\Rightarrow\quad \mathbf{d}(*\eta) = [(-1)^s\nabla_b\eta^b]\varepsilon \ . \tag{3.90}$$
This show that the exterior derivative of the (n-1)-form $*\eta$ on a n-manifold represent the divergence of the vector associated to η.

[5] Note the Hodge operator can be defined more generally as a map associating p-forms with (n-p)-forms $* : \Lambda_p\mathcal{M}\mapsto\Lambda_{n-p}\mathcal{M}$ where
$$(*\omega)_{a_1\dots a_{n-p}} = \frac{1}{p!}\varepsilon^{b_1\dots b_p}_{a_1\dots a_{n-p}}\omega_{b_1\dots b_p} \ . \tag{3.86}$$
applying twice the Hodge returns the original p-form up to a sign determined by the signature of the metric ($s=-1$ for Lorentz metric)
$$**\omega = (-1)^{s+p(n-p)}\omega \ . \tag{3.87}$$
Note that for a 0-form (function) the double Hodge is the function component of the associated n-form up to a sign.

Example 3.12.2. Maxwell equations *can be written in terms of the exterior derivative, the farady tensor is a closed 2-form. Consider $\mathbb{R}^4$ with the Mikowski metric. The Faraday tensor is a 2-form because $F_{[ab]}$. Some of the Maxwell equations can be written as the antisymmetric combination of derivatives of $F_{[ab]}$, hence*

$$0 = \partial_{[a} F_{bc]} = \mathbf{d}F \ . \tag{3.91}$$

It is possible to prove that in Minkowski spacetime all closed forms are exact. Taking this result without proof, implies that there exists a 1-form A such that $F = \mathbf{d}A$. The Maxwell equations for the vector potential 1-form are simply

$$\mathbf{d}^2 A = 0 \ . \tag{3.92}$$

The other Maxwell equations $\partial_a F^{ba} = 4\pi J^b$ also admit an expression in terms of the exterior derivative [exercise],

$$\mathbf{d}(*F) = 4\pi(*J) \ . \tag{3.93}$$

From the above equation, charge conservation now follows taking another exterior derivative and considering the vector associated to the (n-1)-form $*J$:

$$0 = \mathbf{d}^2(*F) = 4\pi\mathbf{d}(*J) \propto \nabla_a J^a \ . \tag{3.94}$$

3.13 Stokes theorem

Stokes theorem generalizes the fundamental theorem of calculus to generic manifolds.

Consider a n-dimensional oriented manifold $\mathcal{M}$ with boundary $\partial\mathcal{M}$. The definition of $\partial\mathcal{M}$ is given in Wald (1984) but there is an intuitive notion which is sufficient here: (i) a manifold with boundary is a manifold in which the charts map to $\mathbb{R}^n$ with $x^1 \geq 0$ ("half" the Euclidean space); (ii) the boundary $\partial\mathcal{M}$ is the the set of points of $\mathcal{M}$ that don't have open neighborhoods which are isomorphic to open sets in $\mathbb{R}^n$, i.e. they are mapped to $x^1 = 0$. Given a $(n-1)$ form on $\omega \in \Lambda_{n-1}\mathcal{M}$, Stokes theorem says that

$$\int_{\mathcal{M}} \mathbf{d}\omega = \int_{\partial\mathcal{M}} \omega \ . \tag{3.95}$$

Stokes theorem can be cast in a more familar expression by considering the vector v associated to the (n-1)-form $\omega = *v$:

$$\mathbf{d}\omega = \nabla_a v^a \varepsilon = \nabla_a v^a \varepsilon = \nabla_a v^a \mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^n = \nabla_a v^a \sqrt{|g|} d^n x \ , \tag{3.96}$$

Thus the r.h.s. of Eq. (3.95) is the integral of the divergence of v. To manipulate the l.h.s. of Eq. (3.95) one must express the (n-1)-form ω in terms of v in the boundary. The boundary is a $(n-1)$-dimensional manifold equipped with a metric γ that is induced by the ambient metric g. The Levi-Civita tensor (surface element) on $\partial\Sigma$ is given by (Carroll (1997) Appendix E, Wald (1984) Appendix B):

$$\varepsilon_{a_1 \dots a_{n-1}} = n^b \varepsilon_{b a_1 \dots a_{n-1}} \ , \tag{3.97}$$

where n is the normal vector to the boundary. The (n-1)-form on the boundary can be written

$$\omega = v_a n^a \sqrt{|\gamma|} d^{n-1} y \ . \tag{3.98}$$

Putting together everything one gets to a familar form:

$$\int_{\mathcal{M}} \nabla_a v^a \sqrt{|g|} d^n x = \int_{\partial\mathcal{M}} n_b v^b \sqrt{|\gamma|} d^{n-1} y \ . \tag{3.99}$$

4. Curvature & Connection

③

These lectures in differential geometry are about curvature, Riemann tensor and connection. The concepts of geodesic and the equation of geodesic deviation are introduced here. Calculations of gravitational redshift in GR are discussed in the end as application.

Suggested readings. *Chap. 3 of Wald (1984); Chap. 3 of Carroll (1997); Chap. 5-6 of Schutz (1985); O'Neill (1983) book.*

4.1 Intuitive/simple definition of curvature of a curve

Consider curves in $\mathbb{R}^2$, and think on "how much they bend" or "how much curved they are":

- To a straight line, one would assign zero curvature, $\kappa \equiv 0$;
- To a circle of radius R, one would assign a constant curvature of $\kappa = 1/R$;
- To a generic curve, one would consider the osculating circle at a given point p (and two other infinitesimally close points) on the curve, and assign $\kappa(p) = 1/R$.

A BSc physicist might proceed in a slightly more sophisticated way by

- Considering the tangent vector $\vec{v}$ at each point of the curve;
- Calculating the curve acceleration as $\vec{a} = d\vec{v}/dt$, where t is the curve parameter;
- Defining the curvature as the modulus of the acceleration, $\kappa' = |\vec{a}|$, as a measure on "how fast" the vector $\vec{v}$ rotates from point to point along the curve.

Interestingly, the two definitions are basically the same. If $\vec{v}$ is normalized,

$$\kappa = \frac{1}{R} = \frac{2\pi}{\text{circumference}} = \frac{1 \text{ arc angle}}{\text{lenght of curve}} \tag{4.1}$$

$$\kappa' = \left|\frac{d\vec{v}}{dt}\right| = \frac{d\theta}{dt} = \frac{\text{infinitesimal arc lenght of rotation}}{\text{lenght of the infinitesimal curve element}} \ . \tag{4.2}$$

Remark 4.1.1. *In order to define $d\vec{v}$ and take the derivative, one must rigidly "transport back" the vector $\vec{v}(dt)$ at point $p + \delta p$ ($t = dt$) to the point p ($t = 0$) and take the difference to compute the differential.*

4.2 Intrinsic and extrinsic curvature of surfaces & relation to parallel transport

To define and study the curvature of 2D surfaces one could think of embedding them in $\mathbb{R}^3$ and study their normal vectors Fig. (4.1). For example, normal vectors along the plane are tangent to curves that never touch each other, i.e. that remain parallel. Tangent curves that focus would be associated to negative curvature, while curves that depart from each other to positive curvature, etc.

However, as usual, we do not want to rely on introducing an $n+1$ Euclidean space to describe something on a n-dimensional manifold. We need to proceed differently but, before that, the above example already suggests that there exist two different types of curvature we might be interested to consider. It is convenient to discuss them immediately.

Consider a n-dimensional manifold $\mathcal{M}$,

- One problem is to define a curvature of a subset (a "section") $\Sigma \subset \mathcal{M}$ of dimension $n-1$ the manifold. The "child" manifold is called an *embedding* and inherits [in a precise sense, (Wald, 1984; Carroll, 1997)] structures from the "parent" $\mathcal{M}$. Characterizing the curvature of an embedding in relation to the ambient space, i.e. study "how σbends in $\mathcal{M}$", leads to the concept of *extrinsic curvature.*
- Another problem is the definition of *instrinsic curvature* of $\mathcal{M}$, without using any ambient space.

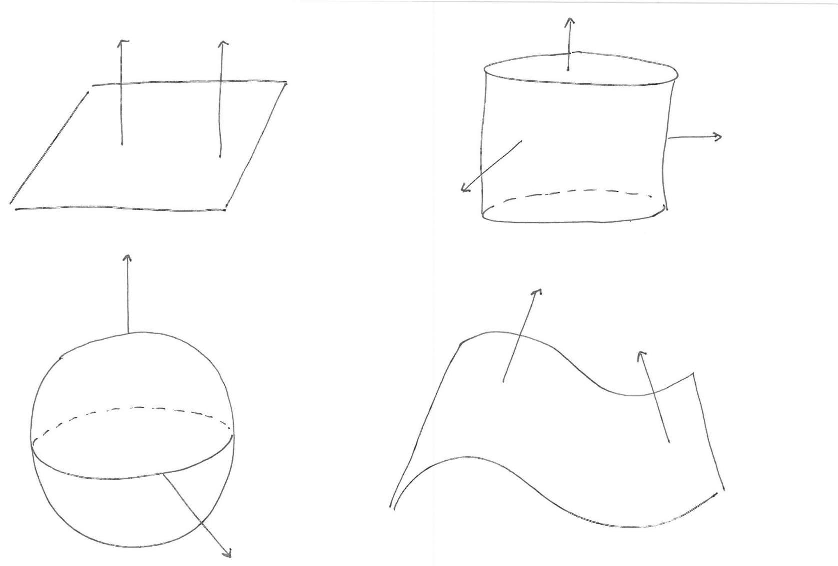

Figure 4.1: Examples of 2D surfaces with their normal vectors.

The two concepts are distinct, as illustrated in the example below, and both useful. We focus here only on the intrinsic curvature, which is the most fundamental one. The intrinsic curvature follows from the latter (Wald, 1984; Carroll, 1997).

Example 4.2.1. Cylinder in $\mathbb{R}^3$. *A cylinder of radius R in $\mathbb{R}^3$ is somehow "round"; in relation to the ambient Euclidean space it is natural to assign an extrinsic curvature $\kappa = 1/R$. However, if we restrict to think of the cylinder as 2D surface, then we could think of unfolding the cylinder by cutting along the height, open it in a plane and identify the two sides of the plane where the cut was. In this plane: parallel lines remain parallel, all Euclid axioms are valid, and the onlyl difference to $\mathbb{R}^2$ is that walking in a direction would bring us back to the starting point. Mathematically, the cylinder has the metric*

$$g = R^2 \mathrm{d}\phi + \mathrm{d}z^2 = \mathrm{d}y^2 + \mathrm{d}z^2 \ , \tag{4.3}$$

where the first expression uses standard cylindrical coordinatres and the second simplly defines $y = R\phi$. The second expression for the metric is exactly the Euclidean metric of $\mathbb{R}^2$. For this reason we assign intrinsic curvature $\kappa = 0$. Note that the periodic condition on the plane is a statement about the topology of the manifold and does not affect the curvature. Moreover, comparing with the 2-sphere, one notice immediately that differently from the cylnder, two parallel lines (e.g. meridians) do meet. Hence, the 2-sphere must have intrinsic curvature not null.

Curvature from "rigidly moving" vectors (parallel transport). Take a plane in 2D and a closed curve on it. If one rigidly transports a vector $\vec{v}$ around the curve in such a way that $\vec{v}$ remains parallel to (point to the same direction of) the previous moved one, then after a lap one obtains the initial vector, Fig. (4.2). Take now a sphere and do the same: (i) the initial and final vector are different; and (ii) different choices of the curve give different result !

We conclude that the intrinsic curvature of a manifold can be detected by rigidly moving, technically *parallel transporting*, vectors and that, in general, paralleling transporting vector depends on the followed path. The main difficulty to overcome is that on a generic manifold the tangent vector spaces at different points are different. It is not possible to compare $v \in T_p\mathcal{M}$ with $u \in T_p\mathcal{M}$ if $p \neq q$. Thus, it is necessary to introduce a method to implement such parallel transport.

4.3 Procedure to define curvature on $\mathcal{M}$

Outline the logic procedure to define the curvature of a manifold.

Observations.

- If we knew how to parallel transport a vector v on a curve γ on $\mathcal{M}$, then we could compare vectors of the same $T_p\mathcal{M}$ and calculate their difference or variation, i.e. a derivative $\nabla v|_p$.
- If we knew how to compute $\nabla v|_p$, then we could define the parallel transport of v along γ as the operation that does not change the vector: $\nabla v|_\gamma = 0$.

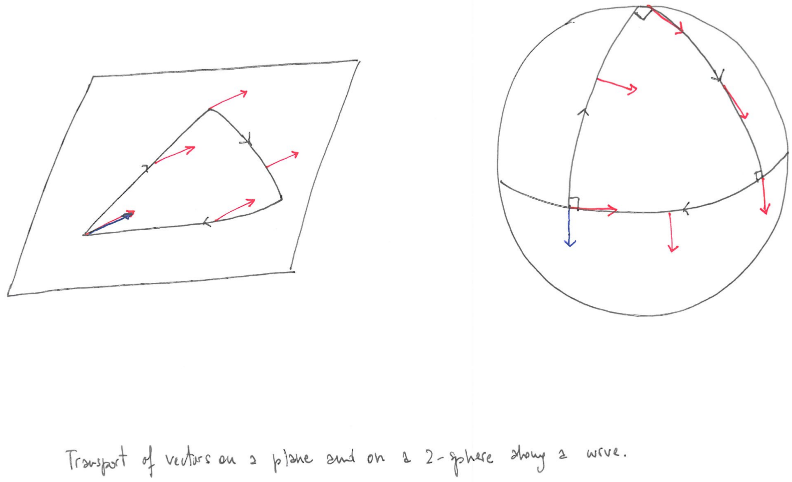

Figure 4.2: Transporting a vector along a curve in a 2D surface.

- A derivative operator ∇ and the concept of parallel transport would allow us to identify or *connect* the tangent spaces $T_p\mathcal{M}$ and $T_q\mathcal{M}$ at two different points.
- Curvature could be then defined from the parallel transport of a vector along an infinitesimal closed curve. In particular, a nonzero intrinsic curvature would correspond to the case where two successive differentiation of v do not commute:

$$[\nabla,\nabla]v = (\nabla\nabla - \nabla,\nabla)v = 0 \quad \Rightarrow \quad \kappa \neq 0\ . \tag{4.4}$$

Procedure. The following key steps will be expanded below:

- Define ∇: the *covariant derivative* of tensors, Sec. 4.4.
- Prove that a metric determines a unique ∇ (*Levi-Civita connection*). The latter defines a map between vectors of $T_p\mathcal{M}$ and those at points infinitesimally close to p. For this reason the covariant derivative is also called a *connection*.
- Define *parallel transport* of a vector v along a curve γ with tangent vector t as the operation such that $t^a\nabla_a v^b = 0$, Sec. 4.5.
- The commutator $[\nabla,\nabla]$ defines the *Riemann tensor* that encodes and quantifies the idea of curvature as "failure of vectors and tensors fields to return to their original values after being transported along an infinitesimal closed loop", Sec. 4.6.
- *Geodesics* are defined as curves whose tangent vector is parallel propagated along itself, $t^a\nabla_a t^b = 0$, Sec. 4.7. They correspond to curves that extremize the lenght between two points of the manifold.
- The *geodesics deviation equation* establishes that the acceleration between two nearby geodesics is zero if the Riemann tensor (curvature) is zero, Sec. 4.8. This corresponds to the intuitive notion that (i) in absence of curvature, "straight lines remain parallel" and in (ii) in presence of curvature "lines focus".

4.4 Connection or Covariant derivative

Definition 4.4.1. *The map* $\nabla : \tau(k,l) \mapsto \tau(k,l+1)$ *is a* covariant derivative *iff* $\forall \alpha,\beta \in \mathbb{R}$ *and* $\forall A, B \in \tau(k,l)$ *the following properties hold:*

1. *Linear:* $\nabla(\alpha A + \beta B) = \alpha\nabla A + \beta\nabla B$*;*
2. *Leibnitz (w.r.t. tensor product):* $\nabla(A\otimes B) = \nabla A \otimes B + A\otimes\nabla B$*;*
3. *Commutation with contraction operator* $C_{(ij)}$*:* $\nabla(C_{(ij)}A) = C_{(ij)}(\nabla A)$*;*
4. *Consistency with derivative of functions:* $\nabla f := \mathrm{d}f\ \forall f \in \mathcal{F}$*;*
5. *Consistency with concept of tangent vector:* $v(f) := v^a\nabla_a f = v^a\partial_a f$*;*
6. *Torsion free:* $[\nabla,\nabla]f = 0\ \forall f\in\mathcal{F}$*.*

Observations. Let us consider the action of the covariant derivative in abstract notation and some properties following the definition:

- The covariant derivative maps $T^{a_1\ldots a_k}_{b_1\ldots b_l} \mapsto \nabla_c T^{a_1\ldots a_k}_{b_1\ldots b_l}$.

- Property 2. is: $\nabla_c(A^{a_1\dots a_k}_{b_1\dots b_l}B^{a_1\dots a_{k'}}_{b_1\dots b_{l'}}) = \nabla_c A^{a_1\dots a_k}_{b_1\dots b_l}\, B^{a_1\dots a_{k'}}_{b_1\dots b_{l'}} + A^{a_1\dots a_k}_{b_1\dots b_l}\,\nabla_c B^{a_1\dots a_{k'}}_{b_1\dots b_{l'}}$.
- Property 3. is: $\nabla_d(A^{a_1\dots c\dots a_k}_{b_1\dots c\dots b_l}) = \nabla_d A^{a_1\dots c\dots a_k}_{b_1\dots c\dots b_l}$.
- Property 6. is : $(\nabla_a\nabla_b - \nabla_b\nabla_a)f = 0 \Rightarrow \nabla_a\nabla_b f = \nabla_{(a}\nabla_{b)}f$ is symmetric.
- Property 5.-6. $\Rightarrow$ the commutator of two vectors can be expressed as

$$
\begin{aligned}
[v,u](f) &= v(u(f)) - u(v(f)) = v(u^a\nabla_a f) - u(v^b\nabla_b f) = v^d\nabla_d(u^a\nabla_a f) - u^c\nabla_c(v^b\nabla_b f) && (4.5a)\\
&= \underbrace{v^d\nabla_d}_{d\to c} u^a\nabla_a f + v^d u^a\nabla_d(\nabla_a f) - u^c\nabla_c \underbrace{v^b\nabla_b}_{b\to a} f - u^c v^b\nabla_c(\nabla_b f) && (4.5b)\\
&= \left(v^b\nabla_b u^a - u^b\nabla_b v^a\right)\nabla_a f - \underbrace{[\nabla,\nabla]f}_{=0} = \left(v^b\nabla_b u^a - u^b\nabla_b v^a\right)\nabla_a f \quad \forall f\ . && (4.5c)
\end{aligned}
$$

- Note that ∇_c is not a dual vector !

Main steps to define a unique connection.

(i) Observe that in a coordinate system the partial derivative ∂ satisfies property 1. – 6. but the resulting object ∂T is not a tensor.
(ii) The difference between covariant derivatives is a tensor.
(iii) Given a metric, there is a unique covariant derivative compatible with the metric, i.e. such that $\nabla g = 0$.
(iv) The unique, metric compatible, covariant derivative can be constructed from the partial derivatives and the *Christoffel symbols* (or connection coefficients).

(i) Partial derivatives. Introduce a coordinate system and the natural basis $\{\partial_\mu\}$ of $T_p\mathcal{M}$. The components of $T\in\tau(k,l)$ are $T^{\mu_1\dots\mu_k}_{\nu_1\dots\nu_l}$ and the quantity

$$\partial_\sigma T^{\mu_1\dots\mu_k}_{\nu_1\dots\nu_l} \tag{4.6}$$

fullfills the properties 1.-6. defining a covariant derivative. However it is not a tensor $\tau_{k,l+1}$ because these components do not transform as those of a tensor under coordinate transformation. For example for a vector,

$$\partial_\sigma v^\mu \mapsto \partial_{\sigma'} v^{\mu'} = \frac{\partial x^\sigma}{\partial x^{\sigma'}}\partial_\sigma\left(\frac{\partial x^{\mu'}}{\partial x^\mu}v^\mu\right) = \underbrace{\frac{\partial x^\sigma}{\partial x^{\sigma'}}\frac{\partial x^{\mu'}}{\partial x^\mu}\partial_\sigma v^\mu}_{\text{ok}} + \underbrace{\frac{\partial x^\sigma}{\partial x^{\sigma'}}v^\mu\frac{\partial x^{\mu'}}{\partial x^\sigma\partial x^\mu}}_{\text{not ok}}\ , \tag{4.7}$$

one does not get the correct transformation because of the second term. A possible approach here is to introduce a symbol with 3 indexes $\Gamma^\sigma_{\mu\mu'}$ and search for its coordinate expression such that

$$\nabla_\mu v^\sigma = \partial_\mu v^\sigma + \Gamma^\sigma_{\mu\lambda}v^\lambda \tag{4.8}$$

transforms as a tensor and is unique. We will basically follow this route, but instead of doing a direct calculation in components we will show (ii) and (iii) and get to the result from those theorems (Wald, 1984).

(ii) Difference of covariant derivatives. Take two covariant derivatives ∇ and $\tilde\nabla$. Property 4. implies that both derivatives are the same when acting on functions $\nabla f = \tilde\nabla f = \mathrm{d}f$. Consider the action on 1-forms

$$\tilde\nabla_a(f\omega_b) - \nabla_a(f\omega_b) = \tilde\nabla_a f\omega_b + f\tilde\nabla_a\omega_b - \nabla_a f\omega_b - f\nabla_a\omega_b = (\mathrm{d}f)_a\omega_b + f\tilde\nabla_a\omega_b - (\mathrm{d}f)_a\omega_b - f\nabla_a\omega_b = f(\tilde\nabla_a - \nabla_a)\omega_b\ . \tag{4.9}$$

The difference of covariant derivatives is linear in f and depends only on objects at point p, $\Rightarrow$ it is a tensor $C\in\tau(1,2)$

$$(\tilde\nabla_a - \nabla_a)\omega_b =: C^c_{ab}\omega_c\ . \tag{4.10}$$

Note that the tensor must be symmetric $C^c_{(ab)}$ because of the torsion-free property 6.; setting $\omega_b = \nabla_b f = \nabla_a f$

$$\underbrace{\nabla_a\nabla_b f}_{\text{sym}} = \underbrace{\tilde\nabla_a\nabla_b f}_{\text{sym}} - C^c_{ab}\nabla_c f\ . \tag{4.11}$$

Calculate now the action of the difference on vector in the following way:

$$
\begin{aligned}
0 &= \tilde\nabla_a(v^b\omega_b) - \nabla_a(v^b\omega_b) = (\tilde\nabla_a - \nabla_a)(\omega_b)v^b + \omega_b(\tilde\nabla_a - \nabla_a)v^b = C^c_{ab}\omega_c v^b + \omega_b(\tilde\nabla_a - \nabla_a)v^b && (4.12a)\\
&= C^b_{ad}\omega_b v^d + \omega_b(\tilde\nabla_a - \nabla_a)v^b = \omega_b\left(C^b_{ad}v^d + (\tilde\nabla_a - \nabla_a)v^b\right) \quad \forall\omega_b \quad \Rightarrow && (4.12b)\\
\nabla_a v^b &= \tilde\nabla_a v^b + C^b_{ad}v^d \quad \text{or} \quad (\tilde\nabla_a - \nabla_a)v^b = -C^b_{ad}v^d\ . && (4.12c)
\end{aligned}
$$

The zero is because $v^b\omega_b$ is a function and the two covariant derivatives act the same way to functions. In the second line the mute indexes have been renamed as $c\to b$ and $b\to d$. The difference of the covariant derivatives applied to vectors is also a tensor. For generic tensor one simply repeat the same calculation by contracting with vectors and forms, to obtain the generic formula

$$\tilde{\nabla}_c T^{a_1\ldots a_k}_{b_1\ldots b_l} = \nabla_c T^{a_1\ldots a_k}_{b_1\ldots b_l} - \sum_i C^{a_i}_{cd} T^{a_1\ldots d\ldots a_k}_{b_1\ldots b_l} + \sum_j C^d_{cb_j} T^{a_1\ldots a_k}_{b_1\ldots d\ldots b_l} \ . \tag{4.13}$$

The above shows that:

- The difference of two covariant derivatives is completely characterized by the symmetric tensor $C^a_{(bc)}$.
- Conversely, given $\tilde{\nabla}$ and a tensor filed $C^a_{(bc)}$, the equation above defines a covariant derivative ∇.
- The covariant derivative is in general **not** unique because it requires to specify $C^a_{(bc)}$ ($n^2(n-1)/2$ independent components).
- In a coordinate basis one can take $\tilde{\nabla}_\mu = \partial_\mu$ (that satifies the properties used to derive the equation above) and specify a C to define the covariant derivative using the formula above. However, in this case one cannot expect that the "symbols with 3 indexes" appearing in the coordinate expression of the equation change as tensor components under coordinate transformation. The reason is that if the basis changes, then one must pick another tensor C' in order to obtain the same result. This shows again that the Γ symbols in Eq. (4.8) are not tensor components, but also that their difference is.

(iii) Unique connection.

Theorem 4.4.1. *Given a metric g_{ab} there is a unique covariant derivative that satisfies $\nabla_a g_{bc} = 0$ (compatible with the metric).*

Proof. *For any derivative operator $\tilde{\nabla}$ explicitely construct the connection compatible with g. Start from*

$$0 = \nabla_a g_{bc} = \tilde{\nabla}_a g_{bc} - C^d_{ab} g_{dc} - C^d_{ac} g_{bd} \quad \Rightarrow \quad \tilde{\nabla}_a g_{bc} = C^d_{ab} g_{dc} + C^d_{ac} g_{bd} = C_{cab} + C_{bac} \ , \tag{4.14}$$

where the last passage takes the contraction of the mute indexes (note it goes in the first position). Now write the last equation substituting index names:

$$\tilde{\nabla}_a g_{bc} = C_{cab} + C_{bac} \quad (abc) \tag{4.15a}$$
$$\tilde{\nabla}_b g_{ac} = C_{cba} + C_{abc} \quad (bac) \tag{4.15b}$$
$$\tilde{\nabla}_c g_{ab} = C_{bca} + C_{acb} \quad (cab) \tag{4.15c}$$

and consider the combination

$$(abc) + (bac) - (cab) = \tilde{\nabla}_a g_{bc} + \tilde{\nabla}_b g_{ac} - \tilde{\nabla}_c g_{ab} \tag{4.16a}$$
$$= C_{cab} + \cancel{C_{bac}} + C_{cba} + \cancel{C_{abc}} - \underbrace{\cancel{C_{bca}}}_{C_{bac}} - \underbrace{\cancel{C_{acb}}}_{C_{abc}} = 2C_{cab} \ , \tag{4.16b}$$

where the symmetry in the last indexes is used. The choice

$$C^c_{ab} = g^{cd} C_{dab} = \frac{1}{2} g^{cd} (\tilde{\nabla}_a g_{bd} + \tilde{\nabla}_b g_{ad} - \tilde{\nabla}_d g_{ab}) \ , \tag{4.17}$$

fixes a unique covariant derivative $\tilde{\nabla}$ and guarantees that it is compatible with the metric.

(iv) Levi-Civita connection & Christoffel symbols. The Levi-Civita connection is constructed by introducing a coordinate basis, taking $\tilde{\nabla} = \partial$ and computing the *Christoffel symbols*

$$\Gamma^\sigma_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} \left(\frac{\partial g_{\nu\sigma}}{\partial x^\mu} + \frac{\partial g_{\mu\sigma}}{\partial x^\nu} - \frac{\partial g_{\mu\nu}}{\partial x^\sigma} \right) \ . \tag{4.18}$$

In terms of the components, the covariant derivatives of tensor follows from the above equations

$$\nabla_\mu v^\nu = \partial_\mu v^\nu + \Gamma^\nu_{\mu\sigma} v^\sigma \tag{4.19a}$$
$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu - \Gamma^\lambda_{\mu\nu} \omega_\lambda \tag{4.19b}$$
$$\nabla_\sigma T^{\mu_1\ldots\mu_k}_{\nu_1\ldots\nu_l} = \partial_\mu T^{\mu_1\ldots\mu_k}_{\nu_1\ldots\nu_l} + \sum_i \Gamma^{\mu_i}_{\sigma\lambda} T^{\mu_1\ldots\lambda\ldots\mu_k}_{\nu_1\ldots\nu_l} - \sum_j \Gamma^\lambda_{\sigma\nu_j} T^{\mu_1\ldots\mu_k}_{\nu_1\ldots\lambda\ldots\nu_l} \ . \tag{4.19c}$$

Notes:

- Γ's are defined by assuming a particular derivative operator and coordinate basis; they do not transform s a tensor if one changes basis.
- Given the basis $e_{\mu_1} \otimes \ldots \otimes e_{\mu_k} \otimes e^{*\nu_1} \otimes \ldots e^{*\nu_l}$ of $\tau(k,l)$, the covariant derivative is expressed as

$$\nabla T = \nabla_\sigma T^{\mu_1\ldots\mu_k}_{\nu_1\ldots\nu_l} e_{\mu_1} \otimes \ldots \otimes e_{\mu_k} \otimes e^{*\nu_1} \otimes \ldots e^{*\nu_l} \otimes e^{*\sigma} \ , \tag{4.20}$$

i.e. the new index goes in the last position. An alternative notation for the components that is more consistent is $\nabla_\sigma T^{\mu_1\ldots\mu_k}_{\nu_1\ldots\nu_l;\sigma}$.

4.5 Parallel transport

Consider a connection ∇ and a curve γ on $\mathcal{M}$ with tange vector t^a.

Definition 4.5.1. *A vector field v^a is* parallel transported *along γ iff $t^a\nabla_a v^b = 0$. Similarly, a tensor is parallel transported along the curve iff $t^c\nabla_c T^{a_1\dots a_k}_{b_1\dots b_l} = 0$.*

Observations.
- Parallel transport depends only on v on the curve (do not need the "full field").
- In a coordinate system:
$$0 = t^\mu\nabla_\mu v^\alpha = t^\mu\partial_\mu v^\alpha + t^\mu\Gamma^\alpha_{\mu\sigma}v^\sigma = \frac{dv^\alpha}{dt} + t^\mu\Gamma^\alpha_{\mu\sigma}v^\sigma \ . \tag{4.21}$$
The above equation is an ODE of first order for the vector components: given a initial value $v^\mu(0)$, the solution exists and it is unique. This implies that a vector a one point p $(t = 0)$ defines uniquely the parallel transport along γ. In other terms, given γ and ∇, the tangent space $T_p\mathcal{M}$ can be connected to $T_q\mathcal{M}$ for any $q \in \gamma$.
- The above definition and observation are not restricted to a Levi-Civita connection (hold also if ∇ is not unique, but specified).

Geometrical meaning of Levi-Civita connection. If the manifold is equipped with a metric, the quantity $g(v, u)$ can be considered the angle between vectors u and v. Thus, the intuitive notion of vectors parallel transported along the curves can be implemented by requiring that

$$g(v, u) = g_{ab}v^a u^b = \text{constant along } \gamma \ . \tag{4.22}$$

If the vectors are parallel transported we have e.g. $t^a\nabla_a v^c = 0$, then one immediately sees that the metric compatibility of the connection is the condition that gives that the derivative of $g(v, u)$ is zero:

$$0 = t^a\nabla_a(g_{cb}v^c u^b) = t^a v^c u^b\nabla_a g_{cb} + g_{cb}\underbrace{t^a\nabla_a v^c}_{=0}u^b + g_{cb}v^c\underbrace{t^a\nabla_a u^b}_{=0} = v^c u^b t^a\nabla_a g_{cb} \ . \tag{4.23}$$

4.6 Riemann tensor

The Riemann tensor measures the noncommutation of covariant derivatives.

Given a connection, a 1-form, and a function one has:

$$\nabla_a\nabla_b(f\omega_c) = \nabla_a(f\nabla_b\omega_c) + \nabla_a(\omega_c\nabla_b f) = \nabla_a\omega_c\nabla_b f + \omega_c\nabla_a\nabla_b f + \nabla_a f\nabla_b\omega_c + f\nabla_a\nabla_b\omega_c \ , \tag{4.24}$$

from which one sees that the commutator

$$[\nabla_a, \nabla_b](f\omega_c) = (\nabla_a\nabla_b - \nabla_b\nabla_a)(f\omega_c) = f(\nabla_a\nabla_b - \nabla_b\nabla_a)\omega_c = f[\nabla_a, \nabla_b]\omega_c \ , \tag{4.25}$$

is a linear map from $T^*_p\mathcal{M}$ to $\tau(0, 3)$. In other terms,

Definition 4.6.1. *It exists a tensor $R \in \tau(1, 3)$ called* Riemann tensor *such that*

$$R_{abc}{}^d\omega_d := [\nabla_a, \nabla_b]\omega_c \ . \tag{4.26}$$

Using the deifnition above, a similar calculation to the one for the connection shows that for vectors

$$[\nabla_a, \nabla_b]v^a = -R_{abd}{}^c v^a \ , \tag{4.27}$$

and in general for tensors

$$[\nabla_a, \nabla_b]T^{c_1\dots c_k}_{d_1\dots d_l} = -\sum_i R_{abe}{}^{c_i}T^{c_1\dots e\dots c_k}_{d_1\dots d_l} + \sum_j R_{abd_j}{}^e T^{c_1\dots c_k}_{d_1\dots e\dots d_l} \ . \tag{4.28}$$

Properties.
1. Antisymmetry in index 1 and 2: $R_{abc}{}^d = -R_{bac}{}^d$. This follows from the definition Eq. (4.26).
2. $R_{[abc]}{}^d = 0$. Direct proof:
$$R_{[abc]}{}^d\omega_d = \nabla_{[a}\nabla_b\omega_{c]} - \nabla_{[b}\nabla_a\omega_{c]} = 2\nabla_{[a}\nabla_b\omega_{c]} = 2\partial_{[a}\partial_b\omega_{c]} = 2\mathbf{d}^2\omega = 0 \ . \tag{4.29}$$
3. If ∇ is metric compatible: $R_{abcd} = -R_{abdc}$. Direct proof using the definition:
$$0 = (\nabla_a\nabla_b - \nabla_b\nabla_a)g_{cd} = R_{abc}{}^e g_{ed} + R_{abd}{}^e g_{ec} \ . \tag{4.30}$$

4. *Bianchi identities.* $\nabla_{[a}R_{bc]d}{}^e = 0$.
   Proof. Take the two expressions

$$\begin{cases}(\nabla_a\nabla_b - \nabla_b\nabla_a)\nabla_c\omega_d & = R_{abc}{}^e\nabla_e\omega_d + R_{abd}{}^e\nabla_c\omega_e \\ \nabla_a(\nabla_b\nabla_c\omega_d - \nabla_c\nabla_b\omega_d) & = \nabla_a(R_{bcd}{}^e\omega_e) = \omega_e\nabla_a R_{bcd}{}^e + R_{bcd}{}^e\nabla_a\omega_e\end{cases} \tag{4.31}$$

   and antisymmetrize them w.r.t. $[abc]$: the l.h.s. become equal, matching the r.h.s. and suing property 2. one gets the result:

$$\underbrace{R_{[abc]}{}^e}_{=0}\nabla_a\omega_d + \cancel{R_{[ab|d|}{}^f\nabla_{|c]}}\omega_f = \omega_e\nabla_{[a}R_{bc]d}{}^e + \cancel{R_{[bc|d}{}^e\nabla_{|a]}\omega_e} \ . \tag{4.32}$$

5. Symmetric in pairs of down indexes: $R_{abcd} = R_{cdab}$ [exercise].
6. In a coordinate basis the Riemann tensor "$R \sim \partial\Gamma - \partial\Gamma + \Gamma\Gamma - \Gamma\Gamma$" [exercise]:

$$R^\sigma_{\mu\nu\rho} = \partial_\nu\Gamma^\sigma_{\mu\rho} - \partial_\mu\Gamma^\sigma_{\nu\rho} + \sum_\alpha(\Gamma^\alpha_{\mu\rho}\Gamma^\sigma_{\alpha\nu} - \Gamma^\alpha_{\nu\rho}\Gamma^\sigma_{\alpha\mu}) \ . \tag{4.33}$$

   Note the upper index is often "shifted back" in this coordinate notation.

Riemann contractions.

Definition 4.6.2. *The* Ricci tensor *is the* $(0,2)$ *symmetric tensor* $R_{ab} := R_{acb}{}^c$.

Definition 4.6.3. *The* Ricci scalar *is the trace* $R := g^{ab}R_{ab} = R^a_a$.

The *contracted Bianchi identities* are obtained from the contraction $C_{(14)}$ of the Bianchi identities $\nabla_{[a}R_{bc]d}{}^e = 0$, i.e. contracting a with e, and using the symmetries of the Riemann:

$$\begin{aligned}
0 &= \nabla_{[a}R_{bc]d}{}^a = \nabla_a R_{bcd}{}^a + \nabla_b R_{cd} - \nabla_c R_{bd} && (4.34a)\\
0 &= \nabla_a R_{bc}{}^{da} + \nabla_b R_c^d - \nabla_c R_b^d \quad \text{(raise } d \text{ with the metric and use metric compatibility)} && (4.34b)\\
0 &= \nabla_a R_c{}^a + \nabla_b R_c^b - \nabla_c R_b^b \quad \text{(contract } b \text{ with } d) && (4.34c)\\
&= 2\nabla_a R_c{}^a - \nabla_c R \quad \Rightarrow && (4.34d)\\
0 &= \nabla^a(R_{ca} - \frac{1}{2}g_{ca}R) =: \nabla^a G_{ac} \ . && (4.34e)
\end{aligned}$$

Definition 4.6.4. *The* Einstein tensor *is the* $(0,2)$ *symmetric tensor* $G_{ab} := R_{ab} - \frac{1}{2}Rg_{ab}$.

Riemann tensor as measure of the deviation of parallel transported vectors. Consider the parallel transport of a vector from p to $q = p + \delta p$ along two different paths. One path passes by intermediate point r and the other by intermediate point s. The coordinates of the four points are

$$p: \ x_0^\mu \ , \quad r: \ x_0^\mu + dx_1^\mu \ , \quad s: \ x_0^\mu + dx_2^\mu \ , \quad q: \ x_0^\mu + dx_1^\mu + dx_2^\mu \ . \tag{4.35}$$

The parallel transport equation for vector v components are

$$0 = t^\mu\partial_\mu v^\alpha + t^\mu\Gamma^\alpha_{\mu\beta}v^\beta = dx^\mu\partial_\mu v^\alpha + dx^\mu\Gamma^\alpha_{\mu\beta}v^\beta \quad \Rightarrow \quad dx^\mu\partial_\mu v^\alpha = -dx^\mu\Gamma^\alpha_{\mu\beta}v^\beta \ . \tag{4.36}$$

The transported vector along $p\bar{r}q$ is (up to second order in dx)

$$\begin{aligned}
v^\alpha(r) &= v^\alpha(p) + dx^\mu\partial_\mu v^\alpha(p) + ... = v^\alpha(p) - \Gamma^\alpha_{\mu\beta}(p)v^\beta(p)dx_1^\mu + ... && (4.37a)\\
v^\alpha(q) &= v^\alpha(r) - \Gamma^\alpha_{\mu\beta}(r)v^\beta(r)dx_2^\mu && (4.37b)
\end{aligned}$$

with

$$\begin{aligned}
\Gamma^\alpha_{\mu\beta}(r) &= \Gamma^\alpha_{\mu\beta}(p) + \frac{\partial\Gamma^\alpha_{\mu\beta}}{\partial x^\nu}(p)dx_1^\mu + \mathcal{O}((dx_1^\mu)^2) \quad \Rightarrow && (4.37c)\\
v^\alpha(q) &= v^\alpha(p) - \Gamma^\alpha_{\mu\beta}v^\beta(p)dx_1^\mu - \Gamma^\alpha_{\mu\beta}v^\beta(p)dx_2^\mu + \Gamma^\alpha_{\beta\nu}\Gamma^\beta_{\sigma\mu}v^\sigma(p)dx_1^\nu dx_2^\mu - \partial_\nu\Gamma^\alpha_{\sigma\mu}v^\sigma(p)dx_1^\mu dx_2^\nu \ . && (4.37d)
\end{aligned}$$

Similarly the transported vector along $p\bar{s}q$ is

$$v'^\alpha(q) = v^\alpha(p) - \Gamma^\alpha_{\mu\beta}v^\beta(p)dx_1^\mu - \Gamma^\alpha_{\mu\beta}v^\beta(p)dx_2^\mu + \Gamma^\alpha_{\beta\mu}\Gamma^\beta_{\sigma\nu}v^\sigma(p)dx_1^\mu dx_2^\nu - \partial_\mu\Gamma^\alpha_{\beta\nu}v^\sigma(p)dx_1^\mu dx_2^\nu \ . \tag{4.38}$$

Note the indexes μ and ν are exchanged in the two expressions and the $\Gamma\Gamma$ and $\partial\Gamma$ contractions are different. Taking the difference the terms $\propto dx$ cancel each other but the terms $proptodx_1dx_2$ do not because different indexes are contracted. At second order in dx one obtains

$$v'^\alpha(q) - v^\alpha(q) = \left[\underbrace{\Gamma^\alpha_{\beta\mu}\Gamma^\beta_{\sigma\nu}v^\sigma(p) - \Gamma^\alpha_{\beta\nu}\Gamma^\beta_{\sigma\mu}v^\sigma(p) + \partial_\nu\Gamma^\alpha_{\beta\mu}v^\sigma(p) - \partial_\mu\Gamma^\alpha_{\beta\nu}v^\sigma(p)}_{=\partial\Gamma-\partial\Gamma+\Gamma\Gamma-\Gamma\Gamma}\right]dx_1^\mu dx_2^\nu \tag{4.39}$$

Summary 4.6.1. *The Riemann tensor measures the noncommutation of covariant derivatives and determines the variation of vector fields parallel trasnported along an infiniteismal curve.*

- *At first order in dx the parallel transport in independent on dx.*
- *At second order in dx the deviation of a parallel transported vector is determined by the Riemann tensor.*
- *In flat space: $R_{abc}{}^{d} \equiv 0$.*

4.7 Geodesics

Intuitively, geodesics are lines "as straight as possible" or lines of minimal lenght. A more technical definition is

Definition 4.7.1. Geodesic $=$ *a curve $\gamma\ :\ \mathbb{R} \mapsto \mathcal{M}$ whose tangent vector t^a maintains its direction, i.e. $t^a\nabla_a t^b = \alpha t^b$* .

It is always possible to change the curve parametrization such that the constant $\alpha = 0$ (*affine parametrization*, see below). Using this parametrization the definition of geodesic is equivalent to

Definition 4.7.2. Geodesic $=$ *curve whose tagent vector is parallel propagaed along itself, $t^a\nabla_a t^b = 0$.*

In a coordinate basis, the curve is $x^\mu(\lambda)$ and the tangent vector components are $t^\mu = \dot{x}^\mu$; the *geodesic equation* reads

$$0 = \frac{dt^\mu}{d\lambda} + \Gamma^\mu_{\sigma\nu} t^\sigma t^\nu = \frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\sigma\nu}\frac{dx^\sigma}{d\lambda}\frac{dx^\nu}{d\lambda} \ . \tag{4.40}$$

The geodesic equation can be also written with index down by using the contracted Christoffel symbols

$$g_{\alpha\sigma}\Gamma^\alpha_{\mu\nu} = \frac{1}{2}\underbrace{g_{\alpha\sigma}g^{\alpha\beta}}_{=\delta^\beta_\sigma}(\partial_\mu g_{\beta\nu} + \partial_\nu g_{\beta\nu} - \partial_\mu g_{\alpha\beta}) = \frac{1}{2}(\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\sigma\nu} - \partial_\sigma g_{\mu\nu}) \ , \tag{4.41}$$

to find

$$0 = g_{\mu\nu}\frac{dt^\nu}{d\lambda} + g_{\mu\nu}\Gamma^\nu_{\alpha\beta}t^\alpha t^\beta = \frac{dt_\mu}{d\lambda} - \partial_\sigma g_{\mu\nu}t^\sigma t^\nu + g_{\mu\nu}\Gamma^\nu_{\alpha\beta}t^\alpha t^\beta \tag{4.42a}$$

$$= \frac{dt_\mu}{d\lambda} - \partial_\sigma g_{\mu\nu}t^\sigma t^\nu + \frac{1}{2}(\partial_\beta g_{\mu\alpha} + \partial_\alpha g_{\mu\beta} - \partial_\mu g_{\alpha\beta})t^\alpha t^\beta \tag{4.42b}$$

$$= \frac{dt_\mu}{d\lambda} - \cancel{\partial_\sigma g_{\mu\nu}t^\sigma t^\nu} + \frac{\cancel{2}}{\cancel{2}}\cancel{\partial_\beta g_{\mu\alpha}t^\alpha t^\beta} - \frac{1}{2}\partial_\mu g_{\alpha\beta}t^\alpha t^\beta \tag{4.42c}$$

$$= \frac{dt_\mu}{d\lambda} - \frac{1}{2}\partial_\mu g_{\alpha\beta}t^\alpha t^\beta \ . \tag{4.42d}$$

Observations.

