University of Trento — anno accademico 2015-2016

General Relativity and Cosmology

Guido Cognola *

The present lecture notes do not substitute the recommended text-books

*e-mail: *guido.cognola@unitn.it*

Contents

Text books

general relativity:

- S. Weinberg, *Gravitation and Cosmology*, John Wiley & Sons, New York (1972).
- L.D. Landau, E.M. Lifshitz, *Teoria dei Campi*, Editori Riuniti, Roma (1976).

cosmology:

- P.J. Peebles, *Principles of Physical Cosmology.* Princenton University Press (1993).
- V. Mukhanov. *Physical Foundations of Cosmology.* Cambridge University Press (2005).
- S. Weinberg, *Cosmology.* Oxford University Press, New York, (2008).

quantity	symbol	value
speed of light in vacuum	c	$2.998 \times 10^{10}\ cm/sec$
Newton constant	G	$6.673 \times 10^{-8}\ cm^3/g\ sec^2$
Planck constant	h	$6.625 \times 10^{-27}\ erg\ sec$
electron charge	e	$4.80 \times 10^{-10}\ \ esu$

Table 1: **universal constants**

quantity	symbol	value
astronomic unit	u.a.	$1.49598 \times 10^{8}\ Km$
parsec	pc	3.2615 light-year $3.0856 \times 10^{13}\ Km$
Hubble constant	H_0	$70\ (Km/sec)\ /Mpc$

Table 2: **astronomic quantities**

Notations and definitions

Indices: Latin indices $i, j, k, ...$ and Greek indices $\alpha, \beta, ...\mu, \nu, ...$ assume the values $0, 1, 2, 3$, while Latin indices $a, b, c, ...$ assume the values $1, 2, 3$ or $2, 3$ as specified in the text. As usual, it is implicitly assumed that repeated indices are summed on the range of they values.

Signature: for the metric we shall use the signature -,+,+,+, which is convenient in the limit processes.

Units: in the last part of the lecture we shall use units in which the speed of light $c = 1$. By a dimensional analysis, the final equations could be written in standard units.

	diameter in Km	mass in g	density in g/cm^3
sun	1392000	2×10^{33}	1.4
earth	12757	6×10^{27}	5.5

Table 3: **sun and earth**

planet	diameter	M/M_T	revolution	rotation	distance from the sun	eccentricity
mercury	4878	0.055	88	59	58	0.2060
venus	12180	0.815	225	243	108	0.0070
earth	12757	1	365	1	149.598	0.0167
mars	6787	0.107	687	1	228	0.0934
jupiter	142200	317.89	4329	0.41	778	0.0485
saturn	119300	95.17	10753	0.43	1427	0.0556
uranus	51200	14.60	30660	0.70	2870	0.0472
neptune	49500	17.2	60152	0.75	4497	0.0086
pluto	2290	0.0022	90410	6.3	5900	0.2500

Table 4: **planets**

Solar system

The *astronomic unit* corresponds to the mean distance between sun and earth. In the tables 3 and 4 there are some (approximated) values of quantities related to the solar system. Units measure are the following: diameter in Km, mass in comparison to the earth mass, period of revolution and rotation in days, mean distance from the sun in $MKm = 10^6 Km$.

We recall that:

- *Eccentricity* e of an ellipse: is the ratio between the focal distance and the major axes. Its value is in the range $[0, 1]$. 0 corresponds to the circle and 1 to the straight line (limit case).
- *Trigonometric parallax* of a star P: is the maximum angle under which the star "sees" the earth-sun system during the whole year. (It is the angle in P of the isosceles triangle built up with the three points star-earth-sun).
- *parsec*: is the distance at which the star has a parallax equal to a second of degree.

1 Introduction

Here we give a schematic description of the ideas concerning physics and geometry, which had great importance in the theoretical development of General Relativity.

The Geometry: from Euclid to Riemann

The Euclidean geometry (Euclid, 325 b.c.), is based on five postulates, one of which (the fifth) (**the parallel postulate**: *given a straight line and an external point, it exists a second straight line, parallel to the first and crossing the given point*), always has been accepted with "extremely care" (from Euclid himself) and many attempts have been made in order to derive it from the previous four, the last one due to Legendre[1].

In the year 1733, the Jesuit Saccheri[2] tries to construct geometries without the use of the parallel postulate, but the so called non-Euclidean geometries effectively appear in the 19^{th} century after the works of the three mathematicians: Gauss[3], Lobachevskii[4] and Bolyai[5], who independently discover the geometry of the hyperbolic plane, which is a surface with negative, constant curvature and for which the fifth postulate does not hold. Such a surface can not be immersed in a $3-dimensional$ space and so it is not possible to see it (as the sphere). It is only possible to characterise it by an intrinsic geometric description.

It is interesting to observe that Gauss proposed an experiment in order o verify which geometry corresponds to the physical $3-dimensional$ space. In fact, using a theodolite, he measured the internal angles of a big triangle, the vertex being the tops of three mountains in Bajern. The result was $180°$ in agreement with Euclidean geometry. It is also remarkable that he implicitly assumed that the light moved on *geodesics* (the curves of minimal distance between two points).

The non Euclidean geometry originally formulated for the surface, has been extended by Riemann[6] to spaces with arbitrary dimensions (Riemannian geometry) and then developed by Christoffel[7], Ricci[8], Bianchi[9] and Levi-Civita[10].

The Physics: from Galileo to Einstein

The Newtonian mechanics[11] is based on a privileged family of reference frames (*inertial frames*), in which the physical laws do not change. (*Galileo Relativity principle-Galileo Transformations*[12]). All inertial reference systems are moving with constant velocity with respect to the *absolute space*, its existence being assumed "a priori".

- *The bucket experiment.* Newton himself proposed an experiment in order to verify the existence of the absolute space. The bucket full of water rotates around its vertical axis and one observes

[1] Adrien Marie Legendre (France) 1752-1833.
[2] Giovanni Girolamo Saccheri (Italia) 1667-1733.
[3] Johann Carl Friedrich Gauss (Germany) 1777-1855.
[4] Nikolai Ivanovich Lobachevskii (Russia) 1792-1856.
[5] János Bolyai (Romania) 1802-1860.
[6] Georg Friedrich Bernhard Riemann (Germany) 1826-1866.
[7] Elwin Bruno Christoffel (Germany) 1829-1900.
[8] Gregorio Ricci-Curbastro (Italia) 1853-1925.
[9] Luigi Bianchi (Italia) 1856-1928.
[10] Tullio Levi-Civita (Italia) 1873-1941.
[11] Isaac Newton (England) 1642-1727.
[12] Galileo Galilei (Italia) 1564-1642.

that the water assumes a parabolic outline when it also rotates with respect to the absolute space, the rotation with respect to the bucket being not important.

Leibniz[13] with other people criticises the Newtonian hypothesis because there are no physical (philosophical) reasons for the introduction of the absolute space. According to him only the relations between the material bodies have an intrinsic physical meaning.

The first constructive attack to the Newtonian hypothesis has been done by Mach[14]. According to him, all celestial bodies have an influence on the reference systems. This means that the water in the bucket assumes the parabolic form when it rotates with respect to the other bodies in the universe and not with respect to a hypothetic absolute space. Mach himself proposed an experiment in order to measure the "inertial forces" which act on a massive body at rest inside a rotating massive shell (*Mach principle*). General relativity gives a "partial answer" to the Mach principle (Lense-Thirring effect), but the problem is still open.

The Galilean relativity principle (more precisely the Galileo transformations) starts to be criticised with the formulation of electromagnetism [15], because electrodymamics laws do not change according to Galileo transformations when one changes the inertial frame. For example, electromagnetic waves move with a constant velocity (the speed of light) independently on the observer or the source. In order to preserve the relativity principle and a constant speed of light it is necessary to change the transformation laws between inertial frames by replacing Galileo transformation with *Lorentz transformations*[16] *(Poincaré*[17]*)* and all physical laws have to be *covariant* with respect to such transformations (*special relativity principle*[18]). Mechanical laws can be easily modified in such a way to satisfy the special relativity principle, but this does not happen for the Newton gravitational law.

The General Relativity

It is a relativistic theory of gravitation, which has been built up in order to satisfy purely theoretic and philosophic requirements (at that time no experimental measure justified the abandonment of Newtonian gravitation law) and which unifies the inertial forces with the gravitational ones (according to Mach ideas), by extending the relativity principle from the inertial to all reference systems.

General relativity can also be seen as a geometrical theory because it is strictly related to the geometry of space-time. The matter bodies create the geometry of the space-time in which they move.

The construction of the whole theory, due to Einstein, has required about ten years, during which, with the help of his friend Grossmann[19], he studied the tensor calculus and the Riemannian geometry, which are extremely important for the development of the theory. Here the principal steps:

1907: Principle of Equivalence;

1913: Principle of General Relativity; gravitational field described by a tensor;

1915: Final form of General Relativity theory;

[13]Gottfried Wilhelm von Leibniz (Germany) 1646-1716.
[14]Ernst Mach (Mähren)) 1838-1916.
[15]James Clerk Maxwell (Scotland) 1831-1879.
[16]Hendrik Antoon Lorentz (Holland) 1853-1928.
[17]Jules Henri Poincaré (France) 1854-1912.
[18]Albert Einstein (Germany) 1879-1955.
[19]Marcel Grossmann (Hungaria) 1878-1936.

1916: Schwarzschild[20] solution (spherical symmetry);

1917: cosmological models by Einstein and de Sitter[21];

1922: cosmological solution by Friedmann[22];

1930: cosmological model by Friedmann-Lemaître[23]-Robertson[24] -Walker[25];

1963: Kerr[26] Solution (axial symmetry).

[20] Karl Schwarzschild (Germany) 1873-1916.
[21] Willem de Sitter (Holand) 1872-1934.
[22] Aleksandr Aleksandrovich Friedmann (Russia) 1888-1925.
[23] Georges Èdouard Lemaître (France) 1894-1966.
[24] Howard Percy Robertson (USA) 1903-1961.
[25] Arthur Geoffrey Walker (England) 1909-2001.
[26] Roy Patrick Kerr (New Zealand) 1934.

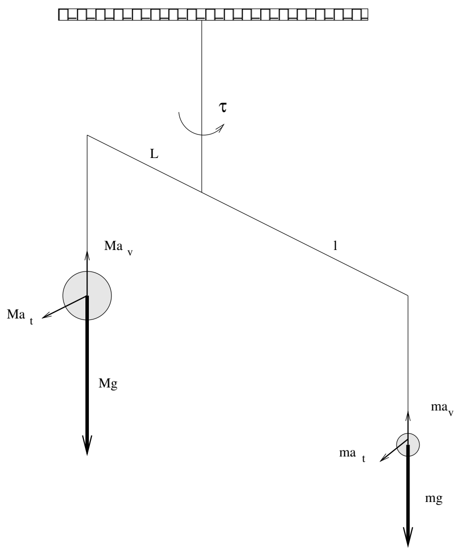

Figure 1: Eötvös experiment

2 The Principle of Equivalence

In General Relativity the equivalence between the mass of a body (*inertial mass*) and its "gravitational charge" (*gravitational mass*) is assumed "a priori" as a fundamental principle. In Newtonian theory such an equivalence is a pure accidental fact, based on the Eötvös[27] experiment (1889), which has a very high precision $(\Delta /m) < 10^{-9}$. 10^{-11} (Dicke, 1960), 10^{-13} (Adelberger et al. , 1999).

Recall that inertial mass m_i and gravitational mass m_g enter dynmical and gravitational laws respectively, that is

$$\vec{F} = m_i \vec{a}\,, \qquad \vec{F}_g = -G\, m_g M_g \vec{r}/r^3\,, \qquad m_i \equiv m_g\,. \tag{2.1}$$

Here $\vec{F}$ is a generic force applied to the particle, while $\vec{F}_g$ is the gravitational force generated by the particles.

2.1 The Eötvös experiment

The apparatus is the one in figure 1. On the two heavy bodies act the gravitational forces, proportional to the gravitational masses m_g, M_g, the inertial forces (due the rotation of the earth), proportional to the inertial masses m_i, M_i and the torque $\mathcal{T}$ due to the presence of the string. When the system is in equilibrium, one has the relation

$$\mathcal{T} = l m_i a_t - L M_i a_t\,, \qquad l(m_g g - m_i a_v) = L(M_g g - M_i a_v)\,,$$

where g is the gravitational acceleration and a_t, a_v the components of centrifugal acceleration, which satisfy $a_t \ll g$, $a_v \ll g$.

From system above it follows

$$\mathcal{T} \sim l a_t m_g \left(\frac{m_i}{m_g} - \frac{M_i}{M_g} \right)\,.$$

The measured value of $\mathcal{T}$ is vanishing and this means that $m_i/m_g \equiv M_i/M_g$, that is, the inertial mass is equivalent to the gravitational mass with an appropriate choice of the unit measure.

[27] Lóránd Baron von Eötvös (Hungaria) 1848-1919.

2.2 Formulation of the principle of equivalence

In every point of space-time and in an arbitrary gravitational field it is always possible to choose a reference frame where, in a "small" region of space-time, *all the laws of nature* are the ones which one has in the absence of gravitation. In other words, locally it possible to choose a Minkowskian reference frame (the locally inertial frame-the frame in free fall). (Is has to be remember that around every point of an arbitrary surface, is is possible to choose coordinates where the Pitagora's theorem holds (normal or Gaussian coordinates).

This formulation of the principle of equivalence is said **strong**, because it is valid for all interactions. There is a **weak** form, restricted to mechanical laws only, which is essentially the same thing as the equivalence between inertial and gravitational mass. Finally there is a third formulation **medium-strong**, which concerns all interactions but the gravitational one. This is due to the fact that the precision of experiment (10^{-13}) tests also the electromagnetic binding energy, but not the corresponding gravitational one.

General relativity is based on the strong principle of equivalence and in the following we shall always consider such a formulation.

2.3 The Motion of a test particle in a gravitational field

As a first application of the principle of equivalence now we derive the differential equation which describe the motion of a free particle in a given gravitational field. We can choose a local inertial frame in which the motion of the particle is the one of Special Relativity. Of course this holds in a small space-time region. If the region can be extended to the whole space-time, the "gravitation" is simply due to "apparent forces" due to the fact that we are using an accelerated frame in the Minkowski space[28]).

If X^μ are the coordinates of the test particle in such a reference, then the equation of motion reads

$$\frac{d^2X^\mu}{c^2\,d\tau^2} = \frac{d^2X^\mu}{ds^2} = 0\,, \qquad ds^2 = -c^2\,d\tau^2 = \eta_{\mu\nu}dX^\mu dX^\nu\,, \qquad \mu,\nu,... = 0,1,2,3,$$

where ds is the invariant interval, $d\tau$ the proper time and $\eta_{\mu\nu}$ the Minkowski metric.

We indicate by x^k the coordinates of the particle in an arbitrary reference frame ($i,j,k,... = 0,1,2,3$). We assume $x^k = x^k(X^\mu)$ to be a "smooth" function (continuous, differentiable, invertible, etc..). Then we can set

$$A^k_\mu \equiv A^k_\mu(X) = \frac{\partial x^k}{\partial X^\mu}\,, \qquad B^\mu_k \equiv B^\mu_k(x) = \frac{\partial X^\mu}{\partial x^k}\,, \qquad A^i_\mu B^\mu_j = \delta^i_j\,, \qquad A^k_\mu B^\nu_k = \delta^\nu_\mu\,. \tag{2.2}$$

The interval ds is a scalar. It has the same value in all reference frames, then we get

$$ds^2 = \eta_{\mu\nu}dX^\mu dX^\nu = g_{rs}dx^r dx^s\,, \qquad g_{rs}(x) = \eta_{\mu\nu}B^\mu_r B^\nu_s\,. \tag{2.3}$$

$g_{rs}(x)$ is the *metric tensor*, which takes into account of the presence of a gravitational field as well as of the choice of the coordinates. It is equivalent to the Minkowski metric only in a local inertial frame.

For the equation of motion we get (we use compact notation $\partial_k = \partial/\partial x^k$)

$$0 = \frac{d^2X^\mu}{ds^2} = \frac{d}{ds}\left(B^\mu_i\,\frac{\partial x^i}{ds}\right) = B^\mu_i\,\frac{\partial^2 x^i}{ds^2} + \partial_i B^\mu_j\,\frac{\partial x^i}{ds}\,\frac{\partial x^j}{ds}\,.$$

[28]Hermann Minkowski (Lithuania) 1864-1909.

Finally, multiplying by the inverse matrix A^k_μ we obtain the equation of motion of the test particle in the arbitrary reference frame, that is

$$\frac{d^2x^k}{ds^2} + \hat{\Gamma}^k_{ij}\,\frac{\partial x^i}{ds}\,\frac{\partial x^j}{ds} = 0\,, \qquad \text{geodesic equation.} \tag{2.4}$$

$$\hat{\Gamma}^k_{ij} = \hat{\Gamma}^k_{ji} = A^k_\mu\,\partial_j B^\mu_i = -B^\mu_i\,\partial_j A^k_\mu = \frac{\partial x^k}{\partial X^\mu}\,\frac{\partial^2 X^\mu}{\partial x^i \partial x^j}\,, \qquad \text{connexion.} \tag{2.5}$$

The quantity $\hat{\Gamma}^k_{ij} = \hat{\Gamma}^k_{ji}$ (*Christoffel symbols* or *Levi-Civita connexion* [29]) plays the role of "gravitational force". It is completely determined by g. This can be seen by deriving the expression of the metric and using the symmetry properties of the connexion. We have

$$\partial_k g_{ij} = \eta_{\mu\nu}\left[\partial_k B^\mu_i\,\partial_s X^\nu + \partial_k B^\nu_s\,\partial_j X^\mu\right]$$

and since

$$\eta_{\mu\nu}\,B^\nu_i = g_{ij}\,A^j_\mu\,,$$

we also obtain

$$\partial_k g_{ij} = \hat{\Gamma}^l_{ki} g_{lj} - \hat{\Gamma}^l_{kj} g_{li}\,. \tag{2.6}$$

In order to invert the latter expression we write it in three different ways obtained by cyclic permutation of the indices, that is

$$\partial_i g_{jk} = \hat{\Gamma}^l_{ij} g_{lk} - \hat{\Gamma}^l_{ik} g_{lj}\,, \tag{2.7}$$

$$\partial_j g_{ki} = \hat{\Gamma}^l_{jk} g_{li} - \hat{\Gamma}^l_{ji} g_{lk}\,. \tag{2.8}$$

By summing (2.7), (2.8), subtracting (2.6) and taking into account of the symmetries properties of $\hat{\Gamma}^k_{ij}$ we finally get

$$\hat{\Gamma}^k_{ij} = \frac{1}{2}\,g^{kl}\left(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}\right)\,. \tag{2.9}$$

Note: The evolution parameter in the geodesic equation (2.4) is not really important in the sense that it will be expressed in terms of the time t when the equations will be explicitly solved. This means that one can use an arbitrary evolution parameter and not necessary the proper time. Using an arbitrary evolution parameter σ, the geodesic equation (2.4) holds also for massless particles for which the proper time is vanishing. It has to be remarked that (2.4) has that form only for parameters related to proper time by an affine transformation (linear transformation).

2.3.1 The Newtonian limit

In order to understand the physical meaning of metric and connexion, on the geodesic equation we perform the *Newtonian limit*. This means that we consider "small velocities" and weak fields independent on time. Then

$$\left|\frac{dx^a}{ds}\right| \ll \left|\frac{dx^0}{ds}\right|\,, \qquad a = 1,2,3\,, \qquad g_{ij} = \eta_{ij} + h_{ij}\,, \qquad |h_{ij}| \ll 1\,, \qquad \partial_0 h_{ij} = 0\,.$$

[29] Tullio Levi-Civita (Italia) 1873-1941.

Using such an approximation the geodesic equation becomes

$$\frac{d^2x^k}{ds^2} + \hat{\Gamma}^k_{00}\left(\frac{dx^0}{ds}\right)^2 \sim 0\,, \qquad \hat{\Gamma}^k_{00} \sim -\frac{1}{2}\, g^{kl}\partial_l g_{00} \sim -\frac{1}{2}\,\eta^{kl}\partial_l h_{00}\,,$$

from which

$$\frac{d^2x^0}{ds^2} \sim -\hat{\Gamma}^0_{00}\left(\frac{dx^0}{ds}\right)^2 \sim 0\,, \qquad \frac{d^2x^a}{ds^2} \sim -\hat{\Gamma}^a_{00}\left(\frac{dx^0}{ds}\right)^2 = \frac{1}{2}\,\eta^{al}\partial_l h_{00}\left(\frac{dx^0}{ds}\right)^2\,.$$

From the first equation above it follows that $x^0 \sim ct$ (t being the time of Newtonian's theory), while the second equation can be set in the form

$$\frac{d^2\vec{x}}{dt^2} = \frac{1}{2}\,c^2\vec{\nabla}h_{00} \tag{2.10}$$

and has to be compared with the analog Newtonian equation

$$\frac{d^2\vec{x}}{dt^2} = -\vec{\nabla}\phi\,, \qquad \phi = -\frac{MG}{r}\,, \tag{2.11}$$

ϕ being the potential (Newtonian) for unit mass, G the gravitational constant and M the mass of the body which generates the gravitational field. We see that (2.10) gives the correct limit (2.11) if

$$g_{00} \sim -\left(1 + \frac{2\phi}{c^2}\right)\,. \tag{2.12}$$

All components g_{ij} of the metric are (remarkable) different from η_{ij} only in the presence of high gravitational fields. For example, on the surface of the earth one has $h_{00} \sim 10^{-6}$.

2.4 Coordinate systems, reference frame and measurable quantities

As it is clear from examples above, the gravitational field modifies the metric tensor g_{ij} and as a consequence, the geometry of space-time will be different from the Minkowskian one. In general it will be a Lorentzian (pseudo-Riemannian) geometry. Only locally it will be possible to choose a coordinate system where special relativity holds.

The coordinates can be chosen in an arbitrary way and they do not have a direct physical meaning. What are physical relevant are invariant quantities (scalars) with respect to general coordinate transformations.

In the following we shall talk about *coordinate system* and *reference system* without distinction, but it has to be remarked that at any reference frame, realised by physical objects (bars and clocks) corresponds a coordinate system, but in general it is not true the "vice-versa". The coordinate system has to satisfy some constraints in order to be realised by physical objects.

Given a coordinate system $\{x^k\}$, the invariant interval is given by

$$ds^2 = g_{ij}dx^i dx^j\,,$$

where g_{ij} depends (in general) on x^k. The fact that g_{ij} is different from Minkowski metric does not necessarily mean that we are in the presence of a gravitational field. For example, the Minkowski interval in a rotating reference frame with angular velocity ω around the z axis assumes the form

$$ds^2 = -c^2dt^2 + dr^2 + r^2d\tilde{\varphi}^2 + dz^2 = -(c^2 - \omega^2r^2)dt^2 + dr^2 + r^2d\varphi^2 + 2\omega r^2 dt d\varphi + dz^2\,, \tag{2.13}$$

where $(r, \tilde{\varphi}, z)$ are the coordinates of the point P attached to the rotating body, while $(r, \varphi = \tilde{\varphi} - \omega t, z)$ are the coordinates of the same point as seen by an observer on the ground. According to the principle of equivalence the metric in (2.13) can be seen as due to a "factious gravitational field".

The proper time $d\tau$ is the one measured by a clock at rest with the event, that is

$$d\tau = -\frac{ds}{c} = \frac{\sqrt{|g_{00}|\, dx^0}}{c} \, .$$

For the specific example above we have

$$d\tau = \sqrt{1 - \frac{\omega^2 r^2}{c^2}}\, dt \, , \tag{2.14}$$

in agreement with the the formula of Special Relativity (the clock in P moves with velocity ωr with respect to the observer on the ground O. At any time we can attach at P an inertial reference frame O' with velocity ωr with respect to O and apply the formulae of Special relativity).

Using the principle of equivalence, the formula (2.14) can be also read in a different way. We can think about two different clocks in the points $P_1 \equiv (r, \tilde{\varphi}, z)$ and $P_2 \equiv (r, \varphi, z)$ of a "gravitational field" given by the metric in (2.13). In such a case the ratio between the proper times measured by the two clocks reads

$$\frac{d\tau_2}{d\tau_1} = \frac{\sqrt{|g_{00}(P_2)|}}{\sqrt{|g_{00}(P_1)|}} \, .$$

For the principle of equivalence we expect such a result to be valid in a generic gravitational field.

Note: in general the integration of the proper time does not have a physical meaning because $d\tau$ is not an exact differential form and so its integration depend on the path. We can integrate $d\tau$ in order to obtain the finite duration of a phenomena if the initial and final events have the same spatial coordinates. On the contrary, it is always possible to integrate the coordinate time dt because it is an exact differential form.

2.5 Time dilation

We consider an arbitrary metric independent on time x^0, two events $E_1 \equiv (x^0, \vec{x})$, $E_2 \equiv (x^0 + dx^0, \vec{x})$ at the same spatial point and two observers $P_1 \equiv (x_1^k)$, $P_2 \equiv (x_2^k)$ at rest with respect to $\vec{x}$.

The coordinate interval between the two events is dx^0 for all the observers, but the interval measured in proper time depends on the observer. We have

$$\frac{d\tau_2}{d\tau_1} = \frac{\sqrt{|g_{00}(P_2)|}}{\sqrt{|g_{00}(P_1)|}}. \tag{2.15}$$

This means that the rate of a clock depends on the gravitational field. In particular, the clock in the presence of gravitation is always in late with respect to the clock in the absence of gravitation. Such an effect has been measured by astronomical and also terrestrial experiments. It has not to be confused with time dilation of Special Relativity, where the effect is due to the relative motion of the two clocks. Of course, if these are moving in a gravitational field, both the effects have to be taken into account.

2.6 Spatial distance and spatial geometry

Now we are going to define the spatial distance between two infinitely closed points A and B. To this aim we consider a light signal living the point A at the time $x^0 + dx^0_1$, reflecting in B at time x^0 and arriving again in A at time $x^0 + dx^0_2$. For the light one has $ds^2 = 0$, which is a second order algebraic equation which permits to derive $dx^0_{1,2}$ in terms of the metric. In fact

$$dx^0_{1,2} = \frac{1}{g_{00}} \left[-g_{0a}dx^a \mp \sqrt{(g_{0a}g_{0b} - g_{00}g_{ab})dx^a dx^b} \right] , \qquad a = 1,2,3,$$

from which

$$\Delta x^0 = dx^0_2 - dx^0_1 = \frac{2}{g_{00}} \sqrt{(g_{0a}g_{0b} - g_{00}g_{ab})dx^a dx^b} , \qquad \Delta\tau = \frac{\sqrt{|g_{00}|}}{c}\Delta x^0 = \frac{2}{c} \sqrt{\gamma_{ab}dx^a dx^b} ,$$

where

$$\gamma_{ab} = g_{ab} - \frac{g_{0a}g_{0b}}{g_{00}} , \qquad \gamma^{ab} = g^{ab} , \qquad \det g = g_{00} \det \gamma$$

and $\Delta\tau$ is the distance (in proper time) between two events in the same spatial point (where the signal starts and arrives).

Now it is natural to define the spatial distance between the points A and B by mean of the equation

$$\Delta\sigma = \frac{c\Delta\tau}{2} \qquad \Longrightarrow \qquad d\sigma^2 = \gamma_{ab}dx^a dx^b .$$

γ_{ab} defines the spatial geometry. Note that if the metric depends on time, then it has no meaning to integrate $d\sigma$ in order to find the distance between far points, because this depends on time.

2.7 Some properties of the metric

The metric of a gravitational field in a physical reference frame must satisfy the following properties:

$\det g < 0$:	real gravitational field ,
$g_{00} < 0$:	physical reference frame built up with material bodies ,
γ_{ab} :	positive quadratic form .

The third condition assures the spatial distance to be always positive.

2.8 Static and stationary fields

A gravitational field is said to be *constant* if it is possible to choose a coordinates system in which g_{ij} does not depend on the coordinate time x^0. More precisely, a constant field is said *static* if $g_{0a} = 0$ and *stationary* if $g_{0a} \neq 0$ $(a = 1,2,3)$. In the latter case the metric is not invariant with respect to time inversion. A static field for which $g_{00} = 1$ is said *ultrastatic* and the corresponding reference frame is said *synchronous*, because for such frames it is possible to synchronise the clocks. Moreover in such frames the coordinate $t = x^0/c$ represents the proper time and the lines $t = constant$ are the geodesic on the 3-dimensional section because $ds^2 = -(dx^0)^2 + \gamma_{ab}dx^a dx^b$.

- In any gravitational field it is always possible to choose coordinates for which $g_{0a} = 0$, but in general these depend on time.

2.9 The shift of spectral lines

Analysing the light arriving from distant stars we can read off atomic spectra similar to the ones which we observe on the earth, but shifted (to the red), the shift depending also on the frequency of the spectral line. Such a phenomena is a direct consequence of time dilation and represents a confirmation of the principle of equivalence. Here we shall give three different derivations of such an important effect, two of them explicitly based on the principle of equivalence,

2.9.1 Derivation I

Let us consider a *stationary* gravitational field (this means that exists a coordinates system in which the metric does not explicitly depend on the coordinate time $x^0 = ct$) and an atom in the point S (the source) which performs a transition and emits photons with proper frequency ν_0 (by definition, the proper frequency is the one measured by a clock at rest with the atom).

In the period $\Delta\tau_S = n/\nu_0$ the atom will emit a number n of waves and these will be received in the point O (the observer) during the period $\Delta\tau_O$. For the observer in O such waves will have a frequency $\nu = n/\Delta\tau_O$. Because the field is stationary, the (coordinate) travel time Δt is the same for all the observers and in particular $\Delta t_S = \Delta t_O$. In this way

$$\frac{\Delta\tau_S}{\Delta\tau_O} = \frac{\sqrt{|g_{00}(S)|}}{\sqrt{|g_{00}(O)|}} \qquad \Longrightarrow \qquad \frac{\nu}{\nu_0} = \frac{\sqrt{|g_{00}(S)|}}{\sqrt{|g_{00}(O)|}} \, .$$

For weak fields $g_{00} \sim -(1 + 2\phi/c^2)$ and so

$$\frac{\Delta\nu}{\nu_0} \sim \frac{\Delta\phi}{c^2} \, , \qquad \Delta\nu = \nu - \nu_0 \, , \qquad \Delta\phi = \phi(S) - \phi(O) \, . \tag{2.16}$$

According to the latter formula, the frequency undergoes a shift (to the red in the case sun-earth because $\Delta\phi < 0$) which is proportional to the frequency itself. Note that it is assumed that the proper frequency is independent on the gravitational field and in some sense this can be considered as a definition of standard clock.

It has to be remarked that it is quite difficult to measure such an effect, because it is "covered" by the Doppler effect[30] due to the motion of the source, the thermal motion which spread the spectral lines and the convective motion of the solar atmosphere.

Moreover, the formula in (2.16), as it stands, connot be used for the sun-earth system, because it has been derived for two points at rest in a stationary gravitational field, but the observer on the earth is in "free fall" with the earth itself in the gravitational field generated by the sun, then the formula has to be corrected by taking into account of doppler effect due to the motion of the observer too.

There are however experiments prepared in the laboratory (Pound e Rebka (1960)), which agrees with (2.16) and confirms the principle of equivalence.

- The formula (2.16) can be also derived by using a semiclassical treatment of photon and the conservation of energy. In fact one has

  $$h\nu_0 \left(1 + \frac{\phi(S)}{c^2}\right) = h\nu \left(1 + \frac{\phi(O)}{c^2}\right) \, ,$$

  from which equation (2.16) directly follows. Here h is the Planck's constant.

[30] Christian Andreas Doppler (Österreich) 1803-1853.

2.9.2 Derivation II

Let us consider an atom at P (the sourse) on a rotating body with angular velocity ω, with respect to an inertial observer O.

At a generic time t, the point P overlaps a point P' of an inertial system O', which moves with velocity ωr with respect to O, r being the radial coordinate of P in the system $O(t, r, \vartheta, \varphi.)$. According to Special Relativity one has

$$d\tau = dt\sqrt{1 - \frac{\omega^2 r^2}{c^2}} = dt\sqrt{1 + \frac{2\phi_c}{c^2}}\,, \qquad \phi_c = -\frac{\omega^2 r^2}{2}\,,$$

where ϕ_c is the centrifugal potential, $d\tau$ the proper time for the observer in P' and dt the proper time for the observer O. During the period $d\tau$ the atom emits n waves with frequency ν_0 and these are received by O during the period dt. Then

$$d\tau\nu_0 = n = dt\,\nu \quad \Longrightarrow \quad \frac{\Delta\nu}{\nu_0} \sim \frac{\phi_c}{c^2} = \frac{\Delta\phi_c}{c^2}\,.$$

Using the principle of equivalence, the result can be extended to an arbitrary gravitational field.

2.9.3 Derivation III

Let us consider an atom in the terrestrial gravitational field situated at a given quote $z = d$ with respect to the ground. The atom emits radiation with proper frequency ν_0 and this arrives at $z = 0$ with frequency ν,

By means of the principle of equivalence, the gravitational field can be "replaced" by a field of inertial forces $\vec{a} = g\vec{k}$ generated by an accelerated lift, $\vec{k}$ being a unit vector. For the observer inside the lift, the signal is emitted at $t = 0$, with frequency ν_0, when the velocity of the atom is zero and it is received at $t \sim d/c$, with frequency ν, when the velocity of the lift is $\vec{v} \sim (gd/c)\vec{k}$. Due to Doppler effect $\nu \neq \nu_0$. The Doppler formula for the general case reads

$$\frac{\nu}{\nu_0} = \frac{\sqrt{1 - v^2/c^2}}{1 - v\cos\alpha/c}\,,$$

where $v = |\vec{v}|$ and α is the angle between the direction of the signal and the velocity of the detector. In this particular case $\alpha = -\pi$ and so

$$\frac{\nu}{\nu_0} = \frac{\sqrt{1 - v^2/c^2}}{1 + v/c} \sim 1 - \frac{v}{c} \quad \Longrightarrow \quad \frac{\Delta\nu}{\nu_0} = -\frac{gd}{c^2} = \frac{\Delta\phi_g}{c^2}\,,$$

$\Delta\phi_g$ being the difference of fields between source and detector.

3 Tensor Analysis

In the following we shall deal with quantities that do not change under coordinate transformations, like pure numbers and functions, but also with quantities that change when we pass from a reference frame to another one, like vectors and tensors. In the present context, the simplest way to define tensors is by means of the transformation rules of their components with respect to general coordinate transformations.

To this aim, in a $N-$dimensional manifold M^N we consider two coordinate systems $x \equiv \{x^k\}$ and $\tilde{x} \equiv \{\tilde{x}^k\}$, $(k = 0, 1, 2, ...N-1)$ the two matrices which relate the two systems

$$a^i_j = \frac{\partial \tilde{x}^i}{\partial x^j}\,, \qquad b^i_j = \frac{\partial x^i}{\partial \tilde{x}^j}\,, \qquad\qquad a^i_k b^k_i = \delta^i_j\,, \qquad b^i_k a^k_i = \delta^i_j\,. \tag{3.1}$$

and the components of tensors (intrinsic invariant quantities) in both the coordinate systems. According to the following transformation rules, we call *scalars*, *vectors* and *tensors* the quantities we are dealing with. More precisely

tensor	**transformation rules**
scalar ϕ	does not change, that is $\tilde{\phi}(\tilde{x}) = \phi(x)$
controvariant vector V^k	$\tilde{V}^k = a^k_j V^j$
covariant vector V_k	$\tilde{V}_k = b^j_k V_j$
controvariant tensor of order p (p indices)	$\tilde{T}^{ijrs...} = a^i_{i'} a^j_{j'} a^r_{r'} a^s_{s'} ... T^{i'j'r's'...}$
covariant tensor $T_{ijrs...}$ *of order q (q indices)*	$\tilde{T}_{ijrs...} = b^{i'}_i b^{j'}_j b^{r'}_r b^{s'}_s ... T_{i'j'r's'...}$
mixed tensor $T^{ij..}_{rs..}$ *of order* (p, q)	$\tilde{T}^{ij...}_{rs...} = a^i_{i'} a^j_{j'} ... b^{r'}_r b^{s'}_s ... T^{i'j'}_{r's'...}$

As in special relativity lower/upper indices represent covariant/controvariant indices respectively, but here the transformations matrices are arbitrary matrices and depend on the considered point.

3.1 Examples

It has to be stressed that the coordinate x^k **is not a vector**. Coordinates are simply *labels* which realise a correspondence (one to one) between the points of a region of the manifold we are dealing with and a region in $\mathbb{R}^4$. In the language of differential geometry this is called a *chart*. In order to cover all the manifold in general more charts are necessary. The collection of the charts is called *atlas* (for example, in order to cover the sphere S^2, at least two charts are necessary).

On the contrary, the differential dx^k is an important example of controvariant vector and we have

$$d\tilde{x}^k = \frac{\partial \tilde{x}^k}{\partial x^j}\, dx^j = a^k_j dx^j\,.$$

In a similar way, the derivative of an arbitrary function (scalar) is a covariant vector. In fact

$$\frac{\partial \tilde{\phi}}{\partial \tilde{x}^k} = \frac{\partial \phi}{\partial x^j}\,\frac{\partial x^j}{\partial \tilde{x}^k} = b^j_k\,\frac{\partial \phi}{\partial x^j}\,.$$

The (tensor) product of a tensor of order p with one of order q gives rise to a tensor of order $p+q$, while the *contraction* (the sum over a covariant-controvariant index) on a tensor of order p gives rise to a tensor of order $p-2$. In particular, the contraction of any covariant vector with a controvariant one gives rise to a scalar. In fact

$$\tilde{T}^{ijk...}_{rsk...} = a^i_{i'} a^j_{j'} a^k_{k'} ... T^{i'j'k'...}_{r's'k''...} b^{r'}_r b^{s'}_s b^{k''}_k ... = a^i_{i'} a^j_{j'} ... T^{i'j'k...}_{r's'k...} b^{r'}_r b^{s'}_s ...$$

$$\tilde{V}^k \tilde{U}_k = a^k_{k'} b^{k''}_k \, V^{k'} U_{k''} = V^k U_k \, .$$

An important example of tensor of order two is given by the metric, which, according to (2.3), is given by

$$g_{ij} = \frac{\partial X^\mu}{\partial x^i} \frac{\partial X^\nu}{\partial x^j} \eta_{\mu\nu} \, , \qquad \tilde{g}_{ij} = \frac{\partial X^\mu}{\partial \tilde{x}^i} \frac{\partial X^\nu}{\partial \tilde{x}^j} \eta_{\mu\nu} \, ,$$

from which directly follows

$$\tilde{g}_{ij} = b^r_i b^s_j g_{rs} \, , \qquad g_{ij} = a^r_i a^s_j \tilde{g}_{rs} \, . \tag{3.2}$$

Then the metric is a covariant tensor of order two.

Now we are going to show that the inverse matrix of the metric is a controvariant tensor. To this aim it is convenient to use a matrix notation, then we shall indicate by $\underline{m}$ a generic matrix with component m_{ij} and by $\underline{m}^T$ the transpose matrix with component m_{ji}. Using this notation, equation (3.2) can be written in the form

$$\underline{g} = \underline{a}^T \, \underline{\tilde{g}} \, \underline{a} \tag{3.3}$$

and as a consequence

$$1 = \underline{g}^{-1} \, \underline{g} = \underline{g}^{-1} \, \underline{a}^T \, \underline{\tilde{g}} \, \underline{a} \, .$$

The desired result can be found by multiplying the latter equation by $[\underline{a}^T \underline{\tilde{g}} \underline{a}]^{-1}$ and recalling that $\underline{a}^{-1} = \underline{b}$. We get

$$\underline{b} \, \underline{\tilde{g}}^{-1} \, \underline{b}^T = \underline{g}^{-1} \quad \Longrightarrow \quad \underline{\tilde{g}}^{-1} = \underline{a} \, \underline{g}^{-1} \, \underline{a}^T \, . \tag{3.4}$$

We see that the matrix $\underline{g}^{-1}$ transforms like a controvariant tensor of order two. As usual, we indicate its components as g^{ij}. So

$$\tilde{g}^{ij} = a^i_r a^j_s \, g_{rs} \, , \qquad g^{ik} g_{kj} = \delta^i_j \, .$$

From the latter equation one sees that the Krönecker symbol δ^i_j is a mixed tensor of order two.

By means of the metric tensor we pass from covariant to controvariant components of a tensor and viceversa. One has

$$T^{ij...} = g^{ii'} g^{jj'} ... T_{i'j'...} \, , \qquad T_{ij...} = g_{ii'} g_{jj'} ... T^{i'j'...} \, .$$

For this reason, on a Riemannian/Lorentzian manifold there is no deep difference between covariant and controvariant tensors.

3.2 Tensor densities

On a Lorentzian manifold an important role is played by the determinant of the metric $g = |\det g_{ij}|$. Such a quantity is not a scalar but a *scalar density of weight* -2, because it transforms like a scalar apart the factor J^{-2}, J being the Jacobian of transformation. In fact we have

$$\tilde{g}_{ij} = b_i^r b_j^s\, g_{rs} \quad \Longrightarrow \quad \tilde{g} = \det \tilde{g} = J^{-2}\, g\,, \qquad\qquad J = \left|\det b_j^i\right|^{-1} = \left|\det a_j^i\right| = \left|\frac{\partial \tilde{x}}{\partial x}\right|\,.$$

We call *tensor density of weight* w a quantity $\mathcal{T}$ which transforms homogeneously like a tensor apart a factor J^w. An arbitrary tensor density can always be expressed in the form $\mathcal{T} = T\, g^{-w/2}$, T being an ordinary tensor.

The square root $\sqrt{|g|}$ is a scalar density of weight -1. This is particularly important because it permits to build up invariant volumes. For a coordinate transformation the infinitesimal volume $d^N x$ is not invariant, but

$$d^N \tilde{x} = J\, d^N x \quad \Longrightarrow \quad \sqrt{|\tilde{g}|}\, d^N \tilde{x} = \sqrt{|g|}\, d^N x\,.$$

Another important example of tensor density of weight -1 is the Levi-Civita symbol $e^{ijrs\dots}$. This is completely antisymmetric and $e^{0123..,N-1} = 1$.

It has to be noted that both tensors and tensor densities transform homogeneously. This means that if they are vanishing in a given reference frame, then they are vanishing in all reference frames.

3.3 Affine connection

We shall deal also with quantities which do not transform homogeneously and so they are not tensors or tensor densities. We have already seen that we can always choose a coordinates system in which all Christoffel symbols are vanishing, but in an arbitrary frame this is not the case. This means that the connection is not a tensor because it transforms non homogeneously. In fact, recalling definition (2.5) and (3.1) (from now on we suppress the “hat”)

$$\Gamma_{ij}^k = A_\mu^k \frac{\partial}{\partial x^i} B_j^\mu\,, \qquad\qquad \tilde{\Gamma}_{ij}^k = \tilde{A}_\mu^k \frac{\partial}{\partial \tilde{x}^i} \tilde{B}_j^\mu\,,$$

$$A_\mu^k = \frac{\partial x^k}{\partial X^\mu}\,, \qquad B_k^\mu = \frac{\partial X^\mu}{\partial x^k}\,, \qquad\qquad A_\mu^i B_j^\mu = \delta_j^i\,, \qquad A_\nu^k B_k^\mu = \delta_\nu^\mu\,,$$

$$\tilde{A}_\mu^k = \frac{\partial \tilde{x}^k}{\partial X^\mu}\,, \qquad \tilde{B}_k^\mu = \frac{\partial X^\mu}{\partial \tilde{x}^k}\,, \qquad\qquad \tilde{A}_\mu^i \tilde{B}_j^\mu = \delta_j^i\,, \qquad \tilde{A}_\nu^k \tilde{B}_k^\mu = \delta_\nu^\mu\,,$$

$$A_\mu^i \tilde{B}_j^\mu = b_j^i\,, \qquad\qquad \tilde{A}_\mu^i B_j^\mu = a_j^i\,,$$

we get

$$\begin{aligned}
\tilde{\Gamma}_{ij}^k &= \tilde{A}_\mu^k \frac{\partial}{\partial \tilde{x}^i} \left(B_s^\mu b_j^s\right) = \tilde{A}_\mu^k\, b_i^r \frac{\partial}{\partial x^r} \left(B_s^\mu b_j^s\right) \\
&= \tilde{A}_\mu^k\, b_i^r\, b_j^s \frac{\partial}{\partial x^s} B_s^\mu + \tilde{A}_\mu^k\, b_i^r\, B_s^\mu \frac{\partial}{\partial x^s} b_j^s \\
&= a_m^k b_i^r b_j^s\, \Gamma_{rs}^m + a_m^k b_i^r \partial_r b_j^m \\
&= \frac{\partial \tilde{x}^k}{\partial x^m} \frac{\partial x^r}{\partial \tilde{x}^i} \frac{\partial x^s}{\partial \tilde{x}^j}\, \Gamma_{rs}^m + \frac{\partial \tilde{x}^k}{\partial x^m} \frac{\partial^2 x^m}{\partial \tilde{x}^i\, \tilde{x}^j}\,.
\end{aligned} \tag{3.5}$$

We see that the connection transforms as a tensor with respect to linear transformations, because in such a case the transformation matrices $\underline{a}$ and $\underline{b}$ are constant quantities, but it has also to be noted that the difference between two different connections is a third order tensor. In fact, if Γ and Σ are two distinct connections we get

$$\left\{\begin{array}{l} \tilde{\Gamma}^k_{ij}(\tilde{x}) = a^k_m b^r_i b^s_j\, \Gamma^m_{rs}(x) + a^k_m b^r_i \partial_r b^m_j\,, \\ \tilde{\Sigma}^k_{ij}(\tilde{x}) = a^k_m b^r_i b^s_j\, \Sigma^m_{rs}(x) + a^k_m b^r_i \partial_r b^m_j\,, \end{array}\right. \quad \Longrightarrow \quad \tilde{\Gamma}^k_{ij}(\tilde{x}) - \tilde{\Sigma}^k_{ij}(\tilde{x}) = a^k_m b^r_i b^s_j \left[\Gamma^m_{rs}(x) - \Sigma^m_{rs}(x)\right] .$$

3.4 Covariant differentiation

Now we want to generalise the concept of differentiation, in such a way that the "differential" of a tensor is again a tensor. We see that the ordinary differentiation does not satisfy such a requirement. For example, for a controvariant vector we have

$$\begin{aligned} \frac{\partial \tilde{V}^k}{\partial \tilde{x}^i} &= b^r_i \partial_r\left(a^k_s V^s\right) = b^r_i a^k_s \partial_r V^s + b^r_i V^s \partial_r a^k_s\,, \\ d\tilde{V}^k &= d\tilde{x}^i\, \frac{\partial \tilde{V}^k}{\partial \tilde{x}^i} = dx^r\, \partial_r\left(a^k_s V^s\right) = a^k_s\, dV^s + V^s\, da^k_s\,. \end{aligned} \tag{3.6}$$

Again we see that the differential is a vector only with respect to linear transformation. This is due to the fact that we are dealing with two vectors V^k and $V^k + dV^k$ in two different points x^k and $x^k + dx^k$ and they transform in different ways when the matrices $\underline{a}\,,\underline{b}$ are not constant. In order to compare the two vectors, we have to "parallel transport" them in the same point, for example, we can transport the vector V^k in the point $x^k + dx^k$ and then compare it with the vector $V^k + dV^k$. Of course, the parallel transport has to be defined. In Euclidean (Minkowskian) manifolds and in Cartesian coordinates, parallel vectors have proportional components, but this is not the case in arbitrary systems.

In an $N-$dimensional manifold M^N, the vectors in a generic point P "live" in the tangent space T_P at P, which is a $N-$dimensional vector space ($\mathbb{R}^N$), while the vectors at the point Q "live" on T_Q, which is another $N-$dimensional vector space. One has to define how compare T_P with T_Q. In classical physics one works in Euclidean manifolds ($M^N \equiv \mathbb{R}^N$) and so the tangent spaces are isomorphic to the base manifold M^N and often one confuses them.

- Example: as an example we consider two "parallel" vectors $\vec{V}$ and $\vec{U}$ in $\mathbb{R}^2$ applied at two different points P_1 and P_2. We use both Cartesian $\{x^k\} \equiv (x, y)$ and polar coordinates $\{\tilde{x}^k\} \equiv (r, \varphi)$, that is

$$\left\{\begin{array}{l} x = r\cos\varphi \\ y = r\sin\varphi \end{array}\right. \quad \Longrightarrow \quad \left\{\begin{array}{l} dx = \cos\varphi dr - r\sin\varphi d\varphi \\ dy = \sin\varphi dr + r\cos\varphi d\varphi \end{array}\right.$$

from which it follows

$$\underline{b} = \left\{\frac{\partial x^i}{\partial \tilde{x}^j}\right\} = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\,\cos\varphi \end{pmatrix} = \begin{pmatrix} b^1_1 & b^1_2 \\ b^2_1 & b^2_2 \end{pmatrix}\,, \tag{3.7}$$

$$\underline{a} = \left\{\frac{\partial \tilde{x}^i}{\partial x^j}\right\} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix} = \begin{pmatrix} a^1_1 & a^1_2 \\ a^2_1 & a^2_2 \end{pmatrix}\,, \tag{3.8}$$

The vectors assume the form

$$\begin{aligned}\vec{V} &= V_1\,\hat{u}_{\,x} + V_2\,\hat{u}_{\,y} = \tilde{V}_1\,\hat{u}_{\,r} + \tilde{V}_2\,\hat{u}_{\,\varphi}\,,\\ \vec{U} &= U_1\,\hat{u}_{\,x} + U_2\,\hat{u}_{\,y} = \tilde{U}_1\,\hat{u}_{\,r} + \tilde{U}_2\,\hat{u}_{\,\varphi}\,,\end{aligned}$$

where $\{\,\hat{u}_{\,x}.\,\hat{u}_{\,y}\}$ is the "repère naturelle" generated by Cartesian coordinates, while $\{\,\hat{u}_{\,r}.\,\hat{u}_{\,\varphi}\}$ is the "repère naturelle" generated by polar coordinates [31]. while V_k, U_k and $\tilde{V}_k.\tilde{U}_k$ $(k = 1, 2)$ are the corresponding components. In order to simplify the computation we assume a constant $\vec{V} \equiv U$ and parallel to the abscissa. Because the frame $\{\,\hat{u}_{\,x},\,\hat{u}_{\,y}\}$ does not depend on the point we have

$$\begin{cases}\vec{V} = v\,\hat{u}_{\,x}\,,\\ \vec{U} = v\,\hat{u}_{\,x}\,,\end{cases} \qquad \Longrightarrow \qquad \begin{cases}V_1 = U_1 = v\,,\\ V_2 = U_2 = 0\,,\end{cases} \qquad\qquad v = constant\,.$$

We see that two equal vectors have equal Cartesian coordinates. The situation completely change in polar coordinates, because the frame $\{\,\hat{u}_{\,r},\,\hat{u}_{\,\varphi}\}$ depends on the point. In fact using (3.7) we have

$$\tilde{V}_k = b^j_k(P_1)V_k\,, \qquad\qquad \tilde{U}_k = b^j_k(P_2)U_k\,,$$

and explicitly

$$\begin{cases}\tilde{V}_1 = b^1_1 V_1 + b^2_1 V_2 = v\cos\varphi_1\\ \tilde{V}_2 = b^1_2 V_1 + b^2_2 V_2 = -v r_1 \sin\varphi_1\end{cases}\,, \qquad \begin{cases}\tilde{U}_1 = b^1_1 U_1 + b^2_1 U_2 = v\cos\varphi_2\\ \tilde{U}_2 = b^1_2 U_1 + b^2_2 U_2 = -v r_2 \sin\varphi_2\end{cases}\,.$$

where $P_1 \equiv (x_1, y_1) \equiv (r_1, \varphi_1)$ and $P_2 \equiv (x_2, y_2) \equiv (r_2, \varphi_2)$. Now we choose the points P_1, P_2 infinitely close then
$(r_1, \varphi_1) \equiv (r, \varphi)$, $(r_2, \varphi_2) \equiv (r + dr, \varphi + d\varphi)$ and

$$\begin{aligned}\delta\tilde{V}_1 \equiv \tilde{U}_1 - \tilde{V}_1 &= -v\sin\varphi d\varphi = \frac{\tilde{V}_2}{r}\,d\varphi\,,\\ \delta\tilde{V}_2 \equiv \tilde{U}_2 - \tilde{V}_2 &= -v\sin\varphi dr = vr\cos\varphi d\varphi = \frac{\tilde{V}_2}{r}\,dr - r\tilde{V}_1\,d\varphi\,.\end{aligned}$$

It is easy to see that the latter relation can be written in the compact form (here we suppress the tilde)

$$\delta V_k = \hat{\Gamma}^j_{ik} dx^i V_j\,, \qquad\qquad i, j, k = 1, 2\,, \qquad (x^1, x^2) = (r, \varphi)\,, \tag{3.9}$$

$\hat{\Gamma}^j_{ik}$ bing the Christoffel symbols related to the metric of $\mathbb{R}^2$ $(ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\varphi^2)$. This means that in general equal vectors in different points have different components when the Christoffel symbols are non vanishing. In Cartesian coordinates all components of the

[31] The coordinate system determines a reference which is called "repère naturelle", which is given by the unitary vectors tangent to the coordinate surfaces in the given point and in general it depends on the point itself. The "repère naturelle" corresponding to Cartesian coordinates does not depend on the point because it is alway parallel to itself.

metric are constant and so the corresponding Christoffel symbols are vanishing and (3.9) gives $\delta V_k = 0$, while in polar coordinates we have

$$g_{11} = g_{rr} = 1\,, \qquad g_{22} = g_{\varphi\varphi} = r^2\,, \qquad g_{ij} = 0\,, \text{ for } i \neq j\,.$$

Using (2.9) one finds that all the Christoffel symbols are vanishing but

$$\hat{\Gamma}^1_{22} = -r\,, \qquad \hat{\Gamma}^2_{12} = \hat{\Gamma}^2_{21} = \frac{1}{r}\,.$$

Looking at (3.9) we see that in general, also in an Euclidean manifold, the "natural components" of a vector changes when we make a parallel transport from a point to another. This is due to the fact that the "repère naturelle" depends on the given point.

3.5 Parallel transport of a vector

The concept of parallel transport has been introduced by Levi-Civita, by considering the original manifold M^N merged in an Euclidean manifold $\mathbb{R}^{N(N+1)/2}$ and subsequently it has been defined axiomatically by Cartan and Weyl directly in M^N.

We consider a vector field $V^k(x)$ and $V^k + dV^k$ in two different points x^k and $x^k + dx^k$ and we indicate by $V^k + \delta V^k$ the vector obtained by parallel transport of V^k in the point $x^k + dx^k$. We expect δV^k to be proportional to the "distance" dx and to the vector itself, in such a way that $\delta(V_1^k + V_2^k) = \delta V_1^k + \delta V_2^k$. Then we write

$$\delta V^k = -\Gamma^k_{ij}\, dx^i V^j\,,$$

where Γ^k_{ij} are quantities, depending on the coordinates, which defines the law of parallel transport. The coefficients Γ^k_{ij} are the components of the *affine connection* and can be defined also in a non-Riemannian manifold.

After the parallel transport we have two vectors at the same point and the difference $DV^k \equiv \nabla V^k$ reads

$$DV^k \equiv \nabla V^k = dV^k - \delta V^k = dV^k + \Gamma^k_{ij}\, dx^i V^j\,.$$

The latter equation defines the *covariant differential* of a controvariant vector and it is a controvariant vector if Γ^k_{ij} transforms as in (3.5), that is

$$\tilde{\Gamma}^k_{ij} = a^k_p b^r_i b^s_j\, \Gamma^p_{rs} + a^k_p \partial_i b^p_j\,.$$

Now we define the *covariant derivative* ∇_i by $D = dx^i \nabla_i$ and so, for a controvariant vector we get

$$\nabla_i V^k = \partial_i V^k + \Gamma^k_{ij} V^j\,. \tag{3.10}$$

By definition, when $\nabla_i V^k = 0$ the vector is parallel transported along the curve x^i.

In principle the connection can be chosen with the only restriction that it satisfies the transformation law (3.10) and so there are infinite possible choices. For example, one can add an arbitrary tensor K^k_{ij} to a given connection $\hat{\Gamma}^k_{ij}$ and the result $\Gamma^k_{ij} = \hat{\Gamma}^k_{ij} + K^k_{ij}$ is a new connection different with respect to the given one. If the given connection is symmetric (for example the one of Levi-Civita), then K^k_{ij} is called *contorsion tensor* and its antisymmetric part defines the *torsion tensor*

$S^k_{ij} = \Gamma^k_{ij} - \Gamma^k_{ji} = K^k_{ij} - K^k_{ji}$. It gives a "measures" of the torsion of the tangent space parallel transported along a coordinate line.

