

1 Historic primer

Imagine, you want to leave a celestial body with radius R . What speed does one need? The total energy E of the rocket needs be positive:

$$\begin{aligned} E &= \frac{1}{2}mv^2 + V(R) \\ 0 &= \frac{1}{2}mv^2 - \frac{GmM}{R} \\ v &= \sqrt{\frac{2GM}{R}} \end{aligned} \tag{1}$$

For Earth, this is 11.2 km/s, for the Sun 618 km/s. Apparently, the bigger body has a larger escape speed. Now imagine, the sphere gets larger, at constant density (unrealistic!) - at what radius would the escape speed reach the speed of light?

$$\begin{aligned} c &= \sqrt{\frac{2G \cdot 4/3\pi r^3 \rho_{\odot}}{r}} \\ &= r \sqrt{\frac{8\pi G}{3} \rho_{\odot}} \\ r &= 485R_{\odot} = 2.5 \text{ AU} \end{aligned} \tag{2}$$

From such a star, light would not be able to escape, and it would need to be black. It was John Michell, a British natural philosopher, who wrote down this thought experiment first in 1783. In his words:

If there should really exist in nature any bodies, whose density is not less than that of the sun, and whose diameters are more than 500 times the diameter of the sun, since their light could not arrive at us; or if there should exist any other bodies of a somewhat smaller size, which are not naturally luminous; of the existence of bodies under either of these circumstances, we could have no information from light; yet, if any other luminous bodies should happen to revolve about them we might still perhaps from the motions of these revolving bodies infer the existence of the central ones with some degree of probability, [...].

Independent and without knowing from from Michell, Simon Laplace wrote in 1796 in his "Exposition du Système du Monde":

The gravitation attraction of a star with a diameter 250 times that of the Sun and comparable in density to the earth would be so great no light could escape from its surface. The largest bodies in the universe may thus be invisible by reason of their magnitude.

He also provided a mathematical proof, i.e. a calculation similar to the above.

It took than more than 100 years until Albert Einstein in 1916 published the "General theory of relativity", and in the same year Karl Schwarzschild published (to the surprise of Einstein) an exact solution to Einstein's equations. The solution worked beautifully for the solar system, but it also predicted that compact objects would be dark stars. But they considered it more a curiosity of the theory than a reality of nature.

It was only in the 1960's that the topic came into focus of the scientific community.

- 1939, work from Richard Tolman, Robert Oppenheimer and George Volkoff showed that an upper limit for the mass of a neutron star exists for it to be stable. For heavier, compact objects no stabilizing force against gravity is known.
- In 1963, the New Zealander Roy Kerr presented a solution that corresponds to a rotating black hole.

- In 1965, Roger Penrose showed that black holes actually can form (and are not an artefact of the symmetry assumed in the calculations). The key concept was that of "trapped surfaces", which was honored with the 2020 Nobel prize in physics.
- 1967 that the American physicist John Wheeler coined the term 'black hole', replacing the term 'completely collapsed objects'.
- Donald Lynden-Bell and Martin Rees proposed in 1971 that in every galaxy an active or dormant, massive black hole resides - and also in our Milky Way.
- 1972 Tom Bolton was able to convincingly identify the first stellar-mass black hole in the Milky Way: Cygnus X-1
- In 2002 the team around Reinhard Genzel determined the mass of Sgr A* from the orbit of a star around it, excluding essentially all other possibilities than that it is a massive black hole with 4 million solar masses.
- 2015 the LIGO gravitational wave experiment discovered its first event, a merger of two black holes of 29 and 36 solar masses
- In 2019, the event horizon telescope collaboration published its first resolved image of a black hole, in the center of the galaxy M87, with a mass of 6.5 billion solar masses.

The scope of this lecture is to understand classical black holes, and get to know the key observations of these objects. One groups black holes typically by mass

- Particle physics scale: Black holes with masses reachable via particle accelerators, or from cosmic ray interactions. Due to their Hawking radiation these should be very bright emitters and short-lived. We don't have any experimental evidence for their existence
- Primordial black holes: Black holes could have formed directly during the big bang. While masses below 4×10^{11} kg should have evaporated since the big bang, masses larger than that would still be around. In particular, planetary masses (10^{24} kg) are being discussed as possible dark matter candidates. No direct evidence for these black holes has been found.
- Stellar mass black holes have been found in many stellar systems - historically mostly in binary systems, where unseen companions were sometimes too heavy to be a neutron star; sometimes also with accretion disks visible in the X-ray regime. Nowadays such black holes are also seen in gravitational waves, when two such objects merge.
- Intermediate mass black holes: Beyond a few 100 up to $10^5 M_{\odot}$ there is some marginal evidence for such black holes, mostly in globular clusters. These objects are attractive to explain the even heavier counterparts in merger trees.
- (super-) massive black holes: Almost all galaxies host in their centers a massive black hole, the mass of which scales with galaxy properties. The most prominent example is Sgr A* in our own Milky Way.

'Black Holes' is a booming field of research (figure 1), and it has diversified into many subbranches. On the theoretical side, black holes might be the entry into the world of quantum gravity, which is beyond the scope of this lecture. On the observational side, black holes are building blocks of the universe, with important roles in galaxy formation and growth regulation. The Galactic Center black hole is used for tests of general relativity, and gravitational wave detections question formation channels of stellar-mass black holes.

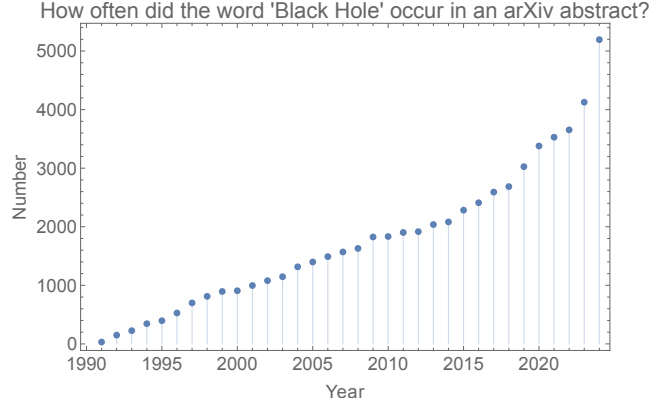


Figure 1: The number of occurrences of the word 'Black Hole' in abstracts submitted to arXiv as a function of year.

2 Tensor algebra - the maths of general relativity

Here is a collection of useful definitions and relations, introducing also the canonic notation for general relativity.

2.1 Euclidean, Cartesian coordinates

Coordinates are

$$\vec{x} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = x^1 \vec{e}_1 + x^2 \vec{e}_2 + x^3 \vec{e}_3 = \sum_{i=1}^3 x^i \vec{e}_i = x^i \vec{e}_i \quad (3)$$

Note that the numbers to top right of the x are not "to the power of", but coordinate indices. Some care and understanding is needed, when one reads a symbol like x^2 . For the unit vectors we have:

$$\vec{e}_i \cdot \vec{e}_j = \delta_{ij} \quad (4)$$

The Kronecker- δ is here the 3D unity matrix. The Euclidean dot product is

$$\begin{aligned} \vec{x} \cdot \vec{y} &= x_1 y_1 + x_2 y_2 + x_3 y_3 \\ &= \sum_{i=1}^3 x_i y_i = x_i y^i \\ &= (x_1, x_2, x_3) \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ &= \vec{x} \cdot g \cdot \vec{y} = g_{ij} x^i y^j \end{aligned} \quad (5)$$

Note the Einstein summation convention: Indices which appear twice, once upper and once lower, are automatically summed over. One can always change the name of such an index pair, as it is "dummy" (like an integration variable in an integral). This notation also gives a convenient way of checking validity of equations: Both sides need to have the same indices in upper and lower positions, after the summations are executed. Further, one has for tensors

$$A^i{}_i = g^{ij} A_{ji} = \text{tr}(A^i{}_j) \quad (6)$$

The use of latin letters for indices indicates 3D, space vectors.

2.2 4D space-time coordinates

The coordinates are

$$x^\mu = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} ct \\ \vec{x} \end{pmatrix} \text{ or for spherical coordinates : } x^\mu = \begin{pmatrix} ct \\ r \\ \theta \\ \phi \end{pmatrix} \quad (7)$$

And with the unit vector e_μ the vector x is:

$$x = x^0 e_0 + x^1 e_1 + x^2 e_2 + x^3 e_3 = \sum_{\mu=0}^3 x^\mu e_\mu = x^\mu e_\mu \quad (8)$$

For 4-vectors, we use greek indices.

2.2.1 Minkowski space

The dot product $x \cdot y$ is defined via

$$x \cdot y = \eta_{\mu\nu} x^\mu y^\nu \quad (9)$$

with the Minkowski metric:

$$\eta_{\mu\nu}^{\text{Cartesian}} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ and } \eta_{\mu\nu}^{\text{Spherical}} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2 \theta \end{pmatrix} \quad (10)$$

For a line element one has

$$\begin{aligned} ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu &= -c^2 dt^2 + dx^2 + dy^2 + dz^2 \\ &= -c^2 dt^2 + dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \end{aligned} \quad (11)$$

Note that there are different conventions found in the literature: The signs of the metric might be opposite, the c^2 could be part of the metric or in the definition of the 0-components of the coordinates (as here), or even an imaginary i is used sometimes to express the opposite sign of the time component compared to the spatial components.

2.2.2 Curved space-time coordinates

The components of a vector are as usual

$$\begin{aligned} x &= x^\mu e_\mu \\ dx &= dx^\mu e_\mu \end{aligned} \quad (12)$$

with the novelty that the base vectors e_μ are not constants, but can be functions of the coordinates. The dot product gets generalized by going from $\eta_{\mu\nu}$ to $g_{\mu\nu}$, which can also be of a more complicated functional form. The dot product for curved manifolds is defined via

$$\begin{aligned} A(x) \cdot B(x) &= g_{\mu\nu}(x) A^\mu(x) B^\nu(x) = A_\nu(x) B^\nu(x) \\ A \cdot B &= g_{\mu\nu} A^\mu B^\nu = A_\nu B^\nu \end{aligned} \quad (13)$$

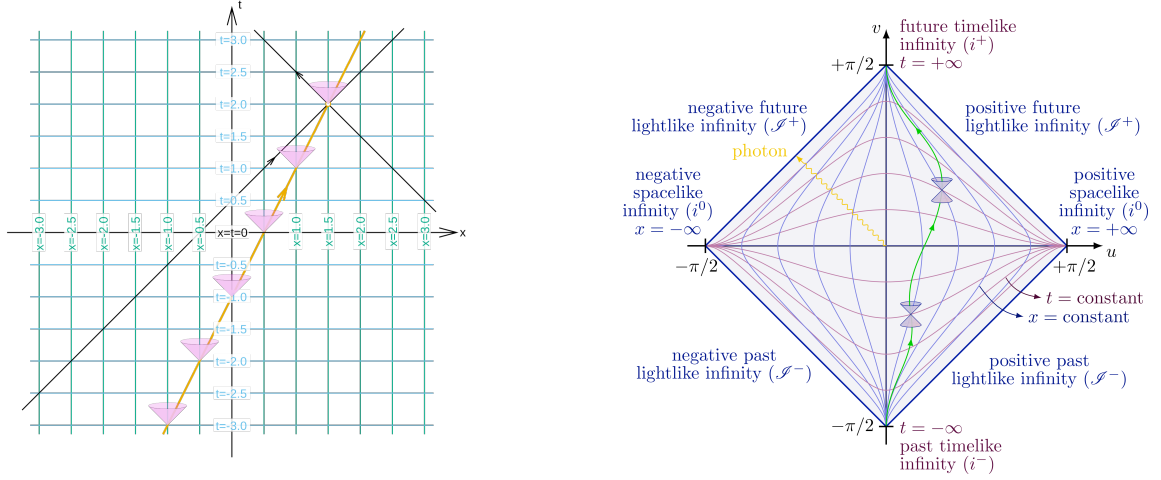


Figure 2: Two representation of a (1-1) Minkowski space-time. Left: Flat coordinates. The worldline of an observer at constant velocity is shown, together with the lightcones, giving the regions of space-time, which causally can connect to the respective event. Light rays travel diagonally, time-like trajectories are more vertical than the light rays. Right: Again Minkowski space-time, plotted in a Penrose diagram with coordinates (u, v) , in which are defined by $r + ct = \tan(u + v)$, $r - ct = \tan(u - v)$. This form of the diagram is useful to describe black holes later. Source: German Wikipedia and TikZ.

Note that the two vectors need to be evaluated at the same space-time point. So one can always calculate a vector length, but (in general) not the cross product of two space vectors X, Y . The metric tensor $g_{\mu\nu}$ is

$$\begin{aligned}
 g_{\mu\nu} &= e_\mu e_\nu = e_\nu e_\mu = g_{\nu\mu} \\
 \delta_\nu^\rho &= e^\rho e_\nu = g^{\rho\mu} e_\mu e_\nu = g^{\rho\mu} g_{\mu\nu} = g_\nu^\rho \\
 g_{\mu\nu} &= (g^{\mu\nu})^{-1}
 \end{aligned} \tag{14}$$

The metric tensor is thus symmetric. The last line follows from the second line, noting that the Kronecker- δ here is the 4D unity matrix. With that one gets the line element

$$ds^2 = (dx^\mu e_\mu)(dx^\nu e_\nu) = e_\mu e_\nu dx^\mu dx^\nu = g_{\mu\nu} dx^\mu dx^\nu \tag{15}$$

which is thus a generalization of the usual Euclidean Pythagorean theorem for infinitesimal paths in curved space-time. The coordinate transformation (changing $x \rightarrow x'$) of a vector A ($\rightarrow A'$) is

$$\begin{aligned}
 A'^\alpha &= \frac{\partial x'^\alpha}{\partial x^\beta} A^\beta \\
 A'_\alpha &= \frac{\partial x^\beta}{\partial x'^\alpha} A_\beta
 \end{aligned} \tag{16}$$

For the example of velocity V :

$$V'^\alpha = \frac{\partial x'^\alpha}{\partial \tau} = \frac{\partial x'^\alpha}{\partial x^\beta} \frac{\partial x^\beta}{\partial \tau} = \frac{\partial x'^\alpha}{\partial x^\beta} V^\beta \tag{17}$$

2.2.3 Note on derivatives

Derivatives are written as:

$$\frac{\partial}{\partial x^\mu} = \partial_\mu = \left(\frac{1}{c} \frac{\partial}{\partial t} \right) \quad (18)$$

The ∂_μ is a very useful notation. In many books, one also finds the notation $X_{\mu,\nu}$ for $\partial_\nu X_\mu$. Here, we don't use it. The 4-gradient is

$$\begin{aligned} \nabla &= (\partial_0, \partial_1, \partial_2, \partial_3) \\ \nabla_{\text{Cartesian}} &= \left(\frac{1}{c} \partial_t, \partial_x, \partial_y, \partial_z \right) \\ \nabla_{\text{Spherical}} &= \left(\frac{1}{c} \partial_t, \partial_r, \partial_\theta, \partial_\phi \right) \end{aligned} \quad (19)$$

And the square of the 4-gradient (the Laplace operator) is:

$$\nabla^2 = g^{\mu\nu} \partial_\mu \partial_\nu = \partial_\mu \partial^\mu \quad (20)$$

It requires thus the inverse of the metric tensor.

2.2.4 Equivalence principle

The equivalence principle states that one can get the same experimental results in any reference frame, i.e. one can also choose a free-falling one, in which locally no gravity is felt, and hence for free falling reference systems one has $g_{\mu\nu} = \eta_{\mu\nu}$.

2.2.5 Example: Magnitude of 4-velocity

$$u^\mu = \frac{dx^\mu}{d\tau} = \left(c \frac{dt}{d\tau}, \frac{d\vec{x}}{d\tau} \right) = \left(\frac{c\gamma}{dt} \frac{dt}{d\tau}, \vec{v}\gamma \right) = \begin{pmatrix} c\gamma \\ \vec{v}\gamma \end{pmatrix} \quad (21)$$

Going to a local inertial frame, one can use the Minkowski metric.

$$\begin{aligned} |u|^2 &= u_\nu u^\nu = \eta_{\mu\nu} u^\mu u^\nu = (c\gamma, \vec{v}\gamma) \cdot \begin{pmatrix} -1 & 0 \\ \vec{0} & \mathbb{1} \end{pmatrix} \cdot \begin{pmatrix} c\gamma \\ \vec{v}\gamma \end{pmatrix} = -c^2\gamma^2 + v^2\gamma^2 \\ &= (v^2 - c^2) \frac{1}{1 - \frac{v^2}{c^2}} = (v^2 - c^2) \frac{c^2}{c^2 - v^2} = -c^2 \end{aligned} \quad (22)$$

Since scalars are Lorentz-invariant, this result holds in any reference frame.

3 Relativistic dynamics

Curves are often parametrized by the "proper time" τ , i.e. the time passing for a particle moving along the space-time curve. τ in general differs from the coordinate time t . This defines the Lorentz factor:

$$\gamma = \frac{dt}{d\tau} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (23)$$

Here, $v = |\vec{v}|$. For a free-falling observer in his local inertial frame $x' = (ct', x'^0, x'^1, x'^3) = (c\tau, \text{const}, \text{const}, \text{const})$, i.e he is at constant coordinates, and time passes at the "proper time". For him, the line elements reads thus

$$ds^2 = -c^2 d\tau^2 \quad (24)$$

which must hold in any reference frame, as it is a scalar relation. This relation holds for any real-world particle, and one calls this a "time-like" trajectory. For light, one has $ds^2 = 0$, leading to "null geodesics".

3.1 Energy and momentum

For particles with rest mass m_0 the 4-momentum p is:

$$p = m_0 u = m_0 \begin{pmatrix} c\gamma \\ \vec{v}\gamma \end{pmatrix} = \begin{pmatrix} E/c \\ \vec{p} \end{pmatrix} \quad (25)$$

Note that E and \vec{p} are different than in Newtonian mechanics. In Minkowski space, its norm is

$$|p|^2 = p_\mu p^\mu = \eta_{\nu\mu} p^\nu p^\mu = -m_0^2 \gamma^2 c^2 + m_0^2 \gamma^2 v^2 = -m_0^2 \gamma^2 c^2 \left(1 - \frac{v^2}{c^2}\right) = -m_0^2 \gamma^2 c^2 \frac{1}{\gamma^2} = -m_0^2 c^2 \quad (26)$$

The 4-acceleration is

$$a^\mu = \frac{d p^\mu}{d\tau m_0} \quad (27)$$

4-velocity and 4-acceleration are orthogonal to each other (when m_0 is constant):

$$u \cdot a = \eta_{\nu\mu} \frac{p^\nu}{m_0} a^\mu = \eta_{\nu\mu} \frac{p^\nu}{m_0} \frac{d p^\mu}{d\tau m_0} = \frac{1}{2m_0^2} \frac{d}{d\tau} p^2 = \frac{1}{2m_0^2} \frac{d}{d\tau} (-m_0^2 c^2) = 0 \quad (28)$$

The energy can be written also in this form:

$$E = -\eta_{\nu\mu} p^\nu u^\mu = -m_0 \eta_{\nu\mu} u^\nu u^\mu = m_0 c^2 \quad (29)$$

But this equation is also valid, if one measures the energy of a particle moving with u in a system moving with v :

$$E = -\eta_{\nu\mu} p^\nu v^\mu = -m_0 \eta_{\nu\mu} u^\nu v^\mu \quad (30)$$

This definition will be carried over to general relativity:

$$E = -g_{\nu\mu} p^\nu u^\mu \quad (31)$$

The kinetic energy K is:

$$K = E - m_0 c^2 = (\gamma - 1) m_0 c^2 \quad (32)$$

Using the series expansion for γ

$$\gamma \approx 1 + \frac{1}{2} \frac{v^2}{c^2} + \frac{3}{8} \frac{v^4}{c^4} \quad (33)$$

one sees that the leading order of K is

$$K = (\gamma - 1)m_0c^2 \approx \frac{1}{2}m_0v^2 \quad (34)$$

From the definitions of E and \vec{p} follows:

$$\begin{aligned} \vec{v}E/c &= \vec{p}c \\ \frac{v^2}{c^2}E^2 &= (|\vec{p}|c)^2 \\ E^2(1 - \frac{v^2}{c^2}) &= E^2 - (|\vec{p}|c)^2 \\ E^2/\gamma^2 &= E^2 - (|\vec{p}|c)^2 \\ (m_0c^2)^2 &= E^2 - (|\vec{p}|c)^2 \end{aligned} \quad (35)$$

This is the relativistic energy-momentum relation. Newton's second law takes the form

$$\vec{F} = \frac{d\vec{p}}{dt} \quad (36)$$

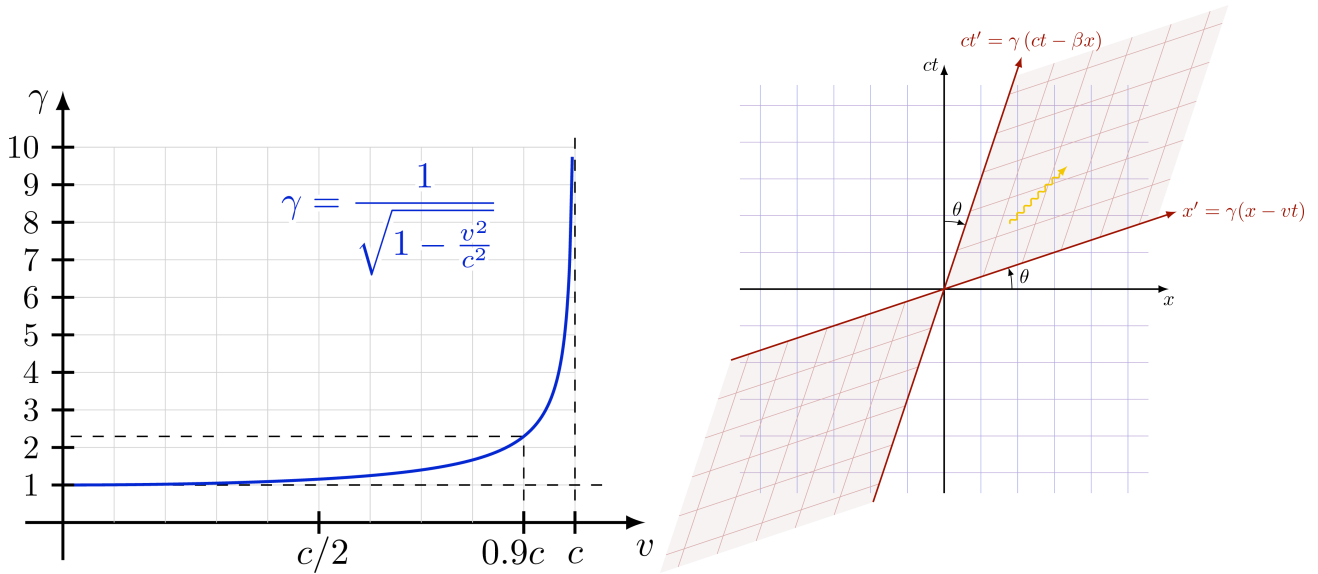


Figure 3: Left: The Lorentz-factor γ as a function of velocity. Right: A Lorentz-transformation of the Minkowski-space time. Source: TikZ

3.2 Lorentz transformations

For an observer moving relative to another one along the x -axis with velocity v_x , the coordinates are

$$\begin{aligned} t' &= \gamma \left(t - \frac{v_x}{c^2} x \right) \\ x' &= \gamma (x - v_x t) \\ y' &= y \\ z' &= z \end{aligned} \tag{37}$$

It is easy to calculate

$$v'_x = \frac{dx'}{dt'} = \frac{d}{dt'} \gamma (x - v_x t) = \frac{d}{dt'} \gamma \left(x - v_x \left(\frac{t'}{\gamma} + \frac{v_x}{c^2} x \right) \right) = -v_x \tag{38}$$

In matrix form, one can write:

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & \frac{v}{c}\gamma & 0 & 0 \\ \frac{v}{c}\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \tag{39}$$

A space-time interval is invariant under a Lorentz transformation:

$$\begin{aligned} s'^2 &= -(ct')^2 + x'^2 + y'^2 + z'^2 \\ &= -c^2 \gamma^2 \left(t - \frac{v_x}{c^2} x \right)^2 + \gamma^2 (x - v_x t)^2 + y^2 + z^2 \\ &= \gamma^2 (-c^2 t^2 + 2c^2 \frac{v_x}{c^2} x - \frac{v_x^2}{c^2} x^2) + \gamma^2 (x^2 - 2x v_x t + v_x^2 t^2) + y^2 + z^2 \\ &= \gamma^2 c^2 t^2 \left(-1 + \frac{v_x^2}{c^2} \right) + \gamma^2 x^2 \left(1 - \frac{v_x^2}{c^2} \right) + y^2 + z^2 \\ &= \gamma^2 c^2 t^2 \frac{-1}{\gamma^2} + \gamma^2 x^2 \frac{1}{\gamma^2} + y^2 + z^2 \\ &= -(ct)^2 + x^2 + y^2 + z^2 = s^2 \end{aligned} \tag{40}$$

The transformation law for velocities follows: The velocity v'_x measured in the primed coordinate system that moves with velocity u is given by the original v_x and u by:

$$v'_x = \frac{dx'}{dt'} = \frac{\gamma(dx - u dt)}{\gamma(dt - \frac{u}{c^2} dx)} = \frac{\frac{dx}{dt} - u}{1 - \frac{u}{c^2} \frac{dx}{dt}} = \frac{v_x - u}{1 - \frac{u v_x}{c^2}} \tag{41}$$

The same transformation can be applied to the energy-momentum vector $(E/c, \vec{p})$ or the wave vector $(\omega/c, \vec{k})$. Let's take E as an example:

$$\begin{aligned}
\frac{E'}{c} &= \frac{1}{c} \frac{m_0 c^2}{\sqrt{1 - v_x^2/c^2}} \\
&= \frac{1}{c} \frac{m_0 c^2}{\sqrt{1 - \left(\frac{v_x - u}{1 - \frac{u v_x}{c^2}}\right)^2 / c^2}} \\
&= \frac{1}{c} \frac{m_0 c^2 \left(1 - \frac{u v_x}{c^2}\right)}{\sqrt{\left(1 - \frac{u v_x}{c^2}\right)^2 - \frac{(v_x - u)^2}{c^2}}} \\
&= \frac{1}{c} \frac{m_0 c^2 \left(1 - \frac{u v_x}{c^2}\right)}{\sqrt{1 - 2 \frac{u v_x}{c^2} + \frac{u^2 v_x^2}{c^4} - \frac{v_x^2}{c^2} + 2 \frac{u v_x}{c^2} - \frac{u^2}{c^2}}} \\
&= \frac{1}{c} \frac{m_0 c^2 \left(1 - \frac{u v_x}{c^2}\right)}{\sqrt{\left(1 - \frac{u^2}{c^2}\right) \left(1 - \frac{v_x^2}{c^2}\right)}} \\
&= \frac{1}{c} \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} (\gamma m_0 c^2 - u \gamma m_0 v_x) \\
&= \frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} \left(\frac{E}{c} - \frac{u}{c} p_x \right) \tag{42}
\end{aligned}$$

So, this is the same as if we would have transformed $(E/c, \vec{p})$ by the same Lorentz boost by u as we did for x .