- The geodesic equation is a system of ODE of 2nd order whose solution exist locally for a given initial data. Given a vector t^a at point p, exists locally a unique geodesic passing through p with tangent t^a.
- *Affine parametrization.* Consider the curve re-parametrization $\lambda(s)$, since

$$\frac{d}{d\lambda} = \frac{ds}{d\lambda}\frac{d}{ds} =: f\frac{d}{ds} \ , \quad \frac{dx^\nu}{d\lambda} = \dot{x}^\nu = f\frac{dx^\nu}{ds} = (x^\mu)' \ , \tag{4.43}$$

  one obtains immediately (index are omitted for brevity)

$$0 = \ddot{x} + \Gamma\dot{x}\dot{x} = x'' + \Gamma x' x' + \frac{f'}{f^2}x' \ . \tag{4.44}$$

  Setting $\alpha := -f'/f^2$, the parametrization such that $\alpha = 0$ is the solution of the equation $f'/f^2 = 0$, i.e. $0 = f' = d^2s/d\lambda^2 \Rightarrow s = a\lambda + b$ with $a, b \in \mathbb{R}$.

Geodesics extremize the length of the curve between two points. Focusing on $n = 4$ and using a metric of signature $(-,+,+,+)$, a geodesic γ is classified according to its tangent vector t^a:

γ	$g(t,t)$	Invariant spacetime interval s
null	$= 0$	$s = \ell = 0$
timelike	< 0	$s = \tau = \int \sqrt{g_{ab}t^a t^b}d\lambda$
spacelike	> 0	$s = \ell = \int \sqrt{-g_{ab}t^a t^b}d\lambda$

Observations:

- s is not defined for a geodesic that changes character from spacelike to timelike.

- Geodesic in a Lorentz manifold **cannot** change from timelike to other type because, from the definition of parallel transport, the norm of the tangent vector must be constant:

$$t^c\nabla_c(g_{ab}t^a t^b) = 2g_{ab}\underbrace{t^c\nabla_c t^b}_{=0} = 0 \quad \Rightarrow \quad g_{ab}t^a t^b = const \ . \tag{4.45}$$

- Proper length or time do not depend on the parametrization.

Take a spacelike geodesic parametrized by $\lambda \in [a,b] \in \mathbb{R}$ such that the norm of the tangent vector norm is unity. Calculate the variation of the length under an infinitesimal coordinate transformation $x^\mu \mapsto x^\mu + \delta x^\mu$ assuming the variation is zero at the boundary $\delta x^\mu(a) = \delta x^\mu(b) = 0$:

$$\delta\ell = \int_a^b \underbrace{(g_{\mu\nu}t^\mu t^\nu)^{-1/2}}_{=1} \left(\underbrace{g_{\alpha\beta}\dot{x}^\alpha \frac{d\delta x^\beta}{d\lambda}}_{b.p.} + \frac{1}{2} \underbrace{\frac{\partial g_{\alpha\beta}}{\partial x^\sigma}\delta x^\sigma \dot{x}^\alpha \dot{x}^\beta}_{\sigma\to\beta;\beta\to\nu} \right) d\lambda = \tag{4.46a}$$

$$= \int_a^b \frac{d}{d\lambda}(g_{\alpha\beta}\dot{x}^\alpha d\delta x^\beta)d\lambda - \int_a^b \frac{d}{d\lambda}(g_{\alpha\beta}\dot{x}^\alpha)\delta x^\beta d\lambda + \frac{1}{2}\int_a^b \frac{\partial g_{\alpha\nu}}{\partial x^\beta}\delta x^\beta \dot{x}^\alpha \dot{x}^\nu d\lambda = \tag{4.46b}$$

$$= \underbrace{[g_{\alpha\beta}\dot{x}^\alpha d\delta x^\beta]_a^b}_{=0} - \int_a^b \left(\frac{d}{d\lambda}(g_{\alpha\beta}\dot{x}^\alpha)\delta x^\beta - \frac{1}{2}\int_a^b \frac{\partial g_{\alpha\nu}}{\partial x^\beta}\delta x^\beta \dot{x}^\alpha \dot{x}^\nu \right) d\lambda = \tag{4.46c}$$

$$= \int_a^b \left(\underbrace{-\frac{d}{d\lambda}(g_{\alpha\beta}\dot{x}^\alpha) + \frac{1}{2}\frac{\partial g_{\alpha\nu}}{\partial x^\beta}\dot{x}^\alpha \dot{x}^\nu}_{\text{geodesic eq.}} \right) \delta x^\beta d\lambda \ . \tag{4.46d}$$

The above calculation proves that a geodesic extremizes the length. A similar calculation can be performed for timelike geodesics. Note that geodesics equations can be derived from the Lagrangian

$$L^2 = g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu \ . \tag{4.47}$$

Riemann normal coordinates. The uniqueness of geodesic allows us to define special coordinates at a point p such that all geodesics passing by p are mapped into straight lines in $\mathbb{R}^n$.

Given $v \in T_p\mathcal{M}$, exists a unique godesics with tangent vector v such that $\gamma_v(\lambda = 0) = p$. This implies that one can associate each point q around p to the vector v that generates the geodesic γ_v from p to q. Specifically, one uses the *exponential map*

$$\mathrm{Exp}_p : T_p\mathcal{M} \mapsto \mathcal{M} \ \text{ with } \ \mathrm{Exp}_p(v) := \gamma_v(\lambda = 1) \ , \tag{4.48}$$

and the point q is identified as the one connected by "unit time" movement along the geodesic [1].

Definition 4.7.3. *The* Riemann coordinates *of point q (around p) are defined as the components of the vector $\psi(q) = (v^1, ..., v^n)$.*

Observations:

- The existance of the exponential map is guaranteed only <u>locally</u> as a consequence of local existance of geodesics. Far from p, two geodesic originating from p can actually cross! In this case the manifold is called *geodesically incomplete*.
- It can be proven that the exponential map is locally a one-to-one map.
- Point p has normal coordinates $\psi(p) = (0, ..., 0)$.
- In normal coordinates the curve γ_v is represented as straight lines in $\mathbb{R}^n$:

$$\gamma_v \ : \ x^\mu = (\lambda v^1, ..., \lambda v^n) \ . \tag{4.49}$$

Because these geodesic are straight lines, they satisfy $\ddot{x}^\mu = 0$. Comparing with Eq. (4.40), one finds that the Christoffel symbols in normal coordinates must be zero:

$$\frac{d^2x^\mu}{d\lambda} = 0 \quad \Rightarrow \quad \Gamma^\mu_{\nu\rho} = 0 \ \text{ at } p \ . \tag{4.50}$$

In turn, this implies that $\nabla_\mu = \partial_\mu$ and by metric compatibility the derivatives of the metric components in normal coordinates must be zero (schematically, omitting indexes):

$$0 = \nabla g = \partial g + \Gamma g = \partial g \ \text{ at } p \ . \tag{4.51}$$

[1] Perhaps the simplest explanation of the name is to think of a map between the tangent to the unit circle and the unit circle $t \mapsto \exp(it)$.

- The above observation is key in GR: one can always use normal coordinates such that the metric at a point reduces to the Minkowski:
$$g_{\mu\nu} = \eta_{\mu\nu} \quad \text{at } p \ , \tag{4.52}$$
while the expansion of the metric around p differs from Mikowski only by second derivatives (curvature terms):
$$g_{\mu\nu} \simeq \eta_{\mu\nu} + \frac{\partial^2 g}{\partial x \partial x} dx dx \ . \tag{4.53}$$

4.8 Geodesics deviation

An intuitive effect of curvature is that geodesics focus. On a 2D plane the curvature is zero and straight lines never meet. On a 2-sphere the curvature is positive $\kappa = 1/R$ and the meridians focus on the poles. Is this effect described by the formalism developed so far and captured by the Riemann tensor?

Consider a one-parameter family of geodesics $\gamma_a(\lambda)$ where λ is the affine parameter and $\sigma \in \mathbb{R}$ control a smooth variation from one geodesic to the next. Assume the geodesics do not cross; $\gamma_\sigma(\lambda)$ is a 2D surface in a n-dimensional manifold. Define the vector fields

$$t^a := (\frac{\partial}{\partial \lambda})^a \quad : \ t^a \nabla_a t^b = 0 \ \ \text{Tangent to } \gamma \tag{4.54a}$$
$$s^a := (\frac{\partial}{\partial \sigma})^a \quad \textit{Deviation vector}\text{, infinitesimal displacement from } \gamma_\sigma \text{ to } \gamma_{\sigma+d\sigma} \ . \tag{4.54b}$$

Properties:
- The two vector commute:
$$[t,s] = t^\mu \partial_\mu (s^\nu \partial_\nu) - s^\mu \partial_\mu (t^\nu \partial_\nu) = \frac{\partial x^\mu}{\partial \lambda} \partial_\mu (\frac{\partial x^\nu}{\partial \sigma} \partial_\nu) - \frac{\partial x^\mu}{\partial \sigma} \partial_\mu (\frac{\partial x^\nu}{\partial \lambda} \partial_\nu) \tag{4.55a}$$
$$= \cancel{\frac{\partial x^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \sigma} \partial_\mu \partial_\nu} + \frac{\partial x^\mu}{\partial \lambda} \frac{\partial}{\partial \sigma} (\underbrace{\frac{\partial x^\nu}{\partial x^\mu}}_{\delta^\nu_\mu}) \partial_\nu - \cancel{\frac{\partial x^\mu}{\partial \sigma} \frac{\partial x^\nu}{\partial \lambda} \partial_\mu \partial_\nu} - \frac{\partial x^\mu}{\partial \sigma} \frac{\partial}{\partial \lambda} (\underbrace{\frac{\partial x^\nu}{\partial x^\mu}}_{\delta^\nu_\mu}) \partial_\nu \tag{4.55b}$$
$$= \frac{\partial x^\mu}{\partial \lambda} \underbrace{\frac{\partial}{\partial \sigma} (\delta^\nu_\mu)}_{=0} \partial_\nu - \frac{\partial x^\mu}{\partial s} \underbrace{\frac{\partial}{\partial \lambda} (\delta^\nu_\mu)}_{=0} \partial_\nu = 0 \ . \tag{4.55c}$$
- $t^a s_a = const$ along the geodesic. Because the commutator is zero, $0 = [t,s] \Leftrightarrow t^a \nabla_a s^b = s^a \nabla_a t^b$, it is immediate:
$$t^c \nabla_c (t^a s_a) = \underbrace{t^c \nabla_c t^a}_{=0} s_a + t^c \underbrace{t^a \nabla_c s_a}_{=t_a \nabla_c s^a} = t_a t^c \nabla_c s^a \underset{[t,s]=0}{=} t_a s^c \nabla_c t^a = \frac{1}{2} s^c \nabla_c (\underbrace{t^a t_a}_{=-1}) = 0 \ . \tag{4.56a}$$
- By a suitable parametrization the constant can be chosen zero $t^a s_a = 0$.

Definition 4.8.1. Geodesic relative velocity = *the rate of change of the deviation vector along* γ_σ, $V^a := t^b \nabla_b s^a$.
Geodesic relative acceleration $A^a := t^b \nabla_b V^a$.

The acceleration of the geodesics is governed by the Riemann tensor, as expressed by the *geodesic deviation equation*:

$$A^a = t^c \nabla_c V^a = t^c \nabla_c (t^b \nabla_b s^a) \tag{4.57a}$$
$$= t^c \nabla_c (s^b \nabla_b t^a) \qquad (\Leftarrow [t,s] = 0) \tag{4.57b}$$
$$= \underbrace{t^c \nabla_c s^b}_{=s^b \nabla_c t^c} \nabla_b t^a + t^c s^b \underbrace{\nabla_c \nabla_b t^a}_{=\nabla_b \nabla_c t^a - R_{cbd}{}^a t^d} \tag{4.57c}$$
$$= s^b \nabla_c t^c \nabla_a t^a + t^c s^b \nabla_b \nabla_c t^a - R_{cbd}{}^a t^d t^c s^b = s^c \nabla_c (\underbrace{t^b \nabla_b t^a}_{=0}) - R_{cbd}{}^a t^d = -R_{cbd}{}^a t^d t^c s^b \ . \tag{4.57d}$$

The Riemann tensor measures how geodesics bends and accelerate toward each other. If the Riemann tensor is zero, then the acceleration betwee two nearby geodesics is zero.

Note the geodesic deviation equation is sometimes written as
$$A^a = t^c \nabla_c (t^b \nabla_b s^a) =: \nabla^2_t s^a = -R_{cbd}{}^a t^c s^b t^d \ . \tag{4.58}$$

Example 4.8.1. Tidal forces in Newtonian gravity. *The Newton equations for two test bodies in gravitational potential* ϕ *with nearby trajectories* $x^i(t)$ *and* $x^i(t) + s^i(t)$ *are (up to* $\mathcal{O}(s^2)$ *terms)*
$$\ddot{x}^i(t) = -(\partial_i \phi)|_{x(t)} \quad \text{and} \quad \ddot{x}^i(t) + \ddot{s}^i(t) = -(\partial_i \phi)|_{x(t)+s(t)} \approx -(\partial_i \phi)|_{x(t)} - (\partial_j \partial_i \phi)|_{x(t)} s^j \ . \tag{4.59}$$

The relative acceleration between the two bodies is thus given by

$$\ddot{s}^i(t) = -\frac{\partial^2 \phi}{\partial x^i \partial x^j}|_{x(t)} s^j \ . \tag{4.60}$$

The quantity $\partial_{ij}\phi$ is known as tidal tensor. *The equation above is similar to the geodesic deviation equation and in fact it is precisely its Newtonian limit Wald (1984), that contains the correpondence*

$$\textit{Riemann tensor} \quad \leftrightarrow \quad \textit{Tidal forces} \ . \tag{4.61}$$

Since the Riemann tensor is not zero in normal coordinates, the above correspondence indicate that the relative acceleration due to tidal forces cannot be transformed away in GR.

4.9 Gravitational redshift

As an application of the geodesic equation, let us consider the problem of gravitational redshift and its formulation in terms of geodesics for (i) a static spacetime with small curvature,, and (ii) a cosmological spacetime.

Weak field. Recall that the simplest approach to the problem is to observe that, because of the Einstein's equivalence principle (EEP), the classical doppler redshift of photons due to the accelerate motion of the emitting source must map to the gravitational problem in which the acceleration is given by the gravitational field. Since the same formulas must apply, the relative difference in wavelength is given by the difference of the gravitational potential between two points (emitter and receiver) at distance d:

$$\frac{\Delta\lambda}{\lambda} \simeq \frac{dg_N}{c^2} = \frac{\Delta\phi}{c^2} \ , \tag{4.62}$$

where g_N is the Newton grav. acceleration. We give a more rigorous formulation and solution to the problem below.

Problem: A photon is emitted at point p with frequency ν_p and absorbed at q with frequency ν_q. Determine the relation between the frequencies mesured by a stationary observer with 4-velocity $u^a = (u^0, 0)$.

Let us assume the gravitational field is described by the metric

$$g = -(1+2\phi)\mathrm{d}t^2 + (1-2\phi)(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) \ , \tag{4.63}$$

where $\phi = \phi(x, y, z)$ is time independent and $|\phi| \ll 1$; we work at first order in ϕ. The metric above is the metric of a static and weak gravitational field, as we shall see later. The coordinates reduce to an inertial Cartesin system in SR for $\phi = 0$. A direct calculation left as exercise shows the Riemann tensor is nonzero and composed of second spatial derivatives of ϕ. We will see later that the scalar ϕ correspond to the Newtonian potential.

Photon trajectories are null geodesics $x^\mu(\lambda)$ in the metric g, where λ is an affine parameter. The tangent vector is the photon 4-momentum $p^\mu = \dot{x}^\mu$. The energy of the photon measured by the observer with 4-velocity u is defined as

$$E = p_\mu u^\mu = g_{\mu\nu} p^\nu u^\mu \ . \tag{4.64}$$

Note that because

$$-1 = u^\mu u_\mu = g_{00} u^0 u^0 \quad \Rightarrow \quad u^0 = (-g_{00})^{-1/2} = (1+2\phi)^{-1/2} \ , \tag{4.65}$$

the energy expression becomes

$$E = p_\mu u^\mu = p_0 u^0 = p_0 (1+2\phi)^{-1/2} \simeq p_0(1-\phi) \ . \tag{4.66}$$

It only remains to compute p_0. Take the geodesic equation for "index-down" 4-momentum and see that the component $\mu = 0$ of the above equation is zero because the metric is static:

$$\frac{dp_\mu}{d\lambda} = \frac{1}{2}\partial_\mu g_{\alpha\beta} p^\alpha p^\beta \quad \Rightarrow \quad \frac{dp_0}{d\lambda} = \frac{1}{2}\partial_0 g_{\alpha\beta} p^\alpha p^\beta = 0 \ . \tag{4.67}$$

Thus $p_0(p) = p_0(q) = const =: \bar{p}_0$. Finally, the ratio of the frequencies/energies gives the expected result:

$$\frac{\nu_q}{\nu_p} = \frac{E_q}{E_p} = \frac{p_\mu u^\mu(q)}{p_\mu u^\mu(p)} = \frac{\cancel{\bar{p}_0} u^0(q)}{\cancel{\bar{p}_0} u^0(p)} \simeq \frac{1-\phi(q)}{1-\phi(p)} \simeq 1 - (\phi(q) - \phi(p)) = 1 - \Delta\phi \ . \tag{4.68}$$

Expanding universe in 2D. Let us repeat the above calculation for the cosmological 2D spacetime discussed previously where the metric is

$$g = -\mathrm{d}t^2 + a^2(t)\mathrm{d}x^2 \ . \tag{4.69}$$

The $\mu = 0$ component of null geodesics is solved by

$$\frac{dp^0}{d\lambda} + \frac{\dot{a}}{a}\left(p^0\right)^2 = \frac{dp^0}{dt}p^0 + \frac{\dot{a}}{a}\left(p^0\right)^2 = 0 \quad \Rightarrow \quad p^0(t) = \frac{w_0}{a(t)} \ , \quad w_0 \in \mathbb{R} \ . \tag{4.70}$$

The energy measured by an observer comving with the Universe $u^\mu = (1, 0)$ is

$$E = -p_\mu u^\mu = -g_{\mu\nu}p^\mu u^\mu = -g_{00}p^0 = +\frac{w_0}{a} \ . \tag{4.71}$$

When $a = 1$ the phothon frequency is $E = \hbar w_0$. The energy of the phothon emitted at t_1 is $E_1 = E(t_1) = w_0/a(t_1) = w_0/a_1$, thus

$$\frac{E_2}{E_1} = \frac{a_1}{a_2} \quad \Rightarrow \quad z = \frac{E_1 - E_2}{E_2} = \frac{a_2}{a_1} - 1 \ . \tag{4.72}$$

If the Universe is expanding $a_2 > a_1$, then $E_2 < E_1$ and the phton is redshifted (cosmological redshift). In an expanding Universe the wavelength grows with time. This gives a way to meaure distances: larger redshifts correspond to larger distances, associating a redshift to a reference distance (standard candle) one can calculate all other distances by resdhift measurement (e.g. spactral absorption line of galaxies). Note this is a distinct effect from Doppler shift.

5. Equivalence Principles

①
This lecture gives an overview of the equivalence principles and the experimental test of the fundations of GR.

Suggested readings. *Will (2014); Di Casola et al. (2015).*

5.1 Role of equivalence principles

Roles of equivalence principles (EPs) according to different views:
- Fundation of the theory;
- Euristic principles based on experimental facts;
- "inspirational" principles.

In all cases they
- Help understanding the theory and its equation, proving us with interpretation/intuition;
- Are the basis for designing experimental tests of the theory and its fundations;
- Constitute a common ground to compare gravity theories and formulations.

5.2 Weak equivalence principle (WEP)

WEP or "universality of free fall". In a gravitational field, test-bodies with negligible self-gravity behave independently of their properties.

Definition 5.2.1. Test-body = *does not back-react on the gravitational field.*

Definition 5.2.2. *The* self-gravity *of a body of mass M and size r is measured by $\sigma = 2GM/(c^2 r)$. Negligible self-gravity means in the limit $\sigma \ll 1$.*

Remark 5.2.1. *The concept of test-body and negligible self-gravity are two different things. The former depends on the external gravitational field, the latter only on its mass and size. There exists test-bodies for which the self-gravity is nonnegligible, and bodies with negligible self-gravity tha cannot be considered test-bodies. Examples:*
- *Pebble in the Earth grav. field: test-body and $\sigma \ll 1$.*
- *Moon in the Earth grav. field: $\sigma \ll 1$, but it cannot be considered a test-body since it affect the Earth grav.field (tides).*
- *Micro black hole with mass $m_{BH} \sim m_{Planck} \approx 22\mu g$ in the Earth grav. field: self-gravity is maximal $\sigma_{BH} = 1$, but it can be considered a test-body on Earth.*

5.3 Newton equivalence principle (NEP)

NEP. In the Newtonian limit, the inertial and gravitational mass are equal.

Observations
- m_g and m_i are quantities defined in Newtonian physics only.
- Any theory of relativity must have the same Newtonian limit.
- WEP $\Rightarrow$ NEP ...
- ... but in general, NEP $\not\Rightarrow$ WEP, because the validity of WEP depends on the equations of motions (EOM). If the EOM are Newton law

$$\ddot{\vec{x}} = -\frac{m_g}{m_i}\frac{GM\vec{x}}{x^3} \ , \tag{5.1}$$

then NEP $\Rightarrow$ WEP. This remain valid for any EOM that depends on the ratio m_g/m_i. But if the EOM contain other combinations of the two masses, then it might not be true.

5.4 Einstein equivalence principle (EEP)

EEP. Fundamental, nongravitational test-particles and fields are locally and at any point of spacetime not affected by the presence of a gravitational field.

This principle encodes the idea that local frames in presence of a gravitational field are equivalent to local frame in absence of the gravitational field. Examples:

- A local nonrotating fre-falling frame in a grav. field = a local inertial frame in absence of gravity.
- A local frame in a grav. field = suitably chosen accelerated frame in absence of gravity (e.g. the rocket simulating the gravitational acceleration).

Meaning of local. Sufficiently small region of spacetime such that a given instrument does not resolve variations of the grav.field and/or tidal forces. Note, however, that the relative fractional acceleration between two free-falling bodies is governed by a tidal tensor

$$\frac{\ddot{s}^i}{s} \sim -\frac{\partial^2 \phi}{\partial x \partial x} \ , \tag{5.2}$$

that does not vanish for $s^i \to 0$ (inhomogeneous field). Geodesic deviation and "composed systems" violate the EEP, in general. The locality requirement should then go together with the request that it applied to fundamental particles and fields. Note there remains a difficulty in the definition of what is fundamental and what is not...

Formulation of EEP for Poincare' invariant physics. The EEP formulated above is general. We know (or assume) that fundamental nongravitational physics laws must be invariant under the Poincare' group (translations and Lorentz transformations). The EEP can be thus re-formulated more specifically as

1. WEP is valid.
2. Local Lorentz invariance (LLI) is valid.
3. Local position invariance (LLP) is valid.

Definition 5.4.1. *LLI = local nongravitational experiments are independent on the velocity of the free-falling frame in which they are performed.*

Definition 5.4.2. *LLP = local nongravitational experiments are independent of where and when in the Universe are performed.*

5.5 Strong equivalence principle (SEP)

The EEP does not include gravitational phenomena. An extension to those is the

SEP. All fundamental test physics is locally not affected by the presence of a gravitational field.

SEP includes tests but self-gravitating $\sigma \sim 1$ experiments. Examples:

- Cavedish experiment (1798), mutual attraction between two light bodies.
- Gravitational-wave detections.
- Any experiment in which a background gravitational field can be identified and the latter does not affect the measure.

An alternative formulations of SEP is

1. WEP is valid for self-gravitating test bodies.
2. Any local test-experiment is independent of the velocity of the fre-falling apparatus.
3. Any local test experiment is independent of when and where in the Universe is perfomed.

5.6 Experimental tests

Experimental tests of WEP. A basic way to test WEP is to define $\alpha := m_g/m_i$ and observe that the relative acceleration between 2 bodies is $\delta a = |a_1 - a_2| \propto |\alpha_1 - \alpha_2|$. Most of the experiments measure the fractional relative acceleration between two bodies, as given by the *Eötvos parameter*

$$\eta := 2\frac{|a_1 - a_2|}{|a_1 + q_2|} = 2\frac{|\alpha_1 - \alpha_2|}{|\alpha_1 + \alpha_2|} = f(\frac{m_g}{m_i}) \ . \tag{5.3}$$

A "null" experiment that verifies $f(1) = 0$ proves the WEP to a certain accuracy. Examples:

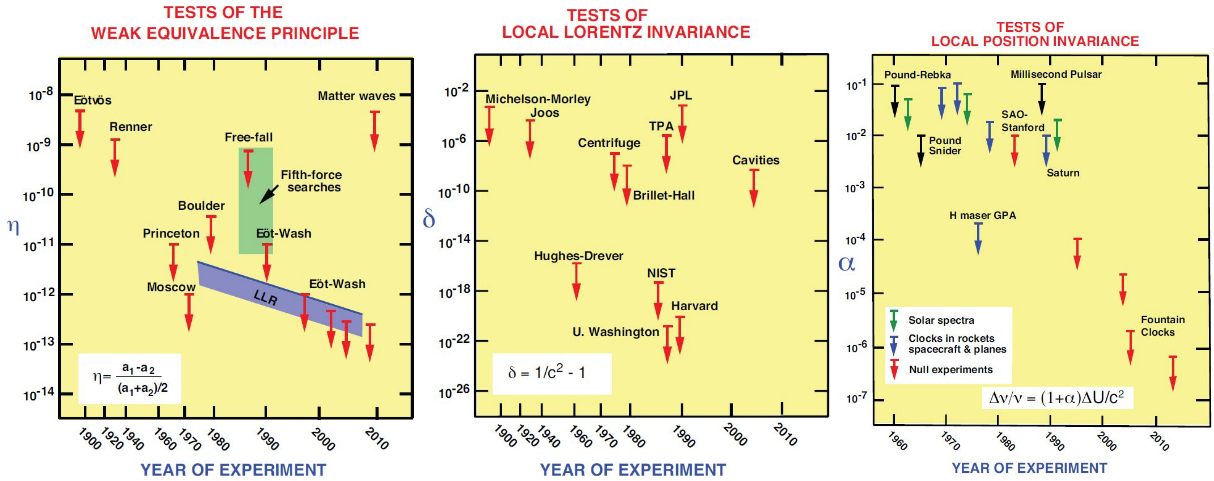

Figure 5.1: Experimental tests of GR, from Will (2014).

- Newton pendulum, $\eta \sim 10^{-2}$.
- Eötvos torsion balance (1885-1909), $\eta \sim 10^{-9}$.
- MICROSCOPE (2018), $\eta \sim 10^{-15}$.

Experimental tests of LLI. Example of tests of LLI are measurment of light speed, for example,
1. Michelson-Morley experiment (1881-1887).
2. FERMI (2009, High-energy astrophysical photons [GeV gamma-ray burst]) [1]

Note these are tests of SR principles and/or quantum gravity. In particular, some quantum gravity theories predict that there exists a fundamental scale

$$E_{\text{Planck}} = \sqrt{\frac{\hbar c^5}{\hbar}} \sim 10^{19} GeV \tag{5.4}$$

at which Lorentz invariance could be violated.

Example 5.6.1. Lorentz violting dispersion relation. *According to SR $E^2 = p^2c^2$ and the speed of the photon is the light speed $v_\gamma = \partial E/\partial p = c$. One couldalternatively postulate phenomenological dispersion relations starting from*

$$E^2 = p^2c^2(1 + corrections) = p^2c^2 + E_{Planck} f_1 |p| c + f_3 E^{-1}_{Planck} f_3 |p|^3 c^3 + \dots , \tag{5.5}$$

such that

$$\frac{v_\gamma}{c} = 1 + corrections . \tag{5.6}$$

The corrections to the photon speed could be constrained by measurements at energy sufficiently high energies (approaching E_{Planck}). For example, phothons of different energies would arrive at different times.

Experimental tests of LLP. Example of tests of LLI are measurements of doppler effect due to the gravitational field:
1. Pound-Rebka experiment (1959), measuring the gravitational redshift $z = (1+\alpha)\delta\phi/c$. The experiment can constraint the deviation α.
2. Shift of spectral lines due to the Sun gravitational field.
3. Clocks on satellites.
4. Global Positioning System (GPS), measuring about $\approx 35\mu s = 49\mu s - 7\mu s$ due to difference of the gravitational doppler shift (GR) and time dilation (SR).

Experimental tests of SEP.
- Violation of WEP for gravitating bodies inducing orbit perturbations.
- Location and frame-dependent effects in the measurement of G.

Example 5.6.2. Nordvedt effect. *The acceleration of a body of grav. mass m_g in an external grav. field could be phenomenologically parametrized as*

$$\ddot{\vec{x}} = -\frac{m_g}{m_i}\nabla\phi = (1 - \eta_N \frac{E_g}{m_i})\nabla\phi , \tag{5.7}$$

where $-E_g < 0$ is the gravitational self-energy of the body and η_N is a parameter. For laboratory experiments o Earth $E_g/m_i \lesssim 10^{-27}$ so there is no effect (e.g. it is negligible in Eövtos experiments). However, for celestial bodies could be masurable since

$$\frac{E_g}{m_i} \sim 10^{-6}\ Sun\ , \quad \frac{E_g}{m_i} \sim 10^{-8}\ Jupiter\ , \quad \frac{E_g}{m_i} \sim 10^{-10}\ Earth, Moon\ . \tag{5.8}$$

[1] https://arxiv.org/abs/0908.1832.

If $\eta_N \neq 0$, then one would for example measure Earth falling towards the Sun with a different acceleration then the Moon. Lunar laser ranging experiments started with the mission Apollo 11 indicate that $\eta_N < 10^{-12}$.

5.7 Consequences of EPs

Let us finally consider the consequences of the above EPs for the development of theories of gravity. We assume that our theory describes "gravity as geometry" and try to imagine how to translate the EPs into geometrical requirements.

WEP. $\Rightarrow$ worldlines of test-bodies depend only on the gravitational field, not on their properties. Test-bodies moves on geodesics: if there is no gravitational field, the geodesics are straight lines like in Newtonian gravity.

EEP. Locally, the physical laws are those of SR. $\Rightarrow$ The spacetime must be locally Minkowski

$$g \sim \eta \ \text{ and } \ \Gamma = 0 \ , \tag{5.9}$$

that implies that the g must be a Lorenzian metric. Because the solutions of the theory for nongravitational phenomena must be locally the same as SR, the equations must be in tensorial form.

In particular LLI and LPI suggest that non-gravitational particles and fields all couple in the same to the gravitational field. This universal coupling hints to gravity as a property of the spacetime rather than generate by a field on it. Gravity must be a metric theory.

In metric theories of gravity only the metric couple to matter and determines the matter motion. However, there could be other fields (scalar, vectors, etc fields) with the role of determining how matter couple to metric and generate gravity. Metric theories of gravity differ in the way additional gravitational field are introduces. Generically they divide into(Will, 2014). (i) *purely dynamical* theories: metric and additional gravity fields are determined dynamically by the field equations; (ii) *prior geometric*: fields or other elements are given a priori.

SEP. $\Rightarrow$ There exists only one gravitational field represented by the metric g. A gravity theory satisfying SEP is a pure metric theory. Any nonpure metric theory predicts that the mass-energy of self-gravitating objects acquires a dependence on the extra gravitational fields. They would produce a force that would make the motion of test bodies nongeodesic.

6. Einstein's Field Equations

③

These lectures introduce and discuss dynamics of fields on generic manifolds and Einstein's equations for spacetime. The Hilbert action formulation of GR and the Cauchy problem in GR are presented. The concept of Killing vector and Lie derivative are introduced here together with their relation to symmetries and dynamics in GR.

Suggested readings. *Chap. 4 of Wald (1984); Chap. 4 of Carroll (1997); Chap. 7-8 of Schutz (1985).*

6.1 GR Postulates

GR postulates [Einstein (1915)]

(i) Gravity is a pure metric geometric theory.
(ii) Spacetime is a 4D manifold $\mathcal{M}$ equipped with a Lorentzian metric g and Levi-Civita connection ∇.
(iii) In local Lorentz frames the non-gravitational laws of physics are those of SR.
(iv) Test-bodies follow geodesics in $\mathcal{M}$ and the equations for matter fields are tensorial equations.
(v) The metric tensor is determined by Einstein's field equations.

Note the postulates could be reduced and made more coincise.

The EOM for particles and fields can be generalized from SR to GR by the formal identification:

$$\text{(SR)} \left\{ \begin{array}{lcl} \mathbb{R}^4 \text{ spacetime} & \longleftrightarrow & \mathcal{M} \text{ spacetime} \\ \eta \text{ metric} & \longleftrightarrow & g \text{ metric} \\ \partial \text{ connection} & \longleftrightarrow & \nabla \text{ connection} \end{array} \right\} \text{(GR)} \tag{6.1}$$

The above rule actually works in most of the cases, although it should be used with some care.

Einstein field equations (EFE) equations will be euristically derived as an extension of Poisson equation for the Newtonian gravitational potential. The equations in vaccum (no matter) will be alternatively derived from the Hilbert action that is the simplest action that (i) is diffeomorphism invariant and (ii) leads to second order tensorial equations for the metric. The EFEs were indeed derived before Einstein by Lorentz and Hilbert using this approach. This also give a natural/generic way to define the stress-energy tensor. However, the treatment of boundary terms in the action variation is nontrivial and requires discussion.

6.2 Equation of motions (EOMs) for particles

Using the scheme in Eq. (6.1) the EOM for a free particle with worldline $x^\mu(\lambda)$ and 4-velocity $u^\mu = \dot{x}^\mu$ in GR is:

$$\text{(SR)} \ \frac{d^2x^\mu}{d\lambda^2} = 0 \quad \longleftrightarrow \quad \frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\sigma\nu}\frac{dx^\sigma}{d\lambda}\frac{dx^\nu}{d\lambda} = 0 \ \text{ or } \ u^\mu\nabla_\mu u^\nu = 0 \ \text{(GR)} \ . \tag{6.2}$$

In presence of an external force f^μ the acceleration particle is nonzero and the EOM is

$$u^\mu\nabla_\mu u^\nu = \frac{f^\nu}{m} \ . \tag{6.3}$$

The 4-momentum of the particle is $p^\mu = mu^\mu$.

Remark 6.2.1. *The* energy *of a particle measured by an observer $\mathcal{O}$ defined by its velocity v^a has, formally, the same expression as in SR:*

$$E = -p_a v^a \ , \tag{6.4}$$

and it is given by the projection of the particle's 4-momentum to the observer worldline with the metric g. However, there is a key difference. In SR: E is the energy measured at a given point p but also the energy measured by any other distant inertial observer with the same 4-velocity v^a because the vectors can be parallel transported anywhere in a path-independent way. In GR: E is only the local *energy measured at point p by $\mathcal{O}$. A distant observer cannot define the energy at point p. There is no global family of inertial observers.*

EOM of a particle in a weak and static grav.field. Check the Newtonian limit of geodesics in a curve spacetime. Problem: relate the Newton equation of a particle in a gravitational potential,

$$\frac{d^2x^i}{dt^2} = -\partial_i\phi \ , \tag{6.5}$$

to the geodesic equation under the hypothesis
(i) Small velocities of the particle, $v/c \ll 1$;
(ii) Weak gravitational field $\phi \ll 1$;
(iii) Static gravitational fields (ϕ and g are time independent).

(i) Small velocity. Use proper time τ to parametrize the worldline, $\dot{x}^\mu = (dt/d\tau, c^{-1}dx^i/d\tau)$. For small velocities the Lorentz factor is about one and the spatial part of the 4-velocity is much smaller than the 0-component:

$$\frac{dt}{d\tau} = \gamma \sim 1 \ \text{ and } \ \frac{1}{c}\frac{dx^i}{d\tau} \ll 1 \quad \Rightarrow \quad \frac{dt}{d\tau} \gg \frac{1}{c}\frac{dx^i}{d\tau} \ . \tag{6.6}$$

The geodesic equation for small velocity is:

$$0 = \frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{00}\frac{dx^0}{d\tau}\frac{dx^0}{d\tau} + 2\Gamma^\mu_{0j}\underbrace{\frac{dx^0}{d\tau}\frac{dx^j}{d\tau}}_{\mathcal{O}(1/c)} + \Gamma^\mu_{ij}\underbrace{\frac{dx^i}{d\tau}\frac{dx^j}{d\tau}}_{\mathcal{O}(1/c^2)} \approx \frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{00}\left(\frac{dx^0}{d\tau}\right)^2 \ . \tag{6.7}$$

(ii) Weak field. Take $g = \eta + h$ where η is the Mikowski metric and h a small perturbation. Problem: g is not positive defined, How can one quantify "small"? One must assume there exists a global coordinate system that correspond to the Cartesian coordinate for $h = 0$ (SR). Then, the requirement is that the components of the tensor h in these coordinates are much smaller than those of η:

$$|h_{\mu\nu}| \ll 1 \ . \tag{6.8}$$

This way it is possible to proceed linearizing in $h_{\mu\nu}$. Indexes are raised and lowered with the Mikowski metric η.

To calculate the geodesic equation in the small velocity limit Eq. (6.7) one needs only

$$\Gamma^\mu_{00} = \frac{1}{2}g^{\mu\alpha}(\underbrace{\partial_0 g_{0\alpha}}_{=0} + \underbrace{\partial_0 g_{0\alpha}}_{=0} - \partial_\alpha g_{00}) = -\frac{1}{2}g^{\mu\alpha}\partial_\alpha g_{00} = \tag{6.9a}$$

$$= -\frac{1}{2}(\eta^{\mu\alpha} + h^{\mu\nu})\partial_\alpha(\underbrace{\eta_{00}}_{=-1} + h_{00}) = -\frac{1}{2}(\eta^{\mu\alpha} + \mathcal{O}(h^{\mu\nu}))\partial_\alpha h_{00} \approx -\frac{1}{2}\eta^{\mu\alpha}\partial_\alpha h_{00} \ . \tag{6.9b}$$

where the first line holds because the metric is time independent (iii), and in the second line the weak field limit is taken (ii). Moreover, the component $\Gamma^0_{00} = 0$ because of (iii): $\partial_0 h_{00} = 0$. The first component of the geodesic equation is thus trivially solved and gives the coordinate time in terms of a linear combination of the proper time; the remaining equations are the spatial equations:

$$0 = \frac{d^2x^0}{d\tau^2} = \frac{d^2t}{d\tau^2} \quad \Rightarrow \quad t = \alpha\tau + \beta \quad \alpha, \beta \in \mathbb{R} \tag{6.10a}$$

$$0 = \frac{d^2x^i}{d\tau^2} - \frac{1}{2}\eta^{ij}\partial_j h_{00}\left(\frac{dx^0}{d\tau^2}\right)^2 \quad \Rightarrow \quad 0 = \frac{d^2x^i}{dt^2} - \frac{1}{2}\delta^{ij}\partial_j h_{00} \ . \tag{6.10b}$$

Direct comparison with Eq. (6.5) identifies the metric component with the Newtonian grav.potential:

$$h_{00} = -2\phi \ . \tag{6.11}$$

6.3 EOMs for fields

Scalar field. The equation

$$\Box_\eta \varphi - m^2\varphi = 0 \ , \tag{6.12}$$

can be easily generalized for Mikowski to generic spacetimes by the substitution

$$\text{(SR) } \Box_\eta = \eta^{ab}\partial_a\partial_b \quad \longleftrightarrow \quad \Box_g = g^{ab}\nabla_a\nabla_b \text{ (GR)} \ . \tag{6.13}$$

Note the generalization is not unique and there are more possibility, e.g. by coupling the field to the curvature by the Ricci scalar

$$\Box_g\varphi - m^2\varphi - \alpha R\varphi = 0 \ . \tag{6.14}$$

There is no general rule, both equations with and without are valid models for scalar fields on $\mathcal{M}$ that have the correct SR limit. Often, a "minimal coupling" principle is invoked: matter fields do not couple to the Riemann tensor (curvature) but only to the metric (pure metric theory).

Electromagnetic field. Maxwell equations in GR are written in terms of the Faraday tensor

$$\nabla_a F^{ba} = J^b \ , \quad \nabla_{[a} F_{bc]} = 0 \ . \tag{6.15}$$

The second equation is the antisymmetric combination of covariant derivatives, and it implies that the Farady tensor can be written in terms of the potential (Note the Christoffel symbols cancel in the antitsymm. combination):

$$F_{ab} = \nabla_{[a} A_{b]} = \nabla_a A_b - \nabla_b A_a = \partial_a A_b - \partial_b A_a \ . \tag{6.16}$$

It is interesting to study the EOM for the potential. In Lorentz gauge one would guess:

$$(\mathrm{SR}) \begin{cases} \partial_a A^a & = 0 \\ \Box_\eta A^a & = -J^a \end{cases} \quad \underset{???}{\longleftrightarrow} \quad (\mathrm{GR}) \begin{cases} \nabla_a A^a & = 0 \\ \Box_g A^a & = -J^a \ . \end{cases} \tag{6.17}$$

Verify the EOM is correct by substituting into the first Maxwell equation:

$$-J_b = \nabla^a F_{ab} = \nabla^a(\nabla_a A_b - \nabla_b A_a) = \nabla^a \nabla_a A_b - \nabla_a \nabla_b A_a = \Box_g A_b - \nabla^a \nabla_b A_a = \Box_g A_b - \nabla_a \nabla_b A^a = \tag{6.18a}$$

$$= \Box_g A_b - \nabla_b \nabla_a A^a - R^c{}_b A_c \ , \tag{6.18b}$$

where in the last line the commutator and Riemann tensor have been used. Imposing the Lorentz gauge $\nabla_a A^a = 0$ does **not** lead to the GR equation guessed above! Indeed, the right equation is the one derived above,

$$(\mathrm{SR}) \begin{cases} \partial_a A^a & = 0 \\ \Box_\eta A^a & = -J^a \end{cases} \quad \longleftrightarrow \quad (\mathrm{GR}) \begin{cases} \nabla_a A^a & = 0 \\ \Box_g A^a - R^c_b A^b & = -J^a \ , \end{cases} \tag{6.19}$$

because the latter is the one that guaranteed charge conservation in the form of a divergence on generic spacetime:

$$\nabla_b J^b = 0 \ . \tag{6.20}$$

In the following, 3 different proofs of the above equation are given.

1. Using the definition of the Riemann and the antisymmetry of the Faraday tensor,

$$[\nabla_a, \nabla_b] F^{ab} = \nabla_a \nabla_b F^{ab} - \nabla_b \nabla_a F^{ab} = \nabla_a \nabla_b F^{ab} + \nabla_b \nabla_a F^{ba} = 2\nabla_a \nabla_b F^{ab} \tag{6.21a}$$

$$= -R_{abc}{}^a F^{cb} - R_{abc}{}^b F^{ac} = +R_{bac}{}^a F^{cb} - R_{abc}{}^b F^{ac} = +R_{bc} F^{cb} - R_{ac} F^{ac} = \tag{6.21b}$$

$$= 2R_{bc} F^{cb} = -2R_{cb} F^{bc} = 0 \tag{6.21c}$$

   the scalar made of the contraction between the antisymmetric double covariant derivative is twice the scalar made of the double covariant derivative contraction (1st line) and it is zero (2nd and 3rd line). This holds for any antisymmetric tensor.

2. For any vector and for any antisymmetric tensor the covariant derivative can be expressed in terms of partial derivatives as [exercise]

$$\nabla_\mu J^\mu = \frac{1}{\sqrt{|g|}} \partial_\mu \left(\sqrt{|g|} J^\mu \right) \ , \quad \nabla_\nu F^{\mu\nu} = \frac{1}{\sqrt{|g|}} \partial_\nu \left(\sqrt{|g|} F^{\mu\nu} \right) \ , \tag{6.22}$$

   where $|g| = -\det g$. The Maxwell equation and the current conservation are thus

$$\frac{1}{\sqrt{|g|}} \partial_\nu \left(\sqrt{|g|} F^{\mu\nu} \right) = -J^\mu \ , \quad \frac{1}{\sqrt{|g|}} \partial_\mu \left(\sqrt{|g|} J^\mu \right) = 0 \ . \tag{6.23}$$

   Multiplying the first equation by $|g|$ and deriving gives the second equation since partial derivatives commute:

$$-\partial_\mu (\sqrt{|g|} J^\mu) = \underbrace{\partial_\mu \partial_\nu}_{\text{sym}} (\underbrace{\sqrt{|g|} F^{\mu\nu}}_{\text{antisym}}) = 0 \ . \tag{6.24}$$

3. Write the Maxwell equations in terms of the exterior derivatives of the 2-form F

$$\mathbf{d}(*F) = (*J) \ , \quad \mathbf{d}F = 0 \ , \tag{6.25}$$

   and note that $\mathbf{d}^2 = 0 \Rightarrow \mathbf{d}(*J) = 0$. The latter equation corresponds to the divergence of $J^a = g^{ab} J_b$ (Example 3.12.1 and discussion around).

