2.5.2023

General relativity

(notes for "Relativity" a.a. 2022/23)

Fiorenzo Bastianelli

1 Foreword

General Relativity is a vast subject with many books available. In this class, I will use two main textbooks to guide you through the basics. The first one is *S. Weinberg: Gravitation and Cosmology", John Wiley & Sons 1972*, which covers tensor analysis and Einstein's equation (chapters 3-7). The second one is *H. Ohanian and R. Ruffini: Gravitation and Spacetime", CUP 2013*, which offers more insights on classical tests, Schwarzschild black hole solution, and gravitational waves.

These notes are supplementary and incomplete. They will only cover some selected topics and fill in some gaps that the textbooks may leave. Therefore, you should rely on the textbooks as your primary source of study.

2 The principle of equivalence of gravitation and inertia

You can find more details about this topic in Chapter 3 of [1], which I recommend you to read.

Newton's law of universal gravitation tells us how massive bodies attract each other with gravitational forces. Suppose we have N particles with inertial masses $m_k^{(I)}$, gravitational masses $m_k^{(G)}$, and positions $\vec{x}_k$, with $k = 1, .., N$. Then, the gravitational force on the k-th particle is given by

$$m_k^{(I)} \frac{d^2 \vec{x}_k}{dt^2} = G \sum_{l \neq k} m_k^{(G)} m_l^{(G)} \frac{\vec{x}_l - \vec{x}_k}{|\vec{x}_l - \vec{x}_k|^3} \ . \tag{1}$$

In addition, it is found that the inertial mass and the gravitational mass of a particle are equal, as confirmed by experiments. This means that we can use the same mass for both the acceleration and the attraction of a particle, i.e. $m_k^{(I)} = m_k^{(G)}$ for any k.

The *principle of equivalence of gravitation and inertia* is based on this equality of masses and states that: *In any gravitational field, we can always find a local inertial frame (a free-falling frame) at any point in spacetime, such that near that point the laws of nature look like the ones in special relativity, where no gravitational field is present.*

This principle helps us to describe how gravity works and to find the equations that govern it (Einstein's equations). To apply this principle, we need to use tensor calculus, which allows us to change coordinates in spacetime in any way we want. This is a branch of mathematics called differential geometry. Einstein illustrated this principle with the example of an elevator that is falling freely under gravity.

The force of gravity on a point particle

Let us use the above principle to find out how one can describe the force of gravity that acts on a point particle of mass m. In the reference frame with coordinates x^μ, one observes a particle that feels a gravitational force. We have to discover how to describe mathematically this force. Thus, we use the principle of equivalence, which assure us that there must exist an inertial frame with coordinates ξ^α (a frame in free fall), such that locally (i.e. in a small neighborhood of the point where the particle is located and for a small amount of time around the time of observation) the particle satisfies the equations of motion of a free particle as known from the theory of special relativity

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

where the proper time τ is computed using the Minkowski metric $\eta_{\alpha\beta}$

$$d\tau^2 = -\eta_{\alpha\beta} d\xi^\alpha d\xi^\beta \ . \tag{3}$$

Then, we can go back to the original frame x^μ and recognize how the gravitational force is described. We use the relations between the coordinate systems, i.e. $\xi^\alpha = \xi^\alpha(x)$ and its inverse $x^\mu = x^\mu(\xi)$, to compute by the chain rule

$$0 = \frac{d^2\xi^\alpha}{d\tau^2} = \frac{d}{d\tau}\Big(\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{dx^\mu}{d\tau}\Big) = \frac{\partial^2\xi^\alpha}{\partial x^\nu \partial x^\mu}\frac{dx^\nu}{d\tau}\frac{dx^\mu}{d\tau} + \frac{\partial\xi^\alpha}{\partial x^\mu}\frac{d^2x^\mu}{d\tau^2} \ . \tag{4}$$

This equation is written more simply by multiplying with $\frac{\partial x^\lambda}{\partial \xi^\alpha}$ with a contraction on the index α, and using

$$\frac{\partial x^\lambda}{\partial \xi^\alpha}\frac{\partial \xi^\alpha}{\partial x^\mu} = \frac{\partial x^\lambda}{\partial x^\mu} = \delta^\lambda_\mu \ . \tag{5}$$

One finds

$$\boxed{\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\lambda}\frac{dx^\nu}{d\tau}\frac{dx^\lambda}{d\tau} = 0} \tag{6}$$

where we have defined the *affine connection*

$$\Gamma^\lambda_{\mu\nu} = \frac{\partial x^\lambda}{\partial \xi^\alpha}\frac{\partial^2\xi^\alpha}{\partial x^\mu \partial x^\nu} \tag{7}$$

and renamed indices. The affine connection is symmetric under exchange of the lower indices

$$\Gamma^\lambda_{\mu\nu} = \Gamma^\lambda_{\nu\mu} \tag{8}$$

because derivatives commute. Finally, one must express the proper time (3) in terms of the new coordinates

$$d\tau^2 = -\eta_{\alpha\beta} d\xi^\alpha d\xi^\beta = -\eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\mu}dx^\mu \frac{\partial\xi^\beta}{\partial x^\nu}dx^\nu \tag{9}$$

and write it in the form

$$\boxed{d\tau^2 = -g_{\mu\nu}(x) dx^\mu dx^\nu} \tag{10}$$

where the *metric tensor* $g_{\mu\nu}$ is defined by

$$g_{\mu\nu}(x) = \eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial\xi^\beta}{\partial x^\nu} \ . \tag{11}$$

There is also a direct relation between the metric tensor (which we interpret as the potential of the gravitational force) and the affine connection (the coefficients that determine the gravitational force on the particle). The relation is found as follows: one differentiates eq. (11) with respects to x^λ

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} = \eta_{\alpha\beta}\frac{\partial^2\xi^\alpha}{\partial x^\lambda \partial x^\mu}\frac{\partial\xi^\beta}{\partial x^\nu} + \eta_{\alpha\beta}\frac{\partial\xi^\alpha}{\partial x^\mu}\frac{\partial^2\xi^\beta}{\partial x^\lambda \partial x^\nu} \tag{12}$$

that using (7) one rewrites it as

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} = \Gamma^\rho_{\lambda\mu} g_{\rho\nu} + \Gamma^\rho_{\lambda\nu} g_{\rho\mu} \, . \tag{13}$$

Then, computing

$$\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\lambda}}{\partial x^\nu} = 2 g_{\rho\nu} \Gamma^\rho_{\lambda\mu} \tag{14}$$

and using the inverse metric $g^{\sigma\nu}$, that satisfies $g^{\sigma\nu} g_{\nu\rho} = \delta^\sigma_\rho$, one finds

$$\Gamma^\sigma_{\lambda\mu} = \frac{1}{2} g^{\sigma\nu} \Big(\frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\lambda\nu}}{\partial x^\mu} - \frac{\partial g_{\lambda\mu}}{\partial x^\nu} \Big) \, . \tag{15}$$

One may use the shorthand notation $\partial_\mu \equiv \frac{\partial}{\partial x^\mu}$, rename indices, and write this formula as

$$\boxed{\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\rho} (\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu})} \, . \tag{16}$$

The Newtonian limit

To relate to Newton's theory, let us look at a simple case of a slow-moving particle in a weak and stationary gravitational field. Since the particle moves slowly, we can ignore all but the zero-component of the 4-velocity in (6)

$$\frac{dx^\mu}{d\tau} = \Big(\frac{dx^0}{d\tau}, \frac{d\vec{x}}{d\tau} \Big) = (c\gamma, \vec{v}\gamma) = (\gamma, \vec{\beta}\gamma) \tag{17}$$

as $|\vec{\beta}| \ll 1$ for a slow motion (in our units $c = 1$), and find

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{00} \frac{dt}{d\tau} \frac{dt}{d\tau} = 0 \, . \tag{18}$$

For a stationary field, the time derivatives of the metric vanish and one is left with

$$\Gamma^\mu_{00} = -\frac{1}{2} g^{\mu\nu} \partial_\nu g_{00} \, . \tag{19}$$

For a weak field we may use nearly Cartesian coordinates and write the metric $g_{\mu\nu}$ as a weak perturbation of the Minkowski metric $\eta_{\mu\nu}$