4 Curvature

4.1 Describing curvature

4.1.1 Covariant derivative and Christoffel symbols

"If I move a vector into the direction of another vector, how do its components change" - a non-trivial question in curved space-time. For a base vector e_μ moving infinitesimal into direction x^ν , the four components for the e_λ are given by the Christoffel symbols (which are not tensors):

$$\begin{aligned}\Gamma^\lambda_{\mu\nu} e_\lambda &= \frac{\partial e_\mu}{\partial x^\nu} \\ \Gamma^\lambda_{\mu\nu} e_\lambda e^\beta &= \frac{\partial e_\mu}{\partial x^\nu} e^\beta \\ \Gamma^\lambda_{\mu\nu} \delta^\beta_\lambda &= \frac{\partial e_\mu}{\partial x^\nu} e^\beta \\ \Gamma^\beta_{\mu\nu} &= \frac{\partial e_\mu}{\partial x^\nu} e^\beta\end{aligned}\tag{43}$$

This allows defining a covariant derivative: "The vector field A not only changes as a function of coordinates, but due to the curvature, there is also a change due to the coordinates changing."

$$\begin{aligned}\nabla_\mu A^\lambda &= \partial_\mu A^\lambda + \Gamma^\lambda_{\mu\nu} A^\nu \\ \nabla_\mu A_\lambda &= \partial_\mu A_\lambda - \Gamma^\nu_{\mu\lambda} A_\nu\end{aligned}\tag{44}$$

We will not use the notation $A^\lambda_{;\mu} = \nabla_\mu A^\lambda$. For a tensor, one has:

$$\nabla_\mu A^{\alpha\beta} = \partial_\mu A^{\alpha\beta} + \Gamma^\alpha_{\lambda\mu} A^{\lambda\beta} + \Gamma^\beta_{\mu\lambda} A^{\alpha\lambda}\tag{45}$$

And for a scalar f , the covariant derivative is the partial one:

$$\nabla_\lambda f = \partial_\lambda f\tag{46}$$

The minus sign in equation 44 appears because the following line should be true (i.e. the covariant derivative should behave like a derivative):

$$(\partial_\lambda A^\mu) B_\mu + A^\mu (\partial_\lambda B_\mu) = \partial_\lambda (A^\mu B_\mu) = \nabla_\lambda (A^\mu B_\mu) = (\nabla_\lambda A^\mu) B_\mu + A^\mu (\nabla_\lambda B_\mu)\tag{47}$$

and the two terms with Christoffel symbols in the covariant derivatives need therefore to cancel. Further, one sees that in that was also the chain rule for the product of two vectors holds. One can show the chain rule also for higher ranked tensors, for example:

$$\begin{aligned}\nabla_\mu (A^{\alpha\beta} B_\gamma) &= \partial_\mu (A^{\alpha\beta} B_\gamma) + \Gamma^\alpha_{\lambda\mu} A^{\lambda\beta} B_\gamma + \Gamma^\beta_{\mu\lambda} A^{\alpha\lambda} B_\gamma - \Gamma^\lambda_{\gamma\mu} A^{\alpha\beta} B_\lambda \\ &= B_\gamma \partial_\mu A^{\alpha\beta} + A^{\alpha\beta} \partial_\mu B_\gamma + B_\gamma \Gamma^\alpha_{\lambda\mu} A^{\lambda\beta} + B_\gamma \Gamma^\beta_{\mu\lambda} A^{\alpha\lambda} - A^{\alpha\beta} B_\lambda \Gamma^\lambda_{\gamma\mu} \\ &= A^{\alpha\beta} (\partial_\mu B_\gamma - B_\lambda \Gamma^\lambda_{\gamma\mu}) + B_\gamma (\partial_\mu A^{\alpha\beta} + \Gamma^\alpha_{\lambda\mu} A^{\lambda\beta} + \Gamma^\beta_{\mu\lambda} A^{\alpha\lambda}) \\ &= A^{\alpha\beta} \nabla_\mu B_\gamma + B_\gamma \nabla_\mu A^{\alpha\beta}\end{aligned}\tag{48}$$

With the chosen definition for the covariant derivative, a property called metric compatibility holds: The covariant derivative of the metric tensor is always zero:

$$\begin{aligned}\nabla_\lambda g^{\alpha\mu} &= 0 \\ \nabla_\lambda g_{\alpha\mu} &= 0\end{aligned}\tag{49}$$

This allows moving the metric tensor in and out of derivatives:

$$\begin{aligned}\nabla_\lambda A^\nu &= \nabla_\lambda(g^{\nu\mu}A_\mu) = A_\mu\nabla_\lambda g^{\nu\mu} + g^{\nu\mu}\nabla_\lambda A_\mu = 0 + g^{\nu\mu}\nabla_\lambda A_\mu = g^{\nu\mu}\nabla_\lambda A_\mu \\ \nabla_\lambda A_\nu &= \nabla_\lambda(g_{\nu\mu}A^\mu) = A^\mu\nabla_\lambda g_{\nu\mu} + g_{\nu\mu}\nabla_\lambda A^\mu = 0 + g_{\nu\mu}\nabla_\lambda A^\mu = g_{\nu\mu}\nabla_\lambda A^\mu\end{aligned}\tag{50}$$

For a torsion-free space-time, we expect that derivatives commute: $\nabla_\mu\nabla_\nu = \nabla_\nu\nabla_\mu$. For a scalar f in flat coordinates the covariant derivative is the partial derivative, and clearly

$$\partial_\mu\partial_\nu f = \partial_\nu\partial_\mu f\tag{51}$$

If the symmetry holds in one coordinate system, it holds in all, hence

$$\begin{aligned}\nabla_\mu\partial_\nu f &= \nabla_\nu\partial_\mu f \\ \partial_\mu\partial_\nu f - \Gamma_{\mu\nu}^\alpha\partial_\alpha f &= \partial_\nu\partial_\mu f - \Gamma_{\nu\mu}^\alpha\partial_\alpha f\end{aligned}\tag{52}$$

and we see that the metric tensor needs to be symmetric in the second and third index.

$$\Gamma_{\mu\nu}^\alpha = \Gamma_{\nu\mu}^\alpha\tag{53}$$

There are thus not $4^3 = 64$ independent Christoffel symbols, but only 40. The Christoffel symbols can be expressed with the metric tensor. This requires some algebra. From the definition we have (equation 43):

$$\begin{aligned}e_\lambda\Gamma_{\alpha\nu}^\lambda &= e_\lambda\Gamma_{\nu\alpha}^\lambda = \frac{\partial e_\nu}{\partial x^\alpha} = \partial_\alpha e_\nu \\ e_\mu e_\lambda\Gamma_{\alpha\nu}^\lambda &= e_\mu\partial_\alpha e_\nu\end{aligned}\tag{54}$$

Consider

$$\partial_\alpha(e_\mu e_\nu) = e_\nu\partial_\alpha e_\mu + e_\mu\partial_\alpha e_\nu \longrightarrow e_\mu\partial_\alpha e_\nu = \partial_\alpha(e_\mu e_\nu) - e_\nu\partial_\alpha e_\mu\tag{55}$$

With that we get:

$$e_\mu e_\lambda\Gamma_{\alpha\nu}^\lambda = \partial_\alpha(e_\mu e_\nu) - e_\nu e_\lambda\Gamma_{\alpha\mu}^\lambda\tag{56}$$

Writing this equation two more times, but with indices relabelled:

$$e_\alpha e_\lambda\Gamma_{\nu\mu}^\lambda = \partial_\nu(e_\alpha e_\mu) - e_\mu e_\lambda\Gamma_{\nu\alpha}^\lambda\tag{57}$$

$$e_\alpha e_\lambda\Gamma_{\mu\nu}^\lambda = \partial_\mu(e_\alpha e_\nu) - e_\nu e_\lambda\Gamma_{\mu\alpha}^\lambda\tag{58}$$

Taking 56 + 57 - 58 yields, using $\Gamma_{\nu\mu}^\lambda = \Gamma_{\mu\nu}^\lambda$ and $\Gamma_{\alpha\mu}^\lambda = \Gamma_{\mu\alpha}^\lambda$:

$$\begin{aligned}
e_\mu e_\lambda \Gamma_{\alpha\nu}^\lambda &= \partial_\alpha(e_\mu e_\nu) + \partial_\nu(e_\alpha e_\mu) - \partial_\mu(e_\alpha e_\nu) - e_\mu e_\lambda \Gamma_{\nu\alpha}^\lambda \\
2e_\mu e_\lambda \Gamma_{\alpha\nu}^\lambda &= \partial_\alpha(e_\mu e_\nu) + \partial_\nu(e_\alpha e_\mu) - \partial_\mu(e_\alpha e_\nu) \\
(e^\mu e^\rho)(e_\mu e_\lambda) \Gamma_{\alpha\nu}^\lambda &= \frac{1}{2}(e^\mu e^\rho)(\partial_\alpha(e_\mu e_\nu) + \partial_\nu(e_\alpha e_\mu) - \partial_\mu(e_\alpha e_\nu)) \\
\delta_\lambda^\rho \Gamma_{\alpha\nu}^\lambda = \Gamma_{\alpha\nu}^\rho &= \frac{1}{2}g^{\mu\rho}(\partial_\alpha g_{\mu\nu} + \partial_\nu g_{\alpha\mu} - \partial_\mu g_{\alpha\nu}) \\
\Gamma_{\mu\nu}^\lambda &= \frac{1}{2}g^{\lambda\alpha}(\partial_\mu g_{\alpha\nu} + \partial_\nu g_{\alpha\mu} - \partial_\alpha g_{\mu\nu}) \\
\Gamma_{\lambda\mu\nu} &= \frac{1}{2}(\partial_\mu g_{\lambda\nu} + \partial_\nu g_{\lambda\mu} - \partial_\lambda g_{\mu\nu})
\end{aligned} \tag{59}$$

With this, we can prove the metric compatibility in few steps:

$$\begin{aligned}
\nabla_\lambda g_{\mu\nu} &= \partial_\lambda g_{\mu\nu} - \Gamma_{\lambda\mu}^\alpha g_{\alpha\nu} - \Gamma_{\lambda\nu}^\alpha g_{\mu\alpha} = \partial_\lambda g_{\mu\nu} - \Gamma_{\nu\lambda\mu} - \Gamma_{\mu\lambda\nu} \\
&= \partial_\lambda g_{\mu\nu} - \frac{1}{2}(\partial_\lambda g_{\mu\nu} + \partial_\mu g_{\lambda\nu} - \partial_\nu g_{\lambda\mu}) - \frac{1}{2}(\partial_\lambda g_{\mu\nu} + \partial_\nu g_{\lambda\mu} - \partial_\mu g_{\lambda\nu}) = 0
\end{aligned} \tag{60}$$

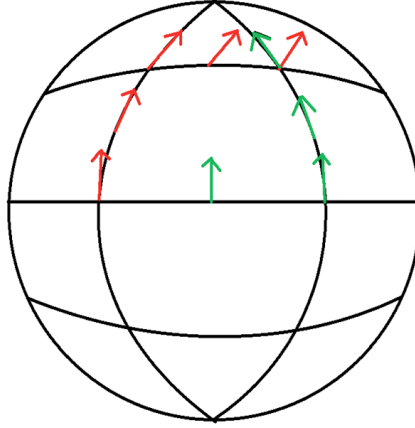


Figure 4: Left: Parallel shifting a vector from bottom left to top right yields different results, depending on the along which coordinate axis one moves first and which second. This is feature of the curvature of the underlying space. Source: medium.com

4.1.2 Riemann tensor, Ricci tensor and Ricci scalar

The idea behind the description of curvature is that parallel shifting a vector in coordinate direction μ and then in ν does not yield the same as first in direction ν and then in μ , see figure 4. Expressed in infinitesimal steps, the

difference for a vector A^ρ is:

$$\begin{aligned}
(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu)A^\rho &= \nabla_\mu(\partial_\nu A^\rho + \Gamma^\rho_{\nu\sigma} A^\sigma) - \nabla_\nu(\partial_\mu A^\rho + \Gamma^\rho_{\mu\sigma} A^\sigma) \\
&= \partial_\mu(\partial_\nu A^\rho + \Gamma^\rho_{\nu\sigma} A^\sigma) - \Gamma^\lambda_{\mu\nu}(\partial_\lambda A^\rho + \Gamma^\rho_{\lambda\sigma} A^\sigma) + \Gamma^\rho_{\mu\lambda}(\partial_\nu A^\lambda + \Gamma^\lambda_{\nu\sigma} A^\sigma) \\
&\quad - \partial_\nu(\partial_\mu A^\rho + \Gamma^\rho_{\mu\sigma} A^\sigma) + \Gamma^\lambda_{\nu\mu}(\partial_\lambda A^\rho + \Gamma^\rho_{\lambda\sigma} A^\sigma) - \Gamma^\rho_{\nu\lambda}(\partial_\mu A^\lambda + \Gamma^\lambda_{\mu\sigma} A^\sigma) \\
&= (\partial_\mu \Gamma^\rho_{\nu\sigma})A^\sigma + \Gamma^\rho_{\nu\sigma} \partial_\mu A^\sigma - \Gamma^\lambda_{\mu\nu} \partial_\lambda A^\rho - \Gamma^\lambda_{\mu\nu} \Gamma^\rho_{\lambda\sigma} A^\sigma + \Gamma^\rho_{\mu\lambda} \partial_\nu A^\lambda + \Gamma^\rho_{\mu\lambda} \Gamma^\lambda_{\nu\sigma} A^\sigma \\
&\quad - (\partial_\nu \Gamma^\rho_{\mu\sigma})A^\sigma - \Gamma^\rho_{\mu\sigma} \partial_\nu A^\sigma + \Gamma^\lambda_{\nu\mu} \partial_\lambda A^\rho + \Gamma^\lambda_{\nu\mu} \Gamma^\rho_{\lambda\sigma} A^\sigma - \Gamma^\rho_{\nu\lambda} \partial_\mu A^\lambda - \Gamma^\rho_{\nu\lambda} \Gamma^\lambda_{\mu\sigma} A^\sigma \quad (61)
\end{aligned}$$

The third terms in both rows cancel, as well as the fourth terms. Term 5 in row 1 after relabelling the summation index λ into σ cancels term 2 in row 2. Term 2 in row 1 after relabelling the summation index σ into λ cancels term 5 in row 2. We are left with

$$\dots = (\partial_\mu \Gamma^\rho_{\mu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + \Gamma^\rho_{\mu\lambda} \Gamma^\lambda_{\nu\sigma} - \Gamma^\rho_{\nu\lambda} \Gamma^\lambda_{\mu\sigma})A^\sigma \quad (62)$$

The term in brackets vanishes, if the parallel shifting of the vector happens in flat space. In curved space-time it does not, and it measures the curvature. It is called the Riemann curvature tensor:

$$R^\lambda_{\alpha\nu\mu} = \partial_\nu \Gamma^\lambda_{\alpha\mu} - \partial_\mu \Gamma^\lambda_{\alpha\nu} + \Gamma^\lambda_{\sigma\nu} \Gamma^\sigma_{\alpha\mu} - \Gamma^\lambda_{\sigma\mu} \Gamma^\sigma_{\alpha\nu} \quad (63)$$

Here it is defined with the "Riemann" sign convention. It has the following symmetries:

$$\begin{aligned}
R_{\lambda\alpha\nu\mu} &= -R_{\alpha\lambda\mu\nu} \\
R^\lambda_{\alpha\nu\mu} &= -R^\lambda_{\alpha\mu\nu} \\
R_{\alpha\beta\mu\nu} &= R_{\mu\nu\alpha\beta} \quad (64)
\end{aligned}$$

One obtains the so-called Ricci tensor as contraction from the Riemann tensor, or expressed in terms of Christoffel symbols:

$$\begin{aligned}
R_{\alpha\mu} &= R^\lambda_{\alpha\lambda\mu} \\
&= \partial_\lambda \Gamma^\lambda_{\alpha\mu} - \partial_\mu \Gamma^\lambda_{\alpha\lambda} + \Gamma^\lambda_{\sigma\lambda} \Gamma^\sigma_{\alpha\mu} - \Gamma^\lambda_{\sigma\mu} \Gamma^\sigma_{\alpha\lambda} \quad (65)
\end{aligned}$$

The Ricci tensor is symmetric

$$R_{\mu\nu} = R^\alpha_{\mu\alpha\nu} = g^{\alpha\beta} R_{\beta\mu\alpha\nu} = g^{\beta\alpha} R_{\alpha\nu\beta\mu} = R^\beta_{\nu\beta\mu} = R_{\nu\mu} \quad (66)$$

The meaning of the Ricci tensor is that it describes the volume changes as a function of coordinates of an infinitesimal space-time element. The Ricci scalar is a contraction of the Ricci tensor:

$$R = R^\alpha_\alpha \quad (67)$$

4.2 Deriving the Bianchi identities

Working for the moment in a special coordinate system - the result will be a scalar, and hence independent of the choice of coordinate system. Going to a local inertial frame. Then, locally the metric is Minkowski, and the Christoffel symbols vanish. But: not the derivatives of the Christoffel symbols! "One can transform away locally the gravitational force, but not the tidal forces". Also, in a local inertial frame, the the covariant derivative becomes a partial derivative. The Riemann curvature in a local inertial frame is:

$$R^\lambda_{\alpha\nu\mu} = \partial_\nu \Gamma^\lambda_{\alpha\mu} - \partial_\mu \Gamma^\lambda_{\alpha\nu} \quad (68)$$

Taking the covariant derivative:

$$\nabla_\sigma R^\lambda_{\alpha\nu\mu} = \partial_\sigma \partial_\nu \Gamma^\lambda_{\alpha\mu} - \partial_\sigma \partial_\mu \Gamma^\lambda_{\alpha\nu} \quad (69)$$

Writing this three times, cycling indices:

$$\begin{aligned} \nabla_\sigma R^\lambda_{\alpha\nu\mu} &= \partial_\sigma \partial_\nu \Gamma^\lambda_{\alpha\mu} - \partial_\sigma \partial_\mu \Gamma^\lambda_{\alpha\nu} \\ \nabla_\nu R^\lambda_{\alpha\mu\sigma} &= \partial_\nu \partial_\mu \Gamma^\lambda_{\alpha\sigma} - \partial_\nu \partial_\sigma \Gamma^\lambda_{\alpha\mu} \\ \nabla_\mu R^\lambda_{\alpha\sigma\nu} &= \partial_\mu \partial_\sigma \Gamma^\lambda_{\alpha\nu} - \partial_\mu \partial_\nu \Gamma^\lambda_{\alpha\sigma} \end{aligned} \quad (70)$$

Adding the three equations, and noting that partial derivatives commute, the right hand sides yields simply 0, which is known as the Bianchi identity

$$\nabla_\sigma R^\lambda_{\alpha\nu\mu} + \nabla_\nu R^\lambda_{\alpha\mu\sigma} + \nabla_\mu R^\lambda_{\alpha\sigma\nu} = 0 \quad (71)$$

This result is valid in any reference frame, as it is a tensor relation.

5 Geodesic equation

5.1 Standard approach

A geodesic is the equivalent of a straight line in Euclidean or Minkowski space. A body with no acceleration will move along such a line, and it is defined by velocity being constant:

$$0 = \frac{d\vec{u}}{d\tau} = \frac{d(u^\mu \vec{e}_\mu)}{d\tau} = \vec{e}_\mu \frac{du^\mu}{d\tau} + u^\mu \frac{d\vec{e}_\mu}{d\tau} \quad (72)$$

From this the geodesic equation can be derived:

$$\begin{aligned} \vec{e}_\mu \frac{du^\mu}{d\tau} &= -u^\mu \frac{d\vec{e}_\mu}{d\tau} \\ \vec{e}_\nu \vec{e}_\mu \frac{du^\mu}{d\tau} &= -\vec{e}_\nu u^\mu \frac{d\vec{e}_\mu}{d\tau} \\ g_{\mu\nu} \frac{du^\mu}{d\tau} &= -\vec{e}_\nu u^\mu \frac{d\vec{e}_\mu}{d\tau} \\ \frac{du_\nu}{d\tau} &= -\vec{e}_\nu u^\mu \frac{d\vec{e}_\mu}{d\tau} \\ g^{\nu\lambda} \frac{du_\nu}{d\tau} &= -g^{\nu\lambda} \vec{e}_\nu u^\mu \frac{d\vec{e}_\mu}{d\tau} \\ \frac{du^\lambda}{d\tau} &= -\vec{e}^{\lambda\mu} u^\mu \frac{d\vec{e}_\mu}{d\tau} \\ \frac{du^\lambda}{d\tau} &= -\vec{e}^{\lambda\mu} u^\mu \frac{d\vec{e}_\mu}{dx^\nu} \frac{dx^\nu}{d\tau} \\ \frac{du^\lambda}{d\tau} &= -\frac{\partial \vec{e}_\mu}{\partial x^\nu} \vec{e}^{\lambda\mu} u^\mu u^\nu \\ \frac{du^\lambda}{d\tau} &= -\Gamma_{\mu\nu}^\lambda u^\mu u^\nu \end{aligned} \quad (73)$$

Hence, we have the two equivalent forms:

$$\begin{aligned} \frac{du^\lambda}{d\tau} + \Gamma_{\mu\nu}^\lambda u^\mu u^\nu &= 0 \\ \frac{d^2 x^\lambda}{d\tau^2} + \Gamma_{\mu\nu}^\lambda \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} &= 0 \end{aligned} \quad (74)$$

It is written here for proper time τ , but actually holds for any parameter.