Remark 6.3.1. Key properties of symmetric/antisymmetric tensors. *Given the* $(0,2)$ *tensors* T_{ab} *(generic),* $A_{[ab]}$ *(antisymm.),* S_{ab} *(symm.), the proofs above show that in general:*

$$S^{ab} A_{ab} = 0 \ , \quad [\nabla_a, \nabla_b] T^{ab} = R_{ab}(T^{ab} - T^{ba}) = R_{ab} 2T^{[ab]} = 0 \ , \quad 0 = [\nabla_a, \nabla_b] A^{ab} = 2\nabla_a \nabla_b A^{ab} \ , \tag{6.26}$$

where R_{ab} *is the Ricci tensor (symm.).*

Conservation law for the stress-energy tensor. The definition of stress-energy tensor T_{ab} in terms of the energy and momentum densities of a continuum distribution of matter as measured by an observer $\mathcal{O}$ defined by its tangent vector v carries over to GR.

In SR the requirement of energy and momentum conservation leads to trequire that the 4-divergence of the stress-energy momentum is zero,

$$\partial_\mu T^{\mu\nu} = 0 \ \ (\text{SR}) \ . \tag{6.27}$$

The above results can be understood as follows. Consider a family of inertial observers with 4-velocities v^α ($v^\mu v_\mu = -1$) such that the velocities are all parallel, $\partial_\mu v^\nu = 0$. The energy current density measured by the observers is

$$J_\mu := -T_{\mu\nu} v^\nu = (E, \varphi_i) \ , \tag{6.28}$$

and the EOM above implies that the 4-divergence of the current it zero

$$\partial_\mu J^\mu = -\partial_\mu (T_{\mu\nu} v^\nu) = -\underbrace{\partial_\mu T_{\mu\nu}}_{=0} v^\nu - T_{\mu\nu} \underbrace{\partial_\mu v^\nu}_{=0} = 0 \ . \tag{6.29}$$

Integrating over a volume Σ and using Gauss's theorem gives the energy conservation,

$$0 = \int_\Sigma \partial_\mu J^\mu = -\int_\Sigma \partial_t (J^0 - \partial_i J^i) = -\frac{d}{dt} \int_\Sigma E + \int_{\partial\Sigma} \varphi^i n_i \ . \tag{6.30}$$

In GR, one would generalize the EOM as

$$\nabla_a T^{ab} = 0 \ \ (\text{GR}) \ , \tag{6.31}$$

and from the definition of stress-energy tensor it is still true that observers with tangent velocity field v^a measure the energy-momentum density at point p

$$J_a := -T_{ab} v^b \ . \tag{6.32}$$

The difference w.r.t. SR is that in GR $\nabla_a T^{ab} = 0$ cannot be interpreted as a conservation law because in general

$$\nabla^a J_a = -\underbrace{\nabla^a T_{ab}}_{=0} v^b - T_{ab} \nabla_a v^b = -T_{ab} \nabla_a v^b \neq 0 \ . \tag{6.33}$$

Energy conservation would be guaranteed for observers such that $\nabla_a v^b = 0$ [1], but in general there are no such global inertial observers (See however Sec. 6.7 below). However, on sufficiently small regions where curvature can be neglected, $R \ll (curvature)^{-1}$, one can still find observers such that $\nabla_a v^b \approx 0$ and the equation above can be interpested as a local conservation law.

Example 6.3.1. *For a* perfect fluid $T_{ab} = (\rho c^2 + p) u_a u_b + p g_{ab}$ *the SR's EOMs are the conservation of mass and momentum of SR hydrodynamics. Moreover, for* dust *(perfect fluid with zero pressure, $p = 0$) the GR EOMs,*

$$0 = \nabla_a T^{ab} = \nabla_a (\rho u^a) u^b + \rho u^a \nabla_a u^b \ . \tag{6.34}$$

imply the geodesic equation. The contraction with u_b gives the continuity equation $\nabla_a(\rho u^a) = 0$:

$$0 = \nabla_a (\rho u^a) \underbrace{u_b u^b}_{-1} + \rho u^a u_b \nabla_a u^b = -\nabla_a (\rho u^a) + \rho u^a \frac{1}{2} \nabla_a (\underbrace{u^b u_b}_{=-1}) = -\nabla_a (\rho u^a) + 0 \ . \tag{6.35}$$

Plugging the continuity equation into the EOMs gives $u^a \nabla_a u^b = 0$, that proves dust moves on geodesics.

6.4 Einstein's field equations (EFE)

The dynamics of the metric must be described by tensorial equations with the correct Newton limit. Let us give an euristic derivation of Einstein's equation starting from these two hypotesis. Newton EOMs relate second derivatives (Laplacian) of the grav.potential to the matter's mass density distribution ρ:

$$\Delta \phi = \delta^{ij} \nabla_j \nabla_i \phi = 4\pi G \rho \ . \tag{6.36}$$

Assuming a weak field metric, we have shown that $g_{00} = -1 + 2\phi$, and we know the matter is described by a symmetric stress-energy tensor. We thus postulate tensorial equations in the form:

$$\text{“} \ \partial^2 g_{ab} \ \text{”} = \kappa G T_{ab} \ , \tag{6.37}$$

clearly, on the l.h.s. one needs a symmetric tensor built on derivative of the metric, while on the r.h.s. the stress-energy tensor must satisfies $\nabla^a T_{ab} = 0$.

Options for the l.h.s.

[1] An equivalent condition is the symmetrized derivative: $T^{ab} \nabla_{(a} v_{b)} = T^{ab} \nabla_{(a} v_{b)} = (T^{ab} \nabla_a v_b + T^{ab} \nabla_b v_a)/2 = T^{ab} \nabla_a v_b$.

(i) $\Box_g g_{ab} = g^{cd}\nabla_c\nabla_d g_{ab} \equiv 0$, not an option :(

(ii) $R_{ab}[g]$, Ricci tensor contains second derivatives, but Bianchi identities are incompatible with the EOM for the matter:

$$0 = \kappa G \nabla^a T_{ab} = \nabla^a R_{ab} = \frac{1}{2}\nabla_b R \neq 0 \ . \tag{6.38}$$

Assuming $\nabla_b R \equiv 0$ would not help either since that would imply the trace of the stress-energy identically constant $T \equiv const$:

$$R = g^{ab}R_{ab} = \kappa G g^{ab}T_{ab} = \kappa G T \quad \Rightarrow \quad 0 = \nabla_b R = \kappa G \nabla_a T \ . \tag{6.39}$$

The latter is not possible since the spacetime of an isolated star has regions where $T \neq 0$ (star's interior) and vacuum regions $T = 0$.

(iii) The Einstein tensor is the right choice for all the cases since Bianchi identities implies the EOM for the matter !

$$G_{ab}[g] = R_{ab}[g] - \frac{1}{2}R[g]g_{ab} = \kappa G T_{ab}[g] \ . \tag{6.40}$$

Note that an equivalent expression (called trace-reverse) can be found by taking trace

$$g^{ab}(R_{ab} - \frac{1}{2}R g_{ab}) = R - \frac{1}{2}R \underbrace{g_{ab}g^{ab}}_{=\mathrm{Tr}(gg^{-1})=4} = \kappa G g^{ab}T_{ab} = \kappa G T \quad \Rightarrow \quad -R = \kappa G T \ , \tag{6.41}$$

and re-inserting into the equation

$$R_{ab}[g] = \kappa G(T_{ab}[g] - \frac{1}{2}T[g]g_{ab}) \ . \tag{6.42}$$

This also shows that the Ricci tensor $R_{ab} = 0$ are the EFE in vacuum.

Weak field limit and determination of κ. Take the static and weak field limit of the trace rever EFE. Start from the metric

$$g = -(1+2\phi)\mathrm{d}t^2 + (1-2\phi)(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) \ , \tag{6.43}$$

and focus on the 00-component:

$$g_{00} = -1 + h_{00} \ , \quad g^{00} = -1 - h_{00} \ . \tag{6.44}$$

The l.h.s. is the Ricci tensor:

$$R_{00} = R^{\mu}_{0\mu 0} = R^{i}_{0i0} = \partial_j \Gamma^{j}_{00} = \partial_j \left[\frac{1}{2} \underbrace{g^{i\lambda}}_{=\eta^{i\lambda}} (\underbrace{\partial_0 g_{\lambda 0}}_{=0} + \underbrace{\partial_0 g_{0\lambda}}_{=0} - \partial_\lambda g_{00})) \right] = -\frac{1}{2}\delta^{ij}\partial_i\partial_j h_{00} = -\frac{1}{2}\Delta h_{00} \ . \tag{6.45}$$

The r.h.s. is $T_{00} - \frac{1}{2}g_{00}T$, where

$$T_{00} = E \simeq \rho \ , \quad \text{(only the small velocity contribution)} \ . \tag{6.46}$$

and

$$g_{00}T = (\eta_{00} + h_{00})(g^{\mu\nu}T_{\mu\nu}) \simeq \underbrace{\eta_{00}}_{=-1}(\underbrace{\eta_{00}}_{=-1}T_{00} + \underbrace{\eta^{ij}}_{=\delta^{ij}}T_{ij}) = +T_{00} - \delta^{ij}T_{ij} \simeq \rho \quad (|T_{00}| \gg |T_{ij}|) \ . \tag{6.47}$$

Putting together things and comparing with Newton law:

$$-\frac{\cancel{1}}{\cancel{2}}\Delta h_{00} = \frac{\cancel{1}}{\cancel{2}}\kappa G\rho \quad \Rightarrow \quad \kappa \equiv 8\pi \ . \tag{6.48}$$

The l.h.s. of EFEs has dimension of $[G_{ab}] = L^{-2}$, the r.h.s. has dimension of energy density $[T_{ab}] = ML^{-1}T^{-2}$ multiplied by G. Correct dimensions are re-established by introducing the proper c factors: $G\kappa = 8\pi G c^{-4}$.

Short discussion on EFE structure. When written in some coordinate system EFEs

$$G_{ab}[g] = R_{ab}[g] - \frac{1}{2}R[g]g_{ab} = \frac{8\pi G}{c^4}T_{ab}[g] \ , \tag{6.49}$$

are a system of 10 coupled nonlinear PDEs for the 10 metric components involving 2nd and 1st derivatives of the metric. As discussed in detail in Sec. 6.6, these are not exactly "10 equations for 10 unknowns". A coordinate transformation can always be choosen so to fix 4 out of 10 metric components, thus there are only 6 metric components to be determined by EFEs. The latter are however not an overdetermined system of equations because the Einstein (Ricci) tensors must satisfy the 4 Bianchi identities.

As explicitely indicated in the above EFEs, the stress-energy tensor depend in general on the metric. As a consequence, it is not possible to specify the matter distribution and dynamics and then calculate the metric (Note the latter is what one would do with currents in electromagnetisms and with the mass distribution for the Newtonian grav. potential). Even more interestingly, EFE contain the EOM for the matter as they imply the local conservation for the stress-energy tensor. Matter fields EOM can be coupled to EFE by specifying appropriate fields in the action from which a T_{ab} can be obtained (see below). A key result is also that EFE contain also the geodesic hypothesis. It can be proven that $\nabla^a T_{ab} = 0$ implies

$$\nabla_a T^{ab} = 0 \quad \Rightarrow \quad u^a \nabla_a u^b = 0 \ , \tag{6.50}$$

for any body with sufficiently weak self-gravity (See Example 6.3.1 for the special case of a perfect fluid).

The best summary about the interpretation of EFE is the famous quote from J.A.Wheeler:

Summary 6.4.1. *Spacetime tells matter how to move; matter tells spacetime how to curve.*

6.5 Hilbert action and Lagrangian formulation

EFE can be derived from an action principle. The latter has some advantages
(i) The action is easier to euristically derive and postulate since it is a scalar;
(ii) The EOM automatically satisfies the symmetries of the theory (diffeomorphism invariant, see Remark 3.4.1 and 6.8.1);
(iii) It leads to a general definition of the stress-energy tensor.

However, the action/Lagrangian formulation of GR has some complications/subtle points w.r.t. the "standard scheme" (Cf. Sec. 2.8):
(a) The integral defining the action is an integral on $\mathcal{M}$ thus the measure contains the metric (field to vary). In turn, the Lagrangian density cannot be a scalar but it is a *tensor density* [2];
(b) Metric compatibility requires $\nabla g = 0$, but the EOM must be 2nd order in g. This implies that we cannot vary with respect to the field derivatives and the Lagrangian must contain second derivatives (rather than first derivatives) of the field;
(c) Also related to the above, the boundary term after the integration by part of the divergence contain ∇g. In general this does not vanish and one needs to treat it with some care (we willdiscuss two ways).

Let us start addressing (a): How does the formalism of Sec. 2.8 translate for fields on manifolds? The action is given by

$$S = \int_{\mathcal{M}} \mathcal{L} = \int_{\mathcal{M}} \mathcal{L}\varepsilon = \int_{\mathcal{M}} \mathcal{L}\sqrt{|g|}\mathrm{d}x^1 \wedge ... \wedge \mathrm{d}x^4 \ . \tag{6.51}$$

There are several possible ways to proceed. One way is to re-define the Lagrangian as a 4-form

$$\mathcal{L} \ \mapsto \ \mathcal{L} = \hat{\mathcal{L}}\varepsilon \ , \tag{6.52}$$

but that complicates the functional derivatives w.r.t. to the fields. Another way is to "helding fixed" the volume element and but write the EOM in terms of the Lagrangian scalar density $\hat{\mathcal{L}}$. Obviously this work for any field ϕ than the metric and leads to the Euler EOM in GR:

$$\frac{\partial \hat{\mathcal{L}}}{\partial \phi} - \nabla_\mu \left(\frac{\partial \hat{\mathcal{L}}}{\partial(\nabla_\mu \phi)} \right) = 0 \ . \tag{6.53}$$

For the metric, the simplest way to proceed is to perform explicitely the variation of the Lagrangian scalar density in components, which is done in the following.

To deal with (b) we decide to vary only w.r.t. to the metric (Hilbert action). One then further observes that $\hat{\mathcal{L}}$ must be a scalar built out of second derivatives of the metric. The Riemann tensor in $n = 4$ has 20 components precisely proportional to the metric's second derivatives. In a local inertial frame one can further perform Lorentz transformations so to eliminate 6 of the Riemann components. The Lagrangian can be constructed with the *curvature invariants* constructed from the contractions of the Riemann; there are $20 - 6 = 14$ possible curvature invariants, but it turns out that only one of those is linear in the second derivatives: the Ricci scalar. For vacuum EFE we thus postulate the

$$S_H[g] = \int \sqrt{|g|} R d^4x \quad \text{(Hilbert action)} \ . \tag{6.54}$$

Variations of the Hilbert action. The variation is better performed in the inverse metric. Is it possible? Yes, stationary points of the action w.r.t. g are stationary points of the action w.r.t. g^{-1}. The two variations are related by

$$0 = \delta(\delta^\mu_\nu) = \delta(g^{\mu\alpha} g_{\nu\alpha}) = \delta g^{\mu\alpha} g_{\nu\alpha} + g^{\mu\alpha}\delta g_{\nu\alpha} \ , \quad \text{multiply by } g_{\mu\rho} \text{ and contract:} \tag{6.55a}$$

$$0 = g_{\mu\rho} g_{\nu\alpha} \delta g^{\mu\alpha} + \underbrace{g_{\mu\rho} g^{\mu\alpha}}_{\delta^\alpha_\rho} \delta g_{\nu\alpha} \quad \Rightarrow \quad \delta g_{\nu\rho} = -g_{\mu\rho} g_{\nu\alpha} \delta g^{\mu\alpha} \ . \tag{6.55b}$$

[2] Tensor densities or pseudotensors of weight w are quantities that under coordinate transformation transform as

$$\tau^{\mu_1 \dots \mu_k}_{\nu_1 \dots \nu_l} = (\det \frac{\partial x^{\alpha'}}{\partial x^\alpha})^w \frac{\partial x^{\mu_1}}{\partial x^{\mu'_1}} \dots \frac{\partial x^{\nu'_l}}{\partial x^{\nu_l}} \tau^{\mu'_1 \dots \mu'_k}_{\nu'_1 \dots \nu'_l} \ .$$

Perform the action variation by vary separately the three terms:

$$0 = \delta S_H = \int \delta(\sqrt{|g|}R) = \int \delta(\sqrt{|g|}g^{\mu\nu}R_{\mu\nu}) \tag{6.56a}$$

$$= \underbrace{\int \sqrt{|g|}R_{\mu\nu}\delta g^{\mu\nu}}_{\text{I.}} + \underbrace{\int \delta(\sqrt{|g|})R}_{\text{II.}} + \underbrace{\int \sqrt{|g|}g^{\mu\nu}\delta R_{\mu\nu}}_{\text{III.}} \tag{6.56b}$$

Term I. is already Ok. Term II makes use of the algebraic identity valid for any symmetric matrix [3]

$$\ln(\det A) = \mathrm{Tr}(\ln A) \quad \Rightarrow \quad \frac{1}{\det A}\delta(\det A) = \mathrm{Tr}(A^{-1}\delta A) \ . \tag{6.57}$$

The above equation in terms of the metric determinant is (epxressing the variation of the metric with the variation of its inverse):

$$\delta \det g = \det g \cdot g^{\mu\nu}\delta g_{\mu\nu} = \det g \cdot g^{\mu\nu}(-g_{\mu\alpha}g_{\nu\beta}\delta g^{\alpha\beta}) = -\det g \cdot \delta^\nu_\alpha \cdot g_{\nu\beta}\delta g^{\alpha\beta} = -\det g \cdot g_{\alpha\beta}\delta g^{\alpha\beta} \ . \tag{6.58}$$

thus

$$\delta\sqrt{|g|} = \delta\sqrt{-\det g} = \frac{1}{2}\frac{1}{\sqrt{|g|}}(-\delta \det g) = -\frac{1}{2}\frac{-\det g}{\sqrt{-\det g}}g_{\alpha\beta}\delta g^{\alpha\beta} = -\frac{1}{2}\sqrt{|g|}g_{\alpha\beta}\delta g^{\alpha\beta} \ . \tag{6.59}$$

Thus,

$$\text{II.} = -\frac{1}{2}\int \sqrt{|g|}Rg_{\alpha\beta}\delta g^{\alpha\beta}. \tag{6.60}$$

The calculation of term III. requires to compute first the variation of the Christoffel symbols and then substitute the variation of the inverse metric; it is lenghty but straightforward and the result is (Carroll, 1997; Wald, 1984):

$$\delta K = g^{\mu\nu}\delta R_{\mu\nu} = \nabla^\alpha \left(\underbrace{\nabla^\beta(\delta g_{\alpha\beta}) - g^{\rho\sigma}\nabla_\alpha(\delta g_{\sigma\rho})}_{=:v_\alpha}\right) \ . \tag{6.61}$$

The key point is that the variation of the Ricci can be expressed in terms of a divergence of a vector. When inserted into the integral, the divergence is a total derivatives and thus produces a boundary term via Stokes theorem

$$\int_M \nabla_\alpha v^a = \int_{\partial\mathcal{M}} v^\alpha n_\alpha \ . \tag{6.62}$$

The latter term is nontrivial: the vector v^α contains covariant derivatives of the varied metric, thus it is not zero (we are varying only the metric, and we are allowed to set to zero only the variation of the fields we are varying). The term has an interesting interpretation: it represent the variation of the extrinsic curvature of the boundary $\partial\mathcal{M}$. Note that putting together the three terms,

$$\delta S_H = \int \sqrt{|g|}\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right)]\delta g^{\mu\nu} - \int \sqrt{|g|}\delta K \ . \tag{6.63}$$

In order to obtain EFE in vacuum for any metric variation we can take

$$\frac{1}{\sqrt{|g|}}\frac{\delta S_H}{\delta g^{\mu\nu}} \ , \tag{6.64}$$

which is the "variation holding the volume element fixed" mentioned at the beginning of the calculation. However, we need to add one more hypothesis; there are two options:

1. Require $\delta K = 0$;
2. Redefine the action: $S_H = \int_\mathcal{M} R + \int_{\partial\mathcal{M}} K$.

Boundary term and Palatini action.

- The proper definition of the boundary term play a relevant role in the Hamiltonian formulation of GR, asymptoticallt flat spacetimes, and the definition of mass-energy of the spacetime.
- An alternative variational approach to GR is to consider the *Palatini action* in which one (i) does not assume the Levi-Civita connection and (ii) varies the connection together with te metric

$$S_P = S_P[g, \nabla] \ \ \text{Palatini action} \ . \tag{6.65}$$

  This is possible because the Ricci tensor can be considered as dependent only on the connection (Cf. expression in terms of Christoffel symbols) and independent on the metric (no explicit presence of the metric). The variation of the Palatini action is not difficult but lengthy Wald (1984). Interestingly, it leads to EFE and the metric compatibility condition for the connection (Levi-Civita) without the need of discussing boundary terms.

[3] A simple way to prove that is to start from $\exp(\mathrm{Tr}(\ln A))$, expressing the matrix A in terms of its diagonal $A = X\Lambda A^{-1}$, use $\ln A = X\ln(\Lambda)X^{-1}$, use the cyclic property of the trace and the fact that the trace of a diagonal matrix is the sum of the eigenvalues. These calculations result in $\exp(\mathrm{Tr}(\ln A)) = \det A$.

Matter terms and Stress-energy tensor. The total action with matter fields ϕ has the form

$$S = \frac{1}{16\pi G} S_H[g] + S_M[g, \phi] \ , \tag{6.66}$$

where S_M is the action for the matter fields. It is immediate from the calculation above the variation of the total action w.r.t. the metric implies EFE for the following definition of the stress-energy tensor:

$$\frac{1}{\sqrt{|g|}} \frac{\delta S}{\delta g^{\mu\nu}} = 0 \quad \Rightarrow \quad \text{EFE} \ , \quad \text{iff} \ \ T_{\mu\nu} := -\frac{2}{\sqrt{|g|}} \frac{\delta S_M}{\delta g^{\mu\nu}} \ . \tag{6.67}$$

This shows how to obtain a symmetric $(0,2)$ stress-energy tensor given an action for the matter fields. For example, for the scalar and electromagnetic fields

$$\mathcal{L} = -\frac{1}{2}\sqrt{|g|}\left(\nabla_a \nabla^a \phi + m^2\phi^2\right) \to \quad T_{ab} = \nabla_a \nabla_b \phi - \frac{1}{2} g_{ab}\left(\nabla_c \nabla^c \phi + m^2 \phi^2\right) \tag{6.68a}$$

$$\mathcal{L} = -\frac{1}{4}\sqrt{|g|} F^{ab} F_{ab} \to \quad T_{ab} = \frac{1}{4\pi}\left(F_{ac}F^c_b - \frac{1}{4} g_{ab} F_{de} F^{de}\right) \ . \tag{6.68b}$$

It is left as [exercise] to compute $\nabla_a T^{ab} = 0$ for the two cases.

6.6 Cauchy problem in GR

The EFE can be written as a 2nd order PDE system of 10 equations by introducing a coordinate system. For example in vacuum one obtains

$$0 = R_{\mu\nu} = -\frac{1}{2} g^{\alpha\beta}\partial_\alpha \partial_\beta g_{\mu\nu} + g^{\alpha\beta}\partial_\alpha \partial_{(\mu} g_{\nu)\beta} - \frac{1}{2} g^{\alpha\beta}\partial_\mu \partial_\nu g_{\alpha\beta} + Q_{\mu\nu}[\partial g, g] \tag{6.69a}$$

$$= -\frac{1}{2} g^{\alpha\beta}\partial_\alpha \partial_\beta g_{\mu\nu} - g_{\alpha(\mu}\partial_{\nu)} H^\alpha + \tilde{Q}_{\mu\nu}[\partial g, g] \tag{6.69b}$$

where $Q_{\mu\nu}$ (and $\tilde{Q}_{\mu\nu}$) represents the *non-principal part* (lower derivatives of $g_{\mu\nu}$), and the second line is re-written introducing the quantity

$$H^\alpha = \partial_\mu g^{\alpha\mu} + \frac{1}{2} g^{\alpha\beta} g^{\rho\sigma} \partial_\beta g_{\rho\sigma} \tag{6.70}$$

for later use. The equations Eq. (6.69a) are not 10 independent equations for the components $g_{\mu\nu}$ of the metric tensor because the 4 Bianchi identities,

$$\nabla_a G^{ab} = 0 \ , \tag{6.71}$$

are further relations between the metric components.

Questions:

(i) What type of equations are EFE?
(ii) How to formulate the initial value/Cauchy problem?
(iii) Is the latter well-posed?

(i) Since the inverse metric is a rational combination of the metric g and its determinant $\det g$ and $Q_{\mu\nu}$ is a rational combination of $g, \partial g$ and $\det g$, the equation $R_{\mu\nu}[g] = 0$ is a *quasilinear* [4] system of 10 coupled 2nd-order PDEs for the metric components. Without further hypothesis EFEs are PDEs of no known type. However, with some physically plausible hypothesis on the spacetime (solutions) it is possibvle to prove well-posedness of the initial value (Cauchy) problem.

To answer (ii) and (iii) we look first at the electromagnetic fields.

Remark 6.6.1. *A PDE problem is well-posed iff exists a unique solution that depends continuously on the boundary data (at least locally in time.)*

Maxwell field equations. We shall see that weak-field (linearized) EFE are analogous to electromagnetism. Let us start recalling the Cauchy or initial value problem (IVP) for Maxwell equations in SR,

$$0 = \partial^\alpha F_{\alpha\beta} = \partial^\alpha(\partial_\alpha A_\beta - \partial_\beta A_\alpha) \ . \tag{6.72}$$

The $\beta = 0$ component is

$$0 = \partial^\alpha(\partial_\alpha A_0 - \partial_0 A_\alpha) = \Box A_0 - \partial_0 \partial^\alpha A_\alpha = \cancel{-\partial_0^2 A_0} + \underbrace{\partial_i \partial^i A_0}_{\Delta A_0} \cancel{+\partial_0^2 A_0} - \partial_0 \partial^i A_i = \tag{6.73}$$

$$= \partial^i(\partial^i A_0 - \partial_0 A_i) = \partial^i F_{i0} = \partial^i E_i =: C \tag{6.74}$$

[4] A PDE is quasilinear if it is linear in all the highest order derivatives of the unknown function.

Note that the electric field can be defined as

$$E_\alpha = F_{\alpha 0} = F_{\alpha\beta} n^\beta \ \text{ with } \ n^\beta = (1, \vec{0}) \ , \tag{6.75}$$

by introducing the timelike vector n^α.

While Eq. (6.72) appears as 4 wave-like equations for the 4 components of A_α, the $C = 0$ equation ($\beta = 0$ component) does not contain 2nd time derivatives. If one tries to take a time derivative of the equation to obtain a dynamical one, one fails and just find

$$\partial^0 C = \partial^0[\partial^\alpha(\partial_\alpha A_0 - \partial_0 A_\alpha)] = \partial^i[\underbrace{\partial^\alpha(\partial_\alpha A_i - \partial_i A_\alpha)}_{\text{l.h.s. of Maxwell eq. for } \beta=i}] \doteq 0 \tag{6.76}$$

as a result of the <u>identity</u>

$$0 \equiv \underbrace{\partial^\alpha\partial^\beta}_{\text{sym}} \underbrace{F_{\alpha\beta}}_{\text{antisym}} = \partial^\alpha(\partial^\beta F_{\alpha\beta}) = \partial^\alpha[\partial^\beta(\partial_\alpha A_\beta - \partial_\beta A_\alpha)] = -\partial^0[\partial^\beta(\partial_0 A_\beta - \partial_\beta A_0)] + \partial^i[\partial^\beta(\partial_i A_\beta - \partial_\beta A_i)] \tag{6.77}$$

that was used in the second passage (The $\doteq$ indicates as usual "on solution"/"on shell"). Hence,

- The equation
$$C = \partial_\alpha E^\alpha = \partial^\alpha(\partial_\alpha A_0 - \partial_0 A_\alpha) = 0 \tag{6.78}$$
is a *constraint.*
- If initially satisfied, it is "transported along the dynamics" because $\partial_0 C = 0$
- The Maxwell equations are *undetermined* (3 equations for the 4 components A_α)
- As formulated above, Maxwell equations do not admint a well-posed IVP: given a solution with initial data on a given spatial surface, $A_\alpha(t=0), \partial_0 A_\alpha(t=0)$, the component A_0 can be arbitrarily specified to obtain another solution.

At this point, one exploits the gauge freedom: two solutions represent the <u>same</u> electric and magnetic fields if they are related by the transformation

$$A_\alpha \ \mapsto \ A_\alpha + \partial_\alpha \phi \ . \tag{6.79}$$

Alternatively, the physical solution is given by the equivalence class of all the A_α related to each other by the gauge transformation above. To proceed, one must fix a gauge. For example, fix

$$\partial^\alpha A_\alpha = 0 \ \ \text{(Lorentz gauge, LG)} \ , \tag{6.80}$$

and obtain from Eq. (6.72)

$$0 = \partial^\alpha F_{\alpha\beta} = \partial^\alpha\partial_\alpha A_\beta - \cancel{\partial_\beta\partial^\alpha A_\alpha} \stackrel{LG}{=} \partial^\alpha\partial_\alpha A_\beta = \Box_\eta A_\beta \ . \tag{6.81}$$

Properties of the Maxwell equations with LG:

- The 4 equations are now dynamical (contain 2nd derivatives)
- The IVP is well posed (4 wave equations)
- For any choice of $A_\alpha(t=0), \partial_0 A_\alpha(t=0)$ that respects the gauge
$$\partial^\alpha A_\alpha(t=0) = 0 \ \text{ and } \ \partial_0\partial^\alpha A_\alpha(t=0) = 0 \ , \tag{6.82}$$
The LG is satisfied for all times because
$$0 \doteq \partial_\beta(\Box A^\beta) = \Box(\partial_\beta A^\beta) \tag{6.83}$$
- The constraint $C = 0$ ($\beta = 0$ eq. before gauge fixing) is sastified along the dynamics
$$C = \Box A_\alpha - \partial^\alpha\partial_0 A_\alpha = \Box A_\alpha - \partial_0\partial^\alpha A_\alpha \stackrel{LG}{\doteq} 0 \ . \tag{6.84}$$

Summary 6.6.1. *Maxwell equations for A_α (Eq. (6.72)) admit a well-posed IVP if one works in an appropriate gauge (e.g. Eq. (6.80)) and if initial data satisfy the constraint Eq. (6.78).*

EFE: constraints and evolution equations The structure of EFEs is similar to Maxwell equations. Let us assume the spacetime possess a timelike vector field n^a that defines a *foliation* of 3D spatial hypersurfaces. With this hypotesis we restrict ourselves to consider, of all the possibile spacetimes, only those called *globally hyperbolic spacetimes* (see below). In vacuum, the projections of EFE along n^b is

$$0 \doteq G_{ab} n^b =: C_a[\partial_i^2 g, \partial g, g] \tag{6.85}$$

and do not depend on 2nd time derivatives: they are 4 constraints. The Bianchi identities (Eq. (6.71)) play the role of the identity in Eq. (6.77) and guarantee that the constraints are transported along the dynamics.

A direct way to see this is to pick coordinates such that $0 = C^\mu = G^{0\mu}$, and because

$$\nabla_\alpha G^{\alpha\mu} = 0 \quad \Rightarrow \quad \partial_0 G^{0\mu} = \underbrace{-\partial_k G^{k\mu} - \Gamma^\mu_{\alpha\beta} G^{\alpha\beta} - \Gamma^\alpha_{\alpha\rho} G^{\mu\rho}}_{\text{this r.h.s contains at most } \partial_0^2 g} , \tag{6.86}$$

the l.h.s. $\partial_0 G^{0\mu}$ contains at most $\partial_0^2 g$, and thus $C^\mu = G^{0\mu}$ contains at most first time derivatives. Moreover, if $C^\mu(t=0) = 0$ then from the equation above $\partial_0 C^\mu = \partial_0 G^{0\mu} = 0$ because $G_{\mu\nu} \doteq 0$ for all times; the constraints are zero all times.

Similarly to Maxwell eqs, one is interested to the EFE solution given by the equivalence class of all the metric $g_{\alpha\beta}$ related to each other by coordinate transformation (diffeomorphism invariance). A way to obtain a well-posed IVP for Eq. (6.69a) is to choose coordinates such that $R_{\mu\nu} \sim \Box g_{\mu\nu}$, i.e.

$$H^\alpha \equiv 0 \quad \text{(Harmonic gauge)} , \tag{6.87}$$

and initial data for $g_{\mu\nu}$ that satisfy the constraints. The name of this coordinate condition follows from:

$$0 = \Box x^\mu = g^{\alpha\beta}\nabla_\alpha\nabla_\beta x^\mu = g^{\alpha\beta}\nabla_\alpha(\partial_\beta x^\mu) = g^{\alpha\beta}[\partial_\alpha(\underbrace{\partial_\beta x^\mu}_{\delta^\mu_\beta}) - \underbrace{\partial_\gamma x^\mu}_{\delta^\mu_\gamma}\Gamma^\gamma_{\alpha\beta}] = 0 - g^{\alpha\beta}\Gamma^\mu_{\alpha\beta} = -H^\mu . \tag{6.88}$$

Causality and globally hyperbolic spacetime. Above, when separating the EFE in constraints and evolution equations, we have implicitly assumed the existance of a <u>global notion of time</u> that determines past and future of each event. In SR the causal structure is simple and given by the light cones: an event can be connected by spacelike, timelike, null curves to other events, and that determines <u>in an absolute sense</u> its future and its past. In the language of hyperbolic PDE, the light cone determines the domain of dependence and the domain of influence of the solution of the wave equation, $\Box_\eta \phi = 0$.

In GR the *global* causal structure of the spacetime is more complex. One can consider a *closed* set of causally-connected events: it is impossible to say which event of the set happened before or after another one. The situation corresponds to the existance of closed timelike curves; an example is given by the Gödel cosmology that satisfies EFE with the cosmological constant. Other examples are discussed in e.g. Chap. 2 of Carroll (1997). One considers "physically realistic" a spacetime in which "causality is well-behaved", i.e. where it is possible to continuously distinguish between past and future as the event p moves in $\mathcal{M}$. Such manifolds are called *time-orientable*.

Some definitions:

- *Achronal set* $S \subset \mathcal{M}$: subset of events that are **not** connected by timelike curves
- *Future domain of dependence of* S $D_+(S)$: the set of events such that every causal curve intersect S in the past
- *Future Cauchy horizon of* S $H_+(S)$: the boundary of $D_+(S)$

All the definitions repeat substituting $+ \mapsto -$ and "past"$\mapsto$ "future". The domain of dependence is $D(S) = D_+(S) \cup D_-(S)$. Finally, a *Cauchy surface* is a spacelike hypersurface $\Sigma \subset \mathcal{M}$ whose domain of dependence is the entire manifold $D(\Sigma) = \mathcal{M}$ ($H(\Sigma) = 0$). Every causal (timelike or null) curve without end-point intersect Σ only once. In other terms, given Σ one can predict past and future. Note Σ are not unique.

Definition 6.6.1. $\mathcal{M}$ *is said* globally hyperbolic spacetime *iff admits a Cauchy surface.*

Many spacetimes of astrophysical and cosmological interest are assumed to be globally hyperbolic. For example, it should be clear that weak-field spacetimes or the spacetime outside a star are of such type.

Remark 6.6.2. *The IVP for the wave equation $\Box_g \phi = 0$ is well posed in a globally hyp. spacetime. See e.g. (Baer et al., 2008).*

6.7 Killing vectors (KVs)

When discussing the local conservation law for the stress-energy tensor it was pointed out that observer such that $\nabla_{(a}v_{b)} = 0$ allowed the definition of energy conservation. Indeed vectors satifies that equation play a special role in GR and are associated with conserved currents.

Definition 6.7.1. Killing vector*: k^a vector field solution of the Killing equation $\nabla_{(a}k_{b)} = 0$.*

Theorem 6.7.1. *The matter current associated to a KV $J^a_k := T_{ab}k^b$ is conserved $\nabla_a J^a_k = 0$.*

Proof (See Eq. (6.33) and discussion around there):

$$\nabla_a J^a_k = \underbrace{\nabla_a T^{ab}}_{=0} k_b + T^{ab}\nabla_a k_b = T^{ab}\nabla_a k_b = \frac{1}{2}T^{ab}\nabla_{(a}k_{b)} = 0 , \tag{6.89}$$

where one uses the KV definition in the last equality, the equation

$$T^{ab}\nabla_{(a}v_{b)} = T^{ab}\nabla_{(a}v_{b)} = \frac{1}{2}(T^{ab}\nabla_a v_b + \underbrace{T^{ab}}_{=T^{ba}}\nabla_b v_a) = T^{ab}\nabla_a v_b , \tag{6.90}$$

that is valid for every symmetric tensor T_{ab}, and the EOM for the stress energy tensor.

In physics, Nöther theorem says that conserved quantities are associated to symmetries. In GR, symmetries are precisely associated to KV. Technically, one says that KV are infinitesimal generators of isometries because the metric remains invariant along their integral curve (see also below discussion on Lie derivatives).

Imagine the spacetime has some symmetry (stationary, spherical symmetry, etc). Then there must exists some coordinates adapted to the symmetry in which the metric component are independent on certain coordinates. For example, the metric

$$g = \mathrm{d}x^2 + f(y)\mathrm{d}y^2 \tag{6.91}$$

is invariant under translations in the x direction and its components do not depend on x; or the metric

$$g = A(r)\mathrm{d}r^2 + B(r)\mathrm{d}\Omega^2 \tag{6.92}$$

is a spherically symetric and its components do not dependend on the angle ϕ. In general, given a symmetry there exists a specific coordinate x^{σ_*} (σ_* is fixed!) such that

$$\partial_{\sigma_*} g_{\mu\nu} = 0 \ \ \forall \mu, \nu \ . \tag{6.93}$$

In this coordinate system the KV is simply given by

$$k^\mu = (\partial_{\sigma_*})^\mu = \delta^\mu_{\sigma_*} \ . \tag{6.94}$$

Geodesics in presence of symmetries. KV are associated to conserved quantities in along geodesics.

Take the geodesic equation for the tangent velocity $u^\mu = \dot{x}^\mu$ with index down in coordinate adapted to the symmetry:

$$u^\alpha \nabla_\alpha u^\mu = 0 \quad \Leftrightarrow \quad \frac{du_\mu}{d\lambda} = \frac{1}{2}\partial_\mu g_{\nu\alpha} u^\nu u^\alpha \ . \tag{6.95}$$

For the special coordinate $\mu = \sigma_*$, there exists a conserved quantity:

$$\mu = \sigma_* \quad \Rightarrow \quad \partial_{\sigma_*} g_{\nu\alpha} = 0 \quad \Rightarrow \quad \frac{du_{\sigma_*}}{d\lambda} = 0 \ . \tag{6.96}$$

Interestingly the conserved quantity can be written in a invariant way as the contraction between the tangent vector and the KV:

$$u_{\sigma_*} = u_\mu \delta^\mu_{\sigma_*} = u^\mu k_\mu = u_\mu k^\mu \ . \tag{6.97}$$

Going back to the geodesic equation, one sees immediately that this implies that (note the last passage the use of symmetry properies discussed in Remark 6.3.1)

$$\frac{du_{\sigma_*}}{d\lambda} = 0 \quad \Rightarrow \quad 0 = u^\alpha \nabla_\alpha (u^\mu k_\mu) = \underbrace{u^\alpha \nabla_\alpha u^\mu}_{=0} k_\mu + u^\alpha u^\mu \nabla_\alpha k_\mu = u^\alpha u^\mu \nabla_{(\alpha} k_{\mu)} \ . \tag{6.98}$$

The Killing equation (and the related antisymmetry of $\nabla_\nu k_\mu$) implies the existance of conserved quantities along the geodesics.

Killing vectors and Riemann tensor. The intuitive idea developed so far is that geometry does not change along the direction identified by a KV. Indeed this is the content of the following

Theorem 6.7.2. *The directional derivative of the Ricci scalar along the KV is zero:* $k^a \nabla_a R = 0$.

Proof (sketch). The proof uses the antisymmetry of the tensor $\nabla_a k_b$ (implied by the Killing equation) and the Bianchi identities. The main steps are:

- Show the derivatives of the KV are actually related to the Riemann tensor by

$$\nabla_a \nabla_b k^c = R_{bad}{}^c k^d \ , \quad \text{and} \quad \nabla_a \nabla_b k^a = R_{bd} k^d \ . \tag{6.99}$$

- Start from contracting the Bianchi identities

$$0 = \nabla^a G_{ab} k^b = (\nabla^a R_{ab} - \frac{1}{2} g_{ab} \nabla^a R) k^b = (\nabla^a R_{ab} - \frac{1}{2}\nabla_b R) k^b \quad \Rightarrow \quad 2k^b \nabla^a R_{ab} = k^b \nabla_b R \ . \tag{6.100}$$

- Show that the l.h.s. of the last equation is zero using the symmetry properties discussed in Rem. 6.3.1 and the relation between the derivative of the KV and Ricci tensor:

$$k^b \nabla^a R_{ab} = \nabla^a (k^b R_{ab}) - \underbrace{R_{ab}}_{\text{sym}} \underbrace{\nabla^a k^b}_{\text{antisym}} = \nabla^a (k^b R_{ab}) + 0 = \nabla^a \nabla^c \nabla_a k_c = [\nabla^a, \nabla^c] \underbrace{\nabla_a k_c}_{\text{antisym}} = 0 \ . \tag{6.101}$$

6.8 Lie derivative

Derivative operators studied so far:
- **d** *Exterior derivative*
  - specific for p-forms
  - no metric required
  - key relevance: Stokes' theorem.
- ∇ *Convariant derivative*
  - for any tensor
  - Levi-Civita connection compatible with the metric
  - key relevance: parallel transport, curvature.

Let us introduce
- $\mathcal{L}$ *Lie derivative* along u
  - for any tensor
  - directional derivative along u^a
  - key relevance: symmetries.

Definition 6.8.1. *Consider two manifolds $\mathcal{M}$ and $\mathcal{N}$, the function $f : \mathcal{N} \mapsto \mathbb{R}$ and the map $\phi : \mathcal{M} \mapsto \mathcal{N}$.*
The pushback *of the function f by ϕ is the function $\phi_* f = f \circ \phi : \mathcal{M} \mapsto \mathbb{R}$.*
The pullforward *of a vector $v \in T_p\mathcal{M}$ is the vector $\phi^* v(f) = v(\phi_* f) = v(f \circ \phi) \in T_{\phi(p)}\mathcal{N}$.*
The pushback *of a 1-form $\omega \in T^*_p\mathcal{N}$ is the 1-form $\phi_*\omega(v) = \omega(\phi^* v) \in T^*_p\mathcal{M}$.*

Note these are "one-way" maps: functions cannot be pulled forward, vectors cannot be pushed back, 1-forms cannot be pulled forward etc. The pushback/pullforward operations extended to $(0, l)$ and $(k, 0)$ tensors respectively.

Pushback/pullforward operations are general transformations connecting points and vectors/duals between manifolds. Given a natural basis of $\mathcal{M}$ $\{\partial_\mu = \partial/\partial x^\mu \ (\mu = 1...m)\}$ and of $\mathcal{N}$ $\{\partial_\alpha = \partial/\partial y^\alpha \ (\alpha = 1...n)\}$ (note the dimensions can be different), the pushback/pullforward operations are

$$\phi^* v(f) = \underline{(\phi^* v)^\alpha}\partial_\alpha f = v^\mu \partial_\mu(\phi_* f) = v^\mu\partial_\mu(f \circ \phi) = v^\mu \partial_\mu(f(\phi)) = v^\mu \frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial f}{\partial y^\alpha} = \underline{v^\mu \frac{\partial y^\alpha}{\partial x^\mu}}\partial_\alpha f \tag{6.102a}$$

$$\phi_*\omega(v) = \underline{(\phi_*\omega)_\mu}\mathrm{d}x^\mu(v) = \omega_\alpha \mathrm{d}y^\alpha(\phi^* v) = \omega_\alpha v^\nu \frac{\partial y^\beta}{\partial x^\nu}\underbrace{\mathrm{d}y^\alpha(\partial_\beta)}_{=\delta^\alpha_\beta} = \omega_\alpha v^\nu \frac{\partial y^\alpha}{\partial x^\nu} =$$

$$= \omega_\alpha v^\nu \frac{\partial y^\alpha}{\partial x^\mu}\delta^\mu_\nu = \omega_\alpha v^\nu \frac{\partial y^\alpha}{\partial x^\mu}\mathrm{d}x^\mu(\partial_\nu) = \underline{\omega_\alpha \frac{\partial y^\alpha}{\partial x^\mu}}\mathrm{d}x^\mu(v) \tag{6.102b}$$

where the underlined expressions in the same line highlight the components of the pushback/pullforward operation in terms of those of the argument. In terms of components, one can think about the action as a transformation matrix, but should note that the transformation in general is **not** invertible.