In order to restrict the possible choices, it is natural to require the scalar product between two arbitrary vectors A^k and B_k to be invariant when they are parallel transported. In this way their transportation laws are related because

$$0 = \delta(A^k B_k) = A^k \delta B_k + B_k \delta A^k = A^k \delta B_k - B_k \Gamma^k_{ij} dx^i A^j \quad \Longrightarrow \quad \delta B_k = \Gamma^j_{ik} dx^i B_j \tag{3.11}$$

and the covariant derivative

$$\nabla_i B_k = \partial_i B_k - \Gamma^j_{ik} B_j \,.$$

The covariant differentiation can be generalised to tensors of arbitrary orders. We start with the special tensor $T^{ij} = A^i B^j$. We have

$$\begin{aligned} \nabla_k T^{ij} &= \nabla_k (A^i B^j) = A^i \nabla_k B^j) + B^j \nabla_k A^i \\ &= A^i (\partial_k B^j + \Gamma^j_{kl} B^l) + Bj(\partial_k A^i + \Gamma^i_{kl} A^l) = \partial_k T^{ij} + \Gamma^i_{kl} T^{lj} + \Gamma^j_{kl} T^{il} \end{aligned}$$

and because this is linear in the tensor it is valid for an arbitrary tensor of order two. Now, with a similar trick we obtain the derivative of an arbitrary tensor of order (p, q). It reads

$$\nabla_k T^{ij\dots}_{rs\dots} = \partial_k T^{ij\dots}_{rs\dots} + \Gamma^i_{kk'} T^{k'j\dots}_{rs\dots} + \Gamma^j_{kk'} T^{ik'\dots}_{rs\dots} - \Gamma^{k'}_{kr} T^{ij\dots}_{k's\dots} - \Gamma^{k'}_{ks} T^{ij\dots}_{rk'\dots} + \dots \tag{3.12}$$

and this is a tensor of order $(p, q+1)$. In particular we have

$$\nabla_k g_{ij} = \partial_k g_{ij} - \Gamma^l_{ki} g_{lj} - \Gamma^l_{kj} g_{il} = 0 \,.$$

The fact that the covariant derivative of the metric vanishes is a direct consequence of the choice (3.11). To see this we choose two arbitrary vector fields obtained by the parallel transport of A^i and B_j, then

$$0 = \nabla_k A^i = \nabla_k B_j = \nabla_k (A^i B_i) \,.$$

But in a Riemannian manifold we also have

$$0 = \nabla_k (A^i B_i) = \nabla_k (g_{ij} A^i B^j) = A^i B^j \nabla_k g_{ij}$$

and due to the arbitrariness of the vectors we obtain

$$\nabla_k g_{ij} = 0 \quad \Longrightarrow \quad \partial_k g_{ij} = \Gamma^l_{ki} g_{lj} + \Gamma^l_{kj} g_{il} \,. \tag{3.13}$$

When the latter condition is satisfied, we say that the connection is *compatible* with the metric.

It has to be noted that from (3.13) it is not possible to determine the connection coefficients as it has been done in (2.9), because in general they are not symmetric ($\Gamma^k_{ij} \neq \Gamma^k_{ji}$), but it is possible if we choose torsion-free connections. In fact we have the following important result:

- *In a Riemannian manifold there is a unique connection compatible with the metric and torsion free and this is the Levi-Civita connection.*

This means that

$$\begin{cases} \nabla_k g_{ij} = 0 \\ S^k_{ij} = 0 \end{cases} \quad \Longrightarrow \quad \Gamma^k_{ij} = \Gamma^k_{ji} = \hat{\Gamma}^k_{ij} = \frac{1}{2} g^{kl} (\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}) \,. \tag{3.14}$$

In the rest of the paper we shall always use such a connection, which, during a parallel transport, preserve the scalar product, the angle between vectors and it is in agreement with the equivalence principle.

3.6 Some useful formulae

Now we derive some useful relations regarding derivative of tensors and determinant of metric.

From $g^{ir} g_{rs} = \delta^i_s$ we immediately get

$$g^{ir} dg_{rs} + g_{rs} dg^{ir} = 0 \quad \Longrightarrow \quad \begin{cases} dg^{ij} = -g^{ir} g^{js} dg_{rs} \,, & \partial_k g^{ij} = -g^{ir} g^{js} \partial_k g_{rs} \,, \\ dg_{rs} = -g_{ir} g_{js} dg^{ij} \,, & \partial_k g_{rs} = -g_{ir} g_{js} \partial_k g^{ij} \,. \end{cases} \tag{3.15}$$

Recalling the definition of the inverse matrix and the determinant g we also get

$$\begin{aligned} \frac{1}{g}\, dg &= g^{ij} dg_{ij} = -g_{ij} dg^{ij} \quad \Longrightarrow \quad \frac{1}{g}\, \partial_k g = g^{ij} \partial_k g_{ij} = -g_{ij} \partial_k g^{ij} \,, \\ \frac{1}{g}\, \partial_k g &= \frac{1}{g}\, \frac{\partial g}{\partial g^{ij}}\, \partial_k g^{ij} = g^{ij} \partial_k g_{ij} \quad \Longrightarrow \quad \frac{1}{g}\, \frac{\partial g}{\partial g^{ij}} = g_{ij} \,. \end{aligned} \tag{3.16}$$

Note that these equations can be easily derived by recalling the relation which exists between the trace of the logarithm of matrix Λ and the logarithm of its determinant. In fact, given a symmetric square matrix Λ with eigenvalues $\lambda_1, \lambda_2, ..., \lambda_N$ one has

$$\mathrm{Tr} \log \Lambda = \sum_{n=1}^{N} \log \lambda_n = \log \prod_{n=1}^{N} \lambda_n = \log \det \Lambda \,.$$

Now, if the matrix depends on a parameter ρ, deriving the latter identity we get

$$\frac{d}{d\rho}\, \mathrm{Tr} \log \Lambda = \Lambda^{-1} \frac{d}{d\rho}\, \Lambda = \frac{d}{d\rho}\, \log \det \Lambda = \frac{1}{\det \Lambda}\, \frac{d}{d\rho}\, \det \Lambda \,.$$

Applying this to the metric, identity (3.16) follows.

From (3.14), by contraction and using (3.13) we have

$$\Gamma^k_{ik} = \frac{1}{2}\, g^{rs} \partial_i g_{rs} = \frac{1}{\sqrt{|g|}}\, \partial_i \sqrt{|g|} = \partial_i \log \sqrt{|g|} \,, \qquad g^{rs} \Gamma^k_{rs} = -\frac{1}{\sqrt{|g|}}\, \partial_i (\sqrt{|g|}\, g^{ik}) \,. \tag{3.17}$$

With the use of formulae above, we can write the (covariant) divergence of a vector A^k and the (covariant) D'Alembertian of a scalar ϕ in the form

$$\nabla_k A^k = g^{rs} \nabla_r A_s = \partial_k A^k + \Gamma^k_{kj} A^j = \partial_k A^k + A^j \partial_j \log \sqrt{|g|} = \frac{1}{\sqrt{|g|}}\, \partial_k (\sqrt{|g|}\, A^k) \,, \tag{3.18}$$

$$\nabla_k \phi = \partial_k \phi \,, \qquad \Box \phi = g^{rs} \nabla_r \nabla_r \phi = \frac{1}{\sqrt{|g|}}\, \partial_r (\sqrt{|g|}\, g^{rs} \partial_s \phi) \,, \tag{3.19}$$

while the derivative of an antisymmetric tensor $F^{ij} = -F^{ji}$ reads

$$\nabla_k F^{kj} = \partial_k F^{kj} + \Gamma^k_{kl} F^{lj} + \Gamma^j_{kl} F^{kl} = \frac{1}{\sqrt{|g|}}\, \partial_k (\sqrt{|g|}\, F^{kj}) \,. \tag{3.20}$$

In particular, for the electromagnetic tensor field we obtain

$$F_{ij} = \nabla_i A_j - \nabla_j A_i = \partial_i A_j - \partial_j A_i \,, \qquad \nabla_k F^{kj} = \Box A^j - g^{js} \nabla_r \nabla_s A^r \,. \tag{3.21}$$

We stress again that we are using the Levi-Civita connection. In the presence of torsion, some of the relations above have more terms.

To finish this section we would like to observe that the integration over a region $V^N \subset M^N$ of the covariant divergence of a vector gives rise to an integral on the boundary S^{N-1} of V^N, via the Gauss theorem. In fact we have

$$\int_{V^N} \nabla_k A^k \sqrt{|g|}\, d^N x = \int_{V^N} \partial_k \left(\sqrt{|g|}\, A^k \right) d^N x = \int_{S^{N-1}} \sqrt{|g|}\, A^k \, d\sigma_k \,. \tag{3.22}$$

This means that, as it happens in special relativity, a continuity equation of the kind $\nabla_k J^k = 0$ will give rise to a conservation law, but, in contrast with special relativity, to a continuity equation of the kind $\nabla_k T^{kj} = 0$ in general will not correspond a conserved vector (see Section 8).

4 The Influence of Gravitation on Physical Systems

Here we are going to see how gravitation modifies the equations of motion of s ome physical systems. To this aim we shall make use of the following **principle of general covariance:**

- If a physical equation is true in the absence of gravitation and it is written in a *covariant* form, then it is valid in an arbitrary gravitational field.

This means that in order to include gravitation in special relativity, we write the equations of motion or the field equations in a covariant form, that is in a form which is invariant under general coordinate transformations. Then we have to use only tensors and covariant operators.

It has to be noted that as regarding the physical contents, the principle of equivalence and the principle of general covariance are the same thing, but the latter provides a powerful mathematical technique which permits to take into account of an arbitrary gravitational field.

4.1 The motion of a test particle

We have already study such a system in the presence of gravitation by using the equivalence principle. Now we shall use the principle of general covariance.

In special relativity the motion of a test particle is determined by the equation

$$\frac{du^k}{d\tau} = 0 \,, \qquad u^k = \frac{dx^k}{d\tau} = -c\,\frac{dx^k}{ds} \,, \qquad u^k u_k = -c^2 \,, \tag{4.1}$$

u^k being the 4-velocity, which with our conventions satisfy the constraint $u^k u_k = -c^2$.

Equation (4.1) is not covariant, because we have seen in the previous section that the differential of a vector is not a vector, but we have also seen that it becomes a vector if we replace differentiation with covariant differentiation. Then we expect the motion of a test particle in the presence of gravitation to be described by the equation

$$\frac{Du^k}{d\tau} = 0 \,, \qquad u^k = \frac{dx^k}{d\tau} \,. \tag{4.2}$$

This is a covariant equation and in the absence of gravitation reduces to (4.1) (the connection is always the one of Levi-Civita). The solutions of (4.2) are called *auto parallel*. In our case (Riemannian manifold equipped with Levi-Civita connection) they coincide with the geodesics already discussed in (2.4). Equation (4.2) holds also for massless particles, but in such a case $d\tau$ is an arbitrary scalar parameter, but not the proper time, which for massless particles is vanishing.

If the particle is not free, but in the presence of an external force $\vec{f}$, on the right hand sides of (4.1) and (4.2) we have f^k/m, m being the mass of the particle and f^k the 4-vector obtained by means of a coordinate transformation applied to $(0, \vec{f})$. Then

$$\frac{Du^k}{d\tau} = \frac{f^k}{m} \quad \Longrightarrow \quad \frac{d^2x^k}{d\tau^2} = \frac{f^k}{m} - \Gamma^k_{ij}\frac{dx^i}{d\tau}\frac{dx^j}{d\tau} \,.$$

The last term can be seen as the gravitational force acting on the particle.

Using the constraints $u^k u_k = -c^2$ and $f^k u_k = 0$, the latter equation can be written in the form

$$\frac{Du^j}{d\tau} = W^{jk} u_k \,, \qquad W^{jk} = -W^{kj} = \frac{u^j f^k - u^k f^j}{mc^2} \,. \tag{4.3}$$

When a vector satisfies an equation of that kind, we say that it is transported *a la Fermi-Walker*[32]. The velocity of a free particle is parallel transported along a geodesic, equation (4.2), while in the presence of an external force it is transported a la Fermi-Walker, equation (4.3).

[32] Enrico Fermi (Italia) 1901-1954.

4.2 The motion of the spin

Now we consider a particle with spin. In the system in which the particle is at rest, $u^k \equiv (-c, 0, 0, 0)$ and the spin vector is $S^k \equiv (0, \vec{S})$ and so $S_k u^k = 0$. Of course, in the absence of external forces it is conserved. Then, in an arbitrary gravitational field we shall have

$$\frac{DS^k}{d\tau} = 0 \,, \qquad\qquad S_k u^k = 0 \,.$$

The situation changes if on the particle acts an external force $\vec{f}$. For simplicity we suppose the force not experiencing any torque. This means that the particle is accelerated, but in a locally inertial frame momentarily at rest with respect to the particle there is no precession of the spin. In such a frame, as above we have

$$\vec{u} = 0 \,, \qquad\qquad \frac{d\vec{S}}{dt} = 0 \,, \tag{4.4}$$

This "no torque" condition can be written in the covariant form

$$\frac{dS^k}{d\tau} = \alpha \, u^k \,, \tag{4.5}$$

which effectively reduces to (4.4) in the reference where $\vec{u} = 0$.

There scalar function α, in general depending on spin and velocity, can be explicitly determined by deriving the identity $S_k u^k = 0$, which holds in any frame. One gets

$$0 = u_k \frac{dS^k}{d\tau} + S^k \frac{du_k}{d\tau} = \alpha \, c^2 + \frac{S_k f^k}{m} \qquad \Longrightarrow \qquad \alpha = \frac{S_k f^k}{mc^2} \,.$$

Then it follows

$$\frac{dS^k}{d\tau} = \alpha u^k = \frac{S_j f^j \, u^k}{mc^2} \,,$$

and so, as it is well known, if the particle is accelerated the spin vector changes direction. This phenomenon is known as the *Thomas precession.*

The latter equation can be immediately generalised to gravitational field by using the principle of general covariance. We have

$$\frac{DS^j}{d\tau} = W^{jk} S_k \,, \qquad\qquad W^{jk} = \frac{u^j w^k - u^k w^j}{c^2} \,, \qquad w^j = \frac{Du^j}{d\tau} = \frac{f^j}{m} \,.$$

As well as the velocity, also the spin is transported a la Fermi-Walker.

4.3 The electromagnetic field

Starting from the potential A_k we define the electromagnetic strength tensor F_{ij} and its dual ${}^*F^{ij}$ by

$$F_{ij} = \nabla_i A_j - \nabla_j A_i = \partial_i A_j - \partial_j A_i \,, \qquad\qquad {}^*F^{ij} = \frac{e^{ijrs}}{2\sqrt{|g|}} F_{rs} \,.$$

e^{ijrs} being the usual Levi-Civita symbol. The current vector reads

$$J^k = \rho_0 u^k \,,$$

ρ_0 being the proper charge density, that is the density in the frame in which the infinitesimal volume we are considering is momentarily at rest. The Maxwell equation in the presence of gravitation read

$$\nabla_j {}^*F^{jk} = 0 \,, \qquad\qquad \nabla_j F^{jk} = -\frac{4\pi}{c}\, J^k \,.$$

Recalling (3.20) one also gets

$$\partial_j \left(\sqrt{|g|} F^{jk}\right) = -\frac{4\pi}{c}\sqrt{|g|} J^k \quad \Longrightarrow \quad \frac{1}{\sqrt{|g|}}\, \partial_k \left(\sqrt{|g|}\; J^k\right) = \nabla_k J^k = 0 \,.$$

From the last (continuity) equation it follows the conservation law of the electric charge.

4.4 Exercise: the action for a test particle

Derive the equation of motion for a test particle using the action principle.

As in special relativity, the only scalar which can be built up using parameters and coordinates of particle is proportional to the invariant interval ds and so the equation of motion is given by

$$\delta \int_{s_1}^{s_2} ds = \delta \int_{\lambda_1}^{\lambda_2} \frac{ds}{d\lambda}\, d\lambda = \delta \int_{\lambda_1}^{\lambda_2} \sqrt{g_{ij} \frac{dx^i}{d\lambda}\frac{dx^j}{d\lambda}}\, d\lambda = 0 \,, \tag{4.6}$$

where an arbitrary evolution paramter λ has been introduced.

According to action principle, the variation is performed on the arbitrary trajectory x^k, but with the constraint $\delta x^k(s_1) = \delta x^k(s_2) = 0$ ($s_1 = s(\lambda_1)$, $s_2 = s(\lambda_2)$). We observe that

$$\delta \frac{ds}{d\lambda} = \delta \sqrt{g_{ij} \frac{dx^i}{d\lambda}\frac{dx^j}{d\lambda}} = \left(\frac{dx^i}{ds}\frac{dx^j}{ds}\, \delta g_{ij} + g_{ij} \frac{dx^i}{ds}\, \frac{d\delta x^j}{ds} + g_{ij} \frac{d\delta x^i}{ds}\, \frac{dx^j}{ds} \right) \frac{ds}{2d\lambda} \,.$$

Using this result in (4.6) we have

$$\delta \int_{s_1}^{s_2} ds = \int_{\lambda_1}^{\lambda_2} d\lambda\, \delta \sqrt{g_{ij} \frac{dx^i}{d\lambda}\frac{dx^j}{d\lambda}} = \int_{s_1}^{s_2} ds \left(\frac{dx^i}{ds}\frac{dx^j}{ds}\, \delta g_{ij} + g_{ij} \frac{dx^i}{ds}\, \frac{d\delta x^j}{ds} + g_{ij} \frac{d\delta x^i}{ds}\, \frac{dx^j}{ds} \right) . \tag{4.7}$$

Integrating by parts and using the fact the variation δx^k is arbitrary, we finally obtain the geodesic equations

$$\frac{d^2 x^k}{ds^2} + \frac{1}{2} g^{kl} \left(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}\right) \frac{dx^i}{ds}\, \frac{dx^j}{ds} = \frac{d^2 x^k}{ds^2} + \hat{\Gamma}^k_{ij}\, \frac{dx^i}{ds}\, \frac{dx^j}{ds} \,. \tag{4.8}$$

Of course here we have the Levi-Civita connection, because only metric appears in the action.

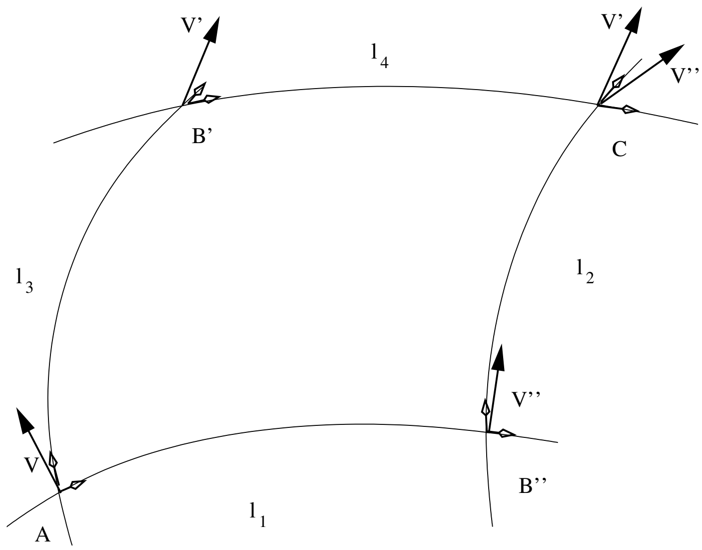

Figure 2: parallel transport of a vector along a "parallelogram"

5 The Riemann Tensor

Here we define and study the properties of *Riemann tensor*, which generalise to manifolds with arbitrary dimensions, the concept of curvature as introduced by Gauss for the surface.

5.1 The parallel transport of a vector along a closed curve

In general, when a vector is transported by parallelism along a closed curve, the final vector differs with respect to the original one. To see this, we chose a vector A^i and an infinitesimal "parallelogram" built up along the coordinate axis (r, s) and then we compare the two vectors obtained by the parallel transport of A^i along the two paths (see figure 2). Then we have

$$\nabla_r \nabla_s A^i - \nabla_s \nabla_r A^i \equiv [\nabla_r, \nabla_s]\, A^i = R^i_{jrs}\, A^j\,, \tag{5.1}$$

$$R^i_{jrs} = \partial_r \Gamma^i_{sj} - \partial_s \Gamma^i_{rj} + \Gamma^i_{rk}\Gamma^k_{sj} - \Gamma^i_{sk}\Gamma^k_{rj}\,, \tag{5.2}$$

where R^i_{jrs} – the *Riemann tensor* – *is built up with the metric and its derivatives (up to second order). In particular, it is linear in the second derivative of the metric and it is possible to show that* R^i_{jrs} *is the unique tensor with such an important feature.*

- Note that in the presence of torsion other terms appear on the right-hand side of (5.1).

A relation similar to (5.1) is valid for a covariant vector, that is

$$[\nabla_r, \nabla_s]\, B_j = [\nabla_r, \nabla_s]\, g_{jk} B^k = g_{jk}\, [\nabla_r, \nabla_s]\, B^k = g_{jk} R^k_{lrs}\, B^l = -R^i_{jrs}\, B_i\,.$$

Since the commutator satisfies the Leibniz rule

$$[\nabla_r, \nabla_s]\, (A^i B^j) = A^i [\nabla_r, \nabla_s]\, B^j + B^j [\nabla_r, \nabla_s]\, A^i\,,$$

for a generic tensor we get

$$[\nabla_r, \nabla_s]\, T^{ij\ldots}_{mn\ldots} = R^i_{krs} T^{kj\ldots}_{mn\ldots} + R^j_{krs} T^{ik\ldots}_{mn\ldots} + \ldots - R^k_{mrs} T^{ij\ldots}_{kn\ldots} - R^k_{nrs} T^{ij\ldots}_{mk\ldots} - \ldots$$

5.2 Sectional curvature

In order to understand the mathematical meaning of the components of the Riemann tensor, in an arbitrary point P we consider two non-parallel vectors a^i and b^j. These define a plane Σ_P in the manifold (a section). We call *sectional curvature* the quantity

$$K(a,b) = \frac{R_{ijrs}a^i b^j a^r b^s}{(g_{mp}g_{nq} - g_{mq}g_{np})a^m b^n a^p b^q}\,. \tag{5.3}$$

This represents the Gauss curvature of the surface which has Σ_P as tangent plane in P.

5.3 Properties of the Riemann tensor

Now we derive the symmetry properties of the Riemann tensor in the absence of torsion(In the presence of torsion some of the properties below have to be modified). To this aim we it is convenient to write it in terms of the metric, using (3.14). By a straightforward calculation we get

$$\begin{aligned} R^k_{jrs} &= \frac{1}{2}\,\partial_r\left[g^{kl}\left(\partial_s g_{jl} + \partial_j g_{sl} - \partial_l g_{sj}\right)\right] + \Gamma^k_{rl}\Gamma^l_{sj} - (r \leftrightarrow s) \\ &= \frac{1}{2}\,\left(\partial_s g_{jl} + \partial_j g_{sl} - \partial_l g_{sj}\right)\,\partial_r\, g^{kl} + \frac{1}{2}\,g^{kl}\left(\partial_r\partial_s g_{jl} + \partial_r\partial_j g_{sl} - \partial_r\partial_l g_{sj}\right) + \Gamma^k_{rl}\Gamma^l_{sj} - (r \leftrightarrow s)\,, \end{aligned}$$

where $(r \leftrightarrow s)$ means that one has to add the same expression by exchanging the two specified indices. Using the following properties for g_{ij}:

$$g_{ik}\partial_r\, g^{kl} = -g^{kl}\partial_r\, g_{ik}\,, \qquad \partial_r\, g_{ik} = \nabla_r\, g_{ik} + \Gamma^l_{ri}g_{lk} + \Gamma^l_{rk}g_{il} = \Gamma^l_{ri}g_{lk} + \Gamma^l_{rk}g_{il}\,,$$

for the completely covariant tensor we get

$$\begin{aligned} R_{ijrs} &= \frac{1}{2}\,\left(\partial_r\partial_s g_{ij} + \partial_r\partial_j g_{si} - \partial_r\partial_i g_{sj}\right) - \Gamma^k_{js}\partial_r\, g_{ik} + g_{ik}\Gamma^k_{rl}\Gamma^l_{sj} - (r \leftrightarrow s) \\ &= \frac{1}{2}\,\left(\partial_i\partial_s g_{jr} + \partial_j\partial_r g_{is} - \partial_i\partial_r g_{js} - \partial_j\partial_s g_{ir}\right) + g_{pq}\Gamma^p_{is}\Gamma^q_{jr} - g_{pq}\Gamma^p_{ir}\Gamma^q_{js}\,. \end{aligned} \tag{5.4}$$

As we already said above, the Riemann tensor is linear in the second derivatives of the metric and all the following properties hold:

$$\begin{aligned} &1.) \qquad R_{ijrs} = -R_{ijsr} = -R_{jirs} = R_{rsij}\,; \\ &2.) \qquad \{R_{ijrs}\}_{(jrs)} \equiv R_{ijrs} + R_{irsj} + R_{isjr} = 0\,; \\ &3.) \qquad \{\nabla_k R^i_{jrs}\}_{(krs)} \equiv \nabla_k\, R^i_{jrs} + \nabla_r\, R^i_{jsk} + \nabla_s\, R^i_{jkr} = 0\,; \qquad \text{(Bianchi Identity)}. \end{aligned} \tag{5.5}$$

where by $\{T_{ijk\ldots}\}_{(ijk)} = T_{ijk\ldots} + T_{jki\ldots} + T_{kij\ldots}$ we indicate the sum over the cyclic permutations of i, j, k. All the symmetry properties in (1.) trivially follow from (5.4) and also the property in (2.) can be easily verified starting from (5.4). In order to verify the *Bianchi identity* in (3.) it is convenient to use a local inertial frame. In fact, because the Bianchi identity is a tensorial expression it has to be valid in an arbitrary reference frame and in particular in a local inertial frame where the metric is the one of Minkowski and the connection is vanishing (in the considered point). In such a case we have

$$\nabla_k R^i_{jrs} = \frac{1}{2}\,\eta^{il}\left(\partial_k\partial_l\partial_s g_{jr} + \partial_k\partial_j\partial_r g_{ls} - \partial_k\partial_l\partial_r g_{js} - \partial_k\partial_j\partial_s g_{lr}\right)$$

and by summing over cyclic permutations the required identity follows.

5.4 Independent components

Due to symmetry properties above, the components of the Riemann tensor are not all independent. To compute the number of independent ones it is convenient to see R_{ijrs} as a symmetric matrix R_{AB}, where $A \equiv (ij)$ and $B \equiv (rs)$ correspond to the couple of antisymmetric indices. If we indicate by N the dimension of the manifold, then the indices A and B (they are antisymmetric matrices) can assume $\mathcal{N} = N(N-1)/2$ values and so the symmetric matrix R_{AB} has $\mathcal{N}(\mathcal{N}+1))/2$ independent elements . By taking into account of all symmetries (properties in 1.), the number of independent components of R^i_{jrs} is then $N(N-1)[N(N-1)+2]/8$, but such components have to satisfy the condition in (2.), which correspond to $N(N-1)(N-2)(N-3)/24$ independent equations. By subtracting such a number from the previous one we finally get

$$\nu_N = \frac{N^2(N^2-1)}{12}\,, \qquad \nu_1 = 0\,, \qquad \nu_2 = 1\,, \qquad \nu_3 = 6\,, \qquad \nu_4 = 20\,.$$

This is the number of independent components of the Riemann tensor.

In one dimension, the Riemann tensor is always vanishing (every one-dimensional space is flat), while in two dimensions there is only one independent component, which has to be proportional to the Gauss curvature ($K = (r_1 r_2)^{(-1)}$, r_1, r_2 being the principal curvature radius).

5.5 Ricci tensor, scalar curvature and Einstein tensor

Starting from Riemann tensor, by contraction we can build up a tensor of order two and a scalar, that is

$$R_{ij} = R^k_{ikj} = R_{ji}\,, \qquad\qquad R = g^{ij} R_{ij}\,. \tag{5.6}$$

The symmetry of the *Ricci tensor* R_{ij} is true only in the absence of torsion.

By contracting the Bianchi identity (property 3. in 5.5), we get

$$0 = \nabla_i\, R^i_j - \frac{1}{2}\nabla_j\, R = \nabla_i (R^i_j - \frac{1}{2}\delta^i_j\, R)\,.$$

The tensor in the brackets above is called *Einstein tensor* and is usually indicated by G^i_j. Then

$$G_{ij} = G_{ji} = R_{ij} - \frac{1}{2}\, g_{ij}\, R\,, \qquad\qquad \nabla_i\, G^i_j = 0\,. \tag{5.7}$$

5.6 Example: the sphere

Compute Riemann, Ricci and scalar curvature for the sphere S^2.

On S^2 we choose (local) coordinates ϑ, φ (latitude, longitude) in such a way that the distance between infinitely closed points reads

$$d\sigma^2 = r^2 (d\vartheta^2 + \sin^2\vartheta d\varphi^2)\,,$$

where r is a constant. If we see the sphere as embedded in R^3, then r represents the radius of the ball having S^2 as frontier and ds is the distance in $\mathbb{R}^3$ in spherical coordinates $(\rho, \vartheta, \varphi)$, but restricted to $\rho = r$.

Using definition of Christoffel symbols for the non-vanishing components we obtain

$$\Gamma^1_{22} = -\frac{1}{2}\sin 2\vartheta\,, \qquad\qquad \Gamma^2_{12} = \Gamma^1_{21} = \cot\vartheta\,,$$

from which it follows

$$R^1_{212} = \sin^2\vartheta\,, \quad \Longrightarrow \quad R_{1212} = R_{2121} = r^2\sin^2\vartheta\,.$$

Finally, by contraction

$$R_{11} = 1\,, \qquad\qquad R_{22} = \sin^2\vartheta\,, \qquad\qquad R_{12} = R_{21} = 0\,, \qquad\qquad R = \frac{2}{r^2}\,.$$

We see that $R = 2K$, $K = 1/r^2$ being the Gauss curvature.

Using (5.3) we obtain the unic sectional curvature

$$K(a,b) = \frac{R_{1212}\,[(a^1b^2)^2 + (a^2b^1)^2 - 2a^1b^1a^2b^2]}{g_{11}g_{22}\,[(a^1b^2)^2 + (a^2b^1)^2 - 2a^1b^1a^2b^2]} = \frac{1}{r^2}\,.$$

5.7 Exercise: Geodesic deviation

All components of the Riemann tensor are vanishing only in the absence of gravitational field (flat manifold). In order to evidenziate a "true" gravitational field then we have to measure the components of that tensor. To this aim let x^k and $y^k = x^k + \xi^k$ be the coordinates of two free test particles in free fall in an arbitrary reference frame, ξ^k being the "small distance" between them.

Show that the vector ξ^k satisfies the following equation of *geodesic deviation*

$$\frac{D^2\xi^k}{d\tau^2} = R^k_{ijs}u^iu^j\,\xi^s\,, \qquad\qquad u^k = \frac{dx^k}{d\tau}\,. \tag{5.8}$$

The first particle moves along the geodesic

$$\frac{Du^k}{d\tau} = \frac{d^2x^k}{d\tau^2} + \Gamma^k_{ij}(x)\,\frac{dx^i}{d\tau}\,\frac{dx^j}{d\tau} = 0\,,$$

but also the second particle moves along a geodesic given by

$$\frac{d^2y^k}{d\tau^2} + \Gamma^k_{ij}(y)\,\frac{dy^i}{d\tau}\,\frac{dy^j}{d\tau} = 0\,.$$

Taking into account that ξ is a small quantity we expand the latter equation in series of Taylor up to first order. We have

$$\frac{dy^k}{d\tau} = \frac{dx^k}{d\tau} + \frac{d\xi^k}{d\tau} = u^k + \frac{d\xi^k}{d\tau}\,,$$

$$\Gamma^k_{ij}(y) = \Gamma^k_{ij}(x) + \xi^r\partial_r\Gamma^k_{ij}(x) + O(\xi^2)\,,$$

and so

$$\frac{d^2y^k}{d\tau^2} + \Gamma^k_{ij}(y)\,\frac{dy^i}{d\tau}\,\frac{dy^j}{d\tau} \;=\; \frac{Du^k}{d\tau} + \frac{d^2\xi^k}{d\tau^2} + 2\Gamma^k_{ij}(x)\,u^i\,\frac{d\xi^j}{d\tau} + \xi^r\partial_r\Gamma^k_{ij}(x)\,u^iu^j + O(\xi^2)\,.$$

At lowest order the vector ξ^k satisfies the equation

$$\frac{d^2\xi^k}{d\tau^2} + 2\Gamma^k_{ij}(x)\,u^i\,\frac{d\xi^j}{d\tau} + \xi^r\partial_r\Gamma^k_{ij}(x)\,u^iu^j = 0\,,$$

which can be written in the covariant form (5.8).

6 The Einstein Equations

Now we are looking for covariant equations which, in the Newtonian limit, that is for small velocities and weak gravitational fields, are going to coincide with the Poisson equation

$$\Delta\,\phi = 4\pi G\mu_0\,, \qquad \phi = - \sim -\frac{c^2(g_{00}+1)}{2}\,, \qquad \mu_0 \sim \frac{T_{00}}{c^2}\,, \tag{6.1}$$

which describe Newtonian gravity. The solution of Poisson equation gives rise to the gravitational potential $\phi = -MG/r^2$, G being Newton constant. μ_0 is the mass density and T^i_j the energy-momentum tensor of the matter which generates gravitation. Using the energy-momentum tensor and recalling (2.12), we can rewrite (6.1) in the form

$$\Delta\, g_{00} \sim -\frac{8\pi G}{c^4}\; T_{00}\,. \tag{6.2}$$

We expect the field equations to be non-linear, second-order differential equations in the metric. The non linearity is due to the fact that the gravitational field has a “gravitational charge” (the energy-momentum) and so it has to be an auto-interacting field.

Now we must find a covariant equation which, in the Newtonian limit, reduces to (6.2). The right-hand side of (6.2) can be immediately generalised by putting the energy-momentum tensor T_{ij}, while on the left-hand side we shall have a tensor $\hat{G}_{ij}$ depending on the second derivative of the metric.

We have seen in Section 5, that only the Riemann tensor and its contractions depend linearly on the second derivative of the metric. Then, in order to build up $\hat{G}_{ij}$, we have at disposal only the following three tensors: R_{ij}, $R\, g_{ij}$ and $\Lambda\, g_{ij}$, Λ being the so called *cosmological constant*. The more general tensor with the properties required will be an arbitrary combination of such three tensors.

In special relativity T_{ij} satisfies the continuity equation $\partial_k T^k_j = 0$, which gives rise to the conservation of energy and momentum and so one expects T_{ij} to satisfy the equation $\nabla_k T^k_j = 0$ and as a consequence $\nabla_k \hat{G}^k_j = 0$. The only possibility is then $\hat{G}_{ij} = G_{ij} + \Lambda\, g_{ij}$, G_{ij} being the *Einstein tensor* defined in Section 5.

Following Einstein here we disregard the cosmological constant and write the equation in the original form

$$G_{ij} = R_{ij} - \frac{1}{2}\, g_{ij}\, R = \frac{8\pi\, G}{c^4}\, T_{ij}\,, \tag{6.3}$$

$$\nabla_i\, G^i_j = \frac{8\pi\, G}{c^4}\, \nabla_i\, T^i_j = 0\,. \tag{6.4}$$

The dimensional constant factor on the right-hand side has been chosen in order to have the correct Newtonian limit.

Note that one can trivially take into account of the presence of the cosmological constant by the replacement

$$T_{ij} \to T_{ij} - \frac{c^4}{8\pi\, G}\,\Lambda\, g_{ij}\,, \tag{6.5}$$

in the final equations.

Before to verify that they have the right Newtonian limit (6.1), we make some observations:

- On the right-hand side of the Einstein equations has to appear the energy-momentum tensor of matter fields in the symmetric form. This means that T_{ij} is not the canonical tensor of quantum field theory, because in general that is not symmetric. A symmetric energy-momentum tensor can be directly obtained by deriving the matter Lagrangian density with respect to the metric g_{ij} (see Section 6.4). Of course in this way one obtains a symmetric tensor which coincide with the one obtained by using the symmetrisation procedure of Belifante.

- By taking the trace $g^{ij}G_{ij}$ of (6.3) we obtain the equations in the equivalent form

$$R_{ij} = \frac{8\pi\, G}{c^4}\,(T_{ij} - \frac{1}{2}\,g_{ij}\,T)\,, \qquad\qquad T = g^{ij}\,T_{ij}\,. \tag{6.6}$$

- As we have already anticipated, the Einstein equations can be generalised by adding the constant term $\Lambda\, g_{ij}$, which can be seen as the contribution due to a perfect fluid spread out in the whole universe. Such a fluid has negative pressure $p = -\Lambda c^4/8\pi G$ and energy density $\varepsilon = -p$. In the original Einstein equations that term was not present, but it was introduced later in order to have static cosmological solutions (see Section 9.3).

6.1 The Newtonian limit

We consider a macroscopic body which generates a gravitational field. Its energy-momentum tensor has the general form

$$T^{ij} = (p+\varepsilon)\,\frac{u^i u^j}{c^2} + pg^{ij}\,, \qquad u^k \equiv (u^0, u^a)\,, \qquad u^k u_k = -c^2\,, \qquad a = 1,2,3\,, \tag{6.7}$$

where u^k is the macroscopic 4-velocity of the fluid, p the pressure and ε the energy density, which takes into account of the interaction energy between the particles which constitute the body. If the relative velocities of such particles are small with respect to c, then the interaction energy and the pressure can be disregarded with respect to the proper energy $\mu_0 c^2$. Then we have

$$T^{ij} = \mu_0\, u^i u^j\,. \qquad \varepsilon \sim \mu_0 c^2\,, \qquad p \sim 0\,. \tag{6.8}$$

It has to be stressed that in principle the macroscopic velocity u^k can be arbitrary, but if the macroscopic motion is non-relativistic, then

$$|u^0| \gg |u^a| \quad \Longrightarrow \quad |T_{00}| \gg |T_{ak}|\,, \qquad T \sim g^{00}T_{00}\,, \qquad a = 1,2,3\,, \qquad k = 0,1,2,3\,.$$

In such an approximation, equations (6.3) and (6.6) notably simplify because

$$0 \sim R_{ka} - \frac{1}{2}\,g_{ka}R = R_{ka} - \frac{1}{2}\,\eta_{ka}R + O(h^2) \quad \Longrightarrow \quad \begin{cases} R_{ka} \sim 0\,, \quad \forall k \neq a\,, \\ R_{11} \sim R_{22} \sim R_{33} \sim \frac{1}{2}\,R\,, \end{cases} \tag{6.9}$$

where we have put $g_{ij} \sim \eta_{ij} + h_{ij}$ and we have used the fact that the Riemann tensor is at least of first order in h_{ij}. We also get

$$R = g^{ij}R_{ij} = (\eta^{ij} - h^{ij} + O(h^2))\,R_{ij} \sim R_{11} + R_{22} + R_{33} - R_{00} + O(h^2)\,.$$

From equations above it follows that the diagonal components of Ricci tensor are all equal, that is $R_{11} \sim R_{22} \sim R_{33} \sim R_{00}$ and so the only independent equation is

$$R_{00} = \frac{8\pi G}{c^4}\left(T_{00} - \frac{1}{2}\,g_{00}T\right) \sim \frac{4\pi G}{c^4}\,T_{00}\,.$$

From definition of Riemann tensor, at leading order in $h_{ij} = g_{ij} - \eta_{ij}$ (weak-static field)) we get

$$R_{00} \sim \partial_k \Gamma^k_{00} \sim -\frac{1}{2}\,\Delta\, g_{00}\,,$$

from which, as required, the Poisson equation (6.1) directly follows.

6.2 The invariance under diffeomorphism

The Einstein tensor is a symmetric tensor in 4 dimensions and so it has 10 independent components ($N[N+1]/2$, $N=4$). Then Einstein equations (6.3) correspond to 10 differential equations of second order in the 10 unknown functions g_{ij}. Nevertheless, by choosing suitable boundary conditions it is not possible to determine a unique solution because such equations are not linearly independent due to contracted Bianchi identity $\nabla_i\, G^i_j = 0$. This means that only 6 of the 10 components of the metric can be determined from field equations, This is related to the fact that the choice of the coordinate system is arbitrary, that is, the metrics g_{ij} and $\tilde{g}_{ij} = b^r_i b^s_j\, g_{rs}$ both are solutions of (6.3). The choice of the 4 arbitrary functions $\tilde{x}^k(x)$ is equivalent to fix 4 arbitrary conditions on the metric, as well as the gauge invariance permits o fix a condition on the electromagnetic potential.

In many problems, a convenient choice of ''gauge" is the following (see Section 8):

$$g^{ij}\,\Gamma^k_{ij} = 0\,, \qquad\qquad \text{de Donder condition.}$$

It is interesting to observe that the quantities G^0_k do not depend on the second derivatives of time parameter and so the corresponding field equations are not "evolution equations", but constraints on initial conditions. In fact one has

$$0 = \nabla_k G^k_j = \nabla_0 G^0_j + \nabla_a G^a_j \quad\Longrightarrow\quad \partial_0 G^0_j = -\partial_a G^a_j - \Gamma^k_{kl} G^l_j + \Gamma^l_{kj} G^k_l\,.$$

The last member in equation above depends almost on second derivative of metric (with respect to time) and so G^0_j depends almost on first derivative of metric (with respect to time).

6.3 The action for gravitation

The action in a curved manifold has to be expressed as the integral of a scalar quantity in the invariant volume $\sqrt{g}\,d^4x$. In order to obtain (6.3), the scalar has to depend on the first derivatives of the metric. The only non-trivial scalar which can be built up with the metric is the scalar curvature R (and of course its powers), which however depends on second derivatives too. Then one could expect field equations depending on the third derivatives of the metric, but this is not the case, because the contribution of the second derivatives to the action is a 4-divergence of a vector and it does not contribute to the field equations if the variation of the metric vanishes on the boundary of the integration domain, as required by the action principle.

By recalling that

$$\Gamma^j_{kj} = \frac{1}{2}\, g^{ij} \partial_k\, g_{ij} = \partial_k \ln\sqrt{g}\,,$$
$$g^{ij}\Gamma^k_{ij} = -g^{kl}\Gamma^j_{lj} - \partial_l g^{kl} = -\partial_l (g^{lk} \ln\sqrt{g})\,,$$

after tedious but straightforward calculations one gets

$$\sqrt{g}\,R = \sqrt{g} g^{ij}\left(\Gamma^r_{is}\Gamma^s_{jr} - \Gamma^l_{kl}\Gamma^k_{ij}\right) + \partial_k\, f^k\,, \qquad\qquad f^k = \sqrt{g}\left(g^{ij}\Gamma^k_{ij} - g^{kj}\Gamma^l_{jl}\right)\,.$$

Then we try the action

$$S_g[g, \partial_k\, g] \;=\; -\frac{c^3}{16\pi G} \int R\,\sqrt{g}\,d^4x \qquad\qquad \text{(Einstein-Hilbert action)}$$
$$\;=\; -\frac{c^3}{16\pi G} \int \sqrt{g} g^{ij} \left(\Gamma^r_{is}\Gamma^s_{jr} - \Gamma^l_{kl}\Gamma^k_{ij}\right)\, d^4x - \frac{c^3}{16\pi G} \int \partial_k\, f^k\, d^4x\,,$$

where the integral is done over a 4-dimensional region and the variation of the metric vanishes on the boundary of such a hyper-surface. The last integral vanishes due to Gauss theorem. The first integral on the right-hand side is called Einstein-Hilbert action[33].

To complete the theory we must add to S_g the action S_m of all other fields. The interaction between matter/radiation fields with gravitation is obtained by replacing ordinary with covariant derivatives (*minimal coupling*) and η_{ij} with g_{ij}. As already anticipated above, the variation of S_m with respect to the metric gives the energy-momentum tensor, that is the right member of (6.3), while the variation of S_g gives rise the first one. The constant factor in front of the integral is chosen in order to get the correct equations when interaction is considered.

The Einstein-Hilbert action or alternatively the non-invariant one give rise to the same field equations. By considering a small variation of the metric $\delta g^{ij}(x)$ we have

$$\delta\, S_g \equiv S[g+\delta g] - S[g] \;=\; -\frac{c^3}{16\pi G} \int \delta(\sqrt{g}\, R)\, d^4x$$
$$\;=\; -\frac{c^3}{16\pi G} \int \left[R\delta\left(\sqrt{g}\right) + \sqrt{g}\,\delta\left(g^{ij} R_{ij}\right)\right]\, d^4x\,.$$

Recall that

$$\frac{dg}{g} = -g_{ij}\, dg^{ij} \quad\Longrightarrow\quad \delta\sqrt{g} = -\frac{1}{2}\,\sqrt{g}\, g_{ij}\delta g^{ij}$$

and

$$\delta R^i_{jrs} \;=\; \delta\partial_r\Gamma^i_{sj} + \delta\left(\Gamma^i_{rk}\Gamma^k_{sj}\right) - (r \leftrightarrow s) = \partial_r\delta\Gamma^i_{sj} + \Gamma^i_{rk}\delta\Gamma^k_{sj} + \Gamma^k_{sj}\delta\Gamma^i_{rk} - (r \leftrightarrow s)$$
$$\;=\; \nabla_r\delta\Gamma^i_{sj} - \nabla_s\delta\Gamma^i_{rj}\,. \tag{6.10}$$

Since we are considering a variation in form of the metric at the given point x, $\delta\Gamma$ is a tensor because it is the difference between two connections at the same point, that is

$$\delta\Gamma^k_{ij}(x) = \Gamma^k_{ij}(g+\delta g) - \Gamma^k_{ij}(g)\,.$$

Contracting (6.10) we obtain (δ^k_i is the Kronecker tensor)

$$\delta R_{ij} = \delta R^k_{ikj} = \nabla_k\left(\delta\Gamma^k_{ij} - \delta^k_i\delta\Gamma^l_{jl}\right)\,, \qquad\qquad g^{ij}\delta\, R_{ij} = \nabla_k\left(g^{ij}\delta\Gamma^k_{ij} - g^{kj}\delta\Gamma^l_{jl}\right) = \nabla_k V^k\,,$$

V^k being a controvariant vector. Recalling (3.22) we see that such a term does not contribute to the field equations, because the variations vanish on the boundary of the region of integration.

Finally we get

$$\delta S_g[g] \;=\; -\frac{c^3}{16\pi G} \int \left(R_{ij} - \frac{1}{2}\, R\, g_{ij}\right)\sqrt{g}\,\delta g^{ij}\, d^4x\,,$$
$$S_m[\Phi, g] \;=\; \frac{1}{2c} \int T_{ij}\,\sqrt{g}\,\delta g^{ij}\, d^4x\,.$$

Due to the arbitrariness of the variations, from $0 = \delta S = \delta S_g + \delta S_m$ the Einstein equations (6.3) follow. By Φ we have indicated the collection of all matter/radiation fields.

[33] David Hilbert (Russia) 1862-1943.

6.4 The matter energy-momentum tensor

Let us consider an infinitesimal coordinate transformation (at first order is a "translation")

$$\tilde{x}^k = x^k + \xi^k(x) \quad \Longrightarrow \quad \begin{cases} a^i_j = \frac{\partial \tilde{x}^i}{\partial x^j} = \delta^i_j + \partial_j \xi^i \,, \\ b^i_j = \frac{\partial x^i}{\partial \tilde{x}^j} = \delta^i_j - \partial_j \xi^i + o([\xi^i]^2) \,. \end{cases}$$

At first order in ξ, the relation between $\tilde{g}^{ij}(\tilde{x})$ and $g^{ij}(x)$ can be obtained in two different ways, that is by means of coordinate transformation or by Taylor expansion. One has respectively

$$\tilde{g}^{ij}(\tilde{x}) \sim g^{ij}(x) + g^{ik}\partial_k \xi^j + g^{jk}\partial_k \xi^i \,,$$

$$\tilde{g}^{ij}(\tilde{x}) \sim \tilde{g}^{ij}(x) + \xi^k \partial_k \tilde{g}^{ij}(x) \sim \tilde{g}^{ij}(x) + \xi^k \partial_k g^{ij}(x) \,,$$

and comparing the two equations above

$$\delta g^{ij} = \tilde{g}^{ij}(x) - g^{ij}(x) = -\xi^k \partial_k g^{ij} + g^{ik}\partial_k \xi^j + g^{jk}\partial_k \xi^i \,, \tag{6.11}$$

which can be written in the covariant form

$$\delta g^{ij} = g^{ik}\nabla_k \xi^j + g^{jk}\nabla_k \xi^i \,, \qquad \delta g_{ij} = -(\nabla_i \xi_j + \nabla_j \xi_i) \,. \tag{6.12}$$

Now we compute the variation of the matter action $S_m[\Phi, g]$ due to an infinitesimal transformation of coordinates $x^k \to x^k + \xi^k$. The action is a scalar quantity and so it is invariant under a coordinate transformation, that is $\delta_\xi S = \delta_\xi S_g = \delta_\xi S_m = 0$, $\delta_\xi S$ being the variation due to a change of coordinates.

The action depends implicitly on coordinates through metric and matter fields. An infinitesimal transformation of coordinates induces a transformation $\delta_\xi g^{ij}$ and $\delta_\xi \Phi$ on metric and matter fields respectively. The first one $\delta_\xi g^{ij}$ is given in (6.12), while the second one $\delta_\xi \Phi$ in principle can be computed for any field but for our aim it is not necessary to have its form explicitly.

We have

$$0 = \delta_\xi S_m = \delta_{g_\xi} S_m + \delta_{\Phi_\xi} S_m = 0 \,, \qquad \delta_\xi g^{ij} = g^{ik}\nabla_k \xi^j + g^{jk}\nabla_k \xi^i \,, \tag{6.13}$$

where by $\delta_{g_\xi} S_m$ and $\delta_{\Phi_\xi} S_m$ we indicate the variations of matter action due to the induced variation of metric and matter fields respectively.

We choose the infinitesimal vectors ξ^k to vanish on the boundary of integration region. With this choice both $\delta_\xi g^{ij}$ and $\delta_\xi \Phi$ will be vanish on the same boundary. Moreover, we assume Φ to be solution of matter fields equations.

By definition, the field equations for matter/radiation are the solutions of $\delta_\Phi S_m = 0$ and so they satisfy the Euler-Lagrange[34] equations

$$\frac{\partial \mathcal{L}}{\partial \Phi} - \nabla_k \frac{\partial \mathcal{L}}{\partial \nabla_k \Phi} = 0 \,,$$

because for an arbitrary variation $\delta\Phi$, vanishing on the boundary of the region of integration, one has

$$\delta_\Phi S_m = \frac{1}{c}\int \sqrt{g}\left(\frac{\partial \mathcal{L}}{\partial \Phi}\,\delta\Phi + \frac{\partial \mathcal{L}}{\partial \nabla_k \Phi}\,\delta\nabla_k \Phi\right) d^4x = \frac{1}{c}\int \sqrt{g}\left(\frac{\partial \mathcal{L}}{\partial \Phi} - \nabla_k \frac{\partial \mathcal{L}}{\partial \nabla_k \Phi}\right)\delta\Phi \, d^4x \,.$$

[34] Leonhard Euler (Switzerland) 1707-1783; Joseph-Louis Lagrange (Italia) 1736-1813.

If Φ is solution of field equations then $\delta_\Phi S_m = 0$. As a consequence also $\delta_{\Phi_\xi} S_m = 0$, because this is a particular variation of fields, and so, for an infinitesimal transformation of coordinates

$$0 = \delta_{g_\xi} S_m \equiv \frac{1}{2c}\int \sqrt{g}\, T_{ij}\bar{\delta} g^{ij}\, d^4x = \frac{1}{c}\int \sqrt{g}\, T^i_j \nabla_i \xi^j\, d^4x = -\frac{1}{c}\int \sqrt{g}\, \nabla_i T^i_j \xi^j\, d^4x\,,$$

where in the last integral an integration by parts has been performed, taking into account that ξ^k vanishes on the boundary.

Due to the arbitrariness of ξ^k it follows

$$\nabla_i T^{ij} = 0\,, \qquad T_{ij} = T_{ji} = \frac{2}{\sqrt{g}}\left[\frac{\partial}{\partial g^{ij}} - \partial_k \frac{\partial}{\partial \partial_k g^{ij}}\right]\left(\sqrt{g}\,\mathcal{L}_m\right)\,. \tag{6.14}$$

The tensor T_{ij} will be identified with the energy-momentum tensor of matter/radiation fields because it follows from the invariance of the action with respect to coordinate transformations.

It has to be observed that it is a symmetric tensor as required by Einstein equations and it is possible to show that it is equal to canonical energy-momentum tensor of quantum field theory, after symmetrisation via Belifante procedure.

6.5 Killing vectors

We have seen that under an arbitrary infinitesimal coordinate transformation $x^k \to x^k + \xi^k$ the metric changes according to

$$\delta g_{ij} = -\left(\nabla_i \xi_j + \nabla_j \xi_i\right)\,, \tag{6.15}$$

and so it is invariant, that is $\delta g_{ij} = 0$, if ξ is a *Killing vector*. This means that it is a solution of the *Killing equation*[35]

$$\nabla_i \xi_j + \nabla_j \xi_i = 0\,. \tag{6.16}$$

The number of Killing vectors is related to the symmetries of the manifold we are dealing with, but in any case it cannot to be greater than $N(N+1)/2$, N being the dimension. In 4-dimensions, the number of Killing vectors is less or equal to ten.

Manifolds having the maximal number of Killing vectors are called *maximally symmetric spaces*. An important example is given by the Minkowski space. It has ten Killing vectors, which generate the four translations and the six rotations (the Poincaré group). Other examples are the spaces with constant curvature (hyper-spheres and hyperbolic manifolds).

6.5.1 Example: energy-momentum tensor for electromagnetic field

Compute energy-momentum tensor for electromagnetic field with arbitrary gravity.

The action can be directly obtained from the one of special relativity by minimal coupling. So

$$S_m = \frac{1}{c}\int \mathcal{L}_m \sqrt{|g|}\, d^4x\,, \qquad \mathcal{L}_m = -\frac{1}{16\pi} F^{rs} F_{rs} = -\frac{1}{16\pi} g^{ij} g^{rs} F_{ir} F_{js}\,.$$

Recalling that $F_{ij} = \nabla_i A_j - \nabla_j A_i = \partial_i A_j - \partial_j A_i$ we get

$$T_{ij} = \frac{2}{\sqrt{g}}\left[\frac{\partial}{\partial g^{ij}} - \partial_k \frac{\partial}{\partial \partial_k g^{ij}}\right]\left(\sqrt{g}\,\mathcal{L}\right) = -\frac{1}{4\pi} g^{rs} F_{ir} F_{js} + \frac{1}{16\pi} g_{ij} F^{rs} F_{rs}\,.$$

Finally we have the known expression

$$T^j_i = \frac{1}{4\pi}\left[-F_{ik} F^{jk} + \frac{1}{4}\delta^j_i F^{rs} F_{rs}\right]\,.$$

[35] Wilhelm Karl Joseph Killing (Germania) 1847-1923.

6.6 Energy-momentum and angular momentum for gravitation

In Minkowski space, the continuity equation $\partial_\mu T^{\mu\nu} = 0$ gives rise to the conservation of energy-momentum and angular momentum of the corresponding field. In fact, integrating on a spatial region V and using the Gauss theorem one has

$$\partial_0 \int_V T^{0\nu}\, d^3x \;=\; \int_V \partial_0 T^{0\nu}\, d^3x = -\int_V \partial_a T^{a\nu}\, d^3x = -\int_\Sigma n_a T^{a\nu}\, d\Sigma\,,$$

where Σ is the boundary of V with unit outward vector $\hat{n} \equiv n^a$ $(a = 1, 2, 3)$. Now, if the energy-momentum tensor has compact support in V, then the last intgral is vanishing and so energy-momentum P^μ and angular momentum $J^{\mu\nu}$ are consrved quantities. They are given by

$$P^\mu = \frac{1}{c}\int_V T^{0\mu}\, dV\,, \qquad\qquad J^{\mu\nu} = \frac{1}{c}\int_V (x^\mu T^{0\nu} - x^\nu T^{0\mu})\, dV\,.$$

In general relativity, the energy-momentum tensor of matter field satisfies the *covariant continuity equation* $\nabla_k T^k_j = 0$, to which in general does not correspond conserved quantities, because

$$\nabla_k T^k_j = \partial_k T^k_j + \Gamma^k_{kl} T^l_j - \Gamma^l_{kj} T^k_l = \frac{1}{\sqrt{g}}\,\partial_k(\sqrt{g}T^k_j) - \Gamma^l_{kj} T^k_l\,.$$

This fact is not surprising because also the gravitational field possesses energy and momentum and so it is reasonable to expect that the whole energy and the whole momentum are conserved (matter+gravity).

We have to look for a quantity τ^{ij}, depending on matter and gravity, satisfying a continuity equation to which correspond conserved quantities. It is easy to understand that we shall fall in trouble, because τ^{ij} cannot be a tensor satisfying a covariant equation. In fact, in such a case we go back to previous result with τ^{ij} in place of T^{ij}. Then we have to renounce to the tensoriality of τ^{ij}, or alternatively to have a covariant continuity equation or both of them. In any case we shall have serious problems with interpretation (the gravitational energy problem). It is not possible to say how much energy there exists in a finite region, because it depends on the reference frame, but nevertheless it is possible to define the total energy of gravitation in the whole universe.