5.2 Lagrangian formulation

There is an elegant way to express the geodesic equation using the Lagrangian \mathcal{L} .

$$\mathcal{L}(x, u) = -mc \sqrt{-g_{\mu\nu}(x) u^\mu u^\nu} \quad (75)$$

Note that formally $\mathcal{L} = -m c^2$. Geodesics are extremal in the sense, that they are the shortest paths between two points. The \sqrt{X} is extremal when X is extremal, and hence one can also vary

$$\mathcal{L}'(x, u) = g_{\mu\nu}(x) u^\mu u^\nu \quad (76)$$

As usual, the coordinates x and velocities u are treated as independent variables. The Euler-Lagrange equation is then

$$\frac{\partial \mathcal{L}'}{\partial x^\lambda} = \frac{d}{d\tau} \frac{\partial \mathcal{L}'}{\partial u^\lambda} \quad (77)$$

This is identical to the geodesic equation. The left side is

$$\frac{\partial \mathcal{L}'}{\partial x^\lambda} = \frac{1}{2} (\partial_\lambda g_{\mu\nu}) u^\mu u^\nu \quad (78)$$

The right side is

$$\begin{aligned} \frac{d}{d\tau} \frac{\partial \mathcal{L}'}{\partial u^\lambda} &= \frac{d}{d\tau} \left(\frac{1}{2} g_{\mu\nu} \delta_\lambda^\mu u^\nu + \frac{1}{2} g_{\mu\nu} u^\mu \delta_\lambda^\nu \right) \\ &= \frac{d}{d\tau} (g_{\mu\lambda} u^\mu) \\ &= \partial_\nu g_{\mu\lambda} u^\nu u^\mu + g_{\mu\lambda} \ddot{x}^\mu \\ &= \frac{1}{2} \partial_\nu g_{\mu\lambda} u^\nu u^\mu + \frac{1}{2} \partial_\mu g_{\nu\lambda} u^\mu u^\nu + g_{\mu\lambda} \ddot{x}^\mu \end{aligned} \quad (79)$$

In the last step, the previous term was split and summation indices have been renamed. The dot-derivative indicates deriving with respect to proper time. Putting things together

$$\begin{aligned} \frac{1}{2} \partial_\lambda g_{\mu\nu} u^\mu u^\nu &= \frac{1}{2} \partial_\nu g_{\mu\lambda} u^\nu u^\mu + \frac{1}{2} \partial_\mu g_{\nu\lambda} u^\mu u^\nu + g_{\mu\lambda} \ddot{x}^\mu \\ 0 &= g_{\mu\lambda} \ddot{x}^\mu + \frac{1}{2} \partial_\nu g_{\mu\lambda} u^\nu u^\mu + \frac{1}{2} \partial_\mu g_{\nu\lambda} u^\mu u^\nu - \frac{1}{2} \partial_\lambda g_{\mu\nu} u^\mu u^\nu \\ 0 &= \ddot{x}_\lambda + \Gamma_{\lambda\mu\nu} u^\mu u^\nu \\ 0 &= g^{\lambda\rho} \ddot{x}_\lambda + g^{\lambda\rho} \Gamma_{\lambda\mu\nu} u^\mu u^\nu \\ 0 &= \ddot{x}^\rho + \Gamma^\rho_{\mu\nu} \dot{x}^\mu \dot{x}^\nu \end{aligned} \quad (80)$$

An advantage of this formulation is that often it avoids using Christoffel symbols.

5.3 The choice of Lagrangian

Since we optimized \mathcal{L}' , why don't we call that Lagrangian? The reason is, that the action

$$S_\tau = \int d\tau \mathcal{L} \quad (81)$$

shall be invariant under transformations from one affine parameter τ to another one σ :

$$\begin{aligned} S_\sigma &= -mc \int d\sigma \sqrt{-g_{\mu\nu}(x) \frac{dx^\mu}{d\sigma} \frac{dx^\nu}{d\sigma}} \\ &= -mc \int \frac{d\sigma}{d\tau} d\tau \sqrt{-g_{\mu\nu}(x) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \frac{d\tau^2}{d\sigma^2}} \\ &= -mc \int d\tau \sqrt{-g_{\mu\nu}(x) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}} = S_\tau \end{aligned} \quad (82)$$

This does not hold for \mathcal{L}' , but the choice of \mathcal{L} in equation 75 is invariant under re-parametrization.

6 Energy-momentum tensor

Recap of continuity equation for charge (expressing conservation of the scalar charge):

$$\frac{d\rho}{dt} + \vec{\nabla} \cdot \vec{j} = 0 \quad (83)$$

Using the 4-current $j^\mu = (c\rho, \vec{j})$ the continuity equation gets

$$\partial_\mu j^\mu = 0 \quad (84)$$

Conserving a scalar quantity is thus an equation with a gradient of a 4-vector. The four-momentum is

$$p^\mu = \left(\frac{E}{c}, \gamma m_0 \vec{v} \right) \quad (85)$$

Conserving energy and momentum, i.e. the four-momentum, being a vector, is thus an equation with a gradient of some 4x4 tensor:

$$\nabla_\mu T^{\mu\nu} = 0 \quad (86)$$

Symbolically, the components (in a local inertial frame) are:

$$T = \partial_\nu \begin{pmatrix} E & p & p & p \\ p & P & S & S \\ p & S & P & P \\ p & S & S & P \end{pmatrix} \quad (87)$$

Note the T_{00} component: It is the energy density, which is the relativistic version of mass density (as in Newtonian mechanics), and all energy components need to be taken into account. So, for the Newtonian limit, $T_{00} = \rho c^2$, which also holds for pressure-less dust. p refers to momentum, P to pressure, and S to shear. The form of T for various cases can easily be looked up, for example for a perfect fluid.

$$T^{\mu\nu} = \left(\rho + \frac{P}{c^2} \right) u^\mu u^\nu + P g^{\mu\nu} \quad (88)$$

In a local inertial frame, this is a diagonal matrix with diagonal $(\rho c^2, P, P, P)$.

7 Derivation of the field equations

7.1 Finding the right tensors

- Independence of coordinate system choice: Tensor equations
- Spacetime curvature should relate to matter
- In a weak-field, slow motion limit, one should recover Newton's equations
- locally, energy and momentum should be conserved

We thus look for a tensor equation. The simplest form in which curvature can occur is a two-index curvature tensor. Matter, in the simplest form, can then be represented by the energy-momentum tensor $T^{\mu\nu}$. So we postulate some equation of type $G^{\mu\nu} \propto T^{\mu\nu}$, with an unknown tensor G . As energy and momentum shall be conserved, i.e. $\nabla_\mu T^{\mu\nu} = 0$, one also has to demand that $\nabla_\mu G^{\mu\nu} = 0$. And we can construct one, starting from the Bianchi identity:

$$\nabla_\lambda R_{\alpha\beta\mu\nu} + \nabla_\nu R_{\alpha\beta\lambda\mu} + \nabla_\mu R_{\alpha\beta\nu\lambda} = 0 \quad (89)$$

Multiplying with $g^{\alpha\mu}$ and swapping indices in the second term:

$$g^{\alpha\mu} \nabla_\lambda R_{\alpha\beta\mu\nu} - g^{\alpha\mu} \nabla_\nu R_{\alpha\beta\mu\lambda} + g^{\alpha\mu} \nabla_\mu R_{\alpha\beta\nu\lambda} = 0 \quad (90)$$

Due to the metric compatibility, one can move the $g^{\alpha\mu}$ into the derivatives, where it will raise the α index to μ (first two terms) and:

$$\nabla_\lambda R^\mu_{\beta\mu\nu} - \nabla_\nu R^\mu_{\beta\mu\lambda} + \nabla^\alpha R_{\alpha\beta\nu\lambda} = 0 \quad (91)$$

In the first and second term, the Ricci tensor appeared, and swapping indices in the last term:

$$\nabla_\lambda R_{\beta\nu} - \nabla_\nu R_{\beta\lambda} - \nabla^\alpha R_{\beta\alpha\nu\lambda} = 0 \quad (92)$$

Multiplying with $g^{\beta\nu}$ and again using the metric compatibility, raising indices, and substituting Ricci tensor and scalar when they occur, one gets

$$\begin{aligned} \nabla_\lambda g^{\beta\nu} R_{\beta\nu} - \nabla_\nu g^{\beta\nu} R_{\beta\lambda} - \nabla^\alpha g^{\beta\nu} R_{\beta\alpha\nu\lambda} &= 0 \\ \nabla_\lambda R^\nu_{\nu} - \nabla_\nu R^\nu_{\lambda} - \nabla^\alpha R^\nu_{\alpha\nu\lambda} &= 0 \\ \nabla_\lambda R - \nabla_\nu R^\nu_{\lambda} - \nabla^\alpha R_{\alpha\lambda} &= 0 \end{aligned} \quad (93)$$

Using $\nabla^\alpha = g^{\alpha\rho} \nabla_\rho$ and using again the metric compatibility:

$$\begin{aligned} \nabla_\lambda R - \nabla_\nu R^\nu_{\lambda} - \nabla_\rho g^{\alpha\rho} R_{\alpha\lambda} &= 0 \\ \nabla_\lambda R - \nabla_\nu R^\nu_{\lambda} - \nabla_\rho R^\rho_{\lambda} &= 0 \end{aligned} \quad (94)$$

Here, ν and ρ are just summation indices, so the second and third term are the same.

$$\nabla_\lambda R - 2\nabla_\nu R^\nu_{\lambda} = 0 \quad (95)$$

Multiplying with $g^{\mu\lambda}$, and again using the metric compatibility and raising indices:

$$\begin{aligned} \nabla_\lambda g^{\mu\lambda} R - 2\nabla_\nu g^{\mu\lambda} R^\nu_{\lambda} &= 0 \\ \nabla_\lambda g^{\mu\lambda} R - 2\nabla_\nu R^{\mu\nu} &= 0 \\ \nabla_\nu g^{\mu\nu} R - 2\nabla_\nu R^{\mu\nu} &= 0 \\ 2\nabla_\nu (R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R) &= 0 \end{aligned} \quad (96)$$

Define $G^{\mu\nu} = R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R$, the last equation states $\nabla_\mu G^{\mu\nu} = 0$. Hence, G is a divergence free, two-index curvature tensor - so it is a viable candidate for the field equation. It is the simple-most such choice. Therefore, we can guess here, that the field equations are

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = \kappa T^{\mu\nu} \quad (97)$$

7.2 Weak fields and slow motions

Note that there is no way to prove the field equations, as they are the theory Einstein **postulated**. But one can connect them to classical mechanics for the limit of a weak field (i.e. slow motions), which can serve at least as a motivation, as it shows that in this limit, we recover what we know.

7.2.1 Connection to Newtonian potential

First, it is useful to see how the Einstein tensor $G^{\mu\nu}$ connects to the Newton potential in the case of a weak field. For a weak field, the metric should be almost Minkowski, such that one can write

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \quad (98)$$

with a small $h \ll 1$. For this weak-field metric, the Christoffel symbols are

$$\begin{aligned} \Gamma_{\mu\nu}^\lambda &= \frac{1}{2}(\eta^{\lambda\alpha} + h^{\lambda\alpha}) [\partial_\mu(\eta_{\alpha\nu} + h_{\alpha\nu}) + \partial_\nu(\eta_{\alpha\mu} + h_{\alpha\mu}) - \partial_\alpha(\eta_{\mu\nu} + h_{\mu\nu})] \\ &= \frac{1}{2}\eta^{\lambda\alpha}(\partial_\mu h_{\alpha\nu} + \partial_\nu h_{\alpha\mu} - \partial_\alpha h_{\mu\nu}) \\ &= \frac{1}{2}(\partial_\mu h_{\nu}^\lambda + \partial_\nu h_{\mu}^\lambda - \partial^\lambda h_{\mu\nu}) \end{aligned} \quad (99)$$

where higher order terms of h are dropped, and the derivatives of the constant Minkowski metric vanish. The Newtonian equations only have spatial derivatives, and thus only $i = 1, 2, 3$ of the 00-component matter:

$$\Gamma_{00}^i = \frac{1}{2}(\partial_0 h_{00}^i + \partial_0 h_{00}^i - \partial^i h_{00}) \quad (100)$$

The time derivatives ∂_0 are 0, as our metric is constant. It remains the spatial one, and we get

$$\Gamma_{00}^i \approx -\frac{1}{2}\partial^i h_{00} \quad (101)$$

Note that since h is small, so is Γ_{00}^i . Next, one evaluates the (spatial part of the) geodesic equation. We are dealing with slow motions, $v \ll c$. In that limit, $\tau \rightarrow t$, and hence

$$\frac{d^2 x^i}{dt^2} + \Gamma_{\mu\nu}^i \frac{dx^\mu}{dt} \frac{dx^\nu}{dt} = 0 \quad (102)$$

The velocity terms $\frac{dx^i}{dt}$ are much smaller than $\frac{dx^0}{dt} = c$ and thus

$$\begin{aligned} \frac{d^2 x^i}{dt^2} + \Gamma_{00}^i \frac{dx^0}{dt} \frac{dx^0}{dt} &= 0 \\ \frac{d^2 x^i}{dt^2} + c^2 \Gamma_{00}^i &= 0 \end{aligned} \quad (103)$$

As expected, only the Γ^i_{00} Christoffel symbol appears. Using the calculation from above

$$\frac{d^2 x^i}{dt^2} = \frac{1}{2} c^2 \partial^i h_{00} = \frac{1}{2} c^2 \vec{\nabla}^i h_{00} \quad (104)$$

In Newton's theory we have

$$\frac{d^2 x^i}{dt^2} = -\vec{\nabla}^i \Phi \quad (105)$$

Therefore, we can identify

$$\begin{aligned} h_{00} &= -\frac{2\Phi}{c^2} \\ g_{00} &= -1 - \frac{2\Phi}{c^2} \end{aligned} \quad (106)$$

7.2.2 Proportionality constant

In Newtonian gravity, the only source of gravity is mass or energy density, which corresponds to the T_{00} -component of the energy-momentum tensor, and all other components being 0:

$$T_{00} = \rho c^2 \quad (107)$$

Hence, we only need to look at the 00-component for determining the proportionality constant κ .

$$G_{00} = R_{00} - \frac{1}{2} g_{00} R = \kappa \rho c^2 \quad (108)$$

Using the Ricci tensor in the form expressed as Christoffel symbols, one sees that in the weak field limit, the third and fourth term can be neglected, as they are squares of Γ , with $\Gamma \sim h \ll 1$. Hence one has

$$R_{\mu\nu} = \partial_\alpha \Gamma^\alpha_{\mu\nu} - \partial_\nu \Gamma^\alpha_{\mu\alpha} \quad (109)$$

The 00-component is

$$R_{00} = \partial_\alpha \Gamma^\alpha_{00} - \partial_0 \Gamma^\alpha_{0\alpha} \quad (110)$$

As our metric tensor is not time variable, ∂_0 yields 0. Thus

$$R_{00} = \partial_i \Gamma^i_{00} - \partial_0 \Gamma^0_{00} \quad (111)$$

where the last term again is 0. Then

$$R_{00} = \partial_i \left(-\frac{1}{2} \partial^i h_{00} \right) = -\frac{1}{2} \partial_i \partial^i \left(-\frac{2\Phi}{c^2} \right) = \frac{1}{c^2} \vec{\nabla}^2 \Phi \quad (112)$$

Given that only the T_{00} component is present in the Newtonian limit, the spatial components G_{ij} are 0 - which allows for a little trick to evaluate the Ricci scalar:

$$R_{ij} - \frac{1}{2} g_{ij} R = 0 \longrightarrow R_{ij} = \frac{1}{2} g_{ij} R \quad (113)$$

Splitting up the contraction sum of the Ricci scalar:

$$R = R^\nu_\nu = g^{\mu\nu} R_{\mu\nu} = g^{00} R_{00} + g^{ij} R_{ij} = g^{00} R_{00} + \frac{1}{2} g^{ij} g_{ij} R = g^{00} R_{00} + \frac{3}{2} R \quad (114)$$

where the identity $g^{ij}g_{ij} = 3$ has been used. This can be solved for R (end of trick):

$$R = -2g^{00}R_{00} \quad (115)$$

With $g^{00} = 1/g_{00}$ the Ricci tensor is

$$R = -2 \left(\frac{1}{-1 - \frac{2\Phi}{c^2}} \right) \frac{1}{c^2} \vec{\nabla}^2 \Phi \approx 2 \left(1 + \frac{2\Phi}{c^2} \right) \frac{1}{c^2} \vec{\nabla}^2 \Phi \approx \frac{2}{c^2} \vec{\nabla}^2 \Phi + O(\Phi^2) \quad (116)$$

We thus have R and R_{00} , and can evaluate the 00-component of the field equation:

$$\begin{aligned} R_{00} - \frac{1}{2}g_{00}R &= \kappa T_{00} \\ \frac{1}{c^2} \vec{\nabla}^2 \Phi - \frac{1}{2} \left(-1 - \frac{2\Phi}{c^2} \right) \frac{2}{c^2} \vec{\nabla}^2 \Phi &= \kappa \rho c^2 \\ \frac{1}{c^2} \vec{\nabla}^2 \Phi + \frac{1}{2} \left(1 + \frac{2\Phi}{c^2} \right) \frac{2}{c^2} \vec{\nabla}^2 \Phi &= \kappa \rho c^2 \\ \frac{1}{c^2} \vec{\nabla}^2 \Phi + \frac{1}{2} \frac{2}{c^2} \vec{\nabla}^2 \Phi + O(\Phi^2) &= \kappa \rho c^2 \\ \frac{2}{c^2} \vec{\nabla}^2 \Phi &= \kappa \rho c^2 \\ \vec{\nabla}^2 \Phi &= \frac{1}{2} \kappa \rho c^4 \end{aligned} \quad (117)$$

Newtonian gravity is given by the Poisson equation

$$\vec{\nabla}^2 \Phi = 4\pi G \rho \quad (118)$$

and hence $\kappa = \frac{8\pi G}{c^4}$. Finally, the Einstein field equations are

$$\begin{aligned} G_{\mu\nu} &= \frac{8\pi G}{c^4} T_{\mu\nu} \\ G^{\mu\nu} &= \frac{8\pi G}{c^4} T^{\mu\nu} \end{aligned} \quad (119)$$

7.3 Summary

What was the path to get to this equation?

- We started with some assumptions: We look for a tensor relation between curvature and matter that conserves energy and momentum, and that recovers Newton's equations.
- The simple-most choice is $G_{\mu\nu} = \kappa T_{\mu\nu}$.
- So we need a tensor for G , for which the covariant derivative vanishes - as $\nabla T = 0$.
- Using the Bianchi identity, we were able to construct one, consisting of the Ricci tensor and Ricci scalar. This required some assumptions on the metric: That it is torsion-free and that the metric compatibility holds.

- Using the geodesic equation of motion for a weak field / slow motion, and neglecting higher order terms, we saw that the 00-component of the metric relates to the Newtonian potential.
- With that we looked at the 00-equation of the proposed tensor relation, and were able to derive κ such that the Newtonian law is recovered in first order.
- That completes the proposal for the field equations. On whether this was a correct proposal, only experiment can judge.
- The resulting equations are complicated. The Ricci tensor and scalar are functions of the Christoffel symbols and metric tensor, and the Christoffel symbols are functions of the metric themselves - including derivatives. See figure 5.

$$\begin{aligned}
& \frac{1}{2}g^{\alpha\beta}\partial_\alpha\partial_\mu g_{\beta\nu} + \frac{1}{2}g^{\alpha\beta}\partial_\alpha\partial_\nu g_{\mu\beta} - \frac{1}{2}g^{\alpha\beta}\partial_\alpha\partial_\beta g_{\mu\nu} - \frac{3}{2}g^{\alpha\beta}\partial_\mu\partial_\nu g_{\alpha\beta} - \frac{1}{2}g^{\beta\lambda}g^{\alpha\rho}\partial_\alpha g_{\rho\lambda}\partial_\mu g_{\beta\nu} \\
& - \frac{1}{2}g^{\beta\lambda}g^{\alpha\rho}\partial_\alpha g_{\rho\lambda}\partial_\nu g_{\mu\beta} + \frac{1}{4}g^{\beta\lambda}g^{\alpha\rho}\partial_\nu g_{\alpha\lambda}\partial_\mu g_{\rho\beta} + \frac{1}{4|g|}g^{\alpha\beta}\partial_\beta |g| \partial_\nu g_{\mu\alpha} - \frac{1}{4|g|}g^{\alpha\beta}\partial_\beta |g| \partial_\alpha g_{\mu\nu} \\
& - \frac{1}{4|g|}g^{\alpha\beta}\partial_\beta |g| \partial_\mu g_{\alpha\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}
\end{aligned}$$

$$\begin{aligned}
& \frac{1}{2}\sum_{\alpha=0}^3\sum_{\beta=0}^3g^{\alpha\beta}\partial_\alpha\partial_\mu g_{\beta\nu} + \frac{1}{2}\sum_{\alpha=0}^3\sum_{\beta=0}^3g^{\alpha\beta}\partial_\alpha\partial_\nu g_{\mu\beta} - \frac{1}{2}\sum_{\alpha=0}^3\sum_{\beta=0}^3g^{\alpha\beta}\partial_\alpha\partial_\beta g_{\mu\nu} - \frac{3}{2}\sum_{\alpha=0}^3\sum_{\beta=0}^3g^{\alpha\beta}\partial_\mu\partial_\nu g_{\alpha\beta} - \frac{1}{2} \\
& \sum_{\alpha=0}^3\sum_{\beta=0}^3\sum_{\rho=0}^3\sum_{\lambda=0}^3g^{\beta\lambda}g^{\alpha\rho}\partial_\alpha g_{\rho\lambda}\partial_\mu g_{\beta\nu} - \frac{1}{2}\sum_{\alpha=0}^3\sum_{\beta=0}^3\sum_{\rho=0}^3\sum_{\lambda=0}^3g^{\beta\lambda}g^{\alpha\rho}\partial_\alpha g_{\rho\lambda}\partial_\nu g_{\mu\beta} + \frac{1}{4}\sum_{\alpha=0}^3\sum_{\beta=0}^3\sum_{\rho=0}^3 \\
& \sum_{\lambda=0}^3g^{\beta\lambda}g^{\alpha\rho}\partial_\nu g_{\alpha\lambda}\partial_\mu g_{\rho\beta} + \frac{1}{4|g|}\sum_{\alpha=0}^3\sum_{\beta=0}^3g^{\alpha\beta}\partial_\beta |g| \partial_\nu g_{\mu\alpha} - \frac{1}{4|g|}\sum_{\alpha=0}^3\sum_{\beta=0}^3g^{\alpha\beta}\partial_\beta |g| \partial_\alpha g_{\mu\nu} - \frac{1}{4|g|}\sum_{\alpha=0}^3 \\
& \sum_{\beta=0}^3g^{\alpha\beta}\partial_\beta |g| \partial_\mu g_{\alpha\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}
\end{aligned}$$

$$\begin{aligned}
& \frac{1}{2}\sum_{\alpha=0}^3\sum_{\beta=0}^3g^{\alpha\beta}\partial_\alpha\partial_\mu g_{\beta\nu} = \frac{1}{2}g^{00}\partial_0\partial_\mu g_{0\nu} + \frac{1}{2}g^{01}\partial_0\partial_\mu g_{1\nu} + \frac{1}{2}g^{02}\partial_0\partial_\mu g_{2\nu} + \frac{1}{2}g^{03}\partial_0\partial_\mu g_{3\nu} \\
& + \frac{1}{2}g^{10}\partial_1\partial_\mu g_{0\nu} + \frac{1}{2}g^{11}\partial_1\partial_\mu g_{1\nu} + \frac{1}{2}g^{12}\partial_1\partial_\mu g_{2\nu} + \frac{1}{2}g^{13}\partial_1\partial_\mu g_{3\nu} + \frac{1}{2}g^{20}\partial_2\partial_\mu g_{0\nu} + \frac{1}{2}g^{21}\partial_2\partial_\mu g_{1\nu} \\
& + \frac{1}{2}g^{22}\partial_2\partial_\mu g_{2\nu} + \frac{1}{2}g^{23}\partial_2\partial_\mu g_{3\nu} + \frac{1}{2}g^{30}\partial_3\partial_\mu g_{0\nu} + \frac{1}{2}g^{31}\partial_3\partial_\mu g_{1\nu} + \frac{1}{2}g^{32}\partial_3\partial_\mu g_{2\nu} + \frac{1}{2}g^{33}\partial_3\partial_\mu g_{3\nu}
\end{aligned}$$