Pushback/pullforward operations are useful in the context of
- Submanifolds and calculation of induced metric, Example 6.8.1;
- Re-interpret coordinate transformation on a manifold as diffeomorphisms (active coordinate transformations), Remark 6.8.1

Example 6.8.1. *Consider the unit 2-sphere S^2 with coordinate $x^\mu = (\theta, \varphi)$ immerse in $\mathbb{R}^3$ with coordinate $y^\alpha = (x, y, z)$ and the map*

$$\phi : S^2 \mapsto \mathbb{R}^3 \ : \ \phi(\theta,\varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) \ . \tag{6.103}$$

The pushforward of the Euclidean metric

$$g_{\alpha\beta} = \mathrm{diag}(1, 1, 1) \tag{6.104}$$

is given by the transformation

$$\frac{\partial y^\alpha}{\partial x^\mu} = \begin{bmatrix} \cos\theta\cos\varphi & \cos\theta\sin\varphi & -\sin\theta \\ -\sin\theta\sin\varphi & \sin\theta\cos\varphi & 0 \end{bmatrix} \tag{6.105}$$

and results is the induced metric on S^2

$$g_{\mu\nu} = \frac{\partial y^\alpha}{\partial x^\mu}\frac{\partial y^\beta}{\partial x^\nu} g_{\alpha\beta} = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix} \tag{6.106}$$

Remark 6.8.1. *Coordinate transformations can be interpreted as diffeomorphisms on the manifold. The the diffeomorphism $\phi : \mathcal{M} \mapsto \mathcal{M}$ (smooth and invertible). Instead of changing from $x^\mu : \mathcal{M} \mapsto \mathbb{R}^n$ to new functions $x^{\mu'} : \mathcal{M} \mapsto \mathbb{R}^n$ (remapping the manifold), one can think of changing the points on the manifold using ϕ and then evaluate the coordinates on the new points using the pullback: $(\phi_* x)^\mu : \mathcal{M} \mapsto \mathbb{R}^3$. This new point of view gives another way to compare tensors at different points on $\mathcal{M}$. (See Remark 3.4.1.)*

Consider a vector field u^a, the associated field lines ϕ_λ (integral curves $\dot{x}^\mu(\lambda) = u^\mu$) constitute a 1-parameter family of diffeomorphisms. Take another vector v^a and ask: What is the variation of v along the u? Clearly, vectors need to be compare at the same point p;

Definition 6.8.2. Lie derivative *along u of v at point p is a vector given by pullback the vector v in a neighbor point along the integral curves of u:*

$$\mathcal{L}_u v := \lim_{\lambda\to 0} \frac{\phi^*_\lambda v - v}{\lambda} \ . \tag{6.107}$$

Introduce a coordinate system y^μ adapted to the vector u, i.e. such that $u^\mu = (1, 0, 0, ..., 0)$. The integral curves close to point p in these coordinates give immediately a coordinate expression for the Lie derivative is

$$\phi_\lambda(p) = (y^1 + \lambda, y^2, ..., y^n) \ , \quad \Rightarrow \quad \mathcal{L}_u v^\mu = \frac{\partial v^\mu}{\partial y^0} \ . \tag{6.108}$$

The expression above is not covariant, but observing that in the same coordinate the commutator between u and v has the same expression

$$[u, v]^\mu = u^\nu \partial_\nu v^\mu - v^\nu \partial_\nu u^\mu = \frac{\partial v^\mu}{\partial y^0} \ , \tag{6.109}$$

should convince that the general expression is:

$$\mathcal{L}_u v = [u, v] \ . \tag{6.110}$$

Properties.

- $\mathcal{L}_u v = -\mathcal{L}_v u$
- The Lie derivative generalizes to any tensor

$$\mathcal{L}_u T^{a_1\dots a_n}_{b_1\dots b_n} = u^c \nabla_c T^{a_1\dots a_n}_{b_1\dots b_n} - \sum_j \nabla_c u^{a_j} T^{a_1\dots c \dots a_n}_{b_1\dots b_n} + \sum_i \nabla_{b_i} u^c T^{a_1\dots a_n}_{b_1\dots c\dots b_n} \ . \tag{6.111}$$

Relation to KV and symmetries. The Lie derivative of the metric along u is the symmetrized covariant derivative of u:

$$\mathcal{L}_u g_{ab} = u^c \underbrace{\nabla_c g_{ab}}_{=0} + \nabla_a u^c g_{ab} + \nabla_b u^c g_{ab} = \nabla_{(a} u_{b)} \ . \tag{6.112}$$

This implies that the Lie derivative of the metric is zero along a KV:

$$\nabla_{(a} k_{b)} \quad \Leftrightarrow \quad \mathcal{L}_k g_{ab} \ , \tag{6.113}$$

i.e. the metric is constant along the integral lines of k. More in general,for any tensor that is invariant with respect the diffeomorphism generated by a vector u, the Lie derivative along u fo the tensor is zero.

7. Weak field and waves

④

These lectures present the weak field limit of GR and the EFE linearized arounf Mikowski spacetime. The equations apply for the description of light deflection, gravitoeletric/magnetic phenomena and gravitational waves. Gravitational-wave propagation properties, their effect on test masses, and the quadrupole formula are discussed. A discussion on the concept of gravitational-wave energy and energy in GR is started here.

Suggested readings. *Chap. 4 of Wald (1984); Chap. 4 of Carroll (1997); Chap. 7-8 of Schutz (1985); Book of Maggiore (2007).*

7.1 Weak field GR

In the regime of weak gravity one assumes that there exists a global inertial Cartesian frame in which the metric can be written as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad \text{with} \quad |h_{\mu\nu}| \ll |\eta_{\mu\nu}| \sim 1 \ . \tag{7.1}$$

Since the component of the perturbation of Mikowski spacetime are "small" (in the sense above), the GR equations can be linearized at linear order in h. Linearized equations apply, for example, to the Solar system where

$$|h_{\mu\nu}| \sim \frac{\phi}{c^2} \lesssim \frac{GM_\odot}{c^2 R_\odot} \sim 10^{-6} \ , \tag{7.2}$$

and in general describe

- Newtonian gravity;
- Gravitoelectric and gravitomagnetic phenomena;
- Propagation of gravitational waves.

Formally, the linearized theory can be regarded as a field theory in which

- $\eta_{\mu\nu}$ is a background metric;
- The grav.field generated by the matter does not backreact on the source;
- $h_{\mu\nu}$ is the main field and transforms as a tensor on flat spacetime under Lorentz transformation (Lorentz covariance). Consider a Lorentz transformation of coordinates ($\Lambda^{\mathrm{T}}\eta\Lambda = \eta$):

$$x^\mu = \Lambda^\mu{}_\nu (x')^\nu = \frac{\partial x^\mu}{\partial x^{\nu'}} x^{\nu'} \quad \Rightarrow \quad g_{\mu'\nu'} = \frac{\partial x^\mu}{\partial x^{\mu'}} \frac{\partial x^\nu}{\partial x^{\nu'}} g_{\mu\nu} = \Lambda^\mu{}_{\mu'} \Lambda^\nu{}_{\nu'} g_{\mu\nu} = \Lambda^\mu{}_{\mu'} \Lambda^\nu{}_{\nu'} (\eta_{\mu\nu} + h_{\mu\nu}) \tag{7.3a}$$

$$= \underbrace{\Lambda^\mu{}_{\mu'} \Lambda^\nu{}_{\nu'} \eta_{\mu\nu}}_{=\eta_{\mu'\nu'}} + \Lambda^\mu{}_{\mu'} \Lambda^\nu{}_{\nu'} h_{\mu\nu} = \eta_{\mu\nu} + \Lambda^\mu{}_{\mu'} \Lambda^\nu{}_{\nu'} h_{\mu\nu} \tag{7.3b}$$

$$\Rightarrow \quad h_{\mu\nu} \to h_{\mu'\nu'} = \Lambda^\mu{}_{\mu'} \Lambda^\nu{}_{\nu'} h_{\mu\nu} \ . \tag{7.3c}$$

7.2 Infinitesimal diffeomorphism invariance

Symmetry of linearized GR. Consider an infinitesimal coordinate transformation:

$$x^\mu \mapsto x^{\mu'} = x^\mu + \xi^\mu(x^\alpha) \quad \text{with} \quad |\partial_\beta \xi^\alpha| \sim |h_{\mu\nu}| \ll 1 \tag{7.4a}$$

$$\frac{\partial x^{\mu'}}{\partial x^\mu} = \delta^{\mu'}_\mu + \partial_\mu \xi^{\mu'} \tag{7.4b}$$

$$\frac{\partial x^\mu}{\partial x^{\mu'}} = \delta^\mu_{\mu'} - \partial_{\mu'} \xi^\mu + \mathcal{O}(|\partial\xi|^2) \tag{7.4c}$$

where one uses the Taylor expansion for the inverse $(1+\delta A)^{-1} \approx 1-\delta A$. Note prime indexes refer to tensor components in primed coordinates x', while the unprimed indexes refer to tensor component in unprime coordinates x; indexes on r.h.s. and l.h.s. take the same values (although they do not "match" on the two side of the equation). For example, $g_{\mu'\nu'} = \eta_{\mu\nu} + h_{\mu\nu} = \eta_{\mu'\nu'} + h_{\mu\nu}$ really means $g_{\mu\nu}(x') = \eta_{\mu\nu} + h_{\mu\nu}(x)$. Using this notation is useful to keep formulas compact. To linear order in h and in $\partial\xi$ the metric change is:

$$
\begin{aligned}
g_{\mu'\nu'} &= \frac{\partial x^\mu}{\partial x^{\mu'}}\frac{\partial x^\nu}{\partial x^{\nu'}} g_{\mu\nu} = (\delta^\mu_{\mu'} - \partial_{\mu'}\xi^\mu)(\delta^\nu_{\nu'} - \partial_{\nu'}\xi^\nu)(\eta_{\mu\nu} + h_{\mu\nu}) && (7.5a)\\
&= (\delta^\mu_{\mu'}\delta^\nu_{\nu'} - \partial_{\mu'}\xi^\mu\delta^\nu_{\nu'} - \partial_{\nu'}\xi^\nu\delta^\mu_{\mu'} + \partial_{\mu'}\xi^\mu\partial_{\nu'}\xi^\nu)(\eta_{\mu\nu} + h_{\mu\nu}) && (7.5b)\\
&= \underbrace{\delta^\mu_{\mu'}\delta^\nu_{\nu'}\eta_{\mu\nu}}_{=\eta_{\mu'\nu'}} - \underbrace{\partial_{\mu'}\xi^\mu\delta^\nu_{\nu'}\eta_{\mu\nu}}_{=\partial_{\mu'}\xi_{\nu'}} - \partial_{\nu'}\xi^\nu\delta^\mu_{\mu'}\eta_{\mu\nu} + \underbrace{\delta^\mu_{\mu'}\delta^\nu_{\nu'}h_{\mu\nu}}_{h_{\mu'\nu'}} - \underbrace{\partial_{\mu'}\xi^\mu\delta^\nu_{\nu'}h_{\mu\nu} - \partial_{\nu'}\xi^\nu\delta^\mu_{\mu'}h_{\mu\nu} + \partial_{\mu'}\xi^\mu\partial_{\nu'}\xi^\nu h_{\mu\nu}}_{=\mathcal{O}(h^2)} && (7.5c)\\
&= \eta_{\mu'\nu'} + h_{\mu'\nu'} - 2\partial_{(\mu'}\xi_{\nu')} && (7.5d)\\
&= \eta_{\mu\nu} + h_{\mu'\nu'} - 2\partial_{(\mu'}\xi_{\nu')} && (7.5e)
\end{aligned}
$$

Any infinitesimal coordinate transformation that maps the perturbed metric field into

$$
h_{\mu\nu} \mapsto h_{\mu'\nu'} = h_{\mu\nu} + 2\partial_{(\mu}\xi_{\nu)} \ , \tag{7.6}
$$

leaves the metric invariant. The weak metric is represented by the equivalence classes of metrics linked by the infinitesimal coordinate transformations above. Note the linearized Einstein tensor is invariant w.r.t the above transformation [exercise].

Observations.

- If one considers ξ^α as the components of a vector, then the transformation can be written in terms of the Lie derivatives, and the infinitesimal coordinate transformation is then interpreted as an *infinitesimal diffeomorphism* generated by ξ:

$$
2\partial_{(\mu}\xi_{\nu)} = \mathcal{L}_\xi\eta_{\mu\nu} \quad \Rightarrow \quad h_{\mu\nu} \mapsto h_{\mu\nu} + \mathcal{L}_\xi\eta_{\mu\nu} \ . \tag{7.7}
$$

  Hence, weak field GR is *invariant under infinitesimal diffeomorphisms.*
- The above transformation is the analogous of gauge transformation for the potentials in electromagnetism:

$$
A_\alpha \mapsto A_\alpha + \partial_\alpha\chi \ . \tag{7.8}
$$

7.3 Weak field equations

Calculation of the Einstein tensor linearized in h. Indexes are raised on lowered with the flat metric η; for example the trace of the perturbation is $h := h^\alpha_\alpha = \eta^{\alpha\beta}h_{\alpha\beta}$. It is left as [exercise] to show:

$$
\begin{aligned}
g^{\mu\nu} &= \eta^{\mu\nu} + h^{\mu\nu} + \mathcal{O}(h^2) && (7.9a)\\
\Gamma^\mu_{\alpha\beta} &= \frac{1}{2}\eta^{\mu\lambda}\left(\partial_\alpha h_{\lambda\beta} + \partial_\beta h_{\lambda\alpha} - \partial_\lambda h_{\alpha\beta}\right) + \mathcal{O}(h^2) && (7.9b)\\
R_{\mu\nu} &= \partial\Gamma - \partial\Gamma + \underbrace{\Gamma\Gamma - \Gamma\Gamma}_{\mathcal{O}(h^2)} = \partial_\alpha\Gamma^\alpha_{\mu\nu} - \partial_\mu\Gamma^\alpha_{\alpha\nu} + \mathcal{O}(h^2) \approx \partial^\alpha\partial_{(\mu}h_{\nu)\alpha} - \frac{1}{2}\partial_\lambda\partial^\lambda h_{\mu\nu} - \frac{1}{2}\partial_\mu\partial_\nu h && (7.9c)\\
R &= \eta^{\mu\nu}R_{\mu\nu} = \frac{1}{2}\left(\eta^{\mu\nu}\partial^\alpha\partial_\mu h_{\nu\alpha} + \eta^{\mu\nu}\partial^\alpha\partial_\nu h_{\mu\alpha}\right) - \frac{1}{2}\partial_\lambda\partial^\lambda(\underbrace{\eta^{\mu\nu}h_{\mu\nu}}_{=h}) - \frac{1}{2}\underbrace{\eta^{\mu\nu}\partial_\mu\partial_\nu}_{\partial_\lambda\partial^\lambda} h = \frac{1}{\cancel{2}}\cancel{2}\partial^\alpha\partial^\nu h_{\nu\alpha} - \partial_\lambda\partial^\lambda h && (7.9d)\\
G_{\mu\nu} &= R_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}R = \partial^\alpha\partial_{(\mu}h_{\nu)\alpha} - \frac{1}{2}\partial_\lambda\partial^\lambda h_{\mu\nu} - \frac{1}{2}\partial_\mu\partial_\nu h - \frac{1}{2}\eta_{\mu\nu}(\partial^\alpha\partial^\beta h_{\alpha\beta} - \partial_\lambda\partial^\lambda h) \ . && (7.9e)
\end{aligned}
$$

The linearized Einstein tensor

$$
G_{\mu\nu} = \underbrace{\partial^\alpha\partial_{(\mu}h_{\nu)\alpha}}_{\text{I.}} - \underbrace{\frac{1}{2}\partial_\lambda\partial^\lambda h_{\mu\nu}}_{\text{II.}} - \underbrace{\frac{1}{2}\partial_\mu\partial_\nu h}_{\text{III.}} - \underbrace{\frac{1}{2}\eta_{\mu\nu}\partial^\alpha\partial^\beta h_{\alpha\beta}}_{\text{IV.}} + \underbrace{\frac{1}{2}\eta_{\mu\nu}\partial_\lambda\partial^\lambda h}_{\text{V.}} \ , \tag{7.10}
$$

can be written in a simpler form considering the <u>trace reverse metric</u>

$$
\bar{h}_{\mu\nu} := h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h \ . \tag{7.11}
$$

Note that

$$
\bar{h} = \eta^{\mu\nu}\bar{h}_{\mu\nu} = \underbrace{\eta^{\mu\nu}h_{\mu\nu}}_{=h} - \frac{1}{2}\underbrace{\eta^{\mu\nu}\eta_{\mu\nu}}_{=4} h = -h \ . \tag{7.12}
$$

Calculate each term:

$$\mathrm{I} = \partial^\alpha \partial_{(\mu} \bar{h}_{\nu)\alpha} + \frac{1}{4}\partial^\alpha \partial_\mu (\eta_{\nu\alpha} h) + \frac{1}{4}\partial^\alpha \partial_\nu (\eta_{\mu\alpha} h) = \underbrace{\partial^\alpha \partial_{(\mu} \bar{h}_{\nu)\alpha}}_{\mathrm{Ia}} + \underbrace{\frac{1}{2}\partial_{(\mu}\partial_{\nu)} h}_{\mathrm{Ib}} \tag{7.13a}$$

$$\mathrm{II} = \underbrace{-\frac{1}{2}\eta_{\alpha\beta}\partial^\alpha\partial^\beta \bar{h}_{\mu\nu}}_{\mathrm{IIa}} \underbrace{- \frac{1}{4}\eta_{\mu\nu}\eta_{\alpha\beta}\partial^\alpha\partial^\beta h}_{\mathrm{IIb}} \tag{7.13b}$$

$$\mathrm{IV} = -\frac{1}{2}\eta_{\mu\nu}\partial^\alpha\partial^\beta(\bar{h}_{\alpha\beta} + \frac{1}{2}\eta_{\alpha\beta} h) = \underbrace{-\frac{1}{2}\eta_{\mu\nu}\partial^\alpha\partial^\beta \bar{h}_{\alpha\beta}}_{\mathrm{IVa}} \underbrace{- \frac{1}{4}\eta_{\mu\nu}\eta_{\alpha\beta}\partial^\alpha\partial^\beta h}_{\mathrm{IVb}} \tag{7.13c}$$

$$0 = \mathrm{V} + \mathrm{IIb} + \mathrm{IVb} \tag{7.13d}$$

$$0 = \mathrm{III} + \mathrm{Ib} \tag{7.13e}$$

$$G_{\mu\nu} = \mathrm{IIa} + \mathrm{Ia} + \mathrm{IVa} = -\frac{1}{2}\eta_{\alpha\beta}\partial^\alpha\partial^\beta \bar{h}_{\mu\nu} + \partial^\alpha\partial_{(\mu}\bar{h}_{\nu)\alpha} - \frac{1}{2}\eta_{\mu\nu}\partial^\alpha\partial^\beta\bar{h}_{\alpha\beta} \tag{7.13f}$$

From the above expression one sees that the last two terms contain the divergence of the metric. Imposing the *Hilbert gauge* (or Lorentz)

$$\partial^\alpha \bar{h}_{\mu\alpha} = 0 \ , \quad \text{(Hilbert gauge)} \tag{7.14}$$

leads to the following equations for linearized GR:

$$\Box_\eta \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu} \ . \tag{7.15}$$

Observations

- By making an infinitesimal coordinate transformation, it is always possible to reduce to Hilbert gauge. The set of 4 functions ξ is given by the solution of 4 inhomogeneous wave equations (Cf. Lorentz gauge in electrodynamics):

$$h_{\mu\nu} \mapsto h_{\mu\nu} + 2\partial_{(\mu}\xi_{\nu)} \tag{7.16a}$$

$$h \mapsto h + \eta^{\alpha\beta}\partial_\alpha\xi_\beta + \eta^{\alpha\beta}\partial_\beta\xi_\alpha = h + 2\partial_\mu\xi^\mu \tag{7.16b}$$

$$\bar{h}_{\mu\nu} \mapsto \bar{h}_{\mu\nu} + 2\partial_{(\mu}\xi_{\nu)} - \eta_{\mu\nu}\partial_\alpha\xi^\alpha \tag{7.16c}$$

$$\partial^\alpha\bar{h}_{\mu\alpha} \mapsto \partial^\alpha\bar{h}_{\mu\alpha} + \Box\xi_\nu + \cancel{\partial^\mu\partial_\nu\xi_\mu} - \cancel{\partial_\nu\partial_\lambda\xi^\lambda} \quad \Rightarrow \quad \Box\xi_\nu = -\partial^\alpha\bar{h}_{\nu\alpha} =: V_\nu \neq 0 \tag{7.16d}$$

- Eq. (7.15) is a linear wave equation for the components of h. At linear order in h, the stress energy tensor does **not** depend on h, so in linear GR one can specify the matter source and solve for the metric. For example, it is possible to calculate solutions using Green functions as in electrodynamics (see below).
- The Bianchi identity in the weak field simplifies, and involves the partial derivative of the Einstein tensor (since ∂ is the connection associated to η). Hence, the weak field equations imply the conservation of the stress-energy tensor on flat background, and that matter does not backreact on the curvature:

$$\partial_\nu G^{\mu\nu} = 0 \quad \Rightarrow \quad \partial_\nu T^{\mu\nu} = 0 \ . \tag{7.17}$$

- In vacuum, linearized EFEs are the equations for a massless spin-2 field propagating in flat spacetime (Chap. 13 Wald (1984)).

7.4 Weak field solutions

Formal solutions of Eq. (7.15) for static and stationary matter distributions.

7.4.1 Static source

A static matter distribution is modeled by a time-independent stress-energy tensor in the form

$$T_{\mu\nu} = \rho t_\mu t_\nu \ \text{ i.e. } \ T_{00} = c^2\rho(x^i) \ , \quad T_{0i} = T_{ij} = 0 \ , \tag{7.18}$$

where $t^\mu = (\partial_t)^\mu$ is the vector along the time direction of the global inertial coordinates. With this prescription, the r.h.s. of Eq. (7.15) is time-independent, thus also the grav. field must be time independent

$$\partial_t \bar{h}_{\mu\nu} = 0 \ , \tag{7.19}$$

and the linearized EFE reduce to Poisson equations for the components of the metric field:

$$\begin{cases} \Delta\bar{h}_{\mu\nu} = -16\pi\rho & , \quad \mu = \nu = 0 \\ \Delta\bar{h}_{\mu\nu} = 0 & , \quad \text{otherwise} \ . \end{cases} \tag{7.20}$$

The solution is immediately given in terms of the Newton grav. potential (formally compare the equations above to Newton's $\Delta\phi = 4\pi\rho$):

$$\begin{cases} \bar{h}_{\mu\nu} &= -4\phi \ , \quad \mu=\nu=0 \\ \bar{h}_{\mu\nu} &= 0 \ , \quad \text{otherwise} \end{cases} \quad \Rightarrow \quad \bar{h}_{\mu\nu} = -4\phi t_\mu t_\nu \ . \tag{7.21}$$

Reversing the trace:

$$\bar{h} = \eta^{\mu\nu}\bar{h}_{\mu\nu} = \eta^{00}\bar{h}_{00} = +4\phi \tag{7.22a}$$

$$h_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\bar{h} = -4\phi t_\mu t_\nu - \frac{1}{2}\eta_{\mu\nu}(4\phi) \tag{7.22b}$$

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} = \eta_{\mu\nu} + \bar{h}_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}\bar{h} = \eta_{\mu\nu}(1-\frac{\bar{h}}{2}) + \bar{h}_{\mu\nu} = \eta_{\mu\nu}(1-2\phi) - 4\phi t_\mu t_\nu \tag{7.22c}$$

or (reintroducing c):

$$g = -c^2\left(1+2\frac{\phi}{c^2}\right)\mathrm{d}t^2 + \left(1-\frac{2\phi}{c^2}\right)\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j \ . \tag{7.23}$$

Note that far from a source of mass M the multipolar expansion of the grav. potential starts with $\phi \approx -\frac{M}{r} + \mathcal{O}(1/r^2)$, so the distant metric is fully specified by the source mass and reduces to Mikowski at $r\to\infty$ if $M\neq 0$, or everywhere if $M=0$. Recalling the discussion in Chap. 4, it is remarkable that this metric describes the motion of particles (geodesic motion) in both Newtonian gravity and SR (using the appropriate limits). Note, however, a subtle point: the geodesic equations on the <u>weak metric</u> (weak gravity field) imply

$$\frac{d^2x^i}{dt^2} = -\partial_i\phi \ , \quad \text{(weak gravity)} \tag{7.24}$$

but they are **not** consistent with the equation $\partial_\mu T^{\mu\nu} = 0$, that imply instead geodesics on <u>flat metric</u> (Mikowksi spacetime, unaccelerated motion, no gravity)

$$\frac{d^2x^i}{dt^2} = 0 \ , \quad \text{(no gravity)} \ . \tag{7.25}$$

This illustrates some of the difficulties/inconsistencies in predicting the EOM of matter from linearized EFE as an expansion on η. Eq. (7.25) is the EOM of the matter that determines curvature, Eq. (7.24) is the EOM of matter in the resulting slightly curved spacetime.

Example 7.4.1. Deflection of light (Einstein's 1915 calculation). *Consider a photon moving in a weak and static gravitational field generated by a mass M. The weak metric is fully defined by Eq. (7.23) and the grav. potential $\phi = GM/rc^2$. The photon moves in the $z=0$ plane in direction x with impact parameter $y=b$. Setting $\mathrm{d}\ell^2 = \delta_{ij}\mathrm{d}x^i\mathrm{d}x^j$ and interpreting the d as differential, the <u>coordinate</u> speed of the photon can be quckly computed from the condition $g=0$,*

$$v = \frac{d\ell}{dt} = c\left(\frac{1-2\phi}{1+2\phi}\right)^{1/2} \approx c(1-2\phi) = c\left(1-\frac{2GM}{c^2r}\right) \ , \tag{7.26}$$

where the square root was expanded in $\phi \ll 1$. The speed of light measured in these non-inertial coordinates decreases the closer the photon approach to the mass. This is analogous to a wave front passing through a medium in which the speed of the wave varies with position. Hence, a beam of light rays is bent towards the mass the closer is to the mass. The deflection can be calculated using Huygen's principle in analogous way to the refraction angle of waves in a medium,

$$\frac{d\theta}{dx} = \frac{1}{c}\frac{dv}{dy} = \frac{2GM}{c^2}\frac{y}{(x^2+y^2)^{3/2}} \quad \Rightarrow \quad \theta = \frac{2GM}{c^2}\int\frac{bdx}{(x^2+b^2)^{3/2}} = \frac{4GM}{c^2b} \ . \tag{7.27}$$

Details and calculations are left as [exercise]. The same solution can be found considering null geodesics, e.g. (Carroll, 1997), and also by starting from the Schwarzschild metric in isotropic coordinates. Note that one can also perform a calculation using Newton gravity (the acceleration does not depend on the photon mass...) and find that GR result is twice the Newtonian prediction.

Observations conducted by Eddington in 1916 and others later, indicate that the Sun deflects photons of an angle $\theta_d \approx 8.5\times10^{-6}$radians (1.75 arcsec) in agreement with the GR formula above ($M_\odot = 1.98847\times10^{33}$ g and $R_\odot = 6.960\times10^{10}$ cm). Accurate measurements of light deflection are also available from the 60s using radio interferometers and astrophysical sources called blazars [Cf. gravitational lensing].

7.4.2 No-stresses source

A matter distribution with mass-energy density current vector J^μ and no stresses is modeled by a stress-energy tensor in the form

$$T_{\mu\nu} = 2J_{(\mu}t_{\nu)} - 2\rho t_\mu t_\nu \ \ \text{i.e.} \ \ T_{0\mu} = cJ_\mu \ , \quad T_{ij} = 0 \ . \tag{7.28}$$

Note that $J^\mu = \rho(W, Wc^{-1}v^i)$, thus the static case considered above is equivalent to consider the slow velocity limit $v/c \ll 1$ for the (nonrelativistic) source. Similarly, taking $T_{ij} = 0$ is equivalent to neglect velocity terms $\mathcal{O}(1/c^2)$ in the source motion.

The linearized equations read

$$\begin{cases} \Box\bar{h}_{0\mu} & = -16\pi T_{0\mu} \ , \quad \mu = 0, ..., 4 \\ \Box\bar{h}_{ij} & = 0 \ , \quad i, j = 1, ..., 3 \ . \end{cases} \tag{7.29}$$

If the spatial component of the metric are assumed time-indpendent, then they are solution of the boundary value problem with the Poisson equation and boundary values $\bar{h}_{ij}|_{r\to\infty} = 0$. This implies they are zero:

$$\partial_t \bar{h}_{ij} = 0 \quad \Rightarrow \quad \begin{cases} \Delta\bar{h}_{ij} = 0 \\ \bar{h}_{ij}|_{r\to\infty} = 0 \end{cases} \quad \Rightarrow \quad \bar{h}_{ij} = 0 \ . \tag{7.30}$$

The linearized equations reduce to those for $\bar{h}_{0\mu}$, that are formally the Maxwell equations in Lorentz gauge for the field

$$A_\mu := -\frac{1}{4}\bar{h}_{0\mu} = -\frac{1}{4}\bar{h}_{\mu\nu}t^\nu \ . \tag{7.31}$$

Once a solution is found, the metric is given by [exercise]

$$g_{00} = -1 + 2A_0 \ , \quad g_{0i} = 4A_i \ , \quad g_{ij} = 1 + 2A_0\delta_{ij} \ , \tag{7.32}$$

If one further assumes that $\bar{h}_{0\mu}$ is time-independent a formal solution can be obtained with the usual Green function method for the Poisson equation

$$\partial_t\bar{h}_{0\mu} = 0 \quad \Rightarrow \quad \begin{cases} A_0 & = -\phi \\ A_i & = \int d^3x' \dfrac{T_{0i}(x')}{|\vec{x} - \vec{x}'|} \end{cases} . \tag{7.33}$$

The above expression indicate that $A_0 = \mathcal{O}(1/c^2)$ and $A_i = \mathcal{O}(1/c)$. Reintroducing the factors c, the metric reads

$$g = -c^2\left(1 + 2\frac{\phi}{c^2}\right)\mathrm{d}t^2 + 4cA_i\mathrm{d}x^i\mathrm{d}t + \left(1 - \frac{2\phi}{c^2}\right)\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j \ . \tag{7.34}$$

Using the same formulas as in electrodynamics, one then defines from A_μ the *gravitoelectric* and *gravitomagnetic* fields and the geodesics EOM reduce to those of a particle subject to the (gravitational) Lorentz force in the small velocity limit. Consider the Lagrangian for a particle in the weak metric Eq. (7.34) and expand in v/c [1] :

$$L == -mc\sqrt{-g_{\mu\nu}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}} = -mc\sqrt{c^2(1 - 2A_0) - 2\cdot 4cA_iv^i - (1 + 2A_0)\delta_{ij}v^iv^j} \tag{7.35a}$$

$$= -mc^2\sqrt{1 - 2A_0 - 8A_i\frac{v^i}{c} - \frac{v^jv_j}{c^2} + 2A_0\frac{v^jv_j}{c^2}} = -mc^2\sqrt{1 + \underbrace{\frac{2\phi}{c^2} - 8A_i\frac{v^i}{c} - \frac{v^jv_j}{c^2}}_{\mathcal{O}(1/c^2)} + \frac{2\phi v^jv_j}{c^4}} \tag{7.35b}$$

$$\approx -mc^2 + \frac{m}{2}v^2 - m\phi + 4mcA_iv^i \ . \tag{7.35c}$$

The above equation implies an EOM with the (gravitational) Lorentz force,

$$\ddot{\vec{x}} = \vec{E} + 4\vec{v}\times\vec{B} \ . \tag{7.36}$$

The difference w.r.t. the EOM for a charge particle are that (i) there is no charge; (ii) there is a factor "4" in front of the gravitomagnetic field.

Example 7.4.2. Lense-Thirring effect. *The spacetime of a "weaky" gravitating planet or star in slow rotation is described by the metric of Eq. (7.32). The grav. field is stationary and the motion of a test body in such field is precisely described by the gravitoeletric/magnetic equations. For example, the precession motion of a gyroscope due to the gravitomagnetic field of the rotating object is precisely given by the spin-precession formula of electromagnetism*

$$\frac{d\vec{s}}{dt} = \vec{\mu}\times\vec{B} = \frac{q}{2m}\vec{s}\times\vec{B} = \vec{s}\times\vec{\Omega} \ , \tag{7.37}$$

where $\vec{\Omega} = -q/m\vec{B}$, q is the particle charge and $\vec{\mu}$ the magnetic moment. The formal substitution

$$q \mapsto m \ , \quad \vec{B} \mapsto 4\vec{B}_g \ , \tag{7.38}$$

[1] Recall that $x^0 = ct$ and $v^i = dx^i/dt$.

maps the eletrogmagnetic problem top the gravitomagnetic one. The precession frequency in the latter is thus $\Omega_g = -2B_g$. *In the grav. field of Earth one obtains that a free-falling body at distance* r *acquires an angular velocity*

$$\Omega = \dot{\varphi} \sim 0.22''/yr \left(\frac{R_\oplus}{r}\right)^3 . \tag{7.39}$$

This effect is called also frame dragging, *and it has been measured by the satellite mission Gravity Probe B in 2004 with 20% confidence.*

An extreme frame dragging phenomenon happens around rotating black holes (But it is not weak field, Eq. (7.32) do not apply). Particles close to the black hole horizon are dragged around at a speed comparable to the hole's rotation $\Omega \sim \Omega_{BH}$. *The Lense-Thirring effect in strong field play an important role to understand high-energy particle emission from matter accreting onto black holes.*

7.5 Gravitational waves (GWs) propagation

The linearized EFE in vacuum are homogeneous wave equations for each component of the metric,

$$0 = \Box_\eta \bar{h}_{\mu\nu} = \eta^{\alpha\beta}\partial_\alpha\partial_\beta \bar{h}_{\mu\nu} . \tag{7.40}$$

Solutions to the above equation can be constructed by superposition of plane waves with (constant) wave vector $k^\mu = (\omega, \vec{k})$ [2] and amplitudes $A_{\mu\nu}$:

$$\bar{h}_{\mu\nu} = A_{\mu\nu} \exp\left(\mathrm{i}k_\mu x^\mu\right) = A_{\mu\nu} \exp\left[\mathrm{i}(-k_0 x^0 + \vec{k}\cdot\vec{x})\right] = A_{\mu\nu} \exp\left[\mathrm{i}(-\omega t + \vec{k}\cdot\vec{x})\right] \tag{7.41a}$$
$$\partial_\mu \bar{h}_{\alpha\beta} = A_{\alpha\beta}\partial_\mu(\exp\left(\mathrm{i}k_\rho x^\rho\right)) = \bar{h}_{\alpha\beta}\partial_\mu(\mathrm{i}k_\rho x^\rho) = \mathrm{i}\bar{h}_{\alpha\beta}k_\rho\delta^\rho_\mu \tag{7.41b}$$

Substituting the plane-wave ansatz into the wave equations, one finds immediately that the the *wave vector is null* (Mikowski metric):

$$0 = \Box\bar{h}_{\alpha\beta} \overset{p.w.}{=} -\eta^{\mu\nu}k_\mu k_\nu \bar{h}_{\alpha\beta} \quad\Rightarrow\quad \eta^{\mu\nu}k_\mu k_\nu = k_\mu k^\mu = 0 \quad\Rightarrow\quad \omega^2 = |\vec{k}|^2 c^2 . \tag{7.42}$$

The last equation is the *dispersion relation* for GWs and indicates GWs propagate at the speed of light. Another way to see this is to consider the worldline of a photon moving along k^μ, and observe that it moves in phase with the wave's phase $\varphi = k_\mu x^\mu$:

$$x^\mu(\lambda) = k^\mu\lambda + x^\mu(0) \quad\Rightarrow\quad k_\mu x^\mu(\lambda) = \underbrace{k_\mu k^\mu}_{=0}\lambda + k_\mu x^\mu(0) = k_\mu x^\mu(0) = const . \tag{7.43}$$

Transverse-traceless (TT) gauge and physical degrees of freedom. Linearized EFE in vacuum are 10 equations. Imposing the Hilbert gauge (4 equations) reduces the problem to $10 - 4 = 6$ equations, but there remains freedom in the choice. Looking at Eq. (7.16), it is immediate to see that the Hilbert gauge is defined up to 4 harmonic functions ξ^μ: any infinitesimal transformation such that $\Box\xi_\nu = 0$ maintains $-\partial^\alpha\bar{h}_{\mu\alpha} = 0$. Let us further fix this gauge choice.

The Hilbert gauge Eq. (7.14) translates into the 4 equations that imply the waves are transverse to the direction of propagation:

$$0 = -\partial^\alpha\bar{h}_{\mu\alpha} = \mathrm{i}k^\mu A_{\mu\nu}\exp\left[\mathrm{i}(k_\rho x^\rho)\right] = \mathrm{i}k^\mu\bar{h}_{\mu\alpha} \quad\Rightarrow\quad k^\mu A_{\mu\nu} = 0 . \tag{7.44}$$

The additional gauge freedom can be fixed by

(i) writing the harmonic function as (note this clearly solves $\Box\xi_\nu = 0$)

$$\xi^\mu = B^\mu \exp\left[\mathrm{i}(k_\rho x^\rho)\right] , \tag{7.45}$$

such that the transformation of the trace reverse metric translates into the following transformation of the plane-wave amplitude,

$$\bar{h}_{\mu\nu} \mapsto \bar{h}_{\mu\nu} + 2\partial_{(\mu}\xi_{\nu)} - \eta_{\mu\nu}\partial_\alpha\xi^\alpha \quad\Rightarrow\quad A_{\mu\nu} \mapsto A_{\mu\nu} - \mathrm{i}2k_{(\mu}B_{\nu)} + \mathrm{i}\eta_{\mu\nu}k_\rho B^\rho . \tag{7.46}$$

(ii) and then fixing the functions B^μ requiring the additional 4 conditions

$$\begin{cases} \bar{h} = 0 = A^\mu_\mu & \text{Traceless condition} \\ \bar{h}_{0\mu} = 0 = A_{0\mu} & \text{Transverse condition .} \end{cases} \tag{7.47}$$

The above conditions give a linear algebraic system for B^μ, which can be inverted to find the solution. The gauge above is called *traceless & transverse (TT) gauge*. The remaining degrees of freedom are $10 - 4 - 4 = 2$, that represent the two physical states of gravity waves.

[2] Note that in general the wave frequency measured by an observer of 4-velocity u^μ is $\omega = k_\mu u^\mu$.

Let us explicitely consider a wave propagating along the $\hat{z}$-direction,

$$k^{\mu} = (\omega, 0, 0, k_3) \ . \tag{7.48}$$

Then

1. Null condition, $k_\mu k^\mu = 0 \Rightarrow -k_3 = \omega$;
2. Phase, $k_\rho x^\rho = \omega(t - z)$;
3. Hilbert gauge, $k^\mu A_{\mu\nu} = 0 \Rightarrow k^0 A_{0\nu} + k^3 A_{3\nu} = \omega A_{0\nu} - \omega A_{3\nu} = 0$ or $A_{0\nu} = A_{3\nu}$;
4. Transverse condition, $A_{0\mu} = 0 \Rightarrow A_{3\nu} = 0$;
5. Traceless condition, $-A_{00} + A_{11} + A_{22} + A_{33} = 0$;

Putting things together (1.-3.) one gets the first and last rows/cols of the amplitude matrix components are zero (Note it is symmetric)

$$A_{\mu\nu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{11} & A_{12} & 0 \\ 0 & A_{12} & A_{22} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{7.49}$$

The trace condition then is $A_{22} = -A_{11} =: A_+$. Setting $A_+ := A_{11}$ and $A_\times := A_{12}$, the plane-wave solution in TT gauge can be written

$$h^{\rm TT}_{\mu\nu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & A_+ & A_\times & 0 \\ 0 & A_\times & -A_+ & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \exp\left[{\rm i}\omega \left(t - \frac{z}{c} \right) \right] \ . \tag{7.50}$$

The two *polarization of the GW* are indicated as "+" and "×"

$$h_+(t - z/c) = A_+ \exp\left[{\rm i}\omega\,(t - z/c)\right] \ , \quad h_\times(t - z/c) = A_\times \exp\left[{\rm i}\omega\,(t - z/c)\right] \ . \tag{7.51}$$

Observations.

- The TT gauge can defined only in vacuum, because in case matter is present $\Box \bar{h}_{\mu\nu} \neq 0$ and, while there is still the freedom to rescale the $\bar{h}_{\mu\nu}$ with an infinitesimal coordinate transformation generated by four harmonic functions, we cannot set to zero the components $\bar{h}_{\mu\nu}$ inside the source.
- In the TT gauge $h_{\mu\nu} = \bar{h}_{\mu\nu}$.
- The *metric* in the TT gauge reads

$$\begin{aligned} g &= -{\rm d}t^2 + {\rm d}z^2 (1 + h_+){\rm d}x^2 + (1 - h_+){\rm d}y^2 + 2h_\times {\rm d}x{\rm d}y & (7.52)\\ &= -{\rm d}t^2 + (\delta_{ij} + h^{\rm TT}_{ij}){\rm d}x^i{\rm d}x^j \ , & (7.53) \end{aligned}$$

  where the first expression holds for a GW along the $\hat{z}$-direction and the second is general.
- Given a wave solution $\bar{h}_{\mu\nu}$ in Hilbert gauge propagating in direction $\hat{n}$, it is possible to obtain the solution in TT outside the source by means of the following projection operator (below summation on repeated indexes is understood but they are not raised)

$$\bar{h}^{\rm TT}_{\mu\nu} = \Lambda_{\mu\nu,\alpha\beta}\bar{h}_{\alpha\beta} \tag{7.54a}$$

$$\Lambda_{\mu\nu,\alpha\beta}(\hat{n}) := P_{\mu\alpha}P_{\nu\beta} - \frac{1}{2}P_{\mu\nu}P_{\alpha\beta} \tag{7.54b}$$

$$P_{\mu\nu}(\hat{n}) := \delta_{\mu\nu} - n_\mu n_\nu \ . \tag{7.54c}$$

  The following properties also hold [exercise]
  1. $P_{\mu\nu}$ is symmetric;
  2. $P_{\mu\nu}$ is transverse, $n^i P_{\mu\nu} = 0$;
  3. $P_{\mu\nu}$ is a projector, $P_{\mu\alpha}P_{\alpha\nu} = P_{\mu\nu}$;
  4. $P_{\mu\nu}$ has trace $P_{\mu\mu} = 2$;
  5. $\Lambda_{\mu\nu,\alpha\beta}$ is a projector, $\Lambda_{\mu\nu,\alpha\beta}\Lambda_{\alpha\beta,mn} = \Lambda_{\mu\nu,mn}$;
  6. $\Lambda_{\mu\nu,\alpha\beta}$ is transverse in al indexes;
  7. $\Lambda_{\mu\nu,\alpha\beta}$ is traceless in $\mu\nu$ and $\alpha\beta$,$\Lambda_{\mu\mu,\alpha\beta} = 0 = \Lambda_{\mu\nu,\alpha\alpha}$;
  8. $\Lambda_{\mu\nu,\alpha\beta}$ is symmetric in $\mu\nu - \alpha\beta$, $\Lambda_{\mu\nu,\alpha\beta} = 0 = \Lambda_{\alpha\beta,\mu\nu}$.
- More in general, for any symmetric tensor $S_{\mu\nu}$ one can obtain a symmetric, transverse and tracefree (STF) tensor using the $\Lambda_{\mu\nu,\alpha\beta}$ projector. STF tensors are a powerful tool to develop multipolar expansions of tensor satisfying wave equations, thus generalizing the multipolar expansion of the Newtonian and electrostatic potentials and scalar wave equations (Thorne, 1980; Maggiore, 2007).

7.6 Effect of GWs on test masses

A simple argument to understand the effect of GW on test masses is given by the following

Example 7.6.1. Distance measurement with the radar method. *Consider two masses p, q at rest and at distance L_0 in absence of GW. Their distance L can be measured by a time measurement by sending a light pulse from p to q, then back from q to p; the clock at p will measure*

$$L = \frac{1}{2}c(t_{p\,2} - t_{p\,1}) \ . \tag{7.55}$$

Let's take $x^i_p = (0,0,0)$ and $x^i_q = L_0 n^i$ ($\eta_{ij}n^i n^j = 1$, direction between the masses) so that $L_0 = \delta_{ij}x^i_q x^j_q$. Calculate the variation of the distance measured at a time $t_p = t_q$ when q receives the light ray in the case a GW is present:

$$L^2 = g_{\mu\nu}(x^\mu_q - x^\mu_p)(x^\nu_q - x^\nu_p) \overset{t_p=t_q}{=} g_{ij}(x^i_q - x^i_p)(x^j_q - x^j_p) \overset{x^i_p=0}{=} g_{ij}x^i_q x^j_q = (\delta_{ij} + h^{TT}_{ij})L_0^2 n^i n^j \tag{7.56a}$$

$$\Rightarrow \quad \frac{\delta L}{L_0} = \frac{L}{L_0} - 1 = \sqrt{1 + h^{TT}_{ij}n^i n^j} - 1 \approx \frac{1}{2}h^{TT}_{ij}n^i n^j \ . \tag{7.56b}$$

The distance's relative variation due to the GW is proportional to the GW amplitude.

A more formal study the effect of GW on test masses employs the geodesic deviation equation

$$u^\mu \nabla_\mu (u^\nu \nabla_\nu s^\alpha) = R^\alpha_{\nu\rho\sigma} u^\nu u^\rho s^\sigma \ , \tag{7.57}$$

where u is the tangent to the particles worldline and s is is the displacement vector between worldlines. Here, one considers the "relative motion" of nearby particles at the passage of the GW using tensorial equations.

