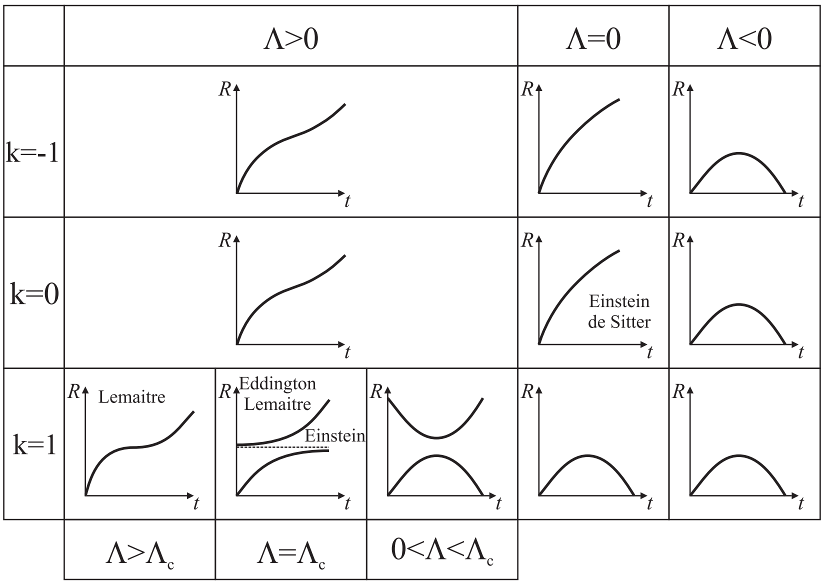
Cosmology I

M. MEZZETTI - *Università degli Studi di Trieste*

academic year 2014/2015



Contents

1 General Relativity

1.1 Introduction

The Newtonian theory of gravity (NG), which allowed us to send probes up to, and beyond, the outskirts of the Solar System, to land on planets and satellites, to rendez-vous with comets and asteroids, is no more adequate when dealing with the structure and evolution of our Universe: as we shall see, Newtonian gravity leads to contradictions. Moreover, NG is not consistent with Special Relativity (SR): when two masses are at rest, Newton's law of gravitation gives the force between them; but, when the masses start moving, their distance, in Newton's law, is the instantaneous distance, so the information on the positions is transmitted with infinite velocity. We know that SR tells us that no information can travel faster than the speed of light c. So, when studying Cosmology, we have to resort to the best theory of gravity on the market today, i.e. General Relativity (GR). Lukily, in order to understand the basic equations governing the cosmic dynamics, we will not need all the technical machinery usually associated to GR. In the following we will just go through the essential concepts needed to understand the meaning of Einstein equations, by making a brief introduction to the basics of this theory.

But, why is GR more complicated than Newtonian gravity? Newtonian gravity is based on the gravitational potential $\Phi(\overline{x})$ at the position $\overline{x}$, which can be derived, given a distribution of mass $\varrho(\overline{x})$, by solving *Poisson equation*

$$\nabla^2\Phi = 4\pi G\varrho, \tag{1}$$

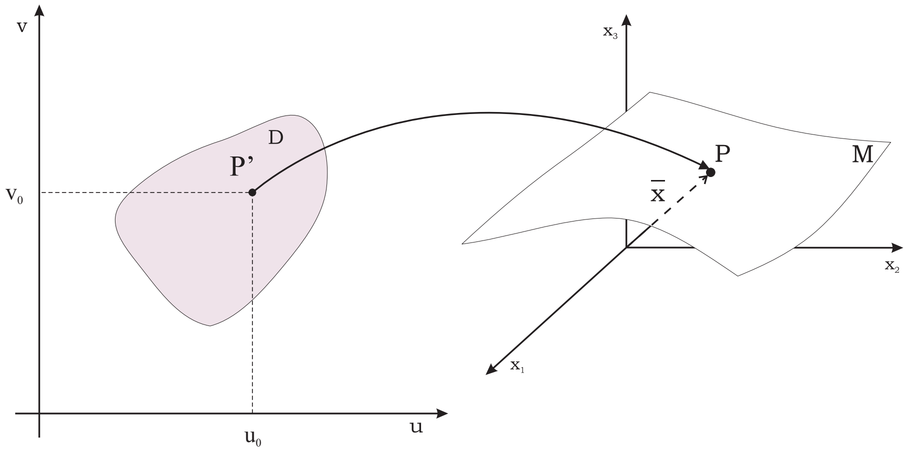
where G is the Newtonian gravitational constant. When we know $\Phi(\overline{x})$ we can derive the acceleration felt by a body at $\overline{x}$: $\overline{g}(\overline{x}) = -\overline{\nabla}\,\Phi(\overline{x})$ and then integrate to find its motion. Eq. (1) is linear in the source term $\varrho(\overline{x})$, so contributions from different bodies simply add together. The gravitational field, as any other field, has some energy associated with it, and SR tells us that energy is equivalent to mass, which in turn produces gravity. So a gravitational field acts in turn as a source of gravitational field and the equations of GR are non linear (while in electromagnetism charges and currents are sources, but not the field itself).

Another difference with electromagnetism is that in this case the acceleration of a body depends on its mass and charge, while in a gravitational field all bodies feel the same acceleration. Einstein thus conceived gravity not as a force like electromagnetism, but as a curvature, a deformation of the geometry of spacetime. So, to understand GR, we must view first some concepts of differential geometry. We start with two-dimensional (2D) surfaces in Euclidean 3D space.

1.2 Surface elements

To be more specific, rather than about surfaces, we will talk about surface elements, as we are interested in their local properties.

We resort to a parametric representation: we consider a bijective function $\overline{x} : D \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ (remember that we work in a three-dimensional Euclidean space $\mathbf{E}^3$).



We define $\overline{x}(u,v) \equiv \big(x_1(u,v), x_2(u,v), x_3(u,v)\big)$. If the surface is expressed in the way $z = f(x,y)$ its parameterization becomes $\overline{x}(u,v) = \big(u, v, f(u,v)\big)$.

A surface is said to be a *regular (smooth) surface* if, having defined the vectors

$$\begin{aligned} \overline{x}_u(u,v) &= \frac{\partial \overline{x}}{\partial u} = \left(\frac{\partial x_1}{\partial u}, \frac{\partial x_2}{\partial u}, \frac{\partial x_3}{\partial u}\right) \\ \overline{x}_v(u,v) &= \frac{\partial \overline{x}}{\partial v} = \left(\frac{\partial x_1}{\partial v}, \frac{\partial x_2}{\partial v}, \frac{\partial x_3}{\partial v}\right) \end{aligned} \tag{2}$$

everywhere (within the domain) $\overline{x}_u \times \overline{x}_v \neq 0$ (cross product).

While keeping fixed $v = v_0$ and by varying u in the neighborhood of a point P' ($\rightarrow P$ on the surface element M) I get a curve on M , whose tangent vector is $\overline{x}_u$. In a similar way, also $\overline{x}_v$ is tangent to a curve on M. These two vectors define the tangent plane to M at the point P.

We can now define a versor $\hat{N}$ perpendicular (normal) to the surface

$$\hat{N} = \frac{\overline{x}_u \times \overline{x}_v}{|\overline{x}_u \times \overline{x}_v|} \qquad \text{and } \hat{N}, \overline{x}_u, \overline{x}_v \text{ form a trihedron.}$$

. . . .

Example: *sphere (in geographic coordinates)*

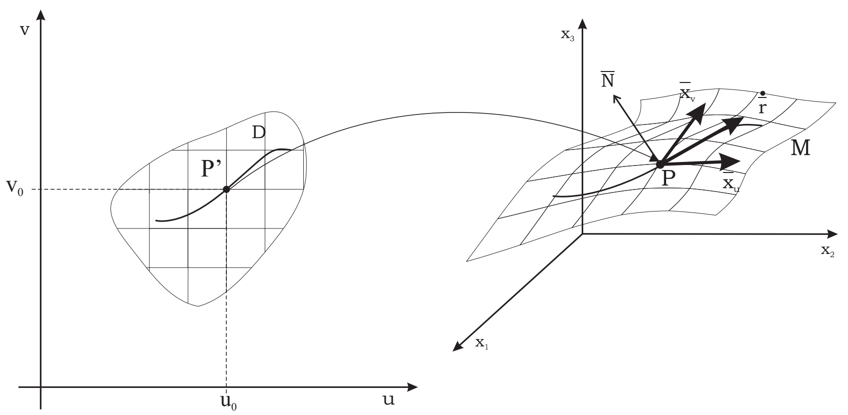
One can describe the surface of the sphere, using the variable u for the longitude ($-\pi \le u \le \pi$) and the variable v for the latitude ($-\frac{\pi}{2} \le v \le \frac{\pi}{2}$), in the following way (most commonly using the colatitude, $\frac{\pi}{2} - v$): $\overline{x}(u,v) = \big(R\cos u \ \cos v, R \ \sin u \ \cos v, R \sin v\big)$.

. . . .

Since in a neighborhood of a point P on M (and of a corresponding point $P' \in D$) the correspondence is bijective, we can think that u and v form, in a neighborhood of P, a system of curvilinear coordinates (like parallels and meridians on a sphere).

If $u = u(t)$, $v = v(t)$ is a curve in D through $P'(u_0, v_0)$, then $\overline{r}(t) = \overline{x}\big(u(t), v(t)\big)$ is a curve on M through $\overline{x}(u_0, v_0)$. The "velocity" vector $\dot{\overline{r}} = \frac{\mathrm{d}\overline{r}}{\mathrm{d}t}$ will be

$$\frac{\mathrm{d}\overline{r}}{\mathrm{d}t} = \dot{\overline{r}} = \frac{\partial \overline{x}}{\partial u}\frac{\mathrm{d}u}{\mathrm{d}t} + \frac{\partial \overline{x}}{\partial v}\frac{\mathrm{d}v}{\mathrm{d}t} \qquad \rightarrow \qquad \dot{\overline{r}} = \overline{x}_u\frac{\mathrm{d}u}{\mathrm{d}t} + \overline{x}_v\frac{\mathrm{d}v}{\mathrm{d}t} \tag{3}$$



The vector $\dot{\overline{r}}$ is also tangent to M and is therefore contained in the tangent plane. Any vector belonging to the tangent plane at P is a linear combination of $\overline{x}_u$ e $\overline{x}_v$ (in $\overline{x}(u_0, v_0)$); conversely, any linear combination $\overline{v} = a\overline{x}_u(u_0, v_0) + b\overline{x}_v(u_0, v_0)$) is the "velocity" vector of a curve on M. The vectors $\overline{x}_u$ e $\overline{x}_v$ form a basis in the tangent plane at the point P.

1.3 The first fundamental form

If $\overline{r}(t) = \overline{x}\big(u(t), v(t)\big)$, with $a \le t \le b$, is a curve on a surface, and if $s = s(t)$ is the arc length (curvilinear abscissa) along $\overline{r}$, from $\overline{r}(a)$ to $\overline{r}(b)$, then the total length L of this curve is obtained by integrating $\frac{\mathrm{d}s}{\mathrm{d}t} = \left|\frac{\mathrm{d}\overline{r}}{\mathrm{d}t}\right|$ on the interval $[a, b]$:

$$L \equiv s(b) = \int_a^b \left|\frac{\mathrm{d}\overline{r}}{\mathrm{d}t}\right| \mathrm{d}t$$

but, since $\dot{\overline{r}} = \overline{x}_u \cdot \dot{u} + \overline{x}_v \cdot \dot{v}$ (with $\dot{u} = \frac{du}{dt}$ e $\dot{v} = \frac{dv}{dt}$)

$$\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 = \left|\frac{\mathrm{d}\overline{r}}{\mathrm{d}t}\right|^2 = \dot{\overline{r}} \cdot \dot{\overline{r}} = (\overline{x}_u \dot{u} + \overline{x}_v \dot{v}) \cdot (\overline{x}_u \dot{u} + \overline{x}_v \dot{v}) = \dot{u}^2(\overline{x}_u \cdot \overline{x}_u) + 2\dot{u}\dot{v}(\overline{x}_u \cdot \overline{x}_v) + \dot{v}^2(\overline{x}_v \cdot \overline{x}_v)$$

Now let $E \equiv \overline{x}_u \cdot \overline{x}_u$, $F \equiv \overline{x}_u \cdot \overline{x}_v$, $G \equiv \overline{x}_v \cdot \overline{x}_v$; ($E = E(u, v,)...$); we obtain:

$$\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)^2 = E\dot{u}^2 + 2F\dot{u}\dot{v} + G\dot{v}^2$$

$$L = \int_a^b \left[E\left(\frac{\mathrm{d}u}{\mathrm{d}t}\right)^2 + 2F\frac{\mathrm{d}u}{\mathrm{d}t}\frac{\mathrm{d}v}{\mathrm{d}t} + G\left(\frac{\mathrm{d}v}{\mathrm{d}t}\right)^2\right]^{\frac{1}{2}} \mathrm{d}t$$

which is shortened writing (it's understood that what matters is the curve, not the parameters used to describe it)

$$L = \int_{\overline{r}} \mathrm{d}s = \int_{\overline{r}} \left[E\mathrm{d}u^2 + 2F\mathrm{d}u\mathrm{d}v + G\mathrm{d}v^2\right]^{\frac{1}{2}}$$

or, in differential form,

$$\mathrm{d}s^2 = E\mathrm{d}u^2 + 2F\mathrm{d}u\mathrm{d}v + G\mathrm{d}v^2 \tag{4}$$

This is the so called **first fundamental form** or **metric form** of a surface.

As we shall see, the metric form determines completely the *intrinsic geometry* of the surface, including its curvature. When we speak of intrinsic geometry we refer to the geometric properties that can be assessed through measures (e.g. distances, but not only) conducted by remaining within the surface, without "going out" from it (that is, without looking at the two-dimensional surface from an Euclidean three-dimensional space). The possibility to define intrinsic properties is essential because, if going from 2 to 3 dimensions, we want to understand the geometry of space that characterizes our universe, we cannot observe it from "outside"!

Notice: Due to the bijective correspondence between the domain $D \in \mathbb{R}^2$ and the surface element M, the curves $u =$ const e $v =$ const form a grid on the surface, and one can think at E, F, e G as functions defined on the surface (and then intrinsic). We may think that the inhabitants of the two-dimensional surface make various measurements of distances between points of the surface to discover the form of the three functions E, F e G, expressed as a function of the curvilinear coordinate grid, perhaps by making assumptions about their possible shape and looking for the best solution.

....

***Example**: the sphere in geographical coordinates:*

$$\begin{aligned} \overline{x}(u, v) &= (R \cos u \cos v, R \sin u \cos v, R \sin v) \\ \overline{x}_u &= (-R \sin u \cos v, R \cos u \cos v, 0) \\ \overline{x}_v &= (-R \cos u \sin v, -R \sin u \sin v, R \cos v) \end{aligned}$$

$$\begin{aligned} E &= \overline{x}_u \cdot \overline{x}_u = R^2 \cos^2 v \sin^2 u + R^2 \cos^2 v \cos^2 u = R^2 \cos^2 v \\ G &= \overline{x}_v \cdot \overline{x}_v = R^2 \sin^2 v \cos^2 u + R^2 \sin^2 v \sin^2 u + R^2 \cos^2 v = R^2 \\ F &= \overline{x}_u \cdot \overline{x}_v = R^2 \cos v \cos u \sin v \sin u - R^2 \cos u \cos v \sin u \sin v = 0 \end{aligned}$$

$$\mathrm{d}s^2 = R^2 \cos^2 v \mathrm{d}u^2 + R^2 \mathrm{d}v^2$$

If we remember that, for $a \leq t \leq b$, $L = \int \mathrm{d}s = \int_a^b \sqrt{(\frac{\mathrm{d}s}{\mathrm{d}t})^2}\mathrm{d}t$, we can write

$$L = \int_a^b \sqrt{R^2 \cos^2 v(\frac{\mathrm{d}u}{\mathrm{d}t})^2 + R^2(\frac{\mathrm{d}v}{\mathrm{d}t})^2}\mathrm{d}t = R \int_a^b \sqrt{\cos^2 v(\frac{\mathrm{d}u}{\mathrm{d}t})^2 + (\frac{\mathrm{d}v}{\mathrm{d}t})^2}\mathrm{d}t$$

and, given the paths $u = u(t)$ and $v = v(t)$ in the domain, we can compute the length L.

. . . .

If $\overline{v} = a\overline{x}_u + b\overline{x}_v$, $\overline{w} = c\overline{x}_u + d\overline{x}_v$, with $a, b, c, d \in \mathbb{R}$, are two vectors tangent to the surface M, then $\overline{v} \cdot \overline{w} = (a\overline{x}_u + b\overline{x}_v) \cdot (c\overline{x}_u + d\overline{x}_v) = acE + adF + bcF + bdG$ which can be written in the matrix form

$$(a, b)\begin{pmatrix} E & F \\ F & G \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} \qquad \text{where} \quad \begin{pmatrix} E & F \\ F & G \end{pmatrix} \quad \text{is the matrix of the first fundamental form}$$

So, if we know the first fundamental form, we are able to compute scalar (dot) products on M: not only lenghts, but also angles!

We remind that being $\overline{x}_u \times \overline{x}_v$ normal to the plane tangent to the surface, the versor $\hat{N} = \frac{\overline{x}_u \times \overline{x}_v}{|\overline{x}_u \times \overline{x}_v|}$ is normal to the surface.

Lagrange identity (important): $|\overline{x}_u \times \overline{x}_v|^2 = (\overline{x}_u \cdot \overline{x}_u)(\overline{x}_v \cdot \overline{x}_v) - (\overline{x}_u \cdot \overline{x}_v)^2 = EG - F^2 = \det\begin{pmatrix} E & F \\ F & G \end{pmatrix}$

Proof: remember that

$$|\overline{x}_u \times \overline{x}_v| = |\overline{x}_u||\overline{x}_v| \ \sin\theta$$
$$\overline{x}_u \cdot \overline{x}_v = |\overline{x}_u||\overline{x}_v| \ \cos\theta$$

so (if we remember that $\sin^2\theta = 1 - \cos^2\theta$) $|\overline{x}_u \times \overline{x}_v|^2 = |\overline{x}_u|^2|\overline{x}_v|^2 \ \sin^2\theta = (\overline{x}_u \cdot \overline{x}_u)(\overline{x}_v \cdot \overline{x}_v) - (\overline{x}_u \cdot \overline{x}_v)^2$ Q.E.D.

The requirement that the surface is smooth implies that $EG - F^2 \neq 0$

At this point we make a change in the symbology used; as we shall see this will lead to a considerable simplification of formulas.

Let's call $g_{11} \equiv E$ $g_{12} = g_{21} \equiv F$ $g_{22} \equiv G$ $\overline{x}_1 \equiv \overline{x}_u$ $\overline{x}_2 \equiv \overline{x}_v$

and let's write $u^1 \equiv u$ $u^2 \equiv v$ (where the superscripts 1 and 2 are upper indices and not exponents).

Then we will have $g_{ij} = \overline{x}_i \cdot \overline{x}_j$ $(i, j = 1, 2)$ and the matrix of the metric form will be:

$$\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix}$$

Remember that $g_{ij} = g_{ij}(u, v) = g_{ij}(u^1, u^2)$.

By defining $g \equiv \det(g_{ij}) = EG - F^2$, from Lagrange identuty $|\overline{x}_1 \times \overline{x}_2|^2 = g$.

In the new notation, the first fundamental form can then be written:

$$\mathrm{d}s^2 = g_{11}(\mathrm{d}u^1)^2 + 2g_{12}\mathrm{d}u^1\mathrm{d}u^2 + g_{22}(\mathrm{d}u^2)^2 = \sum_{i,j} g_{ij}\mathrm{d}u^i\mathrm{d}u^j \tag{5}$$

We used $2g_{12} = g_{12} + g_{21}$ since $g_{12} = g_{21}$; moreover, we will soon understand the reason for we write u^i instead of u_i.

A vector, tangent in P to M, $\overline{v} = a\overline{x}_1 + b\overline{x}_2$ can be written as $\overline{v} = v^1\overline{x}_1 + v^2\overline{x}_2 = \sum_i v^i\overline{x}_i$ (notice that i is a "dummy" variable, and any other letter can be used instead of it.)

If $\overline{v} = \sum_i v^i\overline{x}_i$ and $\overline{w} = \sum_j w^j\overline{x}_j$ are two vectors tangent to M at the same point P, then

$$\overline{v} \cdot \overline{w} = \sum_{i,j}(v^i\overline{x}_i) \cdot (w^j\overline{x}_j) = \sum_{i,j} v^i w^j \overline{x}_i \cdot \overline{x}_j = \sum_{i,j} g_{ij} v^i w^j$$

The vectors $\overline{v}$ and $\overline{w}$ are orthogonal if and only if $\sum_{i,j} g_{ij}v^iw^j = 0$.

We define as g^{ij} the elements of the inverse matrix of (g_{ij}), such that

$$\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}\begin{pmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which, in a more compact way, can be written

$$\sum_j g_{ij} g^{jk} = \delta_i^k \tag{6}$$

where δ_i^k (Kronecker δ) is defined in the following way

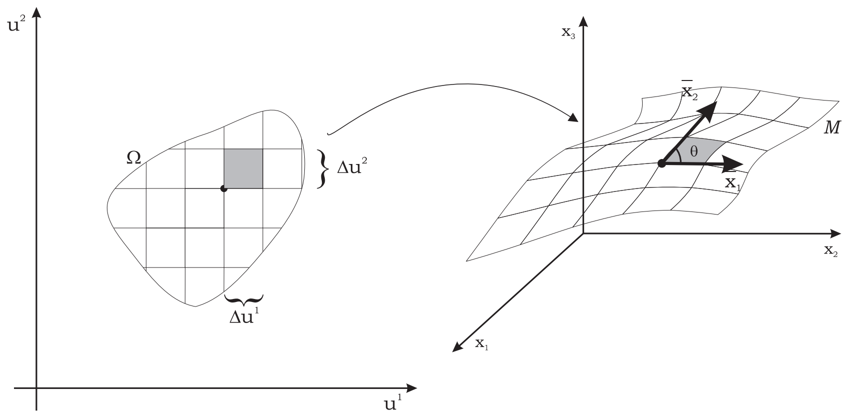
$$\delta_i^k = \begin{cases} 1 & i = k \\ 0 & i \neq k \end{cases} \tag{7}$$

Remembering that the elements of the inverse of a matrix are given by the algebraic complements divided by the determinant of the original matrix, we get:

$$g^{11} = \frac{g_{22}}{g} \qquad g^{12} = g^{21} = -\frac{g_{21}}{g} \qquad g^{22} = \frac{g_{11}}{g}$$

We will now see that the first fundamental form not only allows you to measure distances and angles, but also areas.

Let be $\overline{x} : D \to E^3$ a surface in E^3 and let be $\Omega \in D$ a region of the domain where $\overline{x}$ is bijective. To find the area of $\overline{x}(\Omega)$, we subdivide Ω into rectangular elements by means of lines parallel to the axes u^1 e u^2.



To a small area belonging to Ω, having as sides Δu^1 and Δu^2 corresponds approximately a piece of surface parallelogram-shaped, with sides parallel to the vectors $\overline{x}_1$ e $\overline{x}_2$. These sides have lengths given by $\Delta l_1 \simeq |\overline{x}_1|\Delta u^1$ and $\Delta l_2 \simeq |\overline{x}_2|\Delta u^2$ (Remember that $\overline{x}_1 = \frac{\partial \overline{x}}{\partial u^1}$, and then $\Delta \overline{x}_1 = \frac{\partial \overline{x}}{\partial u^1}\Delta u^1$)

The measure of the small area is given by:

$$\Delta A = |\overline{x}_1|\Delta u^1 \cdot |\overline{x}_2|\Delta u^2 \sin\theta = |\overline{x}_1 \times \overline{x}_2|\Delta u^1 \Delta u^2 = \sqrt{g}\Delta u^1 \Delta u^2$$

where θ is the angle between $\overline{x}_1$ and $\overline{x}_2$, and $g = det(g_{ij})$ as seen above.