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \, . \tag{20}$$

At lowest order in $h_{\mu\nu}$ we find

$$\Gamma^\mu_{00} = -\frac{1}{2} \eta^{\mu\nu} \partial_\nu h_{00} \qquad \to \qquad \begin{cases} \Gamma^0_{00} = 0 \\ \Gamma^i_{00} = -\frac{1}{2} \partial_i h_{00} \end{cases} \tag{21}$$

and the equations of motion (6) simplify to

$$\begin{cases} \dfrac{d^2 x^0}{d\tau^2} \equiv \dfrac{d^2 t}{d\tau^2} = 0 & (22) \\[2ex] \dfrac{d^2 \vec{x}}{d\tau^2} = \dfrac{1}{2} \vec{\nabla} h_{00} \Big(\dfrac{dt}{d\tau} \Big)^2 & (23) \end{cases}$$

The first equation tells us that t is proportional to τ, then using this proportionality we find that the second equation may be rewritten as

$$\frac{d^2\vec{x}}{dt^2} = \frac{1}{2}\vec{\nabla}h_{00} \tag{24}$$

that takes the Newtonian form

$$\frac{d^2\vec{x}}{dt^2} = -\vec{\nabla}\phi \tag{25}$$

under the identification

$$h_{00} = -2\phi \; . \tag{26}$$

For example, consider the case of the Newtonian potential created by a large mass M placed at the origin

$$\phi = -\frac{GM}{r} \tag{27}$$

then the '00' component of the metric takes the form

$$g_{00} = \eta_{00} + h_{00} = -1 - 2\phi) = -1 + \frac{2GM}{r} \; . \tag{28}$$

Note that ϕ is adimensional (we use $c = 1$). Some values are as follows: on the surface of the sun $\phi_{sun} \approx 10^{-6}$, on the earth $\phi_{earth} \approx 10^{-9}$, on a white dwarf $\phi_{wd} \approx 10^{-4}$, on the surface of a proton $\phi_{proton} \approx 10^{-39}$. Note also that $|g_{00}| < 1$, which is related to gravitational time dilation.

<u>Gravitational time dilation</u>

A clock in a gravitational field runs slower than a clock in a flat space. To see why, we can use the principle of equivalence and imagine a locally inertial frame where gravity has no effect on the clock. In this frame, we can choose coordinates ξ'^{α} where the clock is stationary and ticks at regular intervals $\Delta\tau$ as set by the maker. This is the proper time of the clock and we have

$$\Delta\tau \equiv dt' = d\xi'^0 \; . \tag{29}$$

Thus, $\Delta\tau$ is the clock's basic unit. It could be, for example, the period of a wave from a specific atomic transition from an atom at rest and in the absence of gravity. In another inertial frame with coordinates ξ^{α}, where the clock moves, the formula becomes

$$\Delta\tau = \sqrt{(d\xi'^0)^2} = \sqrt{-\eta_{\alpha\beta}d\xi^{\alpha}d\xi^{\beta}} \tag{30}$$

as given by special relativity. In the frame with coordinates x^{μ}, where gravity affects the clock, the space-time interval dx^{μ} between ticks is fixed by

$$\Delta\tau = \sqrt{-\eta_{\alpha\beta}d\xi^{\alpha}d\xi^{\beta}} = \sqrt{-\eta_{\alpha\beta}\frac{\partial\xi^{\alpha}}{\partial x^{\mu}}\frac{\partial\xi^{\beta}}{\partial x^{\nu}}dx^{\mu}dx^{\nu}} = \sqrt{-g_{\mu\nu}(x)dx^{\mu}dx^{\nu}} \; . \tag{31}$$

If the clock moves with velocity $\frac{dx^{\mu}}{dt}$, were $t = x^0$ is the time coordinate of the frame, one may write

$$\Delta\tau = \sqrt{-g_{\mu\nu}(x)\frac{dx^{\mu}}{dt}\frac{dx^{\nu}}{dt}}\, dt \; . \tag{32}$$

For a clock at rest in the gravitational field we set $\frac{d\vec{x}}{dt} = 0$

$$\Delta\tau = \sqrt{-g_{00}(x)}dt \tag{33}$$

to find that

$$dt = \frac{\Delta\tau}{\sqrt{-g_{00}(x)}} \tag{34}$$

is time dilated (recall that in the weak field limit $g_{00} = -1 + \frac{2GM}{r}$ has modulus smaller than 1).

To measure this gravitational time dilation, one has to compare time dilation at different points, otherwise the measuring device suffers the same time delay. Thus, taking two different points x_1^μ and x_2^μ one equates

$$\Delta\tau = \sqrt{-g_{00}(x_1)}dt_1 = \sqrt{-g_{00}(x_2)}dt_2 \tag{35}$$

to find

$$\frac{dt_1}{dt_2} = \sqrt{\frac{g_{00}(x_2)}{g_{00}(x_1)}} \tag{36}$$

that for the frequencies $\nu = \frac{1}{dt}$ corresponding to the periods dt give

$$\frac{\nu_2}{\nu_1} = \sqrt{\frac{g_{00}(x_2)}{g_{00}(x_1)}} \; . \tag{37}$$

In the weak field limit with $g_{00}(x) = -1 - 2\phi(x)$, setting $\nu_1 = \nu$ and $\nu_2 = \nu + \Delta\nu$, we find a change in frequency $\Delta\nu$ given by

$$\frac{\nu_2}{\nu_1} = \frac{\nu + \Delta\nu}{\nu} = 1 + \frac{\Delta\nu}{\nu} = \sqrt{\frac{g_{00}(x_2)}{g_{00}(x_1)}} = \sqrt{\frac{1 + 2\phi(x_2)}{1 + 2\phi(x_1)}} \approx 1 + \phi(x_2) - \phi(x_1) \; . \tag{38}$$

For an application, consider an atomic transition at the earth's surface with position x_1 and frequency ν_1. Observe the same transition at the sun's surface with position x_2 and frequency ν_2. The sun's potential $\phi(x_2)$ is much larger than the earth's potential $\phi(x_1)$, so that the latter can be ignored, and we get

$$\frac{\Delta\nu}{\nu} = \phi(x_2) - \phi(x_1) \approx \phi(x_2) = -\frac{GM_{sun}}{R_{sun}} \approx -2.12 \; 10^{-6} \; . \tag{39}$$

The frequency on the sun is smaller then the frequency seen far away for the same atomic transition: a *gravitational red shift* of the frequency is expected, as verified experimentally by now.

3 The principle of general covariance and tensor analysis

This topic is described in Chapter 4 of [1].
The *principle of equivalence* can be replaced by a more efficient *principle of general covariance*, which allows us to find in an easier way the correct equations valid in the presence of arbitrary gravitational fields. The *principle of general covariance* states that:
"*A physical equation is valid in an arbitrary gravitational field if it satisfies two conditions:*
(i) the equation reduces to the special relativistic form in the absence of gravity, where the metric is the Minkowski metric ($g_{\mu\nu} \to \eta_{\mu\nu}$) and the affine connection vanishes ($\Gamma^\lambda_{\mu\nu} \to 0$).
(ii) the equation is generally covariant, i.e. it keeps the same form under an arbitrary change of coordinates $x^\mu \to x'^\mu(x)$.

We can see that the principle of general covariance implies the principle of equivalence by noting that: point (ii) ensures that the equations are valid in any coordinate system if they are valid in one; point (i) then ensures that they are valid in a locally inertial frame, where gravity is canceled by inertial forces. To apply the principle of general covariance we need to use *tensor analysis*, which deals with arbitrary coordinate transformations and in mathematics is part of differential geometry.

Tensor analysis

Scalars, vectors, and tensors are quantities defined by their behavior under a change of coordinate system. A change of coordinates is specified by functions

$$x^\mu \quad \rightarrow \quad x'^\mu = x'^\mu(x) \tag{40}$$

that are required to be invertible. One can return to the original frame by using the inverse functions

$$x^\mu = x^\mu(x') \ . \tag{41}$$

Invertibility requires that

$$\det \frac{\partial x'^\mu}{\partial x^\nu} \neq 0 \ . \tag{42}$$

Notice that using the chain rule for differentiation and using the transformations in (40)–(41), one finds that

$$\delta^\mu_\nu = \frac{\partial x'^\mu}{\partial x'^\nu} = \frac{\partial x'^\mu}{\partial x^\lambda}\frac{\partial x^\lambda}{\partial x'^\nu} \tag{43}$$

which tells that the matrix (at any given point in spacetime)