In linearized GR, the Riemann tensor is proportional to $h_{\mu\nu}$ and a particle initially at rest (in the global inertial reference system of the backround metric) acquires a velocity due to the perturbation

$$\frac{dx^0}{d\tau} = 1 + \mathcal{O}(h) \ , \quad \frac{dx^i}{d\tau} = \mathcal{O}(h) \quad \Rightarrow \quad u^\mu = c\frac{dx^\mu}{d\tau} \simeq (1,0,0,0) + \mathcal{O}(h) \ . \tag{7.58}$$

Using this velocity and coordinate time instead of proper time, the geodesic equation at lowest order and in TT gauge reduces to

$$\frac{d^2 s^\alpha}{dt^2} = R^\alpha_{00\mu}s^\mu \ , \quad \text{with} \quad R_{\mu 00\nu} = \frac{1}{2}\frac{\partial^2 h^{\rm TT}}{\partial t^2} \ . \tag{7.59}$$

The formula above shows that the two degrees of freedom of the GW are physical: they cannot be gauged away. Restricting to the spatial indexes and calling $s^i_0 = s^i(t=0)$ the deviation vector before the GW arrives, the formula above reduces to

$$\frac{d^2\delta s^i}{dt^2} = \frac{1}{2}\ddot{h}^{\rm TT}_{ij}s^j_0 + \mathcal{O}(h^2) \ , \tag{7.60}$$

where terms $\propto \ddot{h}^{\rm TT}_{ij}s^i(t)$ can be discarded if one works at leading order in the perturbation. Because the particle are initially at rest, $\dot{s}^i(0) = 0$, the solution for the deviation vector is simply

$$s^i(t) = (\delta_{ij} + \frac{1}{2}h^{\rm TT}_{ij})s^j_0 \ . \tag{7.61}$$

Specify for a GW propagating in $\hat{z}$-direction and setting $s^i_0 = (x_0, y_0, z_0)$, one gets

$$\begin{cases} \delta x(t) &= x_0 + \frac{1}{2}(h_+(t)x_0 + h_\times(t)y_0) \\ \delta y(t) &= y_0 + \frac{1}{2}(-h_+(t)y_0 + h_\times(t)x_0) \\ \delta z(t) &= z_0 \ . \end{cases} \tag{7.62}$$

The test masses oscillates in the x-y plane, transverse to the direction of propagation of the GW. The effect of the two polarization is clearly visualized by considering a ring of test masses places in the plane, Fig. (7.1). If initially $(x_0, y_0) = r_0(\cos\phi, \sin\phi)$, the effect of the *plus* polarization is

$$\begin{cases} \delta x(t) &= x_0 + \frac{1}{2}h_+(t)x_0 = r_0(1 + h_+(t))\cos\phi \\ \delta y(t) &= y_0 - \frac{1}{2}h_+(t)y_0 = r_0(1 - h_+(t))\sin\phi \ . \end{cases} \tag{7.63}$$

Take the square of the equations and sum them up:

$$1 = \cos\phi^2 + \sin\phi = \frac{\delta x(t)^2}{r_0^2(1 + h_+(t))^2} + \frac{\delta y(t)^2}{r_0^2(1 - h_+(t))^2} \ , \tag{7.64}$$

this is an ellipsis of semi-axes $a_\pm := r_0(1 \pm h_+)$, and because $h_+(t)$ is an oscillating function with period $T = 2\pi/\omega$, one axis gets shorter and the other longer with the period T.

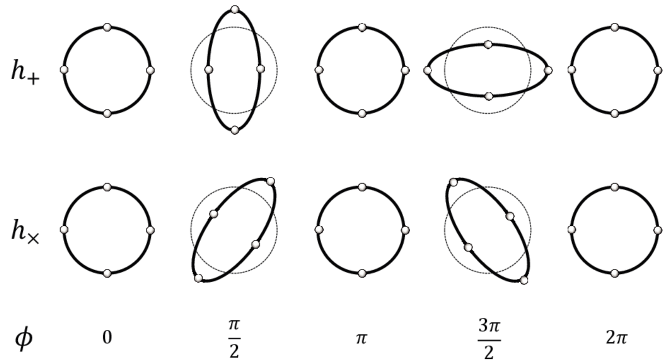

Figure 7.1: Effect on a ring of test masses on the $x-y$ plane at the passage of a GW in perpendiculr direction.

Similarly, for the *cross* polarization

$$\begin{cases} \delta x(t) & = x_0 + \frac{1}{2} h_\times(t) y_0 \\ \delta y(t) & = y_0 + \frac{1}{2} h_\times(t) x_0 \ , \end{cases} \tag{7.65}$$

one can proceed by diagonalizing the r.h.s. matrix with a rotation of $\pi/4$

$$\begin{bmatrix} 1 & h_\times/2 \\ h_\times/2 & 1 \end{bmatrix} \Rightarrow \lambda_\pm = \pm\frac{1}{2} h_\times \ , R = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}\Bigg|_{\alpha=\pi/4} . \tag{7.66}$$

In the rotated frame, the equations are identical to Eq. (7.63). Thus, the cross polarization moves the ring in the same way as the plus but for a phase of $\pi/4$.

7.7 Sources of GWs

A formal solution of the linearized EFE

$$\Box \bar{h}_{\mu\nu} = -16\pi T_{\mu\nu} \ , \tag{7.67}$$

is given in terms of the Green functions with retarded time $t_R := t - |\vec{x} - \vec{x}'|/c$ (See e.g. Jackson (1975))

$$\bar{h}_{\mu\nu}(t, \vec{x}) = -16\pi \int G_R(x^\mu - x^{\mu'}) T_{\mu\nu} d^4x' = 4 \int \frac{T_{\mu\nu}(t_R, \vec{x}')}{|\vec{x} - \vec{x}'|} d^3x' \ , \tag{7.68}$$

where G is the retaded time solution of

$$\Box_{(x)} G(x^\mu, x^{\mu'}) = \delta^{(4)}(x^\mu - x^{\mu'}) \ \Rightarrow \ G_R(x^\mu, x^{\mu'}) = -\frac{1}{4\pi} \frac{1}{|\vec{x} - \vec{x}'|} \delta(t_R - t) \ . \tag{7.69}$$

The physics picture is that the solution at point p (time t and location x^i) is determined by the events in the past lightcone. Note that since these are weak field equations, they apply to a source with negligible selfgravity $\sigma = 2GM/(c^2R) \ll 1$, where M, R are the typical mass and size of the source.

7.7.1 Quadrupole formula

Let us specify the above formula under the conditions
(i) Large distance from a compact source $r = |\vec{x}| = \sqrt{\delta_{ij} x^i x^j} \gg R$;
(ii) Slow velocity, the source motion is slow $v \sim |T_{0i}|/|T_{00}| \ll c$ and $T_{00} \approx \rho c^2$.
The hypothesis (i) implies that

$$|\vec{x} - \vec{x}'| = r|\hat{n} - \frac{\vec{x}'}{r}| = r\sqrt{(\hat{n} - \frac{\vec{x}'}{r}) \cdot (\hat{n} - \frac{\vec{x}'}{r})} = r\sqrt{1 - 2\hat{n} \cdot \frac{\vec{x}'}{r} + \left(\frac{\vec{x}'}{r}\right)^2} \approx r - \hat{n} \cdot \vec{x}' \ , \tag{7.70}$$

and restricting to the spatial indexes (those relevant in the TT gauge) Eq. (7.68) becomes

$$\bar{h}_{ij}(t, \vec{x}) \approx \frac{4}{r} \int T_{ij}(t - \frac{r}{c} + \frac{\hat{n} \cdot \vec{x}'}{c}, \vec{x}') d^3x' \ , \tag{7.71}$$

where one retains only the leading-order term at the denominator $1/|\vec{x} - \vec{x}'| \sim 1/r$.

The hypothesis (ii) allows one to expand the stress-energy tensor in $\hat{n} \cdot \vec{x}'/c$:

$$T_{kl}(t - \frac{r}{c} + \frac{\hat{n} \cdot \vec{x}'}{c}, \vec{x}') = T_{kl}(t - \frac{r}{c}, \vec{x}') + \frac{n_i x'^i}{c} \partial_t T_{kl}(u, \vec{x}') + \frac{n_i x'^i n_j x'^j}{c^2} \partial_{tt} T_{kl}(u, \vec{x}') + \ldots \ , \tag{7.72}$$

where $u := t - r/c$. A simple way to justify the above expansion from the hypothesis (ii) is to look at the expression of T_{kl} in terms of its Fourier transform,

$$T_{kl}(t - \frac{r}{c} + \frac{\hat{n}\cdot\vec{x}'}{c}, \vec{x}') = \int d^4k \tilde{T}_{kl}(\omega, \vec{k}) e^{-\mathrm{i}\omega(t-\frac{r}{c}+\frac{\hat{n}\cdot\vec{x}'}{c})+\mathrm{i}\vec{k}\cdot\vec{x}} \ , \tag{7.73}$$

and realize that the integral of T_{kl} is dominated by the slow characteristic frequencies of the source ($v \sim \Omega R$)

$$\omega\frac{|\vec{x}'|}{c} \sim \Omega\frac{R}{c} \ll 1 \ . \tag{7.74}$$

Hence, the exponential in the Fourier transform can be expanded

$$e^{-\mathrm{i}\omega(t-\frac{r}{c}+\frac{\hat{n}\cdot\vec{x}'}{c})} \simeq e^{-\mathrm{i}\omega u}\left(1 - \mathrm{i}\frac{\omega}{c}{x'}^i n_i - \frac{1}{2}\mathrm{i}^2\frac{\omega^2}{c^2}{x'}^i n_i {x'}^j n_j + ...\right) \ , \tag{7.75}$$

and the above expansion is equivalent to Eq. (7.72). The latter equation is a multipolar expansion of the of T_{kl} in Cartesian coordinates. Truncating the expansion at leading order one gets:

$$\bar{h}_{ij}(t,\vec{x}) \approx \frac{4}{r}\int T_{ij}(t - \frac{r}{c}, \vec{x}')d^3x' \ . \tag{7.76}$$

Focus now on the matter distribution and derive an equation for the integral of the spatial components T_{ij} by using the conservation law on flat background,

$$0 = \partial^\mu T_{\mu\alpha} = \eta^{\mu\nu}\partial_\nu T_{\alpha\mu} = \begin{cases} -\partial_t T_{00} + \partial_i T_{0i} & \alpha = 0 \\ -\partial_t T_{0k} + \partial_i T_{ki} & \alpha = k \ . \end{cases} \tag{7.77a}$$

Derive the $\alpha = 0$ equation in time ∂_t and substitute the $\alpha = k$ equation to obtain

$$0 = -\partial_{tt}T_{00} + \partial_t\partial_k T_{0k} = -\partial_{tt}T_{00} + \partial_l\partial_l T_{kl} \ . \tag{7.77b}$$

Multiply the above equations by $x^i x^j$ and integrate:

$$\begin{aligned}
\frac{d^2}{dt^2}\int T_{00}x^i x^j d^3x &= \int \partial_k\partial_l T_{kl}x^i x^j d^3x && (7.77c)\\
&= \int \partial_k\left(\partial_l T_{kl}x^i x^j\right)d^3x - \int \partial_l T_{kl}\partial_k(x^i x^j)d^3x && (7.77d)\\
&= \underbrace{\oint \partial_l T_{kl}y^i y^j n^k d^2y}_{=0} - \int \partial_l T_{kl}(\delta^i_k x^j + \delta^j_k x^i)d^3x && (7.77e)\\
&= -\int(\partial_l T_{il}x^j + \partial_l T_{kl}x^i)d^3x = -\int(\partial_l(T_{il}x^j) - T_{il}\underbrace{\partial_l x^j}_{=\delta^j_l} + \partial_l(T_{kl}x^i) - T_{kl}\underbrace{\partial_l x^i}_{=\delta^i_l})d^3x && (7.77f)\\
&= -\underbrace{\oint T_{il}x^j n^l d^2y}_{=0} - \underbrace{\oint T_{kl}x^i n^l d^2y}_{=0} + 2\int T_{ij}d^3x && (7.77g)
\end{aligned}$$

The surface integrals in the third and fifth lines are zero since the matter distribution is compact and there is no matter outside a sphere of radius $r > R$. The integral of the spatial components of the stress-energy tensor is thus related to the *moment of inertia* tensor of the matter distirbution ρ

$$2\int T_{ij}d^3x = \frac{1}{c^2}\frac{d^2}{dt^2}\int T_{00}x^i x^j d^3x = \frac{d^2}{dt^2}\int \rho x^i x^j d^3x =: \frac{d^2}{dt^2}I_{ij} \ . \tag{7.77h}$$

Putting together Eq. (7.76) and Eq. (7.77c), one obtains

$$\bar{h}_{ij}(t,\vec{x}) = \frac{2G}{c^4 r}\ddot{I}_{ij}(t - \frac{r}{c}) \ . \tag{7.78}$$

Far from the source, the solution can be projected to the TT gauge to obtain the *quadrupole formula*

$$\bar{h}^{\mathrm{TT}}_{ij}(t,\vec{x}) = \frac{2G}{c^4 r}\Lambda_{ijkl}\ddot{Q}_{kl}(t - \frac{r}{c}) \ , \tag{7.79}$$

where the moment of inertia can be substituted by its traceless version, the *quadrupole moment*:

$$Q_{ij} = I_{ij} - \frac{1}{3}(\underbrace{\delta^{kl}I_{kl}}_{=I})\delta_{ij} = \int \rho(x^i x^j - \frac{1}{3}\vec{x}\cdot\vec{x}\delta_{ij})d^3x \ . \tag{7.80}$$

The quadrupole formula gives the leading order contribution to the GW from a (spatially) compact, slowly moving and nonselfgravitating source. The expansion starts at the quadrupolar order ($\ell = 2$) because of mass and momentum conservation. Indeed, integrating T_{00} and using the conservation law for $T_{\mu\nu}$ one finds immediately $\dot{M} = 0$ and similarly integrating $T_{00}x^i$ gives momentum conservation $\dot{P}^i = 0$. Note that a generic source is decomposed in a infinite series of multipoles.

Remark 7.7.1. *The quadrupole momentum is the tensor that appears in the multipolar expansion of the Newtonian potential*

$$\phi(t, \vec{x}) = -\frac{GM}{r} + \frac{3GQ_{ij}(t)n^i n^j}{2r^3} + \ldots \, . \tag{7.81}$$

The dipolar term (multipole with "one index") is just the center of mass vector that can be removed by using the center of mass frame. The quadrupole is the lowest multipole described by a tensor with 2 indexes. A similar expansion hold for the electrostatic potential, but in that case the dipole cannot be removed and represents the next-to-leading order approximation of the charge distribution.

Using dimensional analysis, $[\ddot{Q}] = ML^2T^{-2}$ and one immediately finds that the GW observed at distance D from the source is

$$h \sim \left(\frac{G}{c^4 D}\right)(Mv^2) = \left(\frac{R}{D}\right)\left(\frac{GM}{c^2 R}\right)\left(\frac{v}{c}\right)^2 \, . \tag{7.82}$$

The formula above indicates thet GW are produced by physical objects that are

- very compact;
- strongly gravitating;
- rapidly moving.

Note that the quadrupole formula does not apply for those objects!

Example 7.7.1. Quadrupole formula for a binary star system. *Consider two masses m_1 and m_2 separated by $\vec{r} = \vec{x}_1 - \vec{x}_2$ in Newtonian gravity. The mass density is $\rho = m_1\delta(\vec{x} - \vec{x}_1) + m_2\delta(\vec{x} - \vec{x}_2)$, the total mass is $m = m_1 + m_2$ and the reduced mass is $\mu = m_1 m_2/m$. The moment of inertia reads*

$$I^{ij} = \int \rho x^i x^j d^3x = m_1 x_1^i x_1^j + m_2 x_2^i x_2^j = m x_{\rm cm}^i x_{\rm cm}^j + \mu r^i r^j \tag{7.83}$$

where $x_{\rm cm}^i = (m_1\vec{x}_1 + m_2\vec{x}_2)/m$ is the center of mass coordinate. In the center of mass frame $I^{ij} = \mu r^i r^j$ and the trace is $I = \mu\delta_{ij}r^i r^j = \mu r^2$. The quadrupole of the 2-body system is thus

$$Q^{ij} = I^{ij} - \frac{1}{3}I\delta^{ij} = \mu(r^i r^j - \frac{1}{3}r^2\delta^{ij}) \, . \tag{7.84}$$

Specialize now for a circular orbit in the $z = 0$ plane, for which $\vec{r} = (x, y, 0) = R(\cos(\Omega t + \pi/2), \sin(\Omega t + \pi/2), 0)$ (the pi/2 phase factors are there for later convenient) and

$$\Omega = \frac{2\pi}{T} = \left(\frac{Gm}{R^3}\right)^{1/2} \tag{7.85}$$

is the orbital frequency as given by Kepler law [3]*. The inertial moment for circular orbits is immediately calculated, the nonzero components are*

$$\begin{cases} I^{11} &= \mu R^2\cos^2(\Omega t + \pi/2) = \mu R^2\frac{1+\cos^2(2\Omega t+\pi)}{2} = \mu R^2\frac{1-\cos^2(2\Omega t)}{2} \\ I^{22} &= \mu R^2\sin^2(\Omega t + \pi/2) = \mu R^2\frac{1-\cos^2(2\Omega t+\pi)}{2} = \mu R^2\frac{1+\cos^2(2\Omega t)}{2} \\ I^{12} &= \mu R^2\cos(\Omega t + \pi/2)\sin(\Omega t + \pi/2) = \mu R^2\frac{1}{2}(\sin(2\Omega t + \pi) + \sin(0)) = -\mu R^2\sin(2\Omega t) \, . \end{cases} \tag{7.86}$$

Taking the derivatives

$$\begin{cases} \ddot{I}^{11} &= 2\mu R^2\Omega^2\cos(2\Omega t) \\ \ddot{I}^{21} &= 2\mu R^2\Omega^2\sin(2\Omega t) \\ \ddot{I}^{22} &= -\ddot{I}^{11} \, . \end{cases} \tag{7.87}$$

From the quadrupole formula one concludes that GW are emitted at frequency $\Omega_{gw} = 2\Omega$. Note this comes from the term $x^i x^j \sim \cos^2(\Omega t)$, hence any monocromatic source emits at 2Ω as a consequence of the quadrupole nature of the GW. Performing the TT projection (first expressions omits constants) one gets (Maggiore, 2007)

$$\begin{cases} h_+ &= \frac{1}{r}(\ddot{I}_{11} - \ddot{I}_{22}) = \frac{G}{rc^2}4\mu R^2\Omega^2\frac{1+\cos^2\theta}{2}\cos(2\Omega t_R + \varphi) \\ h_\times &= \frac{2}{r}\ddot{I}_{12} = \frac{G}{rc^2}4\mu R^2\Omega^2\cos\theta\sin(2\Omega t_R + \varphi), \end{cases} \tag{7.88}$$

where (θ, φ) are the sky location of the source (related to the direction $\hat{n}$ in the STF projector).

[3] Kepler law can be derived from dimensional analysis of the quantities Ω, Gm, R or equating the grav.force to the centripetal force.

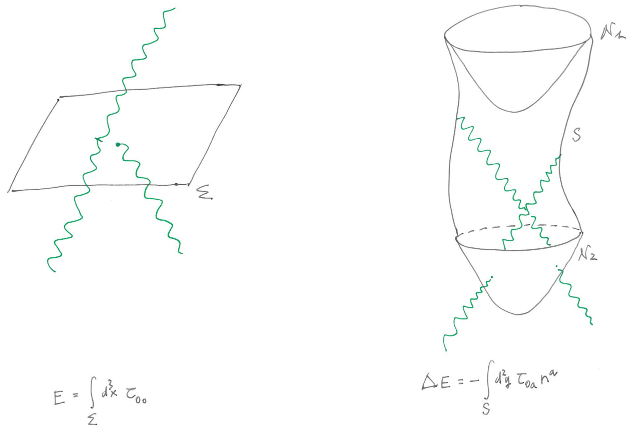

Figure 7.2: Representation of the notion of global energy of a spacelike hypersurface and of energy carried by the GW in asymptotically flat spacetime.

7.8 Energy of GWs

The physical reality of GW has been investigated by Einstein already in 1916 and was finally established in the 60s with the fundamental work of Bondi, Goldberg, Newmann, Penrose, Pirani, Robison, Sachs, Trautman and many others. Note this was theoretical work aiming at clarifying GWs are not coordinate effects and can transport energy [4] Experimental evidence for GWs came afterwards with the Taylor & Hlse pulsar observation starting 1974.

Remark 7.8.1. *The GR definition of energy of the grav. field and energy of the GW is rather complex.*

- *In general, there is* no local definition of energy (density) for the gravitational field *since the metric describes the whole spacetime and cannot be decomposed into a background component (to be used as "reference") and a dynamical component (actually "carrying the energy").*
- *However a notion of* total energy of the spacetime *can be constructed for a class of spacetimes called* asymptotically flat*, and describing the spacetime of isolated systems. This is an advanced topic [Chap. 11 of (Wald, 1984)], but one might suspect a total energy-momentum could be defined for isolated systems as far away from them there is a natural background spacetime that can be usedm, see Fig. (7.2).*
- *Indeed, the Hamiltonian formulation of GR by Arnowitt, Deser e Misner (ADM, 1959) gives a notion of energy for asymptotically flat spacetimes. The ADM energy-momentum can be thought as the energy of a spacelike hypersurface at a given "time" and associated to time translation and boost about* spatial infinity *(*Spi *group). An example of such energy is provided by the mass of the Schwarzschild spacetime (Chap. 8.)*
- *Similarly, a notion of energy-momentum of GW can be given considering asymptotic null hypersurfaces at a given "retarded time". This energy-momentum is the one carried by the gravitational radiation and it is associated to the symmetries of* null infinity *in asymptotically flat spacetimes (*Bondi-Metzner-Sachs (BMS) *group).*

The following discusses how this notions of energy can be defined in linearized gravity, where a flat background metric is present. Considering an isolated system such that asymptotically $g \sim \eta$, one expects

- The energy to be quadratic in the perturbation, so one must consider 2nd order perturbations;
- Any form of energy must generate curvature throughout a stress-energy tensor;
- An energy definition must be gauge invariant.

We thus look for an effective stress-energy tensor that is quadratic in the perturbation and leads to a gauge invariant energy definition. Focus is on vacuum GR. Push the Minkowski expansion to second order:

$$g = \eta + h^{(1)} + \mathcal{O}(2) \tag{7.89a}$$

$$g = \eta + h^{(1)} + h^{(2)} + \mathcal{O}(3) \ , \tag{7.89b}$$

where the notation $h^{(n)}$ indicates that in the global inertial coordinate system the tensor components are

$$|h^{(1)}_{\mu\nu}| \approx \epsilon |\eta_{\mu\nu}| \approx \epsilon \ll 1 \ , \quad |h^{(2)}_{\mu\nu}| \approx \epsilon^2 \ , \quad \text{etc.} \tag{7.89c}$$

[4]The famous conference Chapel Hill in 1957 where this topic was debated triggered the Weber experimental work and also led to a famous thought experiment `https://en.wikipedia.org/wiki/Sticky_bead_argument`.

The Ricci tensor admits a similar expansion

$$R_{\mu\nu} = R^{(0)}_{\mu\nu} + R^{(1)}_{\mu\nu} + R^{(2)}_{\mu\nu} + \mathcal{O}(3) \quad \text{where} \tag{7.90a}$$

$$R^{(0)}_{\mu\nu} = (\mathrm{Ric}[\eta])_{\mu\nu} \sim \eta\partial^2\eta \tag{7.90b}$$

$$R^{(1)}_{\mu\nu} = \left(\mathrm{Ric}^{(1)}[h^{(1)}]\right)_{\mu\nu} \sim \eta\partial^2 h^{(1)} \tag{7.90c}$$

$$R^{(2)}_{\mu\nu} = \left(\mathrm{Ric}^{(1)}[h^{(2)}]\right)_{\mu\nu} + \left(\mathrm{Ric}^{(2)}[h^{(1)}]\right)_{\mu\nu} \sim \eta\partial^2 h^{(2)} + h^{(1)}\partial^2 h^{(1)} \ , \tag{7.90d}$$

where Ric[.] indicates the full expression of the Ricci applied to its argument, $\mathrm{Ric}^{(1)}[.]$ indicates the expression of the Ricci linearized in the metric perturbation applied to its argument, $\mathrm{Ric}^{(2)}[.]$ indicates the expression of the Ricci quadratic in the metric perturbation applied to its argument, etc. For example, the expression for $\mathrm{Ric}^{(1)}$ is given by Eq. (7.9c), while

$$R^{(2)}_{\mu\nu} = \frac{1}{2}h^{\rho\sigma}\partial_\mu\partial_\nu h_{\rho\sigma} - h^{\rho\sigma}\partial_\rho\partial_{(\mu}h_{\nu)\sigma} + \partial^\sigma h^\rho_\nu\partial_{[\sigma]}h_{\rho]\mu} + \frac{1}{2}\partial_\sigma(h^{\rho\sigma}\partial_\rho h_{\mu\nu}) - \frac{1}{4}\partial_\rho h_{\mu\nu}\partial^\rho h - (\partial_\sigma h^{\rho\sigma} - \frac{1}{2}\partial^\rho h)\partial_{(\mu}h_{\nu)\rho} \tag{7.91}$$

The $\mathcal{O}(2)$ term is clearly composed of a term coming from the linearized Ricci applied to the second order metric perturbation $h^{(2)}$ plus a term coming from the second order Ricci applied to the linear perturbation $h^{(1)}$. The EFE in vacuum,

$$0 = R_{\mu\nu} = R^{(0)}_{\mu\nu} + R^{(1)}_{\mu\nu} + R^{(2)}_{\mu\nu} + \mathcal{O}(3) \tag{7.92a}$$

could now be solved hierarchically order-by-order as follows

$$0 = R^{(0)}_{\mu\nu} = (\mathrm{Ric}[\eta])_{\mu\nu} \quad \Rightarrow \quad \text{Trivially satisfied by } \eta \tag{7.92b}$$

$$0 = R^{(1)}_{\mu\nu} = (\mathrm{Ric}^{(1)}[h^{(1)}])_{\mu\nu} \quad \Rightarrow \quad h^{(1)} \quad \text{(up to gauge)} \tag{7.92c}$$

$$0 = R^{(2)}_{\mu\nu} = (\mathrm{Ric}^{(1)}[h^{(2)}] + \mathrm{Ric}^{(2)}[h^{(1)}])_{\mu\nu} \quad \Rightarrow \quad h^{(2)} \quad \text{(up to gauge)}$$
$$\text{etc.} \tag{7.92d}$$

The last equation above can be written as a Einstein equation for the second-order metric as

$$G^{(1)}_{\mu\nu}[h^{(2)}] := \left(\mathrm{Ric}^{(1)}[h^{(2)}] - \frac{1}{2}\mathrm{R}^{(1)}[h^{(2)}]\eta\right)_{\mu\nu} = 8\pi\tau_{\mu\nu} := 8\pi(-G^{(2)}_{\mu\nu}[h^{(1)}]) = -8\pi\left(\mathrm{Ric}^{(2)}[h^{(1)}] - \frac{1}{2}\mathrm{R}^{(2)}[h^{(1)}]\eta\right)_{\mu\nu} , \tag{7.93}$$

where a quadratic-in-h stress-energy tensor constructed from the linear perturbation metric is defined. Note that $\tau_{\mu\nu}$ is

+ Symmetric;
+ Conserved on flat backrground $\partial_\mu\tau_{\mu\nu} = 0$ (Bianchi identities);
+ Quadratic in h, $\mathcal{O}(2)$;
- Not gauge invariant;
- Note unique, as $\tau_{\mu\nu}$ is defined up to a term $\partial^\alpha\partial^\beta U_{\mu\nu\alpha\beta}$ with $U_{\mu\nu\alpha\beta} = \mathcal{O}(2)$ such that $U_{\mu\nu\alpha\beta} = U_{\mu\alpha[\nu\beta]} = U_{[\mu\alpha]\nu\beta} = U_{\nu\beta\mu\alpha}$.

For an asymptotically flat metric satisfying the conditions

$$h^{(1)} \sim \mathcal{O}(1/r) \ , \quad \partial h^{(1)} \sim \mathcal{O}(1/r^2) \ , \quad \partial\partial h^{(1)} \sim \mathcal{O}(1/r^3) \quad \text{for} \quad r \to \infty \ , \tag{7.94}$$

it can be proven that the quantity

$$E := \int_\Sigma d^3x t_{00} \tag{7.95}$$

is <u>gauge invariant</u> under infinitesimal transformations preserving the asymptotically flat conditions,

$$E[h^{(1)}_{\mu\nu}] = E[h^{(1)}_{\mu\nu} + 2\partial_{(\mu}\xi_{\nu)}] \ . \tag{7.96}$$

It is also <u>unique</u> in the sense that it does not change if a term $\partial^\alpha\partial^\beta U_{\mu\nu\alpha\beta}$ is added to the definition of τ_{00}. Note Σ is the 3D <u>spatial</u> hypersurface defined by $t = const$, Fig. (7.2), and that the asymptotic flat conditions guarantee the existance of the integral. Hence, the quantity E can be taken as the total energy associated to the linearly perturbed metric.

In order to define the radiated energy one proceed similarly but considering a situation in which the spacetime is initially time-independent, then go through a time-dependent phase, say between $t_1 < t < t_2$, and then is again time-independent. In the stationary phases $t < t_1$ and $t > t_2$ one takes two 3D surfaces that are asymptotically <u>null</u> and call them $\mathcal{N}_1$ and $\mathcal{N}_2$. If the radiation is measured on $\mathcal{N}_1$ but not on $\mathcal{N}_2$, then it escaped at null infinity. Hence, the energy radiated between the two stationary regime is defined considering the integral of the "density flux" $-\tau_{0\mu}$ over the asymptotically *timelike* surface S between $\mathcal{N}_1$ and $\mathcal{N}_2$ Fig. (7.2)

$$\Delta E := -\int_S \tau_{a0} n^a d^2y \ . \tag{7.97}$$

The asymptoticaly flat conditions Eq. (7.94) are now required on these null hypersurfaces [5]. They guarantee that the

[5] For example, one can think of using the retarded time coordinate $u = t - r$ and taking $r \to \infty$.

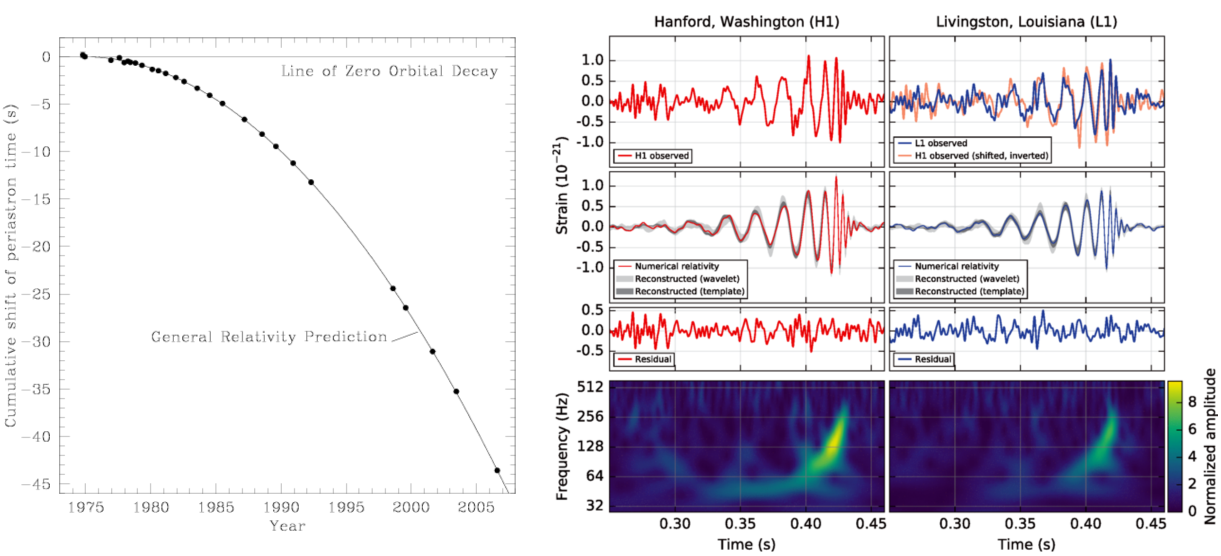

Figure 7.3: Observations of GWs. Left: The decay of the orbital period of the binary system observed in PSR B1913+16; Right: The binary black hole merger waveform GW150914 observed by LIGO in 2015.

integral exists and that the definition is meaningful. Note that the asymptotically flat condition **cannot** be imposed in the nonstationary phase because a wave behaves as $\sim h(t-r)/r$ and its derivative is $\partial_i h \sim \mathcal{O}(1/r)$. The integrand on the r.h.s. of the above equation gives the power (luminosity) for the GW since $\Delta E = \int \dot{E} dt$ (see below).

7.9 GW observations

Observational evidence for GW was found starting 1974 with the measurements of radio signals from the pulsar PSR B1913+16. The source is a binary system in our galaxy at 21000 light years (6400 pc) made of two neutron stars in which one of them is a pulsar emitting period radio pulses. The variation in the time arrival of the pulses allows the identification of the source as a binary system and a precise measurement of the period (~ 7.75 hrs) and the masses. The observations performed during the decades have proven that the period of the orbit decay as predicted by GR due to the emission of GWs Fig. (7.3). Several of these systems have been identified to date and they provide stringent tests of the prediction of GR.

Direct measurements of GW have been possible starting 2015 using gravitational-wave interferometric techniques. The first detection of a GW propagating through the Earth was obtained by the LIGO experiment on 14th September 2015. The source has been identified as a collision (merger) of black holes in circular orbits as predicted by GR and as calculated by means of numerical relativity simulations. The binary was a distance of $\sim$440 Mpc (redshift 0.09) and the two masses of $\sim 35 + 30\ \mathrm{M}_\odot$ formed a black hole of $\sim 62\ \mathrm{M}_\odot$ emitting in GW an energy of $\sim 3\ \mathrm{M}_\odot c^2$. Several detections of GW from mergers of black hole and one neutron star binary merger have been reported since then by the LIGO-Virgo experiments.

7.10 Short-wavelength approximation

The definition of GW energy via $\tau_{\mu\nu}$ is based on the idea that the GW generate curvature. This approach has a conceptual problem if one desires to push the expansion beyond linear order as the Mikowski background has zero curvature.

The above concepts of wave propagations on metric background and the definition of $\tau_{\mu\nu}$ can be generalized to an arbitrary background metric. In particular, one could be interested in defining perturbations or GWs on a background which is curved and dynamical. While in general it is not possible to split background from perturbation metric, in practise, many problems have a clear separation of scales and admit such decomposition. Consider a metric that, in some coordinate, has a typical spatial scale of variation L, then any small amplitude perturbations with wavelength $\lambdabar \ll L$ could be clearly distinguished and separated from the background. Similarly, if the background has a temporal variation up to frequency F, then a small amplitude perturbation with frequencies $f \gg F$ could be clearly distinguished/separated.

Remark 7.10.1. *While the frequency and wavelength of a wave are related by $\lambda = c/f$, the temporal and spatial scales of variation of the background metric are **not** necessarily related. For example, consider a GW with $f \sim 10^2 = 10^3$ Hz, $\lambdabar \sim 500-50$ km and $|h| \sim 10^{-21}$ in the grav. field of Earth $\phi_\oplus \sim GM_\oplus/R_\oplus c^2 \sim 10^{-6}$. $\phi_\oplus$ is not smooth at lengthscales $\lambdabar$ because it has variations of amplitudes $\sim 10^{-9} \gg |h|$ due to e.g. mountains; moreover the length of the laboratory apparatus to detect GW is $\ll \lambda$. It is not possible to separate the length scales of the Earth's grav. field and the GW!*

However, $\phi_\oplus$ is static with frequency $\ll f$. In this sense the grav. field of Earth can be separated from the GW. [GW experiments on Earth measure temporal, not lenght variations]

Let us address the problem of defining the propagation of the short-wavelength perturbation on the background and the generalization of the GW effective stress-energy tensor. Note there are two perturbation parameters:
- the small amplitude $\epsilon \ll 1$
- the wavelength $\lambda\!\!\!^{-} \ll L$, or wavenumber $k = 1/\lambda\!\!\!^{-} \gg 1/L$, of the perturbation.

Starting from the trace reverse EFE

$$R_{\mu\nu} = 8\pi(T_{\mu\nu} - \frac{1}{2}T g_{\mu\nu}) =: 8\pi\bar{T}_{\mu\nu} \ , \tag{7.98}$$

one can repeat the formal expansion in ϵ of the Ricci tensor,

$$R_{\mu\nu} = R^{(0)}_{\mu\nu} + R^{(1)}_{\mu\nu} + R^{(2)}_{\mu\nu} + \mathcal{O}(3) \ , \tag{7.99a}$$

where now the background metric η is to be considered a generic one. The power counting is now something like

$$\eta \sim \mathcal{O}(\epsilon^0) \ , \ \ h^{(n)} \sim \mathcal{O}(\epsilon^n) = \mathcal{O}(n) \ , \ \ \partial\eta \sim \mathcal{O}(1/L) \ , \ \ \partial h^{(1)} \sim \mathcal{O}(\epsilon^1 k) \ , \ \partial^2\eta \sim \mathcal{O}(1/L^2) \ , \ \partial^2 h^{(1)} \sim \mathcal{O}(\epsilon^1 k^2) \ , \tag{7.99b}$$

and thus

$$R^{(0)}_{\mu\nu} \sim \eta\partial^2\eta \sim \mathcal{O}(1/L^2) \quad \text{Long wavelength (low freq.)} \tag{7.99c}$$
$$R^{(1)}_{\mu\nu} \sim \eta\partial^2 h^{(1)} \sim \mathcal{O}(\epsilon k^2) \quad \text{Short wavelength (high freq.)} \tag{7.99d}$$
$$R^{(2)}_{\mu\nu} \sim \eta\partial^2 h^{(2)} + h^{(1)}\partial^2 h^{(1)} \sim \mathcal{O}(\epsilon^2 k^2) \quad \text{Long/Short wavelengths} \ , \tag{7.99e}$$

where the second-order Ricci can contain both long and short wavelenght contributions because the short wavelength combinations $\sim h_{\alpha\beta}h_{\mu\nu}$ can result in a long wavelength mode if the two wavenumbers are comparable but have opposite signs. Hence, the expanded equations can be formally separated into a low/high frequency parts:

$$R^{(0)}_{\mu\nu} = -\left[R^{(2)}_{\mu\nu}\right]^{\text{long}} + 8\pi\left[\bar{T}_{\mu\nu}\right]^{\text{long}} \tag{7.99f}$$
$$R^{(1)}_{\mu\nu} = -\left[R^{(2)}_{\mu\nu}\right]^{\text{short}} + 8\pi\left[\bar{T}_{\mu\nu}\right]^{\text{short}} \tag{7.99g}$$

Based on the results of linearized theory discussed above, one expects
- the long wavelengths equation to correspond to lead to describe the effet of GW on the background curvature (thus leading to the definition of the effective stress-energy tensor for GW), and
- the short wave equation to the propagation of the perturbation on the curved background.

Let us comment on the validity of the above equations. The latter equate terms with different powers of ϵ; to be consistent these powers must be compensated by the other expansion parameter $\lambda\!\!\!^{-}/L$. Considering the long wavelength equation in vaccum $\bar{T}_{\mu\nu} = 0$ or equivalently in a situation where the curvature is dominated by GW, the power counting gives

$$\frac{1}{L^2} \sim \epsilon^2 k^2 = \frac{\epsilon^2}{\lambda\!\!\!^{-2}} \quad \Rightarrow \quad \epsilon \sim \frac{\lambda\!\!\!^{-}}{L} \ . \tag{7.100}$$

In the opposite case, where the $\bar{T}_{\mu\nu}$ dominates over $R^{(2)}_{\mu\nu}$, one must have

$$\frac{1}{L^2} \sim \epsilon^2 k^2 + (\text{matter}) \gg \epsilon^2 k^2 \quad \Rightarrow \quad \epsilon \ll \frac{\lambda\!\!\!^{-}}{L} \ . \tag{7.101}$$

Remark 7.10.2. Breakdown of expansion on Mikowski and of scales separation. *Eq. (7.100) indicates that the metric expansion on Mikowski cannot be pushed beyond linear order. If the background metric is Minkoski, then $1/L = 0$ (stricly zero), and no GW of finite amplitude can exist. In other terms, the expansion in powers of ϵ has no domain of validity. More in general, Eq. (7.101) indicates that if the GW amplitude becomes too large, then the hypothesis of scale separation breaks and it is not possible to define wave-like perturbation on a background. [Note the smallness of the amplitude was assumed above but not really justified.]*

Long wavelengths equation & Isaacson stress-energy tensor. How to actually implement the separation of scales? The idea is to integrate on lengths that are longer that the perturbation wavelength but shorter than the background variation length, $\lambda\!\!\!^{-} < \ell < L$. The specific average operator was introduced by Brill&Hartle and Isaacson in 1968 and for a $(0,2)$ tensor reads

$$\langle S_{\mu\nu}\rangle := \int d^4x\, \eta^{\mu'}_{\alpha}(x,x')\eta^{\nu'}_{\beta}(x,x')S_{\mu\nu}(x')f(x,x')\sqrt{|\eta(x')|} \ , \tag{7.102}$$

where η is the generic background metric, $\eta^{\mu'}_{\mu}(x,x')$ is called the *bivector of geodesic parallel displacement* and is an operator that transports a tensor $S_{\mu\nu}(x')$ in a neighbourg of x'. $\eta^{\mu'}_{\mu}(x,x')$ transforms as tensor in x' in the index μ' and as a tensor in x in the index μ. The function $f(x,x')$ is a weighting function that goes rapidly to zero if the two points x and x' are separated by many wevelengths $\lambda\!\!\!^{-}$. [See Misner et al. (1973) §35.14 and Isaacson PhD thesis appendix for more details.] Operatively, the average operator can be used in calculations with the following rules:

- Covariant derivatives inside the average commute up to $\mathcal{O}(\lambda\!\!\!^{-2}/L^2)$, $\langle h\nabla^\alpha\nabla^\alpha h_{\mu\nu}\rangle = \langle h\nabla^\alpha\nabla^\beta h_{\mu\nu}\rangle$;
- Gradients are zero up to $\mathcal{O}(\lambda\!\!\!^{-}/L)$, $\langle\nabla^\alpha h_{\mu\nu}\rangle = 0$;
- One can integrate by parts, $\langle h\nabla^\alpha\nabla^\beta h_{\mu\nu}\rangle = \langle -\nabla_\beta h\nabla^\alpha h_{\mu\nu}\rangle$;
- Diverge terms average to zero, $\langle\nabla_\rho h^\rho_{\mu\nu}\rangle = 0$.

The long wavelength average of the stress-energy tensor and of the second order Ricci tensor is obtained applying the operator above. In particular the average of the Ricci is calculated from an expression for $R^{(2)}_{\mu\nu}$ similar to Eq. (7.91) in which several terms are in the form of (or can be cast in, integrating by parts) divergences, and thus average to zero. The result specified for a flat background and in TT gauge (far from an isolated source) is

$$\left[R^{(2)}_{\mu\nu}\right]^{\rm long} := \langle R^{(2)}_{\mu\nu}\rangle = -\frac{1}{4}\langle\partial_\mu h^{\rm TT}_{ij}\partial_\nu h^{{\rm TT}\,ij}\rangle \ . \tag{7.103}$$

The *Isaacson tensor* is the effective stress-energy tensor for GW:

$$\tau_{\mu\nu} := \frac{c^4}{32\pi G}\langle\partial_\mu h^{\rm TT}_{ij}\partial_\nu h^{{\rm TT}\,ij}\rangle \ . \tag{7.104}$$

Properties:
- The Isaacson tensor is gauge invariant under infinitesimal diffeomorphisms;
- It should be used to generalize the expressions given above for the GW energy/flux.