Figure 5: Expanding the Einstein equations shows how complicated they are. Top: The full equation in terms of the metric tensor (also including the cosmological constant term). Middle: The same, but the sums explicitly spelled out. Bottom: The first term with the sums expanded. Source: Ville Hirvonen

8 Derivation of the Schwarzschild solution

The simple-most solution one can think of is:

- spherically symmetric spacetime: invariant under rotations and taking the mirror image. The symmetry also suggests using spherical coordinates.
- static spacetime: all metric components are independent of the time coordinate t and under time reversal
- vacuum solution
- asymptotically, it should become flat, such that one can embed it into flat Minkowski space

8.1 Exploiting symmetries

Exploit time reversal coordinate transformation: $(t, \vec{x}) \longrightarrow (-t, \vec{x})$. The metric components should stay unchanged. For $i = 1, 2, 3$:

$$g'_{i0} = \frac{\partial x^\alpha}{\partial x'^i} \frac{\partial x^\beta}{\partial x'^0} g_{\alpha\beta} = \frac{\partial x^\alpha}{\partial x'^i} \frac{\partial x^\beta}{c \partial t'} g_{\alpha\beta} = \frac{\partial x^\alpha}{\partial x^i} \frac{\partial x^\beta}{-c \partial t} g_{\alpha\beta} = -\frac{\partial x^\alpha}{\partial x^i} \frac{\partial x^\beta}{\partial x^0} g_{\alpha\beta} = -g_{i0} \quad (120)$$

Thus $g_{i0} = 0 = g_{0i}$. Exploiting spatial symmetries: $(ct, r, \theta, \phi) \longrightarrow (ct, r, -\theta, \phi)$ and $(ct, r, \theta, \phi) \longrightarrow (ct, r, \theta, -\phi)$ yield

$$\begin{aligned} g_{\mu 2} = g_{2\mu} &= 0 \quad (\mu \neq 2) \\ g_{\mu 3} = g_{3\mu} &= 0 \quad (\mu \neq 3) \end{aligned} \quad (121)$$

Together: $g_{\mu\nu} = 0$ for $\mu \neq \nu$. Thus the metric is diagonal:

$$\begin{aligned} ds^2 &= g_{00}c^2dt^2 + g_{11}dr^2 + g_{22}d\theta^2 + g_{33}d\phi^2 \\ &= g_{tt}c^2dt^2 + g_{rr}dr^2 + g_{\theta\theta}d\theta^2 + g_{\phi\phi}d\phi^2 \end{aligned} \quad (122)$$

Since the metric is static, none of $g_{00}, g_{rr}, g_{\theta\theta}, g_{\phi\phi}$ can depend on t . Also, we can exploit the spherical symmetry: On a radial line (a "hypersurface" of constant t, θ, ϕ) g_{rr} can only depend on r :

$$g_{rr} = A(r) \quad (123)$$

Similarly, g_{tt} can only depend on r :

$$g_{tt} = -B(r) \quad (124)$$

On a hypersurface of constant t_0 and constant r_0 , the metric must be that of a sphere:

$$g_{\theta\theta}d\theta^2 + g_{\phi\phi}d\phi^2 = dl^2 = r_0^2(d\theta^2 + \sin^2\theta d\phi^2) \quad (125)$$

Hence, $g_{\theta\theta} = r_0^2$, $g_{\phi\phi} = r_0^2 \sin^2\theta$. As this holds for any t_0, r_0 , we have

$$\begin{aligned} g_{\theta\theta} &= r^2 \\ g_{\phi\phi} &= r^2 \sin^2\theta \end{aligned} \quad (126)$$

We thus get

$$ds^2 = -B(r)c^2dt^2 + A(r)dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\phi^2 \quad (127)$$

or

$$g_{\mu\nu} = \begin{pmatrix} -B(r) & 0 & 0 & 0 \\ 0 & A(r) & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2\theta \end{pmatrix} \quad (128)$$

8.2 Calculation of the Christoffel symbols

Next, we need to evaluate the $4^3 = 64$ Christoffel symbols $\Gamma_{\mu\nu}^\lambda$. Due to the symmetry in μ, ν , there are actually fewer, namely 40.

8.2.1 $\Gamma_{\mu\nu}^0$

Let's start with $\Gamma_{\mu\nu}^0$:

$$\Gamma_{00}^0 = \frac{1}{2}g^{0\alpha}(\partial_0 g_{\alpha 0} + \partial_0 g_{\alpha 0} - \partial_\alpha g_{00}) = -\frac{1}{2}g^{0\alpha}\partial_\alpha g_{00} = -\frac{1}{2}g^{00}\partial_0 g_{00} = 0 \quad (129)$$

(The second equality holds since the metric static, i.e. $\partial_0 g_{\mu\nu}$ vanishes, the third uses $g^{0i} = 0$ and the last again $\partial_0 \rightarrow 0$.)

$$\Gamma_{10}^0 = \Gamma_{01}^0 = \frac{1}{2}g^{0\alpha}(\partial_0 g_{\alpha 1} + \partial_1 g_{\alpha 0} - \partial_\alpha g_{01}) = \frac{1}{2}g^{00}(0 + \partial_1 g_{00} - 0) = \frac{1}{2}(g_{00})^{-1}\partial_r g_{00} = \frac{B'(r)}{2B(r)} \quad (130)$$

$$\begin{aligned} \Gamma_{20}^0 = \Gamma_{02}^0 &= \frac{1}{2}g^{0\alpha}(\partial_0 g_{\alpha 2} + \partial_2 g_{\alpha 0} - \partial_\alpha g_{02}) = \frac{1}{2}g^{00}(0 + \partial_2 g_{00} - 0) = 0 \\ \Gamma_{30}^0 = \Gamma_{03}^0 &= \frac{1}{2}g^{0\alpha}(\partial_0 g_{\alpha 3} + \partial_3 g_{\alpha 0} - \partial_\alpha g_{03}) = \frac{1}{2}g^{00}(0 + \partial_3 g_{00} - 0) = 0 \end{aligned} \quad (131)$$

(because $g_{00} = -B(r)$ is not a function of θ or ϕ .)

$$\Gamma_{11}^0 = \frac{1}{2}g^{0\alpha}(\partial_1 g_{\alpha 1} + \partial_1 g_{\alpha 1} - \partial_\alpha g_{11}) = -\frac{1}{2}g^{0\alpha}\partial_\alpha g_{11} = -\frac{1}{2}g^{00}\partial_0 g_{11} = 0 \quad (132)$$

(the second equality is because $g_{\alpha 1}$ is $\neq 0$ only for $\alpha = 1$, but then $g^{0\alpha}$ vanishes; the third is because $g^{0\alpha}$ is $\neq 0$ only for $\alpha = 0$, and the last is using $\partial_0 \rightarrow 0$)

$$\begin{aligned} \Gamma_{12}^0 = \Gamma_{21}^0 &= \frac{1}{2}g^{0\alpha}(\partial_1 g_{\alpha 2} + \partial_2 g_{\alpha 1} - \partial_\alpha g_{12}) = (0 + 0 - 0) = 0 \\ \Gamma_{13}^0 = \Gamma_{31}^0 &= \frac{1}{2}g^{0\alpha}(\partial_1 g_{\alpha 3} + \partial_3 g_{\alpha 1} - \partial_\alpha g_{13}) = (0 + 0 - 0) = 0 \\ \Gamma_{23}^0 = \Gamma_{32}^0 &= \frac{1}{2}g^{0\alpha}(\partial_2 g_{\alpha 3} + \partial_3 g_{\alpha 2} - \partial_\alpha g_{23}) = (0 + 0 - 0) = 0 \end{aligned} \quad (133)$$

(Line 1: The first term vanishes as $g_{\alpha 2}$ is $\neq 0$ only for $\alpha = 2$, but then $g^{0\alpha} = 0$. The second as $g_{\alpha 1} \neq 0$ only for $\alpha = 1$, but then $g^{0\alpha} = 0$. The last as $g_{12} = 0$. Line 2: The first term vanishes as $g_{\alpha 3}$ is $\neq 0$ only for $\alpha = 3$, but then $g^{0\alpha} = 0$. The second as $g_{\alpha 1} \neq 0$ only for $\alpha = 1$, but then $g^{0\alpha} = 0$. The last as $g_{13} = 0$. Line 3: The first term vanishes as $g_{\alpha 3}$ is $\neq 0$ only for $\alpha = 3$, but then $g^{0\alpha} = 0$. The second term vanishes as $g_{\alpha 2}$ is $\neq 0$ only for $\alpha = 3$, but then $g^{0\alpha} = 0$. The last as $g_{23} = 0$.)

$$\begin{aligned} \Gamma_{22}^0 &= \frac{1}{2}g^{0\alpha}(\partial_2 g_{\alpha 2} + \partial_2 g_{\alpha 2} - \partial_\alpha g_{22}) = -\frac{1}{2}g^{0\alpha}\partial_\alpha g_{22} = -\frac{1}{2}g^{00}\partial_0 g_{22} = 0 \\ \Gamma_{33}^0 &= \frac{1}{2}g^{0\alpha}(\partial_3 g_{\alpha 3} + \partial_3 g_{\alpha 3} - \partial_\alpha g_{33}) = -\frac{1}{2}g^{0\alpha}\partial_\alpha g_{33} = -\frac{1}{2}g^{00}\partial_0 g_{33} = 0 \end{aligned} \quad (134)$$

(The second equality is because $g^{0\alpha} \neq 0$ only for $\alpha = 0$, but then $g_{\alpha 2/3} = 0$, the third as $g^{0\alpha} \neq 0$ only for $\alpha = 0$, the last as $\partial_0 \rightarrow 0$.) Thus:

$$\Gamma^0_{\mu\nu} = \begin{pmatrix} 0 & \frac{B'(r)}{2B(r)} & 0 & 0 \\ \frac{B'(r)}{2B(r)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (135)$$

8.2.2 $\Gamma^1_{\mu\nu}$

Next, let's do $\Gamma^1_{\mu\nu}$.

$$\Gamma^1_{00} = \frac{1}{2}g^{1\alpha}(\partial_0 g_{\alpha 0} + \partial_0 g_{\alpha 0} - \partial_\alpha g_{00}) = -\frac{1}{2}g^{1\alpha}\partial_\alpha g_{00} = -\frac{1}{2}g^{11}\partial_1 g_{00} = -\frac{1}{2}(g_{11})^{-1}\partial_r g_{00} = \frac{B'(r)}{2A(r)} \quad (136)$$

For $i = 1, 2, 3$:

$$\Gamma^1_{0i} = \Gamma^1_{i0} = \frac{1}{2}g^{1\alpha}(\partial_0 g_{\alpha i} + \partial_i g_{\alpha 0} - \partial_\alpha g_{0i}) = \frac{1}{2}g^{11}(\partial_i g_{10} - \partial_1 g_{0i}) = 0 - 0 = 0 \quad (137)$$

(The second equality uses that only $\alpha = 1$ can contribute and that $\partial_0 \rightarrow 0$. The third holds as $g_{10} = 0$ and $g_{0i} = 0$.)

$$\begin{aligned} \Gamma^1_{11} &= \frac{1}{2}g^{1\alpha}(\partial_1 g_{\alpha 1} + \partial_1 g_{\alpha 1} - \partial_\alpha g_{11}) = \frac{1}{2}g^{11}(\partial_1 g_{11} + \partial_1 g_{11} - \partial_1 g_{11}) \\ &= \frac{1}{2}g^{11}\partial_1 g_{11} = \frac{1}{2}(g_{11})^{-1}\partial_r g_{11} = \frac{A'(r)}{2A(r)} \end{aligned} \quad (138)$$

$$\begin{aligned} \Gamma^1_{12} = \Gamma^1_{21} &= \frac{1}{2}g^{1\alpha}(\partial_1 g_{\alpha 2} + \partial_2 g_{\alpha 1} - \partial_\alpha g_{12}) = \frac{1}{2}g^{11}(\partial_1 g_{12} + \partial_2 g_{11} - 0) = 0 + 0 = 0 \\ \Gamma^1_{13} = \Gamma^1_{31} &= \frac{1}{2}g^{1\alpha}(\partial_1 g_{\alpha 3} + \partial_3 g_{\alpha 1} - \partial_\alpha g_{13}) = \frac{1}{2}g^{11}(\partial_1 g_{13} + \partial_3 g_{11} - 0) = 0 + 0 = 0 \\ \Gamma^1_{23} = \Gamma^1_{32} &= \frac{1}{2}g^{1\alpha}(\partial_2 g_{\alpha 3} + \partial_3 g_{\alpha 2} - \partial_\alpha g_{23}) = \frac{1}{2}g^{11}(\partial_2 g_{13} + \partial_3 g_{11} - 0) = 0 + 0 = 0 \end{aligned} \quad (139)$$

(First line: The second equality uses that only $\alpha = 1$ can contribute and that $g_{12} = 0$. The third uses $g_{12} = 0$ and that $g_{11} = A(r)$ is not a function of θ . Second line: The second equality uses that only $\alpha = 1$ can contribute and that $g_{13} = 0$. The third uses $g_{13} = 0$ and that $g_{11} = A(r)$ is not a function of ϕ . Third line: The second equality uses that only $\alpha = 1$ can contribute and that $g_{23} = 0$. The third uses $g_{13} = 0$ and that $g_{11} = A(r)$ is not a function of ϕ .)

$$\begin{aligned} \Gamma^1_{22} &= \frac{1}{2}g^{1\alpha}(\partial_2 g_{\alpha 2} + \partial_2 g_{\alpha 2} - \partial_\alpha g_{22}) = \frac{1}{2}g^{11}(\partial_2 g_{12} + \partial_2 g_{12} - \partial_1 g_{22}) \\ &= -\frac{1}{2}g^{11}\partial_r g_{22} = -\frac{1}{2}(g_{11})^{-1}\partial_r g_{22} = -\frac{1}{2A(r)}\partial_r r^2 = -\frac{r}{A(r)} \end{aligned} \quad (140)$$

$$\begin{aligned} \Gamma^1_{33} &= \frac{1}{2}g^{1\alpha}(\partial_3 g_{\alpha 3} + \partial_3 g_{\alpha 3} - \partial_\alpha g_{33}) = \frac{1}{2}g^{11}(\partial_3 g_{12} + \partial_3 g_{12} - \partial_1 g_{33}) \\ &= -\frac{1}{2}g^{11}\partial_r g_{33} = -\frac{1}{2}(g_{11})^{-1}\partial_r g_{33} = -\frac{1}{2A(r)}\partial_r r^2 \sin^2 \theta = -\frac{r \sin^2 \theta}{A(r)} \end{aligned} \quad (141)$$

Thus:

$$\Gamma^1_{\mu\nu} = \begin{pmatrix} \frac{B'(r)}{2A(r)} & 0 & 0 & 0 \\ 0 & \frac{A'(r)}{2A(r)} & 0 & 0 \\ 0 & 0 & -\frac{r}{A(r)} & 0 \\ 0 & 0 & 0 & -\frac{r \sin^2 \theta}{A(r)} \end{pmatrix} \quad (142)$$

8.2.3 $\Gamma^2_{\mu\nu}$

Moving on to the $\Gamma^2_{\mu\nu}$.

$$\Gamma^2_{00} = \frac{1}{2}g^{2\alpha}(\partial_0 g_{\alpha 0} + \partial_0 g_{\alpha 0} - \partial_\alpha g_{00}) = -\frac{1}{2}g^{22}\partial_2 g_{00} = 0 \quad (143)$$

(Using $\partial_0 \rightarrow 0$ and that $g_{00} = -B(r)$ is not a function of θ .) For $i = 1, 2, 3$:

$$\Gamma^2_{0i} = \Gamma^2_{i0} = \frac{1}{2}g^{2\alpha}(\partial_0 g_{\alpha i} + \partial_i g_{\alpha 0} - \partial_\alpha g_{0i}) = \frac{1}{2}g^{22}(\partial_i g_{20} - \partial_1 g_{02}) = 0 - 0 = 0 \quad (144)$$

(Using $\partial_0 \rightarrow 0$, that only $\alpha = 2$ can contribute and that $g_{20} = g_{02} = 0$).

$$\Gamma^2_{11} = \frac{1}{2}g^{2\alpha}(\partial_1 g_{\alpha 1} + \partial_1 g_{\alpha 1} - \partial_\alpha g_{11}) = \frac{1}{2}g^{22}(\partial_1 g_{21} + \partial_1 g_{21} - \partial_2 g_{11}) = 0 + 0 - 0 = 0 \quad (145)$$

(Using that only $\alpha = 2$ can contribute and that $g_{11} = A(r)$ is not a function of θ .)

$$\begin{aligned} \Gamma^2_{12} = \Gamma^2_{21} &= \frac{1}{2}g^{2\alpha}(\partial_1 g_{\alpha 2} + \partial_2 g_{\alpha 1} - \partial_\alpha g_{12}) = \frac{1}{2}g^{22}(\partial_1 g_{22} + \partial_2 g_{21} - \partial_2 g_{21}) \\ &= \frac{1}{2}(g_{22})^{-1}\partial_1 g_{22} = \frac{1}{2}\frac{1}{r^2}\partial_r r^2 = \frac{1}{r} \end{aligned} \quad (146)$$

$$\Gamma^2_{22} = \frac{1}{2}g^{2\alpha}(\partial_2 g_{\alpha 2} + \partial_2 g_{\alpha 2} - \partial_\alpha g_{22}) = \frac{1}{2}g^{22}(\partial_2 g_{22} + \partial_2 g_{22} - \partial_2 g_{22}) = 0 + 0 - 0 = 0 \quad (147)$$

($g_{22} = r^2$ does not depend θ).

$$\begin{aligned} \Gamma^2_{33} &= \frac{1}{2}g^{2\alpha}(\partial_3 g_{\alpha 3} + \partial_3 g_{\alpha 3} - \partial_\alpha g_{33}) = \frac{1}{2}g^{22}(\partial_3 g_{23} + \partial_2 g_{23} - \partial_2 g_{33}) \\ &= -\frac{1}{2}(g_{22})^{-1}\partial_\theta g_{33} = -\frac{1}{2r^2}\partial_\theta r^2 \cos^2 \theta = -\cos \theta \sin \theta \end{aligned} \quad (148)$$

$$\Gamma^2_{13} = \Gamma^2_{31} = \frac{1}{2}g^{2\alpha}(\partial_1 g_{\alpha 3} + \partial_3 g_{\alpha 1} - \partial_\alpha g_{13}) = \frac{1}{2}g^{22}(\partial_1 g_{23} + \partial_3 g_{21} - 0) = 0 \quad (149)$$

(only $\alpha = 2$ contributes)

$$\Gamma^2_{23} = \Gamma^2_{32} = \frac{1}{2}g^{2\alpha}(\partial_2 g_{\alpha 3} + \partial_3 g_{\alpha 2} - \partial_\alpha g_{23}) = \frac{1}{2}g^{22}(\partial_2 g_{23} + \partial_3 g_{22} - 0) = 0 \quad (150)$$

(only $\alpha = 2$ contributes, and $g_{22} = r^2$ is not a function ϕ .) Thus:

$$\Gamma^2_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1/r & 0 \\ 0 & 1/r & 0 & 0 \\ 0 & 0 & 0 & -\cos \theta \sin \theta \end{pmatrix} \quad (151)$$

8.2.4 $\Gamma_{\mu\nu}^3$

And finally the $\Gamma_{\mu\nu}^3$.

$$\Gamma_{00}^3 = \frac{1}{2}g^{3\alpha}(\partial_0 g_{\alpha 0} + \partial_0 g_{\alpha 0} - \partial_\alpha g_{00}) = -\frac{1}{2}g^{33}\partial_3 g_{00} = 0 \quad (152)$$

(Using $\partial_0 \rightarrow 0$ and that $g_{00} = -B(r)$ is not a function of ϕ .) For $i = 1, 2, 3$:

$$\Gamma_{0i}^3 = \Gamma_{i0}^3 = \frac{1}{2}g^{3\alpha}(\partial_0 g_{\alpha i} + \partial_i g_{\alpha 0} - \partial_\alpha g_{0i}) = \frac{1}{2}g^{33}(\partial_i g_{30} - \partial_3 g_{0i}) = 0 - 0 = 0 \quad (153)$$

$$\Gamma_{11}^3 = \frac{1}{2}g^{3\alpha}(\partial_1 g_{\alpha 1} + \partial_1 g_{\alpha 1} - \partial_\alpha g_{11}) = \frac{1}{2}g^{33}(\partial_1 g_{31} + \partial_1 g_{31} - \partial_3 g_{11}) = 0 + 0 - 0 = 0 \quad (154)$$

(Using that only $\alpha = 3$ can contribute, that $g_{31} = 0$ and that $g_{11} = A(r)$ is not a function of ϕ .)

$$\Gamma_{12}^3 = \Gamma_{21}^3 = \frac{1}{2}g^{3\alpha}(\partial_1 g_{\alpha 2} + \partial_2 g_{\alpha 1} - \partial_\alpha g_{12}) = \frac{1}{2}g^{33}(\partial_1 g_{32} + \partial_2 g_{31} - \partial_3 g_{21}) = 0 \quad (155)$$

$$\Gamma_{22}^3 = \frac{1}{2}g^{3\alpha}(\partial_2 g_{\alpha 2} + \partial_2 g_{\alpha 2} - \partial_\alpha g_{22}) = \frac{1}{2}g^{33}(\partial_2 g_{32} + \partial_2 g_{32} - \partial_3 g_{22}) = 0 \quad (156)$$

($g_{22} = r^2$ is not a function ϕ .)

$$\begin{aligned} \Gamma_{13}^3 = \Gamma_{31}^3 &= \frac{1}{2}g^{3\alpha}(\partial_1 g_{\alpha 3} + \partial_3 g_{\alpha 1} - \partial_\alpha g_{13}) = \frac{1}{2}g^{33}(\partial_1 g_{33} + \partial_3 g_{31} - 0) \\ &= \frac{1}{2}(g_{33})^{-1}\partial_r r^2 \sin^2 \theta = \frac{1}{2}\frac{1}{r^2 \sin^2 \theta} 2r \sin^2 \theta = \frac{1}{r} \end{aligned} \quad (157)$$

$$\begin{aligned} \Gamma_{23}^3 = \Gamma_{32}^3 &= \frac{1}{2}g^{3\alpha}(\partial_2 g_{\alpha 3} + \partial_3 g_{\alpha 2} - \partial_\alpha g_{23}) = \frac{1}{2}g^{33}(\partial_2 g_{33} + \partial_3 g_{32} - 0) \\ &= \frac{1}{2}(g_{33})^{-1}\partial_\theta r^2 \sin^2 \theta = \frac{1}{2}\frac{1}{r^2 \sin^2 \theta} r^2 2 \cos \theta \sin \theta = \cot \theta \end{aligned} \quad (158)$$

Thus:

$$\Gamma_{\mu\nu}^3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{r} \\ 0 & 0 & 0 & \cot \theta \\ 0 & \frac{1}{r} & \cot \theta & 0 \end{pmatrix} \quad (159)$$

We are done!