Adding all these area elements covering Ω and going to the limit $\Delta u^i \to 0$ we obtain the area of $\overline{x}(\Omega)$:

$$A = \iint_\Omega \sqrt{g}\; \mathrm{d}u^1 \mathrm{d}u^2 \tag{8}$$

We observe that, working in two dimensions, the measure of a set is precisely its area; if we work in three dimensions, the measure will be a volume, and an n-dimensional volume in n dimensions. In all cases, even if we don't prove it here, the measure is obtained by integrating $\sqrt{g}$, where g is the determinant of the n-dimensional metric. This applies in the so-called *Riemannian spaces* (*manifolds*), in which the $ds^2 > 0$. In the *pseudo-Riemannian spaces*, where ds^2 can be positive, negative or equal to zero (such as Minkowski space-time of Special Relativity), some elements of the metric tensor can be negative; since in this case it can be that (as for space-time) $g < 0$, we will use in general the absolute value of g, and we will write $\sqrt{|g|}$ (or, when we know that $g < 0$, $\sqrt{-g}$).

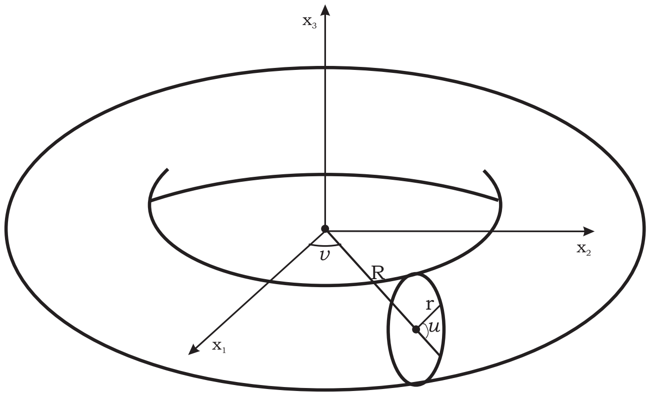
. . . .

Example*: area of the torus*:

$$\overline{x}(u,v) = \left[\, (R + r\cos u)\cos v, (R + r\cos u)\sin v,\ r\ \sin u \right] \qquad \sqrt{g} = r\,(R + r\cos u)$$

$$0 \leq v \leq 2\pi \quad 0 \leq u \leq 2\pi \quad 0 < r < R$$

$$\begin{aligned} S &= \int_0^{2\pi} \left[\int_0^{2\pi} r\,(R + r\cos u)\,\mathrm{d}u \right] \mathrm{d}v = 2\pi r \left[\int_0^{2\pi} R\,\mathrm{d}u + \int_0^{2\pi} r\cos u\,\mathrm{d}u \right] = \\ &= 2\pi r \left[2\pi R + r \int_0^{2\pi} \cos u\,\mathrm{d}u \right] = 4\pi^2 R r \end{aligned}$$



....

1.4 Tensors

Why did we write things like g_{ij} and $\mathrm{d}u^i$ and $\mathrm{d}u^j$? Because we are dealing with tensor quantities, quantities whose properties are related to the way they transform when changing the reference system.

If we switch from the (generally curvilinear) coordinate system u^i $(i = 1, 2, ...) \rightarrow u'^j (j = 1, 2, ...)$ we will get (by means of ... we begin to see how things can be generalized to more than two dimensions)

$$\mathrm{d}u'^j = \sum_i \frac{\partial u'^j}{\partial u^i} \mathrm{d}u^i \qquad (i, j = 1, 2, ...) \tag{9}$$

Every quantity V^j which transforms according to the rule

$$V'^j = \sum_i \frac{\partial u'^j}{\partial u^i} V^i \tag{10}$$

is a *contravariant* tensor (or, to be more precise, its components transform as a *contravariant* tensor); so, also du^i, or u^i, are *contravariant* tensors. A vector is a tensor or rank one. A scalar quantity, the value of which does not change at a given point if we change the coordinate system, is a tensor of rank zero.

We consider now the gradient of a scalar field$\Phi(u^i) = \Phi(u'^j)$. We have:

$$\frac{\partial \Phi}{\partial u'^j} = \sum_i \frac{\partial \Phi}{\partial u^i} \cdot \frac{\partial u^i}{\partial u'^j} = \sum_i \frac{\partial u^i}{\partial u'^j} \cdot \frac{\partial \Phi}{\partial u^i} \tag{11}$$

We see that the gradient of Φ changes differently from du^i! We say that $\frac{\partial \Phi}{\partial u^i}$ is a *covariant* vector, and often we simply write $\partial_i \Phi$ instead of $\frac{\partial \Phi}{\partial u^i}$, with a lower index. When the same index appears both as an upper and a lower index, the sum on that index is implied (Einstein convenction), and we simply write:

$$V'^j = \frac{\partial u'^j}{\partial u^i} V^i \quad \text{and} \quad \frac{\partial \Phi}{\partial u'^j} = \frac{\partial u^i}{\partial u'^j} \frac{\partial \Phi}{\partial u^i}$$

The quantity $\mathrm{d}s^2 = g_{ij} \mathrm{d}u^i \mathrm{d}u^j$ (implying the summation on i and j) is the length, squared, of a segment, and is therefore independent on the reference frame used (it is a scalar). In two different reference frames we will then have

$$\mathrm{d}s^2 = g'_{kl} \mathrm{d}u'^k \mathrm{d}u'^l = g_{ij} \mathrm{d}u^i \mathrm{d}u^j = g_{ij} \frac{\partial u^i}{\partial u'^k} \frac{\partial u^j}{\partial u'^l} \mathrm{d}u'^k \mathrm{d}u'^l \tag{12}$$

$$\text{since} \quad du^i = \frac{\partial u^i}{\partial u'^k} du'^k \quad \text{and} \quad du^j = \frac{\partial u^j}{\partial u'^l} du'^l$$

$$\text{we have} \qquad g'_{kl} = \frac{\partial u^i}{\partial u'^k} \frac{\partial u^j}{\partial u'^l} g_{ij}$$

i.e., the metric tensor is a covariant tensor of rank two.

On the contrary, the g^{ij} tensor is contravariant tensor (of rank two).
We have also seen that the inner product $\overline{v} \cdot \overline{w}$ can be expressed as $\overline{v} \cdot \overline{w} = g_{ij} v^i w^j$.

If we multiply two tensors we also get a tensor:$A_{ij} \cdot C^k = D^k_{ij}$.

If we *contract* a tensor we still have a tensor, but with its rank reduced by two: $T^j_{kmj} = B_{km}$. In fact, e.g.,

$$A'^k_i = \frac{\partial u'^k}{\partial u^j} \cdot \frac{\partial u^l}{\partial u'^i} \cdot A^j_l \Longrightarrow A'^k_k = \frac{\partial u'^k}{\partial u^j} \cdot \frac{\partial u^l}{\partial u'^k} \cdot A^j_l = \frac{\partial u^l}{\partial u^j} \cdot A^j_l = \delta^l_j \cdot A^j_l = A^j_j = A \quad \text{(a scalar)}.$$

If D_i and D^j are the covariant and contravariant components of the same vector (tensor) and we consider a generic vector C^j such that

$$D_i = g_{ij} C^j \; / \cdot D^i \qquad \rightarrow \qquad D_i D^i = g_{ij} C^j D^i,$$

by performing this inner product we obtain, on the left, a scalar that depends on the vector$\overline{D}$, while the right side depends on both $\overline{C}$ and $\overline{D}$; since these two quantities are equal, necessarily $\overline{C} \equiv \overline{D}$, i.e. $D_i = g_{ij} D^j$. We can get this result also in another way. We have seen that a vector $\overline{v}$ can be written as $\overline{v} = v^i \overline{x}_i$, by using its contravariant components; we now define its covariat components v_k in the following way:

$$v_k \equiv \overline{v} \cdot \overline{x}_k = v^i \overline{x}_i \cdot \overline{x}_k = v^i g_{ik} = g_{ik} v^i = g_{ki} v^i \tag{13}$$

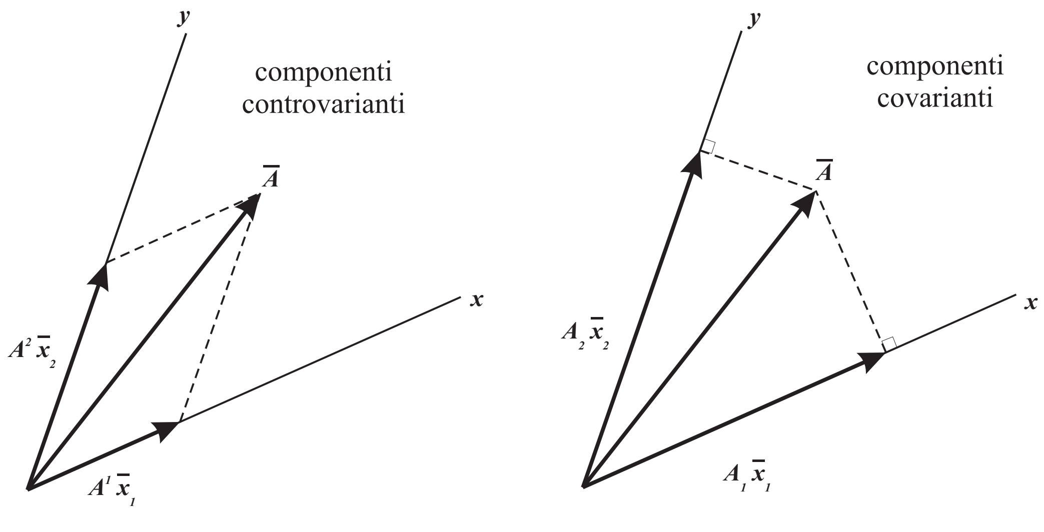
In a similar way we have $D^j = g^{ij} D_i$. **We see that the metric tensor can be used to transform contravariant components into covariant componenets (and vice versa).**

If g_{ij} lowers an upper index, we can use it also to lower an upper index of g^{jk}, obtaining

$$g_{ij} g^{jk} = g^k_i \equiv \delta^k_i \tag{14}$$

on the basis of what has been said above: the metric tensor in the mixed form (i.e. with an upper index and a lower one) is equal to the Kronecker delta.

It is not easy to represent the covariant and contravariant components of a vector in general, but one can give a graphic description in some particular case, for example in the case of rectilinear coordinates. Consider, in the plane, a rectilinear non-orthogonal coordinate system Oxy. Let $\bar{x}_i$ be the basis vectors. If we write the vector $\bar{A}$ as $\overline{A} = A^i \bar{x}_i$, I realize that A^i are the ususal components of a vector, such that the component vectors, having magnitude A^i and direction and versus given by $\bar{x}_i$, add according to the parallelogram rule to give the vector $\bar{A}$. The *contravariant components* correspond to the *parallel projections* on the axes.



Conversely, if I write, as done above, the *covariant components* as $A_i = \bar{A} \cdot \bar{x}_i$, I realize that they are the projections of di $\bar{A}$ along the $\bar{x}_i$ direction; they correspond to the *normal projections* on the axes.

It follows that, if the reference frame is *rectilinear and orthogonal*, parallel and perpendicular projections are the same thing, and covariant and contravariant components are equal. It's no more necessary to distinguish between upper and lower indices.

Notice that a vector (or more generally a tensor), *per se*, is neither covariant nor contravariant, but *its components* are covariant or contravariant.

But every quantity with indices is not necessarily a tensor. For instance, as we shall see, the affine connections Γ^i_{jk} do not represent a tensor, since they do not transform like a tensor.

We can draw an *important conclusion*: **each equation is invariant under a general coordinate transformation if it is expressed as the equality between two tensors with the same upper and lower indexes** :

$$A^{\alpha}_{\beta\gamma} = B^{\alpha}_{\beta\gamma} \to A'^{\alpha}_{\beta\gamma} = B'^{\alpha}_{\beta\gamma} \quad \text{if} \quad A^{\alpha}_{\beta\gamma} \quad \text{and} \quad B^{\alpha}_{\beta\gamma} \quad \text{are tensors.}$$

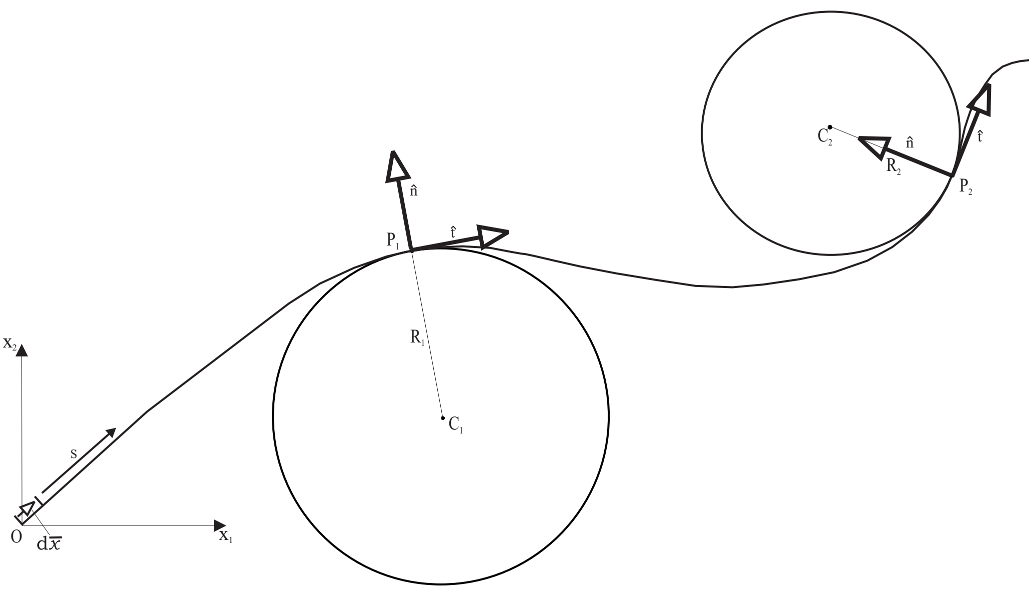
Since also the zero is a tensor of whatever rank (just think that it transforms always into a zero), a relation like $A^{\alpha}_{\beta\gamma} = 0$ will be satisfied in any reference frame.

On the contrary, an equality between quantities that are not tensors with the same upper and lower indices (e.g. $T_{\mu\nu} = 5$; $V^i = B_i$) can be true in some reference frame, but not in all of them.

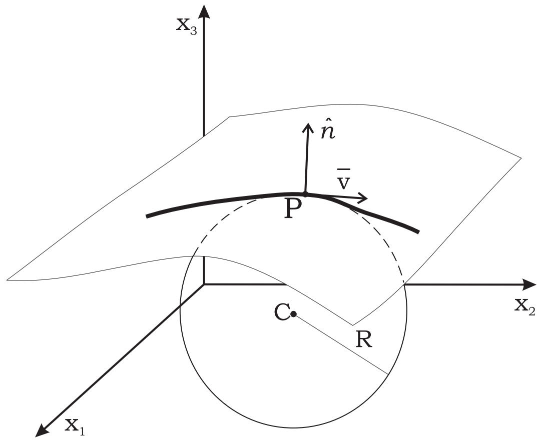
1.5 Curvature of a plane curve and of a surface

Let see first how we can define, in a quantitative way, the curvature of a curve in the plane. A plane curve can be parametrized in the following way: $\bar{x}(t) = \big(x_1(t), x_2(t)\big)$ where t is a parameter, not necessarily the time; the tangent vector (velocity) is $\frac{\mathrm{d}\bar{x}}{\mathrm{d}t}$. The curvilinear abscissa s (arc-length) is defined as

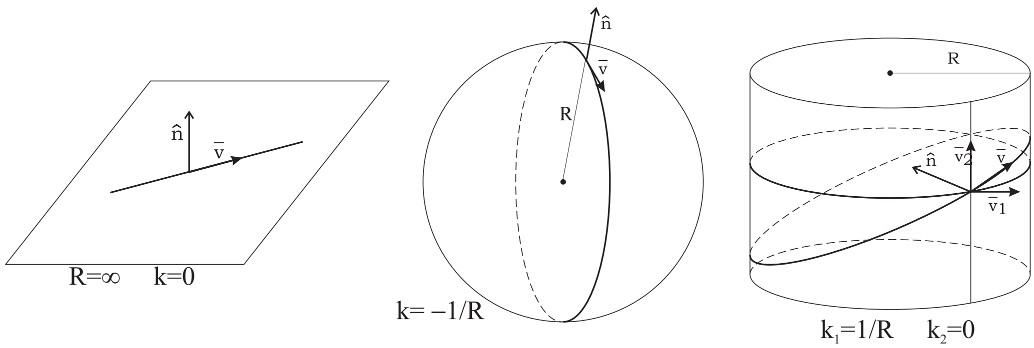
$$O \equiv \bar{x}(t=0) \qquad P \equiv \bar{x}(t) \qquad \mathrm{d}s = |\mathrm{d}\bar{x}| = \left|\frac{\mathrm{d}\bar{x}}{\mathrm{d}t}\right| \mathrm{d}t \quad \to \quad s = \int_0^t \left|\frac{\mathrm{d}\bar{x}}{\mathrm{d}t}\right| \mathrm{d}t = s(t)$$



If we switch parameter from t to s, we notice that $\frac{\mathrm{d}\bar{x}}{\mathrm{d}s} = \dot{\bar{x}}(s)$ has magnitude 1: it's the tangent versor $\hat{t}(s)$. We also define a versor $\hat{n}$ (normal) obtained from a rotation of $\hat{t}$ by 90° in a positive direction (consistent with O, x_1, x_2). At each point P on the curve we can find the osculating circle (the circle which best approximate the curve in an infinitesimal neighbourhood of P) and its radius R; the curvature k of the curve in P is just $k = \pm 1/R$, positive if the center C and the versor $\hat{n}$ are on the same side of the curve with respect to P, negative in the opposite case.



Let's see now how one can extend the notion of curvature to a surface. Let us then consider a point P on a surface, and let $\hat{n}$ be the unit vector normal to the surface in P. If $\overline{v}$ is a vector tangent to the surface at the point P,$\overline{v}$ and $\hat{n}$ define a plane that cuts the surface along a curve which will have, in P, a certain radius of curvature. The curvature in P is given by $k = \pm\frac{1}{R}$, where the sign is taken positive or negative depending on whether the center of curvature C is, with respect to P, on the same side of $\hat{n}$ or on the opposite side (you can also take the opposite choice but, as we will see, things do not change). Let's see, as examples of surfaces, the plane, the sphere and the right cylinder.

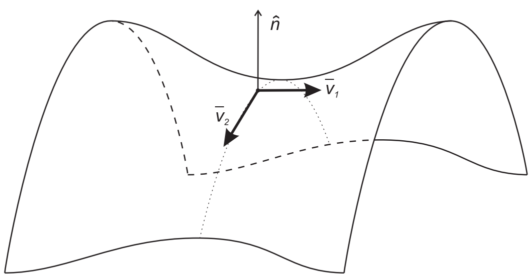


In the case of the cylinder, it can be seen that there are two directions perpendicular to each other and corresponding to the vectors $\overline{v}_1$ and $\overline{v}_2$ which, in turn, correspond to the maximum and the minimum value (k_1 and k_2) of k, the so-called *principal curvatures.* This applies in general, for all smooth surfaces.

The **Gauss curvature** K is defined as the product $k_1 \cdot k_2$. From this we see that K does not depend on the convention adopted for signs of k.

For the plane $K = 0$, for the sphere $K = 1/R^2$, for the cylinder $K = 0$, as for the plane! Although this may appear strange at first sight, actually it reflects the fact that by cutting a right cylinder along a segment parallel to its axis, it can lie on a plane without deforming and without changing lengths and angles of figures drawn on it . The geometry of a cylinder cannot be locally distinguished from that of a plane when we measure angles, lengths, areas, i.e. all those properties that can be measured by moving only along its surface. However, an overall view allows to distinguish a plane from a cylinder: an insect that moves along a circular cross-section (perpendicular to the axis of the cylinder) without turning neither to the right nor to the left, will eventually retrace his steps, but this does not happen on the plane. Also a right circular cone has $K = 0$.

An example of a surface with $K < 0$ is given by a hyperbolic paraboloid (a surface shaped like a saddle) $z = x^2 - y^2$: the two centers of curvature are located on opposite sides with respect to P and then we have $K < 0$.



In general a surface will have $K > 0$ if, with respect to the tangent plane in P, it is "all on one side" (at least locally), while it will have $K < 0$ if the surface is on both sides with respect to the tangent plane in P.

For a torus we have the outer zone with $K > 0$, the inner one with $K < 0$, separated by two circumferences, above and below, with zero curvature.

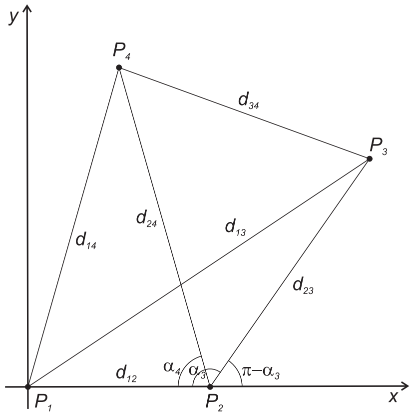
One could define the curvature of a surface in other ways, for example $K' = k_1 + k_2$. In this case, plane and cylinder would be different locally. But the main advantage of the Gauss curvature lies in the fact that, as we shall see, it may be determined by resorting only to measurements carried out on the surface, without the need to "see" the surface in 3 dimensions (as would happen instead for $K' = k_1 + k_2$).

The Gauss curvature is an *intrinsic property* of the surface, and can be determined by knowing the metric tensor $g_{ij}\quad (i,j = 1,2)$. This is the result of the so-called **Theorema Egregium**, so named by the same Gauss.

. . . .

Example: *The mysterious planet* (from Weinberg, 1972)

To have an intuitive idea of how this is possible, consider this example: suppose we have measured on the surface of a celestial body on which we were transported blindfolded (so without seeing it from space!) the distances between four locations P_1, P_2, P_3, P_4 as shown in the figure. Given the values of the six segments, can we tell if the planet's surface is flat or not?



$d_{12} = 780km \qquad d_{13} = 1498km \qquad d_{14} = 1112km \qquad d_{23} = 735km \qquad d_{24} = 960km \qquad d_{34} = 813km$

By Carnot theorem: $d_{13}^2 = d_{12}^2 + d_{23}^2 - 2d_{12}d_{23}\cos\alpha_3$, so that

$$\cos\alpha_3 = \frac{d_{12}^2 + d_{23}^2 - d_{13}^2}{2d_{12}d_{23}}$$

In a similar way

$$\cos\alpha_4 = \frac{d_{12}^2 + d_{24}^2 - d_{14}^2}{2d_{12}d_{24}}$$

Coordinates of the points: $P_1 = (0,0)$; $P_2 = (d_{12}, 0)$;

$P_3 = (d_{12} + d_{23}\cos(\pi - \alpha_3), d_{23}\sin(\pi - \alpha_3) = (d_{12} - d_{23}\cos\alpha_3, d_{23}\sin\alpha_3)$; $P_4 = (d_{12} - d_{24}\cos\alpha_4, d_{24}\sin\alpha_4)$.

$d_{34}^2 = [d_{12} - d_{23}\cos\alpha_3 - d_{12} + d_{24}\cos\alpha_4]^2 + [d_{23}\sin\alpha_3 - d_{24}\sin\alpha_4]^2 = d_{23}^2 + d_{24}^2 - 2d_{23}d_{24}\cos(\alpha_3 - \alpha_4)$

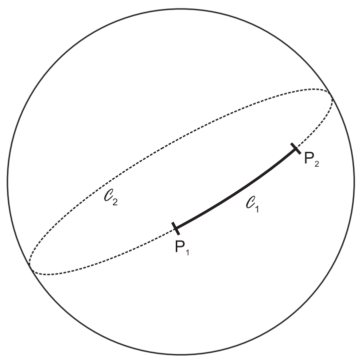
So, if the surface was flat, we would get $d_{34} = 1147.6$, but this is different from the measured value (813!). So we can say thet we are not on a flat planet (if we assume that the surface is a sphere, we could even derive the radius of the planet).