There are several proposal which permit to define energy and momentum of gravitation. All of them are based on "pseudo-tensors" t_{ij} for the gravitational field, which locally are different, but they give the same results for energy, momentum and angular momentum of the whole gravitational field.

Here we first give a definition of t_{ij} for a general gravitational field (Landau-Lifsits pseudo-tensor) and then we shall discuss a second quite simple and "intuitive" definition, which however can be used only for asymptocally flat manifods, because it is based on the existence of a "quasi-Minkowskian" reference frame.

6.6.1 The Landau-Lifsits energy-momentum pseudo-tensor

First of all we observe that in a locally inertial reference frame (more generally in any frame where $\partial_k g_{ij} = 0\,,\ \partial_r\partial_s g_{ij} \neq 0$ in the considered point) one has

$$T^{ij} = \partial_k \eta^{kij} = \frac{c^4}{8\pi G\,|g|}\,\partial_k \mathcal{Q}^{kij} = \frac{c^4}{8\pi G|g|}\,\partial_k \mathcal{Q}^{kji}\,, \qquad \begin{cases} \eta^{ikj} = -\eta^{jki}\,, \\ \mathcal{Q}^{ikj} = -\mathcal{Q}^{jki}\,. \end{cases} \tag{6.17}$$

In fact, using (5.4) and (3.16), after straightforward calculations one gets

$$G^{ij} = \frac{1}{2|g|}\,\partial_k\partial_l\,\Big[|g|(g^{ij}g^{kl} - g^{ik}g^{jl})\Big] = \frac{1}{|g|}\,\partial_k \mathcal{Q}^{kij} = \frac{1}{|g|}\,\partial_k \mathcal{Q}^{kji}\,, \tag{6.18}$$

$$\mathcal{Q}^{kij} = -\mathcal{Q}^{jik} = \frac{1}{2}\,\partial_l\,\left[|g|(g^{ij}g^{kl} - g^{ik}g^{jl})\right]\,, \tag{6.19}$$

which give rise to (6.17) via Einstein equation. It as to be noted that $\mathcal{Q}^{kij}$ is not symmetric with respect to the last two indices.

Since $\mathcal{Q}^{kij}$ is not a tensor, in a general reference frame equation (6.17) is no more valid. We shall have

$$\frac{c^4}{8\pi G|g|}\,\partial_k\mathcal{Q}^{kij} - T^{ij} = t_L^{ij} = t_L^{ji} \neq 0\,,$$

where t_L^{ij} is the *Landau-Lifsits energy-momentum pseudo-tensor*. By definition it is vanishing in a local reference frame and so it depends on first derivative of the metric only. It can be written as a complicated quadratic form in the Levi-Civita connection, which can be derived from the following identity after a long and tedious calculation (see Landau-Lifsits)

$$t_L^{ij} = \frac{c^4}{8\pi G|g|}\,\partial_k\mathcal{Q}^{kij} - G^{ij}\,. \tag{6.20}$$

Ones t_L^{ij} has been computed, Einstein equations assumes the form

$$\partial_k\mathcal{Q}^{kij} = \frac{8\pi G}{c^4}\,\tau_L^{ij}\,, \qquad\qquad \tau_L^{ij} = |g|(T^{ij} + t_L^{ij})\,,$$

Due to the antisymmetry of $\mathcal{Q}^{ikj} = -\mathcal{Q}^{jki}$ one has

$$\partial_j\partial_k\mathcal{Q}^{kij} = \partial_j\partial_k\mathcal{Q}^{kji} = 0 \quad\Longrightarrow\quad \partial_i\tau^{ij} = 0\,,$$

from which we derive the "conserved" quantities

$$P^j = \frac{1}{c}\int_V \tau^{0j}\,dV\,, \qquad\qquad J^{ij} = \frac{1}{c}\int_V (x^i\tau^{0j} - x^j\tau^{0i})\,dV\,. \tag{6.21}$$

To be more precise, momentum and angular momentum are conserved if the integration is done over the whole spatial section and at same time τ^{ij} goes to zero on the boundary. In such case we get, for example

$$\frac{dP^j}{dt} = -\int_V \partial_a\tau^{aj}\,dV = -\int_\Sigma n_a\tau^{aj}\,d\Sigma \to 0\,.$$

It as to be noted that, as well as t^{ij}, P^j and J^{ij} are not tensors with respect to general coordinate transformations, but they are tensors with respect to linear (in particular to Lorentz) transformations. This means that the quantities above, for example P^j, depend on reference frame. Also in the absence of gravitation, t^{ij} is different from zero in non Minkowskian coordinates.

We finally observe that in an asymptotically flat manifold, energy-momentum and angular momentum of gravitational field are determined by the asymtotic behaviour of the field if we choose a reference frame which is Minkowskian at infinity. In fact in such hypothesis it is reasonable to assume $|g_{ij} - \eta_{ij}| \sim 1/r$ and looking at (6.19) one gets $\mathcal{Q}^{kij}| = O(1/r^2)$. Choosing a ball of radius r and boundary $\Sigma(r)$ as integration region in (6.21), we have

$$P^j = \frac{c^3}{8\pi G}\int_V \partial_a\mathcal{Q}^{a0j}\,dV = \frac{c^3}{8\pi G}\int_{\Sigma(r)} n_a\mathcal{Q}^{a0j}d\Sigma = \frac{c^3r^2}{8\pi G}\int_{S_2} n_a\mathcal{Q}^{a0j}d\sigma\,, \tag{6.22}$$

where S_2 is the unitary sphere and $d\sigma = \sin\vartheta\,d\vartheta\,d\varphi$. The last integral above is convergent and in the limit $r \to \infty$ gives the total energy-momentum of the field, which is a Lorentz 4-vector. This corresponds to energy-momentum as measured by the Minkowskian observer at infinity. A Similar equation holds for angular momentum too.

6.6.2 Energy-momentum in asymptotically Minkowskian manifolds

Now we give a different definition of energy-momentum for gravitation, which is more "intuitive" with respect to the previous one, but it works only on manifolds which asymptotically concide with Minkowski. If this is the case, we can choose a "quasi-Minkowskian" reference frame where

$$g_{ij} = \eta_{ij} + h_{ij}\,, \qquad |h_{ij}| \to 0\,,\ \text{(at infinity)}, \qquad g_{ij} \to \eta_{ij}\,,\ \text{(at infinity)}, \tag{6.23}$$

and we can write the Einstein tensor in the form

$$G_{ij}(g_{rs}) = G_{ij}(\eta_{rs} + h_{rs}) = G^{(1)}_{ij} + \left(G_{ij} - G^{(1)}_{ij}\right)\,,$$

where $G^{(1)}_{ij}$ is the lowest contribution in h_{rs} in a Taylor series expansion of Gij. It is linear in the second derivatives of h_{rs}. In fact, at first order one gets

$$\Gamma^k_{ij} = \frac{1}{2}\, g^{kl}\left(\partial_i g_{jl} + \partial_j g_{il} - \partial_l g_{ij}\right) = \frac{1}{2}\,\eta^{kl}\left(\partial_i h_{jl} + \partial_j h_{il} - \partial_l h_{ij}\right) + O(h^2) \tag{6.24}$$

$$\begin{aligned} R^i_{jrs} &= \partial_r \Gamma^i_{sj} - \partial_s \Gamma^i_{rj} + \Gamma^i_{rl}\Gamma^l_{sj} - \Gamma^i_{sl}\Gamma^l_{rj} \\ &= \frac{1}{2}\,\eta^{il}\left[\partial_j\left(\partial_r h_{sl} - \partial_s h_{rl}\right) - \partial_l\left(\partial_r h_{sj} - \partial_s h_{rj}\right)\right] + O(h^2) \end{aligned} \tag{6.25}$$

$$R_{ij} = R^k_{ikj} = -\frac{1}{2}\left[\Box h_{ij} + \partial_i\partial_j h - \partial_i\partial_k h^k_j - \partial_j\partial_k h^k_i\right] + O(h^2) \equiv R^{(1)}_{ij} + O(h^2) \tag{6.26}$$

$$R = g^{ij} R_{ij} = -\Box h + \partial_i\partial_j h^{ij} + O(h^2) \equiv R^{(1)} + O(h^2) \tag{6.27}$$

$$\begin{aligned} G_{ij} &= \frac{1}{2}\left[\partial_i\partial_k h^k_j + \partial_j\partial_k h^k_i - \partial_i\partial_j h - \Box h_{ij} - \eta_{ij}\left(\partial_r\partial_s h^{rs} - \Box h\right)\right] + O(h^2) \\ &\equiv G^{(1)}_{ij} + O(h^2) \end{aligned} \tag{6.28}$$

Here and in the rest of this section, all indices are rised and lowered by using the metric η_{ij} and $\Box$ represents the D'Alembertian operator in Minkowski space, that is

$$h^k_j = \eta^{ki} h_{ij}\,, \qquad h^{rs} = \eta^{ri}\eta^{sj} h_{ij} \qquad h = \eta^{ij} h_{ij}\,, \qquad \Box = \eta^{rs}\partial_r\partial_s\,.$$

By $O(h^2)$ we mean corrections of the order equal or greater than h^2 and its derivatives. More precisely, in (6.24) the quadratic corrections start with $O(h\partial h)$ and so the corrections in (6.25)-(6.28) starts with $O(\partial h\partial h; h\partial^2 h)$.

Using this notation, the Einstein equations can be set in the form

$$G^{(1)}_{ij} = \frac{8\pi G}{c^4}\,\left(T_{ij} + t_{ij}\right) = \frac{8\pi G}{c^4}\,\tau_{ij}\,, \tag{6.29}$$

where

$$t_{ij} = t_{ji} = -\frac{c^4}{8\pi G}\left(G_{ij} - G^{(1)}_{ij}\right)\,, \qquad t_{ij} = O(h^2) \equiv O(\partial h\partial h; h\partial^2 h)\,. \tag{6.30}$$

From (6.29) we see that τ_{ij} is the source of the tensorial field h_{ij} and of course it depends on h_{ij} itself, because the gravitational field is auto-interacting (the 4-momentum is the "charge" of gravitation).

In this approach in which gravitation is considered as a "tensor" field on Minkowski space, the quantity t_{ij} is interpreted as the "energy-momentum tensor" of the h_{ij} field. Note however that $G^{(1)}_{ij}$, h_{ij} and t_{ij} are not true tensors with respect to general coordinate transformations.

As it can be trivially verified using equation (6.28), the quantity $G^{(1)}_{ij}$ satisfies the exact identity (sometime called linearised Bianchi identity) $\eta^{ik}\partial_k G^{(1)}_{ij} = 0$ and so the "energy-momentum tensor" τ_{ij} has to satisfy the continuity equation

$$\partial_i \tau^{ij} = 0\,, \qquad \tau^{ij} = \eta^{ir}\eta^{js}\tau_{rs}\,, \tag{6.31}$$

which gives rise to the (conserved) quantities

$$\begin{cases} P^k = \frac{1}{c}\int_V \tau^{0k}\, d^3x\,, & \frac{dP^k}{dt} = -\int_\Sigma n_a \tau^{ak}\, d\Sigma\,, \\ J^{ij} = \frac{1}{c}\int_V (\tau^{0i}x^j - \tau^{0j}x^i)\, d^3x\,, & \frac{J^{ij}}{dt} = -\int_\Sigma n_a(\tau^{ai}x^j - \tau^{aj}x^i)\, d\Sigma\,, \end{cases} \tag{6.32}$$

which respectively represent the 4-momentum and the angular momentum of matter plus gravitation. The quantities $V, \Sigma, \hat{n} \equiv n_a$ have the same meaning as above.

For the hypothesis done on the metric, for $r \to \infty$ we shall reasonable have $|h_{ij}| \sim 1/r$ and as a consequence $t_{ij} \sim 1/r^4$ because t_{ij} is at least quadratic in the derivatives of h_{ij}. Moreover, T_{ij} has a compact support and so the convergence of integrals in (6.32) is assured and surface integrals do not give contributions, in the limit $r \to \infty$. The momentum and angular momentum of a closed gravitational system are conserved.

It has to be remarked that as well as τ_{ij}, such quantities are not tensors, but nevertheless they are Lorentz covariant. This means that they are tensors with respect to the Lorentz group (we are considering a tensor field theory in Minkowski).

Using (6.18) or directly (6.28) of course one gets

$$G^{ij}_{(1)} = \partial_k\, Q^{kij}_{(1)} = \partial_k\, Q^{kji}_{(1)}\,, \qquad Q^{kij}_{(1)} = Q^{jik}_{(1)}\,, \tag{6.33}$$

where $Q^{kij}_{(1)}$ is the lowest contribution in the series expansion of (6.19).

Performing the integrals in (6.32) in a ball of large radius r with boundary $\Sigma(r)$ and using the Gauss theorem as above we get

$$P^k = \frac{c^3}{8\pi G}\int_V \partial_i Q^{i0k}_{(1)}\, dV = \frac{c^3}{8\pi G}\int_{\Sigma(r)} n_a Q^{a0k}_{(1)}\, d\Sigma = \frac{c^3 r^2}{8\pi G}\int_{S_2} n_a Q^{a0k}_{(1)}\, d\sigma\,, \tag{6.34}$$

which is the total momentum of fields inside the ball. A similar equation can be written for the angular momentum J_{ij}.

- The interesting thing to note is that, while the integrals in (6.32) are done on an 3-dimensional hypersurface $t = constant$, the integral in (6.34) is done on a 2-dimensional spherical surface of radius r. This means that energy, momentum and angular momentum are determined only by the asymptotic behaviour of the field.

- Another important feature of such quantities is that they are invariant with respect to coordinate transformations which leave the metric in the quasi-Minkowskian form (6.23).

6.6.3 Example: energy of Schwarzschild gravitational field

Now we compute the energy for the solution we shall derive in Section 7.2, equation (7.9).

The metric in spherical coordinates is

$$ds^2 = -\left(1 - \frac{r_S}{r}\right) dt^2 + \frac{dr^2}{1 - r_S/r} + r^2 d\sigma^2 \,, \qquad r_S = \frac{2MG}{c^2} \,.$$

In order to compute the energy of the field, first of all we have to write it in "quasi-Minkowskian" coordinates form, that is

$$ds^2 = g_{ij} dx^i dx^j = (\eta_{ij} + h_{ij})\, dx^i dx^j \,, \qquad x^k \equiv (x^0, x^1, x^2, x^3) \equiv (t, x, y, z) \,.$$

To this aim we set $\vec{r} = x\hat{u}_x + y\hat{u}_y + z\hat{u}_z = x^1 u_1 + x^2 u_2 + x^3 u_3$, $\hat{u}_a$ being the unit vectors along the axis, and observe that

$$\begin{cases} r = \sqrt{\delta_{ab}\, x^a x^b} \,, \\ dr = \beta_a dx^a \,, \\ dr^2 + r^2 d\sigma^2 = \delta_{ab} dx^a xs dx^b \,, \end{cases} \qquad \Longrightarrow \qquad r^2 d\sigma^2 = (\delta_{ab} - \beta_a \beta_b)\, dx^a dx^b \,,$$

where

$$\beta^a = \frac{x^a}{r} = \frac{x^a}{\sqrt{\delta_{ab}\, x^a x^b}} \,, \qquad \beta^a \beta_a = 1 \,, \qquad \frac{\partial \beta^a}{\partial x^b} = \frac{1}{r} \left(\delta^a_b - \beta^a \beta_b\right) \,, \qquad \partial_a f(r) = \beta_a f'(r) \,.$$

Note that there are no difference between covariant and controvariant spatial indices since the spatial components of the metric are δ_{ab}.

In these coordinates we get

$$\begin{aligned} ds^2 &= \eta_{ij} dx^i dx^j + \frac{r_S}{r} (dx^0)^2 + \frac{r_S/r}{1 - r_S/r} \beta_a \beta_b dx^a dx^b \\ &= \eta_{ij} dx^i dx^j + \varepsilon(r)\, (dx^0)^2 + \varepsilon(r)\, \beta_a \beta_b\, dx^a dx^b + O(\varepsilon)^2 \,, \end{aligned} \tag{6.35}$$

where we have set $\varepsilon(r) = r_S/r$.

The total energy can be computed using (6.34) or (6.22). Also in this latter case we can approximate the metric up to first order in h_{ij} since all terms in $\mathcal{Q}^{kij}$ which in the limit $r \to \infty$ goes to zero more quickly that $1/r^2$ do not give contributions to the integral. Then we get

$$g_{ij} = \eta_{ij} + h_{ij} \,, \qquad h_{0a} = 0 \,, \qquad h_{00} = \varepsilon(r) \,, \qquad h_{ab} = \varepsilon(r)\, \beta_a \beta_b \,,$$

$$g^{ij} \sim \eta^{ij} - h^{ij} \,, \qquad h^{0a} = 0 \,, \qquad h^{00} = \varepsilon(r) \,, \qquad h^{ab} = \varepsilon(r)\, \beta^a \beta^b \,.$$

From (6.22) it follows

$$P^0 = \frac{c^3 r^2}{8\pi G} \int_{S_2} n_a Q^{a00}\, d\sigma \,, \qquad \hat{n} = \frac{\vec{r}}{r} \quad \Longrightarrow \quad n_a = \frac{x_a}{r} = \beta_a \,.$$

The quantity Q^{a00} has to be computed up to the order $1/r^2 \sim \varepsilon'(r)$. Using (6.19) one obtains

$$\mathcal{Q}^{a00} = \frac{1}{2} \partial_b \left(|g|\, g^{00} g^{ab}\right) \sim -\frac{1}{2} \partial_b \left(|g| \delta^{ab} + h^{00}\, \delta^{ab} - h^{ab}\right) \,.$$

Now

$$\partial_b h^{00} = \partial_b \varepsilon(r) = \beta_b \, \varepsilon'(r) \, ,$$

$$\partial_b h^{ab} = \beta^a \beta^b \partial_b \varepsilon(r) + \frac{\varepsilon(r)\beta^b}{r} \, (\delta^a_b - \beta^a \beta_b) + \frac{\varepsilon(r)\beta^a}{r} \, (\delta^b_b - \beta^b \beta_b) = \beta^a \varepsilon' + 2\beta^a \, \frac{\varepsilon(r)}{r} \, .$$

Recalling (3.16) we also get

$$\partial_b \, |g| = g^{ij} \partial_b \, g_{ij} \sim -\partial_b h_{00} + \delta^{cd} \partial_b h_{cd} = -\partial_b \left[\varepsilon(r) \, (1 - \delta^{ab} \beta^c \beta^d) \right] = 0 \, ,$$

and so

$$\mathcal{Q}^{a00} \sim \beta^a \, \frac{\varepsilon(r)}{r} = \beta^a \frac{r_S}{r^2} \, .$$

Using this in (6.22) we finally obtain

$$P^0 = Mc \quad \Longrightarrow \quad E = Mc^2 \, .$$

As expected, the total energy of the static field is equal to the rest energy of the body.

One can also verify that the momentum $P^a = 0$, as expected. In fact, applying (6.19) to this special case it follows

$$\mathcal{Q}^{a0b} = -\frac{1}{2} \, \partial_0 \, \left(|g| \, g^{00} g^{ab} \right) = 0 \, .$$

7 Exact Solutions of Einstein Equations

Here we shall compute exact solutions of Einstein equations and we shall discuss in some detail the physical consequences.

7.1 Spatial spherical symmetry

We start with solutions which represent the gravitational field generated by sources with spatial spherical symmetry (central potential). This means that it has to be possible to find "quasi Minkowskian coordinates" $(t, \vec{x})$, which permit to write the interval ds in a form invariant with respect to the (spatial) orthogonal group. The interval has to be a quadratic form in dx^k, depending only on time and rotational invariants

$$\vec{x} \cdot \vec{x}\,, \qquad \vec{x} \cdot d\vec{x}\,, \qquad d\vec{x} \cdot d\vec{x}\,.$$

It is convenient to use "polar coordinates" (r, ϑ, φ)

$$r = \sqrt{\vec{x} \cdot \vec{x}}\,, \qquad \vec{x} \cdot d\vec{x} = r\, dr\,,$$

$$d\vec{x} \cdot d\vec{x} = dr^2 + r^2 \left(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2\right) = dr^2 + r^2\, d\sigma^2\,,$$

$d\sigma^2$ being the metric of the unitary sphere. In such coordinates the general form for the interval is

$$ds^2 = \alpha(t,r)\, dt^2 + \beta(t,r)\, dr^2 + \gamma(t,r)\, d\sigma^2 + 2\delta(t,r)\, dr\, dt\,,$$

$\alpha, \beta, \gamma, \delta$ being arbitrary functions of t and r. The coordinates are not yet fixed. In fact, without breaking the symmetry, we can perform coordinate transfomations of the kind

$$(t, r, \vartheta, \varphi) \to (f_1(\tilde{t}, \tilde{r}), f_2(\tilde{t}, \tilde{r}), \vartheta, \varphi)\,,$$

With a suitable choice of the functions f_1, f_2 the metric can be diagonalised and put in the "standard" form

$$ds^2 = -B\, dt^2 + A\, dr^2 + r^2 \left(d\vartheta^2 + \sin^{\vartheta}\, d\varphi^2\right)\,, \tag{7.1}$$

$A = A(t,r)$ and $B = B(t,r)$ being arbitrary functions to be determined by solving Einstein's equations. For physical reasons the signature has to be $(-,+,+,+)$ and so A, B must be positive functions.

Now we show that the transformation which permits to write the metric in standard form effectively exists. First af all, we perform the transformation $(t, r, \vartheta, \varphi) \to (\tilde{t}, \tilde{r}, \vartheta, \varphi)$ by means of $\tilde{t} = t$ and $\tilde{r}^2 = \gamma$. In this way

$$ds^2 = \tilde{\alpha}(\tilde{t}, \tilde{r})\, d\tilde{t}^2 + \tilde{\beta}(\tilde{t}, \tilde{r})\, d\tilde{r}^2 + \tilde{r}^2\, d\sigma^2 + 2\tilde{\delta}(\tilde{t}, \tilde{r})\, d\tilde{r}\, d\tilde{t}\,.$$

By a second transformation $(\tilde{t}, \tilde{r}, \vartheta, \varphi) \to (t, r, \vartheta, \varphi)$ of the kind $(\tilde{t} = f(t,r)\,,\ r = \tilde{r})$, with

$$\frac{df}{dr} = -\frac{\tilde{\delta}}{\tilde{\alpha}}\,, \qquad \frac{df}{dt} \neq 0\,,$$

the standard form directly follows.

7.2 The Schwarzschild solution

This is the first exact solution of Einstein equations (1916). It gives the gravitational field outside a spherical symmetric source (external solution). Due to the symmetry of the source, we can put the metric in the simplified form (7.1) and compute the positive functions $A(t,r), B(t,r)$.

Of course, the energy-momentum tensor of the source has support in a spatial region $r < R_0$, R_0 being the radius of the spherical source, which in principle could depend on time (the source could contract/expand). For $r > R_0$ the energy-momentum tensor is vanishing and so we have to look for solutions of Einstein equations in vacuum, that is with $T_{ij} = 0$. Using (6.6) these read

$$R_{ij} = 0\,, \tag{7.2}$$

which in principle correspond to ten second-order differential equations in the metric, but, as we have already said, only six of them are independent. In our case, due to symmetry, we expect less independent equations, because the metric has only two free parameters.

Now we have to compute all components of Ricci tensor using (3.14), (5.2) and (5.6). The computation is quite tedious but straightforward. It is convenient to distinguish between temporal/radial $(p,q,r,s = 0,1)$ and angular $(a,b,c,d = 2,3)$ indices. As usual $i,j,k,l = 0,1,2,3$.

Since the metric in (7.1) is diagonal we have $g^{ij} = 1/g_{ij}$ and so

$$g_{00} = -B\,, \qquad g_{11} = A\,, \qquad g_{ab} = r^2\hat{g}_{ab}\,,$$
$$g^{00} = -\frac{1}{B}\,, \qquad g^{11} = \frac{1}{A}\,, \qquad g^{ab} = \frac{\hat{g}^{ab}}{r^2}\,,$$

where "hatted" quantities are related to the unitary sphere S^2 (see example 5.6). We obtain

$$\begin{aligned}
\Gamma^c_{ab} &= \tfrac{1}{2}\,\hat{g}^{cd}\left(\partial_a\hat{g}_{bd} + \partial_b\hat{g}_{ad} - \partial_d\hat{g}_{ab}\right) = \hat{\Gamma}^c_{ab}\,,\\
\Gamma^p_{ab} &= -\tfrac{1}{2}\,g^{pq}\partial_q g_{ab}\,,\\
\Gamma^a_{pb} &= \tfrac{1}{2}\,g^{ac}\partial_p g_{bc}\,,\\
\Gamma^a_{pq} &= -\tfrac{1}{2}\,g^{ac}\partial_c g_{pq}\,,\\
\Gamma^p_{aq} &= \tfrac{1}{2}\,g^{pr}\partial_a g_{pr}\,,\\
\Gamma^r_{pq} &= \tfrac{1}{2}\,g^{rs}\left(\partial_p g_{qs} + \partial_q g_{ps} - \partial_s g_{pq}\right)
\end{aligned} \tag{7.3}$$

and the non-vanishing components of connection read

$$\begin{aligned}
&\Gamma^0_{00} = \tfrac{\dot{B}}{2B}\,, && \Gamma^0_{01} = \Gamma^0_{10} = \tfrac{B'}{2B}\,, && \Gamma^0_{11} = \tfrac{\dot{A}}{2B}\,,\\
&\Gamma^1_{00} = \tfrac{B'}{2A}\,, && \Gamma^1_{01} = \Gamma^1_{10} = \tfrac{\dot{A}}{2A}\,, && \Gamma^1_{11} = \tfrac{A'}{2A}\,,\\
&\Gamma^1_{ab} = -\tfrac{r}{A}\,\hat{g}_{ab}\,, && \Gamma^a_{1b} = \Gamma^a_{b1} = \tfrac{1}{r}\,\delta^a_b\,, && \Gamma^c_{ab} = \hat{\Gamma}^c_{ab}\,,\\
&\hat{\Gamma}^2_{33} = -\tfrac{1}{2}\,\sin 2\vartheta\,, && \hat{\Gamma}^2_{23} = \hat{\Gamma}^2_{32} = -\cot\vartheta\,, && \hat{\Gamma}^3_{23} = \hat{\Gamma}^3_{32} = -\cot\vartheta\,,
\end{aligned} \tag{7.4}$$

where "dot" and "prime" means derivative with respect to t and r respectively.

We start with the computation of the R_{01} component. Using previous expressions we get

$$R_{01} = R_{10} = R^k_{1k0} = \partial_p\Gamma^p_{10} - \partial_1\Gamma^p_{p0} + \Gamma^a_{a1}\Gamma^1_{10} + \Gamma^p_{pq}\Gamma^q_{10} - \Gamma^p_{1q}\Gamma^q_{p0} = \frac{\dot{A}}{rA}\,.$$

According to (7.2) the latter quantity has to be vanish and this happens if

$$\dot{A} = 0 \quad\Longrightarrow\quad A = A(r)\,.$$

Such a condition now can be used to simplify other equations. One verifies that all components of R_{ij} with $i \neq j$ are identically vanishing, while the components with $i = j$ read

$$0 = R_{00} = \frac{B''}{2A} - \frac{B'}{4A}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{B'}{rA}\,, \tag{7.5}$$

$$0 = R_{11} = -\frac{B''}{2B} + \frac{B'}{4B}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{A'}{rA}\,, \tag{7.6}$$

$$0 = R_{22} = 1 + \frac{r}{2A}\left(\frac{A'}{A} - \frac{B'}{B}\right) - \frac{1}{A}\,, \tag{7.7}$$

$$0 = R_{33} = \sin^2\vartheta\, R_{22}\,. \tag{7.8}$$

Among the four differential equations above, only two of them are linearly independent and these determine the two unknown functions A and B.

Dividing (7.5) by B and (7.6) by A and summing up the two expressions we obtain

$$\frac{R_{00}}{B} + \frac{R_{11}}{A} = \frac{1}{rA}\left(\frac{A'}{A} + \frac{B'}{B}\right)\,,$$

which is vanishing if

$$\frac{A'}{A} + \frac{B'}{B} = 0 \quad \Longrightarrow \quad \frac{\partial}{\partial r}(AB) = 0 \quad \Longrightarrow \quad AB = f(t)\,.$$

With a transformation of coordinates depending only on the temporal variable $t \to g(t)$ (this does not change the standard form of the metric) it is always possible to put $f(t) = 1$, that is $A = 1/B$, which corresponds to a particular choice of the time parameter. Such a transformation is solution of the equation $f(g(t))\,(\dot{g}(t))^2 = 1$. In fact with such a transformation

$$f(t) = A(t,r)B(t,r) \to A(g(t),r)B(g(t),r)\dot{g}^2(t) = f(g(t))\dot{g}^2(t) = 1\,.$$

Finally, using (7.7) we get

$$R_{22} = 1 - rB' - B = 0 \quad \Longrightarrow \quad \frac{d}{dr}\,(rB) = 1 \quad \Longrightarrow \quad B = 1 - \frac{r_S}{r}\,,$$

where $-r_S$ is an integration constant, which will be fixed by taking the Newtonian limit. For weak field and small velocity in fact we have (see (2.12)

$$-B = g_{00} \sim -\left(1 + \frac{2\Phi}{c^2}\right) = -\left(1 - \frac{2MG}{c^2 r}\right) \quad \Longrightarrow \quad r_S = \frac{2MG}{c^2}\,,$$

where M is the mass of the body which generates the gravitational field. The quantity r_S is called the *Schwarzschild radius.*

The final solution has then form

$$ds^2 = -\left(1 - \frac{r_S}{r}\right)\,dt^2 + \frac{1}{1 - r_S/r}\,dr^2 + r^2(d\vartheta^2 + \sin^2\vartheta\,d\varphi^2)\,. \tag{7.9}$$

This is called the *Schwarzschild metric.* It represents the gravitational field outside a spherical symmetric body. At the spatial infinity $r \gg r_S$ the metric (7.9) will coincide with the one of Minkowski and this means that very far from the source the special relativity holds true.

- The spherical source which generates the gravitational field can have arbitrary dimension R_0. For planets and visible stars one has $R_0 \gg r_S$ and since the Schwarzschild solution is valid for $r > R_0$, the metric (7.9) has no singularities. The situation drastically change when $R_0 < r_S$. In fact, in such a case the *static metric* (7.9), is *singular* for $r = r_S$. However it has to be noted that this is not a "true" singularity of the field because in such a point all invariant quantities are finite (of course Ricci tensor and scalar curvature are vanishing, but for example the scalar $R^{ijrs}R_{ijrs}$ is equal to $12r_S^2r^{-6}$ and so it is singular only at $r = 0$). The singularity of the metric is due to the "particular choice of coordinates". When we cross the surface $r = r_S$ (the *horizon*), both g_{00} and g_{11} change the sign. In particular, for $r < r_S$ g_{00} becomes positive, while g_{11} becomes negative. This means that the time-like Killing vector ∂_t becomes space-like, while the space-like Killing vector ∂_r becomes time-like. Such a feature will be discussed in more detail in chapter 7.8 which is dedicated to the physics of black holes.

  In the case of a "point-like source", the metric in (7.9) has a second singularity at $r = 0$, which effectively corresponds to a "physical" singularity, which is due to the fact the density of matter has to be infinite at the origin ($\varepsilon(r) = M\,\delta(r)$).

- It has also to be noted the curious fact that the classical escape velocity of a particle at a distance $d = r_S$ from the body is equal to the speed of light (this seems a pure coincidence).

7.3 The classical tests of Einstein equations

In the Newtonian theory there are three classical phenomena which are in disagreement with experimental data, even if one takes into account of possible corrections due to special relativity. On the contrary general relativity is in excellent agreement with all that phenomena. Recently other tests have confirmed the validity of general relativity.

1. **Precession of perihelia:** it has been observed that the perihelia (the point on the orbit nearest the sun) of mercury is not fixed with respect to distant stars, but it "advance", the precession angle being $\Delta\varphi_O = 5600.73 \pm 0.41$ seconds per century. This means that the orbit is not exactly elliptic, but it precesses around a focus (see figure 3). Even if one takes into account of perturbations due to the presence of other planets and the rotation of the earth, the Newtonian theory gives a precession angle $\Delta\varphi_N = 5557.62 \pm 0.20$ seconds per century. The discrepancy between the measured value and the one provided by Newtonian theory is then $43.11 \pm 0.45''$ per century. This could be due to a modification of Newtonian law, for example as a consequence of the solar oblateness or other unknown effects. Of course the effect is present in all planets, but for mercury it is more evident because such a planet has a very eccentric orbit and moreover it is quite near the sun.

2. **The deflection of light by the sun:** as well as massive particles, photons move along geodesic and so they interact with gravitational field. Then we expect the light to be deflected by the sun. This effect can be observed by looking at a distant star when the sun is between the earth and the star itself and six month later, when the sun is on other side (see figure (4) (the first data was taken during an eclipse in the year 1919).

   For a massive particle, the deflection angle can be computed using Newtonian theory. It depends on the starting velocity at infinity and the impact parameter, but (of course) not on the mass of the particle. Then in the formula one can put the speed of light and obtains in this way a formula valid also for photons. As we shall explicitly see, general relativity gives a deflection angle which is double than the one computed by using Newtonian theory.

Figure 3: precession of perihelia-aphelia (figure by G. t'Hooft)

3. **The radar echo delay:** this concerns the time which takes an electromagnetic signal to go from the earth to a inner planet and back. As we shall see, this is longer than what expected if light traveled in straight lines at constant velocity (Shapiro experiment (1964).

7.4 The orbit of a test particle in Newtonian theory

Before to find the geodesics of the metric (7.9), we study the motion of a particle in Newtonian theory, then we consider a test particle in a central field $V(r) = m\phi(r) = -mMG/r$. As it is well known, the angular momentum L and the energy E are conserved quantities. They read

$$L = mr^2\dot{\varphi} = \text{const}\,, \qquad E = \frac{1}{2}\,m(\dot{r}^2 + r^2\dot{\varphi}^2) + V = \frac{m\dot{r}^2}{2} + \frac{L^2}{2mr^2} + V = \text{const}\,. \tag{7.10}$$

The differential equation which determines the trajectory of the particle can be easily obtained from equations in (7.10) by eliminating the time variable. This gives

$$d\varphi = \frac{L\,dr}{r^2}\left[2m(E-V) - \frac{L^2}{r^2}\right]^{-1/2} \implies \varphi(r) - \varphi(r_0) = \int_{r_0}^{r} \frac{Ldr}{r^2\sqrt{2m(E-V) - L^2/r^2}}\,. \tag{7.11}$$

For an open orbit (see figure 4) one has $L = mbv_\infty$ and $E = mv_\infty^2/2$ where b is the impact parameter and v_∞ the velocity very far from the scattering center. Choosing the system of coordinates as in figure 4) ($\phi = 0$ corresponds to the point on the trajectory nearest the source), the deflection angle read

$$\chi_N = |2\phi(\infty) - \pi|\,, \qquad \varphi(\infty) = b\int_{r_0}^{\infty} \frac{dr}{r\sqrt{r^2 + 2MGr/v_\infty^2 - b^2}}\,, \tag{7.12}$$

where r_0 is the minimum value (on the trajectory) of the coordinate r. This means that $\phi(r_0) = 0$ and $\frac{dr}{dt}(r_0) = 0$, then r_0 it is the positive solution of the equation

$$2m(E-V) - \frac{L^2}{r^2} = 0 \implies 2mE = \frac{L^2}{r_0^2} - \frac{2m^2MG}{r_0}\,.$$

The integral in (7.12) can be done exactly, the primitive of the integrand function being

$$\varphi(r) = \arcsin\frac{2C + Br}{r\sqrt{\Delta}}\,, \qquad B = \frac{2MG}{v_\infty^2}\,, \qquad C = -b^2\,, \qquad \Delta = B^2 - 4C\,. \tag{7.13}$$

Since the angle does not depend on the mass of the particle, the latter formula can be used for a photon too, by putting $v_\infty = c$. In this way $B = 2MG/c^2 = r_S$.

The experimental impact factor is (more or less) equal to the radius o the sun, that is $b \sim r_0 \sim R_\odot$ and $r_S \ll R_\odot$. For our purposes it is sufficient to find an approximate solution at the first order in $r_S/R_\odot$. From (7.13) then it follows

$$\varphi(\infty) \sim \frac{\pi}{2} + \frac{r_S r_0}{2b^2} \sim \frac{\pi}{2} + \frac{r_S}{R_\odot}$$

and finally $\chi_N \sim r_S/R_\odot$. As we shall see below, this result is exactly one-half with respect to the one given by general relativity.

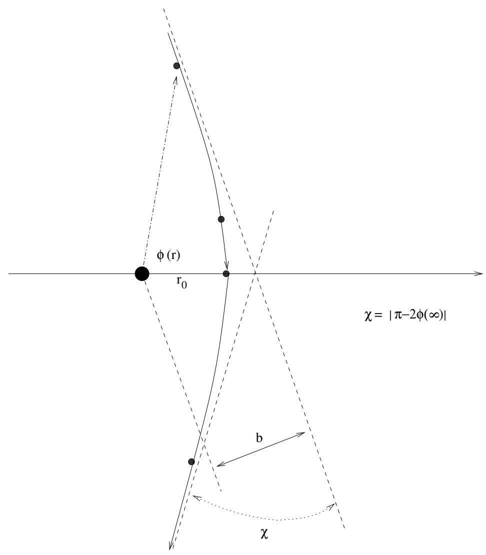

Figure 4: deflection of light

7.5 The orbit of a test particle in general relativity

A test particle moves along a geodesic $x^k(\lambda)$ given by the equation

$$\frac{Du^k}{d\lambda} = \frac{du^k}{d\lambda} + \Gamma^k_{ij}\, u^i\, u^j = 0\,, \qquad u^k = \frac{dx^k}{d\lambda}\,, \tag{7.14}$$

where λ is an arbitrary affine parameter (not the proper time, because we shall compute the trajectories for photons too) and u^k represents the 4-velocity (it is exactly the 4-velocity if $\lambda = \tau$). Now we look for solutions of (7.14) for a generic static and isotropic metric of the form

$$ds^2 = -Bdt^2 + Adr^2 + r^2\, d\sigma^2 = g_{ij} dx^i dx^j\,,$$

where $A = A(r)$ and $B = B(r)$. We have seen above that this metric is solution of Einstein equations in vacuum if $A = 1/B$, but for more generality, we do not use such a result (see the *post-Newtonian approximation in 7.6*).

Then we use connection (7.4) with $\dot{A} = \dot{B} = 0$ and re-write (7.14) in the form

$$\frac{du^0}{d\lambda} + \Gamma^0_{ij} u^i u^j = \frac{du^0}{d\lambda} + \frac{B'}{B} u^0 u^1 = 0\,, \tag{7.15}$$

$$\frac{du^1}{d\lambda} + \Gamma^1_{ij} u^i u^j = \frac{du^1}{d\lambda} + \frac{B'}{2A}(u^0)^2 + \frac{A'}{2A}(u^1)^2 - \frac{r}{A}\left[(u^2)^2 + \sin^2\vartheta (u^3)^2\right] = 0\,, \tag{7.16}$$

$$\frac{du^2}{d\lambda} + \Gamma^2_{ij} u^i u^j = \frac{du^2}{d\lambda} + -\frac{\sin 2\vartheta (u^3)^2}{2} + 2\cot\vartheta u^2 u^3 + \frac{2u^1 u^2}{r} = 0\,, \tag{7.17}$$

$$\frac{du^3}{d\lambda} + \Gamma^3_{ij} u^i u^j = \frac{du^3}{d\lambda} + 2\cot\vartheta u^2 u^3 + \frac{2u^1 u^3}{r} = 0\,. \tag{7.18}$$

As it happens in the classical case, for symmetry reasons we expect the trajectory of the particle to belong to a plane, say (x, y). Then we look for solutions with $\vartheta = \pi/2$. This choice implies $u^2 = d\vartheta/d\lambda = 0$ and so equation (7.17) is identically satisfied. Recalling also that $u^1 = dr/d\lambda$, the other three equations simplify to

$$\frac{d}{d\lambda}\,\log u^0 = -u^1 \frac{d}{dr}\,\log B = -\frac{d}{d\lambda}\,\log B\,, \tag{7.19}$$

$$\frac{du^1}{d\lambda} = -\frac{1}{2A}\left[B'(u^0)^2 + A'(u^1)^2\right] + \frac{r(u^3)^2}{A}\,, \tag{7.20}$$

$$\frac{d}{d\lambda}\,\log u^3 = -\frac{2u^1}{r} = -\frac{d}{d\lambda}\,\log r^2\,. \tag{7.21}$$

Now equations (7.19) and (7.20) can be trivially solved and in fact

$$\begin{aligned} \frac{d}{d\lambda}\,\log(Bu^0) = 0 \quad &\Longrightarrow \quad u^0 = \frac{dx^0}{d\lambda} = \frac{cost}{B}\,, \\ \frac{d}{d\lambda}\,\log(r^2u^3) = 0 \quad &\Longrightarrow \quad r^2u^3 = r^2\frac{d\varphi}{d\lambda} = cost = J\,. \end{aligned} \tag{7.22}$$

The first integration constant is related to the possible choices of the λ parameter and so we put it equal to 1. This means that

$$\frac{dx^0}{d\lambda} = \frac{1}{B} \quad \Longrightarrow \quad \lambda \to x^0 = ct \text{ at large distances where } B \to 0\,.$$

In this case the evolution parameter tends to the Newtonian time very far from the source.

The second integration constant J, which we call *angular momentum*, at large distances effectively becomes the angular momentum per unit mass, in units where $c = 1$.

There exists a second conserved quantity, the energy, which can be derived from (7.20), but more quickly by recalling that the scalar product is invariant under parallel transport. So

$$0 = \frac{D}{d\lambda}\,\left(u^k u_k\right) = \frac{d}{d\lambda}\,\left(u^k u_k\right) \quad \Longrightarrow \quad u^k u_k = \frac{ds^2}{d\lambda^2} = -\mathcal{E}\,.$$

The constant $\mathcal{E}$ is positive for massive particles and vanishing for massless particles and we shall see that it is related to the classical energy.

From previous equations now we get the explicit relation

$$\frac{ds^2}{d\lambda^2} = -\frac{1}{B} + A\left(\frac{dr}{d\lambda}\right)^2 + \frac{J^2}{r^2} = -\mathcal{E}\,.$$

In conclusion, the trajectory is determined by the two equations

$$J = r^2\frac{d\varphi}{d\lambda}\,, \tag{7.23}$$

$$\mathcal{E} = \frac{1}{B} - A\left(\frac{dr}{d\lambda}\right)^2 - \frac{J^2}{r^2} = \frac{1}{B}\left[1 - \frac{A}{c^2B}\left(\frac{dr}{dt}\right)^2\right] - \frac{J^2}{r^2}\,, \tag{7.24}$$

where in the last expression we have explicitly set $\lambda = ct$. At large distances from the source the coordinate time t will coincide with the Minkowskian time.

Now we eliminate the parameter λ and obtain the exact solution

$$\varphi(r) - \varphi(r_0) = J\int_{r_0}^{r} \frac{\sqrt{AB}\,dr}{r^2\sqrt{1 - B\mathcal{E} - BJ^2/r^2}}\,. \tag{7.25}$$

The meaning of the constant of motion can be understood in the Newtonian limit, where the metric becomes the one of Minkowski. We have

$$B = -g_{00} \sim 1 + \frac{2\phi}{c^2}\,, \qquad \lambda \sim ct\,, \qquad \phi = -\frac{MG}{r}\,,$$

$$J \sim \frac{r^2}{c}\frac{d\varphi}{dt} \quad \Longrightarrow \quad J = \frac{L}{c}\,,$$

$$\frac{1}{2}\left[\left(\frac{dr}{dt}\right)^2 + \frac{L^2}{r^2}\right] + \phi = \frac{c^2(1-\mathcal{E})}{2} \quad \Longrightarrow \quad \mathcal{E} = 1 - \frac{2E}{c^2}\,,$$

where L and E are Newtonian, non-relativistic quantities (angular momentum and energy per unit mass).

7.5.1 Particular solutions: closed circular orbits

We look for circular orbits, this means $r = r_0$ =constant. The values of r_0 can be determined by equation (7.24) which becomes

$$-\frac{J^2}{r_0^2} + \frac{1}{B(r_0)} = \mathcal{E}\,.$$

The derivative of (7.24) with respect to r at $r = r_0$ reads

$$-\frac{B'(r_0)}{b^2(r_0)} + \frac{2J^2}{r_0^3} = 0$$

and so

$$\mathcal{E} = \frac{1}{B(r_0)}\left(1 - \frac{r_0 B'(r_0)}{B(r_0)}\right)\,, \qquad J^2 = \frac{r_0^3 B'(r_0)}{B^2(r_0)}\,.$$

From equations above one derives $\mathcal{E}$ and J starting from the dimension r_0 of the orbit.

For massless particles, $\mathcal{E} = 0$ and so $r_0 B'(r_0) = B(r_0)$, from which one gets

$$r_0 = \frac{3}{2} r_S\,.$$

We see that the light can travel along a circular closed orbit with a radius equal to 3/2 the Schwarzschild radius.

Of course all considerations above holds for $r \geq R_\odot$, $R_\odot$ being the radius of the body which generates the field.

7.6 The post-Newtonian approximation

One starts from a generic static and isotropic metric, but without to assume the validity of Einstein equations and one develops the functions A and B in power series of r_S/r, that is

$$ds^2 = -B(r)(dx^0)^2 + A(r)dr^2 + r^2(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2)\,, \tag{7.26}$$

$$A(r) = 1 + \gamma\,\frac{r_S}{r} + \ldots\,, \qquad B(r) = 1 - \alpha\,\frac{r_S}{r} + \frac{\beta - \alpha\gamma}{2}\left(\frac{r_S}{r}\right)^2 + \ldots\,,$$

α, β, γ being arbitrary parameters to be determined from experimental data.

From the principle of equivalence it follows $g_{00} \sim -(1 - 2MG/c^2 r)$ and so $\alpha = 1$. If we assume also the validity of Einstein equations (6.3)), then $\gamma = \beta = 1$. For the moment, for more generality we do not fix β and γ, but we put $\alpha = 1$ as it is imposed by the principle of equivalence.

At the first order in r_S/r we get

$$A(r) \sim 1 + \gamma \frac{r_S}{r}\,, \qquad B(r) \sim 1 - \frac{r_S}{r}\,.$$

Conserved quantities and trajectory for test particle/photon are given by equations (7.23), (7.24) and the (7.25), that is

$$J = r^2 \frac{d\varphi}{d\lambda}\,, \qquad \mathcal{E} = \frac{1}{B} - A\left(\frac{dr}{d\lambda}\right)^2 - \frac{J^2}{r^2}\,. \tag{7.27}$$

$$\varphi(r) - \varphi(r_0) = J \int_{r_0}^{r} \frac{\sqrt{AB}\,dr}{r^2\sqrt{1 - B(r)\mathcal{E} - BJ^2/r^2}}\,. \tag{7.28}$$

7.6.1 The deflection of light

For massless particles we have $\mathcal{E} = 0$ and $J = L/c = b$, where b is the impact parameter (see figure 4). The deflection angle is

$$\chi = |2\varphi(\infty) - \pi|\,, \qquad \varphi(\infty) = J\int_{r_0}^{\infty} \frac{\sqrt{AB}\,dr}{r^2\sqrt{1 - BJ^2/r^2}}\,,$$

r_0 being the minimum value of the r coordinate and $\varphi(r_0) = 0$. From equation $dr/d\lambda = 0$ we get

$$J^2 = \frac{r_0^2}{B(r_0)} = \frac{r_0^2}{1 - r_S/r_0} \sim r_0^2(1+\varepsilon)\,,$$

where we have set $\varepsilon = r_S/r_0$. Putting $x = r_0/r$ we obtain

$$\begin{aligned}
&\sqrt{AB} \sim \sqrt{(1+\gamma\varepsilon x)(1-\varepsilon x)} \sim 1 + \frac{\varepsilon(\gamma-1)}{2}\,,\\
&\left[1 - \frac{B(r)J^2}{r^2}\right]^{-1/2} \sim \left[1 - (1-\varepsilon x)(1+\varepsilon)x^2\right]^{-1/2} \sim \frac{1}{\sqrt{1-x^2}}\left(1 + \frac{\varepsilon x^2}{2(1+x)}\right)\,,
\end{aligned} \tag{7.29}$$

from which

$$\begin{aligned}
\varphi(\infty) &\sim \sqrt{1+\varepsilon}\int_0^1 \left[1 + \frac{\varepsilon(\gamma-1)x}{2} + \frac{\varepsilon x^2}{2(1+x)}\right] \frac{dx}{\sqrt{1-x^2}}\\
&= \sqrt{1+\varepsilon}\left[\arcsin x - \varepsilon\sqrt{\frac{1-x}{1+x}} - \frac{\varepsilon x}{2}\sqrt{\frac{1-x}{1+x}} - \frac{\varepsilon}{2}\arcsin x - \frac{\varepsilon(\gamma-1)}{2}\sqrt{1-x^2}\right]_0^1\\
&\sim \left[\frac{\pi}{2} + \varepsilon + \frac{\varepsilon(\gamma-1)}{2}\right]\,.
\end{aligned} \tag{7.30}$$

Now, choosing $r_0 \sim R_\odot$, at the first order in ε we obtain

$$\chi \sim \frac{2r_S}{R_\odot} + \frac{(\gamma-1)r_S}{R_\odot}\,.$$

As we already said, for general relativity ($\gamma = 1$) such a result is twice the one computed using Newtonian theory. For the sun one gets the value $\chi \sim 1.75$ seconds, which is in good agreement with the experimental data.

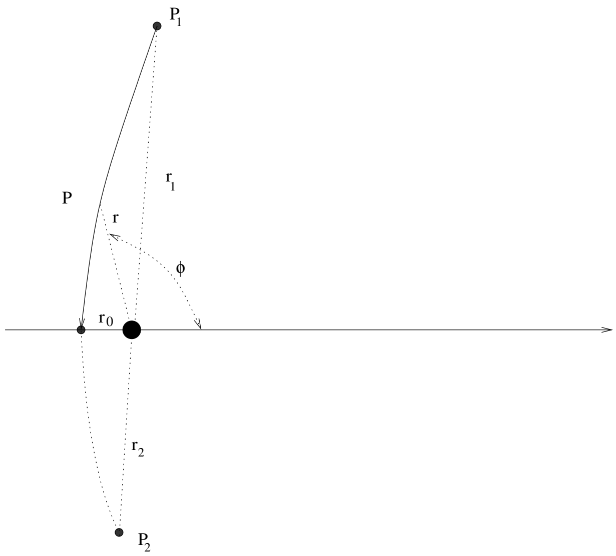

Figure 5: radar echo delay

7.6.2 Radar echo delay

We would like to compute the time $t(r_1, r_2)$ which takes a photon to go from the point $P_1 \equiv (r_1, \pi/2, \varphi_1)$ to the point $P_2 \equiv (r_2, \pi/2, \varphi_2)$ in the presence of the gravitational field generated by a star (see figure 5). In order to make the calculation we first compute the time $t(r, r_0)$ which the signal takes to reach the nearest point to the star $P_0 \equiv (r_0, \pi/2, 0)$, starting from an arbitrary point $P \equiv (r, \pi/2, \varphi)$ on the trajectory (note that we refer to the specular trajectory with respect to the one in the picture 5). The minimal distance r_0 from the star is directly related to the angular momentum because the energy vanishes ($\mathcal{E} = 0$). The coordinate time can be obtained by integrating $dx^0 = d\lambda/B$, which is an exact differential form.This means that the integral does not depend on the path (in a stationary field).

Putting $x = r_0/r$, at the first order in $\varepsilon = r_S/r_0$ we get

$$
\begin{aligned}
dx^0 = \frac{d\lambda}{B} &= \sqrt{\frac{A}{B}}\,\frac{dr}{\sqrt{1 - BJ^2/r^2}} \\
&\sim -\left[1 + \frac{(\gamma+1)\varepsilon x}{2}\right]\left[1 + \frac{\varepsilon x^2}{2(1+x)}\right]\frac{r_0\,dx}{x^2\sqrt{1-x^2}} \,,
\end{aligned}
$$

from which it follows

$$
\begin{aligned}
c\,t(r, r_0) &= \int_{r_0}^{r} \sqrt{\frac{A}{B}}\,\frac{dr}{\sqrt{1 - BJ^2/r^2}} \\
&\sim r_0 \int_{r_0/r}^{1} \left[1 + \frac{(\gamma+1)\varepsilon x}{2} + \frac{\varepsilon x^2}{2(1+x)}\right]\frac{dx}{x^2\sqrt{1-x^2}} \\
&= \sqrt{r^2 - r_0^2} + \frac{r_S(\gamma+1)}{2}\log\left(\frac{r + \sqrt{r^2 - r_0^2}}{r_0}\right) + \frac{r_S}{2}\sqrt{\frac{r - r_0}{r + r_0}} \,.
\end{aligned}
$$

The first term on the right-hand side of the latter equation represents the time which the signal will take if it was moved along a straight line at constant velocity c. As we can see, the true time is greater than that, in contrast with what happens for massive particles.

The equations of motion are invariant with respect to time inversion, then $t(r, r_0) = t(r_0, r)$ and the total time which the signal takes to go from P_1 to P_2 and back is

$$t_{tot} = 2[t(r_1, r_0) \pm t(r_2, r_0)] ,$$

where in the latter equation one has to take the minus sign when the two points are on the same side with respect to the star, while one has to take the plus sign when they are on opposite positions, as in figure 5, In such a latter case one obtains a delay given by

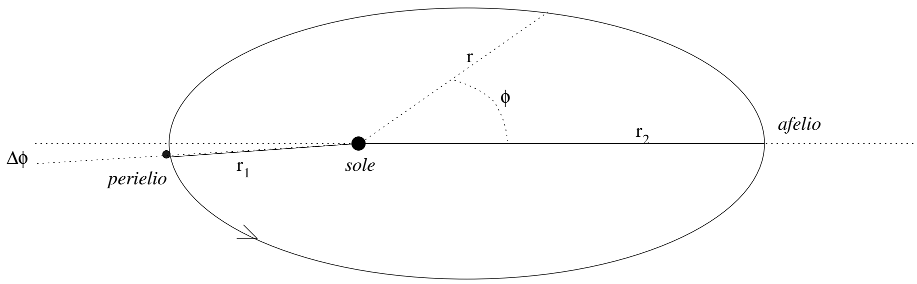

Figure 6: precession of perihelia

$$
\begin{aligned}
\Delta t &= 2\left[t(r_1,r_0)+t(r_0,r_2)-\frac{1}{c}\left(\sqrt{r_1^2-r_0^2}+\sqrt{r_2^2-r_0^2}\right)\right] \\
&= (\gamma+1)\frac{r_S}{c}\left[\log\frac{r_1 r_2}{r_0^2}+\log\left(1+\sqrt{1-\frac{r_0^2}{r_1^2}}\right)+\log\left(1+\sqrt{1-\frac{r_0^2}{r_2^2}}\right)\right] \\
&\quad +\frac{r_S}{c}\left[\sqrt{\frac{r_1-r_0}{r_1+r_0}}+\sqrt{\frac{r_2-r_0}{r_2+r_0}}\right]
\end{aligned} \tag{7.31}
$$

Of course, what one effectively measures is the proper time $\Delta\tau = \sqrt{g_{00}(P_1)}\,\Delta t$.

- In the physical experiment (Shapiro-1964) one measures the time which takes a radar signal to go from the earth to mercury and back and one obtains a delay $\Delta\tau = 240\mu s$, which is in a good agreement with (7.31) if $\gamma = 1$ (general relativity).

7.6.3 The precession of perihelia

Let us consider a particle (mercury) freely moving in the gravitational field generated by a star (sun). We assume the metric to be defined by (7.26) and the orbit of the particle to be a spatial, approximatively closed curve (an ellipse, according to Newtonian theory – see figure 6). Let us indicate by r_1 and r_2 respectively the nearest (perihelia) and the more far away (aphelia) points of the orbit, with respect to the star. In such a points $dr/d\lambda = 0$ and so they are solutions of the equation

$$
f(r) = \frac{1}{B(r)} - \frac{J^2}{r^2} - \mathcal{E} = 0\,, \tag{7.32}
$$

It has to be noted that, in contrast with Newtonian theory, in general the latter algebraic equation has more than two solutions (the orbit is not elliptic), but since for our aim approximated solutions are sufficient, we expand the function $1/B(r)$ in power series up to second order in r_S/r. In such a way (7.32) becomes a second order algebraic equation in the variable $1/r$. In fact

$$
f(r) \sim \left[1-\frac{(\beta-\gamma)}{2}-\frac{J^2}{r_S^2}\right]\frac{r_S^2}{r^2}+\frac{r_S}{r}+1-\mathcal{E} = -C\left(\frac{1}{r}-\frac{1}{r_1}\right)\left(\frac{1}{r}-\frac{1}{r_2}\right)\,, \tag{7.33}
$$

$$
C = J^2 - \left(1-\frac{\beta-\gamma}{2}\right) r_S^2\,.
$$

By the equations $f(r_1)=0$ and $f(r_2)=0$ we get

$$\mathcal{E} = \frac{\frac{r_1^2}{B_1} - \frac{r_2^2}{B_2}}{r_1^2 - r_2^2} \sim 1 + \frac{r_S}{r_1+r_2} + \dots$$
$$J^2 = \frac{\frac{1}{B_1} - \frac{1}{B_2}}{\frac{1}{r_1^2} - \frac{1}{r_2^2}} \sim \frac{r_S L}{2} + \left[1 - \frac{\beta-\gamma}{2}\right] r_S^2 + \dots \qquad (7.34)$$

where we have disregard higher order terms and we have set

$$B_1 = B(r_1)\,, \qquad B_2 = B(r_2)\,, \qquad \frac{1}{L} = \frac{1}{2}\left(\frac{1}{r_1} + \frac{1}{r_2}\right)\,.$$

Using this notation $C = r_S L/2$ and

$$\frac{J^2}{C} \sim 1 + \left[1 - \frac{\beta-\gamma}{2}\right) \frac{r_S^2}{C} = 1 + \left[1 - \frac{\beta-\gamma}{2}\right) \frac{2r_S}{L}\,.$$

The approximated expressions can be obtained directly and more quickly by means of (7.33).