The long wavelength ("coarse grained") are finally

$$R^{(0)}_{\mu\nu} = 8\pi\langle\bar{T}_{\mu\nu}\rangle + 8\pi\tau_{\mu\nu} \ , \tag{7.105}$$

and the Bianchi identies implies local conservation of the r.h.s. with respect to the covariant derivative of the background. Far away from the source $\partial^\mu\tau_{\mu\nu} = 0$. A calculation similar to the one for the quadrupole formula leads to the GW luminosity.

$$0 = \int d^3x\,(\partial_0\tau^{00} + \partial_i\tau^{0i}) = -\dot{E} + \int d^3x\,\partial_i\tau^{0i} = -\dot{E} + \oint d^2y\,n_i\tau^{0i} = -\dot{E} + r^2\oint d^2y\,n_r\tau^{0r} \tag{7.106a}$$

$$= -\dot{E} + r^2\oint d^2y\,\langle\partial^0 h_{ij}\partial_r h^{ij}\rangle = -\dot{E} + r^2\oint d^2y\,\langle\partial_t h_{ij}\partial_t h^{ij}\rangle \ , \tag{7.106b}$$

where in the surface integrals one takes a spherical surface of radius r and then uses the fact that for a wave $h(t-r/c)/r$ the spatial derivative can be written as $\partial_r h(t-r/c) = -\partial_t h(t-r)/(rc) = +\partial^0 h(t-r)/(rc)$. The GW luminosity is then given by

$$\frac{dE}{dt} = \frac{c^3}{32\pi G}r^2\int d\Omega\langle\partial_t h^{\rm TT}_{ij}\partial_t h^{{\rm TT}\,ij}\rangle = \frac{c^3}{16\pi G}\langle\dot{h}^2_+ + \dot{h}^2_\times\rangle = \frac{G}{5c^5}\langle\dddot{Q}_{ij}\dddot{Q}^{ij}\rangle. \tag{7.107}$$

Remark 7.10.3. The two faces of GW luminosity. *The dimension analysis of the luminosity formula starts from*

$$[Q] = aML^2 \ , \quad [\dddot{Q}] = aML^2T^{-3} \sim a\Omega^3ML^2 \ , \quad [G/c^5] = TE^{-1} \ , \tag{7.108}$$

where a dimensionless factor a and an angular frequency Ω are intorduced for later convenience. From the above one notices that the numerically small factor $G/c^5\sim$ in front of the formula is the <u>inverse</u> *of a power; if $\dddot{Q}$ has typical values of laboratory experiments the GW luminosity generated in these experiments is ridiculously small. However, Weber (an optimist) suggested to re-express the formula in terms of*

$$c^5/G \sim 10^{52}\ W \tag{7.109}$$

which is an enormous luminosity factor. We use, as usual, R as the typical size of the sources, $\sigma = GM/c^2R$ as a measure of the source's self-gravity, $M = c^2R\sigma/G$ the mass of the source, $v = \Omega R$ the source's velocity. One gets [6]

$$\dot{E} \sim \frac{G}{c^5}a^2\Omega^6M^2R^4 = a^2\frac{G}{c^5}\left(\frac{v}{c}\right)^6\left(\frac{c}{R}\right)^6\frac{c^4R^2\sigma^2}{G^2}R^4 = a^2\frac{c^5}{G}\left(\frac{v}{c}\right)^6\sigma^2 = a^2\frac{c^5}{G}\left(\frac{v}{c}\right)^6\left(\frac{GM}{c^2R}\right) \ . \tag{7.110}$$

The last formula above shows that a source with strong self-gravity $\sigma\sim 1$ and high-velocity $v\sim c$ can generate GWs corresponding to the most luminous radiation in the Universe.

GW Propagation in curved background. Let us examine the short-wavelength equation considering fist the vacuum case and then the matter dominated case.

In vacuum Eq. (7.100) implies that there exist only one scale since $\mathcal{O}(\epsilon)\sim\mathcal{O}(\lambda\!\!\!^{-}/L)$ and one can use only ϵ. Inspection of the orders reveals immediately that $R^{(2)}_{\mu\nu}$ can be neglected and the equation reduces to the $1/\epsilon$ part of $R^{(1)}_{\mu\nu}$:

$$\begin{cases} R^{(1)}_{\mu\nu} & \sim \eta\partial^2h^{(1)} \sim \mathcal{O}(\epsilon k^2) = \mathcal{O}(\frac{1}{\epsilon}) \\ R^{(2)}_{\mu\nu} & \sim \eta\partial^2h^{(2)} + h^{(1)}\partial h^{(1)} \sim \mathcal{O}(\epsilon^2k^2) = \mathcal{O}(1) \end{cases} \ . \tag{7.111}$$

[6] Note that for a binary system of reduced mass M, orbital separation R and orbital frequency Ω, the first formula for E with $a^2 = 32/5$ is exactly the result of the quadrupole calculation.

The $1/epsilon$ (leading order) part is obtained by substituing the background metric with the flat metric and the background covariant derivatives with the partial derivatives in the flat metric. Hence, this equation reduces exactly to the wave equation in flat background once the Hilbert gauge is imposed

$$0 \simeq \left[R^{(1)}_{\mu\nu}\right]_{1/\epsilon} \overset{H.G.}{=} \Box\bar{h}_{\mu\nu} \ . \tag{7.112}$$

If matter dominate the curvature, the two expansion parameters are not equivalent and from Eq. (7.101) one has $\epsilon \ll \lambda\!\!\!^{-}/L \ll 1$. Hence one can keep only terms linear in ϵ and expand in $\lambda\!\!\!^{-}/L$ keeping leading order and next-to-leading order terms. The short wavelength part of $R^{(2)}_{\mu\nu}$ is negligible w.r.t. $R^{(1)}_{\mu\nu}$ because it contains one more power of ϵ. The short wavelength part of $T_{\mu\nu}$ must contain a short wavelength term $\mathcal{O}(\epsilon)$ because the stress-energy tensor depends in general on the metric. The trace part will have a short wavelength terms $\mathcal{O}(\epsilon)$ given by the multiplication of the (high-frequency) T with the background metric and by the multiplication of the (low frequency) T with the metric perturbation TODO expand. Hence,

$$\begin{cases} R^{(1)}_{\mu\nu} & \sim \eta\partial^2 h^{(1)} \sim \mathcal{O}(\epsilon k^2) \\ \left[R^{(2)}_{\mu\nu}\right]^{\text{short}} & \sim \eta\partial^2 h^{(2)} + h^{(1)}\partial h^{(1)} \sim \mathcal{O}(\epsilon^2 k^2) \\ \left[\bar{T}_{\mu\nu}\right]^{\text{short}} & = \left[T_{\mu\nu} - \frac{1}{2}(\eta_{\mu\nu} + h_{\mu\nu})T\right]^{\text{short}} \sim \mathcal{O}(\frac{\epsilon}{L^2}) \sim \mathcal{O}(\epsilon k^2 \frac{\lambda\!\!\!^{-2}}{L^2}) \end{cases} . \tag{7.113}$$

Both the curvature and matter contributions can be discarded and the equation reduces to the linear Ricci on the <u>curved</u> background metric. A Hilbert gauge can be introduced demanding the covariant divergence of $\bar{h}^{(1)}$ to be zero. The result is that to $\mathcal{O}(\epsilon k^2)$ the perturbation propagates following the wave equation on the curve background:

$$0 \simeq R^{(1)}_{\mu\nu} \overset{H.G.}{=} \Box_\eta\bar{h}_{\mu\nu} \ . \tag{7.114}$$

8. Schwarzschild solution

③
These lectures introduce the solution of GR in vacuum and spherical symmetry found by Schwarzschild 1915.

Suggested readings. *Chap. 5 of Wald (1984); Chap. 5 of Carroll (1997); Chap. 11 of Schutz (1985)*

8.1 Schwarzschild spacetime

Exact solution of GR for a spacetime
(i) Vacuum;
(ii) Spherically symmetric;
(iii) Static.

The last hypothesis is actually not necessary: Birkhoff theorem (Sec. 8.3) says that any spacetime for which (i) and (ii) hold is static. The solution describes the spacetime outside a spherically symmetric mass distribution and, more in general, the vacuum spacetime far away from an isolated source. Asymptotically, the spacetime is flat and reproduces the weak field solution and Minkoswki.

The Schwarzschild metric provides us with the basis for key GR calculations like
- Mercury perihelion precession;
- Light bending;
- Gravitational redshift;
- Shapiro time-delay.

These predictions can be tested in the weak field regime of the Solar system. Moreover, the Schwarzschild metric provides us with some of the unexpected and key predictions/phenomena of GR in the strong field regime
- Black holes;
- Mass limit for compact stars (when combined with the proper interior solution for spherically symmetric mass distribution);
- Gravitational collapse.

8.2 Derivation of the solution

Spherically symmetric and static metric.

Definition 8.2.1. *A metric is* stationary *iff exists a timelike Killing vector field (KV) $T^a = (\partial_t)^a$.*

In the coordinate adapted to the KV, the metric components are "time-independent"

$$\partial_t g_{\alpha\beta} = 0 \quad \Rightarrow \quad g = -g_{00}(x^i)\mathrm{d}t^2 + 2g_{0i}(x^i)\mathrm{d}t\mathrm{d}x^i + g_{ij}(x^i)\mathrm{d}x^i\mathrm{d}x^j \ . \tag{8.1}$$

Definition 8.2.2. *A metric is* static *iff it is stationary and invariant under time-reversal $t \to -t$.*

Observe that
- Stationary = invariance w.r.t. time translations;
- Static = invariance w.r.t. time translations and reflections.

Since the only term that violates time reversal in the expression above is $\mathrm{d}t\mathrm{d}x$, the components g_{0i} must be zero and one can write the metric as

$$g = -g_{00}\mathrm{d}t^2 + g_{ij}\mathrm{d}x^i\mathrm{d}x^j = -N^2\mathrm{d}t^2 + \gamma \ , \tag{8.2}$$

where $N(x^i)$ is a smooth function of the spatial coordinates called *lapse function* and γ is a 3D Riemannian metric at $t = const$ (with signature $(+,+,+)$.) The form above suggests that the manifold can be written as

$$\mathcal{M} = \mathbb{R} \times \Sigma_t \ , \tag{8.3}$$

where Σ_t are spacelike hypersurfaces defined as those orthogonal to the timelike KV T^a and parametrized by the time coordinate and equipped with the metric γ. The surface $t = const$ has normal

$$n_\mu = -N(\mathrm{d}t)_\mu = (-N, 0, 0, 0) \quad \text{or} \quad n^\mu = (\frac{1}{N}, 0, 0, 0) \ , \tag{8.4}$$

implying that any vector on Σ is orthogonal to n^μ.

Remark 8.2.1. *Recalling the discussion in Example 6.8.1 and Remark 6.8.1, one says that Σ_0 is a submanifold of $\mathcal{M}$ formally identified by the map $\phi : \Sigma \mapsto \mathcal{M}$. The latter map is called an* embedding. *The metric on Σ is given by the pullback of the 4D metric $\gamma = \phi^* g$, and it is called the* induced metric.

Definition 8.2.3. *A metric is* spherically symmetric *iff there exist three spacelike KV $R^a_{(i)}$ $i = 1, 2, 3$ satisfying the algebra of $SO(3)$,*

$$[R_{(i)}, R_{(j)}] = \epsilon^{ijk} R_{(k)} \tag{8.5}$$

whose orbits are two dimensional spheres. The induced metric in the orbits is the standard metric in S^2, up to rescalings.

There is an equivalent, simpler and more practical definition

Definition 8.2.4. *A metric is* spherically symmetric *iff there exist a coordinate system $x^\mu = (t, r, \theta, \phi)$ such that*
(i) the surfaces $t = const$, $r = const$ are two-spheres with line usual element $d\Omega^2 = \mathrm{d}\theta^2 + \sin(\theta)\mathrm{d}\phi^2$ (up to a rescaling of the radius);
(ii) the metric can be written

$$g = -e^{2\alpha(t,r)}\mathrm{d}t^2 + e^{2\beta(t,r)}\mathrm{d}r^2 + e^{2\gamma(t,r)}r^2\mathrm{d}\Omega^2 \ . \tag{8.6}$$

Note the metric form with the exponentials guarantee that the metric signature is $(-, +, +, +)$. The agreement of the two definition can be checked by verifying that the above metric admit the three KV

$$R_{(3)} := \partial_\phi \ , \quad R_{(2)} := -\sin\theta\partial_\theta - \cot\theta\cos\phi\partial_\phi \ , \quad R_{(1)} := \cos\theta\partial_\theta + \cot\theta\sin\phi\partial_\phi \ , \tag{8.7}$$

corresponding to rotations about the three Cartesian axes [exercise].

Schwarzschild radial coordinate. Restrict the metric to $t = \bar{t}$ and $r = \bar{r}$ such that the 2-sphere element is $ds^2 = e^{(2\gamma(\bar{t},\bar{r}))}\bar{r}^2 d\Omega^2$. The Schwarzschild radial coordinate r is by definition the coordinate such that the area of the 2-spheres is given by $A = 4\pi r^2$. For this reason it is also called *areal radius*. Given the generic form of the metric (with a generic radial coordinate), it can be computed by performing the coordinate transformation:

$$A = 4\pi e^{2\gamma(\bar{t},\bar{r})}\bar{r}^2 \quad \Rightarrow \quad r^2 = e^{2\gamma(\bar{t},\bar{r})}\bar{r}^2 \ . \tag{8.8}$$

Remark 8.2.2. *It is important to realize that r does* **not** *represent the "distance from the center to the surface of 2-sphere". The areal radius is defined only by the property of the surface (its area); the center is not a point of the 2-sphere and, in general, might not belong to the manifold (Cf. the whormhole solution below).*

Putting together spherical symmetry and static hypothesis and setting $N = e^{(\alpha)}$ one gets

$$g = -e^{2\alpha(r)}\mathrm{d}t^2 + e^{2\beta(r)}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \ . \tag{8.9}$$

Physical interpretation of the lapse. Consider a photon of 4-momentum p^μ moving on null geodesics of g. An observer with 4-velocity $u^\mu = (u^0, u^i) = (u^0, 0, 0, 0) = (N, 0, 0, 0)$ [1] measures the photon energy

$$E = -u^\mu p_\mu = -u^0 p_0 = -N p_0 \ . \tag{8.10}$$

But because the metric has the KV $T^\mu = (1, 0, 0, 0)$, the quantity $T^\mu p_\mu = T^0 p_0 = \bar{p}$ is a constant of motion (same in every point) and the ratio of the photon's energies measured by the observer at two different radial coordinates is

$$\frac{E(r_1)}{E(r_2)} = \frac{N(r_1)\bar{p}}{N(r_2)\bar{p}} = \frac{N(r_1)}{N(r_2)} = \frac{e^{\alpha(r_1)}}{e^{\alpha(r_2)}} \quad \Rightarrow \quad z = \frac{N(r_1)}{N(r_2)} = e^{\alpha(r_1)-\alpha(r_2)} - 1 \ . \tag{8.11}$$

The lapse function (and the coefficient α) is thus directly related to the redshift of photons as measured at different locations. We shall see below that the metric is asymptotically flat and that at large coordinate radii it reduces to the Minkowski one. Thus, $N(r_2) \approx 1$ for $r_2 \to \infty$ and the lapse gives the redshift of photon measured by a distance observer.

[1] The u^0 component is simply calculated from the normalization $-1 = u^\mu u_\mu = g_{00}u^0u^0$.

Determination of α, β. The metric coefficients are determined by EFE in vaccum, $R_{\mu\nu} = 0$ [2] Combine the rr and tt equations gives

$$\begin{cases} R_{tt} & = 0 \\ R_{rr} & = 0 \end{cases} \quad \Rightarrow \quad 0 = e^{2(\beta-\alpha)} R_{tt} + R_{rr} = \frac{2}{r}(\partial_r \alpha \partial_r \beta) \quad \Rightarrow \quad \alpha = -\beta + const \ . \tag{8.12}$$

The constant can be neglected/reabsorbed in a scaling of the time coordinate

$$e^{\alpha} = e^{-\beta} e^{c} \quad \Rightarrow \quad e^{2\alpha} \mathrm{d}t^2 = e^{-2\beta} e^{2c} \mathrm{d}t^2 = e^{-2\beta} \mathrm{d}(e^c t)^2 \ . \tag{8.13}$$

Consider now the $\theta\theta$ equation

$$R_{\theta\theta} = 0 \quad \Rightarrow \quad 1 = e^{2\alpha}(2r\partial_r \alpha + 1) = \partial_r(r e^{2\alpha}) \quad \Rightarrow \quad e^{2\alpha} = 1 - \frac{R_S}{r} \ , \tag{8.14}$$

where R_S is also a constant with dimension of a length. One can verify that with the above choices of α, β all the other EFE in vacuum are identically satisfied.

The metric takes the form

$$g = -\left(1 - \frac{R_S}{r}\right) \mathrm{d}t^2 + \left(1 - \frac{R_S}{r}\right)^{-1} \mathrm{d}r^2 + r^2 \mathrm{d}\Omega^2 \ . \tag{8.15}$$

The constant R can be fixed assuming that the metric describes an isolated system.

- If the isolated system has zero mass, then g must match Minkoswki. One sees immediately the condition $R = 0$ reproduces Mikowski in spherical coordinates.
- If the isolated system has mass M, then the g_{00} and g_{rr} components of the metric above must reduce those of the weak field metric at distances far away from the source. For $r \to \infty$ one must have

$$g_{00} = -\left(1 - \frac{R_S}{r}\right) \approx -(1 + 2\phi) \ , \quad g_{rr} = +\left(1 - \frac{R_S}{r}\right)^{-1} \approx \left(1 + \frac{R_S}{r}\right) \approx +(1 - 2\phi) \ . \tag{8.16}$$

Hence, the *Schwarzschild radius* R_S is defined by:

$$-\frac{R_S}{r} \approx 2\phi \approx -2\frac{GM}{c^2 - r} \quad \Rightarrow \quad R_S := 2M = \frac{2GM}{c^2} \ . \tag{8.17}$$

Observations.

- $g = \eta$ for $M = 0$;
- $g \to g_{\text{weak field}}$ for $r \to \infty$;
- Metric coefficients are singular for $r = 0$ and $r = 2M$.

Q: How should we interpret these singularities? Are they physical or related to the coordinates choice?

A sufficient condition to verify that a singularity is physical is to find a scalar of the curvature that diverges at that point. The Ricci scalar is of no use here, but one can compute the *Kretschmann scalar*

$$R^{\alpha\beta\gamma\delta} R_{\alpha\beta\gamma\delta} = 12 M^2 r^{-6} \ , \tag{8.18}$$

that indicates that $r = 0$ is a physical singularity of the metric. Note that $r = 0$ is not part of the manifold/spacetime because the metric is not defined there.

On the other hand none of the curvature scalars that one can construct diverges at $r = R_S$ suggesting the latter is a coordinate singularity... We will study the behaviour of the metric around $r \sim R_S$ in Sec. 8.5. For the moment it is sufficient to observe that the character of the KV changes below the Schwarzschild radius,

$$g(\partial_t, \partial_t) = g_{tt} \mathrm{d}t(\partial_t) \mathrm{d}t(\partial_t) + 0 = g_{tt} = -\left(1 - \frac{R_S}{r}\right) \geq 0 \ \text{ for } r \leq R_s \quad \Rightarrow \quad \partial_t \text{ is null/spacelike for } r \leq R_s \ . \tag{8.19}$$

Hence, the Schwarzschild metric is valid only for $r > R_s$. Moreover, note that for the Sun $R_\odot \sim 10^6 M_\odot \gg R_{S\odot}$: the Schwarzschild radius of the Sun is located in the interior of the Sun where the Schwarzschild is not valid because there it is not vacuum. This implies that Schwarzschild metric and coordinates can be safely used for the exterior of the Sun and Solar system.

[2] Explicit expressions for the Ricci tensor require trivial but lenghty calculations. The result is, for example, tabulated in the additional material available at `http://sbernuzzi.gitpages.tpi.uni-jena.de/gr/`

8.3 Birkhoff theorem

Theorem 8.3.1. Birkhoff (1923). *The Schwarzschild metric is the unique vacuum solution in spherical symmetry.*

Note the statement above does not mention the word "static" ... but before commenting let us sketch the main steps of the proof:

1. Use the definition of spherically symmetric spacetime based on the existance of the three spacelike rotational KV and show that any spherically symmetric spacetime can be foliated in 2-spheres.
2. The most general form of the metric is
$$g = -e^{2\alpha(t,r)}\mathrm{d}t^2 + e^{2\beta(t,r)}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \ . \tag{8.20}$$
3. Use EFE in vacuum for the metric above and show that the "time-dependence" in α, β can be removed. Specifically, one finds
$$R_{tr} = 0 \quad \Rightarrow \quad \partial_t\beta = 0 \ , \tag{8.21a}$$
$$\begin{cases} \partial_t R_{\theta\theta} & = 0 \\ R_{tr} & = 0 \end{cases} \quad \Rightarrow \quad \partial_t\partial_r\alpha = 0 \quad \Rightarrow \quad \alpha(t,r) = \beta(r) + c(t) \ , \tag{8.21b}$$

   and the term $c(t)$ can be reabsorbed in the definition of the coordinate time exactly as the constant c above.

The last step proves that the Schwarzschild metric derived above is the most general vacuum and spherically symmetric solution, and shows that the that

Theorem 8.3.2. *Any spherically symmetric vacuum spacetime is static.*

The theorem applies for any vacuum spherically symmetric solution. For example the exterior of a spherically symmetric body that is contracting under the (attractive) gravitational forces (gravitational collapse) is static. Physically, the staticity result can be understood with the absence of gravitational monopole radiation (analogous to the fact that the Coulomb solution is the only spherically symmetric solution of Maxwell equations in vacuum).

Remark 8.3.1. *Note that the coordinate system of Eq. (8.15) breaks down at points in which $T^a = 0 = \nabla^a r$ (or T^a and $\nabla^a r$ are collinear). Hence, the specific form of the metric cannot be used in those conditions and the Birkhoff does not apply.*

8.4 Geodesics

The EOM for particles and light can be found following the general procedure of minimizing the Lagrangian and solving the resulting system of 2nd order coupled ODEs for $\ddot{x}^\mu$, schematically (calculations are left as [exercise])

$$L = g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} \quad \to \quad \frac{dt^2}{d\lambda^2} = ... \ , \quad \frac{dr^2}{d\lambda^2} = ... \ , \quad \frac{d\theta^2}{d\lambda^2} = ... \ , \quad \frac{d\phi^2}{d\lambda^2} = ... \ .$$

Solutions for the geodesic equations can be more easily found using the conserved quantities associated to the metric's symmetries. For each KV one has a constant of motion, and additionally the Lagrangian is constant along the geodesics. Schematically,

$$k_\mu\frac{dx^\mu}{d\lambda} = const \text{ for each KV and} \tag{8.22a}$$
$$-s := g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = \begin{cases} +1 & \text{timelike geodesics, particles (taking } \lambda \text{ as proper time)} \\ 0 & \text{null geodesics, photons} \end{cases} \ . \tag{8.22b}$$

There is a third property (beside KV and Lagrangian) that allows one to simplify the solution: exactly as in the 2-body problem in Newtonian gravity, the motion is on a plane. This can be shown by considering the $\ddot{\theta}$ geodesic

$$\ddot{\theta} + \frac{2}{r}\dot{\theta}\dot{r} - \sin\theta\cos\theta\dot{\phi}^2 = 0 \ , \tag{8.23}$$

and observing that if the motion is initially in the plane, $\theta(0) = \pi/2$ and $\dot{\theta}(0) = 0$, it remains in the plane ($\theta(0) = \pi/2$ and $\dot{\theta}(0) = 0$ satisfy the equation at all times).

Let us calculate the constants of motion associated to the time-symmetry and ϕ-rotational KV. Setting $A := (1 - \frac{R_S}{r})$,

$$T^\mu = (1,0,0,0) \ , \ T_\mu = (-A,0,0,0) \ ; \quad R^\mu_\phi = (0,0,0,1) \ , \ R_{\phi\,\mu} = (0,0,0,r^2\sin^2\theta) \ , \tag{8.24}$$

direct calculation/inspection indicates that

$$e := -T_\mu\frac{dx^\mu}{d\lambda} = +A\frac{dt}{d\lambda} \ ; \quad \ell := R_\mu\frac{dx^\mu}{d\lambda} = r^2\frac{d\phi}{d\lambda}|_{\theta=\pi/2} \ . \tag{8.25}$$

are respectively, a first integral of the geodesic $\ddot{t}$ and a first integral of the $\ddot{\phi}$ equatorial geodesic.

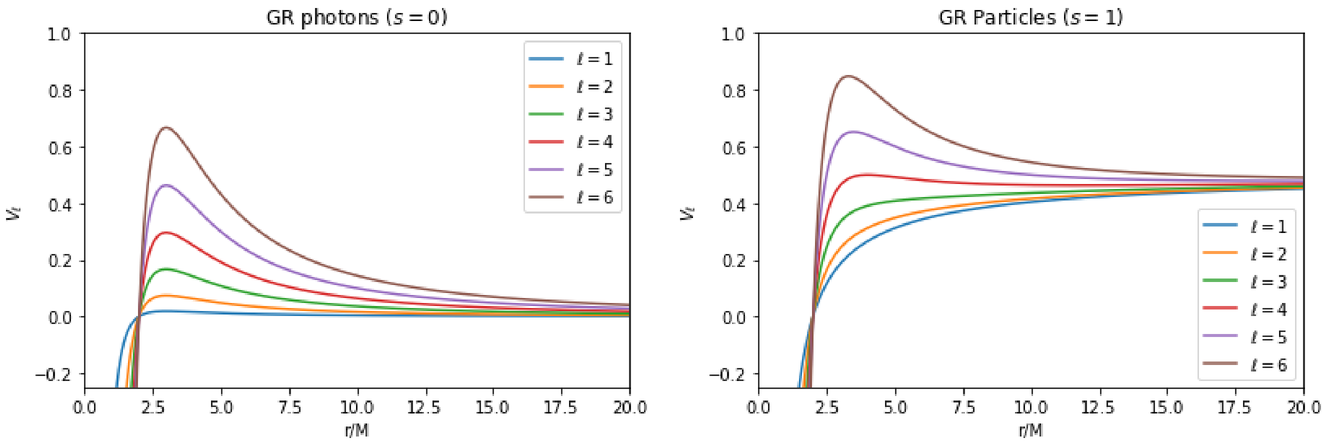

Figure 8.1: Effective potential of photons and particles in units $c = G = M = 1$.

Meaning of integrals of motion.

- For a particle of mass m the quantity e is interpreted as the the total energy (including grav. potential) per unit mass relative to a static observer at infinity. In other terms, e is the energy required by such observer to put the unit-mass particle into the orbit with energy e. The key point here is that e is different from the energy $-p_\mu U^\mu$ measured by any observer with 4-velocity $U_\mu U^\mu = -1$ because U^μ is not a KV. This holds also for a stationary observer ($U^i = 0$) and it is mathematically a consequence of the 4-velocity normalization. Physically, the observer U^μ measures as $-p_\mu U^\mu$ only the kinetic energy of the fre-falling particle. The total energy is the one conserved and can be defined only in presence of a KV (in general it cannot be defined). Obviously, for a photon $\hbar e$ is the total energy of the photon.
- For a particle of mass m the quantity ℓ is interpreted as the angular momentum per unit mass and, similarly, $\hbar\ell$ is the angular momentum of a photon. Note that the expression $\ell = r^2\dot{\phi} = r^2\Omega$ generalizes Kepler 2nd law.

Exercise 8.4.1. *Consider again the* gravitational redshift of photons *with momentum p^μ moving from r_1 to r_2 and as measured by an observer with 4-velocity $u^\mu = (u^0, 0, 0, 0)$ with $u^0 = N = A^{1/2}$ and $u_0 = -A^{-1/2}$. The photon's energy measured by the observer is*

$$E = \hbar\omega = -u_0 p^0 = -u_0 \frac{dt}{d\lambda} = +A^{1/2} u_0 \frac{dt}{d\lambda} = A^{1/2}(A^{-1}e) = A^{-1/2}e = \left(1 - \frac{R_S}{r}\right)^{-1/2} e \ . \tag{8.26}$$

From the above expression one sees immediately that the photon energy measured by u^μ does not correspond to the total photon energy e. Taking the ratio of the energy at two radii, e cancels and one finds immediately the general formula and the correct Newtonian limit for $r \gg M$:

$$\frac{E(r_1)}{E(r_2)} = \frac{\omega(r_1)}{\omega(r_2)} = \left(\frac{A(r_1)}{A(r_2)}\right)^{1/2} = \left(\frac{(r_1 - 2M)r_2}{r_1(r_2 - 2M)}\right)^{1/2} \approx 1 - \frac{M}{r_1} + \frac{M}{r_2} = 1 + \Delta\phi \ , \tag{8.27}$$

Schwarzschild potential. Consider the Lagrangian equation $-s = ...$ restricted to the equatorial plane and multiply by A:

$$-As = -\underbrace{A^2\dot{t}^2}_{=e^2} + AA^{-1}\dot{r}^2 + A\underbrace{r^2\dot{\phi}^2}_{=\ell^2/r^2} \quad \Rightarrow \quad \dot{r}^2 + A(\frac{\ell^2}{r^2} + s) = e^2 \ . \tag{8.28}$$

The above equation can be written as the Newtonian EOM of a unit-mass particle with energy $e^2/2$ moving in a central potential by introducing the *Schwarzschild potential*:

$$V_\ell^{(s)}(r) = \frac{1}{2}A(r)(\frac{\ell^2}{r^2} + s) = \frac{s}{2} - s\frac{M}{r} + \frac{\ell^2}{2r^2} - \frac{M\ell^2}{r^3} \quad \Rightarrow \quad \frac{1}{2}\dot{r}^2 + V_\ell^{(s)}(r) = \frac{e^2}{2} \ . \tag{8.29}$$

This equation is exact in GR and $s = 1(0)$ for timelike(null) geodesics. For timelike geodesics, the potential has a "$1/r$" structure very similar to the Newtonian case

$$V^{(s)} \sim \text{const} + 1/r \text{ Newton potential of a mass } M + 1/r^2 \text{ centrifugal potential} + 1/r^3 \text{ GR term} \ ;$$

because the GR term vanishes faster than all the others for $r \gg M$, the EOM have the correct weak field limit. Note the potential is always postive for $r > R_s = 2M$ and is zero at the Schwarzschild radius and goes to zero ($s = 0$) or to one ($s = 1$) for $r \to \infty$. For $r \to 0$ ($r < 2M$) the potential goes to $V \to -\infty$ while the Newtonian potential goes to $V \to +\infty$. Fig. (8.1) shows examples of particle and photon potential.

Discussion on orbits. Several qualitative features of the orbits on Schwarzschild can be discussed by investigating the potential $V_\ell^{(s)}$. In what follows it is convenient to simplify the radial equation by multiplying by 2, and redefine the potential by omitting the factor 1/2 in front. We will focus on orbits at radii $r \geq R_S$ where the potential is postive. Labels s and ℓ are also omitted.

The radial EOM implies that the potential of the orbit must be always smaller than the total energy (simply because the square velocity is nonnegative)

$$\dot{r}^2 = e^2 - V(r) \geq 0 \quad \Rightarrow \quad V(r) < e^2 \ . \tag{8.30}$$

In turn, this implies that (given the total energy) the motion is restricted to the radii where $V(r) < e^2$ is verified. For example, a particle "moving in" from large radii can continue up to a radius r_* such that $V(r_*) = e^2$; at this point its velocity is zero and its acceleration is positive since ($dr < 0$, turning point)

$$\ddot{r}^2 = -\frac{dV(r)}{dr} \ , \tag{8.31}$$

hence it turns and "moves out" again to infinity. The main features of the orbits are determined by the extrema of the potential, i.e. by the equations

$$0 = \frac{dV(r)}{dr} = sMr^2 - \ell^2 r + 3M\ell^2 \ , \quad \frac{d^2V(r)}{dr^2} = 2sMr - \ell^2 \ . \tag{8.32}$$

Let us look at some of these features.

For photons ($s = 0$) the potential is zero at $r = 2M$ and goes to zero for $r \to \infty$. The potential has a maximum at $r = 3M$ (from Eq. (8.32) with $s = 0$) which is independent on $\ell > 0$ and corresponds to an energy or impact parameter:

$$e_c^2 := V(3M) = \frac{\ell^2}{(3M)^2} - \frac{2M\ell^2}{(3M)^3} = \frac{\ell^2}{27M^2} \quad \text{or} \quad b_c^2 := \frac{\ell^2}{e_c^2} = 27M^2 \ . \tag{8.33}$$

Following the general discussion above,

- Photons with energy $e > e_c$ moving-in from large radii continue moving to $r = 2M$ [and further to $r \to 0$ [3]] (plunge orbits).
- Photons with $e < e_c$ moving-in from large radii hit a turning point at a minimum radius r_* and then move back to large radii (hyperbolic orbits).
- Photons with energy $e = e_c$ have no radial acceleration, i.e. they are on a circular orbit at $r_c = 3M$ called *light ring*. Such orbit exists for every $\ell > 0$ and is unstable since the $r = 3M$ correspond to a maximum of the potential (a small perturbation around the maximum destroys the orbit.)

For particles ($s = 1$) , the potential is zero at $r = 2M$ and goes to one for $r \to \infty$. The extrema of the potential are, from Eq. (8.32) with $s = 1$, at radii

$$r_\pm = \frac{\ell^2 \pm \sqrt{\ell^2(\ell^2 - 12M^2)}}{2M} \ . \tag{8.34}$$

- For $\ell^2 < 12M^2$ there are no extrema.
- For $\ell^2 = 12M^2$ there is a single extremum at $r_c = r_+ = r_- = 6M$, corresponding to a point of inflection of the potential.
- For $\ell^2 > 12M^2$ there are two extrema, corresponding to a maximum (r_-) and a minimum (r_+) of the potential. Note that for large angular momenta, $\ell^2 \gg 12M^2$, their limiting values are $(r_-, r_+) \approx (3M, \ell^2/M)$ corresponding to the light ring and the Newtonian value (Kepler law).

Indicating with $e_\pm^2 = V(r_\pm)$ the energies corresponding to the extrema, one can characterize the orbits as above for a given ℓ FIG.

- Particles moving-in with $e > e_-$ continue to move to $r = 2M$ and $r \to 0$ (plunge orbits).
- Particles moving-in with $e < e_+$ follow hyperbolic orbits.
- Particles with $e = e_- = e_+$ move on a stable circular orbit called *last stable orbit (LSO)* or *innermost stable circular orbit (ISCO)* with angular frequency (squared)

$$\Omega^2 = \left.\frac{\ell^2}{r_+^4}\right|_{r_+=6M} = \left.\frac{M}{r_+^2(r_+ - 3M)}\right|_{r_+=6M} = 2^{-2}3^{-3}M^{-2} \ , \tag{8.35}$$

  where ℓ can be eliminated using $V' = 0$, i.e. $\ell^2 = Mr^2/(r - 3M)$. The energy of the LSO is

$$e^2 = V(6M) = \left.\frac{r_+ - 2M}{r_+^{1/2}(r_+ - 3M)}\right|_{r_+=6M} = \sqrt{8/9} \ . \tag{8.36}$$

[3] In other terms, photons with impact parameter $b < b_c$ are "captured" and move down to $r \to 0$; the capture cross section of the black hole can be defined as $\sigma = \pi b_c^2 = 27M^2\pi$.

- Particles with $e = e_\pm$ move on circular orbits at $r_\pm$; the inner circular orbit at r_- is unstable, the outer circular orbit at r_+ is stable. In other terms, circular orbits with $r_+ \geq 6M$ are stable, and those with $3M < r_- < 6M$ are unstable.
- Particles with $e_+ < e < e_-$ move on bound orbits between $r_\pm$ (not circular). Among these orbits it is interesting to consider those close to circular the orbits r_+ and perform a perturbative analysis:

$$\ddot{r}_+ + \ddot{\delta r} = -\left.\frac{dV}{dr}\right|_{r_+ + \delta r} \approx -\left.\frac{dV}{dr}\right|_{r_+} - \left.\frac{d^2V}{dr^2}\right|_{r_+} \delta r \quad \Rightarrow \quad \ddot{\delta r} = -\left.\frac{d^2V}{dr^2}\right|_{r_+} \delta r \ . \tag{8.37}$$

The perturbed orbit oscillates around the circular one at a frequency

$$\omega^2 = \left.\frac{d^2V}{dr^2}\right|_{r_+} = \frac{M(r_+ - 6M)}{r_+^3(r_+ - 3M)} = \frac{(r_+ - 6M)}{r_+}\frac{M}{r_+^2(r_+ - 3M)} = \frac{(r_+ - 6M)}{r_+}\Omega^2 \ , \tag{8.38}$$

where the angular momentum is again eliminated using $V' = 0$, i.e. $\ell^2 = Mr^2/(r - 3M)$ and where the last expression highlights the relation to the angular frequency of the circular orbit. From the above expression one can verify that $\omega^2 \approx \Omega^2$ for $r \gg 6M$: the orbit is closed, and the particle returns to the same radius after one period. This is consistent with the fact that Newtonian bound orbit are closed ellipsis. However, in general, GR predicts that bound orbits are **not** closed. Rather, they *precess* at a frequency

$$\omega_p = \Omega - \omega = \left[1 - \left(1 - \frac{6M}{r_+}\right)^{1/2}\right]\Omega \approx 3M^{3/2}r_+^{-5/2} \ . \tag{8.39}$$

The last expression above is the leading-order term for $r \gg M$ that is responsible for Mercury perihelion precession.

Exercise 8.4.2. Mercury perihelion. *A more complete calculation for precessing bound orbits in GR should account of eccentricity. Starting from the radial geodesic and (i) restricting to equatorial plane, (ii) including the constant of motions, (iii) multiplying by $(\dot{\phi})^{-2}$, (iv) changing variable to $u = 1/r$, one obtains [exercise]*

$$\frac{d^2u}{d\phi^2} + u = \frac{M}{\ell^2} + \underbrace{3Mu^2}_{GR\ term} \ . \tag{8.40}$$

The above equation is again similar to the Newtonian equations for elliptic orbits but includes a GR term. Without the GR term one has

$$u_N = M\ell^2(1 + \varepsilon\cos\phi) \ , \tag{8.41}$$

where the eccentricity ε is fixed by the initial condition. The GR term can be treated as perturbation when compared to the Newtonian term at the r.h.s. because Mercury's the tangential velocity is

$$\frac{3Mu^2}{M\ell^{-2}} = 3u^2\ell^2 = 3r^{-2}(r^2\dot{\phi})^2 \simeq 3(r\frac{d\phi}{dt})^2 = 3(\frac{v_\perp}{c})^2 \approx 10^{-7} \ . \tag{8.42}$$

Let $u = u_N + v$ and find a linear equation in v:

$$\underbrace{\frac{d^2u_N}{d\phi^2} + u_N - \frac{M}{\ell^2}}_{=0} + \frac{d^2v}{d\phi^2} + v = 3M(u_N^2 + 2u_Nv + v^2) \approx 3Mu_N^2 \ . \tag{8.43}$$

This is an equation for a forced oscillator, the solution is given by the general solution of the homogeneous equations plus a particular solution of the complete equation. A particular solution is given by

$$v = 3M^2\ell^{-4}[1 + \underbrace{\varepsilon\phi\sin\phi}_{secular\ term} + \varepsilon^2(\frac{1}{2} - \frac{1}{6}\cos(2\phi))] \ , \tag{8.44}$$

which is the combination of a constant term, a secular term $\propto \phi$ and an oscillating term. An approximate solution to the perturbation problem is obtained by just picking the secular term,

$$u \approx u_N + v_{secular} = M\ell^{-2}(1 + \varepsilon\cos\phi) + 3M^2\ell^{-4}\varepsilon\phi\sin\phi \simeq M\ell^{-2}\left(1 + \varepsilon\cos(\phi - 3M^2\ell^{-2}\phi)\right) \ , \tag{8.45}$$

where $3M^2\ell^{-2}\phi \sim \sin(3M^2\ell^{-2}\phi)$ for small arguments of the $\sin(.)$ and trigonometric identities were used. The approximate solution above shows that if $\varepsilon \neq 0$, then the orbit is not periodic of 2π and it is not an ellipses. For a revolution of $\phi = 2\pi$, the perihelion shift is

$$2\pi(1 - 3M^2\ell^{-2}) = 2\pi - 6M^2\ell^{-2} \quad \Rightarrow \quad \Delta\phi = 6\pi 3M^2\ell^{-2} \ . \tag{8.46}$$

The angular momentum in the formula above is difficult to measure from astronomical observations and it is best to substitute it with the semi-major axis a. This is done by comparing $u_N(\phi)$ to the generic equation for an ellipsis $u(\phi)$: $ua(1-\varepsilon^2) = (1+\varepsilon\cos\phi)$, which implies $M\ell^2 = a(1-\varepsilon^2)$. Plugging in the angular momentum expression in terms of the semi-major axis, the final result for the perihelion shift is

$$\Delta\phi = \frac{6\pi M}{a(1-\varepsilon^2)} = \frac{6\pi GM}{a(1-\varepsilon^2)c^2} \ . \tag{8.47}$$

Mercury data are $GM_\odot/c^2 \simeq 1.48\cdot 10^5$ cm, $a \simeq 5.79\cdot 10^{12}$ cm, $\varepsilon \simeq 0.20$, $T \simeq 88$ days, resulting in $\Delta\phi \simeq 0.103''$/ orbit $\simeq 43''/100$ yrs. This number fills the discrepancy w.r.t. the Newtonian calculation and the measured value. Note that PSR 1913+16 has a precession of $\Delta\phi \sim 4.2°/100$ yrs, which is $\sim 270\times$ the one of Mercury and indicate the extreme gravity of that binary system [4]

Exercise 8.4.3. Light bending. *A similar calculation as above but for photons gives the trajectory equation*

$$\frac{d^2u}{d\phi^2} + u = 3Mu^2 \ , \tag{8.48}$$

where again the $\propto u^2$ term is small in the Solar system

$$\frac{3Mu^2}{u} = \frac{3R_s}{2r} \leq \frac{R_S}{R_\odot} \sim 10^{-6} \ . \tag{8.49}$$

Neglecting this term gives a straight trajectory for light

$$u_N = b^{-1}\sin\phi \ , \quad or \quad \Delta\phi = 2\arcsin\left(\frac{b}{r}\right) = \pi \ . \tag{8.50}$$

where b is the impact parameter. Note that the total deflection angle is is twice $\Delta\phi$. Inserting the leading order order solution, and solving for the first-order perturbation one finds

$$\frac{d^2v}{d\phi^2} + v = 3Mu_N^2 \quad \Rightarrow \quad u = u_N + v = b^{-1}\sin\phi + \frac{3M}{2b^2}(1 + \frac{1}{3}\cos(2\phi)) \ . \tag{8.51}$$

In the limit of large radii $u \to 0$, the deflection angle $\phi \to 0, \pi$ and because $\sin\phi \approx \phi$ $\cos\phi \approx 1$, one obtains $\Delta\phi \approx 4GM/(bc^2) = 2R_S/b$.

Exercise 8.4.4. Radially infalling particle. *Consider a particle of mass m infalling from large radii on a radial orbit ($\ell = 0$). How long it takes to reach $r = R_s = 2M$? We are interested in computing both the proper time and the coordinate time.*

Consider first proper time; from the radial equation with $\ell = 0$

$$\dot{r}^2 = e^2 - 1 + \frac{R_s}{r} \geq 0 \quad \Rightarrow \quad d\tau = -\frac{dr}{\sqrt{e^2 - 1 + R_s/r}} \ , \tag{8.52}$$

where the minus sign is taken because of the infalling orbit ($dr < 0$). The integral of the r.h.s. is finite for any value of e^2, implying that the particle reaches R_s in a finite proper time.

Consider coordinate time; from the definition of the energy,

$$\frac{dt}{d\tau} = A^{-1}e = (1 - \frac{R_S}{r})^{-1}e \quad \Rightarrow \quad dt = \frac{e\,d\tau}{A(r)} = -\frac{e\,dr}{(1 - R_S/r)\sqrt{e^2 - 1 + R_s/r}} \ . \tag{8.53}$$

Let $\xi = r - R_S = r - 2M$, then

$$dt = -\frac{(\xi + R_S)^{3/2}\,d\xi}{\xi\sqrt{\xi(e^2-1) + e^2 R_S}} \ , \tag{8.54}$$

and one sees that for $\xi \to 0$ ($r \to R_S$) the integral is divergent. Take for example $e^2 \sim 1$ and see that $\int d\xi/\xi \sim \ln\xi$. It should be clear that the divergent term is $A^{-1} \sim 1/\xi$ and it does not depend on e. Hence, the coordinate time diverges for $r \to R_s$. This example further show that the metric singularity is not physical but an effect of coordinates.

8.5 Coordinates, light cones, and extensions

Let us investigate the coordinate singularity in $r = R_S$.

[4] Differently from Mercury, PSR 1913+16 cannot be used this way to verify GR from measured precession because the masses of PSR are not known. The measurement is in fact used to estimate the masses.

Remark 8.5.1. Q: What is a coordinate singularity? *A: Intuitively, coordinate singularities are points in the manifold where the specific coordinates (chart) fail and do not describe properly the geometry. Examples are the points $\theta = 0, \pi$ of the 2-sphere when the usual (θ, ϕ) coordinates are employed. Imagine you do not have an idea of the 2-sphere but you have only a metric expression for the usual line element $d\Omega^2(\theta, \phi)$: the north pole is a point for which $\theta = 0$ but the value of ϕ is undefined. A possible way to discover that the chart is bad is to calculate gauge invariant quantities. In the 2-sphere case, one could consider the circumference of circles at $\theta = \bar{\theta} = const$ and find out that the distance between the points on these circles $(\bar{\theta}, \phi)$ tends to zero for $\bar{\theta} \to 0$. Because the metric is Riemannian two points should be identified as the same point if the distance between them is zero. Thus, the points at $(0, \phi)$ are the same point and it is the coordinate choice that is misleading. In GR the metric is not positive definite and the situation is more complicated, but one possibility to find out about bad coordinates is to study the causal structure of the spacetime.*

Study the Schwarzschild metric slightly below $r = R_S = 2M$. Start by changing variable to $\xi := R_S - r > 0$,

$$g = \frac{\xi}{R_S - \xi}\mathrm{d}t^2 - \frac{R_S - \xi}{\xi}\mathrm{d}\xi^2 \ , \tag{8.55}$$

where we do not write from now on the metric on the sphere since we know for any fixed time and radial coordinate the metric is a 2-sphere. Inside the Schwarzschild radius ($R_s - \xi > 0$)

$$r < R_S \ \Rightarrow \ \xi > 0 \ \Rightarrow \begin{cases} g_{tt} & > 0 \\ g_{\xi\xi} & < 0 \end{cases} \Rightarrow \begin{cases} \partial_\xi \text{ and } \partial_r & \text{are timelike vectors} \\ \partial_t & \text{is spacetime vector} \end{cases} \Rightarrow \begin{cases} \xi \text{ and } r & \text{are timelike coordinates} \\ t & \text{is spacetime coordinates} \end{cases} \tag{8.56}$$

If the above is not clear remember that

$$\mathrm{d}t(\partial_\xi) = 0 \ , \quad g(\partial_\xi, \partial_\xi) = g_{\xi\xi} < 0 \ , \tag{8.57}$$

and similarly for the ∂_t vector.