$$\frac{\partial x'^\mu}{\partial x^\lambda} \tag{44}$$

is the inverse of the matrix

$$\frac{\partial x^\lambda}{\partial x'^\nu} \tag{45}$$

and viceversa. Then, one defines scalar, vectors, and tensors by

$$\begin{array}{lll}
\phi(x) & \rightarrow \phi'(x') = \phi(x) & \text{scalar} \\
V^\mu(x) & \rightarrow V'^\mu(x') = \dfrac{\partial x'^\mu}{\partial x^\nu} V^\nu(x) & \text{contravariant vector} \\
W_\mu(x) & \rightarrow W'_\mu(x') = \dfrac{\partial x^\nu}{\partial x'^\mu} W_\nu(x) & \text{covariant vector} \\
T^{\mu\nu}(x) & \rightarrow T'^{\mu\nu}(x') = \dfrac{\partial x'^\mu}{\partial x^\lambda}\dfrac{\partial x'^\nu}{\partial x^\rho} T^{\lambda\rho}(x) & \text{tensor of rank (2,0)} \\
S^\mu{}_\nu(x) & \rightarrow S'^\mu{}_\nu(x') = \dfrac{\partial x'^\mu}{\partial x^\lambda}\dfrac{\partial x^\rho}{\partial x'^\nu} S^\lambda{}_\rho(x) & \text{tensor of rank (1,1)} \\
\cdots & \rightarrow \qquad \cdots & \qquad \cdots
\end{array} \tag{46}$$

and so on for tensors of rank (m,n), where there are m matrices $\frac{\partial x'^\mu}{\partial x^\nu}$ that rotate the m upper indices, and n matrices $\frac{\partial x^\mu}{\partial x'^\nu}$ that rotate the n lower indices. For example, the $(2,1)$ tensor $F^{\mu\nu}{}_\lambda$ transforms as

$$F^{\mu\nu}{}_\lambda(x) \ \rightarrow \ F'^{\mu\nu}{}_\lambda(x') = \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x'^\nu}{\partial x^\sigma}\frac{\partial x^\lambda}{\partial x'^\tau} F^{\rho\sigma}{}_\tau(x) \tag{47}$$

Note that with these definitions, contraction of an upper index with a lower index produces a scalar

$$V^\mu(x)W_\mu(x) \quad\to\quad V'^\mu(x')W'_\mu(x') = \frac{\partial x'^\mu}{\partial x^\nu}V^\nu(x)\frac{\partial x^\lambda}{\partial x'^\mu}W_\lambda(x) = \underbrace{\frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\lambda}{\partial x'^\mu}}_{\delta^\lambda_\nu}V^\nu(x)W_\lambda(x) = V^\nu(x)W_\nu(x) \tag{48}$$

i.e.

$$V^\mu(x)W_\mu(x) \quad\to\quad V'^\mu(x')W'_\mu(x') = V^\mu(x)W_\mu(x) \tag{49}$$

which is precisely the transformation law of a scalar.

Not all quantities that we have defined so far are tensors: while the metric $g_{\mu\nu}(x)$ is a rank $(0,2)$ tensor

$$g_{\mu\nu}(x) \quad\to\quad g'_{\mu\nu}(x') = \frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial x^\rho}{\partial x'^\nu}g_{\lambda\rho}(x) \tag{50}$$

the affine connection $\Gamma^\lambda_{\mu\nu}$ is not a tensor as it transforms in a more complicated way

$$\Gamma^\lambda_{\mu\nu}(x) \quad\to\quad \Gamma'^\lambda_{\mu\nu}(x') = \frac{\partial x'^\lambda}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\mu}\frac{\partial x^\gamma}{\partial x'^\nu}\Gamma^\alpha_{\beta\gamma}(x) + \frac{\partial x'^\lambda}{\partial x^\alpha}\frac{\partial^2 x^\alpha}{\partial x'^\mu \partial x'^\nu} \tag{51}$$

where the second term, which can also be written as

$$\frac{\partial x'^\lambda}{\partial x^\alpha}\frac{\partial^2 x^\alpha}{\partial x'^\mu \partial x'^\nu} = -\frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\frac{\partial^2 x'^\lambda}{\partial x^\alpha \partial x^\beta} \tag{52}$$

breaks the tensorial behavior.

Thanks to these definitions, tensorial equations, that are defined as equations that equate tensors of the same rank, take the same form in all reference frames: e.g. if $A_{\mu\nu}$ and $B_{\mu\nu}$ are tensors then the tensorial equation

$$A_{\mu\nu}(x) = B_{\mu\nu}(x) \tag{53}$$

maintains the same form in all frames, i.e.

$$A_{\mu\nu}(x) = B_{\mu\nu}(x) \qquad\longleftrightarrow\qquad A'_{\mu\nu}(x') = B'_{\mu\nu}(x')\ . \tag{54}$$

The same equation can be written also as

$$A_{\mu\nu}(x) - B_{\mu\nu}(x) = 0 \tag{55}$$

where the right-hand side is understood to contain the zero tensor of appropriate rank, i.e. the tensor which has all components null (then, it is easily verified that the components vanish in all frames).

Tensors are basically elements of a vector space of suitable dimension. One can verify the following algebraic properties of tensors (*tensor algebra*):

A) A linear combination of tensors of the same rank is a tensor of the same rank.
E.g. if $A_{\mu\nu}$ and $B_{\mu\nu}$ are tensors and a and b scalars then

$$T_{\mu\nu} = aA_{\mu\nu} + bB_{\mu\nu} \tag{56}$$

is a tensor.

B) Tensor product (or direct product) of tensors.
The multiplications of the tensors components give rise to a new tensor of appropriate rank. E.g if $A_{\mu\nu}$ and B^{μ} are tensors, then

$$T_{\mu\nu}{}^{\lambda} = A_{\mu\nu} B^{\lambda} \tag{57}$$

is a tensor of rank $(1, 2)$.

C) Contraction of a contravariant index with a covariant index of a tensor produces a tensor of lower rank. E.g. taking the tensor $T_{\mu\nu}{}^{\lambda}$, one obtains a vector A_{μ} by setting

$$T_{\mu\nu}{}^{\nu} = A_{\mu} \ , \tag{58}$$

and another vector by

$$T_{\mu\nu}{}^{\mu} = B_{\nu} \ . \tag{59}$$

Indeed, we realized in (49) that index contraction gives rise to a scalar.

Finally, indices of a tensor may be raised and lowered by using the metric tensor $g_{\mu\nu}$, and its inverse $g^{\mu\nu}$ that satisfies

$$g_{\mu\nu}(x) g^{\nu\lambda}(x) = \delta_{\mu}^{\lambda}. \tag{60}$$

For example, from the contravariant vector V^{μ} one obtains the covariant vector V_{μ} as

$$V_{\mu} = g_{\mu\nu} V^{\nu} \tag{61}$$

and similarly, from the covariant vector V_{μ} one obtains the contravariant vector V^{μ} by

$$V^{\mu} = g^{\mu\nu} V_{\nu} \ . \tag{62}$$

<u>Covariant derivatives</u>

Derivatives of tensors are not tensors themselves. This causes a problem when defining tensorial equations that should contain derivatives. This problem is solved by using the concept of *covariant derivatives*, which are derivatives that when applied to tensors produce new tensors.

To expose clearly the problem, let us verify that for a vector field $V^{\nu}(x)$ its derivative

$$\partial_{\mu} V^{\nu}(x) = \frac{\partial V^{\nu}(x)}{\partial x^{\mu}} \tag{63}$$

is not a tensor. We compute

$$\begin{aligned}
\partial_{\mu} V^{\nu}(x) \ \ \rightarrow \ \ \partial'_{\mu} V'^{\nu}(x') &= \frac{\partial V'^{\nu}(x')}{\partial x'^{\mu}} = \frac{\partial}{\partial x'^{\mu}} \Big(V^{\beta}(x) \frac{\partial x'^{\nu}}{\partial x^{\beta}} \Big) = \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \frac{\partial}{\partial x^{\alpha}} \Big(V^{\beta}(x) \frac{\partial x'^{\nu}}{\partial x^{\beta}} \Big) \\
&= \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \Big(\frac{\partial V^{\beta}(x)}{\partial x^{\alpha}} \frac{\partial x'^{\nu}}{\partial x^{\beta}} + V^{\beta}(x) \frac{\partial^2 x'^{\nu}}{\partial x^{\alpha} \partial x^{\beta}} \Big) \\
&= \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \frac{\partial x'^{\nu}}{\partial x^{\beta}} \partial_{\alpha} V^{\beta}(x) + V^{\beta}(x) \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \frac{\partial^2 x'^{\nu}}{\partial x^{\alpha} \partial x^{\beta}}
\end{aligned} \tag{64}$$

where the first term would be the one expected for a tensorial transformation, but the second term breaks the tensorial character. A similar transformation appears in the transformation rule of the affine connection, see eqs. (51)–(52), which can be used to introduce the concept of *covariant derivative* of the vector field, defined by

$$\nabla_{\mu} V^{\nu} = \partial_{\mu} V^{\nu} + \Gamma^{\nu}_{\mu\lambda} V^{\lambda} \tag{65}$$

which then transforms as a tensor

$$\nabla_\mu V^\nu(x) \quad \to \quad \nabla'_\mu V'^\nu(x') = \frac{\partial x^\rho}{\partial x'^\mu}\frac{\partial x'^\nu}{\partial x^\sigma}\nabla_\rho V^\sigma(x) \ . \tag{66}$$

Geometrically, the connection connect tangent spaces at different nearby points and allows to define a parallel transport of vectors which are compared in defining the covariant derivative. The concept of a covariant derivative allows to introduce tensorial equations with derivatives, and thus identify the correct differential equations describing physical systems under the force of gravity.