From equation (7.28) of the trajectory we get

$$\begin{aligned}
\varphi(r_2) - \varphi(r_1) &= J \int_{r_1}^{r_2} \frac{\sqrt{A}\,dr}{r^2\sqrt{f}} \sim \frac{J}{\sqrt{C}} \int_{r_1}^{r_2} \frac{\sqrt{A}\,dr}{\sqrt{\left(\frac{1}{r_1} - \frac{1}{r}\right)\left(\frac{1}{r} - \frac{1}{r_2}\right)}} \\
&\sim \sqrt{r_1 r_2}\left[1 + \left(1 - \frac{\beta-\gamma}{2}\right)\frac{r_S}{L}\right] \int_{r_1}^{r_2} \left(1 + \frac{\gamma r_s}{2r}\right) \frac{dr}{r\sqrt{r-r_1}\sqrt{r_2-r}} \\
&= \pi\left[1 + \left(1 - \frac{\beta-\gamma}{2}\right)\frac{r_S}{L}\right]\left(1 + \frac{\gamma r_S}{2L}\right) \\
&\sim \pi + \frac{(2+2\gamma-\beta)\pi r_S}{2L}\,.
\end{aligned}$$

We see that the difference between the angular coordinates of perihelia (r_1) and aphelia (r_2) is greater than π and this means the the two points are not on opposite side with respect to the star (the orbit is not elliptic, it is not closed).

For a complete revolution we obtain

$$\Delta\varphi = 2[\varphi(r_2) - \varphi(r_1)] - 2\pi = \frac{3\pi}{L}\left(\frac{2+2\gamma-\beta}{3}\right)\,.$$

For the system Mercury-Sun, general relativity ($\beta = \gamma = 1$) gives the value $\Delta\varphi = 0.1038$ seconds per revolution. To this value correspond 43.03 seconds per century, which is exactly the value which could not be explained in the framework of Newtonian theory.

- In the numeric computation it has to be taken into account that the revolution is the one of mercury, while the century corresponds to 100 revolutions of the earth equivalent to 415 revolutions of mercury (see table 4).
- As we have already pointed out, the observed precession of mercury is $5600.73''$ per century, but $5025''$ are due to the rotation of the astronomic system of coordinates and $532''$ are due to the perturbation of other planets.

It has also to be noted that general relativity would give negligible corrections (about $\sim 10^{-4}$ seconds per century) to the value $\Delta\varphi_N$ computed in the framework of Newtonian theory as the sum of the two effects above. This means that, in our approximation, it is effectively reasonable to compare the theoretical value $\Delta\varphi$ with $\Delta\varphi_O - \Delta\varphi_N$ (for an exhaustive discussion see Weinberg-1972).

- The last integral above can be easily computed using the residue theorem. To this aim we consider the function

$$F(z) = \frac{1}{z}\left(1 + \frac{\gamma r_S}{2z}\right)\frac{1}{\sqrt{z-r_1}\sqrt{r_2 - z}}$$

and integrate it over a closed path which includes the two singularities r_1 and r_2. The residue theorem gives

$$\oint F(z)\, dz = -2\pi i \sum \text{residue} = -2\int_{r_1}^{r^2} F(r)\, dr\ .$$

In the sum one has to consider all external residue, infinity too, but since the function at infinity goes as $1/z^3$, the residue at that point is vanishing. In order to compute the residue at the origin, where there is a simple pole, it is convenient to make the Laurent expansion. Then

$$\begin{aligned} F(z) &\sim \frac{1}{z}\left(1 + \frac{\gamma r_S}{2z}\right)\frac{1}{\sqrt{-r_1 r_2}}\left(1 + \frac{z}{2r_1}\right)\left(1 + \frac{z}{2r_2}\right) \\ &\sim \frac{1}{\sqrt{-r_1 r_2}}\left(1 + \frac{\gamma r_S}{2L}\right)\frac{1}{z} + \ldots \end{aligned}$$

The integral now read

$$\int_{r_1}^{r^2} F(r)\, dr = \pi \frac{1}{\sqrt{-r_1 r_2}}\left(1 + \frac{\gamma r_S}{2L}\right)\ ,$$

from which the desired result follows.

7.7 The internal solution (spherical symmetry)

Here we shall find a solution of Einstein equations, in the presence of matter, representing a spherical symmetric body (of course in a suitable system of coordinates). The more general form of the metric is given by (7.1), while for the energy-momentum tensor with the required symmetry we choose the one of a perfect fluid, which has the form

$$T^i_j = (p+\varepsilon)\frac{u^i u_j}{c^2} + p\delta^i_j\ , \qquad u^k u_k = -c^2\ ,$$

where ε represents the energy density, p the pressure and $u^k = dx^k/d\tau$ the 4-velocity of the fluid. It is convenient to choose a system of coordinates in which the fluid is at rest, that is $\vec{u} = 0$, $u^k u_k = u^0 u_0 = -c^2$, $u^0 = c/\sqrt{-g_{00}}$ and for simplicity we look for static solutions (the collapse of a star is not considered here). With such assumptions all parameters depend on the r coordinate only and we have (see equation (7.5)-(7.8))

$$ds^2 = -B(r)dt^2 + A(r)dr^2 + r^2 d\sigma^2 \quad \Longrightarrow \quad \begin{cases} R_{00} = \frac{B''}{2A} - \frac{B'}{4A}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{B'}{rA}\,, \\ R_{11} = -\frac{B''}{2B} + \frac{B'}{4B}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{A'}{rA}\,, \\ R_{22} = 1 + \frac{r}{2A}\left(\frac{A'}{A} - \frac{B'}{B}\right) - \frac{1}{A}\,, \\ R_{33} = \sin^2\vartheta\, R_{22}\,. \end{cases}$$

while the non vanishing components of energy-momentum tensor read

$$\begin{cases} T^0_0 = -\varepsilon(r)\,, \\ T^a_b = p(r)\,\delta^a_b\,, \end{cases} \qquad \begin{cases} T_{00} = B(r)\,\varepsilon(r)\,, \\ T_{ab} = p(r)\,g_{ab}\,, \end{cases} \qquad T = T^k_k = 3p(r) - \varepsilon(r)\,.$$

Now we can use (6.3) or alternatively (6.6) to compute the parameters A and B in terms of p and ε. Using (6.3) for G_{00}, after a straightforward calculation we get

$$G_{00} = \frac{8\pi G\, B(r)\,\varepsilon(r)}{c^4} = \frac{B(r)}{r^2}\frac{d}{dr}\left[r\left(1 - \frac{1}{A(r)}\right)\right]. \tag{7.35}$$

The latter equation can be integrated to obtain

$$A(r) = g_{11} = \left[1 - \frac{2M(r)G}{rc^2}\right]^{-1}, \qquad M(r) = \frac{4\pi}{c^2}\int_0^r \varepsilon(y)y^2\,dy\,.$$

The other parameter of the metric can be obtained by solving (6.6) for R_{22}. We have

$$R_{22} = \frac{4\pi G\,(\varepsilon - p)r^2}{c^4} = \frac{r}{2A}\left[\frac{2(A-1)}{r} + \frac{A'}{A} - \frac{B'}{B}\right],$$

and from (7.35)

$$\frac{A'}{A} = \frac{8\pi G\,\varepsilon\, rA}{c^4} - \frac{A-1}{r}\,.$$

Then we get

$$\frac{B'}{B} = \frac{\frac{2M(r)G}{rc^2} + \frac{8\pi G r^2 p(r)}{c^4}}{r\left[1 - \frac{2M(r)G}{rc^2}\right]}\,. \tag{7.36}$$

- It has to be noted that $g_{11} = A(r)$ is “formally” similar to what we have obtained for the external solution, but with M replaced by $M(r)$, which represents “the mass” inside the ball of radius r (not exactly, because it is not the proper mass).

  As well as it happens in Newtonian gravity, we see that the mass outside the surface r does not influence the gravitational field at any point inside that surface, but the same thing is not valid for $g_{00} = -B(r)$.

  For r greater that the dimension of the source ($r > R_0$) one has $p(r) = \varepsilon(r) = 0$ and, as expected, the metric becomes equal to the one of Schwarzschild .

An interesting physical constraint between variables can be obtained by using the four conservation laws (6.4.) One can verify that the unique non trivial one is $\nabla_k T^{k1} = 0$. Using the fact that T^{ij} is diagonal we get

$$\nabla_k T^{k1} = \partial_k T^{k1} + \Gamma^k_{kj} T^{j1} + \Gamma^1_{kj} T^{kj} = \partial_r T^{11} + \Gamma^k_{k1} T^{11} + \Gamma^1_{00} T^{00} + p\, g^{ab}\, \Gamma^1_{ab} \,.$$

Using the Christoffel symbols in (7.3) we have

$$\begin{aligned} 0 &= \partial_r \left(\frac{p}{A}\right) + \left(\frac{B'}{2B} + \frac{A'}{2A} + \frac{2}{r}\right) \frac{p}{A} + \frac{\varepsilon B'}{2AB} + p \left(\frac{A'}{2A^2} - \frac{2}{rA}\right) \\ &= \frac{1}{A} \left(p' + \frac{(p+\varepsilon)B'}{2B}\right) , \end{aligned}$$

and by means of equation (7.36) we finally obtain

$$p' = -(p+\varepsilon)\, \frac{\frac{M(r)G}{rc^2} + \frac{4\pi G r^2 p(r)}{c^4}}{r\left[1 - \frac{2M(r)G}{rc^2}\right]} \,. \tag{7.37}$$

This is the *Tolman-Oppenheimer-Volkoff* equation of the *thermodynamical equilibrium*. Its physical meaning becomes clear in the Newtonian limit $p \ll \varepsilon$, $M(r)G \ll rc^2$ where the equation assumes the simple form

$$p' \sim -\frac{M(r)G\,\varepsilon}{r^2 c^2} \,.$$

The right-hand side of the latter equation represents the gravitational force which the mass $M(r)$ experience on the unitary surface element with mass ε/c^2. The star will contract under the proper gravitational attraction until the internal pressure $p(r)$ will satisfy (7.37). If the mass M of the original star is sufficiently large ($M > M_{TOV}$, $M_{TOV} \sim 3.2\, M_\odot$ being the mass of *Tolman-Oppenheimer-Volkoff* and $M_\odot$ the mass of the sun), then the pressure will not be able to contrast the gravitational force and the collapse will be unavoidable and unstoppable. The radius of the star will became smaller than the Schwarzschild radius and so it will become a *black hole*.

In the usual stars like the sun, the main contribution to the pressure is proportional to the temperature (ordinary pressure of a gas) and as a consequence the thermodynamical equilibrium is broken when the temperature becomes smaller than a critical value depending on the total mass. In such a case the star will start to collapse, the density, as well as the electron degeneracy pressure, will increase. If the value of the mass M is smaller than M_{Ch} ($M_{Ch} \sim 1.44\, M_\odot$ is called the mass of *Chandrasekhar*) then the collapse will stop, the final result being a *white dwarf*. In such a case the gravitational force in balanced by the electron degeneracy pressure.

Also in the case in which the value of the mass M is greater than M_{Ch}, but smaller that M_{TOV} the collaps will stop, the result being a *neutron star*. The density is so high that all electrons are "captured" by protons to form neutrons. The gravitational force in balanced by neutron degeneracy pressure.

- Note that in principle the collapse for stars with $M > M_{TOV}$ could be stopped by quark degeneracy pressure or by unknown effects of quantum gravity.

7.8 Black holes

We recall the Schwarzschild metric

$$ds^2 = -\left(1 - \frac{r_S}{r}\right) dt^2 + \frac{1}{1 - r_S/r} dr^2 + r^2(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2)\,,$$

where $r_S = 2MG/c^2$ and coordinates run over $t \in (-\infty, \infty)$, $\vartheta \in [0, \pi]$, $\varphi \in [-0, 2\pi]$ and $r \in (r_S, \infty)$.

As we have already said above, the singularity of the metric at $r = 0$ corresponds to a physical singularity, because at that point some physical/geometrical quantities diverge, while the singularity of the metric at $r = r_S$ is not a physical singularity because at that point all invariant quantities are finite.

In principle there could be massive bodies with dimensions $R_\odot$ smaller than the corresponding Schwarzschild radius r_S. They are entirely contained inside the spherical surface $r = r_S$, which is called *event horizon*.

From the classical (non-quantistic) point of view, what is inside the event horizon cannot be seen by an observer outside it. For this reason such objects are called *black holes*. According to the principle of equivalence, a very "small" observer in free fall crossing the event horizon will not experience any force, apart tide forces, which in principle could be very small because they are due to the curvature of space-time. In fact, on the horizon one gets

$$R_{ijrs}R^{ijrs} = \frac{12r_S^2}{r^6} \qquad \Longrightarrow \qquad R_{ijrs}R^{ijrs}\Big|_{r=r_S} = \frac{12}{r_S^4} = \frac{3}{4}\,\frac{c^8}{G^4M^4}$$

and the curvature is really small for heavy black holes. For example, on the surface of the earth the "square" of Riemann tensor is of the order $10^{-52}\,cm^{-4}$, while for the the smallest black hole generated by the collapse of a star, this means $M \sim 3M_\odot$, the square of Riemann tensor is of the order $10^{-23}\,cm^{-4}$, but it seams that at the center of galaxies there are black holes with $M \sim 10^8\,M_\odot$. For such supermassive objects the square of Riemann tensor is of the order $10^{-55}\,cm^{-4}$.

Note however that such considerations do not take into account of quantum effects (see the end of this section).

7.8.1 Schwarzschild geometry

The Schwarzschild metric is a solution of Einstein equations in vacuum also for $0 < r < r_S$, this means inside the event horizon, but in such a case the meaning of coordinates become "unclear" because the inner region is not connected with the external one and the interpretation of the results becomes problematic. Temporal and spatial components of the metric change sign and so, when we cross the event horizon, the time-like Killing vector ∂_t becomes space-like, while the space-like Killing vector ∂_r becomes time-like and so they exchange they roles.

Moreover, the geometry of space-time inside the event horizon analysed by means of Schwarzschild coordinates appear to be unphysical, since for example, proper time and proper distance becomes imaginary. The reason of such a very strange behaviour is due to the fact that the Schwarzschild metric describes the geometry of space time as seen by an observer at rest. As we shall see below, inside the event horizon it is impossible to have an observer at rest.

In the external region $r > r_S$, the null radial geodesic ($ds = 0, d\vartheta = 0, d\varphi = 0$) satisfy the equation

$$\frac{dt}{dr} = \pm\frac{1}{1 - r_S/r}\,, \qquad\qquad r > r_S\,.$$

The plus and minus signs refer to outgoing and ingoing geodesic respectively,

The proper time of an observer at rest at the point $r > r_S$ (angular coordinates does not matter), is given by

$$d\tau(r) = \sqrt{|g_{00}|}\, dt = \sqrt{1 - \frac{r_S}{r}}\, dt < dt\,, \qquad\qquad r > r_S\,.$$

This is in agreement with time dilation: the clock in the gravitational field runs more slowly that the clock at infinity where gravitation is absent. In fact, the coordinate time t is also equal to the proper time of the observer at rest at infinity (in units where $c = 1$). This means that for the observer at infinity, the time duration between two events in a point near the horizon will be very large, tending to infinity for $r \to r_S$.

In a similar way, the light emitted by a source at rest at the point $r > r_S$, will arrive at infinity redshifted by the factor $(1 - r_S/r)^{-1/2}$. If λ_0 is the proper wavelength of the light, then the wavelength λ_∞ measured by the observer at rest at infinity is given by

$$\lambda_\infty = \frac{\lambda_0}{\sqrt{1 - r_S/r}}\,,$$

which tends to infinity when r approaches r_S.

The g_{a0} "spatial-temporal" components of the metric are vanishing ($a = 1, 2, 3$) and so the spatial metric, which describes the spatial geometry, is trivially given by

$$\gamma_{ab} = g_{ab} \quad \Longrightarrow \quad d\ell^2 = \frac{dr^2}{1 - r_S/r} + r^2 \left(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2\right)\,.$$

$d\ell$ represents the proper distance between the two points $P_1 = (t, r + dr, \vartheta + d\vartheta, \varphi + d\varphi)$ and $P = (t, r, \vartheta, \varphi)$ as measured by an observer at rest in P at time t. In particular, for two points on the same radial line ($d\vartheta = d\phi = 0$) we have

$$d\ell = \frac{dr}{\sqrt{1 - r_S/r}}\,, \qquad\qquad r > r_S\,,$$

and we see that the proper distance diverge when $r \to r_S$.

In order to study the region inside the horizon it is convenient to use a non singular metric. As we see by looking at above equations, the singularity $r = r_S$ of the metric is not a singularity of the field. One can verify that all invariants quantities, as well as the determinant of the metric, are not singular on the event horizon. Such a surface is a singularity of the metric only and can be eliminated by a suitable choice of coordinates. In fact, there exists a choice of coordinates in which the Schwarzschild metric is regular everywhere, apart the origin $r = 0$, which is a true singularity of the field. The price to pay is to have a non-static metric. There are various metric extensions of that kind, which describe different regions of space-time (Eddington[36] (1924), Lemaître (1938), Finkelstein (1958), Fronsdal (1959)). The maximal extension is due to Kruskal and Szekeres (1960).

7.8.2 The Eddington-Finkelstein coordinates

Starting from Schwarzschild coordinates $(t, r, \vartheta, \varphi)$ one first introduces the Regge-Wheeler or *tortoise* coordinate r_* by

$$\frac{dr}{1 - r_S/r} = dr_* \quad \Longrightarrow \quad r_* = r + r_S \log\left|\frac{r}{r_S} - 1\right|\,, \qquad\qquad r > r_S\,, \tag{7.38}$$

[36] Arthur Stanley Eddington (England) 1882-1944.

which satisfies the properties

$$\lim_{r\to r_s} r_* = -\infty\,, \qquad\qquad \frac{dr^2}{1-r_S/r} = \left(1-\frac{r_S}{r}\right)\, dr_*^2\,.$$

In terms of $(t, r_*, \vartheta, \varphi)$ coordinates the Schwarzschild metric assumes the form

$$ds^2 \;=\; \left(1-\frac{r_S}{r}\right)\left(-dt^2+dr_*^2\right)+r^2(d\vartheta^2+\sin^2\vartheta\, d\varphi^2)\,,$$

from which it follows that the null radial geodesic $(ds=0, d\vartheta=0, d\varphi=0)$ satisfy the equation

$$dt^2 = dr_*^2 \quad\Longrightarrow\quad dt=\pm dr_* \quad\Longrightarrow\quad t\pm r_* = const\,.$$

Then one defines the *ingoing* and *outgoing* null coordinates (also called *advanced* and *retarded* time respectively)

$$v=t+r_*\,,\qquad u=t-r_*\,,\qquad\qquad v\in(-\infty,\infty)\,,\qquad u\in(-\infty,\infty)\,, \tag{7.39}$$

which for $v, u = const$ represent ingoing and outgoing radial null geodesic respectively (straight lines in such coordinates).

In terms of *ingoing/outgoing Eddington-Finkelstein coordinates* the metric reads

$$ds^2 \;=\; -\left(1-\frac{r_S}{r}\right)dv^2+2dv\,dr+r^2(d\vartheta^2+\sin^2\vartheta\, d\varphi^2)\,, \tag{7.40}$$

$$ds^2 \;=\; -\left(1-\frac{r_S}{r}\right)du^2-2du\,dr+r^2(d\vartheta^2+\sin^2\vartheta\, d\varphi^2)\,. \tag{7.41}$$

Of course, for $r > r_S$ both the two metrics above are solutions of Einstein equations in vacuum because they have been obtained from the Schwarzschild one by a regular transformation of coordinates. However, in contrast with Schwarzschild solution, in such coordinate systems there are no metric singularity at $r = r_S$ because both the matrices above and their inverses are regular at $r = r_S$ and so we can analytically extend the coordinates to the values $r \le r_S$, that is

$$0<r<\infty\,;\qquad\qquad -\infty<v<\infty\,;\qquad\qquad -\infty<u<\infty\,.$$

It has to be noted that, after such an extension, the new coordinates (v, r) and (u, r) can not be obtained from the Schwarzschild ones by a regular transformation, because r_* diverges on the horizon. The parameter $t(v, r) = v - r_*$ as a function of the new coordinates, goes to infinity when $r \to r_S$ and goes to zero when $r \to 0$.

By definition, $dv = 0$ represents an ingoing null radial geodesic, that is the geodesic of a massless particle which runs to the singularity at $r = 0$, while $du = 0$ represents an outgoing null radial geodesic, that is the geodesic of a massless particle which runs away from the singularity at $r = 0$.

Now let us first focus our attention to the first extension (7.40) of Schwarzschild metric. In order to study the motion of particles/fotons in such a space it is useful to make a Finkelstein diagram, which is a plot of $t_* = t + (r_* - r) = v - r$ against r (see Figure 7). The light cone of an observer in an arbitrary point $P \equiv (t_*, r, \vartheta, \varphi)$ is defined by means of the radial null geodesics of $(ds = 0, d\vartheta = 0, d\varphi = 0)$, that is

$$dv=0\,,\ \text{(ingoing)},\qquad\qquad dv=\frac{2dr}{1-r_S/r}\,,\ \text{(outgoing)},$$

or, what is the same,

$$\frac{dt_*}{dr} = -1\,, \ \text{(ingoing)}, \qquad\qquad \frac{dt_*}{dr} = \frac{1 + r_S/r}{1 - r_S/r}\,, \ \text{(outgoing)}.$$

By definition of the coordinates, the ingoing null geodesics are straight lines, which in the Finkelstein diagram have angular coefficient equal to -1 at any point, while the outgoing geodesic, as well as its tangent at P, depend on the point considered and the amplitude of the light cone becomes smaller and smaller while $r \to 0$. More precisely, for the outgoing null geodesics we get

$$\lim_{r\to\infty} \frac{dt_*}{dr} = 1\,, \qquad \lim_{r\to r_S} \frac{dt_*}{dr} = +\infty\,, \qquad \lim_{r\to 0} \frac{dt_*}{dr} = -1\,.$$

Of course, very far from the singularity the light cone is the one of an observer in Minkowski space, while for $r < r_S$ the future light cone is entirely contained inside the event horizon.

Now let us consider the radial motion of massive and massless particles, then from (7.40) it follows

$$2dv\,dr = ds^2 + \left(1 - \frac{r_S}{r}\right) dv^2\,.$$

With our conventions $ds^2 \leq 0$ and time increase when $dv > 0$ (past to future).

We see that when $r > r_S$ for a massive particle $dv\,dr$ can be positive, negative or vanishing depending on initial conditions and so the particle outside the event horizon can move in any direction, while for a massless particle $dv\,dr$ can be only positive or vanishing corresponding to outgoing or ingoing null geodesics.

On the contrary, when $r < r_S$

$$dv\,dr < 0 \quad \Longrightarrow \quad dr < 0\,, \qquad\qquad (r < r_S)\,. \tag{7.42}$$

and we see that inside the event horizon the r coordinate can only decrease in time and this means that the particle/foton is destined to fall on the singularity. It can not stay at rest or move backward.

The ingoing metric (7.40), which is defined for $v \in (-\infty, \infty)$, $r \in (0, \infty)$, describes a space-time in which there is a region $r < r_S$ from which particles and light can enter but can not escape. For this reason such a solution of Einstein equations is called *black hole*.

All considerations we have done for the ingoing solution can be repeated for the outgoing solution using $t_* = t - (r_* - r) = u + r$ against r in the plot. In such a case, for $r < r_S$ one gets

$$du\,dr > 0 \quad \Longrightarrow \quad dr > 0\,, \qquad\qquad (r < r_S)\,. \tag{7.43}$$

Inside the horizon the r coordinate always increase in time and this means that particles and light will always go away from the singularity at $r = 0$. Never can remain inside the event horizon.

The outgoing metric (7.41), which is defined for $u \in (-\infty, \infty)$, $r \in (0, \infty)$, describes a space-time in which there is a region $r < r_S$ from which particles and light will escape away. For this reason such a solution of Einstein equations is called *white hole*.

The ingoing Eddington-Finkelstein coordinates can be considered as an extension of the Schwarzschild ones to the region $r < r_S$ and represent the gravitational field created by spherical symmetric collapsed body (black hole), while the outgoing Eddington-Finkelstein coordinates represent the same solution, but with future and past exchanged. Such a solution was mathematically expected because Einstein equations are invariant under time reversal but its physical realisation requires very peculiar and improbable initial conditions.

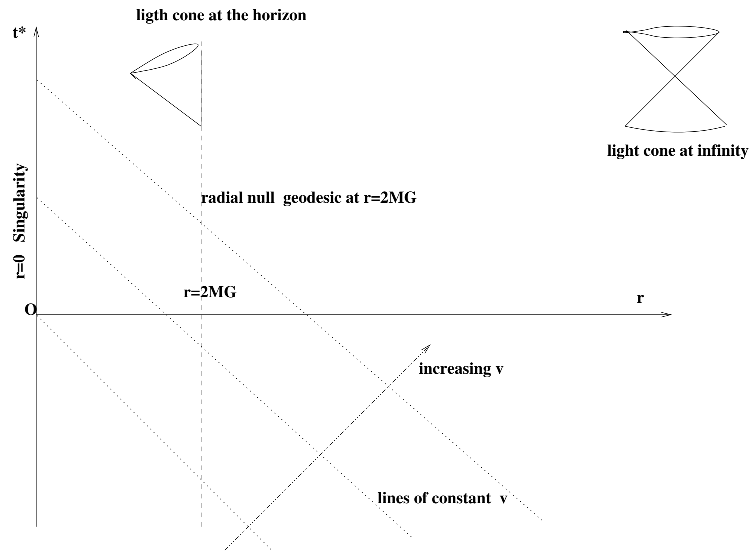

Figure 7: Eddington-Finkelstein diagram

7.8.3 The Kruskal-Szekeres extension

The maximal extension of Schwarzschild metric has been done by Kruskal and Szekeres and describes both black and white holes.

One starts again from Schwarzschild coordinates $(t, r, \vartheta, \varphi)$ defined for $r > r_S$ and make the transformation

$$X^2 - T^2 = \left(\frac{r}{r_S} - 1\right) e^{r/r_S}\,, \qquad \log\frac{X+T}{X-T} = \frac{t}{r_S}\,. \tag{7.44}$$

In these coordinates the metric read

$$ds^2 = \frac{4\,r_S^3\,e^{-r/r_S}}{r}\,(-dT^2 + dX^2) + r^2 d\sigma^2\,, \qquad r = r(T, X)\,,$$

and it is singular only at $r(T, X) = 0$. Then we can extend the coordinates to all values $r(T, X) > 0$, that is $X^2 - T^2 > -1$.

The physical singularity at $r = 0$ in such coordinates is an extended region defined by means of equation

$$X^2 - T^2 = -1 \quad \Longrightarrow \quad T = \pm\sqrt{X^2 + 1}\,.$$

This represents a hyperbola with two branches, one in the future and one in the past (with respect to $T = 0$).

The Kruskal space-time is divided in four natural regions by the *null geodesic* $X^2 - T^2 = 0$. In the Schwarzschild coordinates these correspond to the horizon $r = r_S$. (see figure 8).

The first two regions (I, II) form the *black hole* and correspond to the outer and inner regions of the event horizon. In particular, the outer region (I) is the one described by Schwarzschild coordinates.

The two regions (III, IV) have similar properties to the previous ones, but with past and future exchanged. They form the so called *white hole*.

In the diagram of Kruskal, radial null geodesics are straight lines which form an angle of 45^o with the horizontal axes. This means that an object inside the region II is destined to fall on the

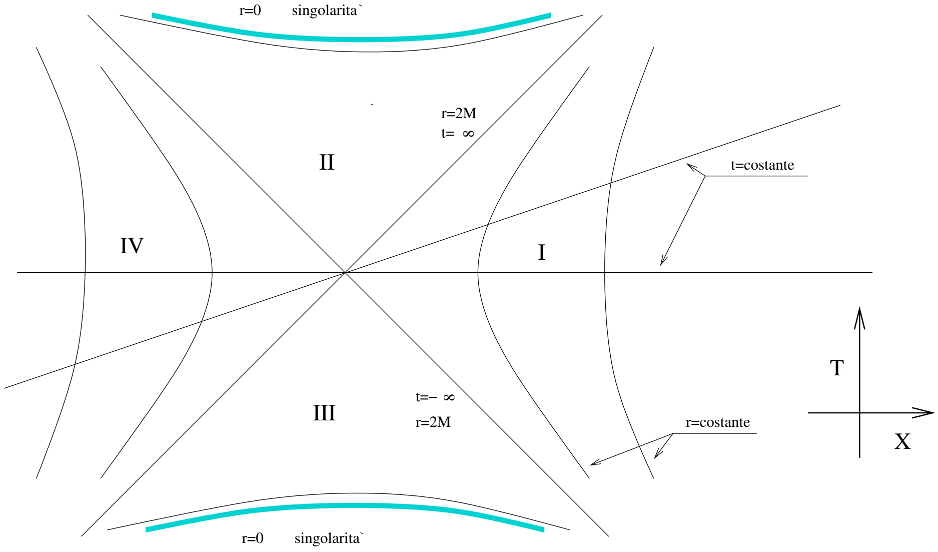

Figure 8: Kruskal extension of Schwarzschild metric

singularity $X = \sqrt{T^2 - 1}$ $(r = 0)$. In order to escape from such a region it would travel with a velocity greater than the speed of light. On the contrary, an object inside the region III will escape away, in regions I or IV.

An observer at rest in region I or IV (in the other two regions nothing can stay at rest) is represented by the hyperbola $X^2 - T^2 = const.$ It can send signals only to infinity or to region II and it can receive signals only from region III. Observers in region I can not communicate with observers in region IV and viceversa.

The Kruskal-Szekeres coordinates are an extension of Eddington-Finkelstein ones. One says that these represent the maximal extension of Schwarzschild coordinates because all geodesic can be prolongated and eventually they stop on the physical singularity.

Sometimes the metric is written in "null" coordinates

$$\begin{cases} V = T + X\,, \\ U = T - X\,, \end{cases} \qquad \Longrightarrow \qquad ds^2 = -\frac{4r_S^3}{r}\, e^{r/r_S}\, dU dV + r^2(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2)\,.$$

To finish the section we just right down the inverse transformation of (7.44), that is

$$\begin{cases} T = \sqrt{\frac{r}{r_S} - 1}\, e^{r/2r_S}\, \sinh\frac{t}{2r_S}\,, \\ X = \sqrt{\frac{r}{r_S} - 1}\, e^{r/2r_S}\, \cosh\frac{t}{2r_S}\,, \end{cases} \qquad r > r_S\,,$$

$$\begin{cases} T = \sqrt{1 - \frac{r}{r_S}}\, e^{r/2r_S}\, \cosh\frac{t}{2r_S}\,, \\ X = \sqrt{1 - \frac{r}{r_S}}\, e^{r/2r_S}\, \sinh\frac{t}{2r_S}\,, \end{cases} \qquad r < r_S\,.$$

- According to quantum mechanics, black holes are not really black. since they can emit particles by tunneling effect. The emission has a *Planckian spectrum* with a temperature

$$T_H = \frac{\hbar c}{4\pi k_B\, r_S}\,,$$

where k_B is the Boltzmann constant and $\hbar = h/2\pi$, h being the Planck constant. T_H is called the *Hawking temperature* and is the one measured by an observer at rest very far from the source.

8 Gravitational Radiation

Here we shall show that Einstein equations can have radiative solutions, which, in the weak field approximation, have properties similar to electromagnetic waves. The similarty is valid for weak fields, since in such a case only we can disregard autointeracting terms and consider linear equations. The "linearised" gravitational waves have an interpretation as a flux of particles, called *graviton*, as well as electromagnetic waves can be seen as a flux of photons. As in Section 6.6.2, we choose a "quasi-Minkowskian" reference frame where

$$g_{ij} = \eta_{ij} + h_{ij}\,, \qquad |h_{ij}| \to 0\,, \text{ (at infinity)}, \qquad g_{ij} \to \eta_{ij}\,, \text{ (at infinity)}, \tag{8.1}$$

and we develop the field equations in a power series up to first order in h_{ij}, around Minkowski metric. In this way the Einstein theory becomes equivalent to the standard theory of a tensorial field on the Minkowski space. Gravitational waves are seen as perturbations of flat space-time.

- In principle one could expand the metric around an arbitrary solution $\hat{g}_{ij}$ of Einstein equation, that is $g_{ij} = \hat{g}_{ij} + h_{ij}$, and interpret gravitational waves as perturbations of such a solution. Such a procedure is necessary for example when one studies cosmological waves which are perturbations of FLRW solution (see Section 9).

By assuming $|h_{ij}| \ll 1$ we get

$$g^{ij} \sim \eta^{ij} - h^{ij}\,, \qquad \det\{g_{ij}\} \sim (1+h)\det\{\eta_{ij}\} \quad \Longrightarrow \quad g = |\det\{g_{ij}\}| \sim |1+h|\,,$$

$$\Gamma^k_{ij} \sim \frac{1}{2}\,\eta^{kl}\left(\partial_i h_{jl} + \partial_j h_{il} - \partial_l h_{ij}\right)\,, \qquad \Gamma^k_{ik} \sim \frac{1}{2}\,\partial_i h\,, \qquad \eta^{ij}\Gamma^k_{ij} \sim \partial_j h^{jk} - \frac{1}{2}\,\eta^{jk}\partial_j h\,,$$

$$R^{(1)}_{ij} = -\frac{1}{2}\left[\Box h_{ij} + \partial_i\partial_j h - \partial_i\partial_k h^k_j - \partial_j\partial_k h^k_i\right]\,, \qquad R^{(1)} = -\Box h + \partial_i\partial_j h^{ij}\,,$$

$$G^{(1)}_{ij} = -\frac{1}{2}\left[\Box h_{ij} - \eta_{ij}\Box h + \partial_i\partial_j h - \left(\partial_i\partial_k h^k_j + \partial_j\partial_k h^k_i - \eta_{ij}\partial_l\partial_k h^{kl}\right)\right]\,,$$

where $\Box = \eta^{ij}\partial_i\partial_j$ and all indices are rised and lowered by means of Minkowski metric as in Section 6.6.2.

In order to simplify calculations now we use the freedom to choice the coordinate system. As we have seen in Section 6.5, by an infinitesimal coordinate transformation of the kind $x^k \to x^k + \xi^k$ one has

$$g_{ij} \to g_{ij} - \nabla_i\xi_j - \nabla_j\xi_i \quad \Longrightarrow \quad h_{ij} \to h_{ij} - \partial_i\xi_j - \partial_j\xi_i + O(h^2)\,,$$

where $\partial_i\xi_j$ is of the order of h_{ij} in order to deal again with a weak field.

- Note that the latter transformation for the field h_{ij} is similar to the gauge transformation which one has for the electromagnetic potential.

We can fix four arbitrary conditions on the metric by choosing a suitable vector field ξ_k. For our aim, a very convenient choice is the so called *de Donder gauge* $\Gamma^k_{ij}g^{ij} = 0$, which at first order in h_{ij} simplifies to

$$\partial_k h^k_j - \frac{1}{2}\partial_k h = \partial_k\psi^k_j = 0\,, \qquad \psi^k_j = h^k_j - \frac{1}{2}\delta^k_j h\,.$$

This is also called *harmonic gauge*, because the coordinates are harmonic functions, that is $\Box x^k = 0$. In such a gauge one has

$$R^{(1)}_{ij} = -\frac{1}{2}\Box h_{ij}\,, \qquad G^{(1)}_{ij} = -\frac{1}{2}\Box\psi_{ij}\,,$$

and the field equations (6.28) read

$$\Box\psi_{ij} = -\frac{16\pi G}{c^4}\tau_{ij} = -\frac{16\pi G}{c^4}(T_{ij} + t_{ij})\,, \qquad \partial_k\psi^k_j = 0\,. \tag{8.2}$$

At first order in h_{ij} we can dropped the grvitational energy-momentum pseudo-tensor t_{ij} since it is at least a second order function in h_{ij} an so, in vacuum ($T_{ij} = 0$) the field equations can be written as

$$R^{(1)}_{ij} \sim 0 \quad\Longrightarrow\quad \Box h_{ij} = 0\,, \qquad \partial_k\left(h^k_j - \frac{1}{2}\delta^k_j h\right) = 0\,. \tag{8.3}$$

Such equations are similar to the ones which satisfy electromagnetic waves in the Lorentz gauge. Their solutions can be obtained by using the same methods, which we briefly recall in the following section.

8.1 Electromagnetic waves

In the Lorentz or covariant gauge one has

$$\Box A_\mu = 0\,, \qquad \partial_\mu A^\mu = 0\,. \tag{8.4}$$

A real solution is given by the plane wave

$$A_\mu = e_\mu\, e^{ik_\alpha x^\alpha} + e^*_\mu\, e^{-ik_\alpha x^\alpha}\,,$$

where

$$k_\mu k^\mu = 0\,, \qquad k^\mu e_\mu = 0\,,$$

which are the wave equation and the Lorentz gauge in the momentum space.

We see that the four quantities e_μ are not independent because they have to satisfy the transversality equation $k^\mu e_\mu = 0$. The Lorentz gauge does not fix completely the field A_μ. In fact, it remains a *residual gauge* which permits to fix a second condition together transversality. Starting from a potential A_μ which satisfy the Lorentz gauge, we can build up a new potential A'_μ by means of a gauge transformation of the kind

$$A'_\mu = A_\mu + \partial_\mu f\,, \qquad \Box f = 0$$

and also this, by construction, satisfy the Lorentz gauge and of course is a solution of the wave equation. This means that effectively the Lorentz gauge selects a whole class of potentials. We can pick up a particular element of that class in the way we are going to describe. We put

$$A'_\mu = e_\mu\, e^{ik_\alpha x^\alpha} + e^*_\mu\, e^{-ik_\alpha x^\alpha} + \partial_\mu f\,, \qquad f = i\varepsilon\, e^{ik_\alpha x^\alpha} - i\varepsilon\, e^{-ik_\alpha x^\alpha}\,.$$

By construction, if ε is an arbitrary constant, then $\Box f = 0$ and so A'_μ is a solution of (8.4). It can be written as

$$A'_\mu = e'_\mu\, e^{ik_\alpha x^\alpha} + e'^*_\mu\, e^{-ik_\alpha x^\alpha}\,, \qquad e'_\mu = e_\mu - \varepsilon\, k_\mu\,.$$

It is always possible to choose ε in such a way that one component of A'_μ vanishes. The independent components then remain only 2.

To be more explicit, we suppose the wave to propagate along the x axis. Then $k^2 = k^3 = 0$, $k^0 = k^1 > 0$ e $k_0 = -k_1$. From this it follows that $e_0 = -e_1$ and with the choice $\varepsilon = e_1/k_1$ one gets $e'_0 = -e'_1 = 0$. The true physical degrees of freedom correspond to e_2 and e_3, that is A_2 and A_3, which are the components orthogonal to the direction of propagation.

The physical meaning becomes clear if we perform a rotation around the propagation axis (x). This is realised by the Lorentz matrix

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\varphi & \sin\varphi \\ 0 & 0 & -\sin\varphi & \cos\varphi \end{pmatrix} .$$

The vectors transform according to the law

$$\tilde{e}_\mu = \Lambda_\mu^\nu \, e_\nu \, ,$$

from which we get

$$\tilde{e}_0 = e_0 \, , \qquad \tilde{e}_1 = e_1 \, ,$$

$$\begin{cases} \tilde{e}_2 = \cos\varphi e_2 + \sin\varphi e_3 \, , \\ \tilde{e}_3 = -\sin\varphi e_2 + \cos\varphi e_3 \, , \end{cases} \Longrightarrow \quad \tilde{e}_\pm = e_\pm \, e^{\mp i\varphi} \, , \qquad e_\pm = e_2 \pm i e_3 \, .$$

The electromagnetic field has been decomposed in parts with *elicity* equal to ± 1 (right and left polarisation) and so we say that the photon has spin equal to 1. The non physical part has elicity 0.

8.2 Gravitational plane waves

In analogy with electromagnetism, we write the solution of (8.3) in the form

$$h_{ij} = e_{ij} \, e^{ik_r x^r} + e^*_{ij} \, e^{-ik_r x^r} \, , \qquad e_{ij} = e_{ji} \, , \tag{8.5}$$

with

$$k_r k^r = 0 \, , \qquad k^i e_{ij} - \frac{1}{2} \, k_j \, e^r_r = 0 \, .$$

e_{ij} is said *polarisation tensor*. In the following we shall see that it has only two independent components. This means that the physical degrees of freedom of the gravitational field are only two.

Four components of e_{ij} can be directly eliminated by choosing the harmonic gauge, while the other four can be eliminated by using the residual gauge. In fact, given a solution satisfying the harmonic gauge, one can do a gauge transformation of the form

$$h'_{ij} = h_{ij} - \partial_i \xi_j - \partial_j \xi_i \, , \qquad \Box \xi^k = 0 \, ,$$

obtaining in this way another solution satisfying the harmonic gauge. Putting

$$\xi^k = i\varepsilon^k \, e^{ik_r x^r} - i\varepsilon^{*k} \, e^{-ik_r x^r} \, , \qquad \Box \xi^k = 0 \, ,$$

with a constant vector ε^k, we get

$$e'_{ij} = e_{ij} + k_i \varepsilon_j + k_j \varepsilon_i \, .$$

Since ε^k is arbitrary, we can fix other four conditions.

Now we consider a wave propagating along the x axis. Then $k^a = 0$, $k^0 = k^1 > 0$ e $k_0 = -k_1$ $(a, b = 2, 3)$ and from harmonic gauge condition it follows

$$e_{0a} + e_{1a} = 0\,, \qquad -e_{00} - e_{01} = e_{01} + e_{11} = \frac{1}{2}(e_{11} + e_{22} + e_{33} - e_{00})\,.$$

Solving the algebraic system we have

$$e_{0a} = -e_{1a}\,, \qquad e_{01} = -\frac{1}{2}(e_{00} + e_{11})\,, \qquad e_{22} = -e_{33}\,. \tag{8.6}$$

Now, by the residual gauge we also obtain

$$\begin{array}{lll} e'_{ab} = e_{ab}\,, & e'_{0a} = e_{0a} - k_1\varepsilon_a\,, & e'_{1a} = e_{1a} + k_1\varepsilon_a\,, \\ e'_{00} = e_{00} - 2k_1\varepsilon_0\,, & e'_{11} = e_{11} + 2k_1\varepsilon_1\,, & e'_{01} = e_{01} + k_1(\varepsilon_1 - \varepsilon_0)\,. \end{array}$$

By taking equation (8.6) into account, we see that with a suitable choice of ε^k all components of e'_{ij} vanish, but $e'_{ab} = e_{ab}$, which remain invariant under such a transformation. Then, of the ten original components of the polarisation tensor e'_{ij} (or equivalently h'_{ij}), only two of them are linearly independent, that is

$$e'_{22} == -e'_{33} = e_{22} = -e_{33}\,, \qquad e'_{23} = e'_{32} = e_{23} = e_{32}\,,$$

$$h'_{22} = -h'_{33} = h_{22} = -h_{33}\,, \qquad h'_{23} = h'_{32} = h_{23} = h_{32}\,,$$

all the other components e'_{0k}, e'_{1k} h'_{0k} h'_{1k} being vanishing. It has to be observed that now $h' = h'^k_k = h'_{22} + h'_{33} = 0$.

As in the previous paragraph, by a rotation generated by the matrix Λ we get

$$\tilde{e}_{ij} = \Lambda_i^r \Lambda_j^s\, e_{rs} \quad \Longrightarrow \quad \begin{cases} \tilde{e}_{00} = e_{00}\,, \\ \tilde{e}_{11} = e_{11}\,, \\ \tilde{e}_{01} = e_{01}\,, \end{cases} \qquad \begin{cases} \tilde{e}_{\pm} = e^{\pm 2i\varphi}\, e_{\pm}\,, \\ \tilde{f}_{\pm} = e^{\pm i\varphi}\, f_{\pm}\,, \end{cases}$$

where we have set

$$e_{\pm} = e_{22} \mp ie_{23} = -e_{33} \mp ie_{23}\,, \qquad f_{\pm} = e_{12} \mp ie_{13} = -e_{02} \pm ie_{03}\,.$$

The components $e_{\pm}$ and $f_{\pm}$ have elicity equal to ± 2, and ± 1 respectively, while e_{00}, e_{11} and e_{01} have elicity 0. The physical components $e_{\pm}$ have elicity ± 2 and this corresponds to a particle with spin equal to 2, said *graviton*.

Transverse-Traceless gauge. — We have seen that it is always possible to choose a gauge in which the wave has only two independent components orthogonal to the direction of propagation. This gauge is called *Transverse-Traceless gauge* (TT-gauge). For a plane wave the components $h^{TT}_{ij} \equiv h'_{ij}$ satisfy the conditions

$$h^{TT}_{0k} = 0\,, \qquad \hat{n}^a h^{TT}_{ab} = 0\,, \qquad \eta^{ij} h^{TT}_{ij} = 0\,, \qquad \eta^{ij}\partial_i h^{TT}_{jk} = 0\,,$$

where $\hat{n}^a$ $(a = 1, 2, 3)$ is the unit 3-vector which points in the direction of propagation.

In general it is very difficult to realised explicitly the required gauge transformation, but to our aim it is sufficient to observe that the physical components are equal to the ones in the harmonic gauge. For example, for a plane wave propagating along the (x) axis, the non vanishing components read

$$h_{22}^{TT} = -h_{33}^{TT} = h_{22} = -h_{33}\,, \qquad h_{23}^{TT} = h_{32}^{TT} = h_{23} = h_{32}\,,$$

and satisfy the conditions

$$\eta^{ij} h_{ij}^{TT} = 0\,, \qquad \eta^{ij}\partial_i h_{jk}^{TT} = 0\,,$$

but

$$\eta^{ij} h_{ij} = h \neq 0\,, \qquad \eta^{ij}\partial_i hjk = \frac{1}{2}\,\partial_k h \neq 0\,.$$

Now we observe that

$$h_{22}^{TT} = -h_{33}^{TT} = \frac{1}{2}\left(h_{22}^{TT} - h_{33}^{TT}\right) = \frac{1}{2}\left(h_{22} - h_{33}\right) = \frac{1}{2}\left(\psi_{22} - \psi_{33}\right)\,,$$

and finally

$$h_{22}^{TT} = -h_{33}^{TT} = \frac{1}{2}\left(\psi_{22} - \psi_{33}\right)\,, \qquad h_{23}^{TT} = h_{32}^{TT} = \psi_{23}\,. \tag{8.7}$$

Just for completeness we write down the gauge transformation which relates h_{ab}^{TT} to ψ_{ab}. It reads

$$h_{ab}^{TT} = \left[(\delta_{ac} - \hat{n}_a\hat{n}_c)(\delta_{bd} - \hat{n}_b\hat{n}_d) - \frac{1}{2}\,(\delta_{ab} - \hat{n}_a\hat{n}_b)(\delta_{cd} - \hat{n}_c\hat{n}_d)\right]\psi^{cd}\,, \qquad a,b,c,d = 1,2,3$$

Choosing $\hat{n}^a \equiv (1,0,0)$ one recovers (8.7).

8.3 Example: test particles in the presence of a gravitational wave

First of all we consider a free test particle initially at rest, that is $\vec{u} = 0$ with respect to the chosen reference system. The particle moves according to geodesic equation

$$\frac{du^k}{d\tau} + \Gamma_{ij}^k u^i u^j = 0\,, \qquad u^k = \frac{dx^k}{d\tau}\,,$$

where Γ_{ij}^k is the connection related to the metric $g_{ij} = \eta_{ij} + h_{ij}$ due to the plane wave. The initial acceleration of the particle reads

$$\left.\frac{du^k}{d\tau}\right|_0 = -\Gamma_{00}^k u^0 u^0 = 0\,, \qquad \Gamma_{00}^k \sim \frac{1}{2}\,\eta^{kj}(2\partial_0 h_{0j} - \partial_j h_{00})\,,$$

and we see that it is vanishing if the TT-gauge, because $h_{0j} = h_{0j}^{TT} = 0$. This means that the particle remains at rest also in the presence of the wave, but this is simply due to the properties of the reference frame corresponding to the TT-gauge. The coordinate of the particle does not change in such a reference frame, but remember that coordinates have no direct physical meaning. In a different coordinate system $h_{0j} \neq 0$ and so the particle is accelerated.

In order to read off the presence of gravitational waves we have to compare the motion of more particles. As an example, at time t consider a first particle at the origin $P_1 \equiv (0,0,0)$ and a second particle at the point $P_2 \equiv (0,0,\varepsilon)$, ε being a small quantity. The proper distance is given by

$$\Delta\,\ell = \int_{P_1}^{P_2} |\gamma_{ab}dx^a dx^b|^{1/2} = \int_0^\varepsilon \sqrt{\gamma_{33}}\,dx^3 \sim \varepsilon\,\sqrt{g_{33}(0)}\,,$$

where γ_{ab} is the metric of spatial geometry. For a plane wave in the TT-gauge $\gamma_{ab} = g_{ab} = \eta_{ab} + h^{TT}_{ab}$ and so

$$\Delta \ell \sim \varepsilon \, [1 + h^{TT}_{33}(0)]^{1/2} \sim \varepsilon \left[1 + \frac{1}{2}\, h^{TT}_{33}\right] .$$

We see that also in the TT-gauge the proper distance between the particles in general changes as a consequence of the gravitational wave.

More information about the motion can be obtained by means of geodesic deviation equation. Let x^k and $y^k = x^k + \xi^k$ be the coordinates of two "sufficiently close" test particles in free fall in the gravitational field due to a plane wave moving along the $x^3 = z$ direction. In such a case the "small quantity" ξ^k satisfies the equation of geodesic deviation (see (5.8))

$$\frac{D^2\xi^k}{D\tau^2} = R^k_{rsj} u^r u^s \, \xi^j \, ,$$

u^k and τ being respectively the 4-velocity and the proper time of the first particle. Of course the equation is valid in any reference system, but for physical interpretation and experimental applications it is natural to choose a Lorentzian frame attached to the first particle. With this choice the particle is at rest in the point P_1 (say the origin) along the whole geodesic and so, at any time

$$u^k = (c, 0, 0, 0) \, , \qquad \Gamma^k_{ij}(P_1) = 0 \, , \qquad \frac{d\Gamma^k_{ij}}{d\tau}(P_1) = 0 \, , \qquad \tau = t \, ,$$

In such a reference frame the geodesic deviation equation becomes

$$\frac{\partial^2\xi^k}{\partial t^2} \equiv \ddot{\xi}^k = c^2 \, R^k_{00j} \xi^j \, ,$$

but since we are not in the TT-gauge, the components of the gravitational wave can be very complicated. Fortunately the linearised Riemann tensor is gauge invariant and so the right-hand side of the latter equation can be evaluated in an arbitrary gauge. At lowest order in h_{ij} then we get

$$\ddot{\xi}_k \sim \frac{1}{2}\, c^2 \partial_0^2 h^{TT}_{kj} \, \xi^j \sim \frac{1}{2}\, \ddot{h}^{TT}_{kj} \xi^j(0) \, ,$$

where on the right-hand side ξ^j has been replaced by its initial value since of course

$$\xi^k \sim \xi^k(0) + O(h) \, , \qquad \xi^k(0) = y^k(0) \equiv (0, \varepsilon \cos\phi, \varepsilon \sin\phi, 0) \, .$$

We have put the second particle in the (x, y) plane near the origin, ε being a small constant quantity and $0 \le \phi < 2\pi$. The third component of the particle does not enter the game because the wave has non vanishing components only in the (x, y) plane, in fact, for a plane wave along the z axis we have $h^{TT}_{11} = -h^{TT}_{22}$ and $h^{TT}_{12} = h^{TT}_{21}$. Then

$$\begin{aligned} \ddot{\xi}_1 &= \frac{1}{2}\, \varepsilon \left(\ddot{h}^{TT}_{11} \cos\phi + \ddot{h}^{TT}_{12} \sin\phi\right) \, , \\ \ddot{\xi}_2 &= \frac{1}{2}\, \varepsilon \left(\ddot{h}^{TT}_{12} \cos\phi - \ddot{h}^{TT}_{11} \sin\phi\right) \, . \end{aligned}$$

The latter equations have the solution

$$\begin{aligned} \xi_1 &= \varepsilon \cos\phi + \frac{1}{2}\, \varepsilon \left(A_{11} \sin\omega t \, \cos\phi + A_{12} \sin\omega t \, \sin\phi\right) \, , \\ \xi_2 &= \varepsilon \sin\phi + \frac{1}{2}\, \varepsilon \left(A_{12} \sin\omega t \, \cos\phi - A_{11} \sin\omega t \, \sin\phi\right) \, , \end{aligned}$$

Figure 9: Motion of a ring of test particles - polarized plane wave along z direction

where the plane wave has been written in the real form

$$h_{ij}^{TT} = A_{ij}\,\sin\omega(t-z)\,, \qquad A_{0k}=0 \qquad A_{3k}=0\,, \qquad A_k^k = 0\,.$$

In order to understand the physical meaning of the latter equations in it convenient to consider separately the two linear polarizations of the wave, that is $A_{11}=0$ or $A_{12}=0$. In such case we see that

$$A_{11}\neq 0\,,\;\; A_{12}=0 \quad\Longrightarrow\quad \begin{cases} \xi_1 = \varepsilon\cos\phi + \frac{1}{2}\,\varepsilon A_{11}\,\sin\omega t\,\cos\phi\,, \\ \xi_2 = \varepsilon\sin\phi - \frac{1}{2}\,\varepsilon A_{11}\sin\omega t\,\sin\phi\,, \end{cases}$$

$$A_{12}\neq 0\,,\;\; A_{11}=0 \quad\Longrightarrow\quad \begin{cases} \xi_1 = \varepsilon\cos\phi + \frac{1}{2}\,\varepsilon A_{12}\,\sin\omega t\,\sin\phi\,, \\ \xi_2 = \varepsilon\sin\phi + \frac{1}{2}\,\varepsilon A_{12}\sin\omega t\,\cos\phi\,. \end{cases}$$

In the first case ($A_{12}=0$) consider for example $\phi=0$ and $\phi=\pi$. Then we see that ξ_2 remains equal to its initial value, while $|\xi^1|$ increases in both the cases in the interval $t\in[0,\pi/2\omega]$ and decreases for $t\in[\pi/2\omega,\pi/\omega]$. On the contrary, if $\phi=\pi/2$ or $\phi=3\pi/2$ then ξ^1 remains unchanged, while $|\xi^2|$ first decreases and then increases during the above periods of time. A similar reasoning holds in the second case too ($A_{11}=0$) but with the choices $\phi=\pi/4$, $\phi=5\pi/4$ and $\phi=3\pi/4$, $\phi=7\pi/4$. This means that a ring of test particles moves as in figure 9.

8.4 Energy and momentum of a plane gravitational wave

The energy-momentum "tensor" of the gravitational field can be defined by means of (6.30), where

$$G_{ij}^{(1)} = R_{ij}^{(1)} - \frac{1}{2}\,\eta_{ij}\,R^{(1)}\,, \qquad R^{(1)} = \eta^{ij}\,R_{ij}^{(1)}\,, \qquad R^{(2)} = \eta^{ij}\,R_{ij}^{(2)}\,.$$

Up to second order in h_{ij} we get

$$G_{ij} = R_{ij}^{(1)} - \frac{1}{2}\eta_{ij}R^{(1)} + R_{ij}^{(2)} - \frac{1}{2}\left[\eta_{ij}R^{(2)} + h_{ij}R^{(1)} - \eta_{ij}h^{rs}R_{rs}^{(1)}\right] + \ldots$$

and so at lowest order one has

$$t_{ij} \sim -\frac{c^4}{8\pi G}\left\{R_{ij}^{(2)} - \frac{1}{2}\left[\eta_{ij}R^{(2)} + h_{ij}R^{(1)} - \eta_{ij}h^{rs}R_{rs}^{(1)}\right]\right\} + O(h^3)\,.$$

We are interested in the energy-momentum "transported" by a plane wave in vacuum. Then equations (8.3) hold and

$$R_{ij}^{(1)} = 0\,, \qquad R^{(1)} = 0\,, \qquad t_{ij} \sim -\frac{c^4}{8\pi G}\left[R_{ij}^{(2)} - \frac{1}{2}\eta_{ij}R^{(2)}\right]\,.$$

Of course, in such an approximation we disregard auto-interacting terms by putting $R_{ij}^{(1)}=0$.