Because particles follow timelike paths, and below the Schwarzschild radius the timelike paths are those in the ∂_ξ direction, particle inside R_S must reach the point $r = 0$

$$\xi \text{ increases } \Rightarrow \ r \text{ decreases } \Rightarrow \text{ particle reaches } r \to 0 \ . \tag{8.58}$$

Moreover, a photon emitted by a infalling particle that just passed $r \sim R_S$ must move "forward in time" according to the observer with the particle. This indicates that also photons in $r < R_S$ must move to $r \to 0$. All the worldline crossing the Schwarzschild radius move to the singularity. However, this is all speculative, since we already know that the metric in Schwarzschild should **not** be used for $r < R_S$.

Light cones. Calculate radial null curves (by taking a shortcut, repeat the proper calculation as [exercise])

$$0 = g(u, u) \ , \quad g = -A(r)\mathrm{d}t^2 + A^{-1}(r)\mathrm{d}r^2 \ \Rightarrow \ \frac{dt}{dr} = \pm A^{-1/2}(r) = \pm\sqrt{1 - \frac{R_S}{r}} \ . \tag{8.59}$$

A shown in FIG the light cones "close up" when moving from large radii ($g \sim \eta$) towards $r \to R_S$. At $r = R_S$ the cone is infinitely thin and the coordinate time $t \to \infty$ (See Exercise 8.4.4). Even considering "valid" radii $r > R_s$ away from the coordinate singularity, "something physical" is happening. Consider an observer falling towards R_S and sending signals (light pulses) to another observer at large radii (FIG). If the falling body sends the pulses at its fixed proper time intervals, the observer far away gets the pulses at increasing intervals as measured by its clocks. (The formal calculation isa simple [exercise] by now.) In particular, the observer far away will never receive the signal sent when the infalling observer reaches R_S.

The radial null curves equation above is solved by

$$t = \pm r_* + const \ , \quad \text{with} \quad r_* := r + R_S \ln(\frac{r}{R_S} - 1) \quad \text{(Tortoise coordinate)} \ . \tag{8.60}$$

The tortoise coordinates $r_*(r) : [2M, \infty) \mapsto (-\infty, \infty)$, maps the exterior spacetime into $\mathbb{R}$. The metric is

$$g = A\left(r(r_*)\right)(-\mathrm{d}t^2 + \mathrm{d}r_*^2) \ , \tag{8.61}$$

and clearly the cones do not close up by approaching $r_* \to -\infty$ ($r = R_S$) but stay at 45° in the spacetime diagram (t, r_*). Since the tortoise coordinates are valid only in the exterior, (t, r_*) remain always timelike and spacelike respectively. The light cones can be easily characterized by introducing the *null coordinates*

$$\begin{cases} u & := t - r_* \ , \quad \text{outgoing} \\ v & := t + r_* \ , \quad \text{ingoing} \end{cases} \quad \text{(Null coordinates)} \ , \tag{8.62}$$

also defined in the exerior $r > R_S$ and in which

$$g = -\frac{1}{2}A(r(u, v))(\mathrm{d}u\mathrm{d}u + \mathrm{d}v\mathrm{d}u) \ , \tag{8.63}$$

where $r(u,v)$ is given by the inverse of

$$\frac{1}{2}(v-u) = r_* = r + R_S \ln(\frac{r}{R_S} - 1) \ , \quad \text{for} \quad r > R_S \ . \tag{8.64}$$

Note the metric is singular at R_S. Because $\mathrm{d}u(\partial_v)\mathrm{d}v(\partial_v) = 0 \cdot 1 = 0$ and similarly for the ∂_u, the vectors (∂_u, ∂_v) are null. Thus,

- Ingoing radial null geodesics are given by $v = const$;
- Outgoing radial null geodesics are given by $u = const$.

Both tortoise and null coordinates are defined in the exterior, so they do not bring new information about the spacetime. However, they are the basis to define new coordinates that allows us to describe together both exterior and interior.

Eddington-Finkelstein (EF, 1923-1958) coordinates, metric extension & the "null membrane". Consider the coordinate transformation from Schwarzschild coordinates to the a system of coordinate made of the ingoing null coordinate and the Schwarzschild radius,

$$(t,r) \mapsto (v,r) \quad \text{(Ingoing EF coordinates)} \quad \Rightarrow \quad g = -A(r)\mathrm{d}v^2 + \mathrm{d}v\mathrm{d}r + \mathrm{d}r\mathrm{d}v \ . \tag{8.65}$$

Note that $g_{vv}(R_s) = 0$ but the metric is invertible at R_S. In the ingoing EF coordinates the metric is regular at R_S and, while they are formally defined in the exterior for $r \in [2M, \infty)$, the metric expression can be analytically continued (extended) to the interior [5]. This is a big step beyond. The radial null geodesics are now given by

$$0 = g = -A(r)\left(\frac{dv}{dr}\right)^2 + 2\frac{dv}{dr} = -\frac{dv}{dr}\left(A(r)\frac{dv}{dr} - 2\right) \quad \Rightarrow \quad \frac{dv}{dr} = \begin{cases} 0 & \text{ingoing null radial geodesics} \\ 2A^{-1} = \dfrac{2r}{r - R_S} & \text{outgoing null radial geodesics} \end{cases} \tag{8.66}$$

Light cones are shown in Fig. (8.2). The future direction at a given spatial point is given by increasing values of $v = t + r_*$. One observes immediately two interesting facts,

1. Light cones remain well-behaved at $r = R_S$, and timelike/null geodesics can be calculate and can cross R_S. This wass expected from R_S being only a coordinate singularity.
2. Light cones "tilt" over for $r < R_S$ and all the future directed paths are in the direction of the $r = 0$ singularity. This show that the spacetime has something special at R_S: the causal structure is such that the 2-sphere $r = R_S$ functions as a "one-way membrane" that shed the interior form the exterior and things can only "flow in" and not out. The surface $r = R_S$ is called an *event horizon* (EH).

Remark 8.5.2. The EH is a null surface. *Take the surfaces $r = const$ defined by the 1-form $n_\mu = (\mathrm{d}r)_\mu = \partial_\mu r = (0,1,0,0)$. Note the inverse metric is (only the 2×2 relevant block is shown)*

$$g_{\mu\nu} = \begin{bmatrix} -A & 1 \\ 1 & 0 \end{bmatrix} \quad \Rightarrow \quad g^{\mu\nu} = \begin{bmatrix} 0 & 1 \\ 1 & A \end{bmatrix} \ . \tag{8.67}$$

The normal vector is the vector ∂_a associated to the 1-form:

$$n^\mu = g^{\mu\nu}\partial_\mu r = g^{vr}\partial_r r + g^{rr}\partial_r r = (1, A(r), 0, 0) \tag{8.68}$$

Hence, the norm of the normal vector is

$$n_\mu n^\mu = g^{\mu\nu} n_\mu n_\nu = g^{\mu\nu}\partial_\mu r \partial_\nu r = g^{rr} = A(r) \ . \tag{8.69}$$

The above expressions show that, among all the $r = const$ surfaces, the only one that is null ($n_\mu n^\mu = 0$) is the s-sphere at $r = R_S$ for which $A(R_S) = 0$. The normal vector to the EH is $n = \partial_v$. Note that normals to null surface cannot be normalized.

Remark 8.5.3. *The extended metric in ingoing EF coordinates admit the KV ∂_v, whose relation with the timelike KV is immediately found by the coordinate transformation, $x^\mu = (v,r) = (t + r_*(r), r)$,*

$$\partial_t = \frac{\partial x^\mu}{\partial t}\partial_\mu = \partial_v \ . \tag{8.70}$$

However, the KV's norm is $g(\partial_v, \partial_v) = g_{vv} = -A(r)$ which is timelike in the exterior, null at the EH and spacelike in the interior.

[5] Since the metric components are real analytical functions of the coordinates, they can be expanded as convergent functions about a point. Since they satisfies EFE in some open set then also the extended metric is a solution of the vacuum EFE.

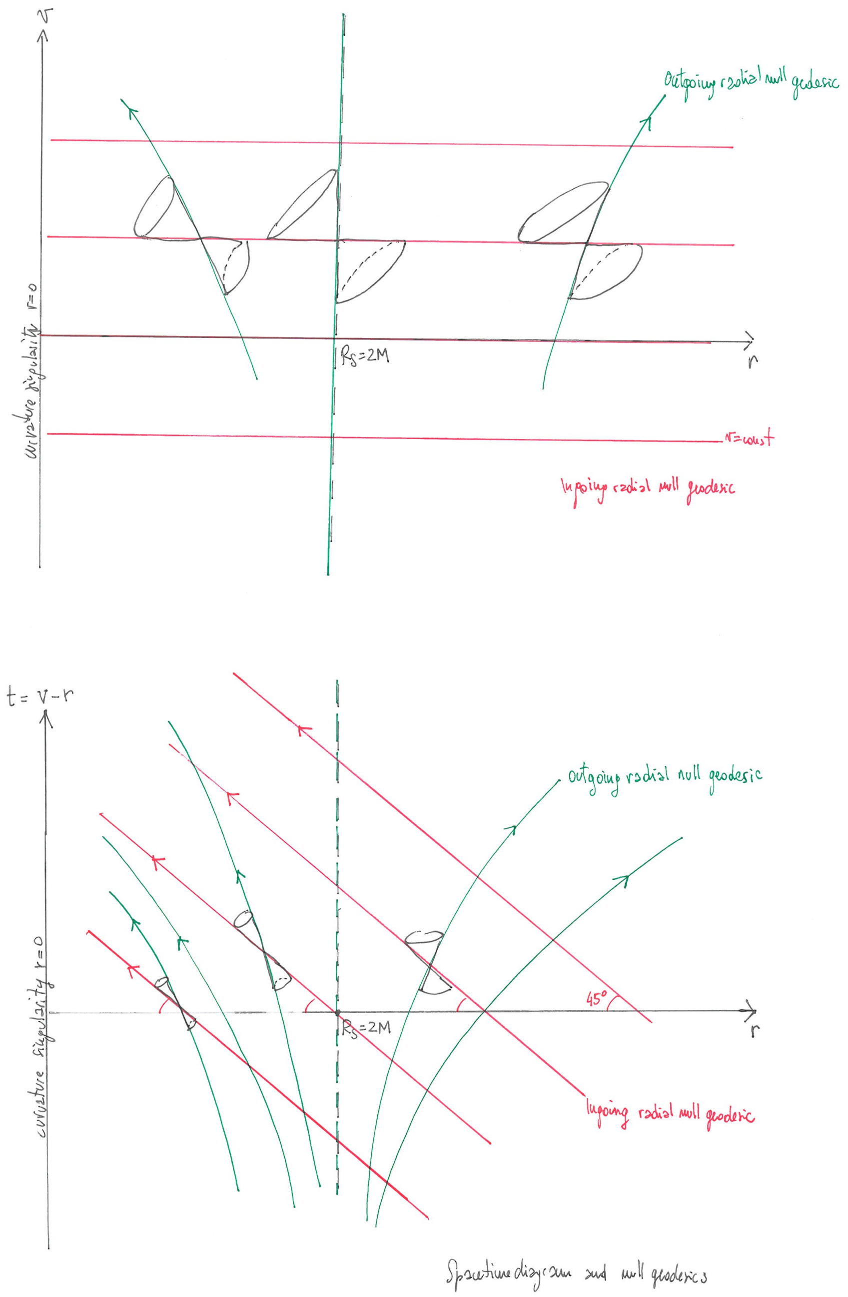

Figure 8.2: Light-cones and radial null geodesics in Eddington-Finkelstein coordinates.

Consider now the tranformation

$$(t, r) \mapsto (u, r) \quad \text{(Outgoing EF coordinates)} \quad \Rightarrow \quad g = -A(r)\mathrm{d}u^2 - \mathrm{d}u\mathrm{d}r - \mathrm{d}r\mathrm{d}u \ . \tag{8.71}$$

It is immediate to repeat all the calculations done for the ingoing EF. But because of the minus sign in the metric above, the outgoing radial null geodesics are those given by $du/dr = 0$ while the ingoing radial null geodesics are those given by $du/dr = 2A^{-1}$. The future directed paths in these coordinates are again those at increasing $u = t - r_*$, but the cones are tilted of 90°. This means that now the one-way membrane can be crossed only moving back in the past: things can "flow out" and not in! Clearly the extension performed with the outgoing EF coordinates does **not** represent the same spacetime (physics) as the extension performed with the ingoing EF coordinates.

Isotropic radial coordinate & the spatial isometry. There is a last step to take that will add information about this spacetime. Consider the radial transformation to *isotropic coordinate* x defined by

$$r = x\left(1 + \frac{M}{2x}\right)^2 =: x\Psi^2(x) \quad \Rightarrow \quad x = \frac{1}{2}\left(r - \frac{M}{2} + \sqrt{r(r - R_s)}\right) \quad \text{for} \quad r \geq R_S \ . \tag{8.72}$$

The Schwarzschild metric in isotropic coordinates (t, x) is

$$g = -\left(1 - \frac{M}{2x}\right)^2 \Psi^2(x)\mathrm{d}t^2 + \Psi^4(x)\left(\mathrm{d}x^2 + x^2\mathrm{d}\Omega^2\right) \ , \tag{8.73}$$

and it is defined for $x \in [M/2, +\infty)$, with the horizon corresponding to $M/2$ and the asymptotic flat end to $x \to \infty$. The function $\Psi(x)$ is called *conformal factor* and in the weak field limit $\Psi \approx 1 - \phi/2$. Note the spatial part of the isotropic metric is the Euclidean 3D diagonal metric multiplied by the conformal factor.

Take now slices $t = const$ and focus on the spatial part of the metric. The spatial metric can be extended to $x \in (0, +\infty)$. The points $x = 0$ do **not** correspond to the $r = 0$ singularity, because we have first mapped the Schwarzschild exterior and then extended the result. The points correspond to another asymptotically flat-end. To see this, take the transformation

$$x \mapsto y = \frac{M^2}{4x} \ , \tag{8.74}$$

and observe that:
- It leaves invariant the 2-sphere of the horizon, $x = y = M/2$;
- It leaves the metric in the same form $g = \Psi(x)(\mathrm{d}x^2 + x^2\mathrm{d}\Omega^2) = \Psi(y)(\mathrm{d}y^2 + y^2\mathrm{d}\Omega^2)$;
- It maps $x = 0$ to $y \to \infty$ and reflects all the points with respect to $M/2$.

The map is an isometry of the metric that leaves invariant the horizon. The region around $x = M/2$ is called the *Einstein-Rosen bridge*. The extension is a spatial slice of Schwarzschild spacetime that connects two asymptotically flat regions without entering below the event horizon.

Remark 8.5.4. *The Schwarzschild metric in isotropic coordinates is usually visualized taking* $t = const$ *and* $\theta = const$ *and plotting the resulting 2D metric as embedded in* $\mathbb{R}^3$*. The extension is then visualized by "gluing" a reflected copy at the Einstein-Roseon bridge.*

Summary 8.5.1. *Let us summarize the results so far:*
- *Schwarzschild coordinates are not appropriate to describe the interior* $r < R_S = 2M$*;*
- *Schwarzschild radius* R_S *is a coordinate singularity but appear to hide a new physical property of the spacetime;*
- *EF coordinates allows one the analytic continuation of the metric to* $r \in (0, \infty)$ *and remove the coordinate singularity at* R_S*;*
- $r = R_S$ *marks a null 2-sphere that characterizes the spacetime with a peculiar causal structure;*
- *In the extension with ingoing EF,* $r = R_S$ *is a EH from which no particles or photons can escape. All the future-oriented timelike or null curves that start from* $r < R_S$ *stay inside the horizon.*
- *In the extension with outgoing EF,* $r = R_S$ *has a similar property, but time reversed. All the future-oriented timelike or null curves go out from the null surface.*
- *There exist a spatial extension in isotropic coordinates that connects two aymptotically flat ends through the Einstein-Rosen bridge.*

8.6 Kruskall-Szekeres (1960) maximal extension

The EF coordinate metric extension suggests that the two pairs of coordinates explore two different regions of a spacetime larger than the initial Schwarzschild exterior $r > R_S$.
Problem: Is is possible to find coordinates that describe the whole spherically symmetric spacetime?
Starting from null coordinates (defined only for $r > R_s$) define new null coordinates

$$\begin{cases} \bar{u} & := -e^{u/2R_S} = -(\frac{r}{R_s} - 1)^{1/2} e^{(r-t)/2R_S} \\ \bar{v} & := +e^{v/2R_S} = +(\frac{r}{R_s} - 1)^{1/2} e^{(r+t)/2R_S} \end{cases} \quad \Rightarrow \quad g = -\frac{16M^3}{r} e^{r/R_S}(\mathrm{d}\bar{u}\mathrm{d}\bar{v} + \mathrm{d}\bar{v}\mathrm{d}\bar{u}) \ . \tag{8.75}$$

Note that for $r = R_S$ one has $\bar{u} = \bar{v} = 0$ and in the metric $r = r(\bar{u}, \bar{v})$. The metric above is regular for $r < R_S$ and can be extended to $r \in \mathbb{R}^+$. The *Kruskall-Szekeres* coordinates are constructed from $(\bar{u}, \bar{v})$ by introducing

$$\begin{cases} T & := \frac{1}{2}(\bar{v} + \bar{u}) \\ R & := \frac{1}{2}(\bar{v} - \bar{u}) \end{cases} \ \text{(Kruskall-Szekeres coordinates)} \quad \Rightarrow \quad g = -\frac{16M^3}{r} e^{r/R_S}(-\mathrm{d}T^2 + \mathrm{d}R^2) \ , \tag{8.76}$$

where $r = r(T, R)$, $T^2 - R^2 = (1 - r/R_S)e^{r/R_S}$, $T/R = \tanh(t/2R_S)$.

Spacetime diagram. See FIG
- Radial null curves are straight lines:

$$g = 0 \quad \Rightarrow \quad \left(\frac{dT}{dR}\right)^2 = 0 \quad \Rightarrow \quad T = \pm R + const \ . \tag{8.77}$$

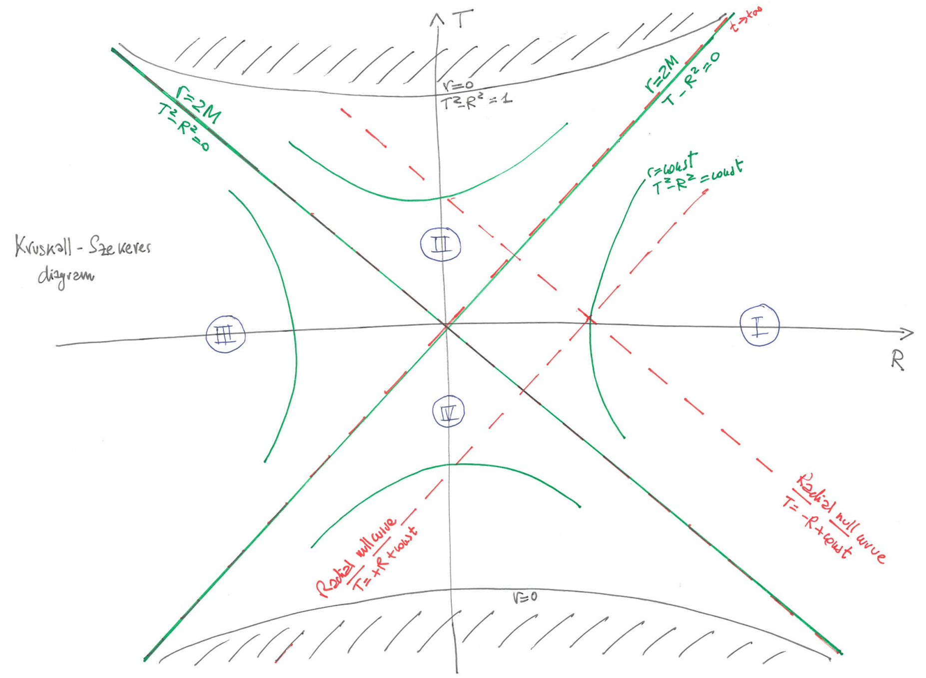

Figure 8.3: Kruskall maximal extension.

- $r = const$ surfaces are hyperbolae,

$$r = const \quad \Rightarrow \quad const = (1 - r/R_S)e^{r/R_S} = T^2 - R^2 \; . \tag{8.78}$$

- The $r = R_S$ is the special “hyperbola” $0 = T^2 - R^2 \Rightarrow T = \pm R$, that clear coincides with a radial null curve. Here vectors T and ∇r are collinear.
- The singularity $r = 0$ correspond to the hyperbola $T^2 - R^2 = 1$ and the spacetime $r \in (0, \infty)$ is mapped to the region $T^2 - R^2 > 1$.
- $t = const$ surfaces are stright lines:

$$t = const \quad \Rightarrow \quad const = \tanh(t/2R_S) = \frac{T}{R} \; . \tag{8.79}$$

- Light cones are at 45° everywhere in the spacetime.

Black hole, white hole and worm hole. Let us explore the Kruskall-Szekeres diagram. There are four regions:

(I) Schwarzschild exterior solution, where (t, r) coordinates are well behaved. Asymptotically flat region for $r \to \infty$. Following future-directed ($t > 0$) null rays one goes from (I) to (II). Following past-directed null rays one goes from (I) to (III).

(II) *BLACK HOLE.* Particles and light can move in, but not out. Once they enter the <u>event horizon</u> at $r = R_S$ they reach the singularity $r = 0$. This is the region explored with the ingoing EF coordinates extension.

(III) *WHITE HOLE.* Time reversal of (II). Things can only move out from the past singularity and cross the past horizon towards the future.

(IV) *WORM HOLE.* Events in this region are spacelike to event in (I); (IV) is causally disconnected from (I). (IV) is a “copy” of (I) and has another asymptotically flat end, the two regions are connected at an istant of time through the Einstein-Rosen bridge but there is no way to cross it without violating

Remark 8.6.1. *The Kruskall-Szekeres spacetime is a valid solution of EFE in vacuum an spherical symetry, but does not necessarily correspond to a region of our Universe, i.e. it is not necessarily physical reality. Indeed, to obtain such a solution a region of the Universe should be initially formed with two asymptotically flat ends I and IV connected by a singularity in III. These would be very special conditions. Physical intuition and astrophysical observations indicate that the most extreme and plausible phenomeon to produce a black hole is the gravitational collapse of massive stars. In this case the initial spacetime is Schwarzschild only in the exterior (after the surface of the collapsing object) while in the interior a regular metric (with no funny causal structure) is determined by the nonvacuum solution of EFE (see Sec. 8.8). Thus, the initial condition is just region (I) glued to an interior nonvacuum spacetime. If the matter collapses under its own gravity and the last surface crosses R_S, then an event horizon can form and the black hole spacetime will be formed by region (I) and (II) of the Kruskall-Szekeres spacetime. There is no clear way to produce white holes or worm holes, that are instead considered unphysical.*

8.7 Conformal infinity & diagrams

Some fundamental questions in GR like the definition of energy and radiated energy and the definition of asymptotical flatness require a precise definition of the spatial and null infinity and a precise procedure to write tensorial equation in those "limiting points". This notion is implemented in the concept of *conformal infinity*: the asymptotic structure of the manifold can be studied by considering a suitable conformal metric (related to physical metric) with the property that it can be extended to infinities and the tensor fields can be evaluated at those points.

A related, more practical, question to start with and motivated by the Kruskall-Szekeres analysis above is the following: Is it possible to describe the spacetime, including infinities, with coordinates with compact support?

Mikowski. Start in standard coordinates $(t, r) \in (-\infty, \infty) \times [0, \infty)$ and in null coordinates $(u, v) = (t - r, t + r) \in (-\infty, \infty) \times (-\infty, \infty)$ with $u \leq v$; the metric is

$$\eta = -\mathrm{d}t^2 + \mathrm{d}r^2 = -\frac{1}{2}(\mathrm{d}u\mathrm{d}v + \mathrm{d}v\mathrm{d}u) \ . \tag{8.80}$$

The null coordinates can be easily compactified using the arctan (.) functions,

$$\begin{cases} U &= \arctan(u) \\ V &= \arctan(v) \end{cases} \quad (U, V) \in (-\pi/2, \pi/2)^2 \ , \quad U \leq V \tag{8.81a}$$

$$\eta = -2\omega^{-2}(\mathrm{d}U\mathrm{d}V + \mathrm{d}V\mathrm{d}U) \ , \quad \omega(U, V) := 2\cos U \cos V \ . \tag{8.81b}$$

The $\omega(U, V)$ function is called the *conformal factor*. From the expression above one sees that the conformal factor is zero for $\pm\pi/2$ that corresponds to the null infinities $u, v \pm \infty$. However, the *conformal metric*

$$\tilde{\eta} := \omega^2\eta \ , \tag{8.82}$$

is well defined at $\pm\pi/2$ and can be extended in those points. Compactified time/spacelike coordinates can be now obtained:

$$\begin{cases} T &= V + U \\ R &= V - U \end{cases} \quad 0 \leq R < \pi \ , \quad |T| + R \leq \pi \tag{8.83a}$$

$$\eta = \omega^{-2}(-\mathrm{d}T^2 + \mathrm{d}R^2) \ , \quad \omega(T, R) := \cos T + \cos R \ . \tag{8.83b}$$

Again, while the the ranges for Mikoski spacetime are indicated above, the conformal metric

$$\tilde{\eta} = \omega^2\eta = -\mathrm{d}T^2 + \mathrm{d}R^2 \ , \tag{8.84}$$

can be extended beyond those ranges. The conformal metric is indeed valid for

$$T \in (-\infty, \infty) \ , \quad R \in [0, \pi] \ , \tag{8.85}$$

and represent a cylinder $\mathcal{M} = \mathbb{R} \times S^1$. One observers that the extension include, in particular, the points $R = 0$ and $T = \pm\pi$ that represents the infinities of Minkoski. Thus, working with tensor fields on the the conformal metric allows one to evaluate them at the inifinities. Note that had we kept the full coordinates the conformal metric would have been

$$\tilde{\eta} = -\mathrm{d}T^2 + \mathrm{d}R^2 + \sin^2 R\mathrm{d}\Omega^2 \ , \tag{8.86}$$

that represents the manifold $\mathbb{R} \times S^3$ and corresponds to Einstein's static universe (Chap. X). Note this manifold has curvature while Minkowski space does not. The conformal infinity of Minkowski can be technically defined as the boundary of the Einstein's static universe.

Returning to the 2D conformal metric, one can visualize the constrution by drawing the cylinder with axis T and identifying the portion of the cylinder corresponding to Minkowski, i.e. the portion delimited by the range of coordinate conditions

$$0 \leq R < \pi \ , \quad |T| + R \leq \pi \ . \tag{8.87}$$

By cutting the cylinder along T and unfolding it, the Minskowki spacetime can be visualized as the interior plus $R = 0$ of a triangle whose other boundaries correspond to infinities (now compactified) Fig. (8.4):

- i^+ Future timelike infinity, $(T, R) = (\pi, 0)$.
- i^0 Spatial intinifty, $(T, R) = (0, \pi)$.
- i^- Past timelike infinity, $(T, R) = (-\pi, 0)$.
- S^+ Future null infinity, $T = \pi - R$, $0 < R < \pi$.
- S^- Past null infinity, $T = -\pi + R$, $0 < R < \pi$.

The i inifinities are points correspondings to the poles of S^3 $(R = 0, \pi)$, while S are null surfaces with topology $\mathbb{R} \times S^2$. The diagram summarizes the causal structure of Mikowski:

- Radial null geodesics are lines at $\pm 45°$.
- All null geodesics begins at S^- and terminate to S^+.
- All timelike geodesics begins at i^- and terminate to i^+.
- All spacelike geodesics begins and end at i^0.

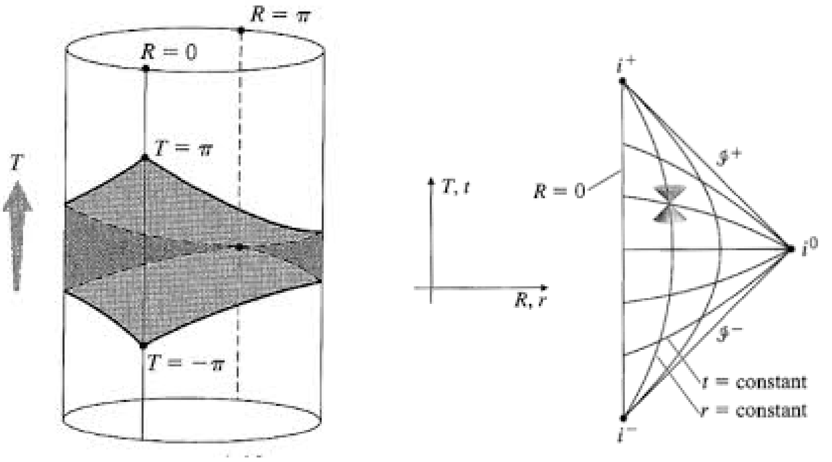

Figure 8.4: Construction of conformal diagram of Minkowski.

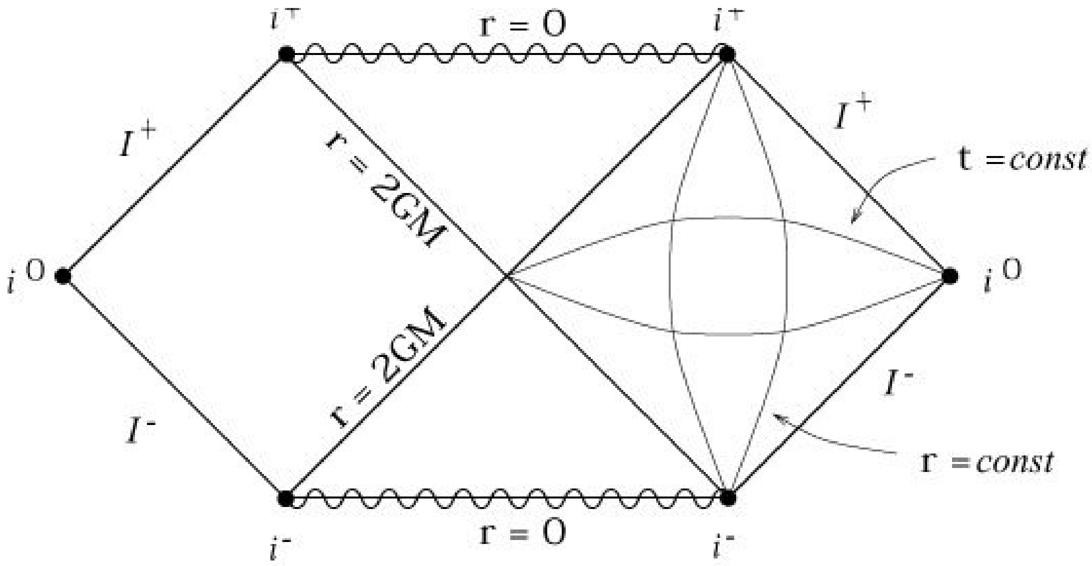

Figure 8.5: Conformal diagram for Schwazrschild.

Schwarzschild. Conformal diagram can be constructed also for Schwarzschild spacetime. The construction proceed similarly, starting from the Kruskall-Szekeres null coordinates

$$\begin{cases} U &= \arctan{(\bar{u}/\sqrt{2M})} \\ V &= \arctan{(\bar{v}/\sqrt{2M})} \end{cases} \quad (U,V) \in (-\pi/2, \pi/2)^2 \ , \quad -\pi/2 < U + V < +\pi/2 \ , \tag{8.88}$$

and finding that at constant angular coordinates the metric is conformally related to the Minkoski metric. The diagram is shown in Fig. (8.5), where the symbols have the same meaning as in the Mikowski conformal diagram. Note in particular that

- Radial null geodesics are lines at $\pm 45°$.
- $i^{\pm}$ are distinct from $r = 0$.
- Conformal infinity is the same as in Mikwoski because Schwarzschild is asymptotically flat.

8.8 Interior nonvacuum solutions: spherical stars

The simplest models of star solutions in GR are usually obtained assuming perfect fluid stress-energy tensor

$$T_{ab} = (\rho + p)u_a u_b + p g_{ab} \ . \tag{8.89}$$

In spherical symmetry the exterior solution is the Schwzraschild metric but the interior must be determined by EFE with matter terms. For a static fluid, the fluid's 4-velocity is assumed to be in the same direction as the timelike KV. In particular one takes for consistency with the metric $u_a = (\mathrm{d}t)_a$. The equations for stellar equilibrium can be found combining EFE and the EOM for the stress-energy tensor. For a perfect fluid, these equations are not sufficient to determine the matter and metric configuration and an *equation of state (EOS)* in the form

$$p = p(\rho) \tag{8.90}$$

must be assumed to close the system. The metric reads

$$g = -e^{-2\alpha(r)}\mathrm{d}t^2 + \left(1 - \frac{2m(r)}{r}\right)^{-1} \mathrm{d}r^2 \ , \tag{8.91}$$

and the metric coefficients and matter fields are determined by the *Tolmann-Oppenheimer-Volkoff (TOV) equations*

$$\frac{dm}{dr} = 4\pi\rho r^2 \tag{8.92a}$$

$$\frac{d\alpha}{dr} = \frac{m + 4\pi r^3 p}{r(r-2m)} \approx \frac{m}{r^2} \tag{8.92b}$$

$$\frac{dP}{dr} = -(p+\rho)\frac{m + 4\pi r^3 p}{r(r-2m)} \approx -\frac{\rho m}{r^2} \ . \tag{8.92c}$$

The above equations are a system of coupled ODE that must be integrated from $r = 0$ to the matter surface R by specifying a central pressure (or density) $P(r = 0) = P_c$, $m(0) = 0$ and an arbitrary constant $\alpha(0) = \alpha_0$. The matter surface is defined by the condition of vanishing pressure $P(R) = 0$. It is immediate to note that if $p \geq 0$ and $dp/d\rho \geq 0$, then the pressure solution of the TOV equation is a monotonically descreasing function The arbitrary constant α_0 is fixed afterwards by demading that $\alpha(R)$ matches with to the Schwarzschild metric (continuity of metric functions). Note that the $m(R) = M$, which is the mass then appearing in the exterior metric.

The meaning of the TOV equations should be clear from their Newtonian limit that are also shown in the r.h.s. of the above equations. The integral expression for the mass function

$$M = m(R) = 4\pi \int_0^R \rho r^2 dr \ , \tag{8.93}$$

is formally identical to the Newtonian equation. However, in GR the proper volume on the spatial hypersurfaces is not $4\pi r^2 dr$ but $4\pi\sqrt{g_{rr}}r^2 dr$. Thus the quantity

$$M_p = 4\pi \int_0^R \rho r^2 \left(1 - \frac{2m(r)}{r}\right)^{-1/2} dr =: M + E_b \ , \tag{8.94}$$

represents the *proper mass* and the difference $E_b = M_p - M > 0$ is often interpreted as the gravitational binnding energy. The equation for α is the generalization of the Poisson equation for the Newtonian potential in spherical symmetry. The equation for the pressure describes the hydrostatic equilibrium. Notably, the r.h.s. of the GR equation is has always larger magnitude than the respective Newtonian equation, indicating that for a given density the GR pressure is always larger.

In general, the TOV equations must be solved numerically for a given EOS. A simple example of numerical integration is linked at the course webpage. However, there are example of analytical solutions, the simplest being for an idealized constant density star and also due to Schwarschild (1916). A simple calculation imposing $\rho(r) \equiv \rho_c$ leads to [exercise]

$$p(r) = \rho \left[\frac{(1-2M/R)^{1/2} - (1-2Mr/R^3)^{1/2}}{(1-2Mr^2/R^3)^{1/2} - 3(1-2M/R)^{1/2}}\right] \ \Rightarrow \ p_c = \rho_0 \left[\frac{1-(1-2M/R)^{1/2}}{3(1-2M/R)^{1/2}}\right] \approx (\frac{\pi}{6})^{1/3} M^{2/3}\rho_0^{4/3} \ , \tag{8.95}$$

where again the Newtonian limit ($r \gg M$) for the central pressure is indicated. The GR equation for p_c however, shows that the denominator becomes infinite for radii $R = 9M/4$. No uniform star can exist with such a radius (mass) given the mass (radius). This result is actually valid for a generic star with EOS such that $\rho \geq 0$, $p \geq 0$ and $dp/d\rho \geq 0$ and known as *Buchdahl limit*. There exist a maximum star mass for a given radius set by

$$M \leq \frac{4}{9}R \ , \tag{8.96}$$

which is independent on the EOS (provided the latter satisfies reasonable consitions).

TOV equations are typically employed for determining the equilibrium startcture of compact stars, i.e. stars with significant self gravity. A primary application is the calculation of mass radius of *neutron stars* that are stars of masses $M \sim M_\odot$ and size $R \sim 10$ km. The self-gravity of neutron stars is thus $\sigma = 2GM/Rc^2 \sim 0.2$. Thus, these systems are the most compact stars that exists and are close to black holes (vacuum) in terms of gravity. Neutron stars were predicted by Landau and Baade and Zwicky in 1933 as a product of the gravitational collapse of massive stars. Their density is comparable of higher than those of nuclei $\rho \sim M/R^3 \sim 10^{15}$ g/ccm and their composition is unknown since no first-principle calculations can be performed for the matter in those density regime. The esistance of neutron star has been confirmed by many and different astrophysical observations, from pulsars to gravitational waves. The mass-radius diagram for neutron stars is shown in Fig. (8.6). Note all the different models have a maximum mass.

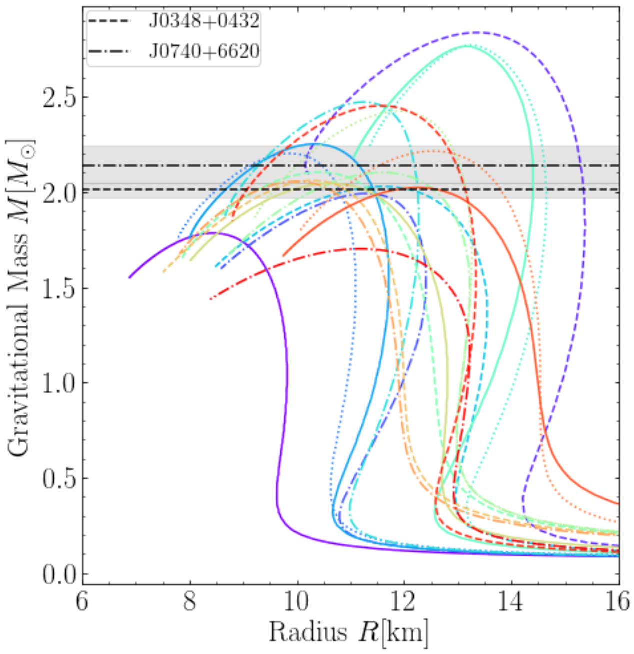

Figure 8.6: Mass-radius diagram for neutron stars with different EOS.

9. Cosmology

③

What spacetime GR predicts for the Universe? These lectures give the basic answer discussing the Robertson-Friedmann-Walker metric and the Big bang.

Suggested readings. *Chap. 5 of Wald (1984); Chap. 8 of Carroll (1997); Chap. 12 of Schutz (1985)*

9.1 Cosmological scales

GR applies to the description of system with mass M and size R such that

$$\sigma = \frac{2GM}{Rc^2} \sim 1 \ . \tag{9.1}$$

The above condition is met for

1. Isolated systems made of compact objects like black holes and neutron stars, eventually in binaries and moving toward each other ($M \sim const$ and R rapidly decreasing);
2. Homogeneous systems at sufficiently large distances, since if the matter density is constant, then $M \sim \rho R^3 \Rightarrow M/R \sim \rho R^2$.

Schutz (1985) defines these two situations as: R becomes smaller faster then M, and M becomes large faster than R. The second case clearly applies to the Universe: at sufficiently large scales one expects that GR should be needed for its description. But, at what scales?

In CGS units: $G \simeq 6.674 \times 10^{-8}$, $c \simeq 2.998 \times 10^{10}$ and $\mathrm{M}_\odot \simeq 1.989 \times 10^{33}$ that imply $G\mathrm{M}_\odot/c^2 \simeq 1.477$ km. Since $1\mathrm{pc} \simeq 3.086 \times 10^{18}$ cm one gets the handy formulas

$$\sigma \approx 2 \times 1.477 \left(\frac{M}{1\ \mathrm{M}_\odot}\right)\left(\frac{1\ \mathrm{km}}{R}\right) \ , \quad \sigma \approx 10^{-13}\left(\frac{M}{1\ \mathrm{M}_\odot}\right)\left(\frac{1\ \mathrm{pc}}{R}\right) \ . \tag{9.2}$$

- Galaxy: $R \sim 15$ kpc $= 10^4$ pc and $M \sim 10^{12}\ \mathrm{M}_\odot$, $\sigma \sim 10^{-5}$. A galaxy is as relativistic as the Solar system.
- Galaxies cluster: $R \sim$ Mpc $= 10^6$ pc with 10^3 galaxies, $\sigma \sim 10^{-4}$. At these distances one either assumes (cosmological principle) or start to observe homogeneity. Hence, the interesting scales for GR are
- Cosmological distances: $R \gtrsim$ Gpc $= 10^9$ pc.

9.2 Observations

The key observation for modern cosmology is Hubble's 1929 measurement (and later ones) of the Universe's expansion. The velocity of galaxies is proportional to their distance from Earth and it is larger the farther the galaxies are (Fig. (9.1)), i.e. galaxies are moving far away from each other and the Universe is expanding,

$$v = H_0 d. \tag{9.3}$$

The recession velocity was measured from the redshift of light from Cepheid variables $z \approx v/c$, while the distance is derived from the intrinsic luminosity of standard candles (supernovae). Note that the GR prediction of an expanding Universe is due to Alexander Friedmann 1922 and independently to Georges Lemaˆitre in 1927 that also predicted the proportionaly between velocity and distance. The currently most precise measurements of H_0 areprovided by the Hubble's SH0ES (Supernovae H0 for the Equation of State) experiment based on observations of Cepheids in six reliable hosts of Type Ia supernovae, and by the Planck experiment on CMB (see below). The measured values are in 5σ disagreement

$$H_0 \sim 73.5 \pm 1.4\ \mathrm{km/s/Mpc}\ (\text{Hubble SH0ES})\ , \quad H_0 \sim 67.4 \pm 0.5\ \mathrm{km/s/Mpc}\ \ (\text{CMB})\ . \tag{9.4}$$

Note these are very different measurements: one refers to the Universe expansion rate today, the other is based on the early Universe expansion.

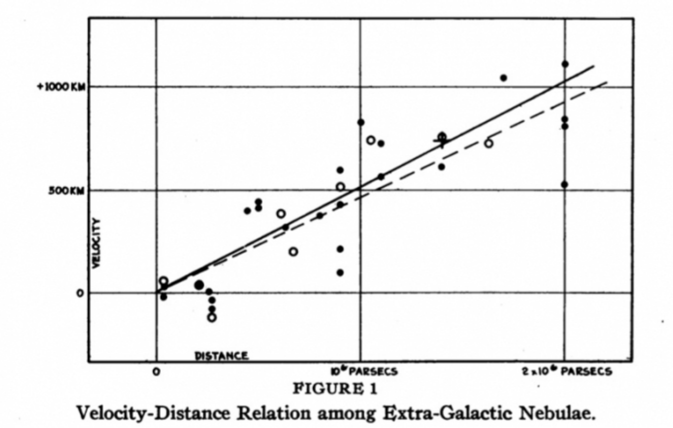

Figure 9.1: Hubble's observation (1929)

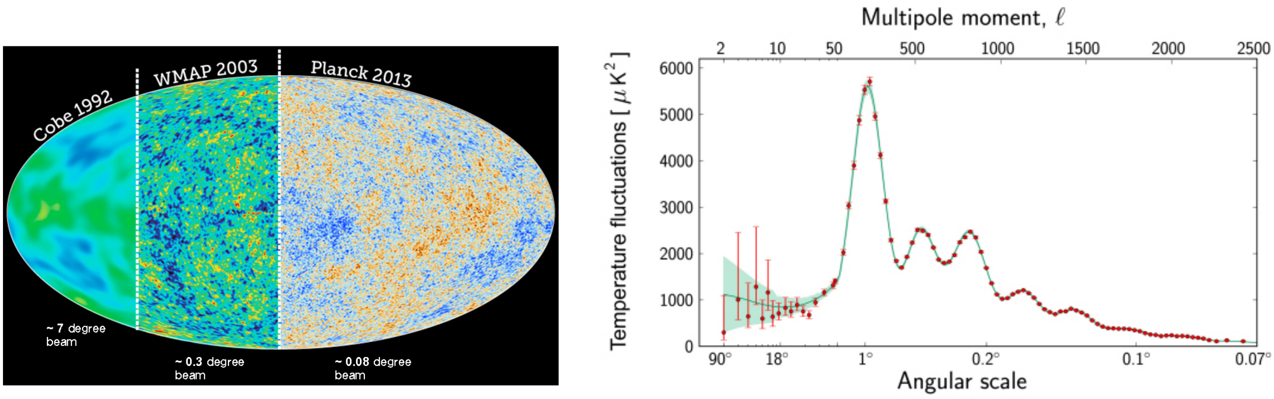

Figure 9.2: Left: cosmic microwave background (CMB) observations form various experiments (1989-2001-2013). Right: Temperature power spectrum measured by Planck.