8.2.5 Summary of Christoffel symbols calculated

For the following, it is handy to have all non-vanishing Christoffel symbols collected.

$$\begin{aligned} \Gamma_{10}^0 = \Gamma_{01}^0 &= \frac{B'}{2B} & \Gamma_{00}^1 &= \frac{B'}{2A} & \Gamma_{12}^2 = \Gamma_{21}^2 &= \frac{1}{r} & \Gamma_{13}^3 = \Gamma_{31}^3 &= \frac{1}{r} \\ & & \Gamma_{11}^1 &= \frac{A'}{2A} & \Gamma_{33}^2 &= -\cos \theta \sin \theta & \Gamma_{23}^3 = \Gamma_{32}^3 &= \cot \theta \\ & & \Gamma_{22}^1 &= -\frac{r}{A} & & & & \\ & & \Gamma_{33}^1 &= -\frac{r \sin^2 \theta}{A} & & & & \end{aligned} \quad (160)$$

8.2.6 Tricks for Christoffels

For contracted Christoffel symbols $\Gamma^\mu_{\mu\nu} = \Gamma^\mu_{\nu\mu}$ (these occur in two of the four terms for the Ricci tensor) exists a neat trick.

$$\begin{aligned}
\Gamma^\mu_{\mu\nu} &= \frac{1}{2}g^{\mu\alpha}(\partial_\mu g_{\alpha\nu} + \partial_\nu g_{\mu\alpha} - \partial_\alpha g_{\mu\nu}) \\
&= \frac{1}{2}g^{\mu\alpha}\partial_\mu g_{\alpha\nu} - \frac{1}{2}g^{\mu\alpha}\partial_\alpha g_{\mu\nu} + \frac{1}{2}g^{\mu\alpha}\partial_\nu g_{\mu\alpha} \\
&= \frac{1}{2}g^{\mu\alpha}\partial_\mu g_{\alpha\nu} - \frac{1}{2}g^{\alpha\mu}\partial_\mu g_{\alpha\nu} + \frac{1}{2}g^{\mu\alpha}\partial_\nu g_{\mu\alpha} \\
&= \frac{1}{2}g^{\mu\alpha}\partial_\nu g_{\mu\alpha}
\end{aligned} \tag{161}$$

In line 3 we changed the names of the summation indices μ to α and α to μ . The trick is based on the linear algebra fact that for any matrix

$$\ln |M| = \text{tr}(\ln M) \tag{162}$$

where $|M|$ is the determinant of the matrix. Take the derivative ∂_ν on the left side:

$$\partial_\nu \ln |M| = \frac{1}{|M|} \partial_\nu |M| \tag{163}$$

And on the right side:

$$\partial_\nu \text{tr}(\ln M) = \text{tr}(\partial_\nu \ln M) = \text{tr}(M^{-1} \partial_\nu M) = M^{\mu\alpha} \partial_\nu M_{\mu\alpha} \tag{164}$$

This is just what we found in equation 161 for the Christoffel symbol, if $M = g$. Thus

$$\Gamma^\mu_{\mu\nu} = \Gamma^\mu_{\nu\mu} = \frac{1}{2|g|} \partial_\nu |g| = \partial_\nu \ln \sqrt{|g|} \tag{165}$$

The formula is particularly useful for diagonal metrics, since $|g| = g_{00}g_{11}g_{22}g_{33}$ is just the product of the diagonal elements. For example, for our ansatz for the Schwarzschild metric, $|g| = A B r^4 \sin^2 \theta$ yields directly

$$\Gamma^0_{01} + \Gamma^1_{11} + \Gamma^2_{21} + \Gamma^3_{31} = \frac{A'}{2A} + \frac{B'}{2B} + \frac{2}{r} \tag{166}$$

Another useful way to get the Christoffel symbols is exploiting section 5. From the Lagrangian \mathcal{L}' (which is essentially $1/2$ times the line element divided by $d\tau^2$!), one gets the equation of motions via the Euler-Lagrange mechanism. These are four equations, for each of the coordinates one for its second derivative. The geodesic equation is identical to these, and in its definition one has the Christoffel symbols. Hence, one writes the four equations for $\ddot{x}^0, \ddot{x}^1, \ddot{x}^2, \ddot{x}^3$, each of which will contain four terms for the squares of velocities and six cross terms. By comparing the coefficients, one can read off the Christoffel symbols. Let's try this. The geodesic equation written out for component μ is

$$\begin{aligned}
-\ddot{x}^\mu &= \Gamma^\mu_{00}(u^0)^2 + 2\Gamma^\mu_{01}u^0u^1 + 2\Gamma^\mu_{02}u^0u^2 + 2\Gamma^\mu_{03}u^0u^3 + \\
&\quad \Gamma^\mu_{11}(u^1)^2 + 2\Gamma^\mu_{12}u^1u^2 + 2\Gamma^\mu_{13}u^1u^3 + \\
&\quad \Gamma^\mu_{22}(u^2)^2 + 2\Gamma^\mu_{23}u^2u^3 + \\
&\quad \Gamma^\mu_{33}(u^3)^2
\end{aligned} \tag{167}$$

Let's look at our ansatz for the Schwarzschild metric:

$$\mathcal{L}' = -\frac{1}{2}B(r)(u^t)^2 + \frac{1}{2}A(r)(u^r)^2 + \frac{1}{2}r^2(u^\theta)^2 + \frac{1}{2}r^2 \sin^2 \theta (u^\phi)^2 \tag{168}$$

For the 0-component, we have

$$\begin{aligned}
\frac{\partial \mathcal{L}'}{\partial x^t} &= 0 \\
\frac{d}{d\tau} \frac{\partial \mathcal{L}'}{\partial u^t} &= -\frac{d}{d\tau}(B(r)u^t) = -B'\dot{r}u^t - B\dot{u}^t = -B'u^r u^t - B\ddot{x}^t \\
-\ddot{x}^0 = -\ddot{x}^t &= \frac{B'}{B}u^t u^r = \frac{B'}{B}u^0 u^1
\end{aligned} \tag{169}$$

Comparing the coefficients with equation 167, one sees that only $\Gamma_{10}^0 = \Gamma_{01}^0$ does not vanish, and we find (like before) $\Gamma_{10}^0 = \Gamma_{01}^0 = B'/(2B)$ as in equation 160. The 1-component is slightly more work:

$$\begin{aligned}
\frac{\partial \mathcal{L}'}{\partial r} &= -\frac{1}{2}B'(u^t)^2 + \frac{1}{2}A'(u^r)^2 + r(u^\theta)^2 + r \sin^2 \theta (u^\phi)^2 \\
\frac{d}{d\tau} \frac{\partial \mathcal{L}'}{\partial u^r} &= \frac{d}{d\tau}(A(r)u^r) = A'\dot{r}u^r + A\dot{u}^r = A'(u^r)^2 + A\ddot{x}^r \\
-\ddot{x}^1 = -\ddot{x}^r &= \frac{A'}{A}(u^r)^2 + \frac{1}{2}\frac{B'}{A}(u^t)^2 - \frac{1}{2}\frac{A'}{A}(u^r)^2 - \frac{r}{A}(u^\theta)^2 - \frac{r}{A}\sin^2 \theta (u^\phi)^2 \\
&= \frac{1}{2}\frac{A'}{A}(u^r)^2 + \frac{1}{2}\frac{B'}{A}(u^t)^2 - \frac{r}{A}(u^\theta)^2 - \frac{r}{A}\sin^2 \theta (u^\phi)^2
\end{aligned} \tag{170}$$

Again, we can confirm by comparing coefficients with equation 167 that equation 160 is correct, now for the 1-component. The 2-component:

$$\begin{aligned}
\frac{\partial \mathcal{L}'}{\partial \theta} &= r^2 \sin \theta \cos \theta (u^\phi)^2 \\
\frac{d}{d\tau} \frac{\partial \mathcal{L}'}{\partial u^t} &= \frac{d}{d\tau}(r^2 u^\theta) = 2r\dot{r}u^\theta + r^2\dot{u}^\theta = 2ru^r u^\theta + r^2\ddot{x}^\theta \\
-\ddot{x}^2 = -\ddot{x}^\theta &= \frac{2}{r}u^r u^\theta - \sin \theta \cos \theta (u^\phi)^2 = \frac{2}{r}u^1 u^2 - \sin \theta \cos \theta (u^3)^2
\end{aligned} \tag{171}$$

Also here, we get by comparing coefficients with equation 167 the correct Christoffel symbols from equation 160. Finally the 3-component:

$$\begin{aligned}
\frac{\partial \mathcal{L}'}{\partial \phi} &= 0 \\
\frac{d}{d\tau} \frac{\partial \mathcal{L}'}{\partial u^\phi} &= \frac{d}{d\tau}(r^2 \sin^2 \theta u^\phi) = 2r u^r \sin^2 \theta u^\phi + 2r^2 \sin \theta \cos \theta u^\theta u^\phi + r^2 \sin^2 \theta \ddot{x}^\phi \\
-\ddot{x}^3 = -\ddot{x}^\phi &= \frac{2}{r}u^r u^\phi + \cot \theta u^\theta u^\phi = \frac{2}{r}u^1 u^3 + \cot \theta u^2 u^3
\end{aligned} \tag{172}$$

And also the final comparison with equation 167 confirms what we had in equation 160. Overall, this way of calculating the Christoffel symbols seems more elegant, faster and less error-prone.

8.3 The vacuum solution

Vacuum means $T_{\mu\nu} = 0$, i.e. we have $G_{\mu\nu} = 0$, which implies $R_{\mu\nu} = 0$ (as then the Ricci scalar also vanishes). We have thus to solve the $4^2 = 16$ field equations. It works simply by brute-force: Plugging into the Ricci tensor the Christoffel symbols just retrieved. What one uses then repeatedly is:

- Many Christoffel symbols are 0.
- Time derivatives ∂_0 are 0 as the metric is static.
- The coordinates are independent, i.e. a function f only depending on θ does not depend on r : $\partial_r f(\theta) = 0$
- The product rule of derivation: $(fg)' = f'g + fg'$
- Expanding out the contracted sums, like $\Gamma^\lambda_{\mu\lambda} = \Gamma^0_{\mu 0} + \Gamma^1_{\mu 1} + \Gamma^2_{\mu 2} + \Gamma^3_{\mu 3}$

All this is not difficult, just cumbersome.

$$\begin{aligned}
R_{00} &= R^\lambda_{0\lambda 0} = \partial_\beta \Gamma^\beta_{00} - \partial_0 \Gamma^\beta_{0\beta} + \Gamma^\beta_{00} \Gamma^\sigma_{\beta\sigma} - \Gamma^\beta_{\sigma 0} \Gamma^\sigma_{0\beta} \\
&= \partial_0 \Gamma^0_{00} + \partial_1 \Gamma^1_{00} + \partial_2 \Gamma^2_{00} + \partial_3 \Gamma^3_{00} - 0 + \dots \\
&= 0 + \partial_r \frac{B'(r)}{2A(r)} + 0 + 0 + \Gamma^0_{00} \Gamma^\sigma_{0\sigma} + \Gamma^1_{00} \Gamma^\sigma_{1\sigma} + \Gamma^2_{00} \Gamma^\sigma_{2\sigma} + \Gamma^3_{00} \Gamma^\sigma_{3\sigma} - \dots \\
&= \frac{B''(r)}{2A(r)} - \frac{B'(r)}{2A^2(r)} A'(r) + 0 + \frac{B'(r)}{2A(r)} (\Gamma^0_{10} + \Gamma^1_{11} + \Gamma^2_{12} + \Gamma^3_{13}) + 0 + 0 - \dots \\
&= \frac{B''}{2A} - \frac{B'}{2A^2} A' + \frac{B'}{2A} \left(\frac{B'}{2B} + \frac{A'}{2A} + \frac{1}{r} + \frac{1}{r} \right) - \dots \\
&= \dots - (\Gamma^0_{00} \Gamma^0_{00} + \Gamma^0_{10} \Gamma^1_{00} + \Gamma^0_{20} \Gamma^2_{00} + \Gamma^0_{30} \Gamma^3_{00} \\
&\quad + \Gamma^1_{00} \Gamma^0_{01} + \Gamma^1_{10} \Gamma^1_{01} + \Gamma^1_{20} \Gamma^2_{01} + \Gamma^1_{30} \Gamma^3_{01} \\
&\quad + \Gamma^2_{00} \Gamma^0_{02} + \Gamma^2_{10} \Gamma^1_{02} + \Gamma^2_{20} \Gamma^2_{02} + \Gamma^2_{30} \Gamma^3_{02} \\
&\quad + \Gamma^3_{00} \Gamma^0_{03} + \Gamma^3_{10} \Gamma^1_{03} + \Gamma^3_{20} \Gamma^2_{03} + \Gamma^3_{30} \Gamma^3_{03}) \\
&= \dots - \left[(0 + \frac{B'}{2B} \frac{B'}{2A} + 0 + 0) + (\frac{B'}{2A} \frac{B'}{2B} + 0 + 0 + 0) + (4 \times 0) + (4 \times 0) \right] \\
&= \frac{B''}{2A} - \frac{A'B'}{2A^2} + \frac{B'^2}{4AB} + \frac{A'B'}{4A^2} + \frac{B'}{rA} - \frac{B'^2}{2BA} \\
&= \frac{B''}{2A} - \frac{A'B'}{4A^2} - \frac{B'^2}{4AB} + \frac{B'}{rA}
\end{aligned} \tag{173}$$

With $R_{00} = 0$ we thus get

$$0 = 2rABB'' - rBA'B' - rAB'^2 + 4ABB' \tag{174}$$

$$\begin{aligned}
R_{11} &= R^\lambda_{\lambda 11} = \partial_\beta \Gamma^\beta_{11} - \partial_1 \Gamma^\beta_{1\beta} + \Gamma^\beta_{11} \Gamma^\sigma_{\beta\sigma} - \Gamma^\beta_{\sigma 1} \Gamma^\sigma_{1\beta} \\
&= \partial_0 \Gamma^0_{11} + \partial_1 \Gamma^1_{11} + \partial_2 \Gamma^2_{11} + \partial_3 \Gamma^3_{11} - \partial_1 (\Gamma^0_{10} + \Gamma^1_{11} + \Gamma^2_{12} + \Gamma^3_{13}) + \dots \\
&= 0 + \partial_r \frac{A'(r)}{2A(r)} + 0 + 0 - \partial_r \left(\frac{B'(r)}{2B(r)} + \frac{A'(r)}{2A(r)} + \frac{1}{r} + \frac{1}{r} \right) + \dots \\
&= -\frac{B''}{2B} + \frac{B'}{2B^2} B' + \frac{2}{r^2} + \Gamma^0_{11} \Gamma^\sigma_{0\sigma} + \Gamma^1_{11} \Gamma^\sigma_{1\sigma} + \Gamma^2_{11} \Gamma^\sigma_{2\sigma} + \Gamma^3_{11} \Gamma^\sigma_{3\sigma} - \dots \\
&= -\frac{B''}{2B} + \frac{B'^2}{2B^2} + \frac{2}{r^2} + 0 + \frac{A'}{2A} (\Gamma^0_{10} + \Gamma^1_{11} + \Gamma^2_{12} + \Gamma^3_{13}) + 0 + 0 - \dots \\
&= -\frac{B''}{2B} + \frac{B'^2}{2B^2} + \frac{2}{r^2} + \frac{A'}{2A} \left(\frac{B'}{2B} + \frac{A'}{2A} + \frac{1}{r} + \frac{1}{r} \right) - \dots \\
&= -\frac{B''}{2B} + \frac{B'^2}{2B^2} + \frac{2}{r^2} + \frac{A'B'}{4AB} + \frac{A'^2}{4A^2} + \frac{A'}{rA} - \dots \\
&= \dots - (\Gamma^0_{01} \Gamma^0_{10} + \Gamma^0_{11} \Gamma^1_{10} + \Gamma^0_{21} \Gamma^2_{10} + \Gamma^0_{31} \Gamma^3_{10} \\
&\quad + \Gamma^1_{01} \Gamma^0_{11} + \Gamma^1_{11} \Gamma^1_{11} + \Gamma^1_{21} \Gamma^2_{11} + \Gamma^1_{31} \Gamma^3_{11} \\
&\quad + \Gamma^2_{01} \Gamma^0_{12} + \Gamma^2_{11} \Gamma^1_{12} + \Gamma^2_{21} \Gamma^2_{12} + \Gamma^2_{31} \Gamma^3_{12} \\
&\quad + \Gamma^3_{01} \Gamma^0_{13} + \Gamma^3_{11} \Gamma^1_{13} + \Gamma^3_{21} \Gamma^2_{13} + \Gamma^3_{31} \Gamma^3_{13}) \\
&= \dots - \left[\left(\frac{B'^2}{4B^2} + 0 + 0 + 0 \right) + \left(0 + \frac{A'^2}{4A^2} + 0 + 0 \right) + \left(0 + 0 + \frac{1}{r^2} + 0 \right) + \left(0 + 0 + 0 + \frac{1}{r^2} \right) \right] \\
&= -\frac{B''}{2B} + \frac{B'^2}{4B^2} + \frac{2}{r^2} + \frac{A'B'}{4AB} + \frac{A'^2}{4A^2} + \frac{A'}{rA} - \frac{A'^2}{4A^2} - \frac{2}{r^2} \\
&= -\frac{B''}{2B} + \frac{B'^2}{4B^2} + \frac{A'B'}{4AB} + \frac{A'}{rA}
\end{aligned} \tag{175}$$

With $R_{11} = 0$ we thus get

$$0 = -2rABB'' + rAB'^2 + rBA'B' + 4B^2A' \tag{176}$$

Adding the two equations from R_{00} and R_{11} one gets

$$0 = 4ABB' + 4B^2A' = 4B(AB' + BA') = 4B(AB)' \tag{177}$$

If the derivative of the product AB is 0, the product needs to be a constant,

$$A(r)B(r) = K_1 \tag{178}$$

Knowing $A(r)$, we also know $B(r)$ up to the constant factor K_1 .

So, let's turn to R_{22} :

$$\begin{aligned}
R_{22} &= R_{2\lambda 2}^\lambda = \partial_\beta \Gamma_{22}^\beta - \partial_2 \Gamma_{2\beta}^\beta + \Gamma_{22}^\beta \Gamma_{\beta\sigma}^\sigma - \Gamma_{\sigma 2}^\beta \Gamma_{2\beta}^\sigma \\
&= \partial_0 \Gamma_{22}^0 + \partial_1 \Gamma_{22}^1 + \partial_2 \Gamma_{22}^2 + \partial_3 \Gamma_{22}^3 - \partial_2 (\Gamma_{20}^0 + \Gamma_{21}^1 + \Gamma_{22}^2 + \Gamma_{23}^3) + \dots \\
&= (0 - \partial_r \frac{r}{A(r)} + 0 + 0) - \partial_\theta (0 - \frac{r}{A(r)} + 0 + \cot \theta) + \dots \\
&= -\frac{1}{A(r)} + \frac{r}{A^2(r)} A'(r) + \frac{1}{\sin^2 \theta} + \dots \\
&= -\frac{1}{A} + \frac{rA'}{A^2} + \frac{1}{\sin^2 \theta} + \Gamma_{22}^0 \Gamma_{0\sigma}^\sigma + \Gamma_{22}^1 \Gamma_{1\sigma}^\sigma + \Gamma_{22}^2 \Gamma_{2\sigma}^\sigma + \Gamma_{22}^3 \Gamma_{3\sigma}^\sigma - \dots \\
&= -\frac{1}{A} + \frac{rA'}{A^2} + \frac{1}{\sin^2 \theta} + 0 - \frac{r}{A} (\Gamma_{10}^0 + \Gamma_{11}^1 + \Gamma_{12}^2 + \Gamma_{13}^3) + 0 + 0 - \dots \\
&= -\frac{1}{A} + \frac{rA'}{A^2} + \frac{1}{\sin^2 \theta} - \frac{r}{A} \left(\frac{B'}{2B} + \frac{A'}{2A} + \frac{1}{r} + \frac{1}{r} \right) + 0 + 0 - \dots \\
&= -\frac{3}{A} + \frac{rA'}{2A^2} + \frac{1}{\sin^2 \theta} - \frac{rB'}{2AB} - \dots \\
&= \dots - (\Gamma_{02}^0 \Gamma_{20}^0 + \Gamma_{12}^0 \Gamma_{20}^1 + \Gamma_{22}^0 \Gamma_{20}^2 + \Gamma_{32}^0 \Gamma_{20}^3 \\
&\quad + \Gamma_{02}^1 \Gamma_{20}^1 + \Gamma_{12}^1 \Gamma_{20}^1 + \Gamma_{22}^1 \Gamma_{20}^2 + \Gamma_{32}^1 \Gamma_{20}^3 \\
&\quad + \Gamma_{02}^2 \Gamma_{20}^2 + \Gamma_{12}^2 \Gamma_{20}^2 + \Gamma_{22}^2 \Gamma_{20}^2 + \Gamma_{32}^2 \Gamma_{20}^3 \\
&\quad + \Gamma_{02}^3 \Gamma_{20}^3 + \Gamma_{12}^3 \Gamma_{20}^3 + \Gamma_{22}^3 \Gamma_{20}^3 + \Gamma_{32}^3 \Gamma_{20}^3) \\
&= \dots - [4 \times 0 + (0 + 0 - \frac{1}{A} + 0) + (0 - \frac{1}{A} + 0 + 0) + (0 + 0 + 0 + \cot^2 \theta)] \\
&= -\frac{1}{A} + \frac{rA'}{2A^2} + \frac{1}{\sin^2 \theta} - \frac{rB'}{2AB} - \cot^2 \theta \\
&= -\frac{1}{A} + \frac{rA'}{2A^2} + 1 - \frac{rB'}{2AB} \tag{179}
\end{aligned}$$

With $R_{\theta\theta} = 0$ we thus get

$$0 = -2AB + rBA' + 2A^2B - rAB' \tag{180}$$

Now we can plug in $B = K_1/A$, which implies $B' = -K_1/A^2 A'$.

$$\begin{aligned}
0 &= -2K_1 + r \frac{K_1}{A} A' + rA \frac{K_1}{A^2} A' + 2K_1A \\
0 &= 2K_1 \left(-1 + r \frac{A'}{A} + A \right) \tag{181}
\end{aligned}$$

Which is a differential equation for A :

$$\frac{dA(r)}{dr} = \frac{1}{r} A(r) (1 - A(r)) \tag{182}$$

Separating variables:

$$\frac{dA}{A(1-A)} = \frac{dr}{r} \tag{183}$$

Integrating:

$$\ln A - \ln(1 - A) = \ln r + K_2 \longrightarrow \frac{A}{1 - A} = K_2 \times r \quad (184)$$

with integration constant K_2 . And from this one has

$$A(r) = \frac{1}{1 + \frac{1}{K_2 r}} \quad (185)$$

We have thus

$$ds^2 = -K_1 \left(1 + \frac{1}{K_2 r}\right) c^2 dt^2 + \left(1 + \frac{1}{K_2 r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \quad (186)$$

For $r \longrightarrow \infty$, the metric shall be flat Minkowski space. In that limit $\left(1 + \frac{1}{K_2 r}\right)^{-1} \longrightarrow 1$, and thus $K_1 = 1$. K_2 is determined from the weak field approximation. With g_{00} from equation 106:

$$g_{00} = -1 + h_{00} = -1 - \frac{2\Phi}{c^2} = -\left(1 - \frac{2GM}{rc^2}\right) = -\left(1 - \frac{r_S}{r}\right) \quad (187)$$

we see that $K_2 = -c^2/(2GM)$. In the last equality we have defined the Schwarzschild radius

$$r_S = \frac{2GM}{c^2} \quad (188)$$

The Schwarzschild metric is

$$ds^2 = -\left(1 - \frac{r_S}{r}\right) c^2 dt^2 + \left(1 - \frac{r_S}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \quad (189)$$

For an illustration: If we choose a constant time, and the equatorial plane with $\theta = \pi/2$ (which is any arbitrary plane due to the spherical symmetry), one gets for the hypersurface:

$$ds_{t_0, \text{plane}}^2 = \left(1 - \frac{r_S}{r}\right)^{-1} dr^2 + r^2 d\phi^2 \quad (190)$$

This looks very similar to flat space in cylindrical coordinates

$$ds_{\text{flat}}^2 = dr^2 + dz^2 + r^2 d\phi^2 = \left(1 + \left(\frac{dz}{dr}\right)^2\right) dr^2 + r^2 d\phi^2 \quad (191)$$

If one chooses

$$z(r) = 2\sqrt{r_S(r - r_S)} \quad (192)$$

one can "embed" the hypersurface in flat space, illustrating the curvature of the Schwarzschild metric (figure 6, Flamm's paraboloid).