....

1.6 Geodesics

Let $\overline{r}(s) = \left(u^i(s)\right)$, con $a \leq s \leq b$, be a curve on a surface (s being the curvilinear abscissa s) between two points P_1 and P_2 ($P_1 = \overline{r}(a)$; $P_2 = \overline{r}(b)$). We say that this curve is a *geodesic* between P_1 and P_2 if its length is stationary for small variations of the curve which cancel at the extremes. The curve that connects, on the surface, P_1 e P_2 along the shortest path is a geodesic, but the opposite is not always true.

For example, on a sphere both $\mathcal{C}_1$ and $\mathcal{C}_2$ (both arcs of a great circle) are geodesics between P_1 e P_2, but the shortest path corresponds to $\mathcal{C}_1$.



From the relation $\mathrm{d}s^2 = g_{jk}\mathrm{d}u^j\mathrm{d}u^k$, if we express the coordinates u^i in parametric form by means of the parameter t (not necessarily the time) we get:

$$\mathrm{d}s^2 = \left(g_{jk}\frac{\mathrm{d}u^j}{\mathrm{d}t}\frac{\mathrm{d}u^k}{\mathrm{d}t}\right)\mathrm{d}t^2$$

By defining $L(u^i, \dot{u}^i, t) = (g_{jk}\dot{u}^j\dot{u}^k)^{1/2}$ (with $g_{jk} = g_{jk}(u^i)$ and $\dot{u}^i \equiv \frac{\mathrm{d}u^i}{\mathrm{d}t}$) the length of a curve between P_1 and P_2 is:

$$S = \int_{P_1}^{P_2} L\mathrm{d}t = \int_{P_1}^{P_2} \mathrm{d}s$$

To find the condition for S to be stationary we use *Euler-Lagrange equations* (see variational calculus):

$$\frac{\partial L}{\partial u^i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{u}^i}\right) = 0 \tag{15}$$

After some algebra and some "tricks" we get the following relation:

$$\frac{\mathrm{d}^2 u^i}{\mathrm{d}s^2} + \Gamma^i_{jk}\frac{\mathrm{d}u^j}{\mathrm{d}s}\frac{\mathrm{d}u^k}{\mathrm{d}s} = 0 \tag{16}$$

This expresses the condition of stationarity, i.e. it is the differential equation that defines a geodesic. The symbol with three indices Γ^l_{jk} is the so-called *affine connection* or *Christoffel symbol of* 2^{nd} *type*, defined as:

$$\Gamma^i_{jk} = \frac{1}{2}g^{il}\left(\frac{\partial g_{lj}}{\partial u^k} + \frac{\partial g_{lk}}{\partial u^j} - \frac{\partial g_{jk}}{\partial u^l}\right) \tag{17}$$

This quantity depends on g_{ij} and on its first derivatives. Moreover, notice that $\Gamma^l_{jk} = \Gamma^l_{kj}$. Often, in order to simplify even more the notation, we use to write:

$$\frac{\partial g_{ij}}{\partial u^k} \equiv \partial_k g_{ij} \equiv g_{ij},_k$$

Note that Γ^i_{jk} is not a tensor, as

$$\Gamma'^l_{mn} \neq \frac{\partial u'^l}{\partial u^i}\frac{\partial u^j}{\partial u'^m}\frac{\partial u^k}{\partial u'^n}\Gamma^i_{jk}$$

In the geodesic equation the term on the left hand side is a tensor of rank 1 (a contravariant vector), although Γ^i_{jk} is not a tensor. So, if it is zero in a reference system, it will also be zero in a generic reference system.

. . . .

***Example**: the plane in cartesian coordinates*

$\mathrm{d}s^2 = \mathrm{d}u^2 + \mathrm{d}v^2$; since g_{ij} is constant, the Γ are all zero, and geodesics are solutions of

$$\frac{\mathrm{d}^2u}{\mathrm{d}s^2} = 0 \qquad \text{and} \qquad \frac{\mathrm{d}^2v}{\mathrm{d}s^2} = 0 \qquad \rightarrow \qquad \begin{array}{c} u = as + b \\ v = cs + d \end{array}$$

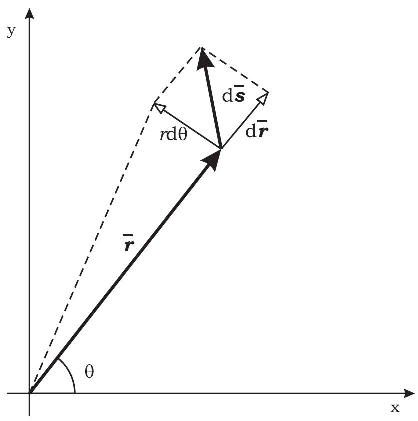
(with a, b, c, d real numbers): those are the parametric equations of a straight line.

. . . .

In a similar, but more complicate, way one can show that arcs of great circle are geodesic lines on the sphere.

. . . .

***Example**: geodesics in the plane in polar coordinates*



$$u^1 \equiv r\,,\; u^2 \equiv \theta\,, \qquad\qquad ds^2 = \mathrm{d}r^2 + r^2\mathrm{d}\theta^2$$

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \qquad g = r^2 \qquad g^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & 1/r^2 \end{pmatrix}$$

$$\frac{\mathrm{d}^2u^i}{\mathrm{d}s^2} + \Gamma^i_{jk}\frac{\mathrm{d}u^j}{\mathrm{d}s}\frac{\mathrm{d}u^k}{\mathrm{d}s} = 0$$

$$\Gamma^i_{jk} = \frac{1}{2} g^{ir} \left(\frac{\partial g_{jr}}{\partial u^k} + \frac{\partial g_{kr}}{\partial u^j} - \frac{\partial g_{jk}}{\partial u^r} \right) \qquad \text{remember the symmetry on } j \text{ and } k$$

$$\Gamma^1_{jk} = \frac{1}{2} g^{11} \left(\frac{\partial g_{j1}}{\partial u^k} + \frac{\partial g_{k1}}{\partial u^j} - \frac{\partial g_{jk}}{\partial u^1} \right) \qquad \text{since } g^{12} = 0$$

$$\Gamma^1_{22} = \frac{1}{2} g^{11} \left(- \frac{\partial g_{22}}{\partial u^1} \right) = -\frac{1}{2} g^{11} \frac{\partial g_{22}}{\partial r} = -\frac{1}{2} \cdot 1 \cdot 2r = -r$$

$$\Gamma^2_{jk} = \frac{1}{2} g^{22} \left(\frac{\partial g_{j2}}{\partial u^k} + \frac{\partial g_{k2}}{\partial u^j} - \frac{\partial g_{jk}}{\partial u^2} \right)$$

$$\Gamma^2_{12} = \frac{1}{2} g^{22} \left(\frac{\partial g_{22}}{\partial u^1} \right) = \frac{1}{2} \cdot \frac{1}{r^2} \cdot 2r = \frac{1}{r} = \Gamma^2_{21}$$

$$\Gamma^1_{11} = \Gamma^1_{12} = \Gamma^2_{11} = \Gamma^2_{22} = 0$$

$$\frac{\mathrm{d}^2 r}{\mathrm{d}s^2} + (-r) \cdot \left(\frac{\mathrm{d}\theta}{\mathrm{d}s} \right)^2 = 0 \qquad \text{(I)}$$

$$\frac{\mathrm{d}^2 \theta}{\mathrm{d}s^2} + \frac{2}{r} \left(\frac{\mathrm{d}r}{\mathrm{d}s} \right) \left(\frac{\mathrm{d}\theta}{\mathrm{d}s} \right) = 0 \qquad \text{(II)}$$

(if $\mathrm{d}\theta/\mathrm{d}s = 0$ we get the staight line passing through the origin); if we put $\mathrm{d}\theta/\mathrm{d}s \equiv \theta'$ and divide (II) by θ' we get:

$$\frac{1}{\theta'} \frac{\mathrm{d}\theta'}{\mathrm{d}s} + \frac{2}{r} \frac{\mathrm{d}r}{\mathrm{d}s} = 0 \qquad \rightarrow \qquad \ln\theta' + \ln r^2 = \ln(\theta' r^2) = \text{const}$$

and then

$$r^2 \frac{\mathrm{d}\theta}{\mathrm{d}s} = h = cost$$

Instead of integrating (I), we use another method. From $\mathrm{d}s^2 = \mathrm{d}r^2 + r^2 \mathrm{d}\theta^2$, dividing by $\mathrm{d}s^2$, we get

$$1 = \left(\frac{\mathrm{d}r}{\mathrm{d}s} \right)^2 + r^2 \left(\frac{\mathrm{d}\theta}{\mathrm{d}s} \right)^2 = \left(\frac{\mathrm{d}r}{\mathrm{d}s} \right)^2 + \frac{h^2}{r^2}$$

You can verify that this relation is an integral of (I). From this we get

$$\frac{\mathrm{d}r}{\mathrm{d}s} = \pm \sqrt{1 - \frac{h^2}{r^2}} = \pm \frac{\sqrt{r^2 - h^2}}{r} \qquad \text{together with} \qquad \frac{\mathrm{d}\theta}{\mathrm{d}s} = \frac{h}{r^2}$$

Dividing the second equation by the first one, to eliminate s, we obtain

$$\frac{\mathrm{d}\theta}{\mathrm{d}r} = \pm \frac{h}{r\sqrt{r^2 - h^2}} = \pm \frac{\mathrm{d}}{\mathrm{d}r} \left[\arccos\left(\frac{h}{r}\right) \right]$$

that is

$$\theta = \pm \arccos\left(\frac{h}{r}\right) + \theta_0 \qquad \rightarrow \qquad \frac{h}{r} = \cos(\theta - \theta_0) \qquad \rightarrow \qquad r\cos(\theta - \theta_0) = h$$

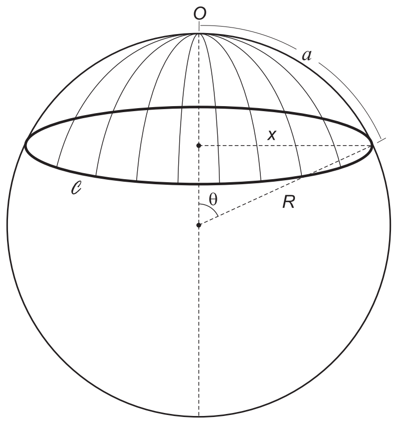
which is precisely the equation of a line in polar coordinates (h is the minimum distance of the line from the origin, obtained for $\theta = \theta_0$).

....

We have defined the geodesics on a surface (which correspond to the lines in the Cartesian plane). We know that, on the plane, the circumference $\mathcal{C}$ of a circle of radius a is $\mathcal{C} = 2\pi a$.

In a similar way, on any surface, to define a circle of radius a and center O, let's draw from this point all the geodesics and let's mark on each of them the point at a distance from O equal to a curvilinear abscissa a; the geometric locus of all these points is the requested circumference. We can now move along this circle (always staying on the surface) and, with the same ruler we used to measure $s = a$, we can measure the length $\mathcal{C}$.

Let's see this for a sphere of radius R.



We will obviously have (we know this because we "see" the sphere in $\mathbf{E^3}$)

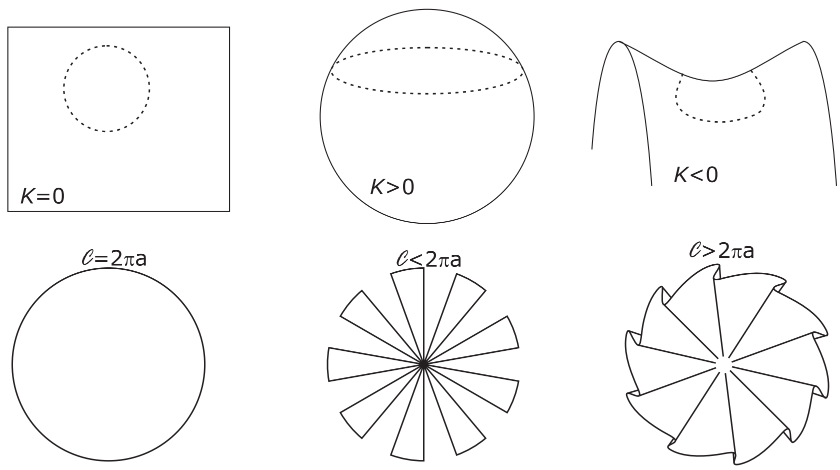
$$C = 2\pi x = 2\pi R\,\sin\big(\frac{a}{R}\big) \simeq 2\pi R\big[\frac{a}{R} - \frac{1}{6}\frac{a^3}{R^3} + ...\big] = 2\pi a - \frac{\pi}{3}\frac{a^3}{R^2} + \mathcal{O}(a^5)$$

But we also know that, for the sphere,$1/R^2 = K$ and, if $a \to 0$, neglecting higher order terms, we can write:

$$K = \frac{3}{\pi}\lim_{a\to 0}\left(\frac{2\pi a - C}{a^3}\right) \tag{18}$$

This result, which is true in general, shows us how to actually derive the Gauss curvature K, with measurements carried out *on the surface*.

For the plane $2\pi a = C$ and $K = 0$; for the sphere $2\pi a > C$ and $K > 0$; around a saddle point $2\pi a < C$ and $K < 0$.



The Gauss curvature is therefore an *intrinsic, local property* of a surface. As the result does not depend on the particular coordinate system used on the surface, K is an *invariant* quantity (such as $\mathrm{d}s^2$, for example), although it may change from point to point on the surface (invariant doesn't mean constant).

While in the Euclidean plane the sum of the internal angles (α, β and γ) of a triangle is equal to π, on a curved surface the result is different. Another property linked to Gauss curvature is that the area A of a triangle formed by geodesic arcs is given by the relation

$$A = \frac{\alpha + \beta + \gamma - \pi}{K}$$

and if $K > 0$, like for the sphere, $\alpha + \beta + \gamma > \pi$, while the opposit holds if $K < 0$.

How does one determine K from g_{ij}? Since the metric tensor contains the information about distances, and measuring these we get K, there must be a link between these two quantities; K must depend on the second derivatives (at least) of g_{ij} at a selected point. This comes from the fact that K is invariant, i.e. does not depend on the coordinate system used, and is a local quantity, that is it depends on the behavior of g_{ij} in an infinitesimal region around the selected point.

It can be proved that, in an infinitesimal neighborhood of an point, we can always choose a coordinate system in which g_{ij} is like $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (actually, g_{ij} can be give the values we want), and in which the derivatives $g_{ij},_k$ are zero, but the second derivatives $g_{ij},_{kl}$ cannot all be set to zero . We name it *locally Euclidean system* .

Since we can always put g_{ij} in the form δ_{ij}, and have $g_{ij},_k = 0$ at a point, the curvature must necessarily depend on the second derivatives of g_{ij}. And the simplest form of dependence would be linear: let's see if we can find some suitable expression. Before doing so, however, we must address another issue.

1.7 Covariant derivative

We have seen that the derivative (the gradient) of a scalar field ϕ, $\partial\phi/\partial u^i$, is a covariant vector. We could then think to perform the derivative of vectoer field $A_i(u^k)$, obtaining in this way a rank two tensor. *But this is not correct!* The differential $\mathrm{d}A_i$ of a vector A_i, basic ingredient of the difference quotient, doesn't in general behave like a tensor. In fact, from the transformation rule

$$A_i = \frac{\partial u'^k}{\partial u^i} A'_k$$

it comes that

$$\mathrm{d}A_i = \frac{\partial u'^k}{\partial u^i}\mathrm{d}A'_k + A'_k \mathrm{d}\frac{\partial u'^k}{\partial u^i} = \frac{\partial u'^k}{\partial u^i}\mathrm{d}A'_k + \frac{\partial^2 u'^k}{\partial u^i \partial u^l} A'_k \mathrm{d}u^l$$

We see that $\mathrm{d}A_i$ is a vector only if $\frac{\partial^2 u'^k}{\partial u^i \partial u^l} = 0$, that is if the u'^i are linear functions of u^i (as it is when we go from a rectilinear coordinate system to another).

But why isn't $\mathrm{d}A_i$ a vector? The reason is that the difference $\mathrm{d}A_i = A_i(u^i + \mathrm{d}u^i) - A_i(u^i)$ is the difference of two vectors that are located in two different points (although infinitely close). The two vectors $A_i(u^i + \mathrm{d}u^i)$ and $A_i(u^i)$ transform then in a different way as the coefficients of the transformations depend on the position. If we want that the difference between two vectors is a tensor, it is necessary that the two vectors are compared at the same point (in this case both, and therefore also their difference, transform in the same way). In order to have a derivative that behaves as a tensor it is necessary to define a new type of derivative, the so-called *covariant derivative.*

For a covariant vector the covariant derivative $\mathrm{D}A_i/\partial u^l$, also written as $A_{i;l}$, is

$$\frac{\mathrm{D}A_i}{\partial u^l} = A_{i;l} = \frac{\partial A_i}{\partial u^l} - \Gamma^m_{il} A_m \tag{19}$$

The covariant derivative for a contravariant vector is:

$$\frac{\mathrm{D}B^i}{\partial u^l} = B^i_{;l} = \frac{\partial B^i}{\partial u^l} + \Gamma^i_{ml} B^m \tag{20}$$

The general rule for the covariant derivative of a tensor of arbitrary rank is to make the partial derivative and then add a term of the type $+\Gamma$ for each contravariant index and a term of type $-\Gamma$ for each covariant index:

$$A^{kl\ldots}_{pq\ldots;j} = \frac{\partial A^{kl\ldots}_{pq\ldots}}{\partial u^j} + \Gamma^k_{mj} A^{ml\ldots}_{pq\ldots} + \Gamma^l_{nj} A^{kn\ldots}_{pq\ldots} + \cdots - \Gamma^r_{pj} A^{kl\ldots}_{rq\ldots} - \Gamma^s_{qj} A^{kl\ldots}_{ps\ldots} - \cdots \tag{21}$$

An important result is that the covariant derivative of the metric tensor is zero: $g_{ik;l} = 0$.

1.8 Parallel transport and curvature tensor

Let $u^i = u^i(s)$ be the parametric equation of a curve, with s curvilinear abscissa measured by starting at a given point on the curve. We know that $\mathrm{d}u^i$ is a vector (from the definition of contravariant vector), $\mathrm{d}s$ is a scalar, and $\mathrm{d}u^i/\mathrm{d}s \equiv v^i$ is then a vector. In particular, v^i is a unit vector, the *versor*[1] *tangent to the* curve.

[1] *To check that $v^i$is a versor, let's see what is his magnitude by means of the sclar product $v_i v^i$:*

$$v_i v^i = g_{ij} v^i v^j = g_{ij} \frac{du^i}{ds}\frac{du^j}{ds} \equiv 1 \quad \Longleftarrow \quad ds^2 = g_{ij}\, du^i du^j$$

If we were in an Euclidean space, to define a geodesic as a segment of a straight line, we would say that the tangent versor does not change with s:

$$\frac{\mathrm{d}v^i}{\mathrm{d}s} = 0$$

If we want now to generalize this relation to any space, 2 or more dimensional, we must not use the normal derivative, but the covariant one, since it is a tensor quantity:

$$\frac{\mathrm{D}v^i}{\mathrm{d}s} = 0$$

Expanding

$$\frac{\mathrm{D}v^i}{\mathrm{d}s} = \frac{\mathrm{D}v^i}{\mathrm{d}u^l}\frac{\mathrm{d}u^l}{\mathrm{d}s} = \frac{\mathrm{d}u^l}{\mathrm{d}s}\left(\frac{\partial v^i}{\partial u^l} + \Gamma^i_{ml}v^m\right) = 0$$

that is

$$\begin{aligned} \frac{\partial v^i}{\partial u^l}\frac{\mathrm{d}u^l}{\mathrm{d}s} + \Gamma^i_{ml}v^m\frac{\mathrm{d}u^l}{\mathrm{d}s} &= 0 \\ \frac{\mathrm{d}v^i}{\mathrm{d}s} + \Gamma^i_{ml}v^m v^l &= 0 \end{aligned}$$

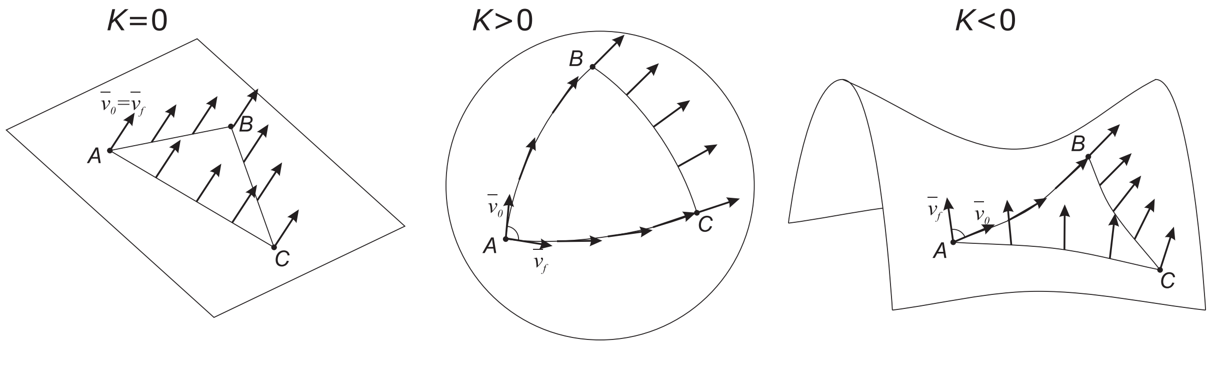
from this, by remembering that $\mathrm{d}u^i/\mathrm{d}s \equiv v^i$, we have

$$\frac{\mathrm{d}^2u^i}{\mathrm{d}s^2} + \Gamma^i_{ml}\frac{\mathrm{d}u^m}{\mathrm{d}s}\frac{\mathrm{d}u^l}{\mathrm{d}s} = 0$$

We find again the geodesic equation (and this is another proof of the fact that, when we leave the Euclidean space, we must switch from usual derivatives to covariant derivatives).

We now introduce the concept of *parallel transport* of a vector along a geodesic: *a vector parallel transported along a geodesic always form the same angle with the tangent to the curve.*

Now imagine we parallel transport a vector $\overline{v}_0$ along a triangle formed by pieces of geodesic. If we are in a Euclidean space (e.g. on a plane) the vector $\overline{v}_f$ we get after closing the path coincides with $\overline{v}_0$.



The same thing does not happen along a spherical triangle: the vector is rotated by an angle which has the same direction of rotation of the direction in which we moved along the spherical triangle. The opposite happens if $K < 0$. We can look at it in another way: imagine we have to go from point A to point B, either directly or through a point C, always along geodesic arcs. In Euclidean space the result of the parallel transport along the two paths is the same, but the same thing does not happen on curved surfaces (what said here for a triangle formed by arcs of geodesic applies to a generic path, which can be thought as consisting of a large number of arches of geodesic). The result is that, unless we are in a Euclidean space, *there is no natural and not ambiguous way to move a vector from one point to another*; we can move it in parallel, but the result depends on the path, and there is not a natural choice for this. So *we can compare two vectors only if they are applied at the same point.* For example, two particles that pass alongside one another have a well-defined relative velocity (and less than c, with c the speed of light), but two particles in different points of a generic space do not have a well-defined, relative velocity.