The quantity t_{ij} is a function of h_{ij} and its derivative and could be exactly computed in a straightforward way, but from the experimental point of view, what is really important is not the energy as a function of time, but the one contained in a given finite volume with dimensions greater that the cube of the typical wave length. For this reason it is sufficient to consider the average

$\langle t_{ij}\rangle$ over a region of space-time with dimensions larger than the wave length λ. In this way, by integration, for the mean value of t_{ij} one obtains the simple expression in the wave vectors

$$\langle t_{ij}\rangle = \frac{c^4\, k_i k_j}{16\pi G}\left(e^{rs}e^*_{rs} - \frac{1}{2}|e^k_k|^2\right) . \tag{8.8}$$

The exact calculation of the latter equation can be done by using the solution (8.5), but it is quite tedious and complicated and we do not report it explicitly. However we observe that $R^{(2)}$ e $R^{(2)}_{ij}$ are real, quadratic functions in h_{ij} depending on its second derivatives and the square of its first derivative and so the average of t_{ij} has to be a real, quadratic function of polarisation e_{ij} and at the same time a quadratic function of wave vector k_i. Since $k_r k^r = 0$ and $k^i e_{ij} = k_j e^r_r/2$, it is not possible to built up a scalar quadratic expression in k_i and e_{ij}. This means that $\langle R^{(2)}\rangle = 0$. With a similar reasoning one finds that a tensor $\langle R^{(2)}_{ij}\rangle$, quadratic in k_i and e_{ij}, must have the form

$$\langle R^{(2)}_{ij}\rangle = k_i k_j(\alpha e^{rs}e^*_{rs} + \beta|e^k_k|^2) , \qquad\qquad \langle R^{(2)}\rangle = 0 ,$$

α, β being constants to be determined by explicit computation.

The energy-momentum "tensor" t_{ij} is not gauge invariant, but its averaged value $\langle t_{ij}\rangle$ is gauge invariant, as it has to do. In fact

$$e'_{ij} = e_{ij} + k_i\xi_j + k_j\xi_i \quad\Longrightarrow\quad \langle t'_{ij}\rangle = \langle t_{ij}\rangle .$$

Finally we write the expression (8.8) for a plane wave propagating along the x axis. It reads

$$\langle t_{ij}\rangle = \frac{c^4\, k_i k_j}{8\pi G}\left(|e_{22}|^2 + |e_{23}|^2\right) = \frac{c^4\, k_i k_j}{16\pi G}\left(|e_+|^2 + |e_-|^2\right) . \tag{8.9}$$

In general the expression of t_{ij} is really complicated and this is true also at lowest order in h_{ij}, but for a plane wave in the TT-gauge, where

$$\eta^{ij}\partial_i h^{TT}_{jk} = 0 , \qquad\qquad \eta^{ij}h^{TT}_{ij} = 0 ,$$

after a straightforward calculation using (6.19) and (6.20), for the Landau-Lifsits pseudo-tensor t^{LL}_{ij} one gets the simple expression

$$t^L_{ij} \sim \frac{c^4}{32\pi G}\,\partial_i h^{rs}_{TT}\partial_j h^{TT}_{rs} . \tag{8.10}$$

By definition, at lowest order the energy-momentum tensor t^L_{ij} ia a quadratic expression in $\partial_k h_{ij}$, but for a plane wave in the TT-gauge only the term in (8.10) will survive. Of course, by taking the average of (8.10) for a plane wave propagating along the x^1 axes one obtains the result in (8.9).

8.5 Emission of gravitational waves

Here we would like to compute the energy emitted as gravitational waves by an arbitrary source. To this aim we consider a weak gravitational field generated by massive bodies which move at small velocities and are confined in a small region around the origin of coordinates. In the harmonic gauge the field equations are given by (8.2), where we can drop t_{ij} because it gives second order contributions. At first order in h_{ij} then we have

$$\Box\psi^i_j \sim -\frac{16\pi G}{c^4}\,T^i_j , \qquad\qquad \partial_k\psi^k_j = 0 , \qquad\qquad \partial_k T^k_j \sim 0 .$$

The solution of the latter equation can be written down by recalling the "retarded potentials solution" which solves the analog electromagnetic equation. If fact in that case one has

$$\Box A_\mu = -\frac{4\pi}{c}\, J_\mu \,, \qquad\qquad A_\mu(t,\vec{x}) = -\frac{1}{c}\int \frac{J_\mu(\vec{y}, t-|\vec{x}-\vec{y}|/c)}{|\vec{x}-\vec{y}|}\, d^3y \,,$$

Since the equations we have to solve are formally similar to the electromagnetic ones we can directly write the solution in the form

$$\begin{aligned}
\psi_j^k(t,\vec{x}) &= -\frac{4G}{c^4}\int \frac{T_j^k(\vec{y}, t-|\vec{x}-\vec{y}|/c)}{|\vec{x}-\vec{y}|}\, d^3y \,. \\
&\sim -\frac{4G}{c^4\rho}\int T_j^k\left(\vec{y}, t-\frac{\rho}{c}+\frac{\vec{x}\cdot\vec{y}}{\rho\, c}\right)\, d^3y \\
&\sim -\frac{4G}{c^4\rho}\int T_j^k\left(\vec{y}, t-\frac{\rho}{c}\right)\, d^3y \,,
\end{aligned}$$

where the integral is extended to the compact support of T_j^k. As usual we have assumed the observer to be very far from the source, that is $|\vec{x}| \gg |\vec{y}|$ and so we have set $|\vec{x}-\vec{y}| \sim \rho$, with ρ a constant. Moreover, since we are considering non relativistic sources, in the last expression above we have also disregarded the term $(\vec{x}\cdot\vec{y}/\rho\, c)$, which is the time necessary for the wave to cross the source.

From the (approximated) conservation law $\partial_k T_j^k = 0$ we are able to get the spatial components of T_{ij} in terms of the temporal component. To this aim we first observe that

$$\partial_0 T_0^0 = -\partial_a T_0^a \,, \qquad\qquad \partial_0 T_b^0 = -\partial_a T_b^a \,, \qquad\qquad a,b,c = 1,2,3.$$

By integration one obtains

$$\partial_0 \int T_a^0 y^b\, dV = \int \partial_0 T_a^0\, y^b\, dV = -\int \partial_c T_a^c y^b\, dV = -\int \partial_c (T_a^c y^b)\, dV + \int T_b^a\, dV \,.$$

From Gauss theorem, all total divergences do not give contributions to the integral, because the integral functions have compact support, then

$$\begin{cases} \int T^{ab}\, dV = \partial_0 \int T^{0a} y^b\, dV \\ \int T^{ab}\, dV = \partial_0 \int T^{0b} y^a\, dV \end{cases} \quad\Longrightarrow\quad \int T^{ab}\, dV = \frac{1}{2}\partial_0 \int \left(T^{0a}y^b + T^{0b}y^a\right)\, dV \,.$$

In a similar way, for the temporal component we have

$$\begin{aligned}
\partial_0 \int T^{00} y^a y^b\, dV &= -\int \partial_c T^{c0} y^a y^b\, dV = \int (T^{0a}y^b + T^{0b}y^a)\, dV \,, \\
\partial_0^2 \int T^{00} y^a y^b\, dV &= \partial_0 \int (T^{0a}y^b + T^{0b}y^a)\, dV = 2\int T^{ab}\, dV \,.
\end{aligned}$$

Finally

$$\int T^{ab}\, dV = \frac{1}{2}\partial_0^2 \int T^{00} y^a y^b\, dV \sim \frac{1}{2}\partial_t^2 \int \mu y^a y^b\, dV \,,$$

where $T^{00} \sim \mu c^2$ has been set, μ being the mass density of the non relativistic source.

At large distances from the source the solution reads

$$\psi^{ab}(t,\vec{x}) \sim -\frac{2G}{\rho\, c^4}\, \partial_t^2 \int \mu(t-\rho/c, \vec{y})\, y^a y^b\, d^3y = -\frac{2G}{\rho\, c^4}\, \ddot{M}^{ab}\Big|_{t-\rho/c} \,, \tag{8.11}$$

where M^{ab} is related to the (mass) quadrupole moment Q^{ab} of the source. It has to be stressed that we have neglected from the very beginning the auto-interacting terms, by replacing the whole "tensor" τ_{ij} with the matter tensor T_{ij}. Of course in some situations this could be dangerous.

We recall that the multipole moments are defined by means of the Taylor expansion

$$f(\vec{x}) = \int \frac{\mu(t,\vec{y})\, dV}{|\vec{x}-\vec{y}|} = \frac{M}{|\vec{x}|} + \frac{D_a\, \hat{u}^{\,a}}{|\vec{x}|^2} + \frac{Q_{ab}\, \hat{u}^{\,a}\, \hat{u}^{\,b}}{2|\vec{x}|^3} + \ldots \qquad \hat{u} = \frac{\vec{x}}{|\vec{x}|}, \qquad |\vec{y}| \ll |\vec{x}|,$$

where

$$\begin{aligned} M(t) &= \textstyle\int \mu(t,\vec{y})\, dV\,, && \text{monopole (total mass),}\\ D_a(t) &= \textstyle\int \mu(t,\vec{y})\, y^a\, dV = M(t)\, \vec{x}_{cm}\,, && \text{dipole,}\\ Q^{ab}(t) &= \textstyle\int \mu(t,\vec{y})(3y^a y^b - |\vec{y}|^2\delta^{ab})\, dV = 3M^{ab} - \delta^{ab}\, M^c_c\,, && \text{quadrupole,}\\ M^{ab}(t) &= \textstyle\int \mu(t,\vec{y})\, y^a y^b\, dV\,, \end{aligned}$$

$\vec{x}_{cm}$ being the coordinates of the center of mass.

- It has to be noted that the monopole (the total mass) depends on time only if the source loses energy, for example by emission of gravitational waves, while the dipole moment is vanishing if one chooses the origin in the center of mass. In any case $\ddot{x}_{cm} = 0$.

Since we are very far from the source, from the practical point of view ψ^{ab} will be a plane wave at least in a small spacetime region. This means that by choosing (x^1) as direction of propagation, the non vanishing components of h^{TT}_{ij} will be $h^{TT}_{22} = -h^{TT}_{33} \neq 0$, $h^{TT}_{23} = h^{TT}_{32} \neq 0$.

Using for simplicity the Landau-Lifsits energy-momentum "tensor" (8.10), for the energy flux along the (x^1) axis then we obtain (recall that for plane waves $\partial_0 = -\partial_1$ and use (8.7))

$$ct^{01}_{LL} \sim \frac{c^3}{16\pi G}\left[\left(\dot{h}^{TT}_{23}\right)^2 + \left(\dot{h}^{TT}_{22}\right)^2\right] = \frac{c^3}{16\pi G}\left[\dot{\psi}^2_{23} + \left(\frac{\dot{\psi}_{22}-\dot{\psi}_{33}}{2}\right)^2\right]. \tag{8.12}$$

From (8.12) we see that the energy flux at large distances is determined by ψ^{ab}, which in the approximation here considered is related to the quadrupole moment by means of (8.11). For technical reasons, in place of M^{ab} it is convenient to use Q^{ab}, which is a traceless spatial tensor.

In terms of Q^{ab} we have

$$\psi_{23} = -\frac{2G}{3\rho\, c^4}\, \ddot{Q}_{23}\,, \qquad \psi_{22} - \psi_{33} = -\frac{2G}{3\rho\, c^4}\, (\ddot{Q}_{22} - \ddot{Q}_{33})\,,$$

from which it follows

$$ct^{01} = \frac{G}{36\pi\rho^2 c^5}\left[\dddot{Q}^2_{23} + \left(\frac{\dddot{Q}_{22} - \dddot{Q}_{33}}{2}\right)^2\right].$$

The latter equation represents the energy density which pass through a unitary surface during the unitary time due to a plane wave. The energy which pass through an infinitesimal surface $dS = \rho^2 d\Omega$ is then

$$ct^{01} dS = \frac{G}{36\pi c^5}\left[\dddot{Q}^2_{23} + \left(\frac{\dddot{Q}_{22} - \dddot{Q}_{33}}{2}\right)^2\right] d\Omega\,, \tag{8.13}$$

$d\Omega$ being the solid angle. This represents the contribution due to a wave propagating along (x^1). In order to take into account of waves propagating in all possible directions, we write (8.13) in a form invariant under spatial rotations, using the (spatial) polarisation tensor e_{ab}, with the properties

$$e_a^a = 0\,, \qquad e_{ab}\hat{n}^b = 0\,, \qquad e_{ab}\, e^{ab} = 1\,, \qquad a,b = 1,2,3\,, \qquad \hat{n}_a\hat{n}^a = 1\,, \tag{8.14}$$

where $\hat{n}^a$ is a unitary vector along the propagation direction and e_{ab} is a unitary three-dimensional tensor (under spatial rotations). By means of e_{ab} we have

$$dW = \frac{G}{72\pi c^5}\left(\dddot{Q}^{ab}\; e_{ab}\right)^2 d\Omega\,, \tag{8.15}$$

which is the contribution to the flux of an arbitrary polarised plane wave.

For example, the two matrices

$$e_{ab} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad e_{ab} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

in (8.15) gives respectively the two contributions in (8.13).

In order to take into account of all polarisations, one makes an "average" of the kind

$$d\bar{W} = 2\,\frac{G}{72\pi c^5}\left(\dddot{Q}^{ab}\,\dddot{Q}^{cd}\;\overline{e_{ab}e_{cd}}\right) d\Omega\,, \tag{8.16}$$

where the factor 2 arises from the two independent polarisation of the wave. The mean value of the product of all polarisations is a constant spatial tensor of order four which depends on the directions. This means that it can be built up by using the vector $\vec{n}$, which takes into account of the direction and the constant tensor δ_{ab}. Note that the other constant tensor e^{abc} (Levi-Civita) can not be used because it has negative parity (it is a pseudotensor). It is easy to verify that the more general constant tensor which satisfy the properties in (8.14) has to be the following:

$$\begin{aligned}\overline{e_{ab}e_{cd}} \;=\; & \frac{1}{4}\left[\hat{n}_a\hat{n}_b\hat{n}_c\hat{n}_d + (\hat{n}_a\hat{n}_b\delta_{cd} + \hat{n}_c\hat{n}_d\delta_{ab})\right. \\ & \left. -(\hat{n}_a\hat{n}_c\delta_{bd} + \hat{n}_a\hat{n}_d\delta_{bc} + \hat{n}_b\hat{n}_c\delta_{ad} + \hat{n}_b\hat{n}_d\delta_{ac}) - \delta_{ab}\delta_{cd} + \delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc}\right]\,.\end{aligned}$$

Using such an expression in (8.16) we get

$$d\bar{W} = \frac{G}{36\pi c^5}\left[\frac{1}{4}\left(\dddot{Q}^{ab}\;\hat{n}_a\hat{n}_b\right)^2 + \frac{1}{2}\left(\dddot{Q}^{ab}\;\dddot{Q}_{ab}\right) - \dddot{Q}^{ac}\;\dddot{Q}_c^{\;b}\;\hat{n}_a\hat{n}_b\right] d\Omega\,.$$

Now, the total emitted energy from he source is obtained by taking the average over all directions, given by the vector $\vec{n}$, and by integrating over the angles. We have at disposal only the tensor δ_{ab} and so

$$\overline{\hat{n}_a\hat{n}_b} = \frac{1}{3}\,\delta_{ab}\,, \qquad \overline{\hat{n}_a\hat{n}_b\hat{n}_c\hat{n}_d} = \frac{1}{15}\left(\delta_{ab}\delta_{cd} + \delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc}\right).$$

After this last average the flux does not depend on the directions and so the integration over the angles gives just a factor 4π, The finally expression read

$$W_{tot} = -\frac{dE}{dt} = \frac{G}{45c^5}\;\dddot{Q}^{ab}\,\dddot{Q}_{ab}\;\;. \tag{8.17}$$

This is called the *quadrupole formula* and represents the total gravitational energy emitted from the gravitational source in the unit time (the total power W_{tot}). In general it is a very small quantity due to the presence of the factor $G/45c^5 \sim 6 \cdot 10^{-55}\, sec^3/Kg\, m^2$, but it could be measured in the presence of high quadrupole moments with high variations (supernova explosion).

The quadrupole formula has been indirectly verified by R.A. Hulse, J.H. Taylor[37] and collaborators (1974) by the observation (for many years) of the binary system PSR 1916+13. One of the two stars is a pulsar (neutron star), which rotates around the partner and emits radio signals, which permit to determine the orbit with high precision. The radius of the orbit decreases systematically in time for the emission of gravitational energy in agreement with the quadrupole formula above.

8.6 Examples

We consider some simple gravitational sources and compute their power by means of (8.17).

8.6.1 Rotating bar

Just to have an idea of the intensity of the gravitational waves with respect to the sources, we first consider a bar rotating around a fixed axis with a constant angular velocity ω. Let us M and L be respectively the mass and the length of the bar. If the transverse dimensions are negligible with respect to the length, then the maximum quadrupole moment is obtained when the rotational axis is orthogonal to the axis of the bar.

Without doing the explicit computation, from dimensional considerations we obtain

$$Q^{ab}Q_{ab} \sim \left[ML^2\right]^2 , \qquad \dddot{Q}^{ab}\, \dddot{Q}_{ab} \sim \left[ML^2\omega^3\right]^2 ,$$

and from (8.17)

$$W_{tot} \sim \frac{G}{c^5}\, M^2L^4\omega^3 \sim \left[10^{-53}W\right] \left[\frac{M}{Kg}\right]^2 \left[\frac{L}{m}\right]^4 \left[\frac{\omega}{sec^{-1}}\right]^6 ,$$

where $1\,W = 1\,Kg\,m^2/sec^3$. It is clear then it is practically impossible to generate gravitational waves in a laboratory. Then one has to look for astronomic or cosmological sources.

8.6.2 Generic bounded gravitational system

We recall that the virial theorem for the Newtonian potential states that

$$\langle K \rangle = -\frac{1}{2}\, \langle U \rangle ,$$

$\langle K \rangle$ and $\langle U \rangle$ being the averaged (in time) kinetic and potential energies of a bounded gravitational system. For such kind of systems we can define a typical angular frequency ω by means of

$$\omega^2 \sim \frac{MG}{L^3} ,$$

M being the total mass of the system and L the spatial size.

[37] Nobel price in Physics (1993).

For example, if we consider the system of two particles with mass m rotating on a circular orbit of radius R_0 with angular velocity ω we have exactly

$$\begin{cases} K = \frac{1}{2}\, I\omega^2 = mR_0^2\omega^2\,, \\ U = -\frac{m^2 G}{2R_0}\,, \end{cases} \quad \Longrightarrow \quad \omega^2 = \frac{mG}{4R_0^3} = \frac{MG}{L^3}\,, \qquad M = 2m\,, \qquad L = 2R_0\,.$$

By dimensional considerations now we get

$$W_{tot} \sim \frac{c^5}{G}\left(\frac{MG}{c^2 L}\right)^5 = \frac{c^5}{G}\left(\frac{r_S}{2L}\right)^5 .$$

The Schwarzschild radius r_S is always smaller than the size of the system and so one obtains a constraint for the power radiated by a bounded gravitational system, that is

$$W_{tot} < \frac{c^5}{G} < 10^{53}\, W\,.$$

8.6.3 A simple binary system

As a simple binary system we consider two bodies with masses m_1, m_2, total mass $M = m_1 + m_2$, moving in the (x, y) plane in a circular orbit under the reciprocal gravitational force. We indicate by ω, and by $\mu = m_1 m_2/M$ the angular velocity and the reduced mass respectively and choose the origin in the center of mass and, as usual, we put $\vec{r} = \vec{r}_1 - \vec{r}_2$. Since the orbit is circular, $r = |\vec{r}|$ is a constant and so

$$\dot{\phi} = \sqrt{\frac{MG}{r^3}} = \omega\,, \qquad \phi(t) = \omega t\,. \tag{8.18}$$

The non vanishing components of tensors M^{ab} and Q_{ab} are

$$\begin{cases} M_{xx} = \mu\, r^2\, \cos^2\omega t = \frac{\mu\, r^2}{2}\,(1 + \cos 2\omega t)\,, \\ M_{yy} = \mu\, r^2\, \sin^2\omega t = \frac{\mu\, r^2}{2}\,(1 - \cos 2\omega t)\,, \\ M_{xy} = \mu\, r^2\, \cos\omega t \sin\omega t = \frac{\mu\, r^2}{2}\, \sin 2\omega t\,, \end{cases} \quad \Longrightarrow \quad \begin{cases} Q_{xx} = \frac{\mu\, r^2}{2}\,(1 + 3\, \cos 2\omega t)\,, \\ Q_{yy} = \frac{\mu\, r^2}{2}\,(1 - 3\, \cos 2\omega t)\,, \\ Q_{zz} = -\mu\, r^2\,, \\ Q_{xy} = \frac{3}{2}\,\mu\, r^2\, \sin 2\omega t\,, \\ Q_{xz} = Q_{yz} = 0\,. \end{cases}$$

From (8.17) it follows

$$W_{tot} = -\frac{dE}{dt} = \frac{32G}{15c^5}\,\mu^2 r^4 \omega^6\,. \tag{8.19}$$

This is the total energy emitted by the system during a unit time. The energy emitted in a period of revolution $T = 2\pi/\omega$ is then

$$E_T = \int_0^T dW_{tot}\, dt = T\, W_{tot} = \frac{64\pi G}{15c^5}\,\mu^2 r^4 \omega^5\,.$$

In the calculation we have assumed the distance r between the two bodies to be a constant and this is a reasonable assumption if E_T is negligible with respect to the absolute value of the total energy E of the system, that is

$$E = -\frac{Gm_1 m_2}{2r}\,, \qquad E_T \ll \frac{Gm_1 m_2}{2r}\,. \tag{8.20}$$

With this assumption we get

$$\frac{dE}{dt} = \frac{Gm_1m_2}{2r^2}\,\frac{dr}{dt} = -W_{tot}\,,$$

from which it follows

$$\frac{dr}{dt} = -\frac{2r^2}{Gm_1m_2}\,W_{tot} = -\frac{64\mu}{15Mc^5}\,r^6\omega^6 = -\frac{K}{r^3}\,, \qquad\qquad K = \frac{64\mu M^2G^3}{15c^5}\,.$$

Here we have replaced ω by means of (8.18). By integrating this latter equation we can compute the total life of a system, with initial dimension $r = r_0$. We get

$$T_{tot} = -\frac{1}{K}\int_{r_0}^{0} r^3\,dr = \frac{r_0^4}{4K}.$$

The latter result has to be considered a rough approximation of the real one, because condition (8.20) does not hold for all time.

Using equation (8.19) for the planets in their motion around the sun we get, for example for earth and jupiter

$$\begin{array}{ll} W_{tot}^{earth} \sim 66\,W\,, & T_{tot}^{earth} \sim 10^{23}\,years\,, \\ W_{tot}^{jupiter} \sim 1700\,W\,, & T_{tot}^{jupiter} \sim 10^{24}\,years\,. \end{array}$$

It is also interesting to analyse the polarisation of the wave emitted by the system. To this aim we recall that

$$\psi^{ab} = -\frac{2G}{\rho\,c^4}\,\ddot{M}^{ab}\,,$$

and the TT-gauge components can be obtained by means of (8.7).

We first consider the waves emitted along the (z) axis (normal to the orbital plane). The non vanishing components of the wave in TT-gauge are given by

$$\begin{cases} h_{xx}^{TT} = -h_{yy}^{TT} = \frac{1}{2}\left(\psi_{xx} - \psi_{yy}\right) = \frac{4Gr^2\omega^2}{\rho}\,\cos 2\omega t\,, \\ h_{xy}^{TT} = h_{yx}^{TT} = \psi_{xy} = \frac{4Gr^2\omega^2}{\rho}\,\sin 2\omega t\,, \end{cases} \quad\Longrightarrow\quad \text{(circular polarisation).}$$

It is clear that due to axial symmetry all the waves with direction in the plane of the orbit have the same properties. Then it is sufficient to consider a wave directed along the (x) axis. We get

$$\begin{cases} h_{yy}^{TT} = -h_{zz}^{TT} = \frac{1}{2}\left(\psi_{yy} - \psi_{zz}\right) = \frac{2Gr^2\omega^2}{\rho}\,\cos 2\omega t\,, \\ h_{yz}^{TT} = h_{zy}^{TT} = \psi_{yz} = 0\,, \end{cases} \quad\Longrightarrow\quad \text{(linear polarisation)}\,.$$

We see that the waves in the direction normal to the plane of the orbit have circular polarisation and twice the amplitude of the waves with direction in the orbital plane. The ratio between the corresponding contribution to W_{tot} is then $1/8$.

8.6.4 The Hulse-Taylor binary system

Now we study the emission of gravitational radiation of the system classified as $PSRB1913+16$, which is constituted by two stars rotating on an elliptic orbit. All parameters of the system are known with high precision and it has been observed that the dimension of the trajectory, as well as the period of revolution, decrease as a function of time. This can be a consequence of the energy loss due to the emission of gravitational waves.

As in the previous example we choose (x, y) as the orbital plane and we indicate by μ, M the reduced and the total mass respectively and by $\vec{r} = (r\cos\phi, r\sin\phi, 0)$, the relative position. In this way, the motion of a star with respect to the other is described by the equations

$$\frac{a(1-e^2)}{r} = 1 + e\cos\phi\,, \qquad \dot{\phi} = \frac{\sqrt{MGa(1-e^2)}}{r^2}\,,$$

where a is the semi-major axis and e the eccentricity. Energy, momentum and revolution period are "nearly conserved" quantities given by

$$E = -\frac{Gm_1m_2}{2a}\,, \qquad L = Gm_1m_2\mu\, a(1-e^2)\,, \qquad T = \frac{2\pi a^{3/2}}{\sqrt{GM}}\,.$$

The quadrupole moments are similar to the ones computed in the previous example, that is

$$\begin{cases} Q_{xx} = \frac{\mu r^2}{2}\,(1 + 3\,\cos 2\omega t)\,, \\ Q_{yy} = \frac{\mu r^2}{2}\,(1 - 3\,\cos 2\omega t)\,, \\ Q_{zz} = -\mu\, r^2\,, \\ Q_{xy} = \frac{3}{2}\,\mu\, r^2\,\sin 2\omega t\,, \\ Q_{xz} = Q_{yz} = 0\,. \end{cases}$$

However now the time dependence of ϕ and r are non trivial and for this reason the computation of the third derivatives of Q_{ab} is quite long and tedious. The final result reads

$$W = -\frac{dE}{dt} = \frac{8G^4Mm_1^2m_2^2}{15c^5a^5(1-e^2)^5}\,(1+e\cos\phi)^4[12(1+e\cos\phi)^2 + e^2\sin^2\phi]\,.$$

The average power irradiated on a period T is given by

$$\langle W\rangle = \frac{1}{T}\int_0^T W\,dt = \frac{1}{T}\int_0^{2\pi}\frac{W\,d\phi}{\dot{\phi}}\,, \tag{8.21}$$

from which the following *Peters-Mathews* result follows:

$$\langle\frac{dE}{dt}\rangle = -\frac{32G^4Mm_1^2m_2^2}{5c^5a^5}\,f(e)\,, \qquad f(e) = (1-e^2)^{-7/2}\left(1 + \frac{73}{24}\,e^2 + \frac{37}{96}\,e^4\right)\,.$$

We see that $f(e)$ is a rapidly increasing function of e which amplifies the effect.

For the period one gets

$$\frac{dT}{dt} = 3\pi\sqrt{\frac{a}{MG}}\,\frac{da}{dt} = \frac{6\pi a^{5/2}}{\mu(MG)^{3/2}}\,\frac{dE}{dt}\,,$$

$$\dot{T} = \langle\frac{dT}{dt}\rangle = \frac{6\pi a^{5/2}}{\mu(MG)^{3/2}}\,\langle\frac{dE}{dt}\rangle = -\frac{192\pi\, m_1m_2\, G^{5/3}}{5c^2\, M^{1/3}}\,\left(\frac{T}{2\pi}\right)^{-5/3}\,f(e)\,.$$

For the system under consideration $e \sim 0.617$ and $f(e) \sim 10$. The other parameters are

$$m_p \equiv m_1 \sim 1.44\mathcal{M}_\odot\,, \qquad m_c \equiv m_2 \sim 1.38\mathcal{M}_\odot\,, \qquad a \sim 1950100\,Km\,.$$

With these data one gets $\dot{T} \sim -2.4024 \times 10^{-12}$. The observed value is $\dot{T}_{obs} \sim -2.4184 \times 10^{-12}$. To this latter value one has to add a small contribution due to the acceleration of the binary system with respect to solar system as a consequence of the rotation of the galaxy. Taking such a correction into account one has $T_{cor} =\sim -2.4056 \times 10^{-12}$ and finally

$$\frac{\dot{T}}{\dot{T}_{cor}} \sim 1.0013\,.$$

9 Cosmological Solutions

These are solutions of Einstein equations, which are able to describe the whole universe. They are based on few, very general axioms, which principally derive from astronomical observations and from the assumption that we are not in a privileged position. Then it is natural o assume the universe to be *spatially homogeneous.*

From our point of view (on a large scale), the universe appears also *spatially isotropic* since all directions are "essentially equivalent" and, because there are not privileged points, it has to be isotropic with respect to any point. The isotropic property which we observe now is assumed to be valid at any time (a non-isotropic model always evolves in an isotropic one). It is clear that homogeneity and isotropy are not local properties, but properties which are valid for a "smeared out" universe averaged on a "sufficiently large scale" of the order $10^8 \sim 10^9$ light-year (this includes many clusters of galaxies).

In the following we shall use units in which $c = 1$ and we shall indicate with ρ and p respectively the density of matter/energy and the pressure of the whole universe.

Cosmological principle: during all its evolution the universe is *spatially homogeneous and isotropic* or, what is the same, it is *isotropic around any point.*

From the mathematical point of view this means that it is possible to choose an *universal time* t in such a way that, at any t, the spatial metric is the same in any point and in any direction. That is, the spatial $3-$*dimensional* space has 6 independent Killing vectors (3 translations and 3 rotations) and so it is a *maximally symmetric manifold* (see 6.5). The metric can be written in the general form

$$ds^2 = -dt^2 + a^2(t)\, d\Sigma^2\,, \qquad d\Sigma^2 = \hat{g}_{ab} dx^a dx^b\,, \qquad a,b,c = 1,2,3,$$

where $d\Sigma^2$ is the metric of a maximally symmetric space and $a(t)$ is a function of time only, which will be determined by solving Einstein equations.

It is clear that this is the more general form for a metric satisfying the cosmological principle. If fact, a dependence of the function a from spatial coordinates trivially breaks homogeneity, but also a dependence of g_{00} from spatial coordinates breaks homogeneity, because in such a case proper time depends on the chosen point and as consequence also $a(t)$, as a function of proper time, will depend on the point. A possible presence of $g_{00}(t)$ can be drop by a redefinition of time.

9.1 Maximally symmetric spaces

These are manifolds with the maximum number of possible Killing vectors and are isotropic and homogeneous, Viceversa, any isotropic and homogeneous space is maximally symmetric. Such kind of manifolds must have constant scalar curvature and so

$$\hat{R}_{abcd} = K(\hat{g}_{ac}\hat{g}_{bd} - \hat{g}_{ad}\hat{g}_{bc})\,, \qquad \hat{R}_{ab} = (n-1)\, K\, \gamma_{ab}\,, \qquad \hat{R} = n(n-1)\, K\,, \tag{9.1}$$

where n is the dimension (3 in our case) and K the Gauss curvature. k can be positive, negative or vanishing. One respectively has the hypersphere S^n, the hyperbolic space H^n and the Euclidean space $\mathbb{R}^n$. The $n-$*dimensional* spaces with constant curvature can be embedded in an Euclidean manifold with $n+1-$*dimensions*. In such a case the Gauss curvature $K = 1/b^2$ becomes the inverse of the square of the "radius" (in the case of S^n, b is really the radius of the ball in $\mathbb{R}^{n+1}$, while in the hyperbolic case b is imaginary).

In the case of General Relativity $n = 3$ and an embedded hypersurface in $\mathbb{R}^4$ is described by the equation

$$x^2 + y^2 + z^2 + u^2 = b^2 \,, \qquad d\Omega^2 = dx^2 + dy^2 + dz^2 + du^2 \,, \tag{9.2}$$

$d\Omega^2$ being the metric of $\mathbb{R}^4$ in Cartesian coordinates. By taking into account of the constraint, we can derive u in terms of the other coordinates and in this way $d\Omega^2$ becomes the metric of the hypersurface. To this aim it is convenient to use hyper-cylindrical coordinates in $\mathbb{R}^4$, that is spherical coordinates in the sub-space $\{x, y, z\}$. We put

$$\begin{array}{rcll} x &=& r\sin\vartheta\cos\varphi \,, & r^2 = x^2 + y^2 + z^2 \,, \\ y &=& r\sin\vartheta\sin\varphi \,, & u^2 = b^2 - r^2 \,, \\ z &=& r\cos\vartheta \,, & du = -\dfrac{rdr}{\sqrt{b^2 - r^2}} \,, \end{array}$$

and so

$$d\Omega^2 \quad \Longrightarrow \quad d\Sigma^2 = dr^2 + r^2(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2) + du^2 = \frac{dr^2}{1 - r^2/b^2} + r^2(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2) \,.$$

Now the whole metric assumes the *FLRW* form (Friedmann-1922, Lemaître-1927, Robertson-Walker-1930)

$$ds^2 = -dt^2 + a^2(t) \left[\frac{dr^2}{1 - k\, r^2} + r^2(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2) \right] \,, \qquad k = \frac{K}{|K|} = 1, 0, -1 \,, \tag{9.3}$$

which is *spatially homogeneous* and *spatially isotropic*, but in general it is not *stationary* for the presence of the factor $a(t)$. In order to get (9.3) we have performed the transformation
$r^2 \to r^2/|K|$, $a^2(t) \to a^2(t)|K|$. In such a case $a(t)$ becomes the "radius" of the spatial part of the universe at cosmological time t.

Reference and coordinates in which the universe appears homogeneous and isotropic are called respectively *co-moving reference frame* and *co-moving coordinate system*. In some physical applications it could be convenient to replace t, r by the *conformal co-moving coordinates* η, χ given by

$$\begin{cases} \eta(t) = \int \frac{dt}{a(t)} & \Longrightarrow \quad dt = a(\eta)\, d\eta \,, \\ d\chi^2 = \frac{dr^2}{1-kr^2} & \Longrightarrow \quad r = \phi_k(\chi) \,, \qquad \chi = f_k(r) \,, \end{cases} \tag{9.4}$$

where by definition $a(\eta) \equiv a(t(\eta))$, while

$$r = \phi_k(\chi) \equiv \begin{cases} \sin\chi \,, & k = 1 \,, & (0 \le \chi < \pi) \,, & (0 \le r < 1) \,, \\ \chi \,, & k = 0 \,, & (0 \le \chi < \infty) \,, & (0 \le r < \infty) \\ \sinh\chi \,, & k = -1 \,, & (0 \le \chi < \infty) \,, & (0 \le r < \infty) \,, \end{cases} \tag{9.5}$$

$$\chi = f_k(r) \equiv \begin{cases} \arcsin r \,, & k = 1 \,, & (0 \le r < 1) \,, & (0 \le \chi < \pi) \,, \\ r \,, & k = 0 \,, & (0 \le r < \infty) \,, & (0 \le \chi < \infty) \,. \\ \operatorname{arcsinh} r \,, & k = -1 \,, & (0 \le r < \infty) \,, & (0 \le \chi < \infty) \,. \end{cases} \tag{9.6}$$

In terms of conformal coordinates the metric reads

$$ds^2 = a^2(\eta) \left[-d\eta^2 + d\chi^2 + \phi_k^2(\chi) \left(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2 \right) \right] \,. \tag{9.7}$$

In the special case $k = 0$, the metric in conformal coordinates is *conformally flat*. This means that it is proportional to the Minkowski metric.

9.2 Friedmann-Lemaître-Robertson-Walker universe

In previous section we have determined the form of the more general homogeneous and isotropic metric. It depends only on an arbitrary fuction of time $a(t)$, which has to be determined in such a way that metric in (9.3) is a solution of Einstein equations with a spatially homogeneous and isotropic energy-momentum tensor (a perfect fluid).

The non vanishing components of the metric read

$$g_{00} = -1\,, \qquad g_{ab} = a^2(t)\hat{g}_{ab}\,, \qquad g^{00} = -1\,, \qquad g^{ab} = \frac{\hat{g}^{ab}}{a^2(t)}\,, \qquad a,b,... = 1,2,3$$

where $\hat{g}_{ab} = \gamma_{ab}$ is the metric of the 3-dimensional hypersurface $(S^3, \mathbb{R}^3, H^3)$, which has constant curvature, Riemann and Ricci tensors given (9.1).

The non vanishing components of Riemannian quantities can be easily computed and read

$$\Gamma^0_{ab} = a\dot{a}\hat{g}_{ab} = \frac{\dot{a}}{a}g_{ab}\,, \qquad \Gamma^a_{0b} = \Gamma^a_{b0} = \frac{\dot{a}}{a}\delta^a_b\,, \qquad \Gamma^c_{ab} = \hat{\Gamma}^c_{ab}\,,$$

$$\begin{aligned}
R^0_{a0b} &= \partial_0\Gamma^0_{ab} - \Gamma^0_{bc}\Gamma^c_{0a} = \frac{\ddot{a}}{a}g_{ab}\,,\\
R^0_{abc} &= a\dot{a}\left(\partial_b\hat{g}_{ac} - \partial_c\hat{g}_{ab} + \Gamma^d_{ac}\hat{g}_{bd} - \Gamma^d_{ab}\hat{g}_{cd}\right) = \hat{\nabla}_b\hat{g}_{ac} - \hat{\nabla}_c\hat{g}_{ab} = 0\,,\\
R^c_{dab} &= \hat{R}^c_{dab} + \frac{\dot{a}^2}{a^2}\left(\delta^c_a g_{bd} - \delta^c_b g_{ad}\right)\,,
\end{aligned}$$

$$\begin{aligned}
R_{00} &= -\frac{3\ddot{a}}{a}\,,\\
R_{ab} &= \frac{\ddot{a}}{a}\,g_{ab} + \hat{R}_{ab} + \frac{2\dot{a}^2}{a^2}\,g_{ab} = \frac{a\ddot{a} + 2\left(k + \dot{a}^2\right)}{a^2}\,g_{ab}\,,\\
R &= \frac{6\left(k + \dot{a}^2 + a\ddot{a}\right)}{a^2}\,.
\end{aligned} \tag{9.8}$$

Finally we also get

$$\begin{aligned}
G_{00} &= 3\left(\frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right)\,,\\
G_{ab} &= -\left(\frac{2\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{k}{a^2}\right)g_{ab}\,.
\end{aligned} \tag{9.9}$$

The general form of the energy-momentum tensor with the required symmetry is the one of a perfect fluid with pressure p and mass/energy density ρ depending on universal time only, that is

$$T^{ij} = pg^{ij} + (p+\rho)u^iu^j\,, \qquad T = T^k_k = 3p - \rho\,, \qquad p = p(t)\,, \qquad \rho = \rho(t)\,.$$

In co-moving reference frame matter is "at rest". This means that

$$u^k \equiv (1,0,0,0)\,, \qquad u^ku_k = -1\,, \qquad u_k \equiv (-1,0,0,0)\,,$$

and so

$$T_{00} = \rho\,, \qquad T_{00} - \frac{1}{2}\,g_{00}T = \frac{\rho + 3p}{2}\,, \qquad T_{ab} = p\,g_{ab}$$

and from Einstein equations (6.3-6.6) it follows

$$\begin{cases} G_{00} = 8\pi G T_{00}\,, \\ R_{00} = 8\pi G\left(T_{00} - \frac{1}{2}\,g_{00}\,T\right)\,, \\ G_{ab} = 8\pi G T_{ab}\,,\,, \end{cases} \quad \Longrightarrow \quad \begin{pmatrix} \left(\frac{\dot a}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G}{3}\,\rho\,, \\ \frac{\ddot a}{a} = -\frac{4\pi G}{3}\,(\rho + 3p)\,, \\ \frac{2\ddot a}{a} + \left(\frac{\dot a}{a}\right)^2 + \frac{k}{a^2} = -8\pi G\,p\,, \end{pmatrix} \tag{9.10}$$

Note that only two of the three equations above are linearly independent. Another independent equation can be obtained by using the conservation law. In particular we have

$$\frac{d}{dt}\,(a^3\rho) = -3pa^2\dot a\,, \tag{9.11}$$

which follows from the equation

$$\begin{aligned} 0 = \nabla_k T^{k0} &= \partial_k T^{k0} + \Gamma^k_{kl} T^{l0} + \Gamma^0_{kl} T^{kl} = \partial_t T^{00} + \Gamma^k_{k0} T^{00} + \Gamma^0_{kl} T^{kl} \\ &= \dot\rho + \frac{3\dot a}{a}\,\rho + \frac{3\dot a}{a}\,p = \frac{1}{a^3}\left[\frac{d}{dt}\,(a^3\rho) + 3pa^2\dot a\right]. \end{aligned}$$

The other three equations are trivially satisfied because

$$\begin{aligned} \nabla_k T^{kb} &= \partial_k T^{kb} + \Gamma^k_{kl} T^{lb} + \Gamma^b_{kl} T^{kl} = \partial_a T^{ab} + \Gamma^c_{ca} T^{ab} + \Gamma^b_{ac} T^{ac} \\ &= \frac{p}{a^2}\left[\partial_a \hat g^{ba} + \hat\Gamma^c_{ca}\hat g^{ab} + \hat\Gamma^b_{ac}\hat g^{ac}\right] = \frac{p}{a^2}\,\hat\nabla_a \hat g^{ac} = 0\,. \end{aligned}$$

Assuming $\rho(t)$ to depend on t by means of $a(t)$, that is $\rho(t) = \rho(a(t))$, equation (9.11) assumes the form

$$\frac{d}{da}\,(a^3\rho) = -3pa^2\,.$$

In order to completely solve the problem one has to postulate also an equation of state for matter, that is a relation $p = p(\rho)$ between pressure and energy, which determines the type of fluid we are dealing with. In summary we have the equations

$$\begin{cases} \frac{2\ddot a}{a} + \left(\frac{\dot a}{a}\right)^2 + \frac{k}{a^2} = -8\pi G p\,, & \text{field equation,} \\ \frac{\ddot a}{a} = -\frac{4\pi G}{3}\,(\rho + 3p)\,, & \text{field equation,} \\ \\ \left(\frac{\dot a}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G}{3}\,\rho\,, & \text{Friedmann equation,} \\ \frac{d}{da}\,(a^3\rho) = -3pa^2\,, & \text{conservation law,} \\ p = p(\rho)\,, & \text{equation of state,} \end{cases} \tag{9.12}$$

where k can assume the values $0, \pm 1$. In has to be stressed that the equations above are not linearly independent and in fact the last three are sufficient for a complete integration.

9.3 Einstein universe

Just for historical reasons now we briefly describe the Einstein cosmological model (1917), which was the first one based on General Relativity. This is a *static*, *spatially homogeneous* and *isotropic* model, which was immediately rejected after the discovery of the cosmological red-shift (see the next Chapter), which is not present in a static universe.

If one does not take into account of a cosmological constant, the *FLRW* metric is the more general one which gives rise to a spatially homogeneous and isotropic universe, but, as it follows from second and third equations in (9.12), for ordinary matter ($\rho > 0$ and $p \geq 0$), the solution is not static. This means that in order to have a static solution we have to consider *exotic matter* with negative pressure or alternatively we have to modify Friedmann equations by taking into account of a cosmological constant Λ. As we said in chapter (6), the cosmological constant Λ gives contributions to the field equations which can be trivially obtained by the replacemet (see (6.5))

$$T_{ij} \to T_{ij} - \frac{c^4}{8\pi\, G}\,\Lambda\, g_{ij} \qquad \Longrightarrow \qquad \begin{cases} \rho \to \rho + \frac{\Lambda}{8\pi G}\,, \\ p \to p - \frac{\Lambda}{8\pi G}\,. \end{cases}$$

From second and third equations in (9.12) then we get

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\,(\rho + 3p) + \frac{\Lambda}{3}\,, \tag{9.13}$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\,\rho - \frac{k}{a^2} + \frac{\Lambda}{3}\,, \tag{9.14}$$

and the static solution $\dot{a} = \ddot{a} = 0$ gives

$$\Lambda = \frac{3k}{a^2} - 8\pi G\rho = 4\pi G(\rho + 3p) > 0\,.$$

We see that there is a static solution only for $\Lambda > 0$ and moreover also k must be positive. The spatial part of the Einstein universe is a hypersphere with constant radius a.

9.4 de Sitter universe

This was introduced by de Sitter in 1917 and is a solution of Einstein equations in the absence of matter, but in the presence of a cosmological constant Λ. It is based on a maximally symmetric manifold in 4-dimensions, which is more than required by physical considerations. Also this model could be seen as a special case of *FLRW*, but due to the maximal symmetry, it is easy to study it directly too. In such a case $a(t)$ is not a constant and so it provides the cosmological red-shift.

The Riemann, Ricci and scalar curvature tensors have the form

$$R_{ijrs} = K(g_{ir}g_{js} - g_{is}g_{jr})\,, \qquad R_{ij} = 3K g_{ij}\,, \qquad R = 12K\,,$$

K being the Gauss curvature of the whole space-time (do not confuse it with the curvature of the spatial section k). Einstein equations become

$$G_{ij} = R_{ij} - \frac{1}{2}\,Rg_{ij} = -3Kg_{ij} = 8\pi G\,T_{ij} - \Lambda g_{ij}\,.$$

It is interesting to note that there are non trivial solutions also in the absence of matter, that is $T_{ij} = 0$. In such a case $K = -\Lambda/3$ and equations (9.13), (9.14) assume the simple form

$$\frac{\ddot{a}}{a} = \frac{\Lambda}{3}\,, \qquad \left(\frac{\dot{a}}{a}\right)^2 = \frac{\Lambda}{3} - \frac{k}{a^2}\,. \tag{9.15}$$

Integrating the fisrt equation in the case with $\Lambda > 0$, which corresponds to de Sitter original model (the solution with $\Lambda < 0$ is called *anti-de Sitter*), we have the general solution

$$a(t) = \frac{1}{2}\left[\alpha\,\exp\left(\sqrt{\frac{\Lambda}{3}}\,t\right) + \beta\,\exp\left(-\sqrt{\frac{\Lambda}{3}}\,t\right)\right]\,, \qquad \frac{\alpha+\beta}{2} = a(0)\,, \tag{9.16}$$

and now from the second equation above we get

$$k = \frac{\alpha\beta\Lambda}{3}\,.$$

We see that the form of $a(t)$ depends on k. In particular, if $k = 0$ then $H(t) = \dot{a}/a = \pm\sqrt{\Lambda/3}$ and the spatial section of the universe is a flat manifold which expands/contracts linearly in time. From (9.16) it follows

$$\begin{aligned} a(t) &= a(0)\,\exp\left(\sqrt{\frac{\Lambda}{3}}\,t\right), && \text{expanding flat solution,} \\ a(t) &= a(0)\,\exp\left(\sqrt{-\frac{\Lambda}{3}}\,t\right), && \text{contracting flat solution.} \end{aligned}$$

Ik $k = 1$ we can choose for example $\alpha = \beta$ and the solution becomes

$$a(t) = a(0)\,\cosh\sqrt{\frac{\Lambda}{3}}\,t\,, \qquad a(0) = \sqrt{\frac{3}{\Lambda}}\,. \tag{9.17}$$

In such a case the spatial part of de Sitter universe is a hypersphere which expands according to (9.17).

Ik $k = -1$ then we can choose for example $\alpha = -\beta$ and the solution becomes

$$a(t) = a(0)\,\sinh\sqrt{\frac{\Lambda}{3}}\,t\,, \qquad a(0) = \sqrt{\frac{3}{\Lambda}}\,. \tag{9.18}$$

In such a case the spatial part of de Sitter universe is a hyperbolic manifold which expands according to (9.18).

The more general solution for $a(t)$ is an arbitrary combination of exponentials or hyperbolic functions. The choice of a particular solution is related to the choice of the parameter t. The solution (9.17) with the hyperbolic cosine above corresponds to a metric without *(cosmological) horizons*, while different choices correpond to metric with horizons, which, as a consequence, do not cover the whole manifold. By transformations and extensions it is always possible to obtain a solution from another one.

The original de Sitter metric was written in the static form

$$ds^2 = -\left[1 - \frac{\Lambda X^2}{3}\right] dT^2 + \frac{dX^2}{1 - \Lambda X^2/3} + X^2 d\sigma^2\,, \tag{9.19}$$

which, in contrast with the *FLRW* metric with $a(t)$ given by (9.17), is singular at $X^2 = 3/\Lambda$ where there is a (cosmological) event horizon. The metric covers only half of the hyperboloid (since we are dealing with Lorentian 4-dimensional spaces, a constant curvature hypersurface is a hyperboloid in 4-dimensions).

Exercise — Verify by a direct computation that the metric in (9.19) satisfies Einstein equation with a cosmological constant Λ.

The Einstein equation with a cosmological constant can be written in the form

$$R_{ij} = \Lambda\, g_{ij}\,.$$

In order to simplify the calculation we use the results already obtained for Schwarzschild . To this aim we put

$$\begin{cases} B = B(X) = 1 - \frac{\Lambda X^2}{3}\,, & \dot{A} = \dot{B} = 0\,, \\ A = A(X) = \frac{1}{1-\Lambda X^2/3} = \frac{1}{B(X)}\,, & \frac{A'}{A} = -\frac{B'}{B}\,, \end{cases}$$

where "dot" and "prime" mean derivative with respect to T and X respectively.

From (7.5)-(7.8) we directly obtain

$$\begin{aligned} R_{00} &= \frac{B''}{2A} - \frac{B'}{4A}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{B'}{rA} = -B\Lambda = g_{00}\Lambda\,, \\ R_{11} &= -\frac{B''}{2B} + \frac{B'}{4B}\left(\frac{A'}{A} + \frac{B'}{B}\right) + \frac{A'}{rA} = \frac{\Lambda}{B} = g_{11}\Lambda\,, \\ R_{22} &= 1 + \frac{r}{2A}\left(\frac{A'}{A} - \frac{B'}{B}\right) - \frac{1}{A} = r^2\Lambda = g_{22}\Lambda\,, \\ R_{33} &= \sin^2\vartheta\, R_{22} = r^2 \sin^{\vartheta}\Lambda = g_{33}\Lambda\,. \end{aligned}$$

Exercise — Verify that the following *Schwarzschild-de Sitter* metric satisfies Einstein equation with a cosmological constant Λ:

$$ds^2 = -\left[1 - \frac{2MG}{r} + \frac{\Lambda r^2}{3}\right] dt^2 + \frac{dr^2}{1 - 2MG/r + \Lambda r^2/3} + r^2 d\sigma^2\,. \tag{9.20}$$

There is an event horizon at the largest real soution of

$$1 - \frac{2MG}{r} + \frac{\Lambda r^2}{3} = 0\,.$$

This corresponds to Schwarzschild-de Sitter black hole. Asymptotically the manifold is de Sitter with $k = 1$.

10 The Standard Cosmological Model

Here we shall discuss the more important consequences of standard cosmology, which is based on the *FLRW* metric. We shall follow the historical point of view first (briefly) considering pure General Relativity and then we will discuss the possible modifications of such a theory in order to take account of recent observational data.

- It has to be noted that the name *standard cosmological model (*SCM*)* for the model of the universe we are going to discuss is not universally used. It has been introduced by Weinberg in analogy with the *standard model* (*SM*) of elementary particles in place of the misleading name *big-bang theory.*

If not otherwise specified, in the following by flat universe we shall mean "spatially flat" universe and by (conformal) coordinates we shall mean "(conformal) comoving coordinates", that is the ones in which the metric has the *FLRW* form. "The observer" O is put in the origin of coordinates and so $O \equiv (t,0,0,0) \equiv (\eta,0,0,0)$, t,η being related by (9.4). A generic point $S \equiv (t,r,\vartheta,\varphi) \equiv (\eta,\chi,\vartheta,\varphi)$, with constant coordinates r,ϑ,φ represents a "fixed" star (galaxy). The coordinates of the galaxy with respect to the observer do not change, but the proper distance increase as a consequece of the expansion. Note that due to homogeneity, all results we shall derive for the observer will be valid for any observer in the universe.

10.1 Cosmological redshift

It has been observed that the spectrum of atoms of distant galaxies is shifted to the red, the shift being proportional to the distance (Hubble[38],1929). It is natural to interpret such a phenomena as due to Doppler effect. This means that all galaxies are moving away with a velocity proportional to their distance. As we shall see, such an effect is a direct consequence of the (non-stationary) *FLRW* metric and gives also an answer to the *Olbers paradox*, because in this way the visible universe is finite.

- **NOTE:** Cosmological redshift is a consequence of the form of the metric and it does not involve the field equations.

It is convenient to introduce the notation

$$H(t) = \frac{\dot a(t)}{a(t)}\,, \qquad H_0 = H(t_0)\,, \qquad a_0 = a(t_0)\,,$$
$$q(t) = -\frac{a(t)\ddot a(t)}{\dot a^2(t)} = -\frac{\ddot a(t)}{H(t)\dot a(t)} = -\frac{\ddot a(t)}{H^2(t)a(t)}\,, \qquad q_0 = q(t_0)\,,$$

where t_0 is the present value of the universal time. The quantity H_0 and the dimensionless parameter q_0 are called respectively the *Hubble constant* and the *deceleration parameter* and can be determined by measuring the red-shift of the spectrum of distant galaxies (see below). Using such notations we get

$$\frac{a(t)}{a(t_0)} = 1 + H_0(t-t_0) - \frac{1}{2}\,q_0H_0^2(t-t_0)^2 + ... \tag{10.1}$$

[38]Edwin Powell Hubble (USA) 1889-1953.

Let $S \equiv (r,0,0)$ be the coordinates of a galaxy with respect to the observer on the earth at a generic time t. In *FLRW* the spatial geometry is represented by the metric

$$dl^2 = a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2)\right],$$

and so the proper distance l_S of the galaxy reads

$$l_S(t) \equiv \int_0^r \frac{a(t)dr}{\sqrt{1-kr^2}} = a(t) f_k(r)\,,$$

where f_k is the function defined in (9.6), but its mathematical form is not important for the calculation. The only important thing is that the proper distance follow a scaling law, with a scale factor $a(t)$ depending on time.

By deriving with respect to time we obtain the speed of the galaxy which is moving away according to

$$v_S(t) = \frac{dl_S}{dt} = \dot{a}(t) f_k(r) = H(t)\, l_S(t) \quad\Longrightarrow\quad v_S(t_0) = H_0 l_S(t_0)\,. \tag{10.2}$$

$v_S(t_0)$ and $l_S(t_0)$ are respectively the velocity and the distance of the galaxy with respect to the observer. Actually the velocity is positive because the measured Hubble constant is positive.

Now we suppose that at time t the galaxy emits a light with frequency $\nu = 1/\Delta t$, which arrives on the earth at time t_0 with frequency $\nu_0 = 1/\Delta t_0$. Since for the light $ds^2 = 0$ we get

$$\frac{dt}{a(t)} = -\frac{dr}{\sqrt{1-kr^2}} \quad\Longrightarrow\quad \int_t^{t_0} \frac{dt}{a(t)} = \int_{t+\Delta t}^{t_0+\Delta t_0} \frac{dt}{a(t)} = -\int_r^0 \frac{dr}{\sqrt{1-kr^2}} = f_k(r)\,, \tag{10.3}$$

from which it follows

$$\int_t^{t+\Delta t} \frac{dt}{a(t)} = \int_{t_0}^{t_0+\Delta t_0} \frac{dt}{a(t)} \quad\Longrightarrow\quad \frac{\Delta t_0}{a(t_0)} = \frac{\Delta t}{a(t)}\,.$$

In deriving the latter equations we have assumed $a(t)$ to be constant during a period of the emitted wave. This is a reasonable assumption because $a(t)$ is a cosmological quantity that changes very slowly. In this way we have

$$\frac{a(t_0)}{a(t)} = \frac{\Delta t_0}{\Delta t} = \frac{\nu}{\nu_0} = \frac{\lambda_0}{\lambda} = 1 + \frac{\lambda_0 - \lambda}{\lambda} = 1 + z\,,$$

where $z = (\lambda_0 - \lambda)/\lambda > 0$ is the *redshift parameter*.

Using expansion (10.1) for $a(t_0)/a(t)$ one obtains

$$z = -H_0(t - t_0) + \left(1 + \frac{1}{2}\, q_0\right) H_0^2 (t - t_0)^2 + \ldots$$

and at the lowest order

$$z \sim H_0(t_0 - t) \sim H_0\, a(t_0) f_k(r)\,, \quad\Longrightarrow\quad z = v_S(t_0) = H_0 l_S(t_0)\,,\ \text{(Hubble law)}.$$

We see that, in agreement with (non-relativistic) Doppler formula z is proportional to the speed of the source. It has to be remarked that the latter expression is valid only if $t_0 - t$ is sufficiently small, which implies that the galaxies have not to be too far. In fact in such a case it is reasonable to assume $a(t)$ to be constant in the interval (t, t_0) and from (10.3) one gets $t_0 - t \sim a(t_0) f_k(r)$.

The redshift parameter can be used as evolution parameter in place of time as a consequence of the relation

$$z = \frac{a(t_0)}{a(t)}\,, \tag{10.4}$$

which is valid for slowly varying $a(t)$.

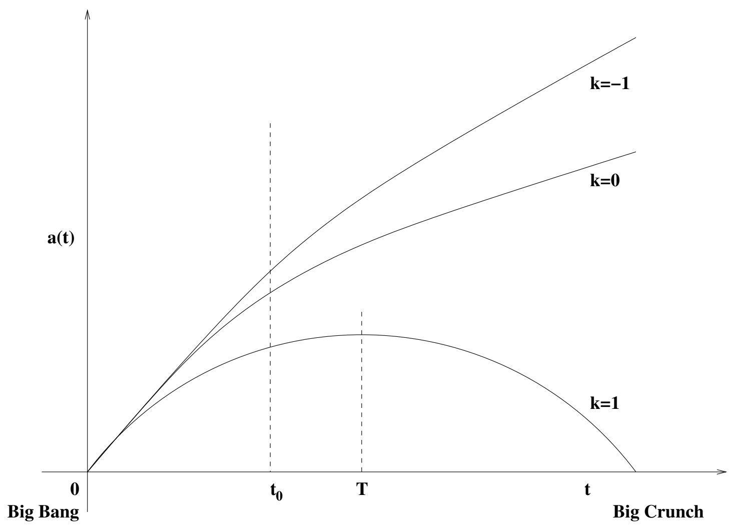

Figure 10: FLRW Model

10.2 Evolution of *FLRW* universe

The evolution of the universe can be determined by studying Friedmann equations in the case of *strong energy dominance condition* $\rho+3p>0$, which is trivially valid for "ordinary" matter/radiation because in such cases both p and ρ are positive (by "ordinary matter/radiation" we mean the one we are usually dealing with, that is atoms, baryons, leptons, photons, neutrinos, etc).

If such hypothesis is satisfied, from second equation in (9.12) it follows that $\ddot{a}/a<0$ and so $a(t)$ is a concave function because $a(t)>0$ by definition. Moreover we know that $\dot{a}(t_0)>0$ because the Hubble constant is positive. This implies that $a(t)$, in a neighbourhood of t_0, is a positive, concave downward, increasing function and so it has to intersect the temporal axis in a point O, which we take as the origin of time (see figure 10). In such a point, called *Big Bang*, the Friedmann equations become singular and so it has no physical sense to cross it.

Due to the concavity of the function (see figure 10) the ratio $a(t_0)/t_0>\dot{a}(t_0)$ and so $t_0<1/H_0$. With a value $H_0\sim 70\,(Km/sec)/Mpc$ the actual age of the universe becomes lower than 14 billion years.

Looking at Friedmann equation in (9.12) one sees that in the case $k=1$, $\dot{a}(t)$ has to be vanish at a given value of $t>t_0$ In such a case $a(t)$ achieves its maximum value at $t=T>t_0$ and so the universe is *spatially closed*, it expands during a period T, which is the solution of the equation

$$\rho(T)a^2(T)=\frac{3}{8\pi G}\,. \tag{10.5}$$

At time $t=T$ the universe achieves its maximum and then it contracts for a period T and reaches again the initial singularity at $t=2T$ (*Big Crunch*).

In the two other cases $k=0,-1$ the universe expands indefinitely. Since the density decreases as a^{-3} or greater, in the limit $t\to\infty$ from (9.12) one obtains $\dot{a}\to 0$ or $\dot{a}^2\to 1$ according to whether $k=0$ or $k=-1$.

10.2.1 Actual critical parameters

In a matter dominated era the Friedmann equations evaluated at $t = t_0$ can be written in terms of the parameters $\rho_0 = \rho(t_0), a_0, H_0, q_0$, $p_0 \sim 0$ and the *critical density*

$$\rho_c = \frac{3H_0^2}{8\pi G} \,. \tag{10.6}$$

From second and third equations (9.12) with $p_0 = 0$ in particular we get

$$\frac{\rho_0}{\rho_c} - 1 = \frac{k}{H_0^2 a_0^2} = 2q_0 - 1 \quad \Longrightarrow \quad \frac{\rho_0}{\rho_c} = 2q_0 \,. \tag{10.7}$$

We see that the value of the spatial curvature depends on the value of the actual density ρ_0 in relation to the critical one ρ_c or, **alternatively**, from the value of the deceleration parameter q_0. In fact

$$\begin{array}{lclcl}
\rho_0 = \rho_c & \Longrightarrow & k = 0 & \Longleftarrow & q_0 = \frac{1}{2} \,, \\
\rho_0 > \rho_c & \Longrightarrow & k = 1 & \Longleftarrow & q_0 > \frac{1}{2} \,, \\
\rho_0 < \rho_c & \Longrightarrow & k = -1 & \Longleftarrow & q_0 < \frac{1}{2} \,.
\end{array}$$

Of course, such results have to be compared with the measured values of the parameters. In the past, a privileged value for q_0 was of the order of unity, giving rise to a closed universe with $k = 1$ and $\rho_0/\rho_c > 1$, while the estimated value of the density in ratio to the critical one was very small ($\rho_0/\rho_c \sim 0.01$). Then one was talking about a *missing mass* (black holes, neutrinos, dusts,...).

At the end of the last century the situation was completely changed as a consequence of very precise new data, but the problem of missing mass, now called *dark matter* and *dark energy*, still remains (see Section 11.2).

10.2.2 The age of the universe

By assuming the universe to be dominated by matter for the most part of his history, we can derive a relation between its age and q_0. In such a case the fourth equation in (9.12) trivially gives

$$\rho(t) = \left(\frac{a_0}{a}\right)^3 \rho_0 = \left(\frac{a_0}{a}\right)^3 2q_0\rho_c \quad \Longrightarrow \quad \frac{\rho_0}{\rho_c} = 2q_0 \left(\frac{a_0}{a}\right)^3 .$$

Now using (10.7) the Friedmann equation in (9.12) assumes the form

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2 \left[1 - 2q_0 + 2q_0\frac{a_0}{a}\right] \left(\frac{a_0}{a}\right)^2$$

and finally, by integration

$$t_0 = \frac{1}{H_0}\int_0^{a_0} \left[1 - 2q_0 + 2q_0\frac{a_0}{a}\right]^{-1/2} \frac{da}{a_0} = \frac{1}{H_0}\int_0^1 \left[1 - 2q_0 + \frac{2q_0}{x}\right]^{-1/2} dx < \frac{1}{H_0} \,, \tag{10.8}$$

that gives the age of the universe in terms of the deceleration parameter q_0. In a flat universe, $k = 0$, $\rho_0/\rho_c = 1$, $q_0 = 1/2$ and so $t_0 = 2/3\, H_0$.

- It has to be noted that the results of the latter paragraph are valid only during the matter dominated era. From recent experimental data we know that recently the expansion started to accelerate and so the actual "deceleration" parameter is negative. In such a case equation (10.8) has to be modified in order to take into account of "repulsive" forces.

10.3 Explicit solutions of Friedman equations

In order to solve the system in (9.12) it is necessary to know the equation of state of matter, which in principle can change during the evolution and could be very complicated, However explicit solutions can be easily obtained when the pressure is proportional to the density, that is $p = w\rho$, with the *barotropic parameter* w to be a constant. Such equation of state cover the two important cases corresponding to low and high velocities of particles and also the cosmological constant case.

Then putting $p = w\rho$, equations in (9.12) assume the form

$$\begin{cases} \frac{2\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = -8\pi G w\rho \,, & \text{field equation,} \\ \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(1+3w\right)\rho \,, & \text{field equation,} \\ \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G}{3}\,\rho \,, & \text{Friedmann equation,} \\ \frac{d}{da}\left(a^3\rho\right) = -3w\rho\, a^2 \,, & \text{conservation law,} \\ p = w\,\rho \,, & \text{equation of state,} \end{cases} \tag{10.9}$$

Equation fourth in (10.9) can be easily solved, the solution being

$$\rho(t) = \frac{const}{a^{3(1+w)}} \,, \qquad p(t) = w\rho(t) \,.$$

The integration constant has to be fixed by initial conditions.

As we already said above for suitable choices of the barotropic parameter the equation of state (10.9) cover physical important cases. With the chosen signature, the trace of the energy-momentum tensor has to be negative (for ordinary matter/energy), then we get

$$0 \le p \le \frac{\rho}{3} \,, \qquad \text{ordinary matter.}$$

The two limits correspond to

$$\begin{aligned} &\text{non relativistic matter} \qquad \Longrightarrow \quad p \ll \rho \,, \\ &\text{ultra-relativistic matter} \qquad \Longrightarrow \quad p \sim \frac{1}{3}\,\rho \,, \end{aligned}$$

which correspond to the choices $w = 0$ and $w = 1/3$ respectively. The choice $w = -1$ corresponds to the cosmological constant, which can be seen as a perfect fluid with negative pressure. In summary we have

$$\text{dust} \quad \Longrightarrow \quad p = 0 \quad \Longrightarrow \quad \rho(t) = \frac{cost}{a^3(t)} \,, \tag{10.10}$$

$$\text{radiation} \quad \Longrightarrow \quad p = \frac{\rho}{3} \quad \Longrightarrow \quad \rho(t) = \frac{cost}{a^4(t)} \,, \tag{10.11}$$

$$\text{cosm. const.} \quad \Longrightarrow \quad p = -\rho \quad \Longrightarrow \quad \rho(t) = const \,. \tag{10.12}$$

The integration constants can be fixed by initial conditions. For example, by measuring matter density $\rho_M(t_0)$ at present time $t = t_0$, we have

$$\rho_M(t) = \frac{\rho_0 a_0^3}{a^3(t)} \,, \qquad \rho_0 = \rho_M(t_0) \,, \qquad a_0 = a(t_0) \,.$$

The density ρ of the universe is probably a "mixture" due to matter, radiation and cosmological constant, but, looking at equations (10.10)-(10.12) we see that, due to the different behaviour during the expansion, we can distinguish three different ere in which one of the three components dominates with respect to the others, starting from the beginning, when radiation dominates, until the present era dominated by the component due to cosmological constant.

Note that the ratio between non-relativistic (dust) and ultra-relativistic (radiation) matter depends on the energy. At the beginning of time, only ultra-relativistic matter was present.

10.3.1 Flat space

According to recent cosmological data. the spatial section of the universe is flat, that is $k = 0$. In this case the Friedmann equations can be easily solved. To this aim it is sufficient to consider equation (10.9) (with $k = 0$) for the two cases $w \neq -1$ and $w = -1$. In this way we get

$$\left[a^{(3w+1)/2}\,\dot{a}\right]^2 = const \quad \Longrightarrow \quad a(t) = a_0\left[\frac{t}{t_0}\right]^{\frac{2}{3(w+1)}}\,, \qquad w \neq -1\,. \tag{10.13}$$

For $w = -1$ the behaviour of the solution completely changes because

$$H(t) = \frac{\dot{a}}{a} = const = H_0 \quad \Longrightarrow \quad a(t) \sim e^{H_0 t}\,. \tag{10.14}$$

We see that in both radiation and matter dominated ere the radius of the universe increase slowly, while in the era dominated by cosmological constant the radius increase exponentially.

The actual age of the universe can be easily computed by assuming that the main contribution is due to matter dominated era. This is a reasonable assumption because radiation dominated for a brief period and cosmological constant starts to dominate very recently. From the solution (10.13) with $w = 0$ we get

$$H(t) = \frac{\dot{a}}{a} = \frac{2}{3\,t} \quad \Longrightarrow \quad t_0 \sim \frac{2}{3\,H_0}\,. \tag{10.15}$$

$1/H_0$ is called the *Hubble time*.

In the following we shall also need the age of the universe as a function of the Hubble parameter in a radiation-dominated universe. From (10.13) with $w = 1/3$ we have

$$H(t) = \frac{\dot{a}}{a} = \frac{1}{2\,t} \quad \Longrightarrow \quad t \sim \frac{1}{2\,H}\,. \tag{10.16}$$

10.3.2 Curved space

Now we solve Friedmann equations in the presence of curvature. To this aim it is convenient to use the conformal coordinates (9.4). In such coordinates, for an arbitrary function $f(t) = f(\eta(t))$ we have

$$\dot{f} = \frac{df}{dt} = \frac{df}{d\eta}\frac{d\eta}{dt} = \frac{f'(\eta)}{a(\eta)}\,, \qquad f' = \frac{df}{d\eta}\,.$$

By setting $h = a'/a$, the first three equations in (10.9) read

$$2h' + h^2 + k = -8\pi G\, w\rho\, a^2\,, \tag{10.17}$$

$$h' = -\frac{4\pi G}{3}\,(3w+1)\rho\, a^2\,, \tag{10.18}$$

$$h^2 + k = \frac{8\pi G}{3}\,\rho\, a^2\,. \tag{10.19}$$

We consider separately the three cases $w = 0$, $w = 1/3$ and $w = -1$.

- $w = 0$. In such a case equation (10.17) can be directly solved because it becomes

$$2h' + h^2 + k = 0 \quad \Longrightarrow \quad h(\eta) = \begin{cases} \cot\frac{\eta}{2}\,, & k = 1\,,\\ \frac{2}{\eta}\,, & k = 0\,,\\ \coth\frac{\eta}{2}\,, & k = -1\,. \end{cases}$$

Now, by integrating the equation $h = a'/a$ we obtain the radius as a function of η, that is

$$a(\eta) = A \begin{cases} 1 - \cos\eta\,, & k = 1\,, \\ \frac{\eta^2}{2}\,, & k = 0\,, \\ \cosh\eta - 1\,, & k = -1\,, \end{cases} \tag{10.20}$$

and finally

$$t(\eta) = \int a(\eta)\, d\eta = T \begin{cases} \eta - \sin\eta\,, & k = 1\,, \\ \frac{\eta^3}{6}\,, & k = 0\,, \\ \sinh\eta - \eta\,, & k = -1\,, \end{cases}$$

where A, T are integration constants to be determined by initial conditions. We see that in all the cases the size of the universe in vanishing at $\eta = 0$ (Big Bang). In both the cases $k = 0, -1$ the universe expands indefinitely, while for $k = 1$ the universe reaches its maximal size for $\eta = \pi$ and then $a(\eta)$ decreases and reaches the value $a(\eta) = 0$ for $\eta = 2\pi$ (Big Crunch).

- $w = 1/3$. In such a case, by summing (10.18) and (10.19) we have

$$h' + h^2 + k = 0 \quad \Longrightarrow \quad h(\eta) = \begin{cases} \cot\eta\,, & k = 1\,, \\ \frac{1}{\eta}\,, & k = 0\,, \\ \coth\eta\,, & k = -1\,, \end{cases}$$

and for $a(\eta)$ and $t(\eta)$ it follows

$$a(\eta) = A \begin{cases} \sin\eta\,, & k = 1\,, \\ \eta\,, & k = 0\,, \\ \sinh\eta & k = -1\,, \end{cases} \qquad t(\eta) = T \begin{cases} 1 - \cos\eta\,, & k = 1\,, \\ \frac{\eta^2}{2}\,, & k = 0\,, \\ \cosh\eta - 1\,, & k = -1\,. \end{cases} \tag{10.21}$$

Also in this case the universe expands indefinitely for $k = 0, -1$ and it is finite for $k = 1$, but now $a(\eta)$ reaches its maximum for $\eta = \pi/2$ and vanishes again for $\eta = \pi$.

- $w = -1$. Such a case has been already considered in section (9.4).
In fact, setting $-p = \rho = \Lambda/8\pi G$ we get (see (9.15)

$$\frac{\ddot{a}}{a} = \frac{\Lambda}{3}\,, \qquad \left(\frac{\dot{a}}{a}\right)^2 = \frac{\Lambda}{3} - \frac{k}{a^2}\,.$$

For $\Lambda > 0$ we obtain

$$a(t) = A \begin{cases} \cosh\sqrt{\frac{\Lambda}{3}}t\,, & k = 1\,, \\ \exp\sqrt{\frac{\Lambda}{3}}t\,, & k = 0\,, \\ \sinh\sqrt{\frac{\Lambda}{3}}t\,, & k = -1\,. \end{cases}$$

Now the universe expands indefinitely in all the cases, but this is not a surprise because all the above solutions correspond to de Sitter space. The first solution covers the whole manifold, while the other two cover only part of it.

If $\Lambda < 0$ (*anti de Sitter*) then there is a solution only in the case $k = -1$. It reads

$$a(t) = A \sin\sqrt{\frac{|\Lambda|}{3}}t\,, \qquad k = -1\,.$$

10.4 Horizons

In *FLRW* can exist both *particle* as well as *event horizons*, which delimit regions inaccessible to the observer. More precisely, a particle horizon delimits a region in the past which cannot be seen by the observer at a given time, but it could be seen in the future, while an event horizon delimits a region in the future which will never be influenced by the observer. In order to see that it is convenient to use conformal coordinates (9.4) and metric (9.7).

Let us consider the observer O in the origin of coordinates who received a light signal at conformal time η emitted by a comoving source S (star/galaxy) at coordinate distance χ_i and at conformal time η_i. Angular coordinates do not enter the game because the light travels along radial geodesics, for which $d\vartheta = d\varphi = 0$, that is

$$ds^2 = 0\,, \qquad \vartheta = const\,, \qquad \varphi = const \quad \Longrightarrow \quad d\eta^2 = d\chi^2\,.$$

These correspond to straight lines at angles $\pm\pi/4$ in the $\{\eta, \chi\}$ plane. For conformal coordinate and proper distance of the source we obtain respectively

$$\chi_i = \chi_i(t, t_i) = \eta - \eta_i = \int_{t_i}^{t} \frac{dt'}{a(t')}\,, \qquad l_i(t, t_i) = a(t)\chi(t) = a(t)\int_{t_i}^{t} \frac{dt'}{a(t')}\,. \tag{10.22}$$

Of course, for a fixed t_i, at time t the observer can not get informations concerning objects at distance $l > l_i$, but in principle, for a fixed t, $l_i(t, t_i)$ could increase indefinitely as a function of $t_i < t$. For example, in Minkowski space $l_i = t - t_i$ and the observer, ant any time, can receive signals from all points because $|t_i|$ can be arbitrary large ($t_i \to -\infty$). The light past cone of any observer in Minkowski space covers the whole spatial manifold.

The situation drastically changes if the integral in (10.22) is finite for any choice of $t_i < t$. In particular, if t_i is the "beginning of time" ($t_i = 0$ in *FLRW*), then the observer can only see the object which lie inside the hypersurface $l \leq l_i(t, t_i) \equiv d_P(t)$, which can be finite. In such a case the hyper-surface $l = d_P(t)$ is called *particle horizon*. This is a dynamical quantity because it depends on time. It could increase but also decrease.

- It has to be noted that our universe was opaque to photons until the time of hydrogen recombination (when universe was more or less 1000 times smaller than now) and so, for physical reasons, there is an "optical horizon" with $d_{opt} < d_P(t_0)$, which "obscures" information about the most interesting stages of the evolution of the early universe.

As well as an observer could not see regions in the past (at present time), there could be regions in the future with which he will not be able to communicate. In fact, the signal sent by the observer at time η, t, will never reach the points for which

$$\chi > \chi_E = \eta_f - \eta \quad \Longrightarrow \quad l > d_E(t) = a(t)\int_{t}^{t_f} \frac{dt'}{a(t')}\,,$$

t_f being the final time, which can be finite or infinite. The hypersurface $l = d_E(t)$ is called *event horizon*. At present time we could never communicate with observers in the region $l > d_E(t_0)$,

Particle horizons arise when the past light cone of the observer O terminates at a **finite conformal time** η_i. Then there will be worldlines of other particles which do not intersect the past of O, meaning that they were never in causal contact. One has

$$\chi_P = \eta - \eta_i\,, \qquad d_P(t) = a(t)\int_{t_i}^{t} \frac{dt'}{a(t')}\,, \tag{10.23}$$

η and t being the conformal and universal time of the observer.

Event horizons arise when the future light cone of the observer O terminates at a **finite conformal time**. Then there will be worldlines of other particles which do not intersect the future of O, meaning that they cannot possibly influence each other. One has

$$\chi_E = \eta_f - \eta\,, \qquad d_E(t) = a(t)\int_t^{t_f} \frac{dt'}{a(t')}\,, \tag{10.24}$$

η and t being the conformal and universal time of the observer.

In order to clarify the meaning of horizons we explicitly compute them in flat *FLRW*, for the three cases considered above (10.13,10.14).

- $w = 1/3$. In the radiation-dominated era we have

$$a(t) = A\sqrt{t}\,, \quad \Longrightarrow \quad d_P(t) = A\sqrt{t}\int_0^t \frac{dt'}{A\sqrt{t'}} = 2t\,. \tag{10.25}$$

  We see that the particle horizon increase linearly in time, while there is no event horizon because $d_E(t) = \infty$ when $t_f = \infty$. At the time t, the observer will see only objects inside the region $l < 2ct$.

- $w = 0$. In the matter-dominated era we have

$$a(t) = A\,t^{2/3}\,, \quad \Longrightarrow \quad d_P(t) = A\,t^{2/3}\int_0^t \frac{dt'}{A\,t'^{2/3}} = 3t\,. \tag{10.26}$$

  Also in this case the particle horizon increase linearly in time and there is no event horizon. At the time t, the observer will see only objects inside the region $l < 3ct$.

- $w = -1$. In such a case, setting $H_\Lambda = \sqrt{\Lambda/3}$ we have

$$a(t) = A\,e^{H_\Lambda t} \quad \Longrightarrow \quad \begin{cases} d_P(t) = e^{H_\Lambda t}\int_{t_i}^t e^{-H_\Lambda t}\,dt = \frac{1}{H_\Lambda}\left(e^{H_\Lambda(t-t_i)} - 1\right)\,, \\ d_E(t) = e^{H_\Lambda t}\int_t^\infty e^{-H_\Lambda t}\,dt = \frac{1}{H_\Lambda}\,. \end{cases} \tag{10.27}$$

  We see that there is an event horizon which is independent on t. The observer can not influence the region out of the sphere $d > d_E$. The particle horizon depends on the initial time t_i. If one considers the whole de Sitter space, then $t_i = -\infty$ and in such a case $d_P = \infty$. This means that the observer can see all the manifold at any time, but if for some physical reason t_i has to be finite, then there is also a particle horizon depending on time.

10.5 Conformal diagrams

As we have seen by general considerations, the metric of a homogeneous and isotropic manifold can be written in the conformal coordinates as

$$ds^2 = a^2(\eta)\left[-d\eta^2 + d\chi^2 + \phi_k^2(\chi)\,d\sigma^2\right]\,, \qquad \phi_k(\chi) = \begin{cases} \sin\chi\,, & k = 1\,, \\ \chi\,, & k = 0\,, \\ \sinh\chi\,, & k = -1\,, \end{cases}$$

$d\sigma^2$ being the metric of the unitary 2−dimensional sphere. The global properties of a manifold with a metric of such a form are completely determined by the radial geodesics and can be conveniently represented by a 2-dimensional *conformal diagram* in which every point corresponds to a 2-sphere. This possibility only depends on the form of the metric and so conformal diagrams are used in other contests too, for example in the physics of black holes. In such cases, the factors a^2 and ϕ could depend of both η, χ.

In general the coordinates η, χ extend to infinite intervals, but it is always possible to perform a further transformation of coordinates which preserves the form of the metric and maps both η, χ in finite intervals. In this way we can represent the manifold by a compact diagram (η, χ), in which the radial geodesics are the straight lines $\chi = \pm\eta + const$. The size of the diagram and the range spanned by coordinates can be altered, but the shape is uniquely determined by the form of the metric. Of course, all metrics related by non singular conformal transformations give rise to the same diagram.

In the construction of the diagram one has to take into account also of the singularities and of the boundaries which are determined by the functions $a(\eta, \chi)$ and $\phi_k(\eta, \chi)$.

Closed radiation dominated universe. As a first example we consider the universe dominated by radiation with $k = 1$. The metric and solution read (see 10.21)

$$ds^2 = -a^2(\eta)\left[-d\eta^2 + d\chi^2 + \phi_1^2(\chi)\, d\sigma^2\right], \qquad \begin{cases} \phi_1(\chi) = \sin\chi\,, \\ a(\eta) = A\,\sin\eta\,, \\ 0 = \eta_i =\leq \eta \leq= \eta_f = \pi\,, \\ 0 \leq \chi \leq \pi\,, \end{cases}$$

A being determined by initial conditions.

All the coordinates have finite range and cover the whole manifold. The conformal diagram is a square with boundaries at $(\eta, 0)$ and (η, π) and physical singularities at $(0, \chi)$ and (π, χ) where the conformal factor vanishes, while energy and pressure diverge. In the period $0 < \eta < \pi/2$ the universe expands, while it contracts for $\pi/2 < \eta < \pi$ (see figure (11).

The particle and event horizons for the observer read respectively

$$\chi_P(\eta) = \eta - \eta_i = \eta\,, \qquad \chi_E(\eta) = \eta_f - \eta = \pi - \eta\,.$$

We see that the particle horizon increase with time and the observer will see all the space at $\eta = \pi$, that is just at the moment of recollapse.

There is an event horizon at any time and this means that there are regions which can not be influenced by the observer at time η.

Closed matter dominated universe. As a second example we consider the universe dominated by matter with $k = 1$. The metric and solution read (see 10.20)

$$ds^2 = -a^2(\eta)\left[-d\eta^2 + d\chi^2 + \phi_1^2(\chi)\, d\sigma^2\right], \qquad \begin{cases} \phi_1(\chi) = \sin\chi\,, \\ a(\eta) = A(1 - \cos\eta)\,, \\ 0 = \eta_i \leq \eta \leq \eta_f = 2\pi\,, \\ 0 \leq \chi \leq \pi\,. \end{cases}$$

Also in this case all the coordinates have finite range and cover the whole manifold, but now the conformal diagram is a rectangle with boundaries at $(\eta, 0)$ and (η, π) and singularities at $(0, \chi)$ and $(2\pi, \chi)$ where $a(\eta)$ vanishes (see figure (11). The universe expands for $0 < \eta < \pi$ and contracts for $\pi < \eta < 2\pi$.

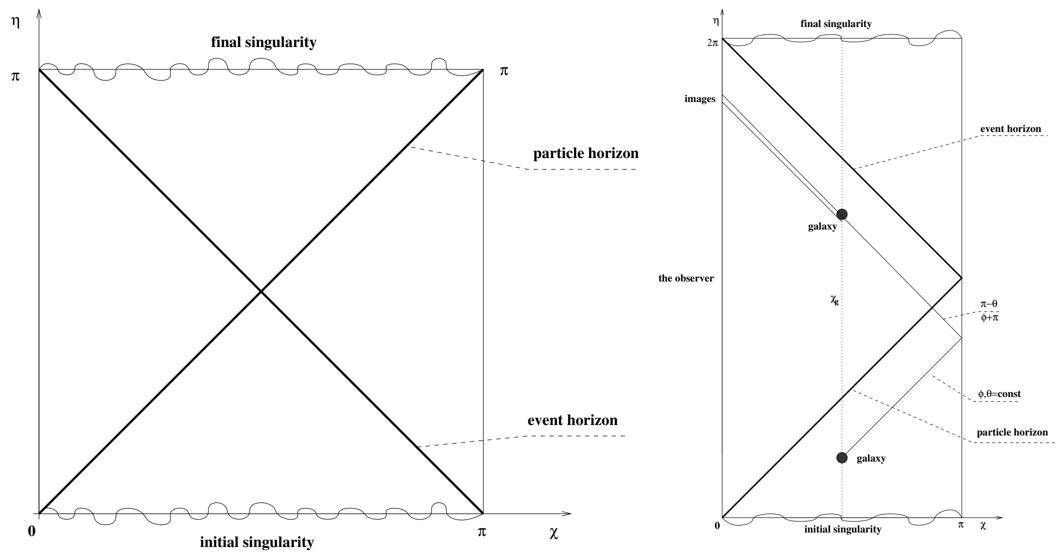

Figure 11: Conformal diagrams: closed radiation (left) and matter (right) flat universes

Particle and event horizons for the observer read respectively

$$\chi_P(\eta) = \eta - \eta_i = \eta\,, \qquad\qquad \chi_E(\eta) = \eta_f - \eta = 2\pi - \eta\,.$$

Looking at the diagram we see that there is a particle horizon but only during the expansion era and an event horizon but only during the contraction era. The observer can influence any region of the universe in the period $\eta < \pi$ and can see any region of the universe in all period $\eta > \pi$.

It has to be noted that in the period $\eta > \pi$ the observer at $\chi = 0$ can see at the same time two copies of the same galaxy, one older than the other, because the light emitted in opposite directions will reach the observer at different times. This is visualized in the diagram by observing that the light signals are "reflected" on the boundaries at $\chi = 0$ and $\chi = \pi$ and move back. This is due to the fact that spatial section is a hypersphere.

As well as $\vartheta = 0$ and $\vartheta = \pi$ correspond to opposite poles of the 2-dimensional sphere $d\sigma^2 = d\vartheta^2 + \sin^2\vartheta\, d\varphi^2$, $\chi = 0$ and $\chi = \pi$ correspond to opposite poles of the 3-dimensional sphere $d\Sigma^2 = d\chi^2 + \sin^2\chi\, d\sigma^2$. It is well known that it is not possible to cover all the sphere with a single chart and in fact our coordinates cover the 3-sphere, but a point, say the south pole. In order to clarify what happens in a neighbourhood of such a point, one has to use another chart.

Consider for example a light signal moving on the 3-sphere along a radial geodesic. Let's $P_+ \equiv (\chi, \vartheta, \varphi)$ be the trajectory, ϑ and φ being constant and $\chi = \chi(\eta)$ increasing in time. When the signal cross the south pole $\chi = \pi$ at crossing time η_S, then $\varphi \to \varphi + \pi$ and $\vartheta \to \pi - \vartheta$ and χ starts to decrease. Note that P_+ and $P_- \equiv (\chi, \pi - \vartheta, \varphi + \pi)$ are specular points (on the geodesic) with respect to the north-south axis. P_+ is the trajectory for $\eta < \eta_S$ and P_- is the trajectory for $\eta > \eta_S$. This means that when the signal cross the south pole the geodesic in the the conformal diagram is represented by a reflected line on the boundary at $\chi = \pi$. Of course the same thing holds for a signal which propagates in the opposite direction. In such a case it is reflected on the boundary at $\chi = 0$.

As we have already said, an observer at $\chi = 0$ will see two copies of the galaxy at $\chi = \chi_g$. Looking at the digram it seams that the observer moving with the galaxy itself will see its younger image at $\eta = \eta_1$, but this is not the case because its coordinates at $\eta = \eta_2$ are $(\chi_g, \vartheta, \varphi)$, while the coordinates of the younger image at $\eta = \eta_2$ are $(\chi_g, \pi - \vartheta, \varphi + \pi)$. We have not to forget that any point of the diagram represents a sphere. This fact becomes clear for the observe at the origin. The light emitted from a galaxy at origin will never intersect the world line of the observer.

11 Thermal history of the universe

(Lemaître–1931) "The evolution of the world can be compared to a display of fireworks that has just ended: some few red wisps, ashes and smoke. Standing on a well chilled-cinder we see the slow fading of the suns, and we try to recall the vanished brilliance of the origin of the world."

The thermal history of the universe has been developed by many people starting from 1930 and it is far to be completed. The standard cosmological picture deals with the universe as it is now and as one can trace its evolution back in time. It is based on the following assumptions:

- On large scale average the mass distributions is close to homogeneous.
- The universe is expanding according to Hubble law.
- The dynamics is described by general relativity and local physics is the same everywhere and at all time.
- The universe expands from a hot state dominated by thermal black body radiation.

The evolution can be parametrized not only by time, but also by the redshift parameter or by the "equilibrium temperature", which is the really important one, because it determines what kind of particles are present and which reactions are permitted. Of course, we assume that all physical laws which we kn ow from particle experiments are also valid during all evolution of the universe.

The relation between temperature of black-body radiation and redshift parameter is very simple (see below), while the relation with time depends on matter content and in general is quite complicated, but for the first period dominated by radiation one can use the (roughly) approximated formula (pure numerical identity–dimensionally wrong)

$$T_{MeV} \sim \frac{1}{\sqrt{t_{sec}}} \,,$$

where time has to be measured in seconds and temperature in MeV. Remember that temperature and energy are related through the relation $E = kT$, k being the Boltzmann constant. In the following there is no possibility to confuse it with the curvature parameter.

11.1 A schematic description

Now we write down a schematic description of the thermal evolution and then we shall analyse some parts in more detail.

$t \sim$ $4.35 \times 10^{17}\, sec \sim 13.8 \times 10^{9}\, years$ –
Actual age of the universe usually indicated with t_0.

$t \sim$ $10^{16} - 10^{17}\, sec$ –
Galaxies and their clusters are formed from small initial inhomogeneities as a result of gravitational instability. Structure formation can be described using Newtonian gravity.

$t \sim$ $10^{12} - 10^{13}\, sec$ –
Nearly all free electrons and protons recombine and form neutral hydrogen. The universe becomes transparent to the background radiation. The cosmic microwave background (CMB) temperature fluctuations, induced by the slightly inhomogeneous matter distribution at recombination, survive to the present day and deliver direct information about the state of the universe at the last scattering surface.

$t \sim$ $10^{11}\, sec\, -\, T \sim\, eV$ –
Matter-Radiation equality which separates the radiation-dominated epoch from the matter-dominated one. The exact value of the cosmological time at equality depends on the constituents of the dark component and it is known at present only up to a numerical factor of order unity.

$t \sim$ $200 - 300\, sec\, -\, T \sim 0.05\, MeV$ –
Nuclear reactions become efficient at this temperature. Free protons and neutrons form helium and other light elements. The abundances of the light elements resulting from *primordial nucleosynthesis* are in very good agreement with available observation data and this strongly supports our understanding of the universe evolution back to the first second after the big bang.

$t \sim$ $1\, sec\, -\, T \sim 0.5\, MeV$ –
Energy of the order of the electron mass. The numerous electron-positron pairs present in the very early universe begin to annihilate when the temperature drops below their rest mass and only a small excess of electrons over positrons remains.

$t \sim$ $0.2\, sec\, -\, T \sim 1 - 2\, MeV$ –
Primordial neutrinos decouple from the other particles and propagate without further scatterings. Ratio of neutrons to protons freezes out because the interactions that keep neutrons and protons in chemical equilibrium become inefficient. The number of the surviving neutrons determines the abundances of the primordial elements.

$t \sim$ $10^{-5}\, sec\, -\, T \sim 200\, MeV$ –
Quark-gluon transition takes place: free quarks and gluons become confined within baryons and mesons (physics not completely understood).

$t \sim$ $10^{-10} - 10^{-14}\, sec\, -\, T \sim 100\, GeV - 10\, TeV$ –
This range of energy scales can still be probed by accelerators. The electroweak symmetry is restored and the gauge bosons become massless.

$t \sim$ $10^{-14} - 10^{-43}\, sec\, -\, T \sim 10\, TeV - 10^{19}\, GeV$ –
This energy range will not be reached by accelerators. The very early universe can give us some rough information about fundamental physics. We can still use General Relativity to describe the dynamics. The main uncertainty is the matter composition. Supersymmetric particles? WIMP (weakly interacting massive particles candidates for dark matter).

The origin of baryon asymmetry in the universe is related to physics beyond the standard model (*SM*). Grand Unification of electroweak and strong interactions takes place at energies about 10^{16} GeV. Topological defects, such as cosmic strings, monopoles, that occur naturally in unified theories might play some role in the early universe, but according to the current microwave background anisotropy data, it seems that they have any significance for large scale structure.

Perhaps the most interesting phenomenon in the above energy range is the accelerated expansion of the universe (*inflation*), which probably occurs somewhere near Grand Unification scales. Fortunately, the most important robust predictions of inflation do not depend substantially on unknown particle physics.

$t \sim\ 10^{-43}\,sec\ -\ T \sim 10^{19}\,GeV\ -$
Near the Planckian scale quantum gravity dominates, spacetime could have no meaning and general ralativity can no longer be trusted.

11.2 Actual cosmological parameters

The Friedmann equations now are usually written by introducing a dimensionless parameter Ω by

$$\Omega = \Omega(t) = \frac{8\pi G\rho(t)}{3H^2(t)}\,, \qquad \Omega_0 = \Omega(t_0) = \frac{\rho_0}{\rho_c}\,, \qquad \rho_c = \frac{3H_0}{8\pi G}\,,$$

ρ_c being the actual critical density already introduced in previous chapter. In some cases Ω is also split in matter, radiation, dark-energy components, but here we will do it only at present time t_0. From equations in (9.12) we get

$$\Omega - 1 = \frac{k}{a^2(t)H^2(t)}\,, \qquad \frac{d\Omega}{da} = (1+3w)\,\frac{\Omega(\Omega-1)}{a}\,, \tag{11.1}$$

where $p = w\rho$ and $dH/da = \dot{H}/\dot{a}$ have been used. We see that the spatial curvature k is positive, vanishing or negative according to whether Ω is larger, equal or smaller than 1.

It is convenient to introduce the following notation for actual values of cosmological parameters:

$H_0 = 100\,h\,(Km/sec)/Mpc$	Hubble constant	$(h \sim 0.68)\,;$
$\tau_0 = \frac{1}{H_0} \sim 0.98 \times\ 10^{10}h^{-1}\,years$	Hubble time	$(\tau_0 \sim 1.45 \times\ 10^{10}\,years)\,;$
$L_0 = \frac{c}{H_0} \sim 3000\,h^{-1}\,Mpc$	Hubble length	$(L_0 \sim 4000\,Mpc)\,;$
$\Omega_M = \frac{8\pi G\rho_M}{3H_0^2}$	matter	$(\Omega_M \sim 0.3)\,;$
$\Omega_k = -\frac{k}{a_0^2H_0^2}$	curvature	$(\Omega_k \sim 0)\,;$
$\Omega_\Lambda = \frac{\Lambda}{3H_0^2}$	dark energy	$(\Omega_\Lambda \sim 0.69)\,;$

$H_0, \Omega_M, \Omega_k, \Omega_\Lambda$ being constant quantities to be measured. H_0 is measured by redshift, while Ω_M is deduced in different ways (galaxy rotation, virial theorem). The best value of Ω_k (as well as Ω_M) is deduced from *CMB*, while Ω_Λ is deduced by Friedmann equation in (11.1) at prensent time, that is

$$\Omega_0 = \Omega_M + \Omega_\Lambda = 1 - \Omega_k \quad \Longrightarrow \quad \Omega_\Lambda = 1 - \Omega_M - \Omega_k\,.$$

We shall see below that the contribution of radiation is actually negligible.

One has to pay attention to the fact that all data in the right hand side of the table above depends on the value of h, related to the Hubble constant. According to recent results (Planck spacecraft 2013), $h \sim 0.68$.

Note that Ω_Λ represents the contribution of *dark energy*, which in the ΛCDM model is related to the cosmological constant by the definition above. The ΛCDM (Λ-Cold Dark Matter) model is a *FLRW* model with cosmological constant Λ, where the matter parameter Ω_M is (quite entirely) due to weakly interacting massive particles (WIMP), that is particle which manifest only through gravitational or weak interactions. They are not visible since they have no charges (electric,color) and they have decoupled very early and so they are could (for non relativistic matter the temperature scales as $T_M \sim 1/a^2$).

The ΛCDM model is in good agreement with experimental data, but since it presents some theoretical problems (see Section (11.7), other possibilities are actually under consideration, where Λ is not given “a priori”, but it emerges as an effective dynamical quantity, for example as a consequence of modified general relativity.

The matter parameter is usually split as

$$\Omega_M = \Omega_B + \Omega_D\,, \qquad \begin{cases} \Omega_B = \frac{8\pi G\rho_B}{3H_0^2} \sim 0.05\,, \\ \Omega_D = \frac{8\pi G\rho_D}{3H_0^2} \sim 0.25\,. \end{cases}$$

where Ω_M is the contribution of all matter detected by gravitational interactions (galaxy rotation, gravitational lensing, CMB, BAO, etc.), while Ω_B is the contribution of ordinary (visible) matter (the most part is due to baryons (protons, neutrons,...) because electrons, photons, neutrinos ... give negligible contributions and Ω_D is the contribution of *Dark matter*. It could be ordinary matter, but not visible (black holes, clouds, neutrinos), but the most part has to be constituted by WIMP.

The actual values of the energy densities are

$$\begin{aligned}
\rho_c &= \frac{3H_0^2}{8\pi G} \sim 1.88\times\, 10^{-29}\, h^2\, g/cm^3 \sim 9.5\times\, 10^{-30}\, g/cm^3\,, \\
\rho_M &= \rho_c\Omega_M \sim 1.88\times\, 10^{-29}\, \Omega_M\, h^2\, g/cm^3 \sim 2.84\times\, 10^{-30}\, g/cm^3\,, \\
\rho_\Lambda &= \rho_c\Omega_\Lambda \sim 1.88\times\, 10^{-29}\, \Omega_\Lambda\, h^2\, g/cm^3 \sim 6.63\times\, 10^{-30}\, g/cm^3\,,
\end{aligned}$$

and all of them correspond to a proton per cubic meter.

11.3 Redshift parameter relations

As we already said above, the evolution can be parametrised by time, temperature or redshift parameter because there is a "one to one" correspondence between them. Here we explicitly derive such relations.

The temperature/energy determines the reactions which preserve the thermal equilibrium and so it is the natural parameter evolution for theoretical developments, while the redshift parameter is important from the experimental point of view.

11.3.1 Time-redshift relation

In order to get the relation between time and redshift we choose a galaxy (the source) $S \equiv (t, r, 0, 0,) = (\eta, \chi, 0, 0)$ and the observer at the origin $O \equiv (t_0, 0, 0, 0) =\equiv (\eta_0, 0, 0, 0)$. We recall that the redshift parameter z is defined by

$$\frac{\lambda_0}{\lambda} = \frac{a_0}{a(t)} = 1+z \quad\Longrightarrow\quad a(t) = \frac{a_0}{1+z}\,, \qquad a_0 = a(t_0)\,, \tag{11.2}$$

where $z = \Delta\,\lambda/\lambda$ is a measurable quantity related to universal time by a one-to-one correspondence. This means that one can use z as evolution parameter. In fact, deriving the latter equation we have

$$dz = -\frac{a_0\dot{a}(t)}{a^2(t)}\, dt = -\frac{a_0 H(t)}{a(t)}\, dt = -(1+z)H(t)\, dt\,, \tag{11.3}$$

and choosing the integration constant so that $z = \infty$ corresponds to $t = 0$ we obtain

$$t(z) = \int_z^\infty \frac{dz}{(1+z)H(z)}\,. \tag{11.4}$$

For an arbitrary function f we also have the relation

$$\dot{f} = \frac{df}{dz}\,\frac{dz}{dt} = -\frac{a_0 H(z)}{a(z)}\,\frac{df}{dz} \quad\Longrightarrow\quad \dot{f}(t_0) = -H_0\,\frac{df}{dz}\,. \tag{11.5}$$

The Hubble parameter as a function of z can be obtained by the Friedmann equation (9.12), which in terms of z reads

$$H^2(z) + \frac{k(1+z)^2}{a_0^2} = \frac{8\pi G\rho(z)}{3} = H_0^2\Omega_0\,\frac{\rho(z)}{\rho_0}\,, \tag{11.6}$$

$$\rho_0 = \rho(t_0) = \rho(z=0)\,, \qquad \Omega_0 = \frac{\rho_0}{\rho_c} = \frac{8\pi G\rho_0}{3H_0^2}\,. \tag{11.7}$$

Here ρ is the whole energy density (matter, radiation, ...).

For $z = 0$ we get

$$\frac{k}{a_0^2} = (\Omega_0 - 1)H_0^2 \quad\Longrightarrow\quad H(z) = H_0\sqrt{(1-\Omega_0)(1+z)^2 + \frac{\Omega_0\rho(z)}{\rho_0}}\,. \tag{11.8}$$

The energy density as a function of z can be obtained from the continuity equation, that is

$$d\rho = -3(p+\rho)d\log a \quad\Longrightarrow\quad \int_{\rho_0}^{\rho(z)} \frac{d\rho}{\rho + p(\rho)} = 3\log(1+z)\,. \tag{11.9}$$

Ones $\rho(z)$ has been computed, one finds $H(z)$ and finally $t(z)$ by (11.4).

The redshift parameter can also be used in place of the distance, because

$$\chi = \eta_0 - \eta = \int_t^{t_0} \frac{dt'}{a(t')} = \frac{1}{a_0}\int_0^z \frac{du}{H(u)}\,. \tag{11.10}$$

However it has to be noted that when $z \to \infty$, χ tends to the particle horizon χ_P and this means that z can measure distances only for $\chi < \chi_P$ (when $\chi > \chi_P$ the velocity of the source exceeds the speed of light).

In a dust dominated universe $\rho(z) = \rho_0(1+z)^3$ and so

$$H(z) = H_0(1+z)\sqrt{1+\Omega_0\, z}\,. \tag{11.11}$$

If the universe is also flat ($\Omega_0 = 1$) one easily obtains

$$t(z) = \frac{2}{3H_0}\,\frac{1}{(1+z)^{3/2}}\,, \qquad \chi(z) = \frac{2}{a_0 H_0}\left(1 - \frac{1}{\sqrt{1+z}}\right)\,, \tag{11.12}$$

from which the known result $t_0 = 2/3H_0$ follows.

11.3.2 Angular diameter-redshift relation

In a static, Euclidean space, the angle which an object with a given transverse size l subtends on the sky is inversely proportional to the distance to this object. In an expanding universe the relation between the distance and the angular size is not so trivial.

Let us consider an extended object with proper transverse size l at a comoving distance χ, the end points of such an object being $P_1 \equiv (\eta, \chi, \vartheta, \varphi)$ and $P_2 \equiv (\eta, \chi, \vartheta + \Delta\,\vartheta, \varphi)$. The observer will measure the angular size $\Delta\,\vartheta$ which is related to the proper length by l

$$l = a(\eta)\phi_k(\chi)\Delta\,\vartheta \quad\Longrightarrow\quad \Delta\,\vartheta = \frac{l}{a(\eta)\phi_k(\chi)}\,. \tag{11.13}$$

Because photons run on radial geodesics, the time η corresponding to the emission of the photons which reach the observer at time η_0 is given by $\eta = \eta_0 - \chi$ and so

$$\Delta\,\vartheta = \frac{l}{a(\eta_0 - \chi)\phi_k(\chi)}\,. \tag{11.14}$$

When the object is quite near the observer one obtains the Euclidean formula because in such a case

$$\chi \ll \eta_0 \quad \Longrightarrow \quad \phi_k(\chi) \sim \chi \quad \Longrightarrow \quad \Delta\,\vartheta \sim \frac{l}{a(\eta_0)\chi} = \frac{l}{d}\,, \tag{11.15}$$

$d = a(\eta_0)\chi$ being the proper distance.

The relation drastically changes if the object is very far away, that is close to the particle horizon. In such a case

$$a(\eta_0 - \chi) \ll a(\eta_0)\,, \qquad\qquad \phi_k(\chi) \sim \phi_k(\chi_P) = const\,, \tag{11.16}$$

and the angular size $\Delta\,\vartheta$ increases with distance because $a(\eta_0 - \chi)$ is a decrescent function of χ. As it approaches the horizon its image covers the whole sky. Of course, the apparent luminosity drops with increasing distance, otherwise remote objects would completely out-shine nearby ones.

- For example, in a closed $1 + 2$-dimensions universe the spatial section is a 2-sphere and the observer at the north pole will see a given object in the south hemisphere subtending an angle $\Delta\,\vartheta$, which increases as the object approaches the south pole.

In terms of redshift one has

$$\Delta\,\vartheta(z) = \frac{(1+z)\,l}{a_0\phi_k(\chi(z))}\,.$$

In particular in a flat universe dominated by matter

$$\phi_0(\chi) = \chi\,, \qquad \chi(z) = \frac{2}{a_0 H_0}\left(1 - \frac{1}{\sqrt{1+z}}\right)\,,$$

and so

$$\Delta\,\vartheta(z) = \frac{H_0\,l}{2}\,\frac{(1+z)^{3/2}}{\sqrt{1+z}-1}\,.$$

If $z \ll 1$ then $\Delta\,\vartheta \sim \Delta\,l/z$, while for $z \gg 1$, $\Delta\,\vartheta \sim z\Delta\,l$. The minimum value is reached at $z = 5/4$.

If curvature is taken into account then

$$\Delta\,\vartheta(z) = \frac{H_0\,l}{2}\,\frac{\Omega_0^2(1+z)^2}{\Omega_0 z + (\Omega_0 - 2)\left(\sqrt{1+\Omega_0 z} - 1\right)}\,. \tag{11.17}$$

In principle, having standard rulers distributed over a range of redshifts we could use the measurements of angular diameter versus redshift to test different cosmological models.

Of course we have not at disposal true standard rulers distributed in the universe, but we can use characteristic lengths at *recombination* time and tray to use them as rulers in order to extract useful physical informations from cosmic microwave background anisotropies. One of this length is the sound horizon l_S , that is the maximum distance that a sound wave in the baryon-radiation fluid can have propagated until recombination. It is of the order of the Hubble length at recombination, that is

$l_S \sim 1/H(z_r)$, $z_r \sim 1000$ being the value of redshift parameter at recombination time. Such a special ruler subtends an angle $\Delta\vartheta$, which depends on the curvature. The temperature autocorrelation function measures how the microwave background temperature in two directions in the sky differs; this temperature difference depends on the angular separation. The power spectrum is observed to have a series of peaks as the angular separation is varied from large to small scales. The first "acoustic peak" is roughly determined by the sound horizon at recombination. Measuring the angular scale of the first acoustic peak one determines the spatial curvature. Our best evidence that the universe is spatially flat ($\Omega_0 = 1$), as predicted by inflation, comes from this test.

11.3.3 Luminosity-redshift relation

A method of recovering expansion history of universe is with the help of the luminosity-redshift relation.

Consider a source of radiation at $P \equiv (t, r, 0, 0) = (\eta, \chi, 0, 0)$ and let L be its total luminosity, that is the energy released by the source per unit time. Then the total energy released in the interval $\Delta\, t$ is

$$\Delta\, E = L\, \Delta\, t = L\, \frac{\Delta\, t}{\Delta\, \eta}\, \Delta\, \eta = L\, a(t)\, \Delta\, \eta\,.$$

All photons emitted in such an interval are located in a shell of width $\Delta\, \chi = \Delta\, \eta$, with radius growing with time. When the photons reach the observer at t_0, the proper width and the area of the shell read respectively

$$\Delta\, l_0 = a_0 \Delta\, \chi = a_0 \Delta\, \eta\,, \qquad\qquad S_0 = 4\pi a_0^2 \phi_k^2(\chi)\,,$$

and due to the redshift of photons the energy in the shell at t_0 is

$$\Delta\, E_0 = \frac{a(t)}{a_0}\, \Delta\, E = \frac{a(t)}{a_0}\, L\, \Delta\, \eta\,.$$

All photons in the shell are detected by the observer in a time $\Delta\, t_0 = \Delta\, l_0 = a_0 \Delta\, \eta$. This means that the *bolometric flux* F measured by the observer, that is the energy per unit area per unit time, is equal to

$$F = \frac{\Delta\, E_0}{S_0 \Delta\, t_0} = \frac{L}{4\pi \phi_k^2(\chi)}\, \frac{a^2(t)}{a_0^4}\,,$$

and as a function of redshift parameter

$$F(z) = \frac{L}{4\pi a_0^2\, \phi_k^2(\chi(z))(1+z)^2}\,.$$

The bolometric flux is related to the *(bolometric) magnitude* m_{bol} by

$$\begin{aligned} m_{bol} = -2.5 \log_{10} F &= 5 \log_{10}(1+z) + 5 \log_{10} \phi_k(\chi(z)) + const \\ &= \frac{5}{\log(10)} \left[\log(1+z) + \log \phi_k(\chi(z))\right] + const\,, \end{aligned}$$

where $\log_{10} x = \log x / \log(10)$. For $z \ll 1$, using (11.10) we have

$$\begin{aligned} \phi_k(\chi) \sim \chi &= \frac{1}{a_0} \int_0^z \frac{du}{H(u)} \sim \frac{1}{a_0 H_0} \int_0^z \frac{du}{1 + u\, H'(0)/H_0 + O(u^2)} \\ &\sim \frac{1}{a_0 H_0} \int_0^z du \left[1 - \frac{u\, H'(0)}{H_0}\right] \sim \frac{z}{a_0 H_0} \left(1 - \frac{z\, H'(0)}{2H_0} + ...\right)\,, \end{aligned}$$

$$\log \phi_k(\chi) \sim \log z + \log\left(1 - \frac{z\,H'(0)}{2H_0}\right) + const \sim \log z - \frac{z\,H'(0)}{2H_0} + const\,,$$

where H' is the derivative with respect to z. Finally, using (11.5)

$$m_{bol} = 5\log_{10} z + \frac{2.5}{\log(10)}\,(1-q_0)z + const\,.$$

q_0 is the deceleration parameter which can be written as

$$q_0 = \frac{1}{2}\Omega_0\left(1 + \frac{3p_0}{\rho_0}\right) = -\left.\frac{\ddot{a}}{aH^2}\right|_{t=t_0} = -1 - \left.\frac{\dot{H}}{H^2}\right|_{t=t_0} = -1 + \frac{H'(0)}{H_0}\,.$$

In principle, measuring the magnitude it is possible to determine the deceleration parameter q_0 (See the Supernova Cosmology ProjectType–SCM).

11.3.4 Temperature-redshift relation

In an expanding universe the frequency of any wave scales as the inverse of the expansion factor $a(t)$, that is

$$\nu \sim \frac{1}{a(t)} \quad\Longrightarrow\quad \lambda \sim a(t)\,.$$

In order to derive this property we consider a packet of waves with definite wavelength and two observers at distinct points P_1 and P_2. At time t the packet passes the first observer who samples the radiation and measures a wavelength $\lambda(t)$ and subsequently at time $t+\Delta t$ the packet passes the second observer at a proper distance $l = \Delta t$ away from the first. According to (10.2), the second observer moves with respect to the first one at a speed $v = Hl = H(t)\Delta t$ and so the frequency he measures is lowered by the Doppler shift. If P_1 is near P_2 then $v \ll 1$ and wavelength is increased according to non relativistic formula

$$\lambda(t+\Delta t) \sim \lambda(t)(1+v) = \lambda(t)\left[1 + \frac{\dot{a}(t)}{a(t)}\,\Delta t\right]\,.$$

On expanding to first order in Δt we obtain the desired result

$$\frac{\dot{\lambda}}{\lambda} = \frac{\dot{a}}{a} \quad\Longrightarrow\quad \lambda(t) \sim a(t) \quad\Longrightarrow\quad \nu(t) \sim \frac{1}{a(t)}\,. \tag{11.18}$$

With the same method one can show that the momentum p of a particle scales as the frequency, that is $p(t) \sim 1/a(t)$ and of course the de Broglie wavelength scales as $\lambda(t)$. It has to be noted that the scaling law is the same for all frequencies or momenta.

Now we shall derive the scaling law of the equilibrium temperature. We will show that

$$\begin{aligned} T_\gamma(t) &\sim \frac{1}{a(t)}\,, && \text{(radiation and/or relativistic matter);} \\ T_M(t) &\sim \frac{1}{a^2(t)}\,, && \text{(non relativistic matter).} \end{aligned} \tag{11.19}$$

From physics of black-body we know that the occupation number (or mean number per mode) of photons at temperature of equilibrium $T = T_\gamma$ is given by the Planck distribution

$$\langle N\rangle = \frac{1}{e^{\hbar\omega/kT} - 1}\,, \tag{11.20}$$

where k is the Boltzmann constant, while $\hbar$ and $\omega = 2\pi\nu$ are respectively the Planck constant over 2π and the angular frequency. This number depends on temperature only and so, assuming thermodynamic equilibrium and no-interactions with other fields, the temperature has to scale has ν, that is

$$T(t) \sim \frac{1}{a(t)} \quad \Longrightarrow \quad T(t) = \frac{T_0 a_0}{a(t)}\,,$$

and for $T(z)$ we obtain the simple relation

$$T(z) = T_0(1+z)\,. \tag{11.21}$$

Of course this is true if the expansion can be considered as an adiabatic reversible transformation.

The relation (11.21) holds also for an ideal relativistic gas of bosons or/and fermions. In such a case one has the Bose-Einstein/Fermi distributions

$$\langle N \rangle = \frac{1}{e^{(\varepsilon-\mu)/kT} \pm 1}\,,$$

ε, μ being the energy and chemical potential of the particle. The energy is proportional to the momentum p and so the temperature has to scale as p.

For non relativistic matter the temperature scales as p^2 because the occupation number at temperature T_M is given by Boltzmann distribution

$$\langle N_M \rangle \sim e^{\frac{p^2}{2mkT_M}} \quad \Longrightarrow \quad T_M(t) \sim \left(\frac{a_0}{a(t)}\right)^2 T_M(t_0)\,.$$

We see that the expansion of the universe tends to break thermal equilibrium between radiation and non relativistic matter due to different cooling laws, but since the heat capacity of non relativistic matter in comparison with the one of radiation is negligible, the radiation will keep its thermal spectrum.

- Note that in this section k is the Boltzmann constant (do not confuse it with spatial curvature), $h \sim 0.68$ is the dimensionless quantity related to Hubble constant (do not confuse it with Planck constant) and $T = T_\gamma$ is the equilibrium temperature of radiation.

11.4 The cosmic microwave background

Energy density of thermal radiation at temperature $T = T_\gamma$ can be obtained from (11.20) by recalling that the number of modes per unit volume in the solid angle $d\Omega$ with angular frequency in the interval $(\omega, \omega + d\omega)$ is

$$d^3N = 2\,\frac{\omega^2 d\omega\, d\Omega}{(2\pi)^3}\,,$$

where the factor of 2 takes account of the two polarisation states.

Now the integration over the solid angle takes account of photon in all directions, then

$$dn_\gamma(\omega) = \frac{1}{\pi^2}\,\frac{\omega^2 d\omega}{e^{\hbar\omega/kT} - 1}\,.$$

This represents the number of photons per unit volume with angular frequency in $d\omega$ and arbitrary direction. The number density (number of photons per unit volume) is then

$$n_\gamma = \frac{1}{\pi^2}\int_0^\infty \frac{\omega^2 d\omega}{e^{\hbar\omega/kT}-1} = \frac{k^3T^3}{\pi^2\hbar^3}\int_0^\infty \frac{x^2dx}{e^x-1} = \frac{2\zeta(3)}{\pi^2}\left(\frac{kT}{\hbar}\right)^3 , \tag{11.22}$$

where the Riemann $\zeta(s)$ function is given by

$$\zeta(s) = \frac{1}{\Gamma(s)}\int_0^\infty \frac{x^{s-1}dx}{e^x-1}\,, \qquad \mathrm{Re}\, s > 1\,, \qquad \zeta(3)\sim 1.20\,, \qquad \zeta(4)=\frac{\pi^4}{90}\,.$$

The energy of a photon with frequency ω is $\hbar\omega$ and so the energy density due to photons with angular frequency in interval $d\omega$ is given by the well known *Planck blackbody radiation spectrum*

$$u(\omega)d\omega = \frac{\hbar}{\pi^2}\frac{\omega^3 d\omega}{e^{\hbar\omega/kT}-1}\,.$$

Finally, integrating on ω one gets the blackbody radiation energy density

$$\begin{aligned} u &= \int_0^\infty u(\omega)\,d\omega = \frac{\hbar}{\pi^2}\int_0^\infty \frac{\omega^3 d\omega}{e^{\hbar\omega/kT}-1} = \frac{k^4T^4}{\pi^2\hbar^3}\int_0^\infty \frac{x^3dx}{e^x-1} \\ &= \frac{6k^4T^4}{\pi^2\hbar^3}\,\zeta(4) = \frac{\pi^2k^4T^4}{15\hbar^3} = a_BT^4\,, \end{aligned} \tag{11.23}$$

where $a_B = \pi^2k^4/15\hbar^3 \sim 7.56\times\ 10^{-15}\,(erg/cm^3)/^oK^4$.

For the radiation, the heat capacity at fixed volume is

$$C_\gamma = \left.\frac{\partial u}{\partial T}\right|_V = 4a_BT^3\,,$$

while the heat capacity of non relativistic matter does not depend on T. In fact, any degree of freedom contributes with $k/2$. For example, if matter consists of atomic hydrogen then $C_M = 3n_Bk/2$, n_B being the mean number of atoms per unit volume. Because actually the matter contributes to the energy density with a proton per cubic meter we get

$$\frac{C_M}{C_\gamma} = \frac{3n_Bk}{8a_BT^3} \sim 4\times\ 10^{-9}\,\Omega_Mh^2\,.$$

This ratio is independent on the redshift because both $T^3 = T_\gamma^3$ and n_B during the expansion scale as $1/a^3$ or, what is the same as $(1+z)^3$. At high redshift n_B is large and so the interaction between matter and radiation is appreciable. This means that matter relaxes to radiation temperature, because radiation has the higher heat capacity. Moreover, due to the very small ratio C_M/C_γ, the spectrum of radiation remains thermal, no matter how strong the interaction.

In contrast with heat capacities, the ratio between energy densities depends on redshift because matter and radiation densities have a different scale behaviour and so

$$\frac{\rho_M(z)}{u(z)} = \frac{(1+z)^3\rho_M}{a_BT^4} = \frac{4\times\ 10^4\,\Omega_Mh^2}{1+z}\,.$$

As expected, at present time ($z=0$) the energy density due to radiation is a very small fraction of the total one. Moreover, when the redshift is not too large, the energy available from annihilation of mass by nuclear burning or by the process accretion by black holes is sufficient to produce appreciable local perturbation to temperature of radiation.

11.4.1 Characteristic quantities for *CMB*

The entropy of a gas of photons is

$$ds_\gamma = \frac{du_\gamma}{T} = 4a_B T^2 dT \quad \Longrightarrow \quad s_\gamma = \frac{4}{3}\, a_B T^3 = \frac{4\pi^2 k}{45}\left(\frac{kT}{\hbar}\right)^3 .$$

The number n_γ of photons per unit volume goes as T^3 and so the ratio

$$\frac{s_\gamma}{kn_\gamma} = \frac{2\pi^4}{45\zeta(3)} \sim 3.6\,,$$

is independent of T as expected because both s_γ and n_γ are conserved in a reversible adiabatic expansion.

The maximum value of the function $u(\omega)$ is reached for $\omega = \omega_m$ given by means of equation

$$\frac{\hbar\omega_m}{kT} = \frac{2\pi\hbar}{kT\lambda_m} \sim 2.82\,,$$

while the frequency $\omega = \omega_h$ at the half-energy point in the spectrum (the integration of $u(\omega)$ from 0 to ω_h corresponds to half of the total energy density) satisfies

$$\frac{\hbar\omega_h}{kT} = \frac{2\pi\hbar}{kT\lambda_h} \sim 3.50\,.$$

Choosing $T = T_0(1+z)$, $T_0 \sim 2.73\,{}^o\!K$ being the present temperature of *CMB*, it follows

$$\begin{aligned}
\lambda_h &\sim 1.50(1+z)\,mm\,,\\
\varepsilon_h &= \frac{2\pi\hbar}{\lambda_h} \sim 3.50kT \sim 1.32\times\, 10^{-15}(1+z)\,erg\,,\\
&\sim 8.2\times\, 10^{-4}(1+z)\,eV\,.
\end{aligned}$$

ε_h is the energy of the photon with frequency $1/\lambda_h$.

The number of photons and energy density in *CMB* follows from (11.22) and (11.23) respectively and reads

$$n_\gamma \sim 420(1+z)^3/cm^3\,, \qquad\qquad \eta = \frac{n_B}{n_\gamma} \sim 2.7\times\, 10^{-8}\,\Omega_B h^2\,,$$

while the energy density is

$$\begin{aligned}
u &\sim 4.2\times\, 10^{-13}\,(1+z)^4\,erg/cm^3\\
&\sim 0.26\,(1+z)^4\,eV/cm^3\\
&\sim n_\gamma\varepsilon_h\,.
\end{aligned}$$

The number of baryons is negligible with respect to the number of photons and this implies that the universe has an enormous entropy compared to its matter content.

11.5 Relic neutrinos

Neutrinos are relativistic fermions with vanishing or negligible masses. The equilibrium occupation number of a gas of neutrinos at temperature T_ν is

$$\langle N\rangle = \frac{1}{e^{(\varepsilon-\mu)/kT_\nu}+1} \leq 1\,,$$

where ε is the energy and μ the chemical potential. Neutrinos and antineutrinos have opposite chemical potential as a consequence of the fact that they can annihilate, for example through the reaction
$\nu_e + \bar{\nu}_e \to e^+ + e^- \to 2\gamma$. This can be seen in the following way.

The Gibbs function G in a mixture of different particles $\alpha_1, \alpha_2, ...$ with chemical potential $\mu_1, \mu_2, ...$ is

$$G = U + pV - TS = \sum_i n_i \mu_i\,,$$

n_i being the number of particle of type α_i. If such particles are in thermal equilibrium through the reaction

$$\alpha_i + \alpha_j \leftrightarrow \alpha_k\,,$$

then G relaxes to an extremum, that is $\delta G = 0$ (chemical equilibrium), if $\delta n_i = \delta n_j = -\delta n_k$ and this implies that

$$\mu_i + \mu_j - \mu_k = 0\,.$$

In particular, if $\mu_i \equiv \mu_k$ then $\mu_j = 0$. This is just the case of photons, because they can be absorbed by a particle α through the reaction $\alpha + \gamma \leftrightarrow \alpha$, then it follows $\mu_\gamma = 0$.

Now, because neutrinos can annihilates through the reaction described above we get
$\mu_{\nu_e} + \mu_{\bar{\nu}_e} = \mu_{e^+} + \mu_{e^-} = 2\mu_\gamma = 0$. Similar considerations hold also for other species of neutrinos. The difference in the number densities of the pairs $\nu, \bar{\nu}$ is determined by the difference in their chemical potentials (eventual masses are the same).

Since the temperature at which the reactions we are going to consider is very high, the chemical potentials (eventually the masses) can be neglected in comparison to the energy kT. Then we can put $\mu = 0$ and $p = \varepsilon$. In this way the couple of particles $\nu, \bar{\nu}$ gives the following contribution to number and energy density:

$$\begin{aligned} n_\nu &= \frac{2}{(2\pi\hbar)^3}\int_0^\infty \frac{4\pi p^2\,dp}{e^{p/kT_\nu}+1} = \frac{3\zeta(3)}{2}\left(\frac{kT_\nu}{\pi^2\hbar}\right)^3\,, \\ u_\nu &= \frac{2}{(2\pi\hbar)^3}\int_0^\infty \frac{4\pi p^3\,dp}{e^{p/kT_\nu}+1} = \frac{7}{8}\,a_B T_\nu^4\,. \end{aligned}$$

The latter relations are a direct consequence of the identity

$$\frac{1}{e^x+1} = \frac{1}{e^x-1} - \frac{2}{e^{2x}-1} \quad\Longrightarrow\quad \int_0^\infty \frac{x^{s-1}dx}{e^x+1} = \Gamma(s)\,(1-2^{s-1})\,\zeta(s)\,.$$

In a plasma at equilibrium, the contribution to the energy density of a family of neutrinos $(\nu, \bar{\nu})$ is lower by a factor 7/8 with respect to the contribution of a photon. However it has to be remember that there are three different species of neutrinos, that is $(\nu_e, \bar{\nu}_e)$, $(\nu_\mu, \bar{\nu}_\mu)$, $(\nu_\tau, \bar{\nu}_\tau)$, and so the total contribution of neutrinos families is $u_\nu = (21/8)u_\gamma$.

What we have derived for neutrinos is also valid for a gas of relativistic electrons-positrons pairs, the only difference is due to the fact that electrons and positrons have both two spin states and so energy and number densities are doubled. For electrons-positrons gas

$$n_e = 3\zeta(3) \left(\frac{kT_e}{\pi^2\hbar}\right)^3 , \qquad u_e = \frac{7}{4}\, a_B T_e^4 \,.$$

As we shall see below, neutrinos ν_e decouple at energies of the order of $5\, m_e$, m_e being the electron mass, while neutrinos ν_μ, ν_τ decouple at energies a little bit higher. Neutrinos ν_μ, ν_τ thermalise through elastic scattering with electrons and positrons, which at these energy are present in a huge number, while the other massive leptons have been already annihilated or decayed and so only few of them can be present.

This means that at energies higher than m_e, all species of neutrinos, electrons and radiation are all in thermal equilibrium at temperature $T = T_\gamma = T_\nu = T_e$. The universe expands, the gas freezes and so neutrinos decouple. They stop to interact with the gas and their temperature $T_\nu(z)$ decreases according to the expansion law of relativistc particles, while the temperature of radiation rapidly increases when electrons-positrons pairs annihilate and release their energy to the microwave background.

After the decoupling of neutrinos but before the annihilation of electron-positron pairs, that is at temperature of the order $kT > 2m_e$, the reaction $\gamma + \gamma \leftrightarrow e^+ + e^-$ produces a sea of electron-positron pairs at the temperature $T = T_\gamma$. The energy density of such a gas of photons and electron-positron pairs is given by

$$u = u_\gamma + u_e = \left(1 + \frac{7}{4}\right) u_\gamma = \frac{11}{4}\, a_B T^4 \,,$$

and the total entropy density of such a gas reads

$$s = s_\gamma + s_e = \left(1 + \frac{7}{4}\right) s_\gamma = \frac{11}{3}\, a_B T^3 \,.$$

Because the annihilation process is reversible, the total entropy in a given comoving volume is conserved.

Now we indicate by T_a, V_a the temperature and the given volume of the gas after annihilation and by T_b, V_b the same quantities before annihilation (V_a and V_b are the same co-moving volume). Before annihilation we have a gas of photons, electrons and positrons whit entropy $S_b = (11/3)\, a_B V_b T_b^3$, while after annihilation we have a gas of photons only, with entropy $S_a = (4/3)\, a_B V_a T_a^3$. Since entropy is conserved we get

$$S_b = \frac{11}{3}\, a_B V_b T_b^3 = S_a = \frac{4}{3}\, a_B T_a^3 V \quad \Longrightarrow \quad V_b T_b^3 = \frac{4}{11}\, V_a T_a^3 \,.$$

Immediately before annihilation of electron-positron pairs, neutrinos have a temperature $T_\nu^b = T_b$ equal to the one of photons, but since they are already decoupled, they evolve independently of photons and their entropy is conserved separately. For entropy of neutrinos we get

$$V_b\,(T_\nu^b)^3 = V_b T_b^3 = V_a\,(T_\nu^a)^3 \quad \Longrightarrow \quad T_\nu^a = \left(\frac{V_b}{V_a}\right)^{1/3} T_b = \left(\frac{4}{11}\right)^{1/3} T_a \,.$$

Summarising we have that after all possible interactions, the temperature of photons and neutrinos are related by

$$T_\gamma(z) = (1+z) T_0 \,, \qquad T_\nu(z) = \left(\frac{4}{11}\right)^{1/3} T_\gamma(z) = \left(\frac{4}{11}\right)^{1/3} (1+z) T_0 \,,$$

where $T_0 \sim 2.73^oK$ is the actual value of temperature of black body radiation in *CMB*. The corresponding actual value for temperature of neutrinos is $T_\nu = T_\nu(0) \sim 1.95^oK$.

The number of neutrinos in a single family after electron-positron pairs annihilation is

$$n_\nu = \frac{3\zeta(3)}{2}\left(\frac{kT_\nu}{\pi^2\hbar}\right)^3 = \frac{6\zeta(3)}{11}\left(\frac{kT_\gamma}{\pi^2\hbar}\right)^3 = \frac{3}{11}\,n_\gamma\,.$$

At present time we obtain $n_\nu \sim 113/cm^3$.

Taking into account of three species of neutrinos the total number of pairs is $n_\nu^{tot} \sim 339/cm^3$ (at present time) and the total energy density of relativistic matter (radiation and neutrinos) becomes

$$u_r = \left[1 + \frac{21}{8}\left(\frac{4}{11}\right)^{4/3}\right] a_B T^4 \sim 1.68\,a_B T^4\,. \tag{11.24}$$

The corresponding Ω_r parameter is

$$\Omega_r = \frac{u_r}{\rho_c} = \frac{8\pi G u_r}{3a_0^2 H_0^2} \sim 4.22\times\,10^{-5}h^2\,.$$

For non-relativistic matter we have $\Omega_M \sim 0.3$ ($\Omega_B \sim 0.04$) and because Ω_M/Ω_r varies as $1/(1+z)$ we get that for a suitable value of z, say $z = z_{eq}$, the mass-radiation densities are equivalent, that is

$$\frac{\Omega_M}{\Omega_r} = 1 \quad\Longrightarrow\quad 1 + z_{eq} \sim 2.37\times\,10^4\,\Omega_M h^2\,.$$

The matter dominated era starts at $z \sim z_{eq}$.

11.5.1 Thermal decoupling of neutrinos

Now we show that $\nu_e, \bar\nu_e$ effectively decoupled at energies higher than m_e. Near decoupling at temperature T, the main thermalisation reaction is

$$\nu + \bar\nu \leftrightarrow e^+ + e^-\,,$$

which has a cross section (G_F is the weak interaction or Fermi constant)

$$\sigma \sim G_F^2 E_\nu^2 \sim 4\times\,10^{-44}\,T_{10}^2\,cm^2\,, \tag{11.25}$$

where for convenience we have introduced the dimensionless quantity T_{10} by

$$T = \frac{T_{10}}{10^{10}}\,{}^oK \quad\Longrightarrow\quad T_{10} = 10^{10}\,T/{}^oK\,.$$

Number density and collision time are respectively given by

$$n_\nu = 1.6\times\,10^{31}\,T_{10}^3\,cm^{-3}\,, \qquad\qquad t_c \sim \frac{1}{\sigma\,n_\nu c}\,.$$

Recall that t_c is the time necessary to have a reaction of the one considered. During a period t the number of reactions will be

$$N_{react} \sim \frac{t}{t_c} \sim \sigma\,n_\nu c\,t\,. \tag{11.26}$$

Of course, if this number is large then neutrinos are is thermal equilibrium with electrons/radiation, but when this number is small thermal equilibrium is broken and neutrinos become free. In (11.26) we can take t equal to the age of the universe at the time of reaction, but also in this case it is clear that at some time during the expansion N_{react} will become smaller that 1, because it would depend on T^3, which scales as $1/a^3$ and becomes smaller and smaller.

The age of the universe can be computed using (10.16) because it is still dominated by relativistic matter, so

$$t = \sqrt{\frac{3}{32\pi G u_r}} \sim \frac{2}{T_{10}^2}\, sec\,.$$

Using this value in (11.26) we get

$$N_{react} \sim 0.04\, T_{10}^3\,,$$

which is of order ~ 1 for $T_{10} \sim 3$ which corresponds to $kT/m_e \sim 5$.

After annihilation of electron-positron pairs neutrinos are practically free. They can interact with protons and neutrons or perform elastic scattering with the few survived massive leptons, but temperature does not change. The same happens for photons.

11.6 Primordial nucleosynthesis

Here we simply report the main important steps which predict the observed abundance of light elements in primordial nebulae, which are essentially made up by hydrogen and helium in the ratio 3 : 1 in mass and traces of other light elements (deuterium, lithium,...).

Primordial nucleosynthesis is considered one of the crucial pieces in favour of the SCM, since by the study of nuclear processes in the background of an expanding cooling universe it yields a remarkable concordance between theory and experiment. It begins at energies of the order of $100\, KeV$ and lower, but the nuclear reactions which produced the abundances of neutron and protons started at higher energies.

During the lepton era (below 100 MeV), electrons, neutrinos, photons and nucleons are in equilibrium. The number density of neutrons n_N is different with respect to the one of protons n_P essentially due to the fact that they have different masses. If fact one has

$$\frac{n_N}{n_P} = e^{-Q/T}\,, \qquad\qquad Q = (m_N - m_P) - (\mu_N - \mu_P) = (m_N - m_P) - (\mu_e - \mu_{\nu_e})\,,$$

m_k, μ_k being respectively mass and chemical potential of the particle considered. Neglecting chemical potentials one gets $Q \sim 1.3\, MeV$.

If temperatures are higher than few MeV, neutron and protons inter-convert principally through the weak interactions

$$N + \nu_e \leftrightarrow P + e^-\,, \qquad\qquad N + e^+ \leftrightarrow P + \bar{\nu}_e\,,$$

but at temperatures below $1\, MeV$, the weak interactions are frozen out, electron-positron pairs are annihilated and neutrons and protons cease to inter-convert. The ratio n_N/n_P is about 1/6, but it will increase since neutrons decays into protons by $n \to p + e^- + \bar{\nu}$. The age of the universe at this epoch is $t \sim 1\, sec$ ($T \sim 1\, MeV$), while neutrons have a lifetime $\tau_n \sim 890\, sec \gg t$, then they start gradually to decay until the temperature reaches a value below $100\, KeV$, when primordial nucleosynthesis begins. At this time the ratio n_N/n_P is approximatively 1/7.

Of all the light nuclei (deuterium D, tritium T, helium ^3He, ^4He, lithium ^7Li, beryllium ^7Be, etc...), the most favorable from the energetically point of view is ^4He. Then we expect that almost all neutrons and an equal number of protons are converted into (ionised) atoms of ^4He. This means that

$$\frac{n_N}{n_P} = \frac{1}{7} = \frac{2}{14} \quad \Longrightarrow \quad \frac{n_{He}}{n_P} = \frac{1}{12}\,,$$

where n_{He} is the number of atoms of ^4He. As we see the process ends up with twelve (ionised) atoms of H per ^4He, which in mass corresponds to a ratio of 3 : 1, as observed.

- We have described nucleosynthesis in a very simple way, but of course the situation is more complicated. First of all, the number density of nucleons is not large enough to produce directly ^4He through a four nucleons interaction. The process will proceed by steps as in figure (12). The final percentages in mass are ~ 0.75 of H, ~ 0.25 of ^4He and traces amounts of other light elements ($D \sim 10^{-5}$, $^3He \sim 10^{-5}$, $^{17}Li \sim 10^{-10}$ atoms per proton).

  Heavier elements are not synthesized in this period but in the later universe during supernovae explosions.

- It has to be observed that, as one might expect, nucleosynthesis do not start at the nuclear binding energy (of the order of $1\,MeV$), because there is a huge number of photons per nucleon and this prevents the process to taking place until energy drops below $100\,KeV$.

- All predictions concerning the abundances of light elements are in good agreement with observations, but at the same time they increase our confidence in the *SM*, because all computations are done by using general relativity, thermodynamics and particle physics.

  For example, if the number of neutrinos species is larger than 3, then at a fixed temperature, energy density of radiation (relativistic particles) will increase and this will give a corresponding increment of neutrons per proton. As a consequence also the abundance of helium will increase in contrast with observation. This means that the number of neutrino species has to be 3, has predicted by standard model of particle physics.

  The abundances of the light elements depend essentially on just one free parameter η, that is the baryon/antibaryon to entropy ratio

  $$\eta = \frac{n_B}{s} = \frac{n_b - n_{\bar{b}}}{s}\,,$$

  where $n_b, n_{\bar{b}}$ are the numbers of baryons and antibaryon per unit volume, while s is the total entropy density. η is independent on the expansion because *baryon asymmetry* n_B as well as entropy density scales as $1/a^3$ and moreover the baryonic number is conserved. In the early universe dominated by radiation, using only general relativity and *SM* physics one can give a strict constraint on the value of η. In fact, the range of η consistent with deuterium and ^3He primordial abundances is

  $$2.6\times\,10^{-10} < \eta < 6.2\times\,10^{-10}\,.$$

  The value of such a parameter has been recently extracted from *CMB* from precise measurements of the relative heights of the first two acoustic peaks. It reads

  $$\eta = 6.1\times\,10^{-10}\begin{cases} +0.3\times\,10^{-10} \\ -0.2\times\,10^{-10} \end{cases}$$

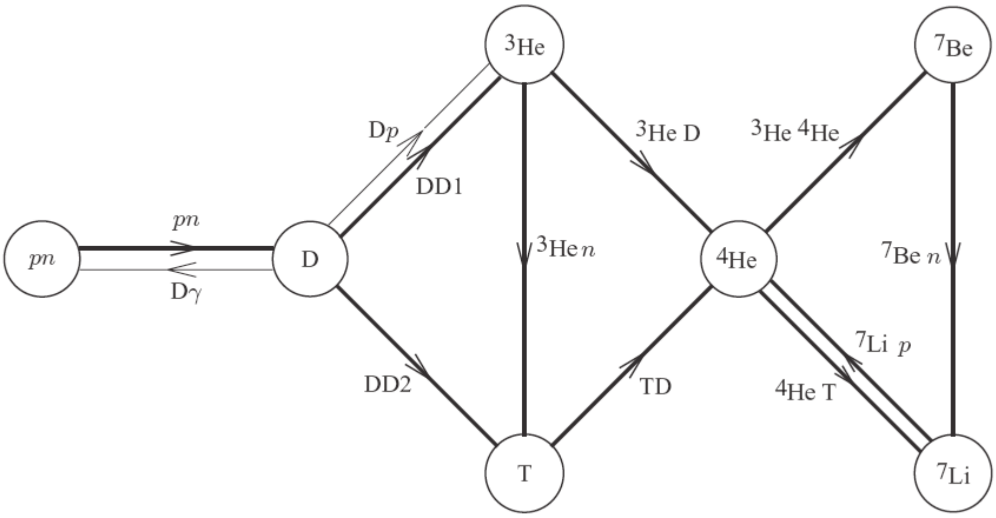

Figure 12: primordial nucleosynthesis (figure by V. Mukhanov)

in astonishing agreement with the predicted theoretical result.

The fact that η is not vanishing means that there is an asymmetry between matter and antimatter and this gives rise to a serious theoretical problem because, in SM physics there are no difference between particles and antiparticles as confirmed in collider experiment. For this reason one expects that *baryogenesis* creates an equal numbers of baryons and antibaryons, but of course, during the evolution, some still unknown process had broken this symmetry. I any case, it is a matter of fact that primordial antimatter is not present in our universe. Possible explanations of this can be found in (discrete) symmetry violations (baryon number B, charge conjugation C and parity CP) or alternatively on a departure from thermal equilibrium.

11.7 Dark matter and dark energy

According to recent experimental data the spatial curvature k is vanishing with high precision (in agreement with the prediction of inflationary models) and, in contrast with as expected from General Relativity, the universe is now in an expanding, accelerated phase (the deceleration parameter q_0 is negative). The presence of an *effective, positive cosmological constant* is then unavoidable. Setting $\Omega_0 = \Omega_k + \Omega_M + \Omega_\Lambda$, $p_0 \sim 0$ and disregarding radiation contribution (actually it is negligible), the Friedmann equation (11.1) becomes

$$\Omega_0 = \Omega_k + \Omega_M + \Omega_\Lambda = 1\,,$$

where

$$\Omega_k = -\frac{k}{a_0^2 H_0^2}\,, \qquad \Omega_M = \frac{8\pi G\rho_M}{3H_0^2} \equiv \Omega_B + \Omega_D\,, \qquad \Omega_\Lambda = \frac{\Lambda}{3H_0^2}\,,$$

Ω_k, Ω_B, Ω_D and Ω_Λ being respectively the contributions due to curvature, baryonic matter, dark matter and dark energy evaluated at present time t_0.

The reason for splitting the matter contribution into two parts derives from the fact that only a very small part of Ω_M is due to "ordinary matter", that is the one we are dealing with every day and which is essentially contained in stars (proton, neutron, electron; for simplicity one refers to that as *baryonic matter*), while the larger part is due to particles which we "feel" only through gravitational interaction (one refers to that as *dark matter* because it is "electromagnetically invisible"). Some small contribution to dark matter is due to ordinary but invisible matter (black holes, clouds,...), but the most part is due to particle which interacts only by gravitational and weak interactions. Such kind of particles are called *WIMP* (weakly interacting massive particles as neutralinos, axions,...). Particles of this kind do not carry electric charge and so they do not interact electromagnetically and do not enter in the formation of stars (could dark matter, *CDM*).

All recent measures are in good agreement with the so called ΛCDM model, which is based on the Friedmann solution with cosmological constant Λ, and with the following values for the parameters

$$\Omega_k \sim 0\,, \qquad \Omega_M \sim 0.03\,, \qquad \Omega_B \sim 0.05\,, \qquad \Omega_D \sim 0.25,\,, \qquad \Omega_\Lambda \sim 0.69\,.$$

We see that the most contribution to Ω_0 is due to dark matter and dark energy.

Dark energy could be generated by a cosmological constant, but it could be the energy of unknown cosmological fields or finally it could be an effect due to a modified theory of gravity. In any case, because all data are in good agreement with ΛCDM model, the equation of state of dark energy has to be of the kind $p_\Lambda \sim -\rho_\Lambda$, that is $w \sim -1$.

In the ΛCDM model the cosmological constant is fixed 'by hand' and could be for example the vacuum energy density of all matter fields, while in other models which considers additional cosmological fields like quintessence ($-1 < w < 1/3$) or phantom ($w < -1$), the dark energy is a time-dependent contribution related to the energy of such kind of fields. Finally, in modified theories of gravity, dark energy is a dynamic quantity which emerges from gravitation.

As we have said above, all data are in good agreement with ΛCDM model, but in principle one could analyse the data by using a different model and in such a case the ratios between the different kind of matter/energy could change.

11.8 The cosmological constant problem

Quantum field theory provides a non-vanishing vacuum energy density, due to all matter fields contribution. In the absence of gravitation such an energy can be neglected because it is a constant, but when gravity is taken into account it has to be carefully considered because it contributes to gravitation. It is natural to relate the cosmological constant to these vacuum energy of quantum fields.

Unfortunately, within the framework of quantum field theory, the evaluation of zero-point energy gives a divergent value and, even if one chooses an ultraviolet cutoff, as it is reasonable, one obtains a very huge value with respect to the observed one ρ_{obs}, in fact

$$\frac{\rho_{vac}}{\rho_{obs}} \sim 10^{55}\,, \qquad \text{cutoff at electroweak scale,}$$
$$\frac{\rho_{vac}}{\rho_{obs}} \sim 10^{123}\,, \qquad \text{cutoff at Planck scale.}$$

Dark energy represents about the 70% of the whole energy density in the universe, but nevertheless it is really small with respect to the vacuum energy provided by quantum field theory. This discrepancy between theory and observation is the content of the old and unsolved *cosmological constant problem*.

As a consequence of an unknown mechanism or maybe unknown matter fields, vacuum energy has to be vanish, but a very small part $\rho_{vac} \to \rho_\Lambda > 0$. From this point of view, $\Lambda = 0$ seams to be more reasonable and less problematic.

There is also an additional puzzle related to the cosmological constant. This is called *the coincidence problem* and consists in the fact that the actual values of ρ_Λ and ρ_M are of the same order of magnitude. The ratio between the contribution to energy density of cosmological constant and matter changes rapidly as the universe expands, that is $\rho_\Lambda(t)/\rho_M(t) \sim a^3(t)$ and this means that $\rho_\Lambda(t)/\rho_M(t) \sim 1$ only for a very brief period, just the one in which we live.

12 Perturbations of metric and energy density

Here we assume the energy density at initial time t_i (after inflation) to be "nearly but not perfectly" homogeneous and then we are going to study what kind of consequences this hypothesis will have on the *CMB* spectrum. We shall see that small inhomogeneities in the initial energy density will reflect as anisotropies in *CMB*.

The first problem consists in the "gauge invariant" classification of perturbations. In fact, energy density depends on the reference frame. By performing a transformation of coordinates, a homogeneous distribution could become inhomogeneous and viceversa, a inhomogeneous distribution could become homogeneous. Of course we have to take into account both perturbations of the metric and of the energy distribution, because they are strictly related.

The problem of pertubations in general relativity and cosmology is a very difficult mathematical task by itself and its connection to *CMB* is also more difficult to treat in detail and is out of the aim of present lectures notes. Here we shall limit our analysis to the classification of cosmological perturbations, independently on their origin, and in Section 13 we shall derive the important Sachs-Wolfe effect, which is due to scalar perturbations of the metric,

12.1 Classification of perturbations

We start with perturbations of the Friedmann metric and for simplicity we consider a flat universe only, the extension to the curved case being quite straightforward. We indicate by δg_{ij} a small perturbation of the metric $\hat{g}_{ij}$, which satisfies Friedmann equations with $k = 0$. Using conformal time and Euclidean coordinates for the spatial section, one has

$$d\hat{s}^2 = \hat{g}_{ij}\, dx^i dx^j = a^2(\eta)\left(-d\eta^2 + \delta_{ab}\, dx^a dx^b\right), \qquad a, b, = 1, 2, 3, \tag{12.1}$$

while for the whole metric $g_{ij} = \hat{g}_{ij} + \delta g_{ij}$

$$ds^2 = g_{ij}\, dx^i dx^j = (\hat{g}_{ij} + \delta g_{ij})\, dx^i dx^j, \qquad \delta g_{ij} \ll \hat{g}_{ij}.$$

Since at any fixed time the spatial background is homogeneous and isotropic, the perturbation can be classified into three distinct types, according to their behaviour with respect to the group of spatial rotation. In fact one has scalar, vector and tensor perturbations.

It is convenient to write the component of the perturbation in the form

$$\begin{aligned} \delta g_{00} &= 2a^2\, \underline{\phi}, \\ \delta g_{0a} &= a^2\, (\partial_a \underline{B} + \underline{S}_a), \\ \delta g_{ab} &= a^2\, (2\underline{\psi}\, \delta_{ab} + 2\partial_a\partial_b \underline{E} + \partial_a \underline{F}_b + \partial_b \underline{F}_a + \underline{h}_{ab}), \end{aligned} \tag{12.2}$$

where all "underlined" quantities are scalars, vectors or tensors with respect to spatial rotations and indices of such quantities are rised or lowered with the Euclidean metric δ_{ab} (there are no difference between covariant or controvariant underlined tensors). With the use of such a notations we do not confuse spatial 3-tensors with spatial parts of 4-tensors. As usual vector and tensors are split in their invariant components (transverse, traceless) and so the following relations hold:

$$\partial_a \underline{S}^a = 0, \qquad \partial_a \underline{F}^a = 0, \qquad \partial_a \underline{h}^{ab} = 0, \qquad \underline{h}^a_a = 0, \qquad \underline{h}_{ab} = \underline{h}_{ba}.$$

Note that in all above expressions we have ordinary partial derivatives because the spatial section of the manifold is flat and we are using Euclidean coordinates.

- We see that scalar perturbations are characterized by the four scalar functions $\underline{\phi}, \underline{\psi}, \underline{B}, \underline{E}$. They are induced by the inhomogeneities of the energy density and are the most important because they exhibit gravitational instability and may lead to the formation of structure in the universe.

- Vector perturbations are described by the two vectors $\underline{S}_a, \underline{F}_a$ and are related to the rotational motions of the fluid. They decay very quickly and are not really important from the point of view of cosmology.

- Tensor perturbations $\underline{h}_{ab}$ have no analog in Newtonian theory. They describe gravitational waves (in the *TT-gauge*), which are the degrees of freedom of the gravitational field itself. In the linear approximation they do not induce any perturbations in the perfect fluid.

12.2 Gauge transformations

By taking into account of all constraints on the quantities $\underline{\phi}, \underline{\psi}, \underline{B}, \underline{E}, \underline{S}_a, \underline{F}_a, h_{ab}$ of course one gets 10 independent functions as the components of g_{ij}, but we know that the arbitrariness in the choice of coordinate system permits to fix other 4 conditions between the components of g_{ij}.

We recall that under the infinitesimal transformation $\tilde{x}^k = x^k + \xi^k(x)$, neglecting higher orders in ξ^k, the metric transforms as (see Section 6.5)

$$\tilde{g}_{ij}(\tilde{x}) \sim g_{ij}(x) - \partial_i \xi_j(x) - \partial_j \xi_i(x) , \qquad |\xi^k| \ll 1 , \tag{12.3}$$

where ∂_k is the derivative with respect to x^k. If computed at the same point, $\tilde{g}_{ij}(x)$ and $g_{ij}(x)$ are related by

$$\tilde{g}_{ij}(x) \sim g_{ij}(x) - \nabla_i \xi_j(x) - \nabla_j \xi_i(x) , \tag{12.4}$$

∇_k being the covariant derivative related to g_{ij}. Now we split both the metrics $g_{ij}(x)$ and $\tilde{g}_{ij}(x)$ into background and perturbation parts, that is

$$g_{ij}(x) = \hat{g}_{ij}(x) + \delta g_{ij}(x) , \qquad \tilde{g}_{ij}(x) = \hat{g}_{ij}(x) + \delta \tilde{g}_{ij}(x) , \tag{12.5}$$

where $\hat{g}_{ij}$ is the metric considered above, which satisfies Friedmann equations with $k = 0$ and homogeneous density distribution. Comparing equations (12.3)-(12.5) and neglecting higher order terms (δg_{ij} as well as ξ^k are infinitesimal quantities) we obtain

$$\begin{aligned} \tilde{g}_{ij}(x) &= \hat{g}_{ij}(x) + \delta \tilde{g}_{ij}(x) \sim g_{ij}(x) - \nabla_i \xi_j(x) - \nabla_j \xi_i(x) \\ &\sim \hat{g}_{ij}(x) + \delta g_{ij}(x) - \hat{\nabla}_i \xi_j(x) - \hat{\nabla}_j \xi_i(x) , \end{aligned} \tag{12.6}$$

and finally

$$\delta \tilde{g}_{ij} \sim \delta g_{ij} - \hat{\nabla}_i \xi_j(x) - \hat{\nabla}_j \xi_i(x) .$$

The latter equation represents the variation in form of the perturbation δg_{ij} as a consequence of an infinitesimal transformation of coordinates. All quantities have to be evaluated at the same point. We have replaced ∇_k with the covariant derivative $\hat{\nabla}_k$ related to $\hat{g}_{ij}$, because they differ for an infinitesimal quantity.

Under the particular infinitesimal transformation $\tilde{x}^k = x^k + \xi^k$, a scalar function f transform as

$$\tilde{f}(\tilde{x}) = f(x) \quad \Longrightarrow \quad \tilde{f}(x) \sim f(x) - \xi^j(x) \partial_j f(x) ,$$

and as above, if we put $f(x)=\hat{f}(x)+\delta f$, $\tilde{f}(x)=\hat{f}(x)+\delta\tilde{f}$. up to higher order terms we get

$$\delta\tilde{f}(x)\sim\delta f(x)-\xi^j(x)\partial_j f(x)\,. \tag{12.7}$$

In a similar way we get the trasformation rules for vectors V^k and V_k. From

$$\begin{aligned}\tilde{V}^k(\tilde{x}) &= \frac{\partial\tilde{x}^k}{\partial x^j}V^j(x)=V^k(x)+V^j(x)\partial_j\xi^k\\ \tilde{V}_k(\tilde{x}) &= \frac{\partial x^j}{\partial\tilde{x}^k}V_j(x)\sim V_k(x)-V_j(x)\partial_k\xi^j\end{aligned}$$

it follows

$$\begin{aligned}\tilde{V}^k(x) &\sim V^k(x)+V^j(x)\partial_j\xi^k-\xi^j(x)\partial_jV^k(x)\,,\\ \tilde{V}_k(x) &\sim V_k(x)-V^j(x)\partial_k\xi^j-\xi^j(x)\partial_jV_k(x)\,,\end{aligned}$$

and finally

$$\delta\tilde{V}^k(x) \sim \delta V^k(x)+V^j(x)\partial_j\xi^k-\xi^j(x)\partial_jV^k(x)\,, \tag{12.8}$$

$$\delta\tilde{V}_k(x) \sim \delta V_k(x)-V_j(x)\partial_k\xi^j-\xi^j(x)\partial_jV_k(x)\,. \tag{12.9}$$

In the folfowing we shall need also the transformation rules for a generic mixed tensor T^i_j. We have

$$\tilde{T}^i_j(\tilde{x})=\frac{\partial\tilde{x}^i}{\partial x^r}\frac{\partial x^j}{\partial\tilde{x}^s}T^r_s\sim T^i_j(x)+T^k_j\,\partial_k\xi^i-T^i_k\,\partial_j\xi^k\,,$$

and so

$$\delta\tilde{T}^i_j(x)\sim\delta T^i_j(x)+T^k_j\,\partial_k\xi^i-T^i_k\,\partial_j\xi^k-\xi^k\partial_kT^i_j\,, \tag{12.10}$$

Using (12.1) one easily computes the non vanishing components of the corresponding connection. They read

$$\hat{\Gamma}^0_{ij}=\frac{a'(\eta)}{a(\eta)}\,\delta_{ij}\,,\qquad\qquad \hat{\Gamma}^i_{0j}=\frac{a'(\eta)}{a(\eta)}\,\delta^i_j\,,$$

where the prime is the derivatives with respect to conformal time η. Recalling (12.2) one also gets

$$\begin{aligned}\delta\tilde{g}_{00} &\sim \delta g_{00}-2\nabla_0\xi_0=\delta g_{00}-2\partial_0\xi_0+\frac{2a'}{a}\,\xi_0=\delta g_{00}+2a\,\partial_\eta(a\xi^0)\,,\\ \delta\tilde{g}_{0a} &\sim \delta g_{0a}-\nabla_0\xi_a-\nabla_a\xi_0=\delta g_{0a}-\partial_0\xi_a-\partial_a\xi_0+\frac{2a'}{a}\,\xi_a\\ &\qquad\qquad = \delta g_{0a}+a^2\left[\partial_a(\xi^0-\underline{\zeta}')-\underline{\xi}'_a\right]\,,\\ \delta\tilde{g}_{ab} &\sim \delta g_{ab}-\nabla_a\xi_b-\nabla_b\xi_a=\delta g_{ab}-\partial_a\xi_b-\partial_b\xi_a+\frac{2a'}{a}\,\delta_{ab}\xi_0\\ &\qquad\qquad = \delta g_{ab}-a^2\left[\frac{2a'}{a}\,\xi^0\delta_{ab}+2\partial_a\partial_b\underline{\zeta}+\partial_a\underline{\xi}_b+\partial_b\underline{\xi}_a\right]\,,\end{aligned} \tag{12.11}$$

where the scalar $\underline{\zeta}$ and the 3-vector $\underline{\xi}^a$ are related to the spatial part of $\xi^k\equiv(\xi^0,\xi^a)$ by

$$\xi_i=g_{ij}\xi^j=a^2\,\eta_{ij}\xi^j\quad\Longrightarrow\quad \xi_0=-a^2\xi^0\,,\qquad\qquad \xi_a=a^2\,\delta_{ab}\xi^b\,,$$

$$\xi^a=\underline{\xi}^a+\delta^{ab}\partial_b\underline{\zeta}\,,\qquad\qquad \partial_a\underline{\xi}^a=0\,,\qquad\qquad \underline{\xi}_a=\delta_{ab}\underline{\xi}^b\,,$$

Now we study separately scalar, vector and tensor perturbations.

12.3 Scalar perturbations

By taking only scalar perturbations into account the metric assumes the form

$$ds^2 = a^2(\eta)\left[-(1-2\underline{\phi})\, d\eta^2 + 2\partial_a \underline{B}\, dx^a\, d\eta + \left([1+2\underline{\psi}]\delta_{ab} + 2\partial_a\partial_b \underline{E}\right) dx^a dx^b\right], \tag{12.12}$$

but the four scalars function are not independent. In fact, comparing (12.2) and (12.11) and considering only the scalar components of perturbation we get

$$\underline{\tilde{\phi}} = \underline{\phi} + \frac{1}{a}\,\partial_\eta(a\xi^0)\,, \qquad \underline{\tilde{B}} = \underline{B} + \xi^0 - \underline{\zeta}'\,, \qquad \underline{\tilde{\psi}} = \underline{\psi} - \frac{a'}{a}\,\xi^0\,, \qquad \underline{\tilde{E}} = \underline{E} - \underline{\zeta}\,, \tag{12.13}$$

and so we can use the arbitrariness of ξ^0 and $\underline{\zeta}$ to fix two constraints between the scalar functions in the metric. The space of scalar pertubations has two dimensions and can be span by choosing

$$\Phi = \underline{\phi} - \frac{1}{a}\,\partial_\eta(a\,\underline{\Omega})\,, \qquad \Psi = \underline{\psi} + \frac{a'}{a}\,\underline{\Omega}\,, \qquad \underline{\Omega} = \underline{B} - \underline{E}'\,,$$

which are invariant functions under gauge transformations. When we are dealing with "fictious" perturbations simply due to the choice of coordinates, both previous functions will vanish, but they will be different from zero in the presence of "physical" perturbations. In this way we are able to distinguish between fictious and physical perturbations.

A gauge invariant formulation has to be done also for the energy-momentum tensor. We indicate by $\rho(x) = \hat{\rho}(\eta) + \delta\rho$ the perturbed energy density, $\hat{\rho}(\eta)$ being the homogeneous unperturbed one satisfying the Friedmann equations with the metric $\hat{g}_{ij}$ and by $u^k(x) = \hat{u}^k(x) + \delta u^k$ the four velocity of the fluid. Here $\hat{u}^k \equiv (1/a, 0, 0, 0)$ ($\hat{u}_k \equiv (-a, 0, 0, 0)$) is the velocity of the homogeneous perfect fluid (recall that $-g_{00} = a(\eta)^2$).

By definition ρ is a scalar quantity and so, using (12.7) and (12.13) we obtain

$$\delta\tilde{\rho} = \delta\rho - \xi^k\partial_k\,\hat{\rho} = \delta\rho - \xi^0\hat{\rho}\,' = \delta\rho + (\underline{\tilde{\Omega}} - \underline{\Omega})\hat{\rho}\,'\,.$$

We see that the difference

$$\overline{\delta\rho} = \delta\rho - \underline{\Omega}\hat{\rho}\,' = \delta\tilde{\rho} - \underline{\tilde{\Omega}}\hat{\rho}\,'$$

does not depend on coordinates and so it characterises in an invariant manner the physical perturbation density.

In a similar way, using (12.9) and (12.13), for the components of the four velocity we get

$$\begin{aligned}
\delta\tilde{u}_0 &= \delta u_0 - \xi^k\partial_k\,\hat{u}_0 - \hat{u}_j\partial_\eta\xi^j = \delta u_0 + \partial_\eta(a\xi^0) = \delta u_0 + \partial_\eta(a\,\underline{\tilde{\Omega}} - a\,\underline{\Omega})\,,\\
\delta\tilde{u}_a &= \delta u_a - \xi^k\partial_k\,\hat{u}_a - \hat{u}_j\partial_a\xi^j = \delta u_a + a\partial_a\xi^0 = \delta u_a + a\partial_a(\underline{\tilde{\Omega}} - \underline{\Omega})\,,
\end{aligned}$$

from which it follows that both the quantities

$$\begin{aligned}
\overline{\delta u_0} &= \delta\tilde{u}_0 - \partial_\eta(a\,\underline{\tilde{\Omega}}) = \delta u_0 - \partial_\eta(a\,\underline{\Omega})\,,\\
\overline{\delta u_a} &= \delta\tilde{u}_a - a\partial_a\underline{\tilde{\Omega}} = \delta u_a - a\partial_a\underline{\Omega}\,,
\end{aligned}$$

are gauge invariant and so characterise the velocity of the fluid in an invariant way. Note that here $u_a = g_{ab}u^b$ are the spatial components of the $4-$vector u_k.

The arbitrariness in the choice of the two scalars $\xi^0, \underline{\zeta}$, permits to fix two conditions on the scalar perturbations $\underline{\phi}, \underline{\psi}, \underline{B}, \underline{E}$ and so only two of them will represent physical perturbations.

12.4 Vector perturbations

By considering only vector perturbations the metric assumes the form

$$ds^2 = a^2(\eta)\left[-d\eta^2 + 2\underline{S}_a dx^a d\eta + (\delta_{ab} + \partial_a \underline{F}_b + \partial_b \underline{F}_a) dx^a dx^b\right] ,$$

and comparing (12.2) and (12.11) the transformations for vectors $\underline{S}_a, \underline{F}_a$ become

$$\tilde{\underline{S}}_a = \underline{S}_a - \underline{\xi}_a{}' , \qquad \tilde{\underline{F}}_a = \underline{F}_a - \underline{\xi}_a ,$$

from which trivially follows that the quantity

$$\overline{V_a} = \tilde{\underline{S}}_a - \tilde{\underline{F}}'_a = \underline{S}_a - \underline{F}'_a$$

is gauge invariant. Of the four independent quantities $\underline{S}_a, \underline{F}_a$, only two of them represent physical peturbations since the other two can be eliminated with a suitable choice of coordinates, that is with a suitable choice of the transverse 3-vector $\underline{\xi}^a$.

12.5 Tensor perturbations

In the case of tensor perturbations only, the metric is quite simple and reads

$$ds^2 = a^2(\eta)\left[-d\eta^2 + (\delta_{ab} - \underline{h}_{ab})\, dx^a dx^b\right] .$$

The perturbation $\tilde{\underline{h}}_{ab} = \underline{h}_{ab}$ is gauge invariant and represents a gravitational wave.

12.6 Cosmological perturbations

We start from Einstein equations

$$G^i_j = 8\pi G\, T^i_j , \qquad \hat{G}^i_j = 8\pi G\, \hat{T}^i_j ,$$

where $\hat{G}^i_j$ is the Einstein tensor in the homogeneous and isotropic background metric $\hat{g}_{ij}$ and of course the energy momentum tensor $\hat{T}^i_j$ has the properties

$$\hat{T}^0_0 = \alpha(t) , \qquad \hat{T}^a_b = \beta(t)\delta^a_b , \qquad \hat{T}^0_b = \hat{T}^a_0 = 0 ,$$

α, be being scalar functions $(-\hat{\rho}, \hat{p}$ for a perfect fluid).

By setting $G^i_j = \hat{G}^i_j + \delta G^i_j$ and similarly $T^i_j = \hat{T}^i_j + \delta T^i_j$ we have the equation for the perturbation in the form

$$\delta G^i_j = 8\pi G\, \delta T^i_j ,$$

which is gauge dependent, but it can be written in a gauge invariant form by introducing the "overlined" quantities as in the section above,in this way

$$\overline{\delta G^i_j} = 8\pi G \overline{\delta T^i_j} .$$

The latter equations assume different forms depending on the perturbation one cosiders.

For scalar perturbations, using (12.10) and (12.13) one has

$$\overline{\delta T^0_0} = \delta \tilde{T}^0_0 - \tilde{\underline{\Omega}}\, \partial_\eta \hat{T}^0_0 = \delta T^0_0 - \underline{\Omega}\, \alpha' , \qquad (12.14)$$

$$\overline{\delta T^0_a} = \delta \tilde{T}^0_a - [(1/3)\hat{T}^c_c - T^0_0]\, \partial_a (\tilde{\underline{\Omega}} = \delta T^0_a - (\beta - \alpha)\, \partial_a \underline{\Omega} , \qquad (12.15)$$

$$\overline{\delta T^a_b} = \delta \tilde{T}^a_b - \tilde{\underline{\Omega}}\, \partial_\eta \hat{T}^a_b = \delta T^a_b - \underline{\Omega}\, \delta^a_b\, \beta' . \qquad (12.16)$$

It is clear that for the tensor G^i_j one obtains similar relations. On the other hand, the perturbation δG^i_j can be explicitly computed because G^i_j is a function of the metric g_{ij} and its derivatives only. This means that under scalar perturbations the invariant quantity $\overline{\delta G^i_j}$ has to depend only on Φ and Ψ. A tedious but straightforward calculation gives

$$\Delta\,\Psi - 3\mathcal{H}(\Psi' + \mathcal{H}\Phi) = 4\pi G\, a^2\, \overline{\delta T^0_0}\,, \tag{12.17}$$

$$\partial_a\,(\Psi' + \mathcal{H}\Phi) = 4\pi G\, a^2\, \overline{\delta T^0_a}\,, \tag{12.18}$$

$$\frac{1}{2}\,(\delta^a_b\Delta\, - \partial^a\partial_b)\,(\Psi - \Phi) - \delta^a_b\,\Big[\Psi'' + \mathcal{H}(2\Psi' + \Phi') + (2\mathcal{H}' + \mathcal{H}^2)\Phi\Big] = 4\pi G\, a^2\, \overline{\delta T^a_b}\,, \tag{12.19}$$

where $\mathcal{H} = a'(\eta)/a(\eta)$, $\partial^a = \delta^{ab}\partial_b$, $\Delta = \delta^{ab}\partial_a\partial_b$ is the Laplace operator and $\overline{\delta T^i_j}$ are the quantities in (12.14)-(12.16) correspondig to scalar perturbations only.

Analog equations to the ones in (12.17)-(12.19) can be obtained for vector and tensor perturbations. In such a cases $\overline{\delta G^i_j}$ will depend respectively on the gauge invariant quantities $\overline{V_a}$ and h_{ab} introduced in previous section. In fact one has

$$\Delta\,\overline{V_a} = 16\pi G\, a^2\, \overline{\delta T^0_a}\,, \qquad \overline{\delta T^0_a} = \delta T^0_a - \underline{F}_b\partial_b\,\hat{T}^0_a - \hat{T}^0_b\partial_a\underline{F}^b\,, \qquad \underline{F}^b = \delta^{ab}\underline{F}_a\,,$$

$$(\partial_\eta + 2\mathcal{H})\,\Big(\partial_a\overline{V_b} + \partial_b\overline{V_a}\Big) = -16\pi G\,\overline{\delta T^a_b}\,, \qquad \overline{\delta T^a_b} = \delta T^a_b - \underline{F}_c\partial_c\,\hat{T}^a_b - \hat{T}^a_c\partial_b\underline{F}^c + \hat{T}^c_b\partial_c\underline{F}^a\,,$$

$$\Big(\partial^2\eta + 2\mathcal{H}\partial_\eta - \Delta\Big)\,\underline{h}_{ab} = 16\pi G\,\overline{\delta T^c_b}\,\delta_{ac}\,.$$

13 Anisotropies in *CMB*

From *CMB* spectrum we know that the universe was very homogeneous and isotropic at the time of recombination, but today it has a well developed nonlinear structure. This structure takes the form of galaxies, clusters and superclusters of galaxies and, on larger scales, of voids, sheets and filaments of galaxies. However, deep redshift surveys show that, when averaged over a few hundred megaparsecs, the inhomogeneities in the density distribution remain small. The explanation of such non linear structures can be found in the primordial inhomogeneities of the energy density and in the natural gravitational instability due to the fact that gravitation is an attractive force.

Independently of their origin (inflation predicts classical inhomogeneities as a consequence of quantum fluctuation of the inflaton condensate – see section 13.5), inhomogeneities in the energy density are necessary in order to accommodate in a reasonable way the observed universe.

We have two kinds of anisotropies of the cosmic microwave background:
i) primary anisotropies due to effects which occur at the last scattering surface and before;
ii) secondary anisotropies due to effects which occur between the last scattering surface and the observer, such as interactions of the background radiation with hot gas or gravitational potentials,

The structure of *CMB* anisotropies is principally determined by two phenomena, that is *acoustic oscillations* and *collisionless damping* (diffusion or Silk damping).

Acoustic oscillations. The pressure of the photons in the photon-baryon plasma in the early universe tends to erase anisotropies, whereas the gravitational attraction of the baryons, moving at speed of sound, makes them to collapse to form dense haloes. These two effects compete to create acoustic oscillations which give the microwave background its characteristic peak structure. Roughly speaking, the peaks correspond to resonances in which the photons decouple when a particular mode is at its peak amplitude.

The peaks contain interesting physical signatures. For example, the angular scale of the first peak determines the curvature of the universe, but not its topology. The locations of the peaks also give important information about the nature of the primordial density perturbations. There are two fundamental brands of density perturbations called *adiabatic* and *isocurvature* perturbations.

In the adiabatic density perturbations the fractional overdensity with respect to the average in each matter component (baryons, photons, ...) is the same (in the considered spot), while in the isocurvature density perturbations the sum of all fractional overdensities with respect to the average is always zero.

Cosmic inflation predicts that the primordial perturbations are adiabatic, while cosmic strings would produce mostly isocurvature primordial perturbations, but in general one could have a mixture of both.

In the *CMB* spectrum adiabatic and isocurvature perturbations produce different peak locations. More precisely, the isocurvature peaks are located at angular scales (l-values of the peaks, see below) in the ratio 1:3:5:..., while adiabatic peaks are located at angular scales in the ratio 1:2:3:... Actual observations of *CMB* anisotropies are consistent with pure adiabatic primordial density, providing key support for inflation.

Collisionless damping. In an expanding universe there are two effects which contribute about equally to the suppression of anisotropies on small scales and give rise to the characteristic exponential damping tail observed in *CMB*. Such effects are due to the fact that the mean free path of the photons increase rapidly, first, because the primordial plasma becomes rarefied and second, because the depth of the last scattering surface is finite (decoupling of photons and baryons does not happen instantaneously, but instead requires an appreciable fraction of the age of the universe up to that era).

13.1 *CMB* power spectrum

The (temperature, angular) power spectrum is a function built up with the fluctuations of the temperature in *CMB*. It relates the anisotropies to important cosmological quantities. One computes the $N-$points correlation functions, which correspond to the $N-$momenta of the ensemble distribution function and then performs the Fourier transform of such correlation obtainig the power spectrum. In our case what is really important is the $2-$point correlation function and since the fluctuations depend on angular variables only, the (angular) power spectrum is obtained by a development in terms of spherical harmonics.

The *CMB* is observed today on the earth where the gravitational potential is $\Phi_0 = \Phi(t_0, x_0^a)$, t_0 being the actual time and $x_0^a = x^a(t_0) = 0$ the origin of the reference frame. One measures the spectrum of photons arriving from all directions and deduces the corresponding black body temperature

$$T(t_0, x_0^a, \mathbf{n}) = T_0 + \delta T(\mathbf{n}) \,, \qquad T_0 = \frac{1}{4\pi}\int d^2\mathbf{n}\, T(t_0, x_0^a, \mathbf{n}) \sim 2.7255^o K \,, \tag{13.1}$$

$\mathbf{n}(\vartheta, \varphi)$ being the unit vector along the direction of observation, $d^2\mathbf{n}$ the integration over all directions and T_0 the mean temperature, which corresponds to the one discussed in section 11.4.
Here $-90^o \leq \vartheta \leq 90^o$ and $0^o \leq \varphi < 360^o$ are galactic coordinates that determine a specific direction, which corresponds to a specific point in the *CMB* (a point on the last scattering surface).

In order to pick out the "underling structure" of anisotropies, one has to compute the so called *correlation functions*, which compare the temperatures of all points having a given angular distance θ. As predicted by inflation, the spectrum of fluctuations is Gaussian and this means that all odd correlation functions are vanishing, while all the even ones are related to the two-point function

$$C(\theta) \equiv \left\langle \frac{\delta T(\mathbf{n}_1)}{T_0}\frac{\delta T(\mathbf{n}_2)}{T_0} \right\rangle \,, \qquad \frac{\delta T(\mathbf{n})}{T_0} = \frac{T(\mathbf{n}) - T_0}{T_0} < 10^{-5} \,, \qquad \mathbf{n}_1 \cdot \mathbf{n}_2 = \cos\theta \,,$$

where the brackets $\langle\rangle$ denote averaging over all directions $\mathbf{n}_1, \mathbf{n}_2$ satisfying the latter condition above. Here θ is the angle between the two considered directions (do not confuse it with the galactic coordinate ϑ). We also observe that

$$\left\langle \left(\frac{T(\mathbf{n}_1) - T(\mathbf{n}_2)}{T_0} \right)^2 \right\rangle = \left\langle \left(\frac{\delta T(\mathbf{n}_1) - \delta T(\mathbf{n}_2)}{T_0} \right)^2 \right\rangle = 2C(0) - 2C(\theta) \,,$$

$$C(0) = \left\langle \left(\frac{\delta T(\mathbf{n})}{T_0} \right)^2 \right\rangle \,, \qquad \left\langle \frac{\delta T(\mathbf{n})}{T_0} \right\rangle = 0 \,,$$

$C(0)$ being the autocorrelation temperature function.

The correlation functions are obtained by averaging the temperature fluctuations measured in all directions on the sky from the earth, the unique vantage point we have access to, but due to homogeneity and isotropy this average has to be close to the *cosmic mean*, which corresponds to the average obtained by all observers in space that measure fluctuations in given directions. This cosmic mean is determined by correlation functions of the random field of inhomogeneities and is the quantity one would have to compute in order to pick up in a correct manner anisotropies and inhomogeneities of the universe. Of course we can only perform local measurements.

The root-mean-square difference between a local measurement and the cosmic mean is known as *cosmic variance*. This difference is due to the poorer statistics of a single observer and depends on the number of appropriate representatives of density inhomogeneities within a horizon. For this reason it is quite tiny at small angular scales but substantial at high angular separations (more than 10 degrees).

According to (11.17), at any angular distance $\Delta\,\theta$ corresponds a region of linear size $\Delta\,l$ on the last scattering surface. A physical process which creates a perturbation density with length scale of the order of the particle horizon, during all evolution of the universe, will "generate" only few observable realisation, that is only few representative points of density inhomogeneity, while physical processes at small length scales will generate more representative points. This means that statistic is better at small angular scales.

Cosmic variance is an unavoidable uncertainty, which is present also for an "ideal" experiment in which one measures temperature in all directions in the sky with arbitrary precision. Of course, for the real observation additional uncertainty is present due to finite number and precision of measurement.

The fluctuations $\delta T(\mathbf{n}) = \delta T(\vartheta,\varphi)$ as well as the correlation function $C(\theta)$ depends on angular coordinates only and so it is convenient to expand them in spherical harmonics Y_{lm} and Legendre polynomials P_{lm}

$$Y_{lm}(\vartheta,\varphi) = \frac{1}{\sqrt{4\pi}}\sqrt{\frac{(2l+1)(l-m)!}{(l+m)!}}P_{lm}(\cos\vartheta)\, e^{im\varphi}\,, \qquad \int_0^{2\pi} d\varphi \int_0^{\pi} d\vartheta\, Y^*_{lm} Y_{l'm'} = \delta_{ll'}\delta_{mm'}\,,$$

by means of

$$\frac{\delta T(\vartheta,\varphi)}{T_0} = \sum_{lm} a_{lm} Y_{lm}(\vartheta,\varphi)\,,$$

where as usual the sum over m runs from $-l$ to l. The sum over l runs from 2 to infinity, since the monopole $l=0$, as well as the dipole $l=1$ contributions are removed "by hand" from the spectrum because they do not have cosmological origin (see next section below).

The expansion coefficients a_{lm} has vanishing mean value, that is $\langle a_{lm}\rangle = 0$, and assuming homogeneity and isotropy their correlation fucntion C_l satisfies the condition

$$\langle a_{lm} a^*_{l'm'}\rangle = \delta_{ll'}\delta_{m,-m'}\, C_l\,, \qquad C_l = \left\langle |a_{lm}|^2 \right\rangle\,.$$

In this way

$$C(\theta) = \left\langle \frac{\delta T(\mathbf{n}_1)}{T_0}\,\frac{\delta T(\mathbf{n}_2)}{T_0} \right\rangle = \sum_{lm} C_l Y_{lm}(\mathbf{n}_1) Y_{l,-m}(\mathbf{n}_2) = \frac{1}{4\pi}\sum_l (2l+1)\, C_l\, P_l(\cos(\theta))\,.$$

From this it follows

$$C_l = \frac{1}{4\pi}\int d^2\mathbf{n}_1\, d^2\mathbf{n}_2\, P_l(\mathbf{n}_1\cdot\mathbf{n}_2) \left\langle \frac{\delta T(\mathbf{n}_1)}{T_0}\,\frac{\delta T(\mathbf{n}_2)}{T_0} \right\rangle\,. \qquad \mathbf{n}_1\cdot\mathbf{n}_2 = \cos\theta\,. \tag{13.2}$$

The multipole coefficients C_l are real and positive and receive their main contribution from fluctuations on angular scale of the order $\theta \sim \pi/l$ and the *CMB power spectrum* defined by $l(l+1)C_l/2\pi$ is about the typical squared temperature fluctuations on this scale. Is is conventional to plot the latter quantity as a function of l in order to point out the contribution of multipoles to the power

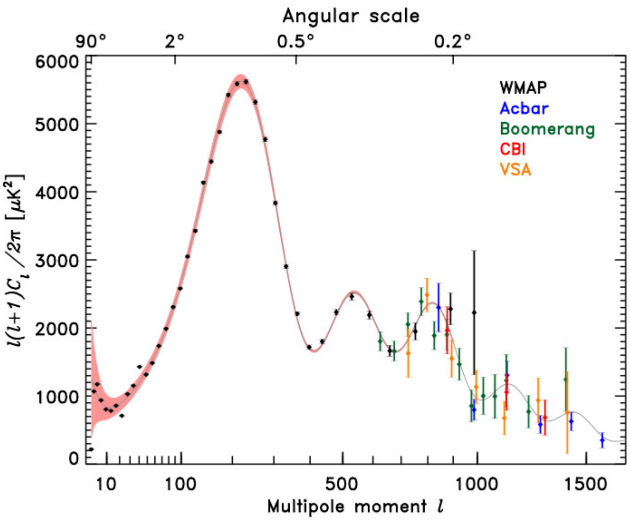

Figure 13: *CMB* power spectrum

spectrum (see figure 13). The monopole $l = 0$ and the dipole $l = 1$ components are excluded in the expansion of $C(\theta)$ as well as in the power spectrum. The first gives a trivial constant contribution, while the second depends on the choice of the reference frame. In the temperature power spectrum measured on the earth, the C_1 dipole contribution is different from zero since the earth moves with respect to the homogeneous-isotropic frame (see next section below).

As we already said above we have access to one vantage point only and so we can perform an average over all directions, but not over all positions in the universe (cosmic mean). This means that the observed multipole coefficients on the earth, say C_l^{obs} are given by (13.2), but without the brackets, that is

$$C_l^{obs} = \frac{1}{4\pi} \int d^2\mathbf{n}_1 \, d^2\mathbf{n}_2 \, P_l(\mathbf{n}_1 \cdot \mathbf{n}_2) \, \frac{\delta T(\mathbf{n}_1)}{T_0} \, \frac{\delta T(\mathbf{n}_2)}{T_0} \, .$$

The difference between the measured value and the hypothetical one obtained by performing the cosmic mean as in (13.2) read

$$\left\langle \left(\frac{C_l - C_l^{obs}}{C_l} \right) \left(\frac{C_{l'} - C_{l'}^{obs}}{C_{l'}} \right) \right\rangle = \frac{2\delta_{ll'}}{2l+1} \, ,$$

from which we see that the cosmic variance goes to zero when l goes to infinity, and so the average on all directions from the earth gives reasonable results for small angles only.

- Note that all theoretical calculations make use of cosmic mean and so the comparison between theoreical previsions and experimental data could be problematic in some cases.

13.2 The dipole anisotropy

This can be easily analysed by looking at the number density of photons in phase space. This quantity is a scalar with respect to Lorentz transformations, because both the phase space volume and the number of photons are scalars. Then we have

$$N_\gamma(\mathbf{p}') = \frac{1}{e^{|\mathbf{p}'|/kT'} - 1} = N_\gamma(\mathbf{p}) = \frac{1}{e^{|\mathbf{p}|/kT} - 1} \, , \qquad (13.3)$$

where $p'^{\mu} = (|\mathbf{p}'|, \mathbf{p}')$ is the 4-momentum of photons in the frame at rest with respect to the cosmic radiation background, where the equilibrium temperature is $T' = T_0$, while $p^{\mu} = (|\mathbf{p}|, \mathbf{p})$ is the 4-momentum in the frame of the earth, which moves with relative velocity v_r with respect to the cosmic background, along the x^1 axis. The 4-momenta are related by the Lorentz transformation

$$p'^{\mu} = \Lambda^{\mu}_{\nu}\, p^{\mu}\,, \qquad \Lambda = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \beta = \frac{v_r}{c}\,, \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}}\,.$$

from which

$$|\mathbf{p}'| = \gamma(1 + \beta\cos\vartheta)\,|\mathbf{p}|\,, \qquad \cos\vartheta = \frac{p_1}{|\mathbf{p}|}\,.$$

From (13.3) it follows

$$T(\vartheta) = \frac{T_0}{\gamma(1+\beta\cos\vartheta)}\,, \tag{13.4}$$

$T(\vartheta)$ being the temperature measured on the earth in the direction ϑ and T_0 the one in (13.1). It is clear from (13.4) that the temperature of photons measured on the earth depends on the angle of observation. We expect to find maximum and minimum values in opposite directions due to blue and red shift, that is $T_{max} = T_0 + \delta T_{dip}$ and $T_{min} = T_0 - \delta T_{dip}$.

WMAP satellite experiment has found $\delta T_{dip} \sim 3.34\,\mu^o K$ at galactic coordinates $(\sim 264^o, \sim 48^o)$. These results indicate a motion of the solar system with a velocity $v \sim 370 km/s$. By taking into account of the motion of the solar system with respect to the center of the galaxy one can deduce a relative velocity $v_r \sim 627 Km/s$ of the local group of galaxies relative to the cosmic background in the galactic direction $(\sim 276^o, \sim 30^o)$.

The temperature in (13.4) can be expanded in powers series of β, the expansion coefficients being functions of Legendre polinomials. The result reads

$$\frac{\delta T(\vartheta)}{T_0} = \frac{T(\vartheta) - T_0}{T_0} = -\beta P_1(\cos\vartheta) - \beta^2\left[\frac{1}{6} - \frac{2}{3}\,P_2(\cos\vartheta)\right] + \ldots$$

Since $\beta \ll 1$, the dominant contribution to the shift of temperature is due to the dipole $P_1(\cos\vartheta)$, but there is also a small quadrupole term which is comparable with the one due to primary anisotropies.

Normally the dipole contribution is removed from the power spectrum of anisotropies because it is a frame-dependent quantity due to the local motion of the earth. The measure of such a quantity permits to determine the "absolute" rest frame in which *CMB* dipole term is vanishing.

13.3 Multipole contributions to *CMB* anisotropies

The temperature fluctuations in the *CMB* at higher multipoles $l \geq 2$ are interpreted as being mostly the result of perturbations in the density and the metric of the very early universe and especially at the surface of the last scattering, even if they do not have exactly the same physical origin. Roughly speaking, depending on the dominant effect which determines the anisotropies, one distinguishes the following regions (see figure 13):

- integrated Sachs-Wolfe rise ($l < 10$, $\theta > 10^o$);
- Sachs-Wolfe plateau ($10 < l < 100$, $10 > \theta > 0.1^o$);

- acoustic peaks ($100 < l < 1000$, $0.1 > \theta > 0.01^o$);
- damping tail ($l > 1000$, $\theta < 0.01^o$).

Note however that for $l < 100$ there are also contributions due to cosmological tensor perturbations in the *FLRW* metric (primordial gravitational waves). All physical effects responsible for the anisotropies are really complicated and we refer the interested reader to the literature for a detailed analysis. Here we only discuss in some detail the Sachs-Wolfe plateau, since for this region it is possible to do some (quite simple) analytical calculations.

13.4 Sachs-Wolfe effect

It consists in the temperature fluctuations of *CMB* due to the fact that photons do not propagate in vacuum, but in the presence of a gravitational potential. As it follows from equivalence principle, the frequency of photons and as a consequence the temperature, depends on the gravtitational potential.

Such an effect is due to scalar perturbations only and so we have to solve equations (12.17)-(12.19), where the energy momentum tensor is $T^i_j = \hat{T}^i_j + \delta T^i_j$, $\hat{T}^i_j$ being those of a perturbed perfect fluid, that is

$$\hat{T}^i_j = (p+\rho)u^i u_j + p\delta^i_j\,, \qquad \hat{T}^0_0 = -\rho(\eta)\,, \qquad \hat{T}^a_b = \delta^a_b\, p(\eta)\,.$$

Note that in contrast with previous section, here ρ, p, u^k represent unperturbed quantities.

By definition the perturbed energy tensor reads

$$T^i_j = \hat{T}^i_j + \delta T^i_j = [(p+\delta p) + (\rho + \delta\rho)](u^i + \delta u^i)(u_j + \delta u_j) + (p + \delta p)\delta^i_j\,,$$

and recalling that in the reference frame we are considering $u^k \equiv (1/a, 0, 0, 0)$, $u_k \equiv (-a, 0, 0, 0)$, at first order in pertubations we get

$$\delta T^0_0 = -\delta\rho\,, \qquad \delta T^0_a = \frac{1}{a}\,(p+\rho)\delta u_a\,, \qquad \delta T^a_b = \delta p\, \delta^a_b\,.$$

From (12.14)-(12.16) then it follows

$$\overline{\delta T^0_0} = -\overline{\delta\rho}\,, \qquad \overline{\delta T^0_a} = \frac{1}{a}\,(p+\rho)\overline{\delta u_a}\,, \qquad \overline{\delta T^a_b} = \overline{\delta p}\, \delta^a_b\,.$$

Using these equations in (12.19) one trivially gets

$$\partial_a \partial_b (\Psi - \Phi) = 0\,, \qquad a \neq b\,,$$

which has the only acceptable solution $\Psi = \Phi$. Note that the general solution $\Psi - \Phi = \alpha + \beta_a x^a$, α being a constant and β_a a constant vector, corresponds to a perturbation if and only if $\alpha = \beta_a = 0$. Now equations (12.17)-(12.19) become

$$\Delta\,\Phi - 3\mathcal{H}(\Phi' + \mathcal{H}\Phi) = -4\pi G\, a^2\, \overline{\delta\rho}\,, \tag{13.5}$$

$$\partial_a\,[a(\eta)\Phi'] = 4\pi G\, a^2\,(p+\rho)\,\overline{\delta u_a}\,, \tag{13.6}$$

$$\Phi'' + 3\mathcal{H}\Phi' + (2\mathcal{H}' + \mathcal{H}^2)\Phi = -4\pi G\, a^2\, \overline{\delta p}\,. \tag{13.7}$$

It is interesting to note that in a non-expanding universe $\mathcal{H} = 0$ and the first equation above becomes equal to the Poisson equation.

In principle, given the equation of state for the fluid, equations above can be solved obtaining the potential Φ and the perturbation $\overline{\delta\rho}, \overline{\delta p}$. We refer the interested reader to the literature and continue to calculate the Sachs-Wolfe effect.

To this aim we consider a gas of photons in the perturbed metric (12.12) (only scalar perturbations) and we fix the coordinates by setting $\underline{B} = \underline{E} = \underline{\Omega} = 0$. This is called *longitudinal or conformal Newtonian* gauge. It is possible to show that this choice fix a unique system of coordinates.

Taking the considerations above into account, the metric simplifies to

$$ds^2 = a^2(\eta)\left[-(1-2\Phi)d\eta^2 + (1+2\Phi)\,\delta_{ab}dx^a dx^b\right]\,, \qquad |\Phi| \ll 1\,. \tag{13.8}$$

The photon propagates along a null geodesic, that is

$$0 = \frac{Dp_k}{d\lambda} = \frac{dp_k}{d\lambda} - \Gamma^r_{ks}p_r p^s = \frac{dp_k}{d\lambda} - \frac{1}{2}\,p^r p^s \partial_k g_{rs}\,, \qquad p^k = \frac{dx^k}{d\lambda}\,, \qquad p^k p_k = 0\,, \tag{13.9}$$

λ being an affine parameter. From the latter equation on the right-hand side above one gets

$$[p^0]^2 = \frac{p^a p_a}{a^2(1-2\Phi)} \sim \frac{1}{a^2}\,(1+2\Phi)\,p^a p_a = \frac{1}{a^4}\,|\mathbf{p}|^2 \quad \Longrightarrow \quad \begin{cases} p^0 = \frac{1}{a^2}\,|\mathbf{p}|\,, \\ p_0 = -(1-2\Phi)\,|\mathbf{p}|\,, \end{cases}$$

where we have set

$$p_a = g_{ab}p^b = a^2(1+2\phi)\delta_{ab}\,p^b\,, \qquad \{p_a\} \equiv (p_1, p_2, p_3)\,, \qquad |\mathbf{p}| = \sqrt{\delta^{ab}p_a p_b}\,.$$

Using the conformal time we also get

$$\frac{dx^a}{d\eta} = \frac{dx^a}{d\lambda}\,\frac{d\lambda}{d\eta} = \frac{p^a}{p^0} = \frac{(1+2\Phi)\delta^{ab}p_b}{|\mathbf{p}|} = \underline{n}^a(1-2\Phi)\,, \tag{13.10}$$

where $\underline{n}^a = \delta^{ab}\underline{n}_b = \delta^{ab}p_b/|\mathbf{p}|$ is a spatial unit vector which determines the direction of propagation.

In the same way from (13.9) at first order in Φ we get

$$\frac{dp_a}{d\eta} = \frac{1}{2p^0}\,p^r p^s \partial_a g_{rs} = |\mathbf{p}| + \frac{a^2\delta_{bc}p^b p^c}{p^0}\,\partial_a\Phi = 2|\mathbf{p}|\,\partial_a\Phi\,. \tag{13.11}$$

In classical mechanics, the Liouville theorem states that the volume in the phase space of a Hamiltonian system is invariant under canonical transformations, or, what is the same, it is conserved along the trajectory of the particle. it is easy to see that such a theorem is valid also in general relativity. To this aim we consider a one-particle system with coordinates (η, x^a) and momentum (p_0, p_a). The volume in the phase space $d^3x d^3p = dx^1 dx^2 dx^3 dp_1 dp_2 dp_3$ is invariant under a general transformation of coordinates. To verify this we set

$$A^i_j = \frac{\partial\tilde{x}^i}{\partial x^j}\,, \qquad B^i_j = \frac{\partial x^i}{\partial\tilde{x}^j}\,, \qquad x^a \to \tilde{x}^a = \tilde{x}^a(\tilde{\eta}, x^b)\,, \qquad p_a \to \tilde{p}_a = B^k_b p_k\,,$$

and observe that at a fixed time $\tilde{\eta}$ the Jacobian of the transformation reads

$$J = \left|\frac{\partial(\tilde{x}^1,\tilde{x}^2,\tilde{x}^3,\tilde{p}_1,\tilde{p}_2,\tilde{p}_3)}{\partial(x^1,x^2,x^3,p_1,p_2,p_3)}\right| = \left|\frac{\partial\tilde{x}^a}{\partial x^b}\right|\,\left|\frac{\partial\tilde{p}_a}{\partial x_b}\right| = \det\left(\left.\frac{\partial\tilde{x}^a}{\partial x^b}\right|_{\tilde{\eta}=const}\right)\,\det\left(\left.\frac{\partial\tilde{x}^a}{\partial x^b}\right|_{\tilde{\eta}=const}\right) = 1\,.$$

The latter result is due the the fact the p_a transforms as a covariant vector and $\tilde{x}^0 = \tilde{\eta} = const$. Since J is trivial, the phase volume is invariant, that is $d^3\tilde{x}d^3\tilde{p} = J\,d^3x d^3p = d^3x d^3p$. Now, the validity of the Liouville theorem directly follows from this result and the principle of equivalence.

Let us consider an ensemble of noninteracting identical particles. The number of particles dN with coordinates in the phase-space volume d^3xd^3p can be written in the form

$$dN = f(\eta, x^a, p_b)\, d^2xd^3p\,,$$

$f(\eta, x^a, p_b)$ representing the density of states at time η. Since the number of particle inside the invariant volume does not change during the evolution, the distribution function has to satisfy the Liouville equation

$$\frac{df(\eta, x^a(\eta), p_b(\eta))}{d\eta} = \frac{\partial f}{\partial \eta} + \frac{\partial f}{\partial x^a}\frac{dx^a}{d\eta} + \frac{\partial f}{\partial p_a}\frac{dp_a}{d\eta} = 0\,. \tag{13.12}$$

This is valid in the absence of particle interactions. If particles interacts, then the total derivative of f is not vanishing, but equal to a term $C(f)$ which takes into acoount of interaction (Boltzmann equation).

Here we are interested in the *CMB*, then we consider a nearly homogeneous isotropic universe filled by slightly perturbed thermal radiation. For a gas of photons f is the Planck dostribution. The energy $E = \omega$ ($\hbar = 1$) of a photon depends on the observer and it is equal to p_0 when measured by an observer at rest in a comoving local inertial frame. It can be written in an invariant way by setting $E = \omega = p_k u^k$, u^k being the velocity of an arbitrary observer.

For such a system the distribution function becomes

$$f = \frac{2}{e^{\omega/T} - 1}\,, \qquad T = T(x^k, \underline{n}^a)\,, \qquad c = \hbar = k_B = 1\,, \tag{13.13}$$

where $T(x^k, \underline{n}^a)$ is the effective temperature which depends on position and time of the observer and on the direction $\underline{n}_a = p_a/|\mathbf{p}|$. Since our universe is nearly isotropic, T will be nearly the equilibrium temperature $\hat{T}$ of the isotropic system, then we set

$$T(x^k, \underline{n}^a) = \hat{T}(\eta) + \delta T\,, \qquad \delta T \ll \hat{T}\,.$$

The fluctuation δT depends on the observer. This is a direct consequence of the fact that f is a scalar and ω is the time component of a vector and so the temperature has to transform as ω. By a change of coordinates $x^k \to \tilde{x}^k$ one gets

$$f(x) \to \tilde{f}(\tilde{x}) = f(x) \quad \Longrightarrow \quad \frac{\tilde{\omega}(\tilde{x})}{\tilde{T}(\tilde{x}} = \frac{\omega(x)}{T(x}$$

In particular, for an infinitesimal change $\tilde{x}^k = x^k + \xi^k$, which relates two observers O and $\tilde{O}$ at rest in their reference we get

$$\omega = p_k u^k = p_0 u^0 = \frac{p_0}{\sqrt{-g_{00}}}\,, \qquad \tilde{\omega} = \tilde{p}_k \tilde{u}^k = \tilde{p}_0 \tilde{u}^0 = \frac{\tilde{p}_0}{\sqrt{-g_{00}}}\,.$$

Using the transformations laws for the 4-vector p_k and definition for the 3-vector l_a, up to higher order we get

$$\tilde{\omega} \sim \omega\,(1 + l_a\,\partial_\eta \xi^a) \quad \Longrightarrow \quad \tilde{T}(\tilde{x}) = \frac{\tilde{\omega}}{\omega}T(x) \sim\; T(x)\,(1 + l_a\,\partial_\eta \xi^a)\;,$$

from which it follows

$$\delta\tilde{T} = \delta T - \xi^0 \partial_\eta \hat{T} + \hat{T}\, l_a\, \partial_\eta \xi^a\,,$$

where

$$\delta\tilde{T} = \tilde{T}(x) - \hat{T}(x)\,, \qquad\qquad \delta T = T(x) - \hat{T}(x)\,.$$

Now we have all elements necessary to solve the Liouville equation (13.12) for the distribution function of the gas of photons in the metric given in (13.8). Using (13.10) and (13.11) we explicitly get

$$\frac{\partial f}{\partial \eta} + \frac{\partial f}{\partial x^a}\frac{dx^a}{d\eta} + \frac{\partial f}{\partial p_a}\frac{dp_a}{d\eta} = \frac{\partial f}{\partial \eta} + (1-2\Phi)\underline{n}^a\,\frac{\partial f}{\partial x^a} + 2|\mathbf{p}|\frac{\partial \Phi}{\partial x^a}\,\frac{\partial f}{\partial p_a} = 0\,.$$

Now we use the fact that the function f in (13.13) depends on the variable $y = \omega/T$ and up to higher order

$$y = \frac{\omega}{T} = \frac{p_0}{T\sqrt{-g_{00}}} \sim -\frac{(1-2\Phi)|\mathbf{p}|}{a(\hat{T}+\delta T)\sqrt{(1-2\Phi}} \sim -\frac{|\mathbf{p}|}{a\hat{T}}\left(1 - \Phi - \frac{\delta T}{\hat{T}}\right)\,.$$

The Liouville equation becomes

$$\frac{\partial y}{\partial \eta} + (1-2\Phi)\underline{n}^a\,\frac{\partial y}{\partial x^a} + 2|\mathbf{p}|\frac{\partial \Phi}{\partial x^a}\,\frac{\partial y}{\partial p_a} = 0\,. \tag{13.14}$$

At zero order in the perturbation this reads

$$\left(\frac{\partial}{\partial \eta} + \underline{n}^a\,\frac{\partial}{\partial x^a}\right)\frac{|\mathbf{p}|}{a\hat{T}} = 0 \quad\Longrightarrow\quad \frac{\partial}{\partial \eta}(a\hat{T}) = 0 \quad\Longrightarrow\quad \hat{T} = \frac{const}{a}\,,$$

which is the result we already have derived in Section 11.3.4.

At first order in the perturbation equation (13.14) gives

$$\begin{aligned} 0 &= \left(\frac{\partial}{\partial \eta} + \underline{n}^a\,\frac{\partial}{\partial x^a}\right)\left[\frac{|\mathbf{p}|}{a\hat{T}}\left(\Phi + \frac{\delta T}{\hat{T}}\right)\right] + 2\Phi\,\underline{n}^a\,\frac{\partial}{\partial x^a}\,\frac{|\mathbf{p}|}{a\hat{T}} - 2|\mathbf{p}|\frac{\partial \Phi}{\partial x^a}\,\frac{\partial}{\partial p_a}\,\frac{|\mathbf{p}|}{a\hat{T}} \\ &= \frac{|\mathbf{p}|}{a\hat{T}}\left[\left(\frac{\partial}{\partial \eta} + \underline{n}^a\,\frac{\partial}{\partial x^a}\right)\left(\Phi + \frac{\delta T}{\hat{T}}\right) - 2\underline{n}^a\,\frac{\partial \Phi}{\partial x^a}\right] \\ &= \frac{|\mathbf{p}|}{a\hat{T}}\left[\left(\frac{\partial}{\partial \eta} + \underline{n}^a\,\frac{\partial}{\partial x^a}\right)\left(\frac{\delta T}{\hat{T}} - \Phi\right) + 2\,\frac{\partial \Phi}{\partial \eta}\right]\,, \end{aligned}$$

from which it follows

$$\frac{d}{d\eta}\left(\frac{\delta T}{\hat{T}} - \Phi\right) = \left(\frac{\partial}{\partial \eta} + \underline{n}^a\,\frac{\partial}{\partial x^a}\right)\left(\frac{\delta T}{\hat{T}} - \Phi\right) = -2\,\frac{\partial \Phi}{\partial \eta}\,. \tag{13.15}$$

The latter equation determines the temperature fluctuations of the microwave background. In the case of practical interest, the universe is matter-dominated after recombination and the main mode in Φ is constant. In such a case along all null geodesics we get the equation

$$\frac{\delta T}{\hat{T}} - \Phi = const\,, \tag{13.16}$$

which describes the influence of the gravitational potential on the microwave background fluctuations. This is known as the *Sachs-Wolfe* effect.

To be more rigorous, one would have to consider a slowly variation in time of the potential Φ, due to the fact that, immediately after ricombination, the radiation is not a completely negligible fraction of the energy density. As a consequence the last term on the right-hand side in (13.15) is not really vanishing, and gives a contribution to $\delta T/\hat{T} - \Phi$ which corresponds to the integral of $\partial_\eta \Phi$ along the null geodesic considered. This is known as the *early integrated Sachs-Wolfe* effect. Another similar contribution to the temperature fluctuation is induced by the time dependence of Φ when, in recent epoch, the universe expansion starts to accelerate (dark-matter dominated era). This is known as the *late integrated Sachs-Wolfe* effect. It is extimated that the contributions due to the integrated Sachs-Wolfe effects do not exceed the 20% of the total amplitude of temperature fluctuations and so for our purposes we can neglect both of them.

13.5 Sachs-Wolfe platou

Now we are going to compute the fluctuation in the temperature of *CMB* measured on the earth due to scalar perturbation Φ. From (13.16) we get

$$\frac{\delta T}{T}(P_f) + \Phi(P_f) = \frac{\delta T}{T}(P_i) + \Phi(P_i)\,, \tag{13.17}$$

where P_i is the point at which the photons are emitted (a point on the surface of last scattering in a given direction) and P_f is the point at which the photons are received (the earth). It is understood that photons run along a geodesic in the direction of observation $\mathbf{n}$ and Doppler effects due to the relative motion of source and receiver are neglected. Moreover, the potential $\Phi(P_f)$ does not depend on direction and so it can be dropped since it gives an isotropic temperature shift (monopole contribution). In this limit equation (13.17) is simply an expression of energy conservation.

We shall see that in the case of adiabatic fluctuations in a critical density, matter dominated universe, the temperature fluctuations are related to the potential $\Phi(P_i)$ only. To this aim we recall that the background temperature T scales as $1/a(t)$ and $a(t) \sim t^{2/3(1+w)}$. This means that, as a consequence of the expansion, there is a temperature fluctuation given by

$$\frac{\delta T}{T} = -\frac{\delta a}{a} = -\frac{2}{3(1+w)}\frac{\delta t}{t}\,, \qquad\qquad p = w\rho\,.$$

In the Newtonian frame where the proper time reads (see (13.8))

$$d\tau = dt\,\sqrt{1-2\Phi} \sim (1-\Phi)\,dt\,,$$

we finally obtain

$$\frac{\delta T}{T} \sim \frac{2}{3(1+w)}\,\Phi\,. \tag{13.18}$$

14 Inflation

The standard cosmological model describes to great accuracy the physical processes through which the universe evolved until the present day. It provides a series of predictions in good agreement with observation and moreover it indirectly confirms the validity of *SM*. However there remain outstanding issues, some of which could be explained in the next future in the framework of *SCM*, for example with the help of quantum gravity, but other problems as homogeneity, flatness and horizon certainly require an extension o *SCM*. These latter mentioned issue are all related to initial conditions.

All the following considerations are obtained essentially by dimensional manipulations and are very far to be rigorous.

14.1 Homogeneity, isotropy and initial conditions

We know by observation that at present time t_0 the domain of homogeneity and isotropy corresponds to a hypersurface Σ_0 of size l_0^3, with l_0 at least as large as the present horizon scale, which roughly corresponds to the Hubble length $L_0 \sim c\,t_0 \sim 10^{28}\,cm$. Such a region was originated by expansion starting from a region Σ_i at time t_i of size l_i^3 given by

$$l_i \sim \frac{a_i}{a_0}\, l_0\,, \qquad a_i = a(t_i)\,, \qquad a_0 = a(t_0)\,, \qquad l_0 \sim L_0\,.$$

The particle horizon at time t_i is of the order $d_P(t_i) \sim ct_i$ (in a radiation dominated flat universe it is exactly $2ct_i$, see 10.25), and so we obtain

$$\frac{l_i}{d_P(t_i)} \sim \frac{t_0 a_i}{t_i a_0}\,. \tag{14.1}$$

This ratio can be arbitrarily large depending on t_i. If we assume *SCM* to be valid until the general relativity is, more or less at Planck scale, then we get a rough estimate of the latter quantity. By choosing initial data at Planck scale (see Appendix A), that is $t_i \sim t_{Pl}$, $T_i \sim T_{Pl}$ we obtain

$$\frac{t_0}{t_i} \sim \frac{t_0}{t_{Pl}} \sim 6 \times 10^{60}\,, \qquad \frac{a_i}{a_0} \sim \frac{T_i}{T_0} \sim \frac{T_{Pl}}{T_0} \sim 0.7 \times 10^{-32}\,, \qquad \frac{l_i}{d_P(t_i)} \sim 10^{28}\,.$$

This means that l_i is built up with 10^{28} causally disconnected intervals and so Σ_i is built up with $(10^{28})^3 = 10^{84}$ causally disconnected regions, but nevertheless it had generated a homogeneous and isotropic domain Σ_0 (actually the fractional variation of the energy density does not exceed $\delta\rho/\rho \sim 10^{-4}$).

We talk about *the initial conditions problem* because we must assume that the energy density was distributed at the very beginning in more than 10^{84} casually disconnected regions in a nearly perfect homogeneous way (see figure 14).

Equation (14.1) can be related to $\dot{a}$ if we assume that $a(t)$ scales as a power of t as it happens in matter and radiation dominated ere. With such an assumption we have $a/t \sim \dot{a}$ and so we get

$$\frac{l_i}{d_P(t_i)} \sim \frac{t_0 a_i}{t_i a_0} \sim \frac{\dot{a}_i}{\dot{a}_0}\,. \tag{14.2}$$

The number of casually disconnected regions is so determined by $\dot{a}_i/\dot{a}_0$.

A similar problem can be formulated for the initial Hubble velocities. In fact it is possible to show that an error in the initial velocities exceeding $10^{-54}\%$ has a dramatic consequence: the universe either recollapses or becomes empty too early.

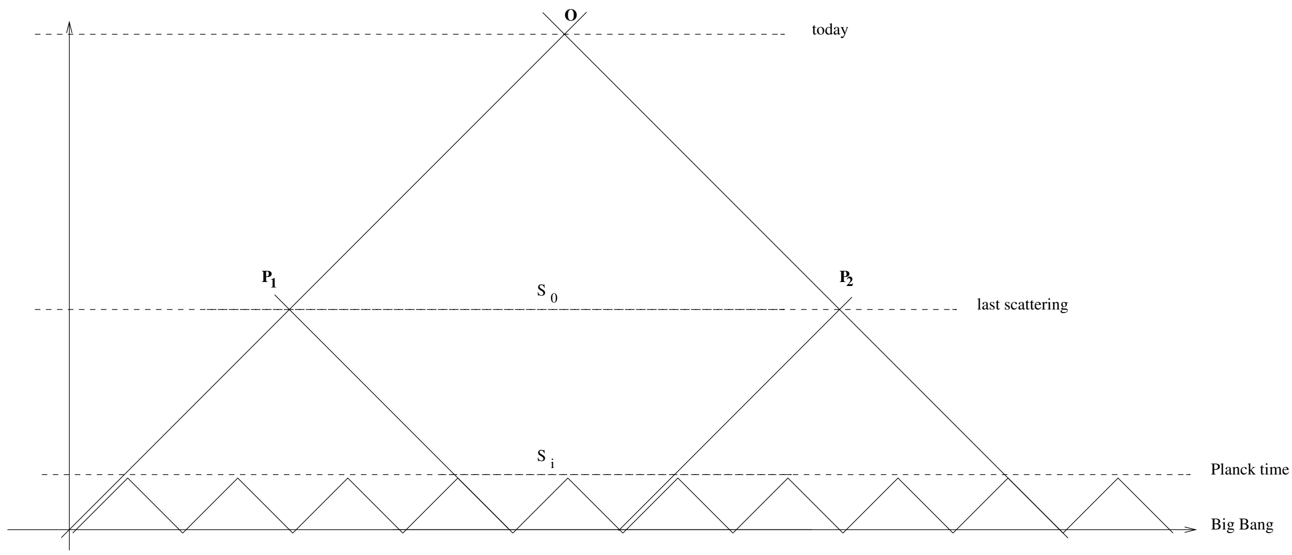

Figure 14: initial conditions (conformal coordinates)

14.2 Flatness problem

Now we are going to study the behaviour of equations in (11.1) from the stability point of view (see Appendix C). They read

$$\Omega - 1 = \frac{k}{a^2 H^2}\,, \qquad \frac{d\Omega}{da} = (1+3w)\,\frac{\Omega(\Omega-1)}{a}\,, \tag{14.3}$$

where

$$\Omega = \Omega(t) = \frac{8\pi G \rho(t)}{3H(t)^2}\,, \qquad H(t) = \frac{\dot a(t)}{a(t)}\,.$$

The second equation in (14.3) has three fixed points, that is $\Omega = 0$, $\Omega = 1$, $\Omega = \infty$. In this case the matrix A has only one component ad so, for $\Omega = 1$ we get

$$A = \frac{d}{d\Omega}\left[(1+3w)\,\frac{\Omega(\Omega-1)}{a}\right]_{\Omega=1} = \frac{1+3w}{a}\,.$$

We see that if $1 + 3w > 0$, then $\Omega = 1$ is a repeller and as a consequence the other two points are attractors, while if $1 + 3w < 0$ then $\Omega = 1$ is an attractor and $\Omega = 0, \infty$ are repellers. This means that in a universe in which the strong energy dominance energy condition is satisfied, independently on the initial values, the system tends to evolve in such a way that $\Omega \to \infty$ or $\Omega \to 0$, which are attractors. However, in contrast to these mathematical considerations, all experimental data are in good agreement with the value $\Omega \sim 1$, which is compatible with $1 + 3w > 0$ if the initial value $\Omega_i = \Omega(t_i)$ was of the order $\Omega_i - 1 \sim 10^{-56}$. In fact we have

$$\begin{cases} \Omega_0 - 1 = \frac{k}{a_0^2 H_0^2} = \frac{k}{\dot a_0^2}\,, \\ \Omega_i - 1 = \frac{k}{a_i^2 H_i^2} = \frac{k}{\dot a_i^2}\,, \end{cases} \Longrightarrow \quad \frac{\Omega_0 - 1}{\Omega_i - 1} = \left(\frac{\dot a_i}{\dot a_0}\right)^2 \sim \left(\frac{l_i}{d_P(t_i)}\right)^2 \sim 10^{56}\,. \tag{14.4}$$

Since $\Omega_0 - 1 \sim 0$, $\Omega_i - 1$ has to be really small too. This remarkable degree of fine tuning is known as the flatness problem, which has no reasonable explanation in *SCM*.

14.3 Horizon problem

This is a theoretical problem which arises as a direct consequence of the fact that *FLRW* cosmology has a particle horizon $d_P(t)$ at any time. In particular, on the last scattering surface there are widely separated points completely outside each others horizons, but nevertheless the *CMB* is isotropic to a high degree of precision. This is related to the fact that the last scattering surface has been generated by causally disconnected region (see figure 14).

In radiation/matter dominated universe the particle horizon is of the order of Hubble length (see (10.25) and (10.26)), that is $d_P(t) \sim 1/H(t)$, but in the presence of inflation the first can be much larger than the latter.

Let us consider a photon moving along a radial geodesic between two arbitrary points $P_1(t_1, r_1, \vartheta, \varphi)$ and $P_2(t_2, r_2, \vartheta, \varphi)$. The coordinate distance traveled by the photon reads

$$\Delta r_{12} = \int_{r_1}^{r_2} \frac{dr}{\sqrt{1-kr^2}} = \int_{t_1}^{t_2} \frac{dt}{a(t)} \,,$$

and the proper distance is obtained by multiplying the latter by the expansion factor $a(t)$.

In a flat, matter dominated universe (see (10.13)) we have

$$a(t) = a_0 \left(\frac{t}{t_0}\right)^{2/3} , \qquad H_0 = \frac{2}{3t_0} \,,$$

and so it follows

$$\Delta r_{12} = \frac{2}{H_0 a_0^{3/2}} \left(\sqrt{a_2} - \sqrt{a_1}\right) , \qquad a_1 = a(t_1) \,, \qquad a_2 = a(t_2) \,. \tag{14.5}$$

In particular, if $t_1 = t_i \sim 0$, $a_1 \sim 0$ then the latter equation gives the (coordinate) particle horizon at time t_2, that is $d_P(t_2) = a(t_2) r_P(t_2)$.

Now let be $t_2 = t_0$ (today) and $t_1 = t_r$ the time at recombination (last scattering surface) more or less for $z \sim 1200$. Then in (14.5) we have $a_2 = a_0$, $a_1 = a_0/(1+z) \sim a_0/1200$ and therefore

$$r_O(t_r) = \frac{2}{H_0 a_0}\left(1 - \frac{1}{\sqrt{1200}}\right) \sim \frac{2}{H_0 a_0} \,.$$

$r_O(t_r)$ represents the coordinate of the photon with respect to the observer at the origin.

On the other hand, the particle horizon at recombination time is obtained again from (14.5) by choosing $a_1 = 0$ and $a_2 \sim a_0/1200$. It follows

$$r_P(t_r) = \frac{2}{H_0 a_0} \frac{1}{\sqrt{1200}} \sim \frac{6 \times 10^{-2}}{H_0 a_0} < r_O(t_r) \,.$$

The comoving distance of a photon in *CMB* and the observer is larger than the particle horizon at recombination time. This means that today we can observe photons in *CMB* which had never been in casual contact, but nevertheless they have the same temperature to high precision (see figure 15).

14.4 Monopole problem

Near the problems we have described above, there are other problems related to grand unified theories (*GUT*). In fact, if such theories are taken seriously, then, in addition to photons and neutrinos, other relics would be present today. Most well known of such kind of *topological defects* are *magnetic*

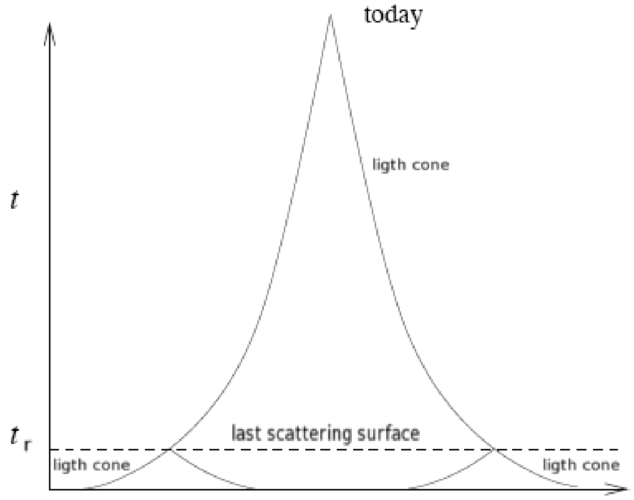

Figure 15: casual structure at recombination

monopoles, which are expected in huge quantity, but they have never been detected and for this reason it is necessary to find some way to dilute their density in the early universe.

Topological defects form when a symmetry is broken and in general they are complicated structures. Monopoles are point-like, but *GUT* predict one dimensional (*cosmic string*), two dimensional (*domain wall*) and more complicated topological defects too.

- It has to be observed that inflation was originally proposed in order to explain the absence on magnetic monopoles in our universe.

14.5 Some considerations about inflation

In order to solve all problems described in previous sections, it is assumed that, at the very beginning, the universe was is a stage of accelerated expansion when gravity acted as a repulsive force. Such a very brief period is called *inflation*. During this stage, the size of the universe had increased enormously in a period of time of the order $t < 10^{-34}\,sec$.

As an effect of this rapid expansion, a small region of space becomes very large and the spatial curvature very small making the universe extremely close to flat. In addition, the horizon size was greatly increased, so that distant points on the *CMB* actually are in causal contact and unwanted relics are tremendously diluted, solving the monopole problem.

Note however that quantum fluctuations made it impossible for inflation to smooth out the universe with perfect precision and for such a reason there is a spectrum of remnant density perturbations. This spectrum turns out to be approximately scale-free, in good agreement with observations of our current universe.

Now we briefly discuss how an inflationary stage can solve the initial value problems. We have seen in (14.2) that the number of casually disconnected regions is determined by the ratio $\dot{a}/\dot{a}_0$, and of course, if gravity is always attractive, such a ratio increases its value with time since $\ddot{a} < 0$. The situation completely changes if for some period gravity becomes repulsive, that is $\ddot{a} > 0$. This can be easily accommodated in *FLRW* by violating the strong dominance energy condition. In such a case $\Omega = 1$ becomes an attractor and the system tends to reach to that value.

A period of accelerated expansion is a necessary condition for solving initial value problems, but whether it is also sufficient depends on the particular model in which this condition is realised. Moreover, because predictions of *SCM* are strongly supported by observation, the accelerated expansion has to stop sufficiently early and in a smooth graceful way, otherwise it would spoil the success of *SCM*. The requirement of the generation of primordial fluctuations, necessary for formation of galaxies, restricts the energy scale of inflation. In simple models inflation should be over $t_f \sim 10^{-34} - 10^{-36}\, sec.$

Inflation explains the origin of the big bang. In fact, because it accelerates the expansion, small initial velocities within a causally connected patch become very large. Furthermore, it can produce the whole observable universe from a small homogeneous region even if the universe was strongly inhomogeneous outside this region. This is due to the fact that in an accelerated universe, there always exists an event horizon given by

$$d_E(t) = a(t) \int_t^{t_{max}} \frac{dt}{a(t)} = a(t) \int_{a(t)}^{a_{max}} \frac{da}{\dot{a}\, a} \,,$$

where t_{max} is the final time and $a_{max} = a(t_{max})$. The integral converges even if $t_{max} = \infty$ or $a_{max} \to \infty$ and so $d_E(t) < \infty$. This means that an observer at a generic time t will never influence the future of the observers which are at a distance $l > d_E(t)$. So, at initial time t_i (the begin of inflation), let us consider two concentric bubbles of spatial dimensions $d_i = d_E(t_i)$ and $2d_i$ and assume the region inside the big bubble $l \leq 2d_i$ to be homogeneous. This means that any point of the small bubble is surrounded by a homogeneous domain of size at least equal to $d_E(t_i)$. Due to the expansion, at time t_f (the end of inflation) the small bubble has a physical size $d_f = (a_f/a_i)\, d_i$, and the region inside $l < d_f$ is still homogeneous, because any point of such a region, at time t_i was outside the event horizon of possible non homogeneous domains.

On the other hand, the particle horizon at t_f is given by

$$\begin{aligned} d_P(t_f) &= a(t_f) \int_{t_i}^{t_f} \frac{dt}{a(t)} = a(t_f) \int_{a_i}^{a_f} \frac{da}{\dot{a}\, a} \\ &= a(t_f) \int_{a_i}^{a_{max}} \frac{da}{\dot{a}\, a} - a(t_f) \int_{a_f}^{a_{max}} \frac{da}{\dot{a}\, a} \sim \frac{a_f}{a_i}\, d_E(t_i) \,. \end{aligned}$$

Because at the end of inflation $a(t_f)$ is very large, the last integral above has been neglected. Then we see that after inflation the size of the homogeneous bubble d_f is of the order of the particle horizon $d_P(t_f)$ and this means that all points inside the small bubble are in casual contact.

Thus, instead of considering a homogeneous universe in many causally disconnected regions, we can begin with a small homogeneous causal domain which inflation blows up to a very large size, preserving the homogeneity irrespective of the conditions outside this domain.

It can be shown that the condition of homogeneity in the original big bubble can be relaxed, because inflation demolishes large initial inhomogeneities and produces in any case a homogeneous, isotropic domain. In order to accommodate *CMB* anisotropies, we have to impose $\dot{a}_i/\dot{a}_0 < 10^{-5}$. In fact in this way, writing (14.4) as

$$\Omega_0 - 1 = (\Omega_i - 1) \left(\frac{\dot{a}_i}{\dot{a}_0}\right)^2 ,$$

we see that $\Omega_0 - 1 \sim 0$ even if $\Omega_i - 1 \sim 1$.

The simplest model which provides an accelerated expansion is the de Sitter universe (see section 9.4), in which the factor $a(t)$ and its derivatives increase exponentially. Unfortunately this can only be a *toy model* because it does not provide a graceful exit from inflation, but nevertheless it can give some interesting insight into its behaviour.

In more realistic models the accelerated phase has to stop very quickly and subsequently the *standard* decelerated phase has to begin. That is

$$\begin{cases} \ddot{a}(t) > 0\,, & \text{for } t < t_f\,, \\ \ddot{a}(t) \sim 0\,, & \text{for } t = t_f\,, \\ \ddot{a}(t) < 0\,, & \text{for } t > t_f\,. \end{cases}$$

This behaviour can be realised with a fluid with an equation of state depending on time and violating the strong dominance energy condition, but only for $t < t_f$, but it can also be obtained with a modified theory of gravity in which the Einstein-Hilbert Lagrangian density is modified by quadratic terms in the Riemann tensor. Such kind on models give rise to de Sitter (like) solutions, but unfortunately they do not provide in a natural way a graceful exit.

Actually, the preferred candidate to drive inflation is a scalar field ϕ called *inflaton*. It has energy and momentum densities given by (spatial derivatives can be neglected)

$$\begin{cases} \rho_\phi(t) = \frac{1}{2}\dot{\phi}^2 - V(\phi)\,, \\ p_\phi(t) = \frac{1}{2}\dot{\phi}^2 - V(\phi)\,, \end{cases} \qquad \Longrightarrow \qquad p + \rho = \dot{\phi}^2\,.$$

We see that $p \sim -\rho$ as required by inflation, if $\dot{\phi}^2 \ll V(\phi)$.

Here we do not enter into the details of inflationary models, but we only recall that one of the aim of inflation it to eliminate all unwanted particles and in fact it redshifts away all unwanted relics, such as magnetic monopoles and other topological defects, but at the same time also any trace of radiation or dust-like matter is similarly redshifted away to nothing. Moreover, inflation is a *supercooling phase* in which the temperature drops by a factor of the order 10^{-5} (from $10^{27}\,{}^oK$ to $10^{22}\,{}^oK$ in some models). Thus, at the end of inflation the universe contains nothing but a cooled scalar field condensate (the inflaton). Then some questions immediately arise:

- How is the universe reheated?
- How does the matter of which we are made arise?
- How does the hot big bang phase of the universe begin?

The physical mechanism which answer to these questions is not yet completely understood due to the unknown nature of the inflaton. The *reheating* after inflation is due to the large energy of the inflaton which decays into particles and radiation. The temperature returns to the pre-inflationary value and the radiation dominated era then finally starts.

Inflation gives also rise to primordial inhomogeneities in the density, which, as a consequence of gravitational instability, generates the observed fluctuations in the *CMB*, as well as all astrophysical structures like clusters and superclusters of galaxies, we see on the largest scales in the universe today.

Before the advent of inflationary cosmology the initial perturbations were postulated and their spectrum was designed to fit observational data, while, according to cosmic inflation, primordial perturbations were originated from quantum fluctuations. These fluctuations have substantial amplitudes only on scales close to the Planckian length, but during the inflationary stage they are stretched to galactic scales with nearly unchanged amplitudes. Thus, inflation links the large-scale structure of the universe to its microphysics. The resulting spectrum of inhomogeneities is not very sensitive to the details of any particular inflationary scenario and has nearly universal shape. This leads to concrete predictions for the spectrum of cosmic microwave background anisotropies. Inflation dilutes away all matter fields, soon after its onset the universe is in a pure vacuum state.

15 Appendices

Here we collect some results we use in the lecture notes.

A Planck units

Here we write down all Planck quantities, which are built up with the Planck and other universal constants. They determine the scale at which quantum gravity effect can not be neglected.

$l_{Pl} = \sqrt{\frac{\hbar G}{c^3}}$	length	$l_{Pl} \sim 1.616 \times\, 10^{-35}\, m$
$M_{Pl} = \sqrt{\frac{\hbar c}{G}}$	mass	$M_{Pl} \sim 2.176 \times\, 10^{-8}\, K_g$
$t_{Pl} = \sqrt{\frac{\hbar G}{c^5}}$	time	$t_{Pl} \sim 5.391 \times\, 10^{-44}\, sec$
$T_{Pl} = \sqrt{\frac{\hbar c^5}{k^2 G}}$	temperature	$T_{Pl} \sim 1.416 \times\, 10^{32}\, {}^oK$

B Boson-Fermion statistic

In table (5) (left) we have collected some results regarding Boson and Fermion statistic. We have indicated by the index i the specie of particle (photon, neutrino, antineutrino, etc.) and by n_i, ρ_i, s_i, p_i the corresponding number density, energy density, entropy density and pressure. T is the equilibrium temperature. In table (5) (rigth) we have collected the number of spin states g_i for all particles we have considered in the lecture notes. The values of Boltzmann and Planck constants read respectively

$$k \sim 8.617 \times\, 10^{-5}\, eV/{}^oK\,, \qquad h \sim 4.135 \times\, 10^{-15}\, eV\; sec\,, \qquad \hbar \sim 6.582 \times\, 10^{-16}\, eV\; sec\,.$$

	Relativistic Bosons	Relativistic Fermions
n_i	$g_i\,\frac{\zeta(3)}{\pi^2}\,\left(\frac{kT}{\hbar c}\right)^3$	$\frac{3}{4}\,g_i\,\left(\frac{kT}{\hbar c}\right)^3$
ρ_i	$g_i\,\frac{\pi^2}{30}\,\frac{(kT)^4}{(\hbar c)^3}$	$\frac{7}{8}\,g_i\,\frac{\pi^2}{30}\,\frac{(kT)^4}{(\hbar c)^3}$
s_i	$g_i\,\frac{2\pi^2}{45}\,\left(\frac{kT}{\hbar c}\right)^3\,k$	$\frac{7}{8}\,g_i\,\frac{2\pi^2}{45}\,\left(\frac{kT}{\hbar c}\right)^3\,k$
p_i	$\frac{1}{3}\,\rho_i$	$\frac{1}{3}\,\rho_i$

specie	γ	ν	$\bar{\nu}$	e^-	e^+
g_i	2	1	1	2	2

Table 5: Boson and Fermion statistic number of spin states

C Autonomous systems

Let us consider a system of first order differential equations of the form

$$\dot{X} = \frac{dX}{dt} = F(X)\,, \qquad X = X(x_1, x_2, ..., x_n)\,, \qquad x_k = x_k(t)\,.$$

This is called an *autonomous system* if the function F does not depend explicitly on t. The particular solution $X = X_c$ which satisfies

$$\dot{X}\Big|_{X=X_c} = F(X_c) = 0\,, \qquad \begin{cases} \dot{X}\Big|_{X=X_c} = 0\,, \\ F(X_c) = 0\,, \end{cases}$$

is called *critical or fixed point*. The behaviour of the system in a neighbourhood of a critical point can be analysed by using the linearised system

$$\dot{X} = A\,X\,, \qquad A = J(X_c)\,, \qquad J(X) = \left(\frac{\partial F(X)}{\partial X}\right)_{X=X_c}\,,$$

where A is the $n \times n$ Jacobian matrix evaluated on the critical point. If all the eigenvalues λ_k of the matrix A_{ij} have a non vanishing real part ($\mathrm{Re}\,\lambda_k \neq 0$), then X_c is said a *hyperbolic critical point*. Depending on the signs of the real parts of λ_k, any fixed point can be classified as an *attractor (silk)*, a *repeller (source)*, a *saddle point*, etc.

For example, in two dimensions $F(x,y) = \{f(x,y), g(x,y)\}$ and if (x_c, y_c) is a fixed point, that is $F(x_c, y_c) = 0$ then

$$A = \begin{pmatrix} \partial_x f & \partial_y f \\ \partial_x g & \partial_y g \end{pmatrix}_{(x,y)=(x_c,y_c)}\,.$$

The critical point is an attractor if both the eigenvalues have negative sign (real part), while it is a repeller in the opposite case. If the eigenvalues have opposite sign then one has to do with a saddle point. If the matrix is real, then one has an attractor if $\mathrm{Tr}\,A < 0$ and $\det A > 0$.