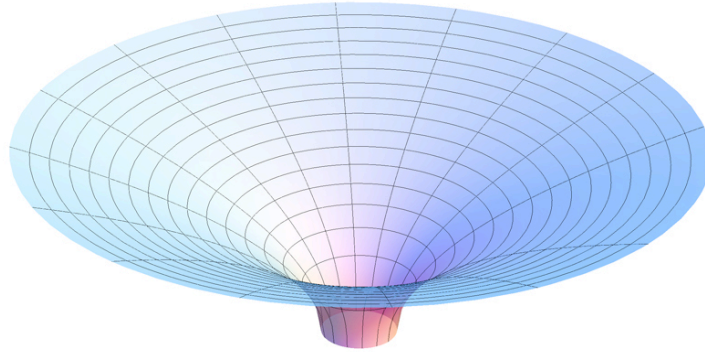


Figure 6: Flamm's paraboloid showing the curvature of the Schwarzschild metric. The 1D-radial function is rotated around the z -axis here for illustration. This is not a gravitational well. The paraboloid does not continue further down, but there is a minimum radius at $r = r_S$, where $z(r) = 2\sqrt{r_S(r - r_S)}$ gets infinitely steep, i.e. $dz/dr = \infty$. Source: Wiki

8.4 The other 13 equations

So far, we have only used R_{00}, R_{11}, R_{22} . There are 13 more field equations.

$$\begin{aligned}
R_{33} &= R_{3\lambda 3}^\lambda = \partial_\beta \Gamma_{33}^\beta - \partial_3 \Gamma_{3\beta}^\beta + \Gamma_{33}^\beta \Gamma_{\beta\sigma}^\sigma - \Gamma_{\sigma 3}^\beta \Gamma_{3\beta}^\sigma \\
&= \partial_0 \Gamma_{33}^0 + \partial_1 \Gamma_{33}^1 + \partial_2 \Gamma_{33}^2 + \partial_3 \Gamma_{33}^3 - \partial_3 (\Gamma_{30}^0 + \Gamma_{31}^1 + \Gamma_{32}^2 + \Gamma_{33}^3) + \dots \\
&= (0 - \partial_r \frac{r \sin^2 \theta}{A(r)} - \partial_\theta (\cos \theta \sin \theta) + 0) - \partial_\phi (0 + 0 + 0 + 0) + \dots \\
&= -\frac{\sin^2 \theta}{A(r)} + \frac{r \sin^2 \theta}{A(r)^2} A'(r) - (\cos^2 \theta - \sin^2 \theta) + \dots \\
&= -\frac{s^2 \theta}{A} + \frac{s^2 \theta r A'}{A^2} - c^2 \theta + s^2 \theta + \Gamma_{33}^0 \Gamma_{0\sigma}^\sigma + \Gamma_{33}^1 \Gamma_{1\sigma}^\sigma + \Gamma_{33}^2 \Gamma_{2\sigma}^\sigma + \Gamma_{33}^3 \Gamma_{3\sigma}^\sigma - \dots \\
&= \frac{s^2 \theta}{A} + \frac{s^2 \theta r A'}{A^2} - c^2 \theta + s^2 \theta + 0 - \frac{s \theta r}{A} \Gamma_{1\sigma}^\sigma - c \theta s \theta \Gamma_{2\sigma}^\sigma + 0 - \dots \\
&= \dots - \frac{s \theta^2 r}{A} (\Gamma_{10}^0 + \Gamma_{11}^1 + \Gamma_{12}^2 + \Gamma_{13}^3) - c \theta s \theta (\Gamma_{20}^0 + \Gamma_{21}^1 + \Gamma_{22}^2 + \Gamma_{23}^3) - \dots \\
&= \dots - \frac{s^2 \theta r B'}{2AB} - \frac{s^2 \theta r A'}{2A^2} - \frac{s^2 \theta}{A} - \frac{s^2 \theta}{A} - c \theta s \theta (0 + 0 + 0 + \cot \theta) - \dots \\
&= \dots - \frac{s^2 \theta r B'}{2AB} - \frac{s^2 \theta r A'}{2A^2} - \frac{2s^2 \theta}{A} - c^2 \theta - \dots \\
&= \dots - (\Gamma_{03}^0 \Gamma_{30}^0 + \Gamma_{13}^0 \Gamma_{30}^1 + \Gamma_{23}^0 \Gamma_{30}^2 + \Gamma_{33}^0 \Gamma_{30}^3 \\
&\quad + \Gamma_{03}^1 \Gamma_{31}^0 + \Gamma_{13}^1 \Gamma_{31}^1 + \Gamma_{23}^1 \Gamma_{31}^2 + \Gamma_{33}^1 \Gamma_{31}^3 \\
&\quad + \Gamma_{03}^2 \Gamma_{32}^0 + \Gamma_{13}^2 \Gamma_{32}^1 + \Gamma_{23}^2 \Gamma_{32}^2 + \Gamma_{33}^2 \Gamma_{32}^3 \\
&\quad + \Gamma_{03}^3 \Gamma_{33}^0 + \Gamma_{13}^3 \Gamma_{33}^1 + \Gamma_{23}^3 \Gamma_{33}^2 + \Gamma_{33}^3 \Gamma_{33}^3) \\
&= \dots - [4 \times 0 + (0 + 0 + 0 - \frac{s^2 \theta}{A}) + (0 + 0 + 0 - c^2 \theta) + (0 - \frac{s^2 \theta}{A} - c^2 \theta) + 0] \\
&= \frac{s^2 \theta}{A} + \frac{s^2 \theta r A'}{A^2} - c^2 \theta + s^2 \theta - \frac{s^2 \theta r B'}{2AB} - \frac{s^2 \theta r A'}{2A^2} - \frac{2s^2 \theta}{A} - c^2 \theta - \frac{s^2 \theta}{A} + \frac{s^2 \theta}{A} + 2c^2 \theta \\
&= \sin^2 \theta \left(-\frac{1}{A} + \frac{r A'}{2A^2} + 1 - \frac{r B'}{2AB} \right) \tag{193}
\end{aligned}$$

This result is exactly the same as for R_{22} , except for the additional factor $\sin^2 \theta$. The $R_{33} = 0$ equation does not yield any new constraints. Next, let's look R_{0i} for $i = 1, 2, 3$ (and due to the symmetry that covers R_{i0} as well, so we cover here 6 equations).

$$\begin{aligned}
R_{0i} = R_{i0} &= R_{0\lambda i}^\lambda = \partial_\beta \Gamma_{0i}^\beta - \partial_0 \Gamma_{i\beta}^\beta + \Gamma_{0i}^\beta \Gamma_{\beta\sigma}^\sigma - \Gamma_{\sigma 0}^\beta \Gamma_{i\beta}^\sigma \\
&= \partial_0 \Gamma_{0i}^0 + \partial_1 \Gamma_{0i}^1 + \partial_2 \Gamma_{0i}^2 + \partial_3 \Gamma_{0i}^3 - \partial_0 (\Gamma_{i0}^0 + \Gamma_{i1}^1 + \Gamma_{i2}^2 + \Gamma_{i3}^3) + \dots \\
&= 0 + 3 \times 0 - 0 + \Gamma_{0i}^0 \Gamma_{0\sigma}^\sigma + \Gamma_{0i}^1 \Gamma_{1\sigma}^\sigma + \Gamma_{0i}^2 \Gamma_{2\sigma}^\sigma + \Gamma_{0i}^3 \Gamma_{3\sigma}^\sigma - \dots \\
&= 4 \times 0 - (\Gamma_{00}^0 \Gamma_{i0}^0 + \Gamma_{10}^0 \Gamma_{i0}^1 + \Gamma_{20}^0 \Gamma_{i0}^2 + \Gamma_{30}^0 \Gamma_{i0}^3 \\
&\quad + \Gamma_{00}^1 \Gamma_{i1}^0 + \Gamma_{10}^1 \Gamma_{i1}^1 + \Gamma_{20}^1 \Gamma_{i1}^2 + \Gamma_{30}^1 \Gamma_{i1}^3 \\
&\quad + \Gamma_{00}^2 \Gamma_{i2}^0 + \Gamma_{10}^2 \Gamma_{i2}^1 + \Gamma_{20}^2 \Gamma_{i2}^2 + \Gamma_{30}^2 \Gamma_{i2}^3 \\
&\quad + \Gamma_{00}^3 \Gamma_{i3}^0 + \Gamma_{10}^3 \Gamma_{i3}^1 + \Gamma_{20}^3 \Gamma_{i3}^2 + \Gamma_{30}^3 \Gamma_{i3}^3) \\
&= 0 - 16 \times 0 = 0 \tag{194}
\end{aligned}$$

Here is also no new information. All R_{0i} and R_{i0} vanish always. So 7 out of 13 equations are done. Hence we are left with R_{12}, R_{13}, R_{23} (and their respective version with indices swapped).

$$\begin{aligned}
R_{12} = R_{21} &= R^\lambda_{1\lambda 2} = \partial_\beta \Gamma^\beta_{12} - \partial_1 \Gamma^\beta_{2\beta} + \Gamma^\beta_{12} \Gamma^\sigma_{\beta\sigma} - \Gamma^\beta_{\sigma 1} \Gamma^\sigma_{2\beta} \\
&= \partial_0 \Gamma^0_{12} + \partial_1 \Gamma^1_{12} + \partial_2 \Gamma^2_{12} + \partial_3 \Gamma^3_{12} - \partial_1 (\Gamma^0_{20} + \Gamma^1_{21} + \Gamma^2_{22} + \Gamma^3_{23}) + \dots \\
&= 0 + 0 + \partial_\theta \frac{1}{r} + 0 - \partial_r (0 + 0 + 0 + \cot \theta) + \dots \\
&= 0 + \Gamma^0_{12} \Gamma^\sigma_{0\sigma} + \Gamma^1_{12} \Gamma^\sigma_{1\sigma} + \Gamma^2_{12} \Gamma^\sigma_{2\sigma} + \Gamma^3_{12} \Gamma^\sigma_{3\sigma} - \dots \\
&= \Gamma^2_{12} (\Gamma^0_{20} + \Gamma^1_{21} + \Gamma^2_{22} + \Gamma^3_{23}) - \dots \\
&= \frac{1}{r} \cot \theta - (\Gamma^0_{01} \Gamma^0_{20} + \Gamma^0_{11} \Gamma^1_{20} + \Gamma^0_{21} \Gamma^2_{20} + \Gamma^0_{31} \Gamma^3_{20} \\
&\quad + \Gamma^1_{01} \Gamma^0_{21} + \Gamma^1_{11} \Gamma^1_{21} + \Gamma^1_{21} \Gamma^2_{21} + \Gamma^1_{31} \Gamma^3_{21} \\
&\quad + \Gamma^2_{01} \Gamma^0_{22} + \Gamma^2_{11} \Gamma^1_{22} + \Gamma^2_{21} \Gamma^2_{22} + \Gamma^2_{31} \Gamma^3_{22} \\
&\quad + \Gamma^3_{01} \Gamma^0_{23} + \Gamma^3_{11} \Gamma^1_{23} + \Gamma^3_{21} \Gamma^2_{23} + \Gamma^3_{31} \Gamma^3_{23}) \\
&= \frac{1}{r} \cot \theta - (4 \times 0 + 4 \times 0 + 4 \times 0 + \frac{1}{r} \cot \theta) = 0
\end{aligned} \tag{195}$$

$$\begin{aligned}
R_{13} = R_{31} &= R^\lambda_{1\lambda 3} = \partial_\beta \Gamma^\beta_{13} - \partial_1 \Gamma^\beta_{3\beta} + \Gamma^\beta_{13} \Gamma^\sigma_{\beta\sigma} - \Gamma^\beta_{\sigma 1} \Gamma^\sigma_{3\beta} \\
&= \partial_0 \Gamma^0_{13} + \partial_1 \Gamma^1_{13} + \partial_2 \Gamma^2_{13} + \partial_3 \Gamma^3_{13} - \partial_1 (\Gamma^0_{30} + \Gamma^1_{31} + \Gamma^2_{32} + \Gamma^3_{33}) + \dots \\
&= 0 + 0 + 0 + \partial_\phi \frac{1}{r} - \partial_r (0 + 0 + 0 + 0) + \dots \\
&= 0 + \Gamma^0_{13} \Gamma^\sigma_{0\sigma} + \Gamma^1_{13} \Gamma^\sigma_{1\sigma} + \Gamma^2_{13} \Gamma^\sigma_{2\sigma} + \Gamma^3_{13} \Gamma^\sigma_{3\sigma} - \dots \\
&= \Gamma^3_{13} (\Gamma^0_{30} + \Gamma^1_{31} + \Gamma^2_{32} + \Gamma^3_{33}) - \dots \\
&= \frac{1}{r} (4 \times 0) - (\Gamma^0_{01} \Gamma^0_{30} + \Gamma^0_{11} \Gamma^1_{30} + \Gamma^0_{21} \Gamma^2_{30} + \Gamma^0_{31} \Gamma^3_{30} \\
&\quad + \Gamma^1_{01} \Gamma^0_{31} + \Gamma^1_{11} \Gamma^1_{31} + \Gamma^1_{21} \Gamma^2_{31} + \Gamma^1_{31} \Gamma^3_{31} \\
&\quad + \Gamma^2_{01} \Gamma^0_{32} + \Gamma^2_{11} \Gamma^1_{32} + \Gamma^2_{21} \Gamma^2_{32} + \Gamma^2_{31} \Gamma^3_{32} \\
&\quad + \Gamma^3_{01} \Gamma^0_{33} + \Gamma^3_{11} \Gamma^1_{33} + \Gamma^3_{21} \Gamma^2_{33} + \Gamma^3_{31} \Gamma^3_{33}) \\
&= 16 \times 0 = 0
\end{aligned} \tag{196}$$

$$\begin{aligned}
R_{23} = R_{32} &= R^\lambda_{2\lambda 3} = \partial_\beta \Gamma^\beta_{23} - \partial_1 \Gamma^\beta_{3\beta} + \Gamma^\beta_{23} \Gamma^\sigma_{\beta\sigma} - \Gamma^\beta_{\sigma 2} \Gamma^\sigma_{3\beta} \\
&= \partial_0 \Gamma^0_{23} + \partial_1 \Gamma^1_{23} + \partial_2 \Gamma^2_{23} + \partial_3 \Gamma^3_{23} - \partial_2 (\Gamma^0_{30} + \Gamma^1_{31} + \Gamma^2_{32} + \Gamma^3_{33}) + \dots \\
&= 0 + 0 + 0 + \partial_\phi \cot \theta - \partial_\theta (4 \times 0) + \dots \\
&= 0 + \Gamma^0_{23} \Gamma^\sigma_{0\sigma} + \Gamma^1_{23} \Gamma^\sigma_{1\sigma} + \Gamma^2_{23} \Gamma^\sigma_{2\sigma} + \Gamma^3_{23} \Gamma^\sigma_{3\sigma} - \dots \\
&= \cot \theta (\Gamma^0_{30} + \Gamma^1_{31} + \Gamma^2_{32} + \Gamma^3_{33}) - \dots \\
&= 0 - (\Gamma^0_{02} \Gamma^0_{30} + \Gamma^0_{12} \Gamma^1_{30} + \Gamma^0_{22} \Gamma^2_{30} + \Gamma^0_{32} \Gamma^3_{30} \\
&\quad + \Gamma^1_{02} \Gamma^0_{31} + \Gamma^1_{12} \Gamma^1_{31} + \Gamma^1_{22} \Gamma^2_{31} + \Gamma^1_{32} \Gamma^3_{31} \\
&\quad + \Gamma^2_{02} \Gamma^0_{32} + \Gamma^2_{12} \Gamma^1_{32} + \Gamma^2_{22} \Gamma^2_{32} + \Gamma^2_{32} \Gamma^3_{32} \\
&\quad + \Gamma^3_{02} \Gamma^0_{33} + \Gamma^3_{12} \Gamma^1_{33} + \Gamma^3_{22} \Gamma^2_{33} + \Gamma^3_{32} \Gamma^3_{33}) \\
&= 16 \times 0 = 0
\end{aligned} \tag{197}$$

The 13 additional equations do not yield any new constraints, such that the solution derived indeed solves the full field equations.

8.5 The Christoffel symbols of the Schwarzschild metric

Using $A(r) = (1 - r_S/r)^{-1}$ and $B(r) = r_S/r - 1$ we can rewrite the Christoffel symbols

$$\begin{aligned}
 \Gamma^0_{10} = \Gamma^0_{10} &= \frac{1}{2r} \frac{r_S}{r-r_S} & \Gamma^1_{00} &= \frac{r_S(r-r_S)}{2r^3} & \Gamma^2_{12} = \Gamma^2_{21} &= \frac{1}{r} & \Gamma^3_{13} = \Gamma^3_{31} &= \frac{1}{r} \\
 \Gamma^1_{11} &= -\frac{1}{2r} \frac{r_S}{r-r_S} & \Gamma^2_{33} &= -\cos\theta \sin\theta & \Gamma^3_{23} = \Gamma^3_{32} &= \cot\theta & & \\
 \Gamma^1_{22} &= r_S - r & & & & & & \\
 \Gamma^1_{33} &= (r_S - r) \sin^2\theta & & & & & &
 \end{aligned} \tag{198}$$

9 Observing a black hole - the Galactic Center

9.1 Atmospheric turbulence

9.1.1 Kolmogorov model

Incoming stellar light is to an extremely good degree a plane electromagnetic wave - outside of Earth's atmosphere. The atmosphere is turbulent, and has thus temperature fluctuations, which lead to refractive index variations, which in turn affect the phases of the initially plane waves (as the local speed of light varies). In the following, we will derive the basics for the wavefront deviations imposed onto stellar light by the turbulence. The dimensionless Reynolds number \mathcal{R} governs, whether a flow of a fluid/gas is turbulent or laminar.

$$\mathcal{R} = LV/\nu \quad (199)$$

where L is a length scale, V the fluid velocity and ν the kinematic viscosity (measured in m^2/s). For air $\nu = 1.5 \times 10^{-5} \text{ m}^2/\text{s}$, air speeds are typically a few m/s , and hence for L in the range of meters to kilometers, $\mathcal{R} \approx 10^6$, so air is mostly turbulent. Turbulence is generated on large scales by macroscopic perturbations of the flow, creating turbulence cells, so-called 'eddies' that cascade down into smaller and smaller ones, down to the regime where the \mathcal{R} gets small enough such the flow can become laminar. So there is an outer scale L_0 for turbulence in air that is (probably) around a few hundred meters, and an inner scale l_0 in the mm-range. In the range between L_0 and l_0 one can describe the turbulence strength as a function of eddie size with a simple model, which is named after Kolmogorov. Beyond the viscosity, the only other parameter it has is the energy creation rate per unit mass $\epsilon \sim E/(m\tau) \sim V^2/\tau$. The unit is $\text{J}/(\text{kg s}) = \text{m}^2/\text{s}^{-3}$. This energy is injected at the outer scale L_0 and then is given to smaller and smaller eddies until one reaches l_0 .

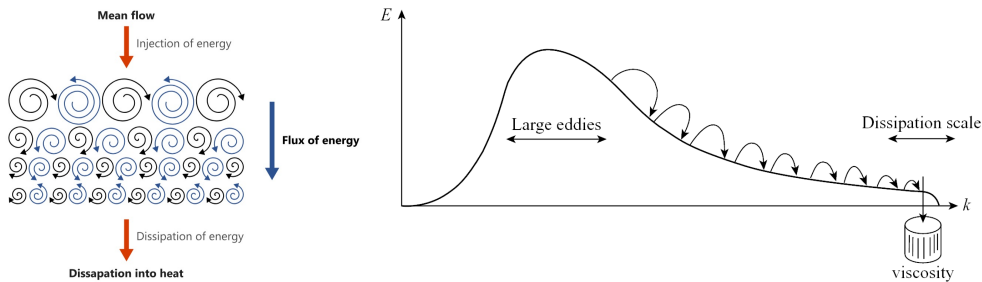


Figure 7: Turbulence cascade

Turbulence can be described by a structure function: The variation of some measurable A between two points \vec{r}_1 and \vec{r}_2 is

$$\begin{aligned} D_A(\vec{r}_1, \vec{r}_2) &= \langle |A(\vec{r}_1) - A(\vec{r}_2)|^2 \rangle \\ D_A(\vec{r}) := D_A(\vec{r}, \vec{0}) &= \langle |A(\vec{r}) - A(\vec{0})|^2 \rangle \end{aligned} \quad (200)$$

(with $\vec{r} = \vec{r}_1 - \vec{r}_2$). The structure function of a turbulent velocity field can only depend on $|\vec{r}_1 - \vec{r}_2| = |\vec{r}| =: r$. The unit of D_V is m^2/s^2 , so one can write

$$D_V(r) = \alpha f(r/\beta) \quad (201)$$

where f is a dimensionless function of a dimensionless argument, and $[\alpha] = \text{m}^2/\text{s}^2$ and $[\beta] = \text{m}$. Both parameters α and β will be functions of ν and ϵ . By dimensional analysis there is (up to constants) only one way to achieve that:

$$\alpha \propto \nu^{1/2}\epsilon^{1/2} \quad \text{and} \quad \beta \propto \nu^{3/4}\epsilon^{-1/4} \quad (202)$$

In the range between L_0 and l_0 this shall be independent of viscosity (which only sets in at l_0). Hence, one needs to have

$$\begin{aligned} D_V(r) &\propto \alpha(r/\beta)^{2/3} \\ &\propto \nu^{1/2}\epsilon^{1/2}r^{2/3}(\nu^{-3/4}\epsilon^{1/4})^{2/3} \\ &\propto \epsilon^{2/3}r^{2/3} \\ &=: C_V^2 r^{2/3} \end{aligned} \quad (203)$$

The result is simple: The turbulence spectrum as a function of length only has one free parameter, C_V . The gas motions move around air parcels, which are in pressure equilibrium, but differ in temperature and density. This induces refractive index variations. With $N := n - 1 \propto \rho$, N will thus follow the same turbulence spectrum

$$D_N(\vec{r}) = C_N^2 |\vec{r}|^{2/3} \quad (204)$$

The refractive index of air is empirically

$$\begin{aligned} N \times 10^6 = (n - 1) \times 10^6 &= 77.6 \left(1 + \frac{7.52 \times 10^{-3}}{\lambda/\mu\text{m}} \right) \left(\frac{P/\text{mbar}}{T/\text{K}} \right) \\ \delta N = \delta n &= -7.76 \times 10^{-5} \left(1 + \frac{7.52 \times 10^{-3}}{\lambda/\mu\text{m}} \right) \left(\frac{P/\text{mbar}}{(T/\text{K})^2} \right) \delta T/\text{K} \end{aligned} \quad (205)$$

There is no δP as pressure differences are quickly removed, as long as the velocities are subsonic. This is thus 'isobaric'.