Let's now quantify what we said above in a qualitative way. Moving along a closed path formed by arcs of geodesic, a vector A_k parallel transported will undergo, returning to the starting point, a variation ΔA_k. If we assume that the surface element bounded by the closed curve is infinitesimal (any finite surface element can be divided into infinitesimal elements), the integrand is constant and, by neglecting infinitesimals of higher order, ΔA_k can by shown to be proportional to a tensor containing products and derivatives of the affine connections:

$$\Delta A_k \quad \propto \quad \frac{\partial \Gamma^i_{km}}{\partial u^l} - \frac{\partial \Gamma^i_{kl}}{\partial u^m} + \Gamma^n_{km}\Gamma^i_{nl} - \Gamma^n_{kl}\Gamma^i_{nm}$$

This tensor is named **Riemann-Christoffel tensor** or **curvature tensor**:

$$R^i_{\ klm} = \frac{\partial \Gamma^i_{km}}{\partial u^l} - \frac{\partial \Gamma^i_{kl}}{\partial u^m} + \Gamma^n_{km}\Gamma^i_{nl} - \Gamma^n_{kl}\Gamma^i_{nm} \tag{22}$$

(*Warning*: you can find it defined with the signs interchanged!) If, in a volume of space, $R^i_{\ klm} = 0$, then $\Delta A_k = 0$: The parallel transport along a closed curve keeps the vector unchanged, and that volume of space is said to be **flat**. This happens in a Euclidean space, as well as in any (volume of) space in which g_{ij} is constant, because the affine connections are null and so also the curvature tensor; and since a tensor equal to zero in a coordinate system is zero in any coordinate system, then $R^i_{\ klm} = 0$ in any refernce frame. On the contrary, if $R^i_{\ klm} \neq 0$ the parallel transport depends on the path, and the space (or the volume of space) is said, by contrast, **curved** (this is the reason for the Riemann-Christoffel tensor is also named curvature tensor).

1.9 Properties of the curvature tensor

It can be proved[2] that $R^i_{\ klm}$ is the only tensor that can be constructed from the metric tensor and its first and second derivatives, and which is linear in the second derivatives (and also quadratic in the first derivatives). The metric tensor allows to write it in the totally covariant form $R_{jklm} = g_{ji}R^i_{\ klm}$.

The Riemann tensor has its own properties, let's see them in the fully covariant form$R_{jklm} = g_{ji}R^i_{\ klm}$:

- *Symmetry* properties
$$R_{jklm} = R_{lmjk}$$
- *Antisymmetry* properties
$$R_{jklm} = -R_{kjlm} = -R_{jkml} = R_{kjml}$$
- *Cyclic* properties
$$R_{jklm} + R_{jmkl} + R_{jlmk} = 0.$$

From the Riemann tensor, by contraction, we can get a rank 2 tensor, the *Ricci tensor*, defined as:

$$R_{km} \equiv R^i_{\ kim} \tag{23}$$

(indices i and l of $R^i_{\ klm}$ are contracted). Ricci tensor is symmetric:

$$R_{mk} = R^i_{\ mik} = g^{ir}R_{rmik} = g^{ir}R_{ikrm} = R^r_{\ krm} = R_{km}$$

It is the only symmetric tensor of rank 2 that can be obtained from $R^i_{\ klm}$. From the Ricci tensor one can obtain the *Ricci scalar* or *curvature scalar*:

$$R = g^{km}R_{km} \tag{24}$$

It is the only scalar that can be obtained from $R^i_{\ klm}$.

All these properties of the Riemann tensor reduce the number of its independent components and, in N dimensions, this number is $\mathcal{N} = \frac{N^2(N^2-1)}{12}$. In particular:

- For $N = 1$, $\mathcal{N} = 0$ and $R_{1111} \equiv 0$ always: a curve has always (intrinsic) curvature zero, we do not have information on how the curve is "embedded" in a space with 2 or more dimensions.
- For $N = 2$, $\mathcal{N} = 1$. There is only one independent componenet, for instance R_{1212}.

[2]See, e.g., *Weinberg*, 1972

- For $N = 3$, $\mathcal{N} = 6$, as many as the components of the (symmetrical) Ricci tensor. So for $N = 3$ it is sufficient to know R_{km} to describe the space cuvatura.
- For $N = 4$, $\mathcal{N} = 20$, while R_{km} has only 10 components. One must use the complete $R^i_{\ klm}$ tensor (apart from situations of particular symmetry, and we'll see that it is so in the case of the homogeneous and isotropic universe).

If we apply covariant derivatives to the curvature tensor we get the so called *Bianchi identities*, which can be expressed in different forms. A particularly important one is the following:

$$R^l_{\ j;l} - \frac{1}{2}R_{;j} = 0. \tag{25}$$

The quantity $R^l_{\ j;l}$ is the (covariant) *divergence* of the Ricci tensor; since R is a scalar (it does not depend on the reference system used) its covariant derivative coincides with the simple partial derivative. Now consider the mixed tensor

$$R^l_{\ j} - \frac{1}{2}\delta^l_{\ j}R$$

Its covariant divergence is (for the rule of the derivation of a product and being $\delta^l_{\ j;l} = 0$ [3]

$$R^l_{\ j;l} - \frac{1}{2}\delta^l_{\ j}\frac{\partial R}{\partial u^l} = R^l_{\ j;l} - \frac{1}{2}\frac{\partial R}{\partial u^j} = 0$$

as seen just above. So the (covariant) divergence of this tensor is equal to zero. If we switch to its covariant components we get

$$g_{il}R^l_{\ j} - \frac{1}{2}g_{il}\delta^l_{\ j}R = R_{ij} - \frac{1}{2}g_{ij}R \equiv G_{ij} \tag{26}$$

where G_{ij} is the so-called *Einstein tensor.* This tensor has very relevant properties: it is *symmetric*, has *vanishing divergence* and, since it comes from Riemann tensor, it contains *terms linear in the second derivatives* of the metric and *quadratic in its first derivatives.*

1.10 The Theorema Egregium

In 2 dimensions the *Theorema Egregium* of Gauss states that the Gauss cirvature K can be derived from the metric tensor; in particular $K = R_{1212}/g$ (remember that in two dimensions the curvature tensor has only one independent component). It can also be shown that $R_{1212}/g = R/2$, so even if R_{1212} is a component of a tensor, and g is not a scalar, their ratio is a scalar

We observe that the relation which expresses the Theorema Egregium, $K = R_{1212}/g = R/2$, is a relationship between tensors of rank zero, i.e. scalars. If it is true in a particular frame of reference, it applies in any frame of reference.

The curvature tensor is related to the Gauss curvature even in spaces with any number of dimensions. Given a point P in one of these spaces, and two vectors a^μ and b^μ applied at the point P, we can draw a family of geodesic curves $x^\mu(s, \alpha, \beta)$ through P, with α and β real numbers. All these geodesics, which have as their initial tangent vector $\mathrm{d}x^\mu/\mathrm{d}s = \alpha a^\mu + \beta b^\mu$, form a two-dimensional surface through P, with Gauss curvature given by[4]

$$K(a,b) = \frac{R_{\lambda\mu\nu\kappa}a^\lambda b^\mu a^\nu b^\kappa}{(g_{\lambda\nu}g_{\mu\kappa} - g_{\lambda\kappa}g_{\mu\nu})\, a^\lambda b^\mu a^\nu b^\kappa}$$

....

***Example**: estimation of Gauss curvature*

Given, for a surface element, the following metric

$$\mathrm{d}s^2 = \mathrm{d}u^2 + e^{\frac{2u}{k}}\,\mathrm{d}v^2$$

estimate K (intrinsic Gauss curvature).

[3] $\delta^l_{\ j;l} = \frac{\partial \delta^l_{\ j}}{\partial u^l} + \Gamma^l_{lk}\delta^k_{\ j} - \Gamma^m_{jl}\delta^l_{\ m} = \Gamma^l_{lj} - \Gamma^l_{jl} = 0$

[4] See *Weinberg* 1972, Section 6.9

We know that $K = R_{1212}/g$

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & e^{2u/k} \end{pmatrix} \quad \rightarrow \quad g = e^{2u/k} \quad \rightarrow \quad g^{ij} = \begin{pmatrix} 1 & 0 \\ 0 & e^{-2u/k} \end{pmatrix}$$

Considering the particular values of g_{ij} and g^{ij} we get $R_{1212} = g_{1k}R^k_{\ 212} = R^1_{\ 212}$

$$R^1_{\ 212} = \frac{\partial \Gamma^1_{22}}{\partial u^1} - \frac{\partial \Gamma^1_{21}}{\partial u^2} + \Gamma^r_{22}\Gamma^1_{r1} - \Gamma^r_{21}\Gamma^1_{r2} = \frac{\partial \Gamma^1_{22}}{\partial u} - \frac{\partial \Gamma^1_{21}}{\partial v} + \Gamma^1_{22}\Gamma^1_{11} + \Gamma^2_{22}\Gamma^1_{21} - \Gamma^1_{21}\Gamma^1_{12} - \Gamma^2_{21}\Gamma^1_{22}$$

Then

$$\Gamma^1_{22} = \frac{1}{2} g^{11}\left(\frac{\partial g_{21}}{\partial u^2} + \frac{\partial g_{21}}{\partial u^2} - \frac{\partial g_{22}}{\partial u^1}\right) = -\frac{1}{k}e^{2u/k}$$

$$\Gamma^1_{21} = 0 \qquad \Gamma^1_{11} = 0 \qquad \Gamma^2_{21} = \frac{1}{k} \qquad \frac{\partial \Gamma^1_{22}}{\partial u} = -\frac{2}{k^2}e^{2u/k}$$

$$R^1_{\ 212} = -\frac{2}{k^2}e^{2u/k} - \left(\frac{1}{k} \cdot -\frac{1}{k}e^{2u/k}\right) = -\frac{1}{k^2}e^{2u/k} \equiv R_{1212}$$

$$K = \frac{R_{1212}}{g} = -\frac{1}{k^2}\, e^{2u/k} \Big/ e^{2u/k} = -\frac{1}{k^2}$$

The metric of this example defines a surface of revolution named *pseudosphere*; it is obtained by *revolving a tractrix* [5] *about its asymptote,* and has the shape of a trumpet.

....

[5] Tractrix (from the Latin verb trahere "pull, drag"; plural: tractrices) is the curve along which an object moves, under the influence of friction, when pulled on a horizontal plane by a line segment attached to a tractor (pulling) point that moves at a right angle to the initial line between the object and the puller at an infinitesimal speed. It is therefore a curve of pursuit. It was first introduced by Claude Perrault in 1670, and later studied by Sir Isaac Newton (1676) and Christiaan Huygens (1692). The revolution of a tractrix about its asymptote produces the surface called pseudosphere. The name derives from the fact that the curvature is constant, as for the sphere, but has the opposite sign.

1.11 Minkoswki space

In Special Relativity, passing from one frame of reference to another, the infinitesimal distance between two events:

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - (\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2) \tag{27}$$

is preserved (= ***is invariant***). If we define $x^0 = ct$; $x^1 = x$; $x^2 = y$; $x^3 = z$ we can write

$$\mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta \qquad \text{with} \qquad \eta_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{28}$$

We have then the metric of Minkoswki space, which is "pseudo-Euclidean", but it is flat: in fact the $\eta_{\alpha\beta}$ are constant, therefore Γ^i_{jk} and R^h_{ijk} are zero. In the following we will use, by convention, the Greek indices α, β, γ, ... if these vary from 0 to 3, while we will use italic indices i, j, k, ... if they vary from 1 to 3. Warning: in literature also the opposite convention is used. Even $\eta_{\alpha\beta}$ is often defined with opposite signs, i.e. with the *signature* $(-1, 1, 1, 1)$ instead of $(1, -1, -1, -1)$.

Moreover, we say that the interval $\mathrm{d}s^2$ is:

- ***time-like*** if $\mathrm{d}s^2 > 0$ (corresponding to a physical trajectory with $v < c$)
- ***space-like*** if $\mathrm{d}s^2 < 0$
- ***light-like, null*** if $\mathrm{d}s^2 = 0$ (corresponding to the motion of particles, like photons, which move with speed $v = c$)

Each observer has with him a ruler and a clock: the time marked by this clock is the proper time τ. An observer, who sees two events (physically connected) occur at different times but at the same place ($\mathrm{d}x = \mathrm{d}y = \mathrm{d}z = 0$) obtains $\mathrm{d}s^2 = c^2\mathrm{d}\tau^2$: $\mathrm{d}s$ and $\mathrm{d}\tau$ are proportional.

The distance $\mathrm{d}s$ between the same two events, both for an observer who sees them occurring at the same point, and for another observer who sees them occurring at a distance $\mathrm{d}l$, is the same:

$$\mathrm{d}s^2 = c^2\mathrm{d}\tau^2 = c^2\mathrm{d}t^2 - |\mathrm{d}\bar{l}^2| \qquad \rightarrow \qquad \mathrm{d}\tau^2 = \mathrm{d}t^2\left(1 - \frac{1}{c^2}\frac{\mathrm{d}\bar{l}}{\mathrm{d}t}\cdot\frac{\mathrm{d}\bar{l}}{\mathrm{d}t}\right) = \mathrm{d}t^2\big(1 - \frac{v^2}{c^2}\big)$$

where v is the particle speed for the observer who sees it moving, and also the relative velocity between the two observers. Defining $\beta \equiv v/c$ and $\gamma \equiv 1/\sqrt{1-\beta^2}$ we get $\mathrm{d}t = \gamma\mathrm{d}\tau$. Since $\gamma \geq 1$, then $\mathrm{d}t \geq \mathrm{d}\tau$: the interval between two "ticks" of a clock is shorter for the “proper” clock; moving clocks appear slower (think about the twin paradox).

The velocity four-vector (**four-velocity**) is defined as $u^\alpha \equiv \frac{\mathrm{d}x^\alpha}{\mathrm{d}s}$; it is a vector since $\mathrm{d}x^\alpha$ is a vector and $\mathrm{d}s$ is a scalar.

In a generic reference frame, not at rest with a particle which has a *3D* velocity $\overline{v} \equiv \frac{\mathrm{d}\overline{x}}{\mathrm{d}t}$, we have

$$u^0 = \frac{\mathrm{d}x^0}{\mathrm{d}s} = \frac{\mathrm{d}(ct)}{c\mathrm{d}\tau} = \frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma$$
$$u^i = \frac{\mathrm{d}x^i}{c\mathrm{d}\tau} = \frac{1}{c}\frac{\mathrm{d}x^i}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}\tau} = \gamma\frac{v^i}{c} = \gamma\beta^i$$

and we can write $u^\alpha = \gamma(1, \overline{\beta})$. If the particle is at rest we have $u^\alpha = (1, 0, 0, 0)$.

The quantity $u^\alpha u_\alpha$ is invariant: $u^\alpha u_\alpha = \eta_{\alpha\beta}u^\alpha u^\beta = u^0u^0 - (u^1u^1 + u^2u^2 + u^3u^3) = \gamma^2 - (\gamma^2 v^2/c^2) = 1$; u^α is the unit vector (versor) tangent to the trajectory of the particle (in the 4-D space-time).

The **four-momentum** is defined as $P^\alpha = m_0 u^\alpha$where m_0 is the rest mass of the particle. If we remember that $\overline{P} = m\overline{v} = \gamma m_0\overline{v}$; $E = mc^2 = m_0c^2\gamma$ we get:

$$P^0 = \gamma m_0 = E/c^2 \qquad\qquad P^i = \gamma m_0\frac{v^i}{c} = m\frac{v^i}{c}$$

$$P^\alpha P_\alpha = \gamma^2 m_0^2 - \gamma^2 m_0^2 \frac{v^2}{c^2} = \gamma^2 m_0^2 \left(1 - \frac{v^2}{c^2}\right) = m_0^2$$

$$P^\alpha P_\alpha = m_0^2 = \frac{E^2}{c^4} - \frac{1}{c^2}\overline{P}\cdot\overline{P} \qquad \rightarrow \qquad m_0^2 c^2 = \frac{E^2}{c^2} - |\overline{P}|^2$$

If $\frac{dP^\alpha}{ds} = 0$ then $P^\alpha = const \Rightarrow E = const$ and $\overline{P} = const$: this is the energy and momentum conservation for a single particle.

The **four-acceleration** is $\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}s^2} = \frac{\mathrm{d}u^\alpha}{\mathrm{d}s}$. The **geodesic** equation has always the same form:

$$\frac{\mathrm{d}^2 x^\alpha}{\mathrm{d}s^2} + \Gamma^\alpha_{\beta\gamma}\frac{\mathrm{d}x^\beta}{\mathrm{d}s}\frac{\mathrm{d}x^\gamma}{\mathrm{d}s} = 0 \tag{29}$$

If the metric tensor is simply $\eta_{\alpha\beta}$, then the $\Gamma^\alpha_{\beta\gamma}$ vanish, so that $\mathrm{d}^2x^\alpha/\mathrm{d}s^2 = 0$, i.e. $x^\alpha = a^\alpha \cdot s + b^\alpha$, or

$$\begin{cases} ct = a^0 \cdot s + b^0 \\ \overline{x} = \overline{a}\cdot s + \overline{b} \end{cases}$$

and the trajectory is a straight line covered with uniform, rectilinear motion. In a curved space-time geodesics correspond to “straight lines”.

While in 3-D Euclidean space the geodesic between two points is a straight line, so it is the shortest distance between two points, in Minkovski geometry the quantity $\int_A^B \mathrm{d}s$ is a maximum on a time-like (physical) trajectory (remember that $\Delta\tau = \Delta s/c$, and you can think at the twin paradox, where the elapsed time is maximized for the twin who remained on the Earth), while it is a minimum on a space-like line.

1.12 The energy-momentum tensor

To deal with General Relativity and Cosmology we need an "object" that describes enegy and momentum conservation for a continuous medium. We know from analytical mechanics that, for a physical system, invariance with respect to time (i.e. $\partial/\partial x^0 = 0$) corresponds to energy conservation, while invariance with respect to space (i.e. $\partial/\partial x^i = 0$) corresponds to momentum conservation. Since we are working within Special Relativity, the invariance will be expressed as $\partial/\partial x^\alpha = 0$, $\alpha = 0, 1, 2, 3$. You probably met such a kind of object in Quantum Field Theory: the so-called stress tensor or energy-momentum tensor. Let’s look for a similar object in the case of a fluid.

Let’s consider first the case of incoherent matter, whose particles (for the moment) do not interact (*"dust"*). The matter field will be described at any point by the four-velocity $u^\alpha = \gamma(1, \overline{v}/c)$ and by its *proper density* $\rho_0(x)$, i.e. that measured by an observer who follows the fluid. With these quantities one can form a symmetric tensor of rank 2 in the simplest way as:

$$T^{\alpha\beta} = \rho_0 c^2 u^\alpha u^\beta \tag{30}$$

Let’s see how this tensor is made in detail:

$$T^{00} = \rho_0 c^2 \gamma^2 = \gamma^2 \rho_0 c^2 = \rho c^2 \qquad \text{by writing} \quad \rho = \gamma^2 \rho_0$$

To interpret this result remember that the mass is $m = \gamma m_0$ (m_0= rest mass) and that a volume element in motion appears contracted by a factor $1/\gamma$, and its density grows by another factor γ. So if the proper density is ρ_0, an observer with respect to which the fluid has velocity $\overline{v}$ measures a density $\gamma^2\rho_0$.

T^{00} represents the mass-energy density (in this case the only contribution to the energy comes from matter motion).

The components of $T^{\alpha\beta}$ can be written:

$$T^{\alpha\beta} = \rho\, c^2 \cdot \begin{pmatrix} 1 & v_x/c & v_y/c & v_z/c \\ v_x/c & v_x^2/c^2 & v_x v_y/c^2 & v_x v_z/c^2 \\ v_y/c & v_y v_x/c^2 & v_y^2/c^2 & v_y v_z/c^2 \\ v_z/c & v_z v_x/c^2 & v_z v_y/c^2 & v_z^2/c^2 \end{pmatrix} \tag{31}$$

We now derive the motion equations from the expression $\partial_\beta T^{\alpha\beta} = 0$, the four-divergence of $T^{\alpha\beta}$(remember we are in Minkowski space-time, and covariant derivatives are simply partial derivatives).

- For $\alpha = 0$ we have $\partial_\beta T^{0\beta} = 0 \quad \Leftrightarrow \quad \frac{\partial T^{0\beta}}{\partial x^\beta} = 0$ which can be expanded:

$$\frac{1}{c}\frac{\partial(\rho c^2)}{\partial t} + \frac{\partial(\rho c v_x)}{\partial x} + \frac{\partial(\rho c v_y)}{\partial y} + \frac{\partial(\rho c v_z)}{\partial z} = 0$$

  and then simplified to

$$\frac{\partial \rho}{\partial t} + \overline{\nabla} \cdot (\rho \overline{v}) = 0 \tag{32}$$

  which is the *continuity equation* for a fluid, expressing mass-energy conservation.

- For $\alpha = 1, 2, 3$ we have

$$\frac{1}{c}\frac{\partial(\rho c v_x)}{\partial t} + \frac{\partial(\rho v_x v_x)}{\partial x} + \frac{\partial(\rho v_x v_y)}{\partial y} + \frac{\partial(\rho v_x v_z)}{\partial z} = 0 \qquad (\alpha = 1)$$

$$\frac{1}{c}\frac{\partial(\rho c v_y)}{\partial t} + \frac{\partial(\rho v_y v_x)}{\partial x} + \frac{\partial(\rho v_y v_y)}{\partial y} + \frac{\partial(\rho v_y v_z)}{\partial z} = 0 \qquad (\alpha = 2)$$

$$\frac{1}{c}\frac{\partial(\rho c v_z)}{\partial t} + \frac{\partial(\rho v_z v_x)}{\partial x} + \frac{\partial(\rho v_z v_y)}{\partial y} + \frac{\partial(\rho v_z v_z)}{\partial z} = 0 \qquad (\alpha = 3)$$

If we multiply the first by $\hat{\imath}$ (unit vector of the x-axis), the second by $\hat{\jmath}$ and the third by $\hat{k}$ and then add them toghether they can be summarized in the expression

$$\frac{\partial}{\partial t}(\rho \overline{v}) + \frac{\partial}{\partial x}(\rho v_x \overline{v}) + \frac{\partial}{\partial y}(\rho v_y \overline{v}) + \frac{\partial}{\partial z}(\rho v_z \overline{v}) = 0$$

which, by expanding and by using continuity equation, becomes

$$\rho\frac{\partial \overline{v}}{\partial t} + \overline{v}\left[\frac{\partial \rho}{\partial t} + \overline{\nabla} \cdot (\rho \overline{v})\right] + \rho v_x \frac{\partial \overline{v}}{\partial x} + \rho v_y \frac{\partial \overline{v}}{\partial y} + \rho v_z \frac{\partial \overline{v}}{\partial z} = 0$$

$$\text{that is} \qquad \rho\left[\frac{\partial \overline{v}}{\partial t} + \left(\overline{v} \cdot \overline{\nabla}\right)\overline{v}\right] = 0 \quad (I) \quad \Leftrightarrow \quad \rho\frac{\mathrm{d}\overline{v}}{\mathrm{d}t} = 0 \quad (II) \tag{33}$$

This equation, typical of fluid dynamics, is the motion equation for a fluid without pressure, viscosity and external forces. Therefore it expresses the conservation of momentum. In particular, in the form (I) one imagines to observe the fluid at a fixed point and to see how its motion evolves (the so-called *Eulerian* point of view), while in the form (II) one imagines to follow in their motion the particles of fluid (the so-called *Lagrangian* point of view).

Thus we see that the tensor $T^{\alpha\beta}$ expresses the energetic and dynamical properties of the fluid (dust, in this case). $T^{\alpha\beta}$ is the **stress-energy tensor** or the **energy-momentum tensor**.