A second key observation is the discovery of *cosmic microwave background* (CMB) by Arno Penzias and Robert Wilson (1964), the oldest radiation electromagnetic radiation in the Universe with a temperature of $T \sim 2.74$ K and a black body spectrum. This is considered generated in a early epoch when electrons and protons formed neutral hydrogen atoms ("recombination") in an excited state, the electron decayed to the ground state and the photon escaped. The process of photons escaping from these new nuclei is called radiation decoupling. The CMB was predicted in 1948 da George Gamow, Ralph Alpher and Robert Herman and it is considered key evidence for the Big Bang model [1]. The CMB is studied by looking at the power spectrum of the multipoles $|c_{\ell m}|$ of the temperature fluctuation,

$$\frac{\Delta T}{T}(\theta, \phi) = \sum_{\ell m} c_{\ell m} Y^{\ell m}(\theta, \phi) \ . \tag{9.5}$$

It is found that the radiation is isotropic in one part over 100,000, temperature fluctuations are of the order $\sim \mu K$. The temperature fluctuations correlate to density variations in the Universe at 370,000 years old, which in turn relate to the structure of galaxies and galaxy clusters today (13.8 billion years later). Temperature fluctuations can have various origins and can be very well predicted using the ΛCDM (Λ-Cold Dark Matter) model and studying perturbations of the homogeneous background cosmological spacetime (the FRW spacetime discussed below), Fig. (9.2).

9.3 Homogeneity & Isotropy

The concept of homogeneity and isotropy can be made formal by considering the symmetries of the metric. Recalling previous discussions on symmetries and diffeomorphisms, the following definition should not sound new or complicated:

Definition 9.3.1. *A diffeomorsphim $\phi : \mathcal{M} \mapsto \mathcal{M}$ is called an* isometry *iff $\phi_* g_{ab} = g_{ab}$, i.e. the metric does not change under the active coordinate transformations implemented by ϕ.*

Focusing on the spatial section of the manifold $\Sigma_t \subset \mathcal{M}$, the manifold is

[1] The Big Bang nucleosynthesis was proposed by Alpher, Bethe and Gamov. The story of the paper is interesting: `https://en.wikipedia.org/wiki/Alpher-Bethe-Gamow_paper`.

Definition 9.3.2. *Spatially* homogeneous *iff every point $p \in \Sigma_t$ can be connected by an isometry ϕ,*

$$\forall\ p,q \in \Sigma\ \ \exists\ isometry\ \ \phi\ :\ q = \phi(p). \tag{9.6}$$

Definition 9.3.3. *Spatially* isotropic *at each point iff there exists an isometry that*
(i) leaves any point $p \in \Sigma$ invariant;
(ii) given a family (congruence) of timelike curves with tangent vector u^a, leaves u^a invariant; and
(iii) transforms/connects/rotates the vectors orthogonal to u^a one into another.
The vector field u^a defines the worldline of isotropic observers*.*

Observations.
- At one point there is a unique observer that sees the Universe as isotropic.
- Homogeneous+isotropic $\Rightarrow$ a unique Σ_t must be perperdicular to the isotropic observers. If Σ_t is unique and not perpendicular to u^a, then the projection of u^a on Σ_t identifies a preferred direction.
- Because of the isometry it is impossible to construct a preferred direction on the Σ_t.
- A metric can be homogeneous but nowhere isotropic. Example: $\Sigma = \mathbb{R} \times S^2$.
- A metric can be isotropic around a point, but not homogeneous. Example: the cone.
- A metric isotropic everywhere is homogeneous.
- A metric isotropic around a point and homogeneous is everywhere isotropic.

9.4 Robertson-Walker metric

An homogeneous+isotropic spacetime must have structure $\mathcal{M} = \mathbb{R} \times \Sigma_t$ and metric

$$g_{ab} = u_a u_b + \gamma_{ab}(t)\ , \tag{9.7}$$

where u^a define the isotropic observers and γ is the spatial metric on the Σ_t.

Spatial metric. The hypotesis of homogeneity+isotropy further constraint the metric γ. One can prove that

Theorem 9.4.1. *Σ_t is a* maximally symmetric manifold *where the Riemann tensor can be written as*

$$^{(3)}R_{ijkl} = \kappa \gamma_{k[i}\gamma_{j]l}\ , \tag{9.8}$$

and κ is a <u>constant</u> with dimension $[\kappa] = L^{-2}$ proportional to the Ricci scalar. Indeed, from the equation above it follows immediately that

$$^{(3)}R_{ij} = \gamma^{ik\,(3)}R_{ijkl} = {}^{(3)}R^k{}_{jkl} = 2\kappa\gamma_{jl}\ ,\quad {}^{(3)}R = \gamma^{ij\,(3)}R_{ij} = 6\kappa\ . \tag{9.9}$$

Maximally symmetric metrics are determined by the value of κ but there are only three relevant cases

$$\begin{cases} \kappa = 0 & \Sigma = \mathbb{R}^3 \quad \text{3-Euclidean space} \\ \kappa > 0 & \Sigma = S^3 \quad \text{3-sphere} \\ \kappa < 0 & \Sigma = \mathbb{H}^3 \quad \text{3-hyperboloid}\ . \end{cases} \tag{9.10}$$

Since one can always normalize the constant to one (see below) the cases correspond to $\kappa = 0, \pm 1$. The $\kappa = 0$ case is trivial. The metric for the two non trivial cases can be found by immersion in $\mathbb{R}^4$ with the following trick. Start from the 4D Euclidean metric where the 4th coordinate u is constrained on a 3-sphere or 3-hyperboloid

$$\gamma = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \pm \mathrm{d}u^2 \quad \text{with} \quad \pm 1 = x^2 + y^2 + z^2 \pm u^2 = \delta_{ij}x^i x^j \pm u^2\ , \tag{9.11}$$

differentiate the surface

$$0 = 2xdx + 2ydy + 2zdz \pm 2udu = 2\delta_{ij}x^i dx^j \pm 2udu\ , \tag{9.12}$$

and write

$$\pm\, du^2 = \pm\frac{(udu)^2}{u^2} = \pm\frac{(\delta_{ij}x^i dx^j)^2}{1 \mp \delta_{ij}x^i x^j}\ . \tag{9.13}$$

Substitute the above expression in the metric and write it terms of κ

$$\gamma = \sum_i \mathrm{d}x_i^2 + \frac{\kappa}{1-\kappa x^2}(\delta_{ij}x^i \mathrm{d}x^j)\ . \tag{9.14}$$

where $x^2 = \delta_{ij}x^i x^j$. Note the expression is valid also for $\kappa = 0$.

Remark 9.4.1. *The definition of maximally symmetry spacetime extend to arbitrary dimension by extending appropriately the formulas for the Riemann. Expression for the metrics of the n-sphere and n-hyperboloid can be found by immersion in $n+1$ Euclidean space with an analogous trick to the one above. Maximally symmetric spacetime with positive curvature ($\kappa = 1$) are called* de Sitter spacetimes, *those with negative curvature are called* anti de Sitter *spacetime.*

Changing coordinates improves the feeling with this metric. Take first spherical coordinates:

$$\sum_i \mathrm{d}x_i^2 = \mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \ , \quad \delta_{ij}x^i\mathrm{d}x^j = r\mathrm{d}r \ , \tag{9.15a}$$

and substitute them inside the metric

$$\gamma = \sum_i \mathrm{d}x_i^2 + \frac{\kappa}{1-\kappa x^2}(\delta_{ij}x^i\mathrm{d}x^j) \tag{9.15b}$$

$$= \mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 + \frac{\kappa}{1-\kappa r^2}r^2\mathrm{d}r^2 = \frac{\mathrm{d}r^2}{1-\kappa r^2} + r^2\mathrm{d}\Omega^2 \ . \tag{9.15c}$$

Perform now the coordinate transformation

$$d\chi^2 = \frac{dr^2}{1-\kappa r^2} \quad \Rightarrow \quad r = S_\kappa(\chi) = \begin{cases} \sin(\chi) & \kappa = 1 \\ \chi & \kappa = 0 \\ \sinh(\chi) & \kappa = -1 \end{cases} \tag{9.16}$$

to obtain the final form

$$\gamma = \mathrm{d}\chi^2 + S_\kappa(\chi)^2\mathrm{d}\Omega^2 \ . \tag{9.17}$$

Note that

- for $\kappa = 0$ this is the Euclidean metric in spherical coordinates;
- for $\kappa = 1$ one should not think of χ as a “radial coordinate”: χ is the third angle on S^3.
- for $\kappa = -1$ the situation is similar: χ is the hyperboloidal coordinate.

Full metric. The full metric can be constructed as follows

- For a given t, choose γ in the above form;
- Transport the spatial coordinates along the isotropic observers, in such a way each isotropic observer is at located fixed spatial coordinates;
- Label each surface Σ_τ with the proper time (clocks) of the isotropic observers.

The result is the *Robertson-Walker metric*

$$g = -\mathrm{d}\tau^2 + a(\tau)^2\left[\frac{\mathrm{d}r^2}{1-\kappa r^2} + \mathrm{d}\Omega^2\right] \ , \tag{9.18}$$

or

$$g = -\mathrm{d}\tau^2 + a(\tau)^2\left[\mathrm{d}\chi^2 + S_\kappa(\chi)^2\mathrm{d}\Omega^2\right] \ , \tag{9.19}$$

where $a(\tau) > 0$ is called the *scale factor* that determines the volume of $\tau = const$ spatial regions (volume comoving with isotropic observers). Note that Eq. (9.18) is invariant under the transformation

$$a \to \lambda a \ , \quad r \to r/\lambda \ , \quad \kappa \to \lambda^2\kappa \ , \tag{9.20}$$

that implies that the curvature κ can be normalized. If the curvature is normalized, then the scale factor has dimension of length and the radial coordinate r (or χ) is dimensionless. Alternatively, it is possible to work with a dimensionless scale factor, e.g. normalized to the “current” value $a(t) \to a(t)/a(t_0) = a(t)/a_0$, use a radial coordinate with the dimension of distance, $r \to a_0 r$, and work with a curvature parameter with dimension of inverse squared length, $\kappa \to \kappa/a_0^2$.

9.5 Cosmological redshift

The Hubble observation on the expansion of the Universe is based on the redshift of spectral lines of distant galaxies. Consider two isotropic observers $\mathcal{O}_{1,2}$ and the following problem. Observer $\mathcal{O}_1$ emits a photon of momentum k^a in p_1 at τ_1, observer $\mathcal{O}_2$ receives the photon in p_2 at τ_2. Compute the redshift of the photon.

The problem can be solved with few symmetry considerations, (Wald, 1984).

(i) Because of isotropy, there exists a KV that points to the direction of the projection of k^a on Σ_1 and Σ_2. For example, for $\kappa = 0$ the KV are $\partial_x, \partial_y, \partial_z$. If ∂_x is the direction of the k-projection on Σ_1, then $k^a(\partial_y)_a = 0 = k^a(\partial_z)_a$ on Σ_1. But because ∂_i are KVs the product with k^a, which is tangent to null geodesics, must be constant. Hence, $k^a(\partial_y)_a = 0 = k^a(\partial_z)_a$ on Σ_2.

(ii) Call ξ^a the KV discussed in (i), e.g. $\xi = \partial_x$. The norm of the KV is proportional to the scale factor,

$$\frac{\sqrt{\xi_a \xi^a(p_1)}}{\sqrt{\xi_a \xi^a(p_2)}} = \frac{a(\tau_1)}{a(\tau_2)} \ . \tag{9.21}$$

This is obvious for $\kappa = 0$ and true in general.

(iii) The photon's frequency observed by an isotropic observer of 4-velocity u is $\omega = -k^a u_a$. Since k is a null vector, the projection against against a unit timelike vector must be minus the projection against a unit spatial vector,

$$0 = k_a k^a \quad \Rightarrow \quad k_a u^a = -k_a \frac{\xi^a}{\sqrt{\xi_b \xi^b}} \ . \tag{9.22}$$

Inserting (iii)-(i)-(ii) in the ratio of the frequencies gives

$$\frac{\omega_2}{\omega_1} = \frac{k_a u^a(p_2)}{k_a u^a(p_1)} = \frac{\cancel{k_a \xi^a}}{\cancel{k_a u^a}} \frac{\sqrt{\xi_b \xi^b(p_1)}}{\sqrt{\xi_b \xi^b(p_2)}} = \frac{\sqrt{\xi_b \xi^b(p_1)}}{\sqrt{\xi_b \xi^b(p_2)}} = \frac{a(\tau_1)}{a(\tau_2)} \ . \tag{9.23}$$

For an expanding Universe, $a(\tau_2) > a(\tau_1)$ and the photon wavelength (frequency) increases (decreses) proportionally. The redshift is given by

$$z = \frac{\Delta \lambda}{\lambda} = \frac{\Delta a}{a} \ . \tag{9.24}$$

For nearby galaxies, $\Delta\tau = \tau_2 - \tau_1 \approx d/c$ and $a(\tau_2) \approx a(\tau_1) + \Delta\tau \dot{a}(\tau_1) + ...$, which gives the Hubble redshift,

$$z = \frac{a(\tau_2) - a(\tau_1)}{a(\tau)_1} \approx \Delta\tau \frac{\dot{a}(\tau_1)}{a(\tau_1)} \approx cd \frac{\dot{a}}{a} = dH \ . \tag{9.25}$$

Cosmological distance measurement. Consider the measurement of the distance between $\mathcal{O}_1$ and $\mathcal{O}_2$. Light signals emitted by $\mathcal{O}_1$ will be measured by $\mathcal{O}_2$ at a later time $\tau_2 > \tau_1$. In a flat and static spacetime, $\mathcal{O}_2$ measures an electromagnetic flux

$$F = \frac{\text{instrinsic luminosity of the source}}{4\pi\chi^2} = \frac{\dot{E}}{4\pi\chi^2} \ . \tag{9.26}$$

In a RW spacetime the formula needs to be changed to account of several effects

- The area of the spherical wave front is not χ^2 but $S^2_\kappa(\chi)$.
- Each photon is redshifted $E_{\rm rec} = E_{\rm emt}/(1+z)$.
- Delay in the photon's time arrival due to the expansion. Two photons emitted at time interval dt arrive at time interval $dt(1+z)$.

The resulting formula is

$$F = \frac{\dot{E}}{4\pi S^2_\kappa(\chi)(1+z)^2} =: \frac{\dot{E}}{4\pi d_L^2} \ . \tag{9.27}$$

The quantity d_L is called *luminosity distance*. The formula above can be used to determine cosmological parameters given the observable $(F, z, ...)$ and *standard candles*. The latter are sources for which $\dot{E}$ is either known or can be estimated accurately.

9.6 Friedman-Robertson-Walker (FRW) equations

The FRW equations determine the scale factor by solving EFE using the RW metric as ansatz. The solution corresponds to a spacetime

- Homogeneous;
- Isotropic;
- Nonvacuum;

where the matter content is modeled by a perfect fluid with the same velocity as the isotropic observers in such a way that the fluid is at rest in coordinates comoving with isotropic observers.

$$T_{\mu\nu} = \mathrm{diag}(\rho, p, p, p) \ , \quad T^\mu_\nu = \mathrm{diag}(-\rho, p, p, p) \ , \quad T = T^\mu_\mu = -\rho + 3p \ . \tag{9.28}$$

The necessary EFE read

$$\begin{cases} G_{\tau\tau} &= 8\pi T_{\tau\tau} = 8\pi\rho \\ G_\perp &= 8\pi T_\perp = 8\pi p \end{cases} , \tag{9.29}$$

where $\perp$ indicate <u>any</u> equation spatially projected: because of isotropy any projection of the Einstein tensor along a unit spatial direction, $G_\perp = G_{ab}\hat{s}^a\hat{s}^b$, gives equivalent equations.

Sketch of the calculation for the $\kappa = 0$ case. Christoffel:

$$\Gamma^{\tau}_{11} = a\dot{a}\ ,\quad \Gamma^{i}_{1\tau} = \frac{\dot{a}}{a}\ . \tag{9.30a}$$

Ricci:

$$R_{\tau\tau} = -3\frac{\ddot{a}}{a}\ ,\quad R_{11} = a\ddot{a} + 2\dot{a}\ . \tag{9.30b}$$

Spatial projection:

$$s^\mu = (0,1,0,0,0)\ ,\quad s^\mu s_\mu = a^2\ ,\quad \hat{s}^\mu := a^{-2}s^\mu\ ,\quad R_\perp = R_{\mu\nu}\hat{s}^\mu\hat{s}^\nu = a^2\ . \tag{9.30c}$$

Ricci scalar

$$R = -R_{\tau\tau} + 3R_\perp = 6(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a}) \tag{9.30d}$$

Putting things together one gets to the final result (given below). The EOM for matter fields are given by

$$0 = \nabla_\mu T^\mu_0 = \partial_\mu T^\mu_0 + \Gamma^\mu_{\mu\alpha}T^\alpha_0 - \Gamma^\alpha_{\mu 0}T^\mu_\alpha = -\partial_\tau\rho - 3\frac{\dot{a}}{a}(\rho + p)\ . \tag{9.30e}$$

Computing the general case for κ gives the FRW equations

$$\begin{cases} \left(\frac{\dot{a}}{a}\right)^2 + \frac{\kappa}{a^2} & = \frac{8\pi}{3}\rho \\ \frac{\ddot{a}}{a} + \frac{4\pi}{3}(\rho + 3p) & = 0 \\ \dot{\rho} + 3(\rho + p)\frac{\dot{a}}{a} & = 0 \end{cases} \qquad \text{(FRW equations)} \tag{9.31}$$

that need to be solved together with an equation of state in the form $p = p(\rho)$.

9.7 FRW solutions

Let us discuss some general consequences of the FRW equations at early and late times. Explicit analytic solution describing these features can be found in e.g. (Wald, 1984).

Dynamical Universe & Hubble law. The FRW equations in Eq. (9.31) predict that for positive energy density and pressure matter the Universe cannot be static but must expand or contract,

$$\begin{cases} \rho & \geq 0 \\ p & \geq 0 \end{cases} \quad \Rightarrow \quad \ddot{a} < 0 \quad \Rightarrow \quad \dot{a} \lessgtr 0\ . \tag{9.32}$$

Remark 9.7.1. *Expansion or contraction in the cosmological context are refereed to the distance between two isotropic observers at the same τ.*

At fixed τ the *comoving radial distance* between two points is

$$d = a(\tau)\chi \tag{9.33}$$

and the expansion velocity is expressed by the *Hubble law*

$$v = \dot{d} = \dot{a}\chi = \dot{a}\frac{d}{a} = \frac{\dot{a}}{a}d =: Hd\ , \tag{9.34}$$

where it is defined the key quantity (note it is not a constant):

$$H := \frac{\dot{a}}{a} \ \text{ Hubble parameter}\ . \tag{9.35}$$

The comoving radial distance and the expansion velocity are coordinate quantities and can be interpreted as physical only for sufficiently close objects. If the distance is smaller with respect to the Hubble parameter at the moment of the observation $d \ll cH^{-1}(\tau = 0)$, then d can be used as a measure of the spatial distance between two isotropic observers. The inverse of the Hubble parameter is called *Hubble radius*. However, if the distance is larger one must take into account that the two objects are at different τ and d is meaningless. The expansion velocity, in particular, can be larger than the speed of light for distance objects.

Big-bang. Hubble observed that today ($\tau = 0$) the galaxies are moving apart from each other, i.e. $H(0) > 0$ and the derivative of the scale factor is positive. Indicating with a subscript the value at $\tau = 0$,

$$H_0 > 0 \quad \Rightarrow \quad \dot{a}_0 > 0 \ . \tag{9.36}$$

The above result combined with the fact that the scale factor's acceleration is negative (for standard matter), $\ddot{a} < 0$, says that the Universe was expanding faster at earlier times. Hence, the scale factor must have been zero at some earlier time. One can set a lower limit to the time at which $a(\tau) = 0$ by simply extrapolating back at the current rate of expansion, to find

$$T_H := \frac{a_0}{\dot{a}_0} = H_0^{-1} \ \text{Hubble time} \ . \tag{9.37}$$

Sometime at $\tau > T_H$ the Universe was in a state in which $a = 0$ that means
- The distance between all points on Σ is zero;
- Curvature is infinite, $R \sim a^{-2}$;
- Matter had infinite density.

This instant is called *Big Bang* [2].

Matter and radiation. A simple EOS considered in cosmology is

$$p(\rho) = w\rho \quad \text{with} \quad w = \begin{cases} 1/3 & \text{radiation} \\ 0 & \text{dust} \\ -1 & \text{vacuum energy} \end{cases} \tag{9.38}$$

Dust is the simplest model of baryonic matter and can be used for stars and galaxies distributions in which the pressure is negligible. The EOS for radiation is derived in in the example below. Vacuum energy is discussed below. By using the above EOS, the matter equation is solved by

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a} \quad \Rightarrow \quad \rho \propto a^{-3(1+w)} \quad \text{or} \quad \rho a^{3(1+w)} = const \ . \tag{9.39}$$

In particular
- For *dust matter*: $\rho_M a^3 = const.$ This express the conservation of mass: the number density of particles (baryons) must decrease as the Universe expand.
- For *radiation*: $\rho_R a^4 = const.$ The energy density of photons decreases more rapidly than the increase in volume because photons lose energy due to redshift.
- For *vacuum*: $\rho_\Lambda a^0 = const.$

Hence, there is a clear hierarchy
- Radiation is the dominant contribution to matter sources at early times;
- As the Universe expands, the matter contribution decay slower and matter must become dominant. The current observations indicate
$$\frac{\rho_M}{\rho_R} \sim 10^3 \ , \tag{9.40}$$
but fitting these observations require to assume vacuum energy (cosmological constant).
- Both matter and radiation decay faster than a^{-2}, hence in the FRW the term $\rho a^2 \to 0$ as time advance.
- Vacuum energy (if present) dominates over matter and radiation at later times.

Example 9.7.1. Equation of state for radiation. *Radiation is often considered as a gas of relativistic particles described modeled by the perfect fluid stress-energy tensor. At the same time one can consider the radiation as described by the stress-energy tensor of electromagnetism. Simply combining these two modesl gives the radiation EOS. Take the trace of both tensors*

$$T_{ab} = (\rho + 3p)u_a u_b + p g_{ab} \ , \quad T_{ab} = F_{ac}F_b{}^c - \frac{1}{4}g_{ab}F^{dc}F_{dc} \tag{9.41}$$

and equal them to obtain the result,

$$-\rho + 3p = T = 0 \ . \tag{9.42}$$

Open-flat-closed. Consider the FRW equation

$$\dot{a}^2 = \frac{8\pi}{3}\rho a^2 - \kappa \geq 0 \ , \tag{9.43}$$

with $\rho > 0$. The scale factor is currently increasing $a_0 > 0$, but a negative $\ddot{a}$ implies that $\dot{a}$ must decrease. For large times and standard matter and radiation $\rho a^2 \to 0$ and

[2] The name is attributed to physicist Hoyle, an opponent of the Big Bang teory, that used it in 1949 on a BBC radio interview.

- $\kappa = 0 \Rightarrow \dot{a} \to 0$: the Universe is flat.
- $\kappa = -1 \Rightarrow \dot{a} \to 1$: expansion continues, the Universe is open.
- $\kappa = +1$: because the r.h.s. has to remain positive there exists a maximum size $a \geq a_c$, after which the scale factor starts to decrease. The critical value cannot be reached asymptotically because $\ddot{a} < 0$. Hence, the Universe is closed.

Another way to arrive to the above conclusion, is to substitute $\rho = b/a^3$ (b is a constant) in Eq. (9.43) and consider the equation in the form

$$\dot{a}^2 = -V_M(a) - \kappa \geq 0 \quad \text{with} \quad V_M := -\frac{8\pi}{3}\frac{b}{a} \ . \tag{9.44}$$

Then, one can use the method of the effective potential: the Universe exists only in regions (values of a) where $-\kappa$ exceeeds the value of the potential V_M (Cf. the discussion on Schwarschild's orbit). For $\kappa = -1$, the Universe extends to $a(\tau) \to \infty$ with finite asymptotics velocity; for $\kappa = 0$ the Universe extends to $a(\tau) \to \infty$ reaching zero velocity; for $\kappa = +1$ the Universe extends to a maximum scale $a_c = 8/3\pi b$ (the interesection $V_M(a_c) = -1$) at which it reaches a turning point and afterwards re-collapse.

The FRW equation is usually written by defining the *critical density* $\rho_c := 8\pi H^{-2}/3$ and the *density parameter* as

$$\Omega := \frac{\rho}{\rho_c} = \frac{8\pi}{3}\frac{\rho}{H^2} \quad \Rightarrow \quad \Omega - 1 = \frac{\kappa}{H^2 a^2} \ . \tag{9.45}$$

The future geometry of the Universe is then summarized by the following table

κ	ρ	Ω	Geometry
< 0	$< \rho_c$	< 1	open
$= 0$	$= \rho_c$	$= 1$	flat
> 0	$> \rho_c$	> 1	closed

9.8 Horizons

How much Universe can be observed from a point of the Universe? The question can be reformulated by asking: given an event p, which isotropic observers could have sent a signal that reached an isotropic observer at or before p?

Note that the question is not trivial since the Universe started at finite time. However, one would expect that, since $a \to 0$ at the big bang, an isotropic observer could communicate with all the others. The situation is in fact more complicated. To answer the questions above one must determine the particle horizon.

Definition 9.8.1. Particle horizon *at p = boundary of the region that contains worldlines of particles that intersect the past light cone of p.*

Consider a flat universe $\kappa = 0$ and make the coordinate transformation from proper time to *conformal coordinate time*

$$\tau \mapsto t := \int \frac{d\tau}{a(\tau)} \quad \Rightarrow \quad g = a^2(t)(-\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) \ . \tag{9.46}$$

The metric is now *conformally flat*, i.e. proportional to the Minkowski metric via the conformal factor given by the square of the scaling factor. Such a conformally flat metric has the general property that a vector is time-like/null/spacelike iff it has the same property in the flat metric. Hence, the causal structure of the $\kappa = 0$ Universe is the same as the one of Mikowski ... as far as the conformal transformation is valid !

Let us study the validity of the transformation in Eq. (9.46). To this end, it is useful to shift the time of the Big bang to $\tau = 0$, such that $a(0) \to 0$ [3].

- If the integral <u>diverges</u> approaching the big bang ($a \to 0$), then then the RW metric is related to Mikowski all the way down to $t \to -\infty$ and NO particle horizon can exist.
- If the integral <u>converges</u>, then particle horizons can occur.

The integral diverges if $a(\tau \to 0) \propto \tau$ (or in general if $a(\tau \to 0)$ is a linear or slower function). The presence of particle horizon depends on the particular solution of the FRW equations. A summary of such solutions for standard matter content can be found in (Wald, 1984). They show that

- For most of the solutions particle horizons are present.
- For closed Universe $\kappa = +1$, particle horizons cease to exist at the moment of maximum expansion for dust matter, but continue to exist also afterwards for radiation.

Example 9.8.1. *Specifying FRW equations to the EOS $P = w\rho$ and the matter solutions $\rho a^{3(1+w)} = const$, one obtains [exercise]*

$$\dot{a}^2 - \frac{C^{(p)}}{a^{1+p}} + \kappa = 0 \ , \tag{9.47}$$

[3] Only in this section ! In the rest of the lecture $\tau = 0$ is today.

where $p = 0, 1$ *for dust and radiation respectively and* $C^{(p)} = 8\pi/3\rho a^{3+p}$ *are constants. Solutions to the above equation in the flat case* ($\kappa = 0$) *are*

$$a(\tau) = (9C^{(0)}/4)^{1/3}\tau^{2/3} \quad , \quad a(\tau) = (4C^{(1)})^{1/4}\tau^{1/2} \tag{9.48}$$

for dust and radiation respectively. The integral for the conformal time converges also for dust, so there are particle horizons. Note dust is the worst case since for positive pressure a *is larger than for zero pressure.*

9.9 Cosmological constant

Before Hubble observation Einstein looked for a static solution for the Universe. The only way to obtain such a solution is to modify EFE by a term proportional to the metric

$$G_{ab} + \Lambda g_{ab} = 8\pi T_{ab} \ . \tag{9.49}$$

The constant Λ is called the *cosmological constant* and was introduced in 1917.

Observations.

- EFE with cosmological constant are the most general equations for a $(0,2)$ symmetric tensor which has zero divergence and it is build out of the metric's second derivatives.
- The presence of Λ makes EFE incompatible with Newton equations.
- For this reason, the Λ constant is often moved to the r.h.s. and interpreted as *vacuum energy*

$$G_{ab} = 8\pi(T_{ab} - \frac{\Lambda}{8\pi} g_{ab}) = 8\pi(T_{ab} + T_{ab}^{\text{vacuum}}) \ . \tag{9.50}$$

- The vaccum energy $\rho = -\Lambda/8\pi$ can be interpreted as:
  - Fundamental constant of nature
  - Energy of quantum fields in vacuum state
  - Energy of a classical field (*dark energy*)
  - ???
- A positive cosmological constant $\Lambda > 0$ is currently necessary to explain observations of SNa Type Ia, that indicate an expanding universe (Riess et al., 1998). Thus, Λ is employed in the standard cosmological models, but its physical interpretation remains an open problem.

FRW with cosmological constant and static Universe. In presence of cosmological constant the FRW equations read

$$\begin{cases} \left(\dfrac{\dot{a}}{a}\right)^2 & = \frac{8\pi}{3}\rho + \frac{\Lambda}{3} - \dfrac{\kappa}{a^2} \\ \dfrac{\ddot{a}}{a} & = -\frac{4\pi}{3}(\rho + 3p) + \frac{\Lambda}{3} \end{cases} \tag{9.51}$$

A static Universe has $\ddot{a} = 0 = \dot{a}$. Imposing these conditions in the above equations for a dust matter ($p = 0$, $\rho > 0$) gives immediately that (i) the cosmological constant is positive and the (ii) the Universe is spherical ($\kappa > 0$):

$$\begin{cases} \ddot{a} = 0 & \Rightarrow \quad \Lambda = 4\pi\rho > 0 \\ \dot{a} = 0 & \Rightarrow \quad a^2 = \frac{\kappa}{4\pi\rho} \quad \Rightarrow \quad a = +\sqrt{\frac{\kappa}{4\pi\rho}} \end{cases} \tag{9.52}$$

This is *Einstein's spherical static Universe.*

9.10 ΛCDM models

Realistic cosmological models are based on the FRW equations and include radiation, matter, the cosmological constant and perturbations of the primordial homogeneous plasma [4]. We discuss in the following some basic elements.

FRW equation in terms of density parameters. The first FRW equation reads

$$H^2 = \frac{8\pi}{3}(\rho_R + \rho_M + \rho_\Lambda) - \frac{\kappa}{a^2} \ . \tag{9.53}$$

Introducing the *critical density* $\rho_c := 3H^2/(8\pi)$ and the *density parameters* $\Omega_i := \rho_i/\rho_c$ for each matter source $i = R, M, \Lambda$ the above equation reads (divide by H^2)

$$1 = \Omega_R + \Omega_M + \Omega_\Lambda - \frac{\kappa}{a^2H^2} \ , \quad \text{or} \quad \Omega_k := -\frac{\kappa}{a^2H^2} = 1 - \Omega_R - \Omega_M - \Omega_\Lambda \ . \tag{9.54}$$

[4] See `https://lambda.gsfc.nasa.gov/education/graphic_history/univ_evol.cfm` for a summary and Fig. (9.4) below for an illustration.

To interpret observations is useful to have a formula with quantities that are measured today. Use a subscript $_0$ for those

$$a_0 = 1 \ , \quad H_0 \ , \quad \rho_{c0} = \frac{3}{8\pi} H_0^2 \ , \quad \Omega_{i0} = \frac{\rho_{i0}}{\rho_{c0}} \ , \quad \Omega_{k0} = 1 - \sum_i \Omega_{i0} \ , \tag{9.55}$$

and re-express the ρ's in the Friedmann equation in terms of the density parameter measured today by making explicitly the depedence on the scale factor:

$$\frac{8\pi}{3}\rho_R = \frac{8\pi}{3}\rho_{R0}\frac{a_0^4}{a^4} = H_0^2\Omega_{R0}a^{-4} \ , \quad \frac{8\pi}{3}\rho_M = H_0^2\Omega_{M0}a^{-3} \ , \quad \frac{8\pi}{3}\rho_\Lambda = H_0^2\Omega_{\Lambda 0} \ , \quad -\frac{\kappa}{H^2} = H_0^2\Omega_{k0}a^{-2} \ . \tag{9.56}$$

The Friedmann equation then becomes

$$\frac{H^2}{H_0^2} = \Omega_{R0}a^{-4} + \Omega_{M0}a^{-3} + \Omega_{k0}a^{-2} + \Omega_{\Lambda 0} \ , \tag{9.57}$$

and allows one to connect observed quantities with the Hubble parameter and the scale factor.

Evolution. The Universe's evolution is determined by the relative influence of the density parameters:

$$\Omega_R a^{-4} \propto \Omega_M a^{-3} \propto \Omega_k a^{-2} \propto \Omega_\Lambda \ , \tag{9.58}$$

where the curvature density parameter is given by the constraint

$$\Omega_k = 1 - \Omega_R - \Omega_M - \Omega_\Lambda \ . \tag{9.59}$$

The future evolution of the Universe is determined by the Ω_Λ. If $\Omega_\Lambda < 0$ (vacuum energy is negative), then the Universe will decelerate and collapse. If $\Omega_\Lambda \geq 0$, then the Universe will expand unless the matter term will be sufficiently large to halt the expansion before the Ω_Λ takes over. If $\Omega_\Lambda = 0$, then the Universe expand forever $\Omega_M \leq 1$ or collapses if $\Omega_M > 1$.

Remark 9.10.1. *In the presence of a cosmological constant, there is no relationship between the spatial curvature and the fate of the universe: any spatial geometry can expand or recollapse.*

The various possibilities can be investigated by neglecting radiation and studying the equation above at the turning point $H = 0$ that represents the collapse treshold. This gives a cubic equation for a

$$\Omega_{\Lambda 0}a^3 + (1 - \Omega_{M0} - \Omega_{\Lambda 0})a + \Omega_{M0} = 0 \ ; \tag{9.60}$$

real solutions are admitted for

$$\Omega_{\Lambda 0} = \begin{cases} 0 & 0 \leq \Omega_{M0} \leq 1 \\ 4\Omega_{M0}\cos[\frac{1}{3}\arccos(\frac{1-\Omega_{M0}}{\Omega_{M0}}) + \frac{4\pi}{3}]^3 & \Omega_{M0} > 1 \end{cases} \tag{9.61}$$

and are represented as black line in Fig. (9.3). Above the black line the Universe expands, below it contracts. The straight blue line represent a flat universe, $\kappa = 0$. On the right the curvature is positive, $\kappa = 1$; on the left it is negative $\kappa = -1$. The figure shows as arrows the direction of evolution of the parameters in an expanding universe. Comments:

- The attractor poin $(0, 1)$ is a de Sitter space: a universe with no matter density, dominated by a cosmological constant, and with scale factor growing exponentially with time;
- The saddle point at $(0, 0)$ corresponds to an empty universe (not ours);
- The repulsive point $(1, 0)$, is known as the Einstein-de Sitter solution.
- A Universe at a point of the diagram at a given time can return on that point following the same trajectory by expanding to infinity and recollapsing.
- A universe with initial conditions located at a generic point on the diagram will, after several expansion times, flow to de Sitter space if it began above the recollapse line, and flow to infinity and back to recollapse if it began below that line.

Since our universe has expanded by many orders of magnitude since early times, it must have begun at a non-generic point in order not to have evolved either to de Sitter space or to a collapse. Inflation provides a mechanism whereby the universe can be driven to the line $\Omega_M + \Omega_\Lambda = 1$ (flatness), thus favouring Einstein-de Sitter geometry with $\Lambda = 0$.

Current observations indicate

- $\Omega_{R0} \sim 10^{-4}$
- $\Omega_{M0} \sim 0.3$
  - $\Omega_{M0} \sim 0.03$ baryons
  - $\Omega_{M0} \sim 0.27$ dark matter
- $\Omega_{\Lambda 0} \sim 0.7$
- $\Omega_{k0} \lesssim 0.01$

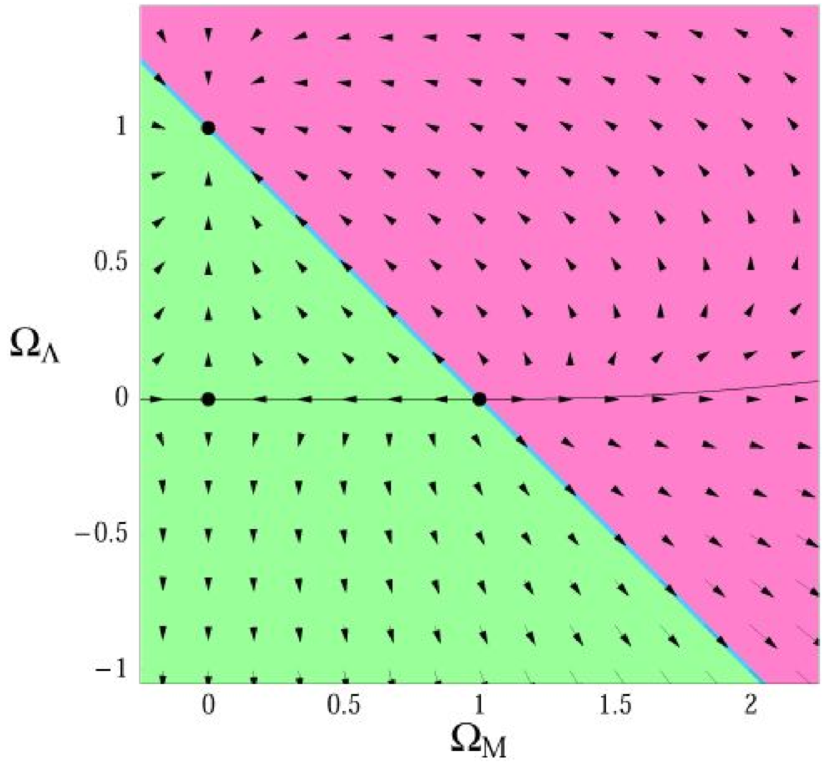

Figure 9.3: Dynamics of an expanding Universe in the $\Omega_\lambda - \Omega_M$ plane. The arrows indicate the direction of evolution of the parameters in an expanding universe. From (Carroll, 2001).

A favoured point in Fig. (9.3) is thus $(0.3, 0.7)$. Note however, the above numbers are inferred from different datasets and very different observations (not from a single "fit"). For example, the density of baryonic matter is estimated from the light elements abundances (Big bang nucleosynthesis), observation of Lyman-alpha forest absoprtion lines (absorption of light emitted from very distant quasars by intervening gas) and CMB. Dark matter density is estimated from gravitational measurements like Galaxies rotational curves and galaxy cluster: weak lensing, structure formations, etc. Several, but not all, observations are compatibile with a flat (or close-to-flat) Universe. Often, matter densities are constrained by assuming a flat geometry.

9.11 Inflation

Two main problems arise confronting the FRW model above with observations:

1. *Flatness problem.* The points $\Omega \sim 1$ ($\kappa \sim 0$) are unstable points in the diagram of Fig. (9.3): any small deviation is expected to grow rapidly and bring the Universe to another geometry. In other terms, the Friedmann equation predicts an expanding Universe dominated by curvature since the relative weight of the r.h.s. terms in absence of vacuum energy ($\rho_\Lambda = 0$) is

$$\frac{\kappa a^{-2}}{8\pi(\rho_M + \rho_R)/3} \sim \frac{\kappa a^{-2}}{a^{-3}} \gg 1 \quad (\text{for} \quad \kappa \neq 0) \ . \tag{9.62}$$

   A flat Universe is not expected.
2. *Horizon problem.* CMB is isotropic to a high degree of accuracy. The natural explanation for this is that the radiation in the Universe had the possibility to interact and thermalize during the recombination epoch. However, this is incompatible with the presence of particle horizons in FRW solutions, which are generically predicted in that epoch.

A solution to both problem is provided by assuming that a early times (*inflation period*) the dynamics of the Universe was characterized by a fast expansion with $\ddot{a} > 0$. This was proposed by Guth in 1980 and further explored by Linde, Albrecht and Steinhardt. A rapid early expansion would provide a mechanism to (i) drive $\Omega \to 1$ quickly, for example with an effective density $\rho_\phi \sim a$; and (ii) "spread out" regions with the same matter/radiation conditions, initially close, to large distances by keeping the same matter/radiation conditions. Moreover, inflation could provide a mechanism to seed the structures observed today in the Universe. Quantum fluctuation during inflations could generate CMB anisotropies and should be thus consistent with those observations. However, the inflation must be driven by some field other than standard matter and radiation, for example scalar fields, vacuum energy, etc. whose precise origin has not been identified.

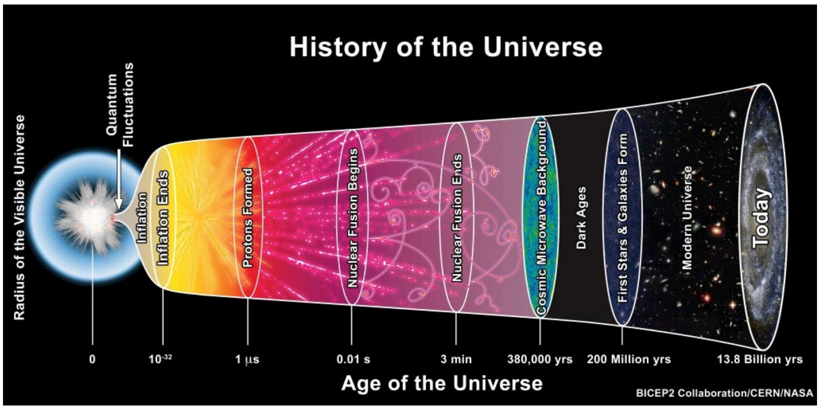

Figure 9.4: Picture of standard model for cosmology. See `https://lambda.gsfc.nasa.gov/education/graphic_history/univ_evol.cfm`.

9.12 Discussion on ΛCDM

Points for discussions on ΛCDM.
SUCCESSES:

- ✓ Hubble observation
- ✓ Type Ia observations
- ✓ GR ($\tau \gtrsim 1$ s quantum effects over)
- ✓ Nucleosynthesis (first 3')
- ✓ He abundancies
- ✓ re-combination epoch (radiation decouple, matter dominated)
- ✓ CMB
- ✓ Galaxies formation and structure
- ✓ Constraints on weakly interacting particle masses
- ✓ Flatness+horizon problem → Inflation

DIFFICULTIES:

- ✗ Origin of dark matter
- ✗ Inflation mechanism
- ✗ Origin of Λ/dark energy
- ✗ Tension among estimated cosmological parameters from different observations, e.g. H_0.
- ✗ Cosmological constant problem $\rho_{\rm vacuum}/\rho_\Lambda \sim 10^{120}$
- ✗ Asymmetry matter/antimatter
- ✗ Early phase, curvature $\sim 1/\ell_{\rm Planck}$ ($\tau \sim 10^{-43}$ s): GR does not apply, quantum gravity ?

Bibliography

Baer, C., Ginoux, N., and Pfaeffle, F. (2008). Wave Equations on Lorentzian Manifolds and Quantization. *arXiv e-prints*, page arXiv:0806.1036.

Carroll, S. M. (1997). Lecture notes on general relativity.

Carroll, S. M. (2001). The Cosmological constant. *Living Rev. Rel.*, 4:1.

Di Casola, E., Liberati, S., and Sonego, S. (2015). Nonequivalence of equivalence principles. *Am. J. Phys.*, 83:39.

Jackson, J. D. (1975). *Classical electrodynamics; 2nd ed.* Wiley, New York, NY.

Landau, L. D. and Lifschits, E. M. (1975). *The Classical Theory of Fields*, volume Volume 2 of *Course of Theoretical Physics*. Pergamon Press, Oxford.

Maggiore, M. (2007). *Gravitational Waves. Vol. 1: Theory and Experiments.* Oxford Master Series in Physics. Oxford University Press.

Misner, C. W., Thorne, K. S., and Wheeler, J. A. (1973). *Gravitation.* W. H. Freeman, San Francisco.

O'Neill, B. (1983). *Semi-Riemannian Geometry With Applications to Relativity, 103, Volume 103 (Pure and Applied Mathematics).* Academic Press.

Riess, A. G. et al. (1998). Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.*, 116:1009–1038.

Schutz, B. F. (1985). *A First Course in General Relativity.* Cambridge Univ. Pr., Cambridge, UK.

Thorne, K. S. (1980). Multipole Expansions of Gravitational Radiation. *Rev. Mod. Phys.*, 52:299–339.

Wald, R. M. (1984). *General relativity.* The University of Chicago Press, Chicago.

Will, C. M. (2014). The Confrontation between General Relativity and Experiment. *Living Rev. Rel.*, 17:4.