In general, covariant derivatives of contravariant and covariant vectors are defined by

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda} V^\lambda \tag{67}$$
$$\nabla_\mu V_\nu = \partial_\mu V_\nu - \Gamma^\lambda_{\mu\nu} V_\lambda \tag{68}$$

and similarly for more general tensors, which will have a connection for each index, e.g.

$$\nabla_\mu V^{\nu\lambda}{}_\rho = \partial_\mu V^{\nu\lambda}{}_\rho + \Gamma^\nu_{\mu\alpha} V^{\alpha\lambda}{}_\rho + \Gamma^\lambda_{\mu\alpha} V^{\nu\alpha}{}_\rho - \Gamma^\alpha_{\mu\rho} V^{\nu\lambda}{}_\alpha \ . \tag{69}$$

The covariant derivative satisfies the Leibniz rule for taking the derivative of products of tensors. For example, one may verify that on a scalar the covariant derivative reduces to the usual derivative consistently with the Leibniz rule: taking the scalar $V^\mu W_\mu$, its derivative can be expanded as

$$\partial_\nu (V^\mu W_\mu) = \partial_\nu V^\mu \, W_\mu + V^\mu \, \partial_\nu W_\mu \tag{70}$$

which of course must be correct, while using covariant derivatives

$$\begin{aligned}
\partial_\nu (V^\mu W_\mu) &= \nabla_\nu (V^\mu W_\mu) = \nabla_\nu V^\mu \, W_\mu + V^\mu \, \nabla_\nu W_\mu \\
&= (\partial_\nu V^\mu + \Gamma^\mu_{\nu\lambda} V^\lambda) W_\mu + V^\mu (\partial_\nu W_\mu - \Gamma^\lambda_{\nu\mu} W_\lambda) \\
&= \partial_\nu V^\mu \, W_\mu + V^\mu \, \partial_\nu W_\mu
\end{aligned} \tag{71}$$

as renaming indices it is verified that the terms with connections cancel each other.

The covariant derivative of the metric vanishes

$$\nabla_\lambda g_{\mu\nu} = \partial_\lambda g_{\mu\nu} - \Gamma^\sigma_{\lambda\mu} g_{\sigma\nu} - \Gamma^\sigma_{\lambda\nu} g_{\mu\sigma} = 0 \ , \tag{72}$$

a fact which is referred to by saying that the metric is covariantly constant. This can be verified using eq. (16), that relates the affine connection to the derivative of the metric. In turn, eq. (72) can be used to rederive relation (16). Let us show this point. Since the metric is covariantly constant we may write

$$\begin{aligned}
0 &= \nabla_\lambda g_{\mu\nu} + \nabla_\mu g_{\lambda\nu} - \nabla_\nu g_{\lambda\mu} \\
&= \ \partial_\lambda g_{\mu\nu} - \Gamma^\sigma_{\lambda\mu} g_{\sigma\nu} - \Gamma^\sigma_{\lambda\nu} g_{\mu\sigma} \\
&\quad + \partial_\mu g_{\lambda\nu} - \Gamma^\sigma_{\mu\lambda} g_{\sigma\nu} - \Gamma^\sigma_{\mu\nu} g_{\lambda\sigma} \\
&\quad - \partial_\nu g_{\lambda\mu} + \Gamma^\sigma_{\nu\lambda} g_{\sigma\nu} + \Gamma^\sigma_{\nu\mu} g_{\lambda\sigma} \\
&= \partial_\lambda g_{\mu\nu} + \partial_\mu g_{\lambda\nu} - \partial_\nu g_{\lambda\mu} - 2\Gamma^\sigma_{\lambda\mu} g_{\sigma\nu}
\end{aligned} \tag{73}$$

where we have used the symmetry on the first two indices of the affine connection. Then, using the inverse metric we find

$$\Gamma^\sigma_{\lambda\mu} = \frac{1}{2} g^{\sigma\nu} (\partial_\lambda g_{\mu\nu} + \partial_\mu g_{\lambda\nu} - \partial_\nu g_{\lambda\mu}) \tag{74}$$

as expected.

Finally, let us consider a vector field defined only along a curve, rather then all over spacetime. The covariant derivative of a vector $V^\mu(\tau)$ defined along a curve parametrized by $x^\mu(\tau)$, where τ is the parameter, takes the form

$$\frac{DV^\mu}{d\tau} = \frac{dV^\mu}{d\tau} + \frac{dx^\rho}{d\tau}\Gamma^\mu_{\rho\sigma}V^\sigma \ . \tag{75}$$

As an example, consider a worldline parameterized by $x^\mu(\tau)$. Its derivative with respect to the parameter τ gives the four-velocity, i.e. the tangent vector to the curve

$$\frac{dx^\mu}{d\tau} \tag{76}$$

which is easily verified to have a vector transformations law (just use the chain rule for differentiation). Its covariant derivative takes the form

$$\frac{D}{d\tau}\frac{dx^\mu}{d\tau} = \frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma}\frac{dx^\rho}{d\tau}\frac{dx^\sigma}{d\tau} \tag{77}$$

which is again a vector defined along the curve. With the notation

$$\frac{dx^\mu}{d\tau} \equiv \dot{x}^\mu \tag{78}$$

we usually write it as

$$\frac{D}{d\tau}\frac{dx^\mu}{d\tau} = \ddot{x}^\mu + \Gamma^\mu_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma \ . \tag{79}$$

The equation

$$\frac{D}{d\tau}\frac{dx^\mu}{d\tau} = 0 \tag{80}$$

is known as the geodesic equation, already encountered in (6) and used to describe the motion of a particle under the force of gravity.

4 Effects of gravitation

This topic is described in Chapter 5 of [1].

We now make use of the tensor calculus introduced earlier to study how gravity affects the equations of mechanics and electromagnetism studied in special relativity. We make use of the principle of general covariance. We take the equations we know from special relativity and rewrite them in a form that makes them generally covariant. This last step is achieved by identifying the tensors that enter the equations and which describe the physical quantities, substitute the Minkowski metric $\eta_{\mu\nu}$ by $g_{\mu\nu}$, and then substitute derivatives of tensors by their covariant derivatives. In this way we reach equations that are generally covariant, and we have succeded in introducing the force of gravity according to the principle of general covariance.

Particle mechanics

We know that the motion of a free particle is described in special relativity by eqs. (2)–(3). Using the particle coordinates x^μ they read

$$\frac{d^2x^\mu}{d\tau^2} = 0 \ , \qquad d\tau^2 = -\eta_{\mu\nu}dx^\mu dx^\nu \ . \tag{81}$$

These equations are easily covariantized. The first derivative with respect to the proper time $\frac{dx^\mu}{d\tau}$ defines the 4-velocity, that is easily verified to be a vector. Then, we use the covariant derivative to keep its vector character

$$\frac{D}{d\tau}\frac{dx^\mu}{d\tau} = 0 \ . \tag{82}$$

This is the correct covariant equation. In standard notations it takes the form

$$\ddot{x}^\mu + \Gamma^\mu_{\rho\sigma}\dot{x}^\rho\dot{x}^\sigma = 0 \ . \tag{83}$$

Similarly, the proper time is covariantized by substituting $\eta_{\mu\nu}$ by the metric $g_{\mu\nu}$

$$d\tau^2 = -g_{\mu\nu}dx^\mu dx^\nu \ . \tag{84}$$

To summarize, the covariant extension of eqs. (81) is

$$\frac{D}{d\tau}\frac{dx^\mu}{d\tau} = 0 \ , \qquad d\tau^2 = -g_{\mu\nu}dx^\mu dx^\nu \tag{85}$$

which is interpreted as including the force of gravity acting on the particle.