In a *locally inertial frame at rest (LIRF)* with respect to the fluid, in which $u^\alpha = (1, 0, 0, 0)$, $T^{\alpha\beta}$ has the particularly simple form

$$T^{\alpha\beta}_{LIRF} = \begin{pmatrix} \rho_0 c^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

We now come to consider the case in which the particles interact in the simplest way, that is through collisionss due to their thermal motion: in this case the fluid has a *pressure*. We assume that there is no transport of energy by conduction or radiation and there is no viscosity. The **fluid** so defined is said to be **perfect**.

If we are now in the *LIRF*, $T^{\alpha\beta}$ will be no more that one written just above, with only $T^{00} \neq 0$. The particles now have random motions around the zero of their positions and velocities. We must then refer back to the previous form (Eq. (31)) of $T^{\alpha\beta}$, in which however the terms that appear will be mediated on the distribution of particle velocity.

But this gives us immediately an important information: all the off-diagonal terms contain elements as v_x, v_y or v_z or their products; when we average $\langle v_x \rangle = 0$ and also $\langle v_x v_y \rangle = \langle v_x \rangle \langle v_y \rangle = 0$ (assuming that v_x and v_y are not correlated). Then $T^{\alpha\beta}$ is diagonal in the *LIRF*.

T^{00}_{LIRF} (expressing the mass-energy density) will be no longer $\rho_0 c^2$, but rather ρc^2, with $\rho > \rho_0$ to take account of the fact that the particles have velocities different from zero even in *LIRF* and their mass-energy density is greater than in the case of pure dust. For the other diagonal terms we have $\langle \rho v_x^2 \rangle, \langle \rho v_y^2 \rangle, \langle \rho v_z^2 \rangle$, which represent the pressure p (also for a relativistic gas).

So, for a perfect fluid:

$$T^{\alpha\beta}_{LIRF} = \begin{pmatrix} \rho c^2 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \tag{34}$$

where ρ takes into account also the mass-energy due to thermal motions.

It's easy to check that, in the *LIRF*, all this can be summarized in the relation

$$T^{\alpha\beta}_{LIRF} = (p + \rho c^2) u^\alpha u^\beta - p\eta^{\alpha\beta}$$

For instance, for T^{00}, by considering that $u^0 = 1$ and $\eta^{00} = 1$, we get $T^{00} = p + \rho c^2 - p = \rho c^2$

But this expression is a tensor, and then will hold in any frame of reference, with $u^\alpha \neq (1, 0, 0, 0)$ and the appropriate metric tensor instead of $\eta^{\alpha\beta}$:

$$T^{\alpha\beta} = (p + \rho c^2) u^\alpha u^\beta - p g^{\alpha\beta} \tag{35}$$

In this case the relation

$$\partial_\beta T^{\alpha\beta} = \left[(p + \rho c^2) u^\alpha u^\beta - p\eta^{\alpha\beta} \right]_{,\beta} = 0$$

leads, for $\alpha = 0$, to

$$\frac{\partial}{\partial t} \left[(p + \rho c^2) \gamma^2 \right] + \nabla \cdot \left[(p + \rho c^2) \gamma^2 \overline{v} \right] = \frac{\partial p}{\partial t} \tag{36}$$

If the overall motion of particles is not relativistic, we have $\gamma \approx 1$ and if $p \ll \rho c^2$(as for non relativistic matter) this relation reduces again to continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \left[\rho \overline{v} \right] = 0.$$

For $\alpha = 1, 2, 3$, in a way similar to that followed in the *"dust"* case, we get

$$\left(\frac{p}{c^2} + \rho \right) \gamma^2 \frac{\mathrm{d}\overline{v}}{\mathrm{d}t} = -\left[\overline{\nabla} p + \overline{v} \frac{\partial (p/c^2)}{\partial t} \right] \tag{37}$$

which is a generalization of the fluid-dynamics relation $\rho \frac{\mathrm{d}\overline{v}}{\mathrm{d}t} = -\overline{\nabla} p$ (*the so-called Euler equation*). As one can see, $(\rho + p/c^2)$ plays the role of "inertial mass density".

Within the perfect fluid there is no exchange of energy by conduction (or radiation), nor is there dissipation. From the first law of thermodynamics, in the frame that follows the fluid and for a volume V, $\mathrm{d}Q = \mathrm{d}U + p\mathrm{d}V$ and $U = \rho c^2 \cdot V$. Then

$$\mathrm{d}Q = \rho c^2 \mathrm{d}V + V\mathrm{d}(\rho c^2) + p\mathrm{d}V = (p + \rho c^2)\mathrm{d}V + V\mathrm{d}(\rho c^2) = T\mathrm{d}S \tag{38}$$

Since $\mathrm{d}Q = 0 \quad \rightarrow \quad \mathrm{d}S = 0$, the entropy of the volume V is conserved.

If we write the equation of state as $p = w\rho c^2$ (with w constant, although, in general, it may be that $w = w(T)$),

$$(1 + w)\rho c^2 \mathrm{d}V = -V \ \mathrm{d}(\rho c^2) \tag{39}$$

and, if $w = const$, we have $\mathrm{d}\rho/\rho = -(1 + w) \ \mathrm{d}V/V$, that is $\rho V^{1+w} =$ constant.

We will meet three interesting cases in cosmology:

1. For a non-relativistic gas $p \ll \rho_0 c^2$ ($\rho \approx \rho_0$) so that $w \simeq 0$ and $\rho_0 V \simeq const.$ If L is the edge of a cubic volume $V = L^3$, we obtain $\rho \propto 1/L^3$

2. For a gas of photons (and in general for a relativistic gas) $\rho_{rad} \propto aT^4$ and $p = \frac{1}{3}\rho c^2$; $w = \frac{1}{3}$:

$$T^4 V^{4/3} = const \qquad TV^{1/3} = const \qquad V \propto L^3 \quad \rightarrow \quad T \propto \frac{1}{L}$$

$$\rho_{rad} V^{4/3} = const \qquad V \propto L^3 \quad \rightarrow \quad V^{4/3} \simeq L^4 \quad \rightarrow \quad \rho_{rad} \simeq \frac{1}{L^4}$$

3. If $p = -\rho c^2 \qquad (w = -1) \quad \rightarrow \quad \rho V^0 = const$ that is ρ does not depend on V and L and remains constant if V changes.

We can express the first principle in another useful way by writing $V \propto L^3$

$$\left(\rho + \frac{p}{c^2}\right)\mathrm{d}V + V\mathrm{d}\rho = 0 \qquad \rightarrow \qquad \left(\rho + \frac{p}{c^2}\right) \cdot 3L^2\mathrm{d}L + L^3\mathrm{d}\rho = 0$$

which gives

$$3\left(\rho + \frac{p}{c^2}\right)\frac{\mathrm{d}L}{L} + \mathrm{d}\rho = 0$$

and, taking into account a possible dependence of L on time,

$$3\left(\rho + \frac{p}{c^2}\right)\frac{\dot{L}}{L} + \dot{\rho} = 0 \qquad (40)$$

We wrote $\partial_\alpha T^{\beta\alpha} = 0$ in Minkowski space; but, if the $\Gamma^\alpha_{\beta\gamma}$ do not all vanish, and this is the general case, instead of the simple partial derivative we must use the covariant derivative:

$$T^{\alpha\beta}{}_{;\beta} = 0$$

that expresses the conservation laws in a generic frame of reference.

1.13 Mach principle

According to Newton's dynamics the inertial properties of a body depend on its motion with respect to absolute space. *Ernst Mach*[6] suggested instead that the inertia is related to the motion with respect to the total distribution of matter in the universe. The motion is only relative to other bodies: operationally we can only measure the motion of matter in relation to other matter, not with respect to the absolute space of Newton. If there were only one body in the universe, its motion would not be defined: without other matter we can not say if this body is at rest or is accelerating. And since the reaction of matter to the acceleration is the only way to determine the inertia, this body does not possess a well defined inertia. The idea that masses and positions of celestial bodies define the inertia and inertial systems is called **Mach principle**. Several objections can be moved to this idea: for instance, no observer can be in an empty universe and verify the ideas of Mach, and inertia may exist even in an empty universe.

Anyway, the ideas of Mach influenced, by his own admission, Einstein himself. According to Newtonian physics, in a volume without interactions, the bodies should remain at rest or move with uniform motion. But since the universe is permeated by gravitational fields that can not be shielded, all bodies move along curved paths due to these fields. Then a question arises: if we say that a path is curved, we assume that we know how to define a straight line. But how can we do this if no body, not even photons, as we shall see, follows a straight line? So we try to do without the concept of straight line, and assume that there are no physical entities such as "gravitational forces" curving the trajectories of heavenly bodies, but that the geometry of space is modified by gravitation in such a way that the trajectories observed correspond to free, inertial motions of bodies. In a more dramatic way we could say that **space-time is the gravitational field**[7]. But how do we express this link between inertial motion and gravitation?

Special Relativity can be described by a geometry of Minkowski $\mathrm{d}s^2 = \eta_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta$and, from the properties of invariance of $\mathrm{d}s^2$ between inertial systems, we derive all the results of this theory (time dilation, length contraction, ...). How do we move to a metric $\mathrm{d}s^2 = g_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta$ in the presence of a gravitational field? What are the links between $g_{\alpha\beta}$ and the gravitational field, and between $g_{\alpha\beta}$ and the gravitation according to Newton? If seen in this way, General Relativity turns out to be a geometrical theory of gravitation.

[6] *Ernst Mach (1838-1916) was professor of physics and then philosophy at the University of Vienna. His ideas have had a precursor in the English bishop and philosopher George Berkeley, in 1710, when Newton was still alive.*

[7] This is the point of view of *Loop Quantum Gravity*, one of the major candidates for a quantum theory of gravitation.

1.14 Locally inertial frames

If our aim is to give a geometric description of space-time, we can now use what we learned about surface elements and generalize it to 4 dimensions. In particular, we have seen that, in a neighborhood of a generic point, one can transform $g_{\alpha\beta}$ in such a way that it can assume the required form and its first derivatives are zero. So, in the neighborhood of an event E, we can always write $g_{\alpha\beta} \equiv \eta_{\alpha\beta} + \mathcal{O}(|x|^2)$, where x represents the displacement in space-time from E: to the first order the geometry is the same as that of Special Relativity. In the (infinitesimal) neighborhood of each event the laws of physics are those that hold in an inertial frame of reference. In a neighborhood of each event we can define a **locally inertial reference frame**.

In the presence of gravitational fields, as mentioned above, local deviations from Special Relativity occur only at the level of the second derivatives of $g_{\alpha\beta}$ which, remember, are related to the curvature tensor $R^{\alpha}_{\beta\gamma\delta}$. In this sense **gravity curves space-time**. But what are these locally inertial reference frames?

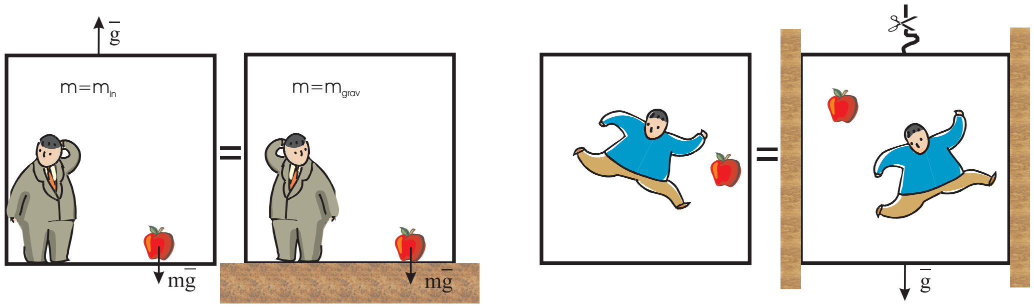
1.15 The Principle of Equivalence

The evidence that all bodies fall (in the absence of air resistance) in the same way under the effect of gravity, led to conclude, with great precision, that inertial mass m_{in} and gravitational mass m_{grav} are mutually proportional (and are, in practice, the same, by including the constant of proportionality within the gravitational constant G). Einstein assumed that, *by definition*, $m_{in} \equiv m_{grav}$. This leads to the famous thought experiment of **Einstein elevator**: an observer, equipped with scientific instruments and locked up into an elevator without the possibility to see what is happening around him, will not be able to distinguish, by his experiments in mechanics, between the two situations:

- he is at rest in a gravitational field with gravitational acceleration $\overline{g}$
- he is in empty space, and the elevator is accelerated upward with constant acceleration $\overline{g}$

Similarly, since all bodies fall in the same way in a gravitational field, the observer will not be able to distinguish between the situations of:

- uniform rectilinear motion in the vacuum
- free fall in a gravitational field



This allows us to say what are the *locally inertial frames: those in free fall.* Then, in a free falling frame, the laws of Special Relativity hold locally (and to the first order in $g_{\alpha\beta}$).

The Principle of Equivalence requires that all the laws of physics (not just those of mechanics) are the same both in a locally inertial frame and in Special Relativity.

Since the effects of gravitation disappear in a system in free fall, the phenomena occurring there are totally independent from the presence of nearby masses. However, according to the point of view of Mach, a large, nearby mass should introduce an anisotropy of the inertial mass. Effects due to the Sun or our Galaxy have been searched, but not found within $\Delta m/m \sim 10^{-20}$, for which the Principle of Equivalence seems favored over the assumptions of Mach (so they are not completely consistent with General Relativity, apart from the inspiration provided to Einstein[8]).

[8]Einstein conceived his theory of General Relativity trying to incorporate the idea of Mach according to which the inertia is due to gravitational interactions with all matter in the universe. But, as admitted by himself, he was only partially successful, since he obtained a solution of his field equations in which a single particle, immersed in a completely empty universe, had inertial properties.

1.16 The Principle of General Covariance

This principle tells us how to write the equations of physics in the presence of a gravitational field, when we know how they are made in the absence of gravity.

In order for an equation, expressing a physical law, applies in a gravitational field it is necessary that:

1. It is "covariant", i.e. does not change shape changing reference frame, and this happens when it is expressed as a relationship between tensors.

2. The equation applies in the absence of gravity, i.e. when $g_{\alpha\beta} \equiv \eta_{\alpha\beta}$ and $\Gamma^{\alpha}_{\beta\gamma} \equiv 0$.

There can be many covariant equations which reduce, in absence of gravity, to the same equation of Special Relativity. However, as the Principle of General Covariance and the Priciple of Equivalence operate on small scales, we expect that only $g_{\alpha\beta}$ and its low order derivatives come into play. This also obeys a principle of simplicity (*Occam's razor*).

In this way$T^{\alpha\beta}{}_{,\beta} \rightarrow T^{\alpha\beta}{}_{;\beta}$ (covariant derivative) or, for instance, for the free fall equation,

$$\frac{\mathrm{d}^2x^\alpha}{\mathrm{d}s^2} = 0 \rightarrow \frac{\mathrm{d}^2x^\alpha}{\mathrm{d}s^2} + \Gamma^{\alpha}_{\beta\gamma}\frac{\mathrm{d}x^\beta}{\mathrm{d}s}\frac{\mathrm{d}x^\gamma}{\mathrm{d}s} = 0$$

We have seen that at each point (event) we can define a locally inertial system, and in it the second derivatives of $g_{\alpha\beta}$ are in general $\neq 0$: i*t is therefore at the level of the second derivatives of the metric tensor that the gravitational field comes into play.*

Similarly, in Newtonian physics, in a system in free fall, what can be measured is the difference in gravitational acceleration between two bodies $\Delta g/\Delta x$. This is the kind of phenomenon we call **tide**. But $\overline{g} = -\overline{\nabla}\Phi_{grav}$ and then $\partial g/\partial x \propto \partial^2\Phi_{grav}/\partial x^2$. What can be measured are therefore the second derivatives of Φ_{grav}, as in General Relativity are the second derivatives of $g_{\alpha\beta}$. Then we see that there is an analogy between $g_{\alpha\beta}$ and Φ_{grav}: the $g_{\alpha\beta}$ take the place of the Newtonian gravitational potential.

1.17 The Einstein equations

In Newton's theory of gravitation the potential Φ satisfies Poisson equation: $\nabla^2\Phi = 4\pi G\rho_0$ and $\overline{g} = -\overline{\nabla}\Phi$. Special Relativity teaches us that all forms of energy are equivalent to mass, and then a relativistic theory of gravity will have as sources of the gravitational field all forms of energy, and not just ρ_0. In particular, the energy density of the gravitational field itself is proportional to $(\overline{\nabla}\Phi)^2$ in the Newtonian case [9](think, by analogy, that the energy density of the electromagnetic field is proportional to E^2).

If, therefore, we move to the left hand side, in Poisson equation, the term $\propto (\overline{\nabla}\Phi)^2$ which would result from the gravitational energy density, we obtain a non-linear differential equation (which is linear in the second derivatives and quadratic in the first ones) for the gravitational field.

Formally we will have an equation such as:

$$F(g_{\alpha\beta}) \sim \kappa T$$

where $g_{\alpha\beta}$ is the metric tensor (corresponding to Φ), F is a differential operator (likely something linear in the second derivatives and quadratic in the first derivatives) which reduces to ∇^2 when the field is weak[10], κ is a proportionality constant that contains the Newtonian gravitational constant G, T is a quantity that describes all forms of non-gravitational energy, and that, in the non-relativistic case, should essentially be reduced to ρ_0.

A natural candidate for T is the component T^{00} of the stress-energy tensor. But keeping as a source of the field only one component of a tensor would not produce an invariant theory: we should adopt a particular reference frame to calculate T^{00}. So Einstein conceived the idea that the source is the entire $T^{\alpha\beta}$: pressure, stresses (if

[9] We can see that the energy density of the gravitational field is proportional to $(\overline{\nabla}\Phi)^2$, that is to the square of the intensity of the gravitational field g^2, in the following way. The gravitational potential energy of a mass M of radius R is given by $E = -GM^2/R$. If you think that this energy is distributed in the field ($g \propto M/r^2$) created by M, between R and ∞, we see that, by calling δ_G the density of gravitational energy, for it to be $-\int_R^\infty \delta_G(r)\cdot 4\pi r^2 \mathrm{d}r \approx -M^2/R$, $\delta_G(r) \sim (M/r^2)^2 \sim g^2$ is required.

[10] We will see soon what does this mean; at the moment you can think that it is when Newton's law holds

$T^{\alpha\beta}$ is not diagonal), etc. ... , all acts as a source. But if T is a tensor, then the left-hand side of the equation must also be a tensor function of $g_{\alpha\beta}$.

We know that $T^{\alpha\beta}$ is symmetrical, and has vanishing covariant divergence: $T^{\alpha\beta}{}_{;\beta} = 0$. Then the left-hand side must share these properties. Moreover, we expect it to be linear in the second derivatives of $g_{\alpha\beta}$ and quadratic in the first derivatives.

We have already met a tensor with these properties, and we have seen that it is unique: the Einstein tensor

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R$$

So Einstein proposed, as a possible equation of the gravitational field,

$$R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} = \kappa T_{\alpha\beta} \tag{41}$$

If we think to include also the derivatives of order zero in the differential operator, being $g_{\alpha\beta}$ symmetric and with vanishing covariant divergence, we can add a term proportional to $g_{\alpha\beta}$and we get:

$$R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta} - \Lambda g_{\alpha\beta} = \kappa T_{\alpha\beta}, \tag{42}$$

where Λ and κ are constant; Λ is the so-called **cosmological constant**.

1.18 The Newtonian limit (weak field)

Once written Einstein's equations, we must check that, within the limits of validity of classical physics, they reduce to Newton's law; we must also find what is the constant κ that appears in the equations.

Let us suppose that the field is stationary (i.e. its time derivative is zero), the velocities of the particles are small ($v \ll c$) and that, at large distances from the masses that generate the field, the metric tensor is asymptotically flat: $g_{\alpha\beta} \to \eta_{\alpha\beta}$. We also assume that the field is weak: the deviations from metric $\eta_{\alpha\beta}$ are small:

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \qquad \text{with } |h| \ll 1 \tag{43}$$

By means of the geodesic equation it can be shown that

$$h_{00} \simeq \frac{2\Phi}{c^2} \qquad \to \qquad g_{00} \simeq 1 + \frac{2\Phi}{c^2} \tag{44}$$

where Φ is the Newtonian gravitational potential, The weak field hypothesis, $|h| \ll 1$, implies then that $|2\Phi/c^2| \ll 1$.

In the case of a mass M with density distributed with spherical symmetry, the external potential, at a distance r from the centre, is given by $\Phi = -GM/r$, according to Newton. The assumption that the field is weak, $|2\Phi/c^2| \ll 1$, implies that

$$\frac{2GM}{rc^2} \ll 1 \qquad \Rightarrow \qquad r \gg \frac{2GM}{c^2}$$

For a black hole or a generic spherical body, $R_S \equiv 2GM/c^2$ is the so-called *Schwarzschild radius*, corresponding, for a non-rotating and electrically neutral black hole, to the *event horizon*, the zone from which nothing can escape (apart from quantum effects of evaporation). In this case we see that the condition of weak field is

$$\frac{R_S}{r} \ll 1 \qquad \Rightarrow \qquad r \gg R_S$$

For our Sun, $R_S \sim 3\,\text{km}$.

Let's see now, with the same assumptions made above, that the Einstein equations reduce to Poisson equation $\nabla^2\Phi = 4\pi G\rho_0$. Then we will determine the value of the constant κ. Let us take Einstein equation with the term containing Λ brought to the right:

$$R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R = \kappa T_{\alpha\beta} + \Lambda g_{\alpha\beta}$$

If we multiply it by $g^{\alpha\gamma}$ we get:

$$R^{\gamma}_{\beta} - \frac{1}{2}R\delta^{\gamma}_{\beta} = \kappa T^{\gamma}_{\beta} + \Lambda\delta^{\gamma}_{\beta}$$

Let us put $\gamma = \beta$ (that is, we add on $\beta = \gamma = 0, 1, 2, 3$, $\delta^{\beta}_{\beta} = \delta^0_0 + \delta^1_1 + \delta^2_2 + \delta^3_3 = 1 + 1 + 1 + 1 = 4$) and contract tensors; since $R = R^{\gamma}_{\gamma}$ we get:

$$R - \frac{1}{2}R \cdot 4 = \kappa T^{\gamma}_{\gamma} + 4\Lambda \qquad \rightarrow \qquad R = -\kappa T^{\gamma}_{\gamma} - 4\Lambda$$

Substituting this result into the starting equation, it becomes:

$$R_{\alpha\beta} = \kappa T_{\alpha\beta} + \Lambda g_{\alpha\beta} + \frac{1}{2}g_{\alpha\beta}\big(-\kappa T^{\gamma}_{\gamma} - 4\Lambda\big) = \kappa\big(T_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}T^{\gamma}_{\gamma}\big) - \Lambda g_{\alpha\beta} \tag{45}$$

The Ricci tensor, if we neglect higher order terms, can be written (by using a suitable reference frame and some "tricks"):

$$R_{\beta\delta} \simeq -\frac{1}{2}\eta^{\alpha\sigma}\frac{\partial^2 h_{\beta\delta}}{\partial x^{\alpha}\partial x^{\sigma}} \tag{46}$$

We evaluate now the component 00 of $R_{\beta\delta}$:

$$R_{00} \simeq -\frac{1}{2}\eta^{\alpha\sigma}\frac{\partial^2 h_{00}}{\partial x^{\alpha}\partial x^{\sigma}}$$

but if $\alpha = 0$ and/or $\sigma = 0$ the derivative is zero (for stationarity $\partial/\partial x^0 = 0$); then only the terms with indices 1, 2, 3 (and $\eta^{11} = \eta^{22} = \eta^{33} = -1$) are left:

$$R_{00} \simeq \frac{1}{2}\left[\frac{\partial^2 h_{00}}{\partial x^1\partial x^1} + \frac{\partial^2 h_{00}}{\partial x^2\partial x^2} + \frac{\partial^2 h_{00}}{\partial x^3\partial x^3}\right] \simeq \frac{1}{2}\nabla^2 h_{00}$$

On the other hand

$$R_{00} = \kappa\left[T_{00} - \frac{1}{2}g_{00}T^{\gamma}_{\gamma}\right] - \Lambda g_{00}$$

The fully covariant energy-momentum tensor is

$$T_{\mu\nu} = g_{\mu\alpha}g_{\nu\beta}\,(p + \varrho c^2)u^{\alpha}u^{\beta} - g_{\mu\alpha}g_{\nu\beta}\,p\,g^{\alpha\beta} \tag{47}$$

We have assumed that motions are not relativistic, so $u^i \ll u^0 \simeq 1$, and

$$T_{\mu\nu} \cong g_{\mu 0}g_{\nu 0}(p + \varrho c^2) - p\,g_{\mu\nu} \cong \eta_{\mu 0}\eta_{\nu 0}(p + \varrho c^2) - p\,\eta_{\mu\nu}, \tag{48}$$

where we have neglected terms which are of first and second order in $h_{\mu\nu}$ with respect to $\eta_{\mu\nu}$. So we see that $T_{00} \cong \varrho c^2$. The tensor in the mixed form is

$$T^{\gamma}_{\nu} = g^{\gamma\mu}T_{\mu\nu} \cong \eta^{\gamma\mu}T_{\mu\nu} \cong \eta^{\gamma\mu}\eta_{\mu 0}\eta_{\nu 0}(p + \varrho c^2) - p\,\eta^{\gamma\mu}\eta_{\mu\nu} \cong \delta^{\gamma}_{0}\eta_{\nu 0}(p + \varrho c^2) - p\,\delta^{\gamma}_{\nu}. \tag{49}$$

If we now want the trace T^{γ}_{γ} we get finally

$$T^{\gamma}_{\gamma} \cong \delta^{\gamma}_{0}\eta_{\gamma 0}(p + \varrho c^2) - p\,\delta^{\gamma}_{\gamma} \cong p + \varrho c^2 - 4 \cdot p \cong \varrho c^2 - 3 \cdot p. \tag{50}$$

We then have

$$\frac{1}{2}\nabla^2 h_{00} \simeq \kappa\big[\rho c^2 - \frac{1}{2}\left(\varrho c^2 - 3 \cdot p\right)\big] - \Lambda \qquad \rightarrow \qquad \nabla^2 h_{00} \simeq \kappa(\rho c^2 + 3\,p) - 2\Lambda.$$

But $h_{00} = 2\Phi/c^2$ and then:

$$\nabla^2\Phi \simeq \frac{\kappa c^4}{2}(\rho + 3\,\frac{p}{c^2}) - \Lambda c^2 \tag{51}$$

Poisson equation tells that $\nabla^2\Phi = 4\pi g\rho_0$; if pressure and cosmological constant are negligible, and velocities are non relativistic, so that $\rho \cong \rho_0$, the two relations coincide if

$$\kappa = \frac{8\pi G}{c^4} \tag{52}$$

Finally we arrive to the complete Einstein equation

$$R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R - \Lambda g_{\alpha\beta} = \frac{8\pi G}{c^4}T_{\alpha\beta} \tag{53}$$

The above conditions require that $\varrho \gg 3p/c^2$ and, for Λ,

$$|\Lambda| \ll \frac{4\pi G\rho_0}{c^2} = \Lambda_E$$

In 1916, when Einstein derived the equations of General Relativity, he was not aware of cosmic expansion (still to be discovered), and sought a static solution for his model of universe. We see that if $\Lambda = \Lambda_E$ and ρ_0 is the density of the (homogeneous and static) universe, and p is negligible, from Eq. (51) we have $\nabla^2\Phi = 0$, $\Phi = const$, $\overline{g} = -\overline{\nabla}\Phi = 0$.

A similar result, as we shall see, comes from the equations of General Relativity. This static model, however, is unstable: just a small density fluctuation and locally we have expansion or contraction.

According to dimensional analysis $[c^2\Lambda] = [4\pi g\rho_0] = [\nabla^2\Phi]$ that is $[\Lambda] = [\nabla^2(\Phi/c^2)] = L^{-2}$(remember that Φ/c^2 is adimensional).

From the relation $\nabla^2\Phi = 4\pi G[\rho_0 + 3p/c^2 - c^2\Lambda/4\pi G]$ we can think that Λ corresponds to the mass-energy of vacuum, i.e. when mass-energy and pressure are removed.

After Hubble's discovery, in 1929, of the expansion of the universe, Einstein considered the cosmological constant as "the biggest blunder" of his life but, as we shall see, it has come back strongly in vogue in recent years.. For long time it was only possible to set upper limits to the value of Λ, just by assuming that there was no evidence of its effects. But recent observations (1997) based on Type Ia supernovae in distant galaxies, and the study of the cosmic microwave background, have allowed to obtain no longer an upper limit, but an estimate of $\Lambda \sim 10^{-56} cm^{-2}$.

1.19 Gravitational waves

We have seen, treating the weak field, that Ricci tensor can be written as in Eq. (46):

$$R_{\beta\delta} \simeq -\frac{1}{2}\eta^{\alpha\sigma}\frac{\partial^2 h_{\beta\delta}}{\partial x^\alpha \partial x^\sigma} = -\frac{1}{2}\left[\frac{1}{c^2}\frac{\partial h^2_{\beta\delta}}{\partial t^2} - \left(\frac{\partial^2 h_{\beta\delta}}{\partial x^2} + \frac{\partial^2 h_{\beta\delta}}{\partial y^2} + \frac{\partial^2 h_{\beta\delta}}{\partial z^2}\right)\right] \equiv -\frac{1}{2}\Box^2 h_{\beta\delta}$$

where $\Box^2$ is the d'Alembert operator or d'Alembertian.

We have seen that the Einstein equations can be written also as in Eq. (45). If we are in vacuum, and we neglect Λ, they become:

$$R_{\beta\delta} \equiv 0 \qquad \Rightarrow \qquad \Box^2 h_{\beta\delta} \equiv 0$$

which is the equation of **a wave propagating at the speed of light**. This shows the existence of **gravitational waves**. They are transverse waves with two components (polarizations).

1.20 Weak field metric and gravitational lenses

Let's use again Eq. (46) which, for a stationary field, simply becomes

$$R_{\beta\delta} \simeq \frac{1}{2}\nabla^2 h_{\beta\delta}. \tag{54}$$

If pressure and cosmological constant are negligible, and velocities are non relativistic, so that $\rho \cong \rho_0$, as before, $T_{\alpha\beta}$ has only one element different from zero: $T_{00} \cong \rho_0 c^2$, and $T^{\gamma}_{\gamma} \cong \rho_0 c^2$. We use now Eq. (45); for $\beta \neq \delta$ both $T_{\beta\delta}$ and $g_{\beta\delta}$ are equal to zero, and by using Eq. (54), Einstein equations reduces to

$$\nabla^2 h_{\beta\delta} = 0 \tag{55}$$

which is Laplace equation. If the solution is equal to zero at the infinity (as we have assumed), it is zero also everywhere: $h_{\beta\neq\delta} = 0$. For $\beta = \delta = 1$ we have

$$R_{11} = \frac{8\pi G}{c^4}\left[T_{11} - \frac{1}{2}\eta_{11}\cdot\rho_0 c^2\right] = \frac{4\pi G\rho_0}{c^2} = \frac{\nabla^2\Phi}{c^2} = \frac{1}{2}\nabla^2 h_{11}$$

which gives $h_{11} = 2\Phi/c^2$; the same holds for $h_{22} = h_{33} = 2\Phi/c^2$.

The complete metric, for a stationary weak gravitational field, is then

$$\mathrm{d}s^2 = \left(1 + \frac{2\Phi}{c^2}\right)c^2\mathrm{d}t^2 - \left(1 - \frac{2\Phi}{c^2}\right)\left(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\right) \tag{56}$$

This relation allows us to obtain another very interesting result. For a light ray $\mathrm{d}s^2 = 0$ and, assuming $\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \equiv \mathrm{d}l^2$, we have

$$\left(1 + \frac{2\Phi}{c^2}\right)c^2\mathrm{d}t^2 = \left(1 - \frac{2\Phi}{c^2}\right)\mathrm{d}l^2$$

from which

$$\left(\frac{\mathrm{d}l}{\mathrm{d}t}\right)^2 = c^2\frac{\left(1 + \frac{2\Phi}{c^2}\right)}{\left(1 - \frac{2\Phi}{c^2}\right)} \equiv v_{eff}^2 \equiv \frac{c^2}{n_g^2}$$

where v_{eff} is the *effective* speed of propagation of the luminous wave and n_g can be thought as an index of refraction of gravity. it is

$$n_g = \sqrt{\frac{\left(1 - \frac{2\Phi}{c^2}\right)}{\left(1 + \frac{2\Phi}{c^2}\right)}} \sim \sqrt{\left(1 - \frac{2\Phi}{c^2}\right)\left(1 - \frac{2\Phi}{c^2}\right)} \qquad \rightarrow \qquad n_g \simeq 1 - \frac{2\Phi}{c^2} \tag{57}$$

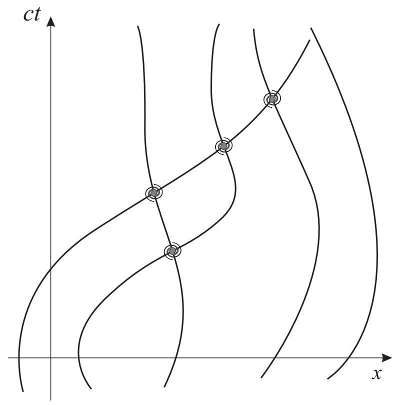
If $\Phi = 0$ to the infinity and is negative near a mass, $n_g > 1$ and $v_{eff} < c$. This relation show us that space, as a consequence of gravitation, behaves as a refractive medium: this is the basis of those phenomena known as **gravitational lenses**.

2 The Robertson-Walker metric

2.1 The cosmological principle

If we want to apply General Relativity (intended as the best available theory to describe the motion of bodies due to the distribution of matter) to the study of the cosmos, we would expect that, in general, the geometry of space-time is not static, but depends on time. This is also suggested by the observational evidence of a general motion of galaxies away from us (Hubble's law).

But if the world lines (trajectories in space-time) of galaxies, considered as "building blocks" of the universe and "tracers" of its evolution, were as they appear in the figure, things would be very complicated: there would be no order in the evolution, and there would be collisions where the lines cross.



Fortunately, astronomical observations are comforting. The expansion of the universe appears to be quite regular. Due to the presence of inhomogeneities (as groups, clusters of galaxies) there are perturbations in the motions of galaxies, gravitationally induced by these inhomogeneities. But these perturbations , corresponding to a speed on the order of $100 \div 1000\ km/s$, are "small" if compared to the recession velocity[11] of galaxies, which can be a significant fractions of the speed of light. These motions also appear generally not systematic.

It is also more and more evident that, on scales of the order of $100\, h^{-1} Mpc$, the universe is on average similar to itself. The high degree of isotropy of the microwave background radiation, which is of the order of 10^{-5}, and the not relevant position occupied by our Galaxy in the large-scale structure, make it reasonable to assume that this isotropy is not characteristic of our position, but is typical of every point in space: a further application of the so-called "Copernican" point of view, according to which the Earth no longer occupies the geometric center of the Universe. But ***isotropy around each point of space implies homogeneity***[12] (in an inhomogeneous space it would be impossible to get isotropy everywhere). We observe that, on the contrary, the reverse is not necessarily true: one can have a space which is homogeneous but not isotropic (as an example imagine a homogeneous universe, but in rotation around a certain spatial axis).

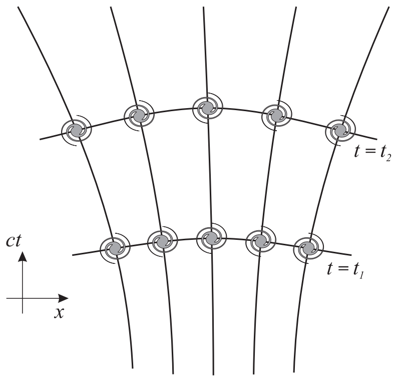
On the basis of what we said above, we can state the Cosmological Principle: "**At any given epoch the universe looks the same at every point, apart from local irregularities**".[13]

This principle makes it possible to greatly simplify the study of cosmology. We can imagine to smooth any irregularity and local motion, on scales of the order of $100\, h^{-1} Mpc$, obtaining a substrate that evolves in a uniform way everywhere, at a given time. We say "at a given time", suggesting the existence of a time definable unequivocally: since the universe is homogeneous at any point the physical parameters evolve in the same way (one can imagine to link the passing of time, at each point, to the local density of matter, equal densities correspond to equal times).

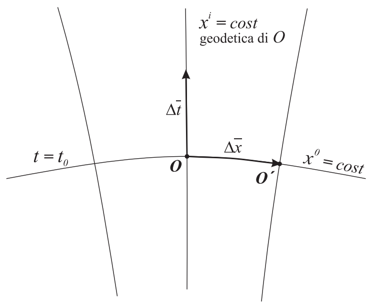
[11]Rather than a real speed, it is the rate of variation of the proper distance, as we shall see; but, albeit improperly (and dangerously), this term is used habitually, as we do in this page.

[12]This is true if the matter distribution is not fractal.

[13]There is also a *Perfect Cosmological Principle*, according to which the universe appears the same at every point and at all times, in which the density and the various cosmological parameters do not change with time. This principle has been implemented, in the so-called *Steady State* model of the universe, by *Hoyle*, *Bondi* and *Gold*.



Imagine that we fill this smoothed and homogeneous space with observers, all at rest with respect to the mean motion of the surrounding matter, each one with his own clock and ruler. The world lines (geodesics) of these observers do not intersect, except possibly at a singular point in the past and, perhaps, in the future. There is only one geodesic passing through a point in space-time, and then the material possesses, at each point, a well-defined speed. This smoothed substrate behaves as a perfect fluid. The regularity of the motion of observers (*Weyl's postulate*) allows us to define, for each value of the cosmic time, a **spatial section** $t = cost$ of space-time. These spatial sections are perpendicular to the geodesic followed by observers (see below).



In fact, if we consider one of these observers O at rest with respect to the mean motion of local matter, its geodesic is defined for him by the conditions $x^i = cost \;\; (i = 1, 2, 3)$; if we consider a close observer, which is on the same surface $t = t_0 = cost$, i.e. $x^0 = cost$, of O, the vector $\Delta\overline{x}$ connecting the event O to the event O' is normal to the vetctor $\Delta\overline{t}$ parallel to the geodesic for O and to the four-velocity with components $(1, 0, 0, 0)$. If $\Delta\overline{t} \cdot \Delta\overline{x} \neq 0$ the events O and O' would no longer be simultaneous, since $\Delta\overline{x}$ would have a non-zero component along the time axis of O.

This allows us to simplify the choice of the metric for the observer O, that will be in general

$$\mathrm{d}s^2 = g_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta$$

But, since vectors of the type $(1, 0, 0, 0)$ (like $\Delta\overline{t}$) and vectors of the type $(0, 1, 0, 0)$ like $\Delta\overline{x}$ are perpendicuolar, their dot procut is:

$$\Delta\overline{t} \cdot \Delta\overline{x} = 0 = g_{0i}\Delta t^0 \Delta x^i \qquad \forall \Delta t^0,\ \forall \Delta x^i \quad \Rightarrow \quad g_{0i} = 0$$

and the metric will be

$$\mathrm{d}s^2 = g_{00}(\mathrm{d}x^0)^2 + g_{ij}\mathrm{d}x^i \mathrm{d}x^j \qquad (i, j = 1, 2, 3)$$

We recall that, for the chosen type of observers, at rest with respect to the mean motion of the local universe, the spatial components $\mathrm{d}x^i$ are zero, i.e. $x^i = const$: the values of spatial coordinates assigned to the observer remain constant over time. These coordinates are called **co-moving**. The fact that the coordinates of the observers are constant does not imply that mutual distances are constant, since $g_{\alpha\beta}$ depend, in general, also on time.

Let us consider a co-moving observer O. His spatial coordinates are $x^i = cost$, and so $\mathrm{d}x^i = 0$; the interval $\mathrm{d}s^2$ between two two successive events along the world line of O is then $\mathrm{d}s^2 = g_{00}(\mathrm{d}x^0)^2$, but this is also equal, by definition, to $c^2\mathrm{d}\tau^2$, where τ is the proper time associated to O:

$$c^2\mathrm{d}\tau^2 = g_{00}(\mathrm{d}x^0)^2 \tag{58}$$

Space is homogeneous, and this relationship must hold for any observer, no matter what its coordinates x^i are, so g_{00} must depend only on x^0. W can then define a new scale of cosmic time such that

$$c\mathrm{d}t = \sqrt{g_{00}}\mathrm{d}x^0$$

that coincides with the proper time of co-moving observers and we write, using t to denote the proper time,

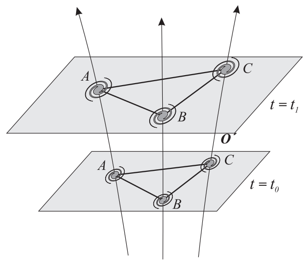
$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 + g_{ij}\mathrm{d}x^i\mathrm{d}x^j \tag{59}$$

A reference frame in which $g_{00} \equiv 1$ and $g_{0i} \equiv 0$ is named *synchronous*. In this case the world lines $x^i = const.$ are geodesic lines. In fact, the four-vector tangent to the world line $u^\alpha \equiv dx^\alpha/ds$ has the components equal to (1, 0, 0, 0) and automatically satisfies the geodesics equation because

$$\frac{du^\alpha}{ds} + \Gamma^\alpha_{\beta\gamma}u^\beta u^\gamma \simeq \Gamma^\alpha_{00}$$

but, since $g_{00} = 1 = const.$ and $g_{0i} = 0$,

$$\Gamma^\alpha_{00} = \frac{1}{2}g^{\alpha\sigma}\left(\frac{\partial g_{\sigma 0}}{\partial x^0} + \frac{\partial g_{\sigma 0}}{\partial x^0} - \frac{\partial g_{00}}{\partial x^\sigma}\right) = 0$$



Let us consider now a triangle formed by three particles both at time $t = t_0$ and at a later time $t = t_1$. The two triangles will not generally be the same, but will necessarily be similar, because if it were not so there would exist inhomogeneity and/or anisotropy in the universe, in contrast to the Cosmological Principle. From this it follows that the dependence of the g_{ij} from time must be similar for all of them, and does not depend on the spatial coordinates. So it will be:

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 + a(t)^2 g_{ij}\mathrm{d}x^i\mathrm{d}x^j \tag{60}$$

where the dependence on time is all in the function $a(t)$ named **scale factor**, and the g_{ij} do not depend on time. The ratio $a(t_1)/a(t_0)$ gives the enlargement at time t_1, with respect to time t_0, of a length (as the side of a triangle) measured along the two surfaces $t = t_1$ and $t = t_0$.

Notice that, for events whic lie on a surface $t = cost$ ($\mathrm{d}t = 0$), and are therefore simultaneous, $\mathrm{d}s^2$ will be *space-like* and then < 0, i.e. $a^2(t)g_{ij}\mathrm{d}x^i\mathrm{d}x^j < 0$; if we write $\tilde{g}_{ij} = -g_{ij}$ then $a^2(t)\tilde{g}_{ij}\mathrm{d}x^i\mathrm{d}x^j > 0$ and

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\tilde{g}_{ij}\mathrm{d}x^i\mathrm{d}x^j \tag{61}$$

similar to $\mathrm{d}s^2 = c^2\mathrm{d}t^2 - (\mathrm{d}\bar{l})^2$ (as we used to write in Special Relativity), where $\mathrm{d}l^2$ is the interval in a 3-D space with $t = const = t_0$

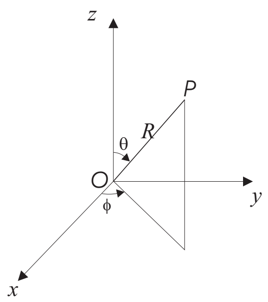
$$\mathrm{d}l^2 = a^2(t_0)\tilde{g}_{ij}\mathrm{d}x^i\mathrm{d}x^j \qquad i = 1, 2, 3 \tag{62}$$

We must now define $\tilde{g}_{ij}$ for an isotropic and homogeneous 3D space.

2.2 The Robertson-Walker metric

We use first of all the fact that we have a spherical symmetry due to isotropy, and we choose a spherical coordinate system, which reflects this symmetry. We remain, for now, in the Euclidean space and define:

$$\begin{cases} x = R\sin\theta\cos\phi \\ y = R\sin\theta\sin\phi \\ z = R\cos\theta \end{cases}$$



The metric tensor can be easily derived (we extend to three dimensions what we learned about surface elements):

$$\begin{aligned}\overline{x}_R &= (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)\\ \overline{x}_\theta &= (R\cos\theta\cos\phi, R\cos\theta\sin\phi, -R\sin\theta)\\ \overline{x}_\phi &= (-R\sin\theta\sin\phi, R\sin\theta\cos\phi, 0)\end{aligned}$$

$$^3g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & R^2 & 0 \\ 0 & 0 & R^2\sin^2\theta \end{pmatrix}$$

$$\mathrm{d}l^2 = \mathrm{d}R^2 + R^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2) = \mathrm{d}R^2 + R^2\mathrm{d}\Omega^2$$

For $R = const$ we have $\mathrm{d}l^2 = R^2\mathrm{d}\Omega^2$ with $^2g_{ij} = \begin{pmatrix} R^2 & 0 \\ 0 & R^2\sin^2\theta \end{pmatrix}$. Remember that the area of a surface can be obtained by using Eq. (8):

$$\mathrm{d}A = \sqrt{^2g}\ \mathrm{d}\theta\mathrm{d}\phi = R^2\sin\theta\mathrm{d}\theta\mathrm{d}\phi$$

so the area of the sphere is

$$A = \int_0^{2\pi}\int_0^{\pi} R^2\sin\theta\mathrm{d}\theta\mathrm{d}\phi = 4\pi R^2$$

This relation can be generalized to three dimensions obtaining in this case a volume:

$$\mathrm{d}V = \sqrt{^3g}\ \mathrm{d}R\mathrm{d}\theta\mathrm{d}\phi = R^2\sin\theta\mathrm{d}R\mathrm{d}\theta\mathrm{d}\phi$$

and then

$$V = \frac{4}{3}\pi R^3$$

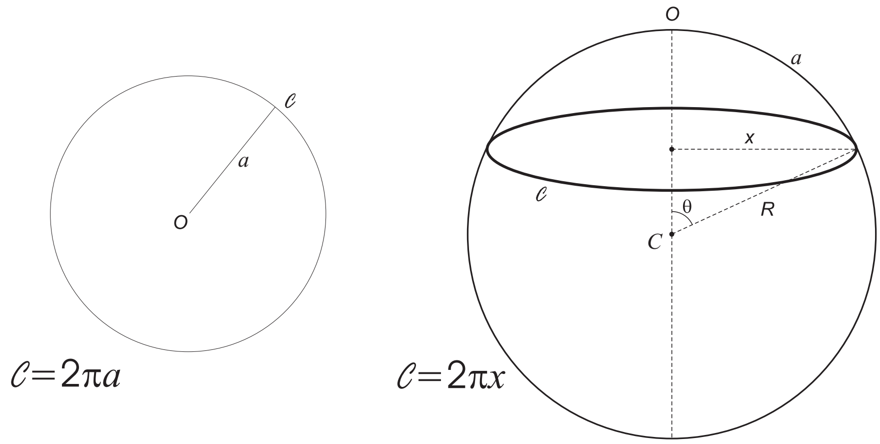
All this is true in Euclidean space, in which the area of the sphere is $4\pi\times$(the coefficient of $\mathrm{d}\Omega^2$). In a generic, non-Euclidean space, but spherically symmetric, the deviation from Euclidean space will be felt (because of isotropy) only in the radial direction. Each point will be on a two-dimensional spherical surface, whose line element is

$$\mathrm{d}l^2 = g(r')(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2) = g(r')\mathrm{d}\Omega^2$$

where $g(r')$ is a function of the third, radial coordinate, which we named r'. On the sphere $g(r') = const.$ and the area of the sphere is $4\pi \times g(r')$.

Since we can freely define the coordinate system, we redefine the radial coordinate r so that $r^2 \equiv g(r')$, with a transformation $r' \to r$. So again, a surface $r = cost$ has area $4\pi r^2$, but now r no longer corresponds to the proper, radial distance (we imagine to freeze expansion and to measure it with a ruler, at $t = const.$) from the center of the spherical surface, although of course there will be a link between the two variables.

We can see it in an intuitive way by reducing the dimensions of space from 3 to 2, and by speaking of a circumference instead of an area. In the Euclidean plane we have $C = 2\pi a$, but on a sphere $C = 2\pi f(a) = 2\pi x$ with $x = f(a) = R\sin\theta = R\sin(a/R)$. Now x is not the proper distance of the circle from the center, which is equal to a, but it is a coordinate like the others, and we can use it if useful.



Let's go back to the sphere in 3D space, with r as radial coordinate (linked to proper distance, but not coincident with it). The line element, in space, is then

$$\mathrm{d}l^2 = g_{rr}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \qquad (\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)$$

Terms like $g_{r\theta}$ or $g_{r\phi}$ are equal to zero since the radial coordinate r is, due to the imposed spherical symmetry, perpendicular to the surfaces $r = const$.

Going now back to space-time we have

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a(t)^2\left[f(r)\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2\right]$$

with the unknown function $f(r)$ to be determined. We fix the time coordinate: $t = t_0$, so that $a(t) = a(t_0) = a_0 = const$; the metric tensor of the spatial part is (in spherical coordinates r, θ, ϕ):

$$g_{ij} = \begin{pmatrix} a^2f(r) & 0 & 0 \\ 0 & a^2r^2 & 0 \\ 0 & 0 & a^2r^2\sin^2\theta \end{pmatrix} \qquad g^{ij} = \begin{pmatrix} \frac{1}{a^2f(r)} & 0 & 0 \\ 0 & \frac{1}{a^2r^2} & 0 \\ 0 & 0 & \frac{1}{a^2r^2\sin^2\theta} \end{pmatrix} \qquad g = a^6f(r)r^4\sin^2\theta$$

We want to impose the condition that space is homogeneous, which means that the curvature of space is constant everywhere. We have just one function to find out, $f(r)$, so we need only one condition: the Ricci scalar of the space section at constant cosmic time, ${}^3R$, is constant in space.

Remember that ${}^3R = g^{\alpha\beta}\ {}^3R_{\alpha\beta}$ and ${}^3R_{\alpha\beta} =^3 R^{\gamma}_{\alpha\gamma\beta}$ (we use the upper index ^{3} in front of R to mean that we refer to the spatial part, not to the complete space-time).

The calculation is long, boring but trivial. Here are just a few steps. It starts as usual from the affine connections; 18 of them are independent, but only 7 are different from zero:

$$\Gamma^1_{11} = \frac{1}{2}\frac{1}{f}\frac{\mathrm{d}f}{\mathrm{d}r} \qquad \Gamma^1_{22} = -\frac{r}{f}$$

$$\Gamma^1_{33} = -\frac{r\sin^2\theta}{f} \qquad \Gamma^2_{12} = \Gamma^2_{21} = \frac{1}{r}$$

$$\Gamma^3_{13} = \Gamma^3_{31} = \frac{1}{r} \qquad \Gamma^2_{33} = -\sin\theta\cos\theta$$

$$\Gamma^3_{23} = \Gamma^3_{32} = \frac{\cos\theta}{\sin\theta}$$

Since ${}^3R = g^{\beta\delta}R_{\beta\delta} = g^{11}\ {}^3R_{11} + g^{22}\ {}^3R_{22} + g^{33}\ {}^3R_{33}$ it turns out that:

$${}^3R_{11} = \frac{1}{r}\cdot\frac{1}{f}\frac{\mathrm{d}f}{\mathrm{d}r} \qquad {}^3R_{22} = 1 - \frac{1}{f} + \frac{1}{2}\frac{r}{f^2}\frac{\mathrm{d}f}{\mathrm{d}r} \qquad {}^3R_{33} = \sin^2\theta \cdot^3 R_{22}$$

and from this, by imposing that ${}^3R = constant = K$

$${}^3R = K = \frac{2}{a^2r^2}\left[1 - \frac{1}{f} + \frac{r}{f^2}\frac{\mathrm{d}f}{\mathrm{d}r}\right] = \frac{2}{a^2r^2}\left[1 - \frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r}{f}\right)\right] = \frac{2}{a^2r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left[r\left(1 - \frac{1}{f}\right)\right]$$

which gives

$$\mathrm{d}\left[r\left(1 - \frac{1}{f}\right)\right] = \frac{Ka^2r^2}{2}\mathrm{d}r$$

that is

$$r\left(1-\frac{1}{f}\right)=\frac{Ka^2r^3}{6}+A \qquad \Rightarrow \qquad f(r)=\frac{1}{1-\frac{Ka^2r^2}{6}-\frac{A}{r}}$$

But, if $r \to 0$, the metric tends to be euclidean, so that $f(r) \equiv 1$; this gives $A=0$ and

$$\mathrm{d}s^2=c^2\mathrm{d}t^2-a^2(t)\left[\frac{\mathrm{d}r^2}{1-\frac{Ka^2(t)r^2}{6}}+r^2\mathrm{d}\Omega\right] \tag{63}$$

We said before that the dependence on time of the space part of the metric is all contained in the function $a^2(t)$ in front of the square braket, and everything inside the square brakets is independent on time. This means that $Ka^2(t)$ does not depend on time, and this can be seen also from the relation defining K, which tells us that Ka^2 is a function of r. This implies that $K=K(t)$.

We define a further change of scale for r such that $\frac{Ka^2r^2}{6} \equiv k\tilde{r}^2$, where $k=0$ if $K=0$, otherwise k has the same sign of K, but magnitude 1. This gives

$$r^2=\frac{6k}{Ka^2}\tilde{r}^2 \qquad \to \qquad r=\tilde{r}\sqrt{\frac{6k}{Ka^2}} \qquad \mathrm{d}r=\sqrt{\frac{6k}{Ka^2}}\mathrm{d}\tilde{r}$$

and then

$$\mathrm{d}l^2=a^2(t)\left[\frac{6k}{K(t)a^2(t)}\cdot\frac{\mathrm{d}\tilde{r}^2}{1-k\tilde{r}^2}+\frac{6k}{K(t)a^2(t)}\tilde{r}^2\mathrm{d}\Omega^2\right]=\frac{6k}{K(t)}\left[\frac{\mathrm{d}\tilde{r}^2}{1-k\tilde{r}^2}+\tilde{r}^2\mathrm{d}\Omega^2\right]$$

If we prefer writing $\mathrm{d}l^2=\tilde{a}^2(t)[...]$, we define $\frac{6k}{K(t)} \equiv \tilde{a}^2(t)$ and we obtain:

$$\mathrm{d}l^2=\tilde{a}^2(t)\left[\frac{\mathrm{d}\tilde{r}^2}{1-k\tilde{r}^2}+\tilde{r}^2\mathrm{d}\Omega^2\right]$$

where $K(t)={}^3R(t) \equiv \frac{6k}{\tilde{a}^2(t)}$.

Finally, if we drop the (inessential) tilde we can write the metric for our universe in the following way:

$$\mathrm{d}s^2=c^2\mathrm{d}t^2-a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2}+r^2\left(\mathrm{d}\theta^2+\sin^2\theta\mathrm{d}\phi^2\right)\right] \tag{64}$$

which is the so-called ***Robertson-Walker* metric** (or line element). The Ricci scalar, giving the curvature of the space part, is given by $K=6k/a^2(t)$, where t is the cosmic time (proper time of co-moving observers), θ and ϕ are angular coordinates and r is linked to the radial distance.

2.3 Topology of the Universe

Let's see now in detail the topologiacal properties of the cosmological models corresponding to the three cases $k=0,+1,-1$.

2.3.1 The $k=0$ case

If $k=0$ the space section at constant cosmic time is an euclidean (flat) space $\mathbf{E}^3$, with $0<r<\infty$; space is infinite. Surface areas and volumes are written in the usual way.

2.3.2 The $k=+1$ case

If $\mathrm{d}\theta=\mathrm{d}\phi=0$ we have $\mathrm{d}l=a(t)\frac{\mathrm{d}r}{\sqrt{1-r^2}}$, we see that $|r|<1$ and the metric diverges if $r \to 1$. We can eliminate this divergence by choosing a new coordinate χ instead of r, such that $r=\sin\chi$ and

$$\mathrm{d}r=\cos\chi\mathrm{d}\chi=\sqrt{1-\sin^2\chi}\mathrm{d}\chi=\sqrt{1-r^2}\mathrm{d}\chi$$

$$\mathrm{d}l^2=a^2(t)\left[\frac{(1-r^2)\mathrm{d}\chi^2}{1-r^2}+\sin^2\chi\mathrm{d}\Omega^2\right]=a^2(t)\left[\mathrm{d}\chi^2+\sin^2\chi\mathrm{d}\Omega^2\right] \tag{65}$$

with $0 \le \chi \le \pi$, $0 \le \theta \le \pi$, $0 \le \phi \le 2\pi$.

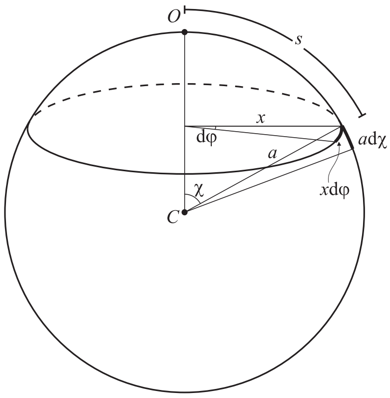
To better understand this metric let's step back to the 2-D sphere $\mathbf{E}^3$. In this case $\chi = s/a$; $s = a\chi$, $a\sin\chi = x$

$$\mathrm{d}l^2 = a^2\mathrm{d}\chi^2 + x^2\mathrm{d}\phi^2 = a^2\left(\mathrm{d}\chi^2 + \sin^2\chi\mathrm{d}\phi^2\right)$$

Moreover, we define $u \equiv x/a$, $u = \sin\chi$, $\mathrm{d}u = \cos\chi\mathrm{d}\chi = \sqrt{1-\sin^2\chi}\mathrm{d}\chi = \sqrt{1-u^2}\mathrm{d}\chi$ and the metric becomes:

$$\mathrm{d}l^2 = a^2\left[\frac{\mathrm{d}u^2}{1-u^2} + u^2\mathrm{d}\phi^2\right] \tag{66}$$

We see that the r coordinate in Robertson-Walker metric with $k = +1$ corresponds to x/a for the 2-D sphere; χ varies between 0 and π.



Let's go back to the R&W metric in the form $\mathrm{d}l^2 = a^2(t)\left[\mathrm{d}\chi^2 + \sin^2\chi\mathrm{d}\Omega^2\right]$ and, remembering how we evaluate the elements of surface and volume, we just calculate them.

For a 2-D sphere, by assuming $\mathrm{d}\chi = 0$, we have

$${}^2g_{ij} = \begin{pmatrix} a^2\sin^2\chi & 0 \\ 0 & a^2\sin^2\chi\sin^2\theta \end{pmatrix} \qquad \rightarrow \qquad \sqrt{{}^2g} = a^2\sin^2\chi\sin\theta$$

$$A(\chi) = \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} a^2\sin^2\chi\sin\theta\mathrm{d}\theta\mathrm{d}\phi = 4\pi a^2(t)\sin^2\chi$$

This has a minimum both for $\chi \to 0$ and $\chi \to \pi$, and has a maximum at the "equator" $\chi = \pi/2$. It is always, as expected for $K > 0$, $k = +1$:

$$\frac{\text{Surface Area}}{(\text{radius})^2} = \frac{4\pi a^2\sin^2\chi}{a^2\chi^2} < 4\pi$$

The proper radius is derived from the metric by imposing $\mathrm{d}\theta = \mathrm{d}\phi = 0$ and is $r_p = a\chi$.

To calculate the volume of the space we have

$${}^3g_{ij} = \begin{pmatrix} a^2 & 0 & 0 \\ 0 & a^2\sin^2\chi & 0 \\ 0 & 0 & a^2\sin^2\chi\sin^2\theta \end{pmatrix} \qquad \rightarrow \qquad \sqrt{{}^3g} = a^3\sin^2\chi\sin\theta$$

The volume within the "radial" coordinate $\chi$is then:

$$V(\chi) = \int_{\chi=0}^{\chi}\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} a^3\sin^2\chi\sin\theta\mathrm{d}\theta\mathrm{d}\phi\mathrm{d}\chi = 2\pi a^3\left(\chi - \frac{\sin2\chi}{2}\right)$$

which monotonically increases and has a maximum, finite value for $\chi = \pi$, which is $V = 2\pi^2a^3$.

The total volume of the space is proportional to a^3, so $a(t)$ is sometimes named *"radius of the universe"*. The volume is finite, even if there are no physical boundaries and the topology of this space is named $\mathbf{S}^3$.

2.3.3 The $k = -1$ case

In this case there are no discontinuities in r ($0 \leq r \leq \infty$) and

$$\mathrm{d}l^2 = a^2\left[\frac{\mathrm{d}r^2}{1+r^2} + r^2\mathrm{d}\Omega^2\right]$$

We introduce here again the variable χ and define $r \equiv \sinh\chi$ so that ($\cosh^2\chi - \sinh^2\chi = 1$) $\mathrm{d}r = \cosh\chi\mathrm{d}\chi = \sqrt{1+\sinh^2\chi}\mathrm{d}\chi = \sqrt{1+r^2}\mathrm{d}\chi$ and

$$\mathrm{d}l^2 = a^2\left[\mathrm{d}\chi^2 + \sinh^2\chi(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right] \tag{67}$$

This space is named $\mathbf{H}^3$. Similarly to what we did for $k = +1$ we can calculate the surface area of the sphere of radius $a\chi$ (proper radius):

$$A(\chi) = \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} a^2\sinh^2\chi\sin\theta\mathrm{d}\theta\mathrm{d}\phi = 4\pi a^2\sinh^2\chi$$

and, since $\sinh\chi \geq \chi$, we get

$$\frac{\text{Surface Area}}{(\text{radius})^2} = \frac{4\pi a^2\sinh^2\chi}{a^2\chi^2} > 4\pi$$

The volume, since $\chi \to \infty$, is infinite. Note the fact that the surface of the sphere increases more rapidly than in the Euclidean space $\mathbf{E}^3$, while for $\mathbf{S}^3$ the area increases less rapidly than in $\mathbf{E}^3$.

We close the paragraph noting that the metric of R&W, as well as in the way it was presented above, namely

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2\right)\right],$$

can be also written in the equivalent form

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - R^2(t)\left[d\chi^2 + S_k^2(\chi)\left(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2\right)\right] \tag{68}$$

with $R(t)$ scale factor, and the function $S_k(\chi)$ defined as:

$$S_k(\chi) = \begin{cases} \sin(\chi) & (k = +1) \\ \chi & (k = 0) \\ \sinh(\chi) & (k = -1) \end{cases} \tag{69}$$

But, **be careful**, in some textbooks and scientific papers yuo can find r in place of χ, so you need to understand from the context which of the two relations is used! We will generally use the first of the two forms.

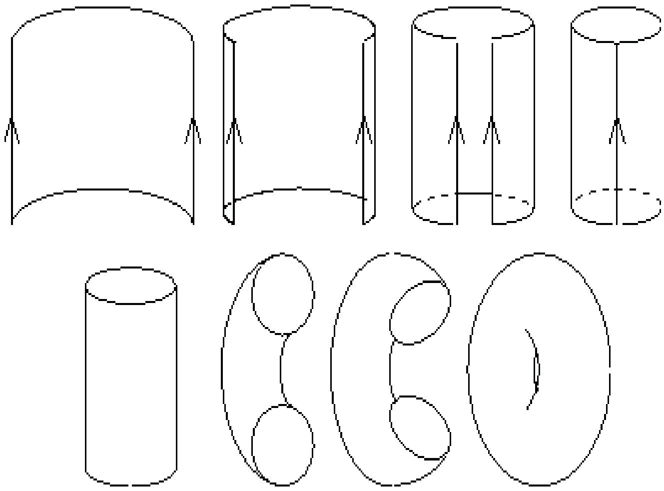
2.3.4 More complex topologies

In the three cases above we have seen the three simplest topologies: $\mathbf{E}^3$, $\mathbf{S}^3$ed $\mathbf{H}^3$. But, actually, General Relativity is a **local theory**, and our assumption of local isotropy and homogeneity implies that the space is **locally** that of $\mathbf{E}^3$, $\mathbf{S}^3$ and $\mathbf{H}^3$.

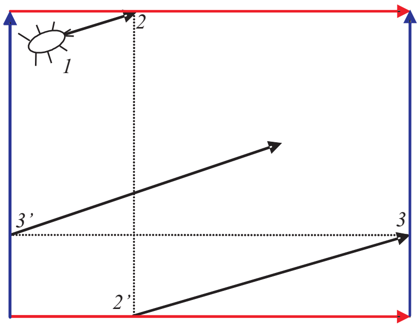
However, more complex topologies are possible[14]. Let's see some examples.

If we start initially in 2 dimensions (to help intuition) we can construct a 2-D Torus ($\mathbf{T}^2$) from a flat rectangular surface (Euclidean). Points belonging to the edges of the rectangle are suitably identified, and this can be visualized imagining to perform bending and gluing as shown below (but the curvature remains zero, while the donut shown below has not zero curvature in $\mathbf{E}^3$!):

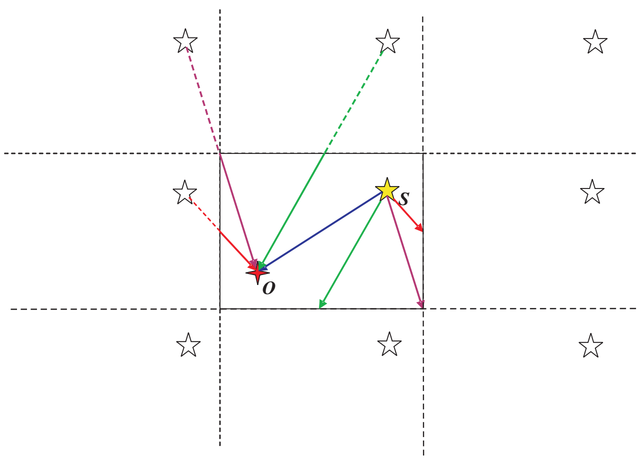
[14] See, for example, the articles *The Mathematics of Three-Dimensional Manifolds* by *W.P. Thurston* and *J.R. Weeks*, Scientific American, July 1984, p. 94, and *La forma dell'universo* by *C. Adams* and *J. Shapiro*, Le Scienze, 414, p. 72 (translation of *The Shape of the Universe: Ten Possibilities*, which appeared in American Scientist in 2001). The book *La segreta geometria del cosmo* by *J.-P. Luminet*, 2004, Raffaello Cortina Editore (see also the paper *arXiv:astro-ph/0310253*), and the site of *Jeffrey Weeks* www.geometrygames.org



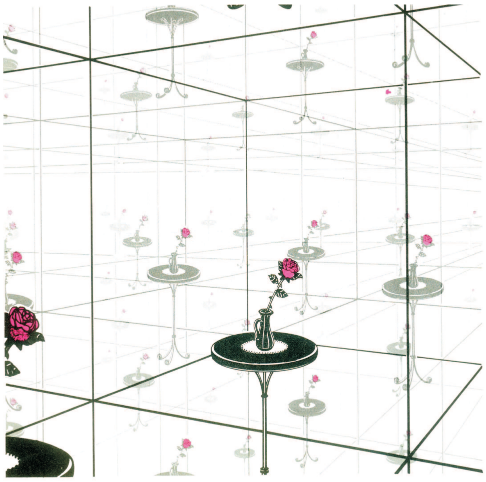
Now imagine an insect walking on the surface of the Torus. The insect crosses the upper boundary in 2 and re-enters from the bottom in 2', goes out in 3 on the right and re-enters in 3', to the left. The Torus is equivalent to a rectangle whose edges are identified two by two. Despite being finite, the surface has no boundaries.



Another typical effect of this type of compact topology is the presence of *ghosts*, that is multiple images of the same object S, arriving in O from different directions. Since the paths, and then the travel times, are different, images of the same object show it at various stages of its evolution (so it is not trivial to recognize it!).

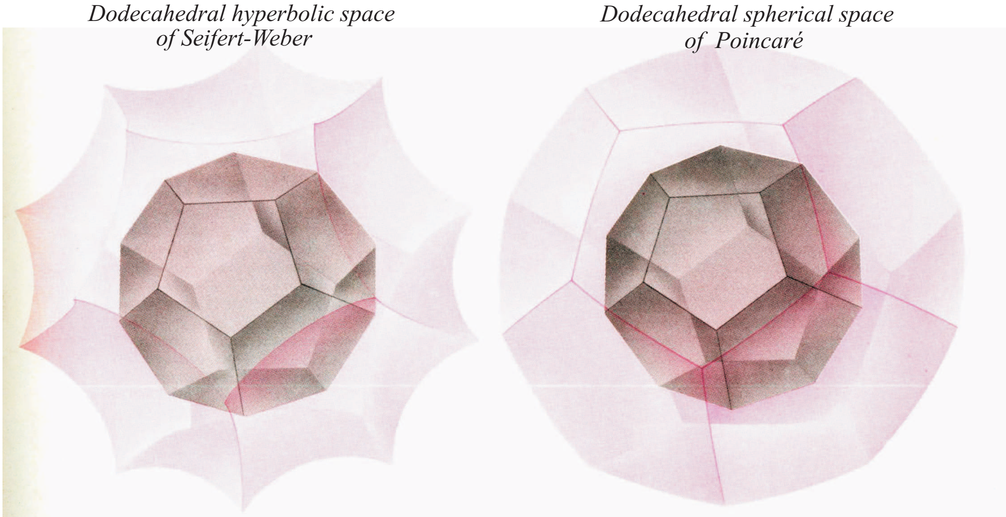


The analogue of $\mathbf{T}^2$ in 3D is the 3D Torus, $\mathbf{T}^3$. An observer placed inside has the impression of being in a room with walls, floor and ceiling covered with mirrors that do not reverse the image. Here too, for each real object, we see its ghosts in all directions.



What we said above is true for an Euclidean space (2D or 3D). This can be represented not only by a cell of parallelepiped shape, but also by a prism-shaped cell with hexagonal base. For an Euclidean 3D space there exist 10 possible compact varieties able to represent the universe, which apparently do not have borders, such as the 2D or 3D Tori.

There are also compact manifolds in non-Euclidean spaces, with positive and negative curvature. These include the dodecahedral, hyperbolic space of *Seifert-Weber*, a compact variant (i.e. with finite volume) of $\mathbf{H}^3$, obtained by pasting each face of the dodecahedron to the opposite face after a rotation of 108°(three tenths of a round angle). A compact variant of the hypersphere $\mathbf{S}^3$ is represented by the dodecahedral spherical space of *Poincaré*, obtained by pasting each face of the dodecahedron to the opposite face after a rotation of 36°(a tenth of a round angle).



Friedmann, in 1924, and Lemaître, in 1927, realized that Einstein's equations did not allow, alone, to decide whether the universe is finite or infinite. Friedmann showed how space can be made finite if points are suitably identifyied; he also realized that this allowed the existence of ghosts and observed that a positive-curvature space is always finite. Lemaître pointed out that the spaces with negative curvature admit topologies with finite volume.

J.-P. Luminet [15] interpreted the lack of fluctuations in the cosmic microwave background (CMB) on angular scales greater than 60° as due to the finite size of our universe. The cell which agrees with the experimental data (those of the *WMAP* satellite) would be the Poincaré spherical, dodecahedral space.

On the basis of what we said above, the claims that if the universe is finite its geometry must be locally spherical, and that if the geometry is locally hyperbolic or Euclidean the universe must be infinite, are wrong. But all spaces of constant curvature and locally spherical ($k = +1$) are compact.

[15] J.-P. Luminet et al., 2003, Nature **425**, 593; also *arXiv:astro-ph/0310253*

curvature	topology (space volume)
sferical $k = +1$	finite
euclidean $k = 0$	finite or infinite
hyperbolic k=-1	finite or infinite

2.4 Hubble's law

Let us consider a co-moving observer taken as the origin, a point characterized by the co-moving coordinates (r, θ, ϕ) and a beam of light (the most efficient way to exchange information) that joins them radially ($\mathrm{d}\theta = \mathrm{d}\phi = 0$). We know that for photons $\mathrm{d}s^2 = 0$.

From R&W metric:

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\frac{\mathrm{d}r^2}{1-kr^2} \equiv 0 \tag{70}$$

A light signal emitted from the point (r, θ, ϕ) at $t = 0$ (assuming that there is an initial instant of the universe, as in the Big-Bang model) arrive at $r = 0$ (the observer) at time t such that (note that if $\mathrm{d}t$ is positive, $\mathrm{d}r$ is negative, that is, as time goes by the beam passes through points gradually closer to us):

$$\int_0^t \frac{c\mathrm{d}t'}{a(t')} = \int_0^r \frac{\mathrm{d}r'}{\sqrt{1-kr'^2}} \tag{71}$$

Let us see the meaning of the term to the right. Imagine to measure with a ruler, at a fixed instant t (by freezing the expansion during the measurement), the radial distance between the origin and the point of coordinates (r, θ, ϕ); from the space part ($\mathrm{d}t \equiv 0$) of R&W metric this distance, named **proper distance** d_{pr}, will be given by:

$$d_{pr}(t) = \int_0^r \frac{a(t)\cdot \mathrm{d}r'}{\sqrt{1-kr'^2}} = a(t)\int_0^r \frac{\mathrm{d}r'}{\sqrt{1-kr'^2}} = a(t)\cdot f_k(r). \tag{72}$$

If we derive with respect to time the relation $d_{pr}(t) = a(t)\cdot f_k(r)$ we obtain the rate of variation of d_{pr} in time which, dimensionally, is a speed, and is named **recession velocity** v_r:

$$\frac{\mathrm{d}d_{pr}(t)}{\mathrm{d}t} \equiv v_r(t) = \dot{a}(t)\, f_k(r) = \frac{\dot{a}(t)}{a(t)}d_{pr}(t) \equiv H(t)\, d_{pr}(t) \tag{73}$$

$$v_r(t) = H(t)\, d_{pr}(t) \tag{74}$$

From what we saw above, when we made the changes of coordinate $r = \sin\chi$, $r = \chi$, $r = \sinh\chi$:

$$f_k(r) = \left\{ \begin{array}{ll} \arcsin r \simeq r + r^3/6 + \cdots & (k=+1) \\ r & (k=0) \\ \mathrm{arcsinh} r \simeq r - r^3/6 + \cdots & (k=-1) \end{array} \right\} \simeq r + k\, r^3/6 + \cdots \tag{75}$$

Actually, $d_{pr}(t)$ cannot be directly measured. Its relation with $d_{pr}(t_0)$ ($t = t_0$ corrisponds to today) comes from the fact that

$$\frac{d_{pr}(t)}{a(t)} = \frac{d_{pr}(t_0)}{a(t_0)} = f_k(r) \qquad \text{since } r = const \text{ in time}$$

$$d_{pr}(t) = \frac{a(t)}{a_0}d_{pr}(t_0) \qquad (a_0 = a(t_0)\ )$$

So $d_{pr}(t)$ depends on time through $a(t)$. The quantity $f_k(r)$, or also $a_0 f_k(r)$, time-invariant, is named **co-moving distance** (it corresponds to the proper distance today).

This is **Hubble's law,** and the quantity $H(t) = \dot{a}(t)/a(t)$ is named **Hubble parameter**. If we write this relation for the present time t_0 we get $v_r(t_0) = H_0\, d_{pr}(t_0)$, where $H_0 \equiv H(t_0)$ is the **Hubble constant**. The uncertainty on its value is parametrized by writing $H_0 = 100\, h\, km\, s^{-1}\, Mpc^{-1}$, with $0.5 \leq h \leq 1.0$. H_0 has dimension $time^{-1}$, and approximately $1/H_0 \simeq 3\cdot 10^{17} h^{-1}\, s$. After decades of disputes, the value of H_0 seems today quite well defined; the value given by the devoted *Key Program* of the *Hubble Space Telescope* is $H_0 = 72 \pm 8\, km\, s^{-1}\, Mpc^{-1}$, while the analysis of the cosmic microwave background (*CMB*) gives $H_0 = 70 \pm 2\, km\, s^{-1}\, Mpc^{-1}$ (WMAP9) and $H_0 = 67 \pm 1.5\, km\, s^{-1}\, Mpc^{-1}$ (Planck).

An important point has to be made on Hubble's law. Keeping H_0 fixed, if d_{pr} increases, v_r may become greater than the speed of light. The proper distance correspondig to $v_r = c$ is named **Hubble radius**, R_H, which is then defined by the relation

$$R_H(t) \equiv \frac{c}{H(t)} \tag{76}$$

and, like H, depends on time.

The fact that $v_r > c$ for $d_{pr} > R_H$, can create some confusion, but this is not in contrast with Special Relativity because, when referred to co-moving observers, the velocity of any object is, *locally*, always less than c. No information travels with $v > c$. The distance between observers, the space interposed between them, grows more rapidly than c, but this does not correspond to a transmission of information. In addition, to evaluate the velocity of an object relatively to an observer, we must move the two velocity vectors (of the object and of the observer) to the same locationt and make a difference; in Euclidean space this implies a parallel transport, but in a curved space the result depends on the path followed. Therefore, in a curved space (or space-time) the relative velocity of two objects, not located in the same position, is ambiguous and meaningless.

In addition to Hubble's parameter, which depends on the time derivative of the scale factor, we define also the so-called **deceleration parameter** q, related to $\ddot{a}$, always positive in a universe without cosmological constant, but begative when expansion accelerates. By definition

$$q(t) \equiv -\frac{\ddot{a}(t)a(t)}{\dot{a}(t)^2}, \tag{77}$$

with value q_0 for $t = t_0$ (at the present time). The two parameters H_0 and q_0 are useful for a **series expansion** of $a(t)$ around $t = t_0$:

$$a(t) = a(t_0) + \dot{a}(t_0)(t - t_0) + \frac{1}{2}\ddot{a}(t_0)(t - t_0)^2 + \ldots = a(t_0)\left[1 + H_0(t - t_0) - \frac{1}{2}q_0 H_0^2 (t - t_0)^2 + \ldots\right] \tag{78}$$

2.5 Conformal time - Redshift

We wrote R&W metric as

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)\left[\mathrm{d}\chi^2 + S_k^2(\chi)\mathrm{d}\Omega^2\right] \qquad \text{with} \quad S_k(\chi) = \begin{cases} \sin\chi & k = +1 \\ \chi & k = 0 \\ \sinh\chi & k = -1 \end{cases}$$

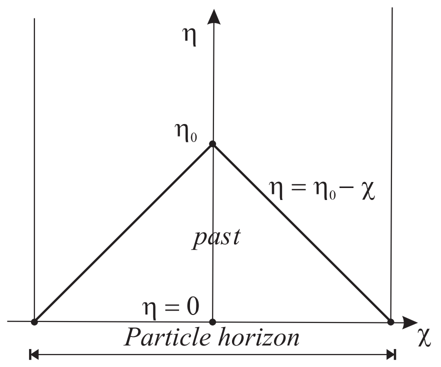
In this case we use the so-called *syncronous gauge*. But sometimes it is useful to factorize completely the scale factor. To do that we define the *conformal time* η in such a way that $\mathrm{d}\eta = c\mathrm{d}t/a(t)$, and the metric can be written as (*conformal gauge*):

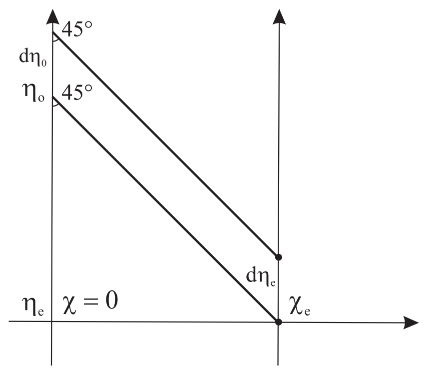
$$\mathrm{d}s^2 = a^2(\eta)\left[\mathrm{d}\eta^2 - \left(\mathrm{d}\chi^2 + S_k^2(\chi)\mathrm{d}\Omega^2\right)\right] \tag{79}$$

If a photon travels toward us radially, with $\theta = \phi = const$, since $\mathrm{d}s^2 = 0$, we have

$$a^2(\eta)[\mathrm{d}\eta^2 - \mathrm{d}\chi^2] = 0 \tag{80}$$

that is $\mathrm{d}\eta = \pm\mathrm{d}\chi$. This represents the light cone, with the rays inclined at 45°. If $\eta = 0$ represents the beginning of the universe, We see that there are values of χ such that no information has yet arrived from those points: we say that there is a *particle horizon* (we will discuss better this point in next section). The equation of motion of a photon traveling toward the observer ($\chi = 0$) is $\eta = \eta_0 - \chi$ with $\eta_0 =$ "η today".





Two signals, emitted at times η_e and $\eta_e + \mathrm{d}\eta_e$ from a co-moving source at $\chi = \chi_e$, will be received in $\chi = 0$ at times η_0 and $\eta_0 + \mathrm{d}\eta_0$.

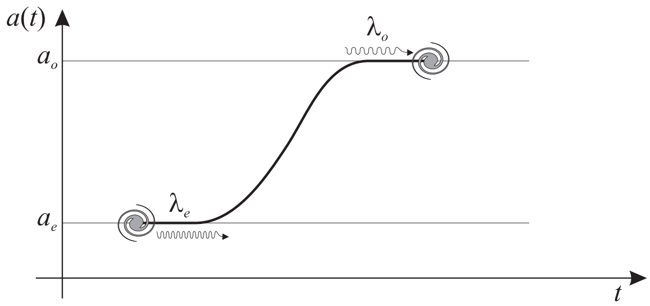
In conformal gauge the motion of photons is always inclined at 45°, so $\mathrm{d}\eta_e \equiv \mathrm{d}\eta_0$ and

$$\frac{c\mathrm{d}t_e}{a(t_e)} = \frac{c\mathrm{d}t_0}{a(t_0)}$$

If $\mathrm{d}t_e$ is the period of an electromagnetic wave with frequency $\nu_e = 1/\mathrm{d}t_e$, the observed frequency $\nu_0 = 1/\mathrm{d}t_0$ is related to ν_e by $\nu_e a(t_e) = \nu_0 a(t_0) = \nu_0 a_0$, that is:

$$\frac{\nu_e}{\nu_0} = \frac{a_0}{a(t_e)} \qquad \rightarrow \qquad \frac{\lambda_e}{\lambda_0} = \frac{a(t_e)}{a_0} \tag{81}$$

The wavelength undergoes a "dilation" equal to that of the scale factor. This effect is produced by the variation of the scale factor due to the expansion, and not by the relative velocity between source and receiver; thus it is improper to call it a Doppler effect. To illustrate this fact let us suppose that, in an ideal cosmological model, a photon is emitted when $a(t)$ is constant, then a phase of expansion from a_e to a_0 follows, and finally there is a new phase of $a = cost = a_0$, during which the photon is received by an observer. Source and observer are at rest with rescpect to the universe when the photon is emitted and received, so there is no Doppler effect, but the cosmological redshift is present and $\lambda_0/\lambda_e = a_0/a_e$!



The **redshift** z is defined as

$$z \equiv \frac{\lambda_0 - \lambda_e}{\lambda_e} = \frac{\lambda_0}{\lambda_e} - 1 = \frac{a_0}{a(t_e)} - 1 \tag{82}$$

and

$$\frac{\lambda_0}{\lambda_e} = 1 + z = \frac{a_0}{a(t_e)} \qquad \Rightarrow \qquad a(t) = \frac{a_0}{1-z} \tag{83}$$

This relation is also useful to link the redshift (*observable*) to the scale factor.

The cosmological redshift is not due to a simple Doppler effect, but *locally*, to the first order in v/c, it can be understood in this way. In fact, from the formula of the Doppler effect, we have:

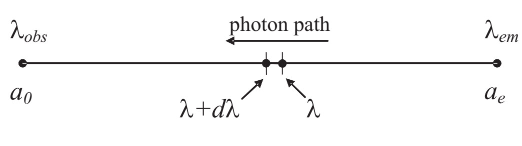
$$\frac{v}{c} = \frac{\Delta\lambda}{\lambda} = \frac{\lambda_0 - \lambda_e}{\lambda_e} = \frac{\lambda_0}{\lambda_e} - 1 = \frac{a_0 - a(t_e)}{a(t_e)}$$

but $a(t_e) \simeq a_0 + H_0 a_0 (t_e - t_0) + \mathcal{O}(\Delta t^2)$ and so

$$v \simeq c\frac{-H_0 a_0 (t_e - t_0)}{a_0 + H_0 a_0 (t_e - t_0)} \simeq cH_0(t_0 - t_e)\left[1 + H_0(t_0 - t_e)\right] \simeq cH_0(t_0 - t_e) \simeq cH_0\Delta t$$

But, if $\mathrm{d}s^2 = 0$ and $\mathrm{d}\theta = \mathrm{d}\phi = 0$, $c^2\mathrm{d}t^2 = a^2\mathrm{d}\chi^2$ and $c\Delta t = a_0\Delta\chi = d_{pr}$ and so $v = H_0 d_{pr}$ which is just Hubble's law. Then, working backward from this, we find the formula of the Doppler effect. *Locally*, therefore, the cosmological redshift can be seen as a Doppler effect due to the differential motion between two nearby, co-moving observers.

But we can also interpret it as *a large-scale integrated Doppler effect*, the sum of many different effects on the path of the photon from the source to us.



Consider the co-moving observers placed along the path the photon takes from its emission ($a = a_e$) to its arrival to us ($a = a_0$). Consider two of these observers, separated by an infinitesimal proper distance dl that the photon travels in a time dt (d$l = c$dt), which move with relative velocity dv due to the expansion (d$v = H$dl), and who observe the photon, emitted with $\lambda = \lambda_{em}$, with wavelengths $\lambda$and $\lambda+$dλ, respectively.If we remember that $H(t) = \dot{a}(t)/a(t)$, the *local* Doppler effect implies:

$$\frac{\mathrm{d}\lambda}{\lambda} = \mathrm{d}z = \frac{\mathrm{d}v}{c} = \frac{H\mathrm{d}l}{c} = \frac{H c\mathrm{d}t}{c} = \frac{\mathrm{d}a}{\mathrm{d}t}\frac{1}{a}\mathrm{d}t = \frac{\mathrm{d}a}{a}$$

By integrating this relation we get

$$\int_{\lambda_e}^{\lambda_0} \mathrm{d}ln\lambda = \int_{a_e}^{a_0} \mathrm{d}ln\ a \qquad \rightarrow \qquad ln\frac{\lambda_0}{\lambda_e} = ln\frac{a_0}{a_e} \qquad \rightarrow \qquad \lambda_0 = \lambda_e\frac{a_0}{a_e} = \lambda_e(1+z)$$

We can understand the cosmological redshift simply as a global effect due to the different relative speeds of the co-moving observers placed along the path of the photon coming to us. There is no need to think of a "stretching effect" of the wavelength of the photon due to the expansion of the space!

2.6 Horizons

The difficulty in defining the global topology of the universe resides also in a limit to observations, which is not merely instrumental, but physical. One question we can ask is: *Is there a maximum distance from which I got so far information?* As we have repeatedly emphasized, the observational evidence suggests that the universe had an origin in time (and most of the theoretical models support this evidence).

We have seen that a luminous signal, emitted at the co-moving coordinate $r = r_H$ at $t = 0$ arrives at the observer ($r = 0$) at time t according to the relation

$$\int_0^t \frac{c\mathrm{d}t'}{a(t')} = \int_0^{r_H} \frac{\mathrm{d}r}{\sqrt{1-kr^2}} \tag{84}$$

According to the definition of proper distance we have

$$\int_0^t \frac{c\mathrm{d}t'}{a(t')} = \int_0^{r_H} \frac{\mathrm{d}r}{\sqrt{1-kr^2}} = f_k(r_H) \equiv \frac{d_{pr}(t, r_H)}{a(t)} \tag{85}$$

The quantity

$$d_H(t) \equiv d_{pr}(t, r_H) = a(t)\int_0^t \frac{c\mathrm{d}t'}{a(t')} \tag{86}$$

represents the *maximum proper distance* from which, at time t, we received light signals. If $d_H(t)$ is finite, there is a part of the universe from which we have not yet received a light signals and there is a so-called **particle horizon** (*PH*). The fact that $d_H(t)$ is finite depends on the behavior of $a(t)$. We will see that, for reasonable cosmological models, $d_H \propto t$ and is therefore finite. For cosmological models without singularity (such as the *Steady State*), the lower limit of integration should be placed not to 0 but to $-\infty$.

If instead we look forward, we can ask: *From what distance we may in the future receive signals that start today?* The answer is obtained by integrating between t and ∞(or $t = t_{max}$ if the universe recollapses) instead of between 0 and t.

$$d_E(t) \equiv a(t)\int_t^{\infty} \frac{c\mathrm{d}t^{'}}{a(t')} \tag{87}$$

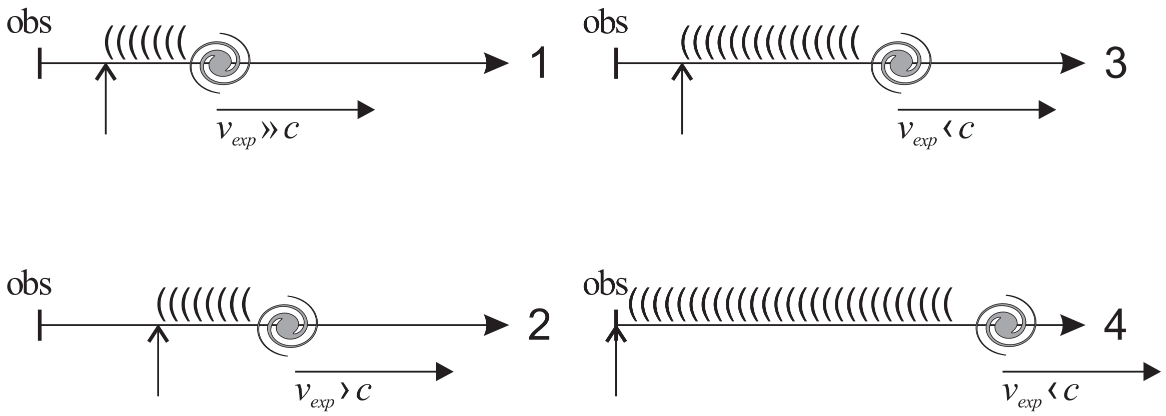
If the integral diverges we have just to be patient enough to see any event, otherwise there are distances from which we will never receive information. In this case, there is an **event horizon** (*EH*). For this to happen it is necessary that $a(t)$ grows faster than t. For example, if $a(t) \propto e^{Ht}$, with $H = const$, we have:

$$d_E(t) = e^{Ht}\cdot\int_t^{\infty} \frac{c\mathrm{d}t'}{e^{Ht'}} = c\, e^{Ht}\left[-\frac{1}{H}e^{-Ht}\right]_t^{\infty} = c\, e^{Ht}\cdot\frac{e^{-Ht}}{H} = \frac{c}{H} = const \tag{88}$$

as in models dominated by a cosmological constant and in the *Steady State* model. But while $d_E = const$, the distances of galaxies grow on their own as $a(t) \propto e^{Ht}$ and then, as time goes on, they "go out" from d_E: do we expect a lonely destiny in this cosmological model? Not exactly, since galaxies going beyond *EH* are always visible, but with a story, as seen by us, more and more slowed, and with photons increasingly reddened. In fact, from galaxies at the edge we will receive photons at $t \to \infty$,when $a(t) \to \infty$ and then with a redshift $z \to \infty$ ($\lambda_{oss} \to \infty$).

At first glance it may seem strange that in an expanding universe, initially very "small" compared to today, there is a particle horizon. The cause of this resides in the fact that the rate of expansion defined by the recession velocity $v_r(t)$ can be, and in some cases is, much larger than c (see the above). This does not violate Special Relativity because the recession velocity does not correspond to a transmission of information, but measures only how the space between co-moving observers increases.

Because of this expansion, which in the early stages of the *Big Bang* occurs with $v_r \gg c$, a photon emitted towards us in the early stages of the universe initially moves away from us because the space to cover increases, per unit time, faster than the speed of the photon; this may finally come closer to us only when the velocity of the expansion becomes smaller than c.



Let us return briefly to the growth of the mutual distances between co-moving observers. It is usual to express this fact by saying that the space "expands", but one should not think that expanding space carries with it the galaxies, such as a cake that, brewing, remove the candies from each other. Immediately after the Big Bang, the elementary particles present in the early universe (and which, later, formed galaxies) found themselves in a state of mutual separation (for the moment we must take this as an initial condition of the motion, but the mechanism may be similar to that producing inflation). The evolution of this mutual distancing, described by the scale factor $a(t)$, results from the application of Einstein's equations. This mutual separation of objects can also be given a description in terms of expansion of the *cosmic substrate*, but it is more correct and safe to think that the objects, given the initial conditions, move under the effect of the gravitational interaction (in this we include also the cosmological constant) with all sources of mass-energy of the universe, and not because they are dragged or "elongated" and "stretched" by a mysterious expanding space (which, among the other, would require to introduce a new type of interaction producing this "stretching").

At the conclusion of these observations on the geometry of the universe, we come to the conclusion that physical limitations prevent us from accessing to the entire universe: we know only a part, and extrapolate our "local" knowledge to the whole. Recall also that the universe, by definition, is the set of all existing objects and physical concepts (such as stars, galaxies, atoms, space and time). There is not a physical space "outside" of a finite universe. In this sense, certain "naive" representations of the universe like a balloon (2D) that swells in space (3D) can be confusing. In addition, if the universe is spatially infinite, it was that even in the past, when $a(t) \to 0$. Finally, there is not a point from which we can say that everything expanded. Each point is equivalent, and the Big Bang happened at any point: if, at any point, we come back in time, the density and the local temperature increase and tend to infinity when $t \to 0$.

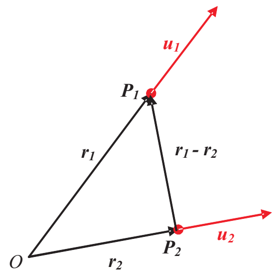
2.7 Milne's model

We defined "cosmic time" in a way that seems obvious but, as we shall see, some consequencies of this assumption (together with the finite value of the speed of light) are far less obvious. We will show this by means of the cosmological model of E. A. Milne (1932), a most ingenious und simple model universe, which nicely illustrates many of the features shared by the more complicated models. And, though it wiIl not be immediately apparent, Milne's model satisfies the Cosmological Principle. It does not use General Relativity, but only Special Relativity.

Furthermore, while this model is not satisfactory and in some way unphysical, it helps us to understand some basic elements of the cosmic expansion.

Consider an empty Minkowski space-time. Totally neglecting gravity, Milne considered an infinite number of test particles (no mass, no volume) shot out (for reasons unknown), in all directions and with all possibie speeds, at a unique creation event $\mathcal{C}$. Let us look at this situation in some particular inertial frame S, and suppose $\mathcal{