9.1.2 Effect on wavefronts

The following lengthy calculation retrieves the effect on a plane wave of the density fluctuations. At some point in time, one has for the electric field E with phase Φ :

$$E \propto \Psi(\vec{x}) = \exp(i\Phi(\vec{x})) \quad (206)$$

The wave number be $k = 2\pi/\lambda$. The phase shift from going through one layer of atmosphere of thickness δh at height h will be at $(x, y) =: \mathbf{x}$ (bold symbols indicate 2D vectors) given by:

$$\Phi|_h(x, y) = \Phi(\mathbf{x}) = k \int_h^{h+\delta h} N(\mathbf{x}, z') dz' \quad (207)$$

Note that this relates the 3D refractive index variations with the 2D phase error of the wavefront. The coherence function $B_{\Psi, h}(\mathbf{x})$ of a wavefront is the auto-correlation of $\Psi|_h(\mathbf{x})$.

$$\begin{aligned} B_{\Psi, h}(\mathbf{x}) &= \int \Psi|_h(\mathbf{r} + \mathbf{x}) \Psi|_h^*(\mathbf{r}) d^2r = \langle \Psi|_h(\mathbf{r} + \mathbf{x}) \Psi|_h^*(\mathbf{r}) \rangle \\ &= \langle \exp(i(\Phi(\mathbf{r} + \mathbf{x}) - \Phi(\mathbf{r}))) \rangle \\ &= \exp\left(-\frac{1}{2} \langle |\Phi(\mathbf{r}) - \Phi(\mathbf{r} + \mathbf{x})|^2 \rangle\right) \\ &= \exp\left(-\frac{1}{2} D_{\Phi, h}(\mathbf{x})\right) \end{aligned} \quad (208)$$

The key here was the third equal sign, which uses a relation valid for Gaussian signals P_G :

$$\begin{aligned}
P_G(x) &= \frac{1}{2\pi\sqrt{\sigma}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \\
\langle \exp(\alpha x) \rangle &= \int \exp(\alpha x) P_G(x) dx \\
&= \frac{1}{2\pi\sqrt{\sigma}} \int \exp\left(\alpha x - \frac{x^2}{2\sigma^2}\right) dx \\
&= \frac{1}{2\pi\sqrt{\sigma}} \int \left(-\frac{(x - \alpha\sigma^2)^2}{2\sigma^2} + \frac{\alpha^2\sigma^4}{2\sigma^2}\right) dx \\
&= \exp\left(\frac{1}{2}\alpha^2\sigma^2\right) \frac{1}{2\pi\sqrt{\sigma}} \int \left(-\frac{(x - \alpha\sigma^2)^2}{2\sigma^2}\right) dx \\
&= \exp\left(\frac{1}{2}\alpha^2\sigma^2\right) \frac{1}{2\pi\sqrt{\sigma}} \int \left(-\frac{\tilde{x}^2}{2\sigma^2}\right) d\tilde{x} \\
&= \exp\left(\frac{1}{2}\alpha^2\sigma^2\right) = \exp\left(\frac{1}{2}\alpha^2 \langle x^2 \rangle\right)
\end{aligned} \tag{209}$$

Next, let's look at the structure function $D_{\Phi,h}(\mathbf{x})$.

$$\begin{aligned}
D_{\Phi,h}(\mathbf{x}) &= \langle |\Phi(\mathbf{x}) - \Phi(\mathbf{0})|^2 \rangle \\
&= \int (\Phi(\mathbf{x} + \mathbf{r}) - \Phi(\mathbf{r})) \times (\Phi(\mathbf{x} + \mathbf{r}) - \Phi(\mathbf{r}))^* d^2r \\
&= \int (\Phi(\mathbf{x} + \mathbf{r})\Phi^*(\mathbf{x} + \mathbf{r}) + \Phi(\mathbf{r})\Phi^*(\mathbf{r}) - \Phi(\mathbf{x} + \mathbf{r})\Phi^*(\mathbf{r}) - \Phi(\mathbf{r})\Phi^*(\mathbf{x} + \mathbf{r})) d^2r \\
&= \int \Phi(\mathbf{r}')\Phi^*(\mathbf{r}') d^2r' + \int (\Phi(\mathbf{r})\Phi^*(\mathbf{r}) - \Phi(\mathbf{x} + \mathbf{r})\Phi^*(\mathbf{r}) - \Phi(\mathbf{r})\Phi^*(\mathbf{x} + \mathbf{r})) d^2r \\
&= \langle \Phi(\mathbf{r}' + \mathbf{0})\Phi^*(\mathbf{r}') \rangle + \langle \Phi(\mathbf{r} + \mathbf{0})\Phi^*(\mathbf{r}) \rangle - \langle \Phi(\mathbf{r} + \mathbf{x})\Phi^*(\mathbf{r}) \rangle - \langle \Phi^*(\mathbf{r} + \mathbf{x})\Phi(\mathbf{r}) \rangle \\
&= 2B_{\Phi,h}(\mathbf{0}) - B_{\Phi,h}(\mathbf{x}) - B_{\Phi,h}^*(\mathbf{x}) \\
&= 2(B_{\Phi,h}(\mathbf{0}) - B_{\Phi,h}(\mathbf{x}))
\end{aligned} \tag{210}$$

(The last equality is because $\Phi(\mathbf{x})$ is a real function.) So we need $B_{\Phi,h}$:

$$\begin{aligned}
B_{\Phi,h}(\mathbf{x}) &= \langle \Phi(\mathbf{x})\Phi(\mathbf{x} + \mathbf{r}) \rangle = \left\langle k \int_h^{h+\delta h} N(\mathbf{x}, z') dz' k \int_h^{h+\delta h} N(\mathbf{x} + \mathbf{r}, z'') dz'' \right\rangle \\
&= k^2 \int_h^{h+\delta h} \int_h^{h+\delta h} \langle N(\mathbf{x}, z') N(\mathbf{x} + \mathbf{r}, z'') \rangle dz' dz'' \\
&= k^2 \int_h^{h+\delta h} \int_{h-z'}^{h-z'+\delta h} \langle N(\mathbf{x}, z') N(\mathbf{x} + \mathbf{r}, z' + z) \rangle dz' dz \\
&= k^2 \int_h^{h+\delta h} dz' \int_{h-z'}^{h-z'+\delta h} B_{N,h}(\mathbf{x}, z) dz \\
&= k^2 \delta h \int_{-\infty}^{\infty} B_{N,h}(\mathbf{x}, z) dz
\end{aligned} \tag{211}$$

The second equality is just the definition of Φ , the fourth changes the integration variable $z := z'' - z'$, the fifth uses the covariance $B_{N,h}$, and the last assumes that the turbulence scale is smaller than the layer one looks at. So we get

$$\begin{aligned}
D_{\Phi,h}(\mathbf{r}) &= 2(B_{\Phi,h}(\mathbf{0}) - B_{\Phi,h}(\mathbf{r})) \\
&= 2k^2\delta h \int_{-\infty}^{\infty} (B_{N,h}(\mathbf{0}, z) - B_{N,h}(\mathbf{r}, z)) \\
&= 2k^2\delta h \int_{-\infty}^{\infty} (B_{N,h}(\mathbf{0}, 0) - B_{N,h}(\mathbf{r}, z)) - (B_{N,h}(\mathbf{0}, 0) - B_{N,h}(\mathbf{0}, z)) \\
&= k^2\delta h \int_{-\infty}^{\infty} (D_{N,h}(\mathbf{r}, z) - D_{N,h}(\mathbf{0}, z)) \\
&= k^2\delta h C_N^2 \int_{-\infty}^{\infty} (|\mathbf{r}|^2 + z^2)^{1/3} - z^{2/3} dz \\
&= k^2\delta h C_N^2 \frac{2\Gamma(1/2)\Gamma(1/6)}{5\Gamma(2/3)} |\mathbf{r}|^{5/3} = 2.914 k^2\delta h C_N^2 |\mathbf{r}|^{5/3}
\end{aligned} \tag{212}$$

With that we get

$$B_{\Psi,h}(\mathbf{r}) = \exp \left[-\frac{1}{2} \left(2.914 k^2\delta h C_N^2 |\mathbf{r}|^{5/3} \right) \right] \tag{213}$$

For the whole atmosphere (multiplying the coherences for each layer with each other) we get thus

$$B_{\Psi}(\mathbf{r}) = \exp \left[-\frac{1}{2} \left(2.914 k^2 |\mathbf{r}|^{5/3} \int_0^{\infty} C_N^2(h) dh \right) \right] \tag{214}$$

One introduces the Fried parameter r_0 as

$$r_0 = \left[0.423 k^2 (\sec Z) \int_0^{\infty} C_N^2(h) dh \right]^{-3/5} \tag{215}$$

(with the $\sec Z$ term taking into account the zenith angle Z) such that one can write

$$B_{\Psi}(\mathbf{r}) = \exp \left[-3.44 \left(\frac{|\mathbf{r}|}{r_0} \right)^{5/3} \right] \text{ and } D_{\Phi}(\mathbf{r}) = 6.88 \left(\frac{|\mathbf{r}|}{r_0} \right)^{5/3} \tag{216}$$

This is a remarkable simple result - the spatial coherence of wavefronts can be described by one length parameter, r_0 . The only unknown in it is the vertical turbulence profile function $C_N(h)$. Surprisingly, there are even devices that can measure this profile (figure 8), they are called "scidars".

9.1.3 The meaning of the Fried parameter

It turns out (beyond the scope of this lecture) that r_0 has very profound meaning for working with telescopes in Earth's atmosphere.

- The mean square wavefront error across a telescope with aperture D is

$$\sigma_{\Phi,D}^2 = 1.03 \left(\frac{D}{r_0} \right)^{5/3} \tag{217}$$

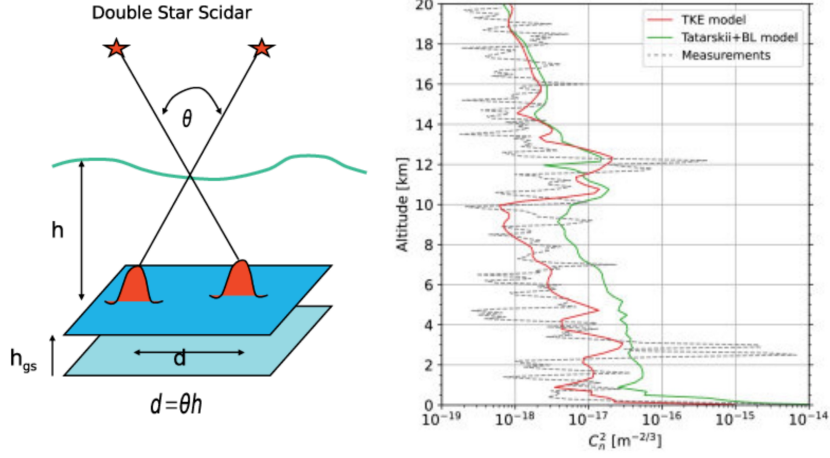


Figure 8: Left: Measurement principle of a scidar. Right: Example of a vertical C_N profile measured at the Paranal observatory

- Therefore, the effective aperture of a telescope in terms of its resolving power is set by r_0 , i.e. the point spread function is $\sim \lambda/r_0$. For $D > r_0$ one would thus want to use adaptive optics. Or in other words: The seeing limits the sharpness of a simple telescope to one that has $D = r_0$.
- The Fried parameters scales with wavelength as $r_0 \propto (k^2)^{-3/5} \propto \lambda^{6/5}$ (equation 215 with $k = 2\pi/\lambda$). Therefore, diffraction-limited operation is easier to achieve at longer wavelengths.
- During observing, the seeing is estimated by the size of seeing disk of a star. Its angular size is

$$\alpha = \frac{\lambda}{r_0} \propto \lambda^{-1/5} \quad (218)$$

This value is usually called 'the seeing'.

- With larger zenith angles, the light passes through more atmosphere, and equation 215 captures the resulting worse seeing by the larger value for r_0 then.
- One can think of the Fried parameter as the typical size of a turbulence cell in the atmosphere.
- In a simplified picture, the spatially turbulent structure is moved by some windspeed V across the telescope aperture. This leads to the definition of the coherence time as $\tau_0 = r_0/V$.
- Consider two rays from slightly different directions (two stars, like a binary), different by a (small) angle θ . Then $|\mathbf{r}| = \theta h \sec Z$. Putting this in D_Φ and integrating over the atmosphere we get

$$\begin{aligned} \sigma_{\Phi, \theta}^2 &= \int_0^\infty \sec Z D_\Phi dh = \int_0^\infty \sec Z 2.914 k^2 C_N^2(h) (\theta h \sec Z)^{5/3} dh \\ &= 2.914 k^2 (\sec Z)^{8/3} \theta^{5/3} \int_0^\infty C_N^2(h) h^{5/3} dh \\ &= \left(\frac{\theta}{\theta_0} \right)^{5/3} \end{aligned} \quad (219)$$

with the so-called isoplanatic angle

$$\theta_0 = \left[2.914 k^2 (\sec Z)^{8/3} \int_0^\infty C_N^2(h) h^{5/3} dh \right]^{-3/5} \quad (220)$$

This is similar to the definition of r_0 , and indeed

$$\theta_0 = 0.314 \frac{r_0}{H} \cos Z \quad (221)$$

with the effective height of the turbulence

$$H = \left(\frac{\int_0^\infty C_N^2(h) h^{5/3} dh}{\int_0^\infty C_N^2(h) dh} \right)^{3/5} \quad (222)$$

So, one can think of the turbulence effective occurring at this height H . The function $C_N(h)$ acts like a weight for the different heights, and again only the vertical turbulence profile $C_N(h)$ enters. The meaning of θ_0 is that at that angle $\sigma_{\Phi, \theta}^2 = 1 \text{ rad}^2$, i.e. the (instantaneous) point spread function will differ at that angle distance between the two stars.

- If one wants to correct the wavefront, one needs roughly as many sensing and correcting elements as there are turbulence cells across the aperture, i.e. D^2/r_0^2 . Larger telescopes and shorter wavelengths thus need more such elements.
- Typical values for r_0 at good astronomical sites are around 20 cm at $0.5 \mu\text{m}$ wavelength. The seeing is typically $1''$, and the isoplanatic angle a few arcseconds.
- The so-called Strehl ratio Sr (ratio of peak of PSF to peak of PSF if the image would be diffraction-limited) can be approximated by (figure 9)

$$Sr = e^{-\sigma_{\Phi, D}^2} \quad (223)$$

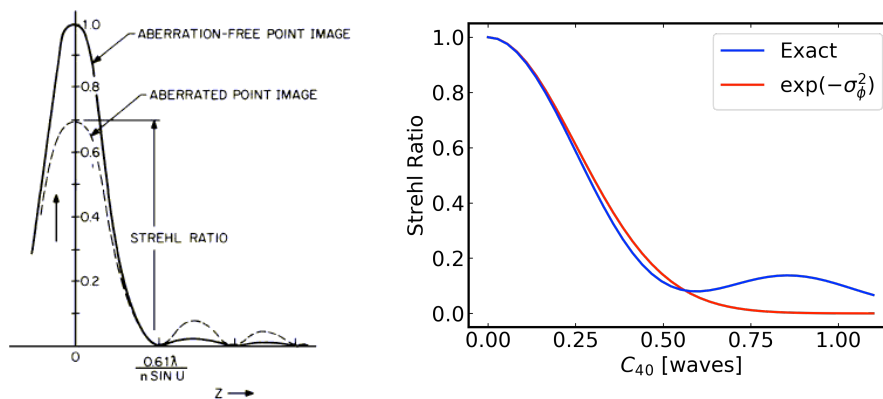


Figure 9: Left: Definition of Strehl ratio. Right: Approximation of the Strehl ratio via equation 223 for different degrees of defocus.

9.2 Designing an AO system

The idea of adaptive optics is to undo the phase modulations, the atmosphere imposed before. For that, one needs to sample the wavefront. One can think of decomposing the wavefront into at least D^2/r_0^2 pieces, which can then be described by a constant phase. Let's get some numbers for an AO working in the infrared H-band.

- At $\lambda = 1.5 \mu\text{m}$, we have $r_0 = 75 \text{ cm}$.
- For the VLT ($D = 8 \text{ m}$) we need thus at least around 100 elements, i.e. 10×10 . In practice, oversampling by a small factor in 1D leads to number of elements around 500 – 1000.
- For typical wind speeds of 200 km/h in the upper atmosphere, the coherence time is around 10 msec. This will be our sensor exposure time, and our loop needs to run at 100 Hz. Actual systems typically run a bit faster still, for example at 300 Hz.
- Each sensor elements will collect light from an area of around r_0^2 . The K-band flux density of a star with $m_K = 0$ is $F_\lambda = 4 \times 10^{-10} \text{ W/m}^2/\mu\text{m}$, the K-band has a width of $0.4 \mu\text{m}$. Hence, for such a star one receives $P = 6 \times 10^{-11} \text{ W}$ per sensor (assuming we have 500), or around 10^7 photons ($E = h\nu = hc/\lambda$) per coherence time. Let's detect that over around $5 \times 5 = 25$ pixels, so we have 2×10^5 photons per pixel. If we have detector with 5 electrons read noise, and an efficiency (transmission, quantum efficiency, ...) of total 50%, the SNR is dominated by the photo-electron noise ($\sqrt{10^5} = 300$). If we go to fainter stars, let's say $m_K = 10$, we have 10^4 less photons - a mere 20 photons, 10 photo-electrons, and hence now the read-noise of 5 electrons will dominate. The SNR is then only marginal. The numbers here are very rough, but illustrate that the requirements AO sets for the guide star are a true worry. Also, detector performance (i.e. noise properties) obviously matters.
- The isoplanatic angle determines for which field of view the AO correction will be valid. With $\theta_0 = 6.5''$ we thus expect around $13''$ corrected field-of-view.
- If one measures a wavefront at ≈ 500 positions, one should also correct it at around the same number of positions, i.e. have similar degrees of freedom. One can think of the problem of measuring 500 values and calculating from them 500 values. Inbetween the two vectors is the so-called control matrix, which has thus 25000 elements. For applying this matrix, one needs $500 \times (500 \text{ multiplications and } 500 \text{ sums})$, i.e. around 5×10^5 operations. At 300 Hz this are 150 million operations per second, or 150 MHz. Nowadays, computers are easily fast enough, but in the 1990s when AO was realized first, the speed of the real-time computer was an important parameter to take into account.
- In many cases, one actually senses the stellar light at visible wavelengths (where exquisite, fast detectors are available), but corrects for the infrared bands.

9.3 Elements of an AO system

We need three basic elements (figure 10):

- a wavefront sensor, detecting light from some star in the field
- a deformable mirror, being able to move quickly enough and far enough to compensate the phase fluctuations.
- a real-time system, running a control loop between the sensor and the mirror.

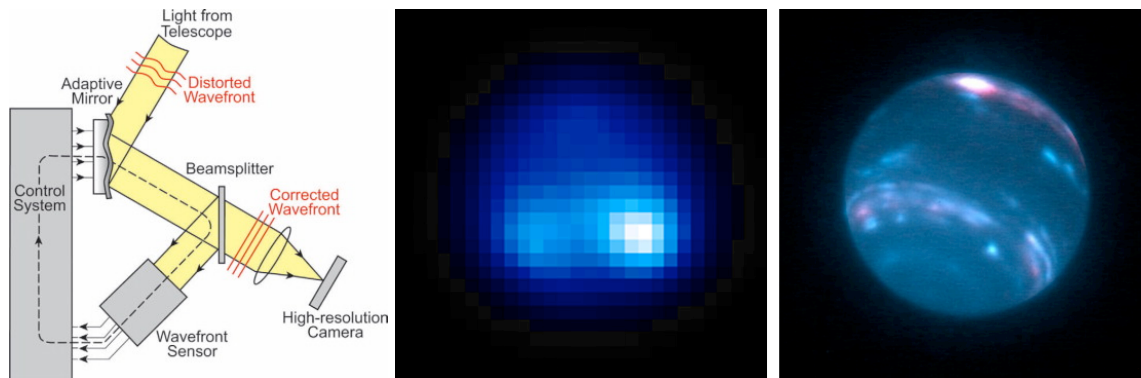


Figure 10: Left: Basic elements of an AO system. Right: An image of Neptune without (left) and with AO (right), obtained at the 10m Keck telescope. Source: Wizinowich

9.3.1 Wavefront sensors

Measuring a wavefront means that one needs to find out, how the phase changes as a function of spatial coordinate. Usually, that requires some interferometric means. However, the derivative, the gradient of the wavefront can be measured by a lens: The image displacement corresponds to the slope of the wavefront. Hence, one can sub-divide the aperture into many small parts, each of which gets lensed by its own small lens, and retrieves in that way the gradient of the wavefront (figure 11). This type of sensor is called Shack-Hartmann wavefront sensor, and it is in widespread use.