Klein-Gordon equation

The Klein-Gordon equation is a relativistic equation for a scalar field $\phi(x)$. In natural units ($c = \hbar = 1$) it reads

$$(\partial^\mu\partial_\mu - m^2)\phi(x) = 0 \tag{86}$$

The first derivative of a scalar is already covariant, as it gives rise to a covariant vector without the need of a connection

$$\nabla_\mu\phi = \partial_\mu\phi \ . \tag{87}$$

The second derivative requires instead a connection

$$\nabla_\mu\nabla_\nu\phi = \partial_\mu\nabla_\nu\phi - \Gamma^\lambda_{\mu\nu}\nabla_\lambda\phi = \partial_\mu\partial_\nu\phi - \Gamma^\lambda_{\mu\nu}\partial_\lambda\phi \ . \tag{88}$$

Finally, indices are raised covariantly with the inverse metric $g^{\mu\nu}$ so that the covariantized Klein-Gordon equation becomes

$$(g^{\mu\nu}\nabla_\mu\nabla_\nu - m^2)\phi = 0 \tag{89}$$

also written as

$$(\nabla^\mu\nabla_\mu - m^2)\phi = 0 \ . \tag{90}$$

It is customary to denote the covariant d'Alambert operator by a *box*

$$\Box = \nabla^\mu\nabla_\mu \tag{91}$$

and write the covariant equation also as

$$(\Box - m^2)\phi = 0 \ . \tag{92}$$

Electrodynamics

Maxwell's equations can be written in special relativity as

$$\begin{aligned} &\partial_\mu F^{\mu\nu} = -J^\nu \\ &\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \end{aligned} \tag{93}$$

with the second one solved in terms of the 4-potential A_μ by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \ . \tag{94}$$

Let us see how to covariantize these equations. The second equation in (93) and its solution (94) are covariantized by substituting derivatives with covariant derivatives

$$\begin{aligned} &\nabla_\mu F_{\nu\lambda} + \nabla_\nu F_{\lambda\mu} + \nabla_\lambda F_{\mu\nu} = 0 \\ &F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu \end{aligned} \tag{95}$$

so that hey become manifestly covariant. However, one may verify that the connection drops out in all of them, so that the original equations

$$\begin{aligned} &\partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} + \partial_\lambda F_{\mu\nu} = 0 \\ &F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \end{aligned} \tag{96}$$

were nevertheless covariant. In a sense, these equations do not feel the force of gravity.

As for the first one in (93), one must first raise indices with the general metric $g_{\mu\nu}$

$$F^{\mu\nu} = g^{\mu\rho} g^{\nu\sigma} F_{\rho\sigma} \tag{97}$$

and then write it using a covariant derivative

$$\nabla_\mu F^{\mu\nu} = -J^\nu \ . \tag{98}$$

This is a covariant equation if J^ν is a contravariant vector, which we assume to be the case.

The Lorentz force and gravity

Finally, let us include gravity in the Lorentz force equation (with $c = 1$)

$$m \frac{d^2 x^\mu}{d\tau^2} = e F^{\mu\nu} \frac{dx_\nu}{d\tau} \tag{99}$$

by covariantizing it. We obtain

$$m \frac{D}{d\tau} \frac{dx^\mu}{d\tau} = e F^{\mu\nu} g_{\nu\lambda} \frac{dx^\lambda}{d\tau} \tag{100}$$

recalling however that one must be careful with indices that are lowered and raised now with $g_{\mu\nu}$ and $g^{\mu\nu}$. To better expose the places where the metric sits, we rewrite the covariant equation as

$$m \frac{D}{d\tau} \frac{dx^\mu}{d\tau} = e g^{\mu\nu} F_{\nu\lambda} \frac{dx^\lambda}{d\tau} \tag{101}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, or more explicitly as

$$m(\ddot{x}^\mu + \Gamma^\mu_{\nu\lambda} \dot{x}^\nu \dot{x}^\lambda) = e g^{\mu\nu} F_{\nu\lambda} \dot{x}^\lambda \ . \tag{102}$$

5 Curvature

This topic is described in Chapter 6 of [1].

The equations of motion for the gravitational field describe the behavior of the metric tensor $g_{\mu\nu}(x)$, which acts as the potential for the gravitational force. To ensure covariance, these equations must be constructed using tensors. However, it can be proven that no tensor can be constructed solely from the metric $g_{\mu\nu}$ and its first derivatives $\partial_\lambda g_{\mu\nu}$. This is because the covariant derivatives of the metric vanish, $\nabla_\lambda g_{\mu\nu} = 0$, while the affine connection $\Gamma^\lambda_{\mu\nu}$, defined as

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}), \tag{103}$$

is not a tensor. Both $\Gamma^\lambda_{\mu\nu}$ and $\partial_\lambda g_{\mu\nu}$ contain 40 independent components, and the former is equivalent to the latter.

One can construct tensors by including also second derivatives of the metric. To find them one can use covariant derivatives. This provides a quick way of identifying such tensors.

Covariant derivatives do not commute and they may be used to define implicitly the Riemann curvature tensor ${R_{\mu\nu}}^\lambda{}_\rho$ by the relation

$$[\nabla_\mu, \nabla_\nu] V^\lambda = {R_{\mu\nu}}^\lambda{}_\rho V^\rho \; . \tag{104}$$

The left-hand side is a tensor, so must be the right-hand side, and in particular the object ${R_{\mu\nu}}^\lambda{}_\rho$ must indeed be a tensor. It is manifestly antisymmetric under the exchange of the first two indices μ, ν. A direct calculation shows that

$${R_{\mu\nu}}^\lambda{}_\rho = \partial_\mu \Gamma^\lambda_{\nu\rho} - \partial_\nu \Gamma^\lambda_{\mu\rho} + \Gamma^\lambda_{\mu\sigma}\Gamma^\sigma_{\nu\rho} - \Gamma^\lambda_{\nu\sigma}\Gamma^\sigma_{\mu\rho} \; . \tag{105}$$

A mnemonic for remembering this structure is to write

$${R_{\mu\nu}}^\lambda{}_\rho = \bar\nabla_\mu \Gamma^\lambda_{\nu\rho} - (\mu \leftrightarrow \nu) \tag{106}$$

where $\bar\nabla_\mu$ contains a connection for the upper index only (in general, covariant derivatives are defined only for tensors).

Algebraic properties of the Riemann tensor are best written lowering the upper index with the metric, $R_{\mu\nu\lambda\rho} = g_{\lambda\sigma} {R_{\mu\nu}}^\sigma{}_\rho$. They are the following ones

$$R_{\mu\nu\lambda\rho} = R_{\lambda\rho\mu\nu} \qquad (\textit{symmetry}) \tag{107}$$

$$R_{\mu\nu\lambda\rho} = -R_{\nu\mu\lambda\rho} = -R_{\mu\nu\rho\lambda} \qquad (\textit{antisymmetry}) \tag{108}$$

$$R_{\mu\nu\lambda\rho} + R_{\lambda\mu\nu\rho} + R_{\nu\lambda\mu\rho} = 0 \qquad (\textit{cyclicity}) \; . \tag{109}$$

A way of proving them is to write them down explicitly in terms of the metric using (105) and (103). This is very laborious, but correct.

Additional tensors can be constructed by index contraction and are the Ricci tensor and the curvature scalar

$$R_{\mu\nu} = {R_{\lambda\mu}}^\lambda{}_\nu \qquad \text{(Ricci tensor)} \tag{110}$$

$$R = g^{\mu\nu} R_{\mu\nu} \qquad \text{(Ricci scalar or curvature scalar)} \; . \tag{111}$$

Other contractions do not give rise to independent tensors. From (107) it follows that the Ricci tensor is symmetric

$$R_{\mu\nu} = R_{\nu\mu} \; . \tag{112}$$

One can compute the number of independent components C_D of the Riemann tensor in arbitrary dimensions D. They are given by

$$\begin{aligned} C_D &= \frac{1}{2}\Big(\frac{1}{2}D(D-1)\Big)\Big(\frac{1}{2}D(D-1)+1\Big) - \frac{D(D-1)(D-2)(D-3)}{4!} \\ &= \frac{1}{12}D^2(D^2-1) \ . \end{aligned} \tag{113}$$

This values is obtained by considering the Riemann tensors as a symmetric matrix R_{AB} where A and B stands for the ordered pair of indices $(\mu\nu)$ with $\mu < \nu$. The last term subtracts the independent relations in (109), which is an expression that can be proven to be completely antisymmetric under exchange of indices. A few values are reported in the following table

D	D^4	C_D
1	1	0
2	16	1
3	81	6
4	256	20
5	625	50

Of course, we use $D = 4$ for our purposes: we see that the Riemann tensor has 20 components while the Ricci tensor has 10 components like the metric.

5.1 Bianchi identities

The Riemann tensors satisfies the following differential Bianchi identities

$$\nabla_\mu R_{\nu\lambda\alpha\beta} + \nabla_\nu R_{\lambda\mu\alpha\beta} + \nabla_\lambda R_{\mu\nu\alpha\beta} = 0 \ . \tag{114}$$

The sum over cyclic permutation of the first three indices makes the total tensor antisymmetric in those indices.

One may contract the Bianchi identities on the indices (ν, α) (i.e. multiplying by $g^{\nu\alpha}$) to find

$$\nabla_\mu R_{\lambda\beta} + \nabla^\alpha R_{\lambda\mu\alpha\beta} - \nabla_\lambda R_{\mu\beta} = 0 \tag{115}$$

and contracting once more the indices (λ, β) one finds

$$\nabla_\mu R - 2\nabla^\alpha R_{\mu\alpha} = 0 \qquad \to \qquad \nabla^\mu \Big(R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R\Big) = 0 \ . \tag{116}$$

It is customary to define the Einstein tensor $G_{\mu\nu}$ by

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \tag{117}$$

which is covariantly conserved as seen form eq. (116), $\nabla^\mu G_{\mu\nu} = 0$.

Exercizes

These exercises help in proving some of the symmetry properties of the Riemann tensor.

Ex.1 *Recalling that the metric is covariantly constant* ($\nabla_\mu g_{\alpha\beta} = 0$) *use* $[\nabla_\mu, \nabla_\nu] g_{\alpha\beta} = 0$ *to prove the antisymmetry in the last two indices of the Riemann tensor,* $R_{\mu\nu\alpha\beta} = -R_{\mu\nu\beta\alpha}$.

Ex. 2 *Rewriting the Bianchi identities for electromagnetism using covariant derivatives, show* *the cyclic property of the Riemann tensor.*

Ex. 3 *From the Jacobi identity valid for arbitrary operators* A, B, C

$$[A,[B,C]]+[B,[C,A]]+[C,[A,B]]=0$$

which is a consequence of the associativity of the multiplication of operators, consider the case with $(A,B,C)\equiv(\nabla_\mu,\nabla_\nu,\nabla_\lambda)$ *acting on a vector field* V^ρ, *i.e.*

$$\Big([\nabla_\mu,[\nabla_\nu,\nabla_\lambda]]+[\nabla_\nu,[\nabla_\lambda,\nabla_\mu]]+[\nabla_\lambda,[\nabla_\mu,\nabla_\nu]]\Big)V^\rho=0$$

and prove the Bianchi identities in (114).

6 Einstein's equations of general relativity

This topic is described in Chapter 7 of [1].

We now come to the Einstein's field equations and write down the main equations in our notations.

Einstein's equations (the equivalent for the metric of Maxwell's equations for the potential A_μ) can be identified by using the principle of general covariance, which embodies the principle of equivalence. We know that any gravitational field can be made sufficiently small in a small region by using a local inertial frame (that in fact makes the gravitational field vanish at a point).

A *weak* and *static* field due to non-relativistic matter with mass density $\rho(x)$ is described by a newtonian potential ϕ, embedded in the component g_{00} of the metric as

$$\nabla^2\phi = 4\pi G\rho \tag{118}$$

$$g_{00}\approx-(1+2\phi) \tag{119}$$

where $G = 6.67\ 10^{-11} Nm^2/Kg^2$ is the Newton gravitational constant. For example, a point-like particle of mass M at rest has a mass density

$$\rho(x) = M\delta^3(\vec{x}) \tag{120}$$

and it gives rise to a potential that satisfies the equation

$$\nabla^2\phi = 4\pi GM\delta^3(\vec{x}) \qquad \to \qquad \phi(x) = -\frac{GM}{r}\ . \tag{121}$$

In special relativity mass and energy are equivalent, so that one can take $\rho(x)$ as the energy density, which appears as the T_{00} component of the energy-momentum tensor (also named stress tensor) of the matter system, and rewrite the equation for the gravitational potential as

$$\nabla^2 g_{00} = -8\pi G T_{00}\ . \tag{122}$$

Special relativity implies that there must be a tensor $G_{\alpha\beta}$ (tensor under Lorentz transformations) with component $G_{00}=-\nabla^2 g_{00}$ (the minus sign is conventional) that can be constructed with second derivatives of the metric, so that the Lorentz invariant extension of (122) becomes

$$G_{\alpha\beta} = 8\pi G T_{\alpha\beta} \tag{123}$$

where the complete energy-momentum tensor $T_{\alpha\beta}$ appears on the right-hand side. So far, this is just a consequence of special relativity. Finally, general relativity is obtained by searching for a general covariant extension that must take the general covariant form

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \ . \tag{124}$$

The conservation of $T_{\alpha\beta}$, namely $\partial^\alpha T_{\alpha\beta} = 0$ is covariantized to $\nabla^\mu T_{\mu\nu} = 0$, and by consistency also $G_{\mu\nu}$ must be covariantly conserved, i.e. $\nabla^\mu G_{\mu\nu}$. The weak and static limit identifies it uniquely with the Einstein tensor.

These considerations lead to the Einstein's equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu} \tag{125}$$

which are generally covariant field equations for the metric $g_{\mu\nu}$. The tensor $T_{\mu\nu}$ is the energy-momentum tensor of the matter that gravitates. An equivalent way of writing these equations is to first take the trace (by multiplying with $g^{\mu\nu}$) to find (in four spacetime dimensions)

$$R - 2R = 8\pi G T^\mu{}_\nu \qquad \to \qquad R = -8\pi G T^\mu{}_\mu$$

so that Einstein's equations take the form

$$R_{\mu\nu} = 8\pi G \Big(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T^\lambda{}_\lambda \Big) \ . \tag{126}$$

In vacuum, these equations reduce to

$$R_{\mu\nu} = 0 \ . \tag{127}$$

An additional term with a dimensionful coupling constant Λ with positive mass dimensions, the so-called cosmological constant, can be added to the equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \ . \tag{128}$$

Originally introduced and then rejected by Einstein, nowadays it allows to describe the presence of dark energy in the universe.

Finally, reintroducing by dimensional analysis the speed of light c, Einstein's equations takes the form

$$\boxed{R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}} \ . \tag{129}$$

We will continue to use units with $c = 1$.

7 Harmonic gauge

The gauge symmetry associated to the arbitrary change of coordinates can be used to simplify the analysis of Einstein's equations.

The gauge symmetry implies that given a solution $g_{\mu\nu}(x)$, also $g'_{\mu\nu}(x)$ will be a solution if the functions in $g'_{\mu\nu}$ are obtained by a change of coordinates

$$g'_{\mu\nu}(x') = g_{\alpha\beta}(x) \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} \ . \tag{130}$$

Infinitesimally, under the change of coordinates $x'^\mu = x^\mu - \xi^\mu(x)$, the metric varies as

$$\begin{aligned}\delta g_{\mu\nu}(x) \equiv g'_{\mu\nu}(x) - g_{\mu\nu}(x) &= \xi^\alpha \partial_\alpha g_{\mu\nu} + (\partial_\mu \xi^\alpha) g_{\alpha\nu} + (\partial_\nu \xi^\alpha) g_{\mu\alpha} \\ &= \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu\end{aligned} \tag{131}$$

The previuos gauge symmetries can be fixed by requiring the harmonic gauge (or De Donder gauge) conditions

$$\Gamma^\mu \equiv g^{\nu\lambda}\Gamma^\mu_{\nu\lambda} = 0 \qquad \leftrightarrow \qquad \partial_\nu(\sqrt{g} g^{\nu\mu}) = 0 \tag{132}$$

These four conditions specify a gauge in which the coordinates are harmonic functions, just like the cartesian coordinates of flat spacetime, and are sometimes called quasi-cartesian coordinates.

8 Linearized Einstein's equations

To study the Einstein's equations in a linearized apporximation around flat spacetime, one sets the metric as

$$g_{\mu\nu}(x) = \eta_{\mu\nu} + h_{\mu\nu}(x) \tag{133}$$

and considers $|h_{\mu\nu}(x)| \ll 1$. Then one may raise and lower indices with the Minkowski metric

$$h^{\mu\nu} = \eta^{\mu\alpha}\eta^{\nu\beta} h_{\alpha\beta} \tag{134}$$

and define for simplicity the "trace" of $h_{\mu\nu}$

$$h = \eta^{\mu\nu} h_{\mu\nu} \ . \tag{135}$$

Then, one may compute at the linear order in $h_{\mu\nu}$

$$g^{\mu\nu}(x) = \eta^{\mu\nu} - h^{\mu\nu}(x) \ , \quad g = |\det g_{\mu\nu}| = 1 + h \ , \quad \sqrt{g} = 1 + \frac{1}{2} h \tag{136}$$

The Christoffel symbols linearize as

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2}\eta^{\rho\sigma}(\partial_\mu h_{\nu\sigma} + \partial_\nu h_{\mu\sigma} - \partial_\sigma h_{\mu\nu}) = \frac{1}{2}(\partial_\mu h_\nu{}^\rho + \partial_\nu h_\mu{}^\rho - \partial^\rho h_{\mu\nu}) \ , \tag{137}$$

the Riemann tensor as

$$R_{\mu\nu}{}^\rho{}_\sigma = \partial_\mu \Gamma^\rho_{\nu\sigma} - \partial_\nu \Gamma^\rho_{\mu\sigma} + ... = \frac{1}{2}\partial_\sigma(\partial_\mu h_\nu{}^\rho - \partial_\nu h_\mu{}^\rho) - \frac{1}{2}\partial^\rho(\partial_\mu h_{\nu\sigma} - \partial_\nu h_{\mu\sigma}) \tag{138}$$

and the Ricci tensor

$$R_{\nu\sigma} = R_{\mu\nu}{}^\mu{}_\sigma = \frac{1}{2}(\partial_\nu \partial^\mu h_{\sigma\mu} + \partial_\sigma \partial^\mu h_{\nu\mu} - \partial_\nu \partial_\sigma h - \Box h_{\nu\sigma}) \tag{139}$$

where now $\Box = \partial^\mu \partial_\mu = \eta^{\mu\nu}\partial_\mu \partial_\nu$.

Then, Einstein's equations in vacuum take the linearized form

$$\Box h_{\mu\nu} + \partial_\mu \partial_\nu h - \partial_\mu \partial^\sigma h_{\sigma\nu} - \partial_\nu \partial^\sigma h_{\sigma\mu} = 0 \ . \tag{140}$$

One can verify that they are gauge invariant under the gauge symmetry, that to lowest order in $h_{\mu\nu}$ simplifies to

$$\delta h_{\mu\nu} = \partial_\mu \xi_\nu + \partial_\nu \xi_\mu \tag{141}$$

where the four components of ξ_μ are arbitrary functions. These symmetries can be used to set four gauge-fixing conditions, that may be take to be the linearzied harmonic (De Donder) gauge

$$\partial^\sigma h_{\sigma\mu} = \frac{1}{2}\partial_\mu h \tag{142}$$

which simplify Einstein's equations to

$$\Box h_{\mu\nu} = 0 \tag{143}$$

which evidently support plane waves solutions (gravitational waves).

It can be shown that only two independent polarizations of the gravitational waves can exist, just like the electromagnetic waves.

8.1 Electromagnetic waves and physical polarizations

First, let us review the case of the electromagnetic waves which have only two degrees of freedom, the two possible polarizations of the waves. The introduction of the four-potential A_μ solves half of the Maxwell equations. The remaining ones in vacuum take the form

$$\partial^\mu F_{\mu\nu} = \partial^\mu(\partial_\mu A_\nu - \partial_\nu A_\mu) = 0 \tag{144}$$

and are gauge invariant under

$$\delta A_\mu = \partial_\mu \theta \tag{145}$$

with θ an arbitrary function of spacetime. The gauge freedom allows to impose the Lorenz gauge $\partial^\mu A_\mu = 0$. In this gauge the equations simplify to

$$\begin{aligned} \Box A_\mu &= 0 \\ \partial^\mu A_\mu &= 0 \ . \end{aligned} \tag{146}$$

Plane wave solution are found using the ansatz (up to an overall normalization) by setting

$$A_\mu(x) = \epsilon_\mu(k)\, e^{ik\cdot x} + c.c. \tag{147}$$

where $\epsilon_\mu(k)$ is an arbitrary polarization depending on the wave vector k^μ, and the exponent contains the Lorentz invariant phase $k \cdot x = k_\mu x^\mu = \eta_{\mu\nu} k^\mu x^\nu = -k^0 x^0 + \vec{k} \cdot \vec{x}$. The notation *c.c.* stands for complex conjugation, and makes the solution real. Plugging this ansatz into the equations (146), one finds a solution when

$$k^\mu k_\mu = 0 \ , \qquad k^\mu \epsilon_\mu(k) = 0 \ . \tag{148}$$

Thus, only three polarizations $\epsilon_\mu(k)$ are possible. However, one of these polarizations is not physical, the one with $\epsilon_\mu(k) \sim k_\mu$. It does not carry electric and magnetic fields, and thus no energy and momentum. It can be removed by a gauge transformation. The gauge transformations that removes it has the form in (145) with

$$\theta(x) \sim e^{ik\cdot x} \tag{149}$$

that satisfies $\Box\theta(x) = 0$, and thus does not ruin the Lorenz gauge condition. The gauge transformation becomes

$$\delta A_\mu = \partial_\mu \theta \sim i k_\mu\, e^{ik\cdot x} \tag{150}$$

and shows that the polarization $\epsilon_\mu(k) \sim k_\mu$ is not physical, as can be removed by an appropriate gauge transformation. Only two physical polarizations remain.

Let us exemplify this considering the motion along the z axis. We can take

$$k^\mu = (k^0, \vec{k}) = (\omega, 0, 0, \omega) \tag{151}$$

which solves $k^\mu k_\mu = 0$ and producing the phase $e^{ik\cdot x} = e^{i\omega(z-t)}$. The two expected polarizations can be taken as

$$\begin{aligned} \epsilon^1_\mu &= (0,1,0,0) \\ \epsilon^2_\mu &= (0,0,1,0) \end{aligned} \tag{152}$$

which indeed satisfy

$$k^\mu \epsilon^i_\mu = 0 \;, \qquad \epsilon^i_\mu \neq \alpha k_\mu \;. \tag{153}$$

Considering for example the solution with ϵ^1_μ, plugging it into (147), and multiplying with an arbitrary amplitude A_0 one finds

$$\begin{aligned} \vec{A} &= A_0 \cos(\omega z - \omega t)\, \hat{x} \\ \vec{E} &= -\frac{\partial \vec{A}}{\partial t} = E_0 \sin(\omega z - \omega t)\, \hat{x} \\ \vec{B} &= \vec{\nabla} \times \vec{A} = B_0 \sin(\omega z - \omega t)\, \hat{y} \end{aligned} \tag{154}$$

where $E_0 = B_0 = \omega A_0$, and $\hat{x}, \hat{y}, \hat{z}$ the usual unit vectors.

The above plane waves do not carry angular momentum. Plane waves carrying angular momentum are obtained using the circular polarization defined by

$$\epsilon^\pm_\mu = \epsilon^1_\mu \pm i\epsilon^2_\mu \;. \tag{155}$$

They are also said to correspond to the helicity $h = \pm 1$, as in a quantum interpretation they are related to photons carrying angular momentum $\pm\hbar$ along the direction of motion (helicity), and with a wavefunction of the form

$$A_\mu(x) = \epsilon^\pm_\mu(k) e^{ik_\nu x^\nu} = \epsilon^\pm_\mu(k) e^{\frac{i}{\hbar} p_\nu x^\nu} \tag{156}$$

where $p^\mu = \hbar k^\mu$ is the 4-momentum of the photon.

8.2 Gravitational waves and physical polarizations

We can now consider in a similar way the gravitational waves. We have seen that they satisfy the equations

$$\begin{aligned} \Box h_{\mu\nu} &= 0 \\ \partial^\mu h_{\mu\nu} &= \frac{1}{2}\partial_\nu h \end{aligned} \tag{157}$$

with the second one describing the harmonic gauge. Plane wave solution can be found using the ansatz (up to a normalization) by setting

$$h_{\mu\nu}(x) = \epsilon_{\mu\nu}(k)\, e^{ik\cdot x} + c.c. \tag{158}$$

where $\epsilon_{\mu\nu}$ is an arbitrary polarization tensor depending on the wave vector k^μ, and the exponent contains the Lorentz invariant phase $k \cdot x = k_\mu x^\mu = \eta_{\mu\nu} k^\mu x^\nu = -k^0 x^0 + \vec{k} \cdot \vec{x}$. The notation *c.c.* stands for complex conjugation, and makes the solution real. Plugging this ansatz into the equations (157), one finds a solution if

$$k^\mu k_\mu = 0 \ , \qquad k^\mu \epsilon_{\mu\nu}(k) = \frac{1}{2} k_\nu \epsilon^\sigma{}_\sigma \ . \tag{159}$$

Thus, only 6 polarizations $\epsilon_{\mu\nu}(k)$ are possible. However, 4 of these polarizations, the ones with $\epsilon_{\mu\nu}(k) \sim k_\mu \epsilon_\nu(k) + k_\nu \epsilon_\mu(k)$ for some $\epsilon_\mu(k)$ are not physical, and can be removed by gauge transformations. The latter have the form in (141), but with ξ_μ of the form

$$\xi_\mu(x) \sim \epsilon_\mu(k) e^{ik\cdot x} \tag{160}$$

which satisfies $\Box \xi_\mu(x) = 0$, and thus does not ruin the harmonic gauge condition (141). The gauge transformation becomes

$$\delta h_{\mu\nu} = \partial_\mu \xi_\nu + \partial_\nu \xi_\mu \sim i(k_\mu \epsilon_\nu(k) + k_\mu \epsilon_\nu(k))\, e^{ik\cdot x} \tag{161}$$

and shows that these types of polarizations are not physical, and can be removed by an appropriate gauge transformations. Thus, only two physical polarizations remain.

Let us exemplify this again by considering the motion along the z axis. We can take

$$k^\mu = (k^0, \vec{k}) = (\omega, 0, 0, \omega) \tag{162}$$

which solves $k^\mu k_\mu = 0$ and gives the phase $e^{ik\cdot x} = e^{i\omega(z-t)}$. The two expected polarizations can be taken as (using the previous em polarizations)

$$\begin{aligned} \epsilon^{\oplus}_{\mu\nu} &= \epsilon^1_\mu \epsilon^1_\nu - \epsilon^2_\mu \epsilon^2_\nu \\ \epsilon^{\otimes}_{\mu\nu} &= \epsilon^1_\mu \epsilon^2_\nu + \epsilon^2_\mu \epsilon^1_\nu \end{aligned} \tag{163}$$

which indeed satisfy

$$k^\mu \epsilon^i_{\mu\nu} = 0 \ , \qquad \epsilon^i_{\mu\nu} \neq \alpha(k_\mu \epsilon_\nu + k_\nu \epsilon_\mu) \tag{164}$$

for $i = (\oplus, \otimes)$. Considering for example the solution with $\epsilon^{\oplus}_{\mu\nu}$, plugging it into (158), and multiplying with an arbitrary amplitude h_0 one finds

$$h_{\mu\nu}(z - t) = h_0 \cos(\omega z - \omega t)\, \epsilon^{\oplus}_{\mu\nu} \tag{165}$$

which inserted into the linearized metric $g_{\mu\nu}(x)$ give the line element

$$\begin{aligned} ds^2 &= (\eta_{\mu\mu} + h_{\mu\nu}(z-t)) dx^\mu dx^\nu \\ &= -dt^2 + (1 + h_{11}(z-t)) dx^2 + (1 - h_{11}(z-t)) dy^2 + dz^2 \end{aligned} \tag{166}$$

which is interpretable as deforming periodically invariant lengths as in the figure 1 (from [2]).

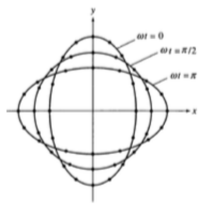

Figure 1: Polarization $\epsilon^{\oplus}_{\mu\nu}$

The polarization $\epsilon^{\otimes}_{\mu\nu}$, does much of the same, but rotated by 45 degrees, see fig. 2

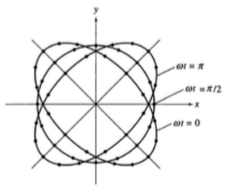

Figure 2: Polarization $\epsilon^{\otimes}_{\mu\nu}$

9 The Schwarzschild solution

Finding exact solutions of the Einstein's field equations is very difficult. One strategy is to use conjectured symmetries of possible solutions, and use these symmetries to restrict the functional form of the metric that is expected to solve the equations. This simplifies Einstein's equations, which then become more tractable and hopefully solvable.

This strategy is the one adopted for finding the Schwarzschild solution. The Schwarzschild metric is obtained by asking for a *static* and *isotropic* solution of the Einstein equations in vacuum, a situation that is realized outside a source that is supposed to be spherical symmetric and static. To implement the required symmetries, time translation and rotational invariance, one assumes the existence of coordinates $x^\mu = (t, \vec{x})$ such that the metric takes the form

$$ds^2 = -F(r)\, dt^2 + 2E(r)\, dt\, \vec{x}\cdot \vec{dx} + D(r)\, (\vec{x}\cdot d\vec{x})^2 + C(r)\, d\vec{x}\cdot d\vec{x} \tag{167}$$

where $r = \sqrt{\vec{x}\cdot\vec{x}}$. This is the most general ansatz consistent with the symmetries. The form of the metric can be further simplified by making changes of coordinates. First of all, one may pass to spherical coordinates (r,θ,ϕ) for $\vec{x}$, and using $\vec{x}\cdot\vec{dx} = rdr$ one rewrites

$$ds^2 = -F(r)\, dt^2 + 2E(r) r\, dtdr + D(r) r^2 dr^2 + C(r)\, [dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2]. \tag{168}$$

Then, one may redefine the time by

$$t \to t' = t + \Phi(r) \tag{169}$$

so that

$$dt' = dt + \frac{d\Phi(r)}{dr} dr \tag{170}$$

and the first two terms inside ds^2 become

$$ds^2 = -F(r)\left(dt' - \frac{d\Phi(r)}{dr}dr\right)^2 + 2E(r) r\left(dt' - \frac{d\Phi(r)}{dr}dr\right) dr + \dots \tag{171}$$

that rearranges to

$$ds^2 = -F(r)\, dt'^2 + 2\Big[rE(r) + F(r)\frac{d\Phi(r)}{dr}\Big] dt' dr - \Big[F(r)\Big(\frac{d\Phi(r)}{dr}\Big)^2 + 2rE(r)\frac{d\Phi(r)}{dr}\Big] dr^2 + \dots \tag{172}$$

Now one can fix the function $\Phi(r)$ to satisfy

$$\frac{d\Phi(r)}{dr} = -\frac{rE(r)}{F(r)} \tag{173}$$

so that the mixed term $dt'dr$ vanishes, and the remaining part takes the form

$$ds^2 = -F(r)\, dt'^2 + G(r) dr^2 + C(r)\, [dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta\, d\phi^2] \tag{174}$$

where

$$G(r) = r^2\Big(D(r) + \frac{E^2(r)}{F(r)}\Big) \,. \tag{175}$$

Now one could redefine the radius $r \to r'$ by setting

$$r'^2 = C(r)r^2 \tag{176}$$

so that one gets the so-called *standard form* of the metric

$$ds^2 = -B(r')\,dt'^2 + A(r')dr'^2 + r'^2(d\theta^2 + \sin^2\theta\,d\phi^2) \tag{177}$$

with

$$\begin{aligned} B(r') &= F(r) \\ A(r') &= \Big(1 + \frac{G(r)}{C(r)}\Big)\Big(1 + \frac{r}{2C(r)}\frac{dC(r)}{dr}\Big)^{-2} \,. \end{aligned} \tag{178}$$

Dropping the primes one finds the static and isotropic metric in the standard form

$$\boxed{ds^2 = -B(r)\,dt^2 + A(r)dr^2 + r^2(d\theta^2 + \sin^2\theta\,d\phi^2)} \,. \tag{179}$$

It is put into Einstein's equations, which are solved to produce the Schwarzschild solution

$$\boxed{ds^2 = -\Big(1 - \frac{2GM}{r}\Big)dt^2 + \Big(1 - \frac{2GM}{r}\Big)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta\,d\phi^2)} \,. \tag{180}$$

The same solution is obtained by relaxing the hypothesis of time invariance (staticity). A more general ansatz for the solution still lead to the same Schwarzschild metric. This is captured by *Birkhoff's theorem*, that states that any spherically symmetric solution of the vacuum field equations must be static and asymptotically flat. This theorem guarantees that the assumption of staticity may be dropped, and still the exterior solution for the spacetime metric outside of a spherical, nonrotating, gravitating body must be given by the Schwarzschild metric.

10 Black holes

The Schwarzschild solution indicates the existence of an event horizon and leads to the concept of a black hole. The recommended treatment is the one presented in [2], see chapter 8.

References

[1] S. Weinberg, "Gravitation and Cosmology", John Wiley & Sons, 1972.

[2] H. Ohanian and R. Ruffini, "Gravitation and Spacetime", Cambridge University Press, 2013.