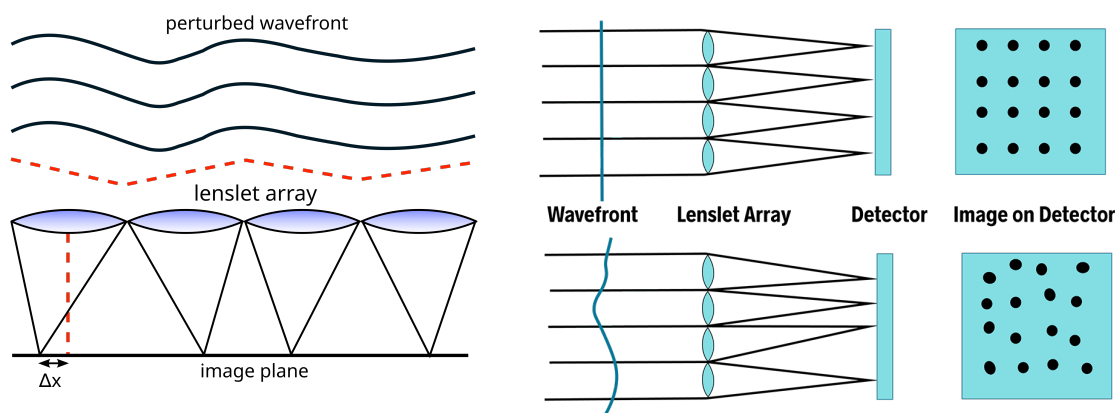


Figure 11: Working principle of a Shack-Hartmann wavefront sensor. Sources: English Wikipedia, A. Tokovinin

The problem of measuring the wavefront is thus converted into one of determining centroid positions of spots on a detector. One can even work with a "quad-cell" for that (figure 12). One can the estimate the position by

comparing the intensities in the four quadrants:

$$\begin{aligned} \Delta X &= r \frac{(I_{BR} + I_{TR}) - (I_{BL} + I_{TL})}{I_{BR} + I_{TR} + I_{BL} + I_{TL}} \\ \Delta Y &= r \frac{(I_{TL} + I_{TR}) - (I_{BL} + I_{TBR})}{I_{BR} + I_{TR} + I_{BL} + I_{TL}} \end{aligned} \tag{224}$$

A very clever way of implementing quad-cells is the so-called pyramid wavefront sensor (figure 13 right). A

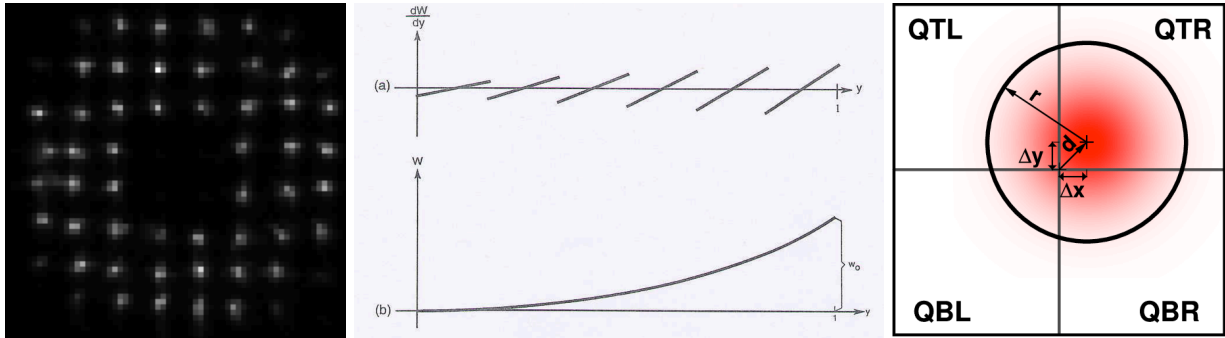


Figure 12: Left: A low-order Shack-Hartmann sensor in operation. Middle: Reconstruction of the wavefront from its slopes. Right: A quad-cell detector and the image of a point source. Sources: AO course Claire Max, Esper-Chain 2016

pyramidal shaped piece of glass splits the stellar light in four (i.e. the star is focussed on the tip of the pyramid). The intensities between the four images are then like in the quad-cell. Each quadruple of corresponding pixels is thus equivalent to one sub-aperture of a Shack-Hartmann sensor in quad-cell mode (figure 13 left). Another type

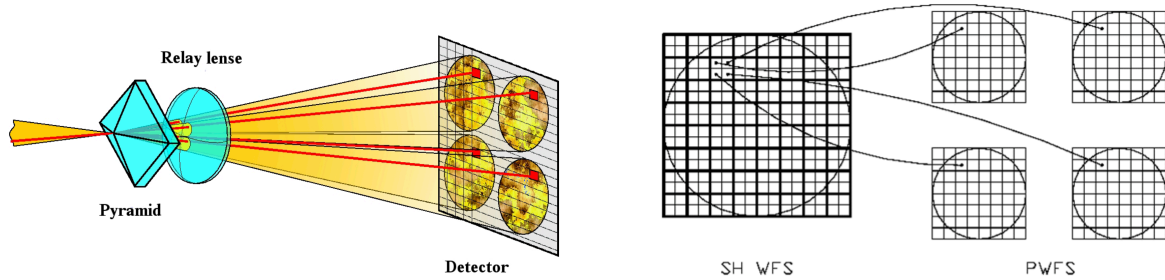


Figure 13: Left: Optical setup for a pyramid wavefront sensor. Right: Correspondence between Shack-hartmann sensor and pyramid sensor. Sources: AO course Claire Max, Peter 2008

of wavefront sensor is the curvature sensor (figure 14). For that, one needs to detect image in front of and behind a focal plane. The brightness difference between the two planes then measures actually the curvature of the wavefront - and hence one can again reconstruct its form. Practically, one moves a mirror back and forth at kHz frequencies around a focal point and records the images, with a suitably fast detector. Often, photo diodes have been employed for that.

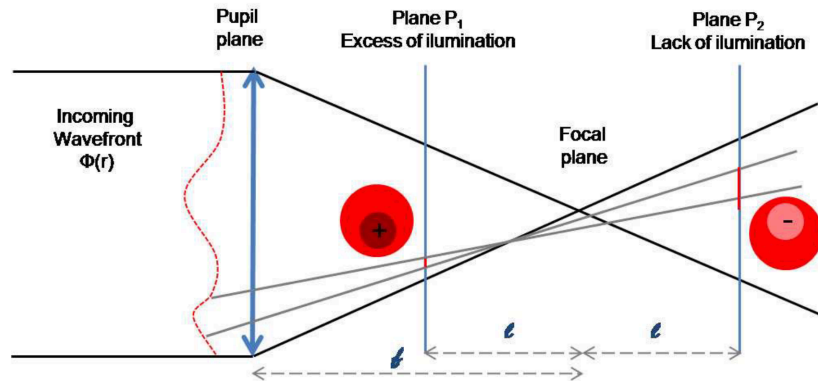


Figure 14: Working principle of a curvature wavefront sensor. Source: Aller-Carpentier 2011

9.3.2 Deformable mirrors

The correcting element is a mirror, the surface shape of which needs to be adapted quickly and constantly. There are two basic types, namely segmented mirrors and continuous mirrors (figure 15, left). In both cases, a large number of actuators sits behind the reflective surface and they are commanded such that the shape of the mirror flattens the incoming wavefronts, creating thus sharp, diffraction-limited images. The actuators need to move by amounts similar to the wavelength, i.e. microns, within the time constant of the control loop, i.e. milliseconds. Hence, Piezo actuators are often chosen. The reflective surface needs to be thin and elastic enough, to allow for

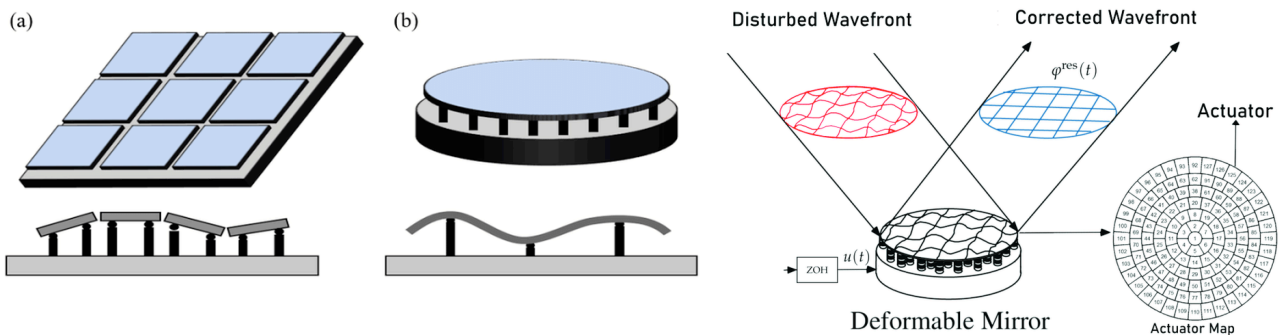


Figure 15: Basic types of deformable mirrors: Segmented mirrors and continuous mirrors

and follow the commanded motions. This leads to quite challenging designs, for example the adaptive secondary mirror of the VLT UT4 is 1.12m across, but the glass shell is only 2 mm thin (figure 16).

9.3.3 Control system

The main task of the control system is to receive the information from the wavefront sensor, calculate the required response, and send the corresponding actuator commands to the deformable mirror. Conceptionally, one can decompose any wavefront into a set of polynomials. For circular apertures, the so-called Zernike polynomials are

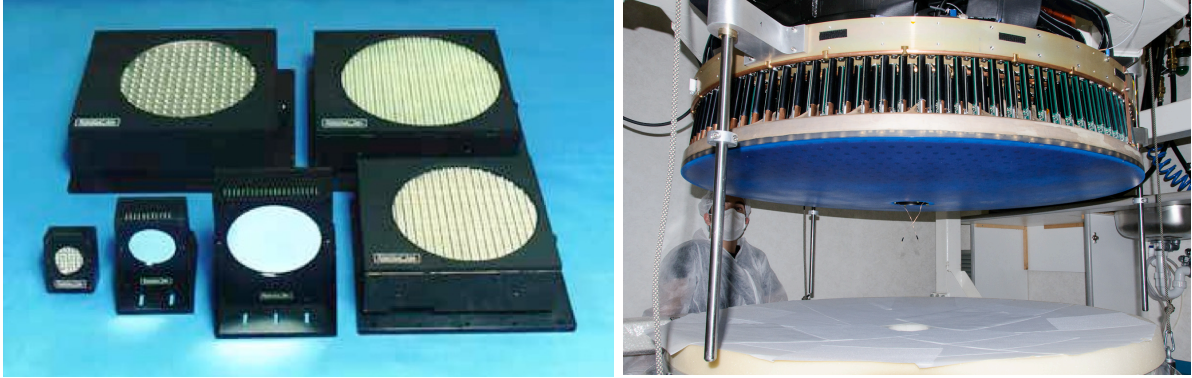


Figure 16: Left: Some small deformable mirrors. Right: The adaptive secondary of UT4 of the VLT, measuring 1m across Source: ESO/Microgate

particularly useful (figure 17). They are defined as

$$\begin{aligned}
 Z_n^m(\rho, \phi) &= R_n^m(\rho) \cos(m\phi) \quad \text{for } n \geq m \geq 0 \\
 Z_n^{-m}(\rho, \phi) &= R_n^m(\rho) \sin(m\phi) \quad \text{for } n \geq m \geq 1 \\
 R_n^m(\rho) &= \sum_{k=0}^{(n-m)/2} \frac{(-1)^k (n-k)!}{k!((n+m)/2-k)!((n-m)/2-k)!} \rho^{n-2k} \quad \text{for } n \geq m, n-m \text{ even} \\
 R_n^m(\rho) &= 0 \quad \text{for } n-m \text{ odd}
 \end{aligned} \tag{225}$$

This yields the following set of radial functions:

$$\begin{aligned}
 1 &= R_0^0(\rho) && \text{"piston"} \\
 \rho &= R_1^1(\rho) && \text{"tip/tilt"} \\
 2\rho^2 - 1 &= R_2^0(\rho) && \text{"astigmatism"} \\
 \rho^2 &= R_2^2(\rho) && \text{"focus"} \\
 3\rho^3 - 2\rho &= R_3^1(\rho) && \text{"trifoil"} \\
 \rho^3 &= R_3^3(\rho) && \text{"coma"}
 \end{aligned} \tag{226}$$

The function set is orthogonal and thus a good choice as a basis. While it is theoretically appealing to decompose wavefronts, it turns out that for real-time control it is more efficient to omit the step of internal representation, and directly calibrate input signals and output signals. This yields the so-called interaction matrix, showing how much each actuator needs to be moved for each sensor. It is calibrated experimentally.

9.4 Advanced concepts

The AO system described so-far is an "SCAO" system, for "single-conjugated adaptive optics". It looks at one natural guide star. Many more ideas have been developed:

- The need for a rather bright guide star within the vicinity is a true limitation in practice. It has led to development of laser guide stars, that create artificial stars in the atmosphere, for example by exciting sodium

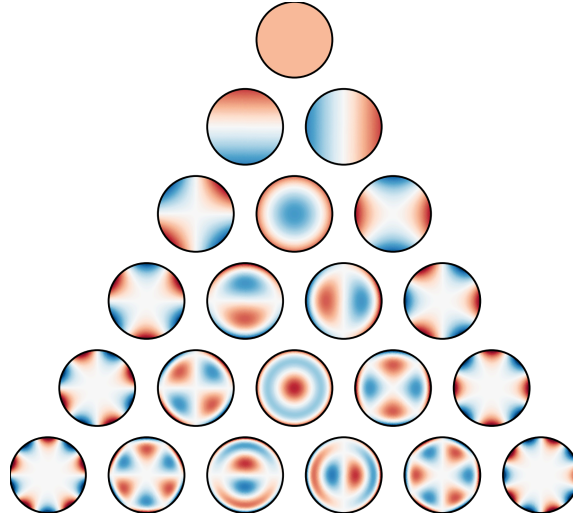


Figure 17: Graphical illustration of the lowest order Zernike polynomials. Source: Wiki

atoms in a natural layer at around 100 km height. The tip/tilt correction cannot be obtained from a laser star (as in geometric optics light paths up and down are identical), only higher order are sensed in this way. Hence, the need for a natural star for tip/tilt correction remains - but it can be much fainter, as only two Zernike orders need to be determined from it.

- A laser guide star shines down in a cone - not covering the full cylinder of atmosphere the telescope looks through. Thus, multiple laser stars can be used to mitigate that.
- Also, multiple natural guide stars can be employed. This is often a trade-off between how sharp the images shall be, and how much field of view one corrects.
- Typical C_N profiles show often two dominant layers: A ground layer, and one at the jet stream height. Hence, one can conjugate multiple deformable mirrors to each layer one wants to address.

9.5 Observing the Galactic Center

Since the discovery of quasars, it was suspected that the central engines is accretion onto massive black holes. This idea was based on two facts: The quick variability within days, coupled with a luminosity that can be equal to the whole host galaxy. For the first quasars, the host galaxy was not even directed, as the central source outshined the the stellar light, and high-resolution techniques were not yet available. Still, the enormous energy output was obvious, since the the quasar spectra contained the well-known hydrogen lines, but redshifted to a degree unseen before. Together with Hubble's law, this implied a large distance, and consequently the extreme luminosities. The suspicion was then that actually almost all galaxies contain central black holes, with very different levels of accretion. If so, the closest massive black hole would reside only ≈ 8 kpc away in the Milky Way center. Observing the latter is hampered by the presence of interstellar dust, preventing the use of optical light. But radio waves and X-ray radiation would pass the extinction screen. Unfortunately, in both bands, one does not see regular stars. But it turns out that at infrared wavelengths, there is a sweet spot for observing the Galactic Center (figure 18), at around $\lambda = 2 \mu\text{m}$. Luckily, this is also the regime, in which adaptive optics is feasible.

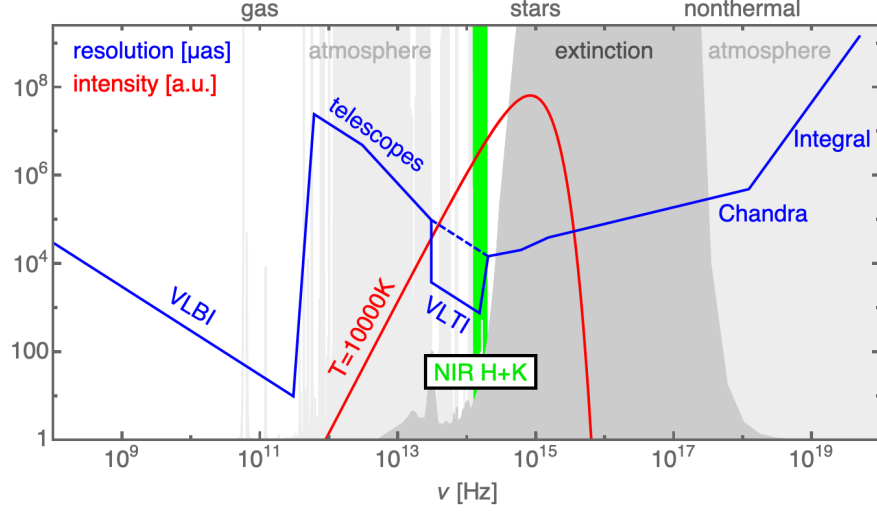


Figure 18: The resolution achievable as a function of observing wavelength (blue line). The red line is the emission of a black-body with 10^4 K. The darker gray shaded area marks the regime where interstellar dust prevents radiation from the Galactic Center reaching us. The lighter gray bands indicate the transmission of Earth's atmosphere. The green band marks the infrared H- and K-band, which is the sweet spot for observing stars in the vicinity of Sgr A*.

Adaptive optics images of the Galactic Center region show a plethora of stars (figure 19). Monitoring the field over three decades has revealed in the very central part the motions of stars, which cannot be described by straight lines - but require curved lines (i.e. the presence of an acceleration) or in some spectacular cases even full Keplerian orbits, like for the star S2 (figure 20, left). From the orbit of S2 alone ($a = 125$ mas, $P = 16$ yr) one can get a rather precise estimate of the central mass (assuming a distance):

$$M = \frac{4\pi^2 a^3}{G P^2} = \frac{4\pi^2}{6.67 \times 10^{-11} \frac{\text{N m}^2}{\text{kg}^2}} \frac{(125 \text{ mas} * 8300 \text{ pc})^3}{(16 \text{ yr}^2)} = 4.3 \times 10^6 M_{\odot} \quad (227)$$

The actual measurement is done via a fit to the data. The data are a set of positions ($i = 1..N_A$) and radial velocities ($j = 1..N_V$): $(x_i \pm \Delta x_i)$, $(y_i \pm \Delta y_i)$, $(v_{z,j} \pm \Delta v_{z,j})$. The model \mathcal{M} is the Keplerian orbit model, for an orbit tilted and rotated in space. The model has as parameters the initial conditions of the star $(x_0, y_0, z_0, v_{x,0}, v_{y,0}, v_{z,0})$, expressed in the classical orbital elements $(a, e, i, \Omega, \omega, t_P)$. Further, the mass of the central object, its position (x_C, y_C) - since we don't see it! - in 2D and distance R_0 . One can also more generally think of the positions and velocities being in an observer-defined coordinate system, and one needs to determine the transformation to the central-mass rest-frame. This can be only positional offsets, but also velocities, angles, rotations, shear, etc. But the key is that for any moment in time t , one can evaluate $\mathcal{M}(t|M, R_0, x_C, y_C, \dots, a, e, i, \Omega, \omega, t_P)$ to get $x(t), y(t), v_z(t)$. The path to evaluate the position-velocity vector for a given time t is the following:

1. Calculate the orbital period P

$$P = 2\pi \sqrt{\frac{a^3}{GM}} \quad (228)$$

2. Calculate the 'mean anomaly' M_A

$$M_A = 2\pi \frac{t - t_P}{P} \quad (229)$$



Figure 19: Left: Pseudo-color image of the Galactic Center, obtained with NACO (VLT) in the H-, K- and L-bands. The image is adaptive optics corrected, and has a Strehl ratio of around 40%. Right: Zoom into the central area, after image deconvolution. The motions of the stars visible in this image can be tracked from the AO images. S2 is the bright star in the center top. It orbits on an elliptical orbit with $P = 16$ yr. Sources: ESO, MPE

3. Calculate the 'eccentric anomaly' E by solving the so-called Kepler equation:

$$M_A = E - e \sin E \quad (230)$$

This needs to be done numerically. Newton's numerical scheme to find a root works well, and one can help the search by giving it a good starting value.

4. Calculate the 'true anomaly' ν

$$\tan \frac{\nu}{2} = \sqrt{\frac{1+e}{1-e}} \tan \frac{E}{2} \quad (231)$$

5. The distance to the central body is

$$r_c = a(1 - e \cos E) \quad (232)$$

6. The position and velocity in the orbital plane are

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = r_c \begin{pmatrix} \cos \nu \\ \sin \nu \\ 0 \end{pmatrix} ; \quad \begin{pmatrix} v_{x,c} \\ v_{y,c} \\ v_{z,c} \end{pmatrix} = \frac{\sqrt{GMa}}{r_c} \begin{pmatrix} -\sin E \\ \sqrt{1-e^2} \cos E \\ 0 \end{pmatrix} \quad (233)$$

7. Rotate into the perspective of the observer. With

$$R_x(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & -\sin \phi \\ 0 & \sin \phi & \cos \phi \end{pmatrix}; \quad R_z(\phi) = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (234)$$

the Cartesian coordinates are

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R_z(\Omega) \cdot R_x(i) \cdot R_z(\omega) \cdot \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix}; \quad \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = R_z(\Omega) \cdot R_x(i) \cdot R_z(\omega) \cdot \begin{pmatrix} v_{x,c} \\ v_{y,c} \\ v_{z,c} \end{pmatrix} \quad (235)$$

The angles in Galactic Center publications are defined such that the x here corresponds to the Dec. direction, and y to the RA direction (i.e. a right-handed coordinate system). There is actually two ways how to implement this recipe for fitting:

- When one knows that the orbits are Keplerian, one can for each measurement time t_i (astrometry) and t_j (radial velocities) simply execute the recipe, and get $x(t_i), y(t_i), v_z(t_j)$. This requires solving the Kepler equation numerically $N_A + N_V$ times.
- More general is to convert the elements once into Cartesian coordinates (one single Kepler equation), and integrate the equations of motion numerically, and then evaluate the solution function at all t_i, t_j . This scheme can also be used for non-Keplerian orbits, which means that the orbital elements have then the meaning of "osculating" elements at the time chosen for the conversion.

The best-fit orbit is then found by minimizing

$$\chi^2(M, R_0, x_C, y_C, \dots, a, e, i, \Omega, \omega, t_P) = \sum_{i=1}^{N_A} \left(\frac{x_i - x(t_i)}{\Delta x_i} \right)^2 + \sum_{i=1}^{N_A} \left(\frac{y_i - y(t_i)}{\Delta y_i} \right)^2 + \sum_{j=1}^{N_V} \left(\frac{v_{z,j} - v_z(t_j)}{\Delta v_{z,i}} \right)^2 \quad (236)$$

with respect to the parameters $(M, R_0, x_C, y_C, \dots, a, e, i, \Omega, \omega, t_P)$. This yields the best-fitting values and uncertainties, i.e. part of the result is for example the mass estimate and the distance estimate.

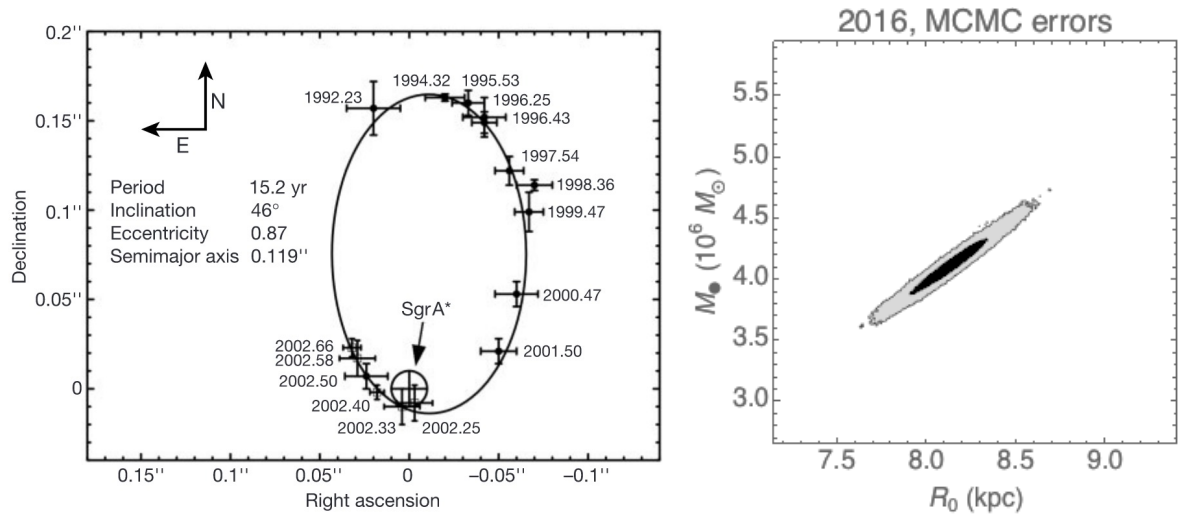


Figure 20: Left: Measured S2 positions as of Schödel et al. 2002, together with the best-fit orbital trace. Right: The posterior distribution for the parameters M and R_0 of an S2-orbit fit as of Gillessen et al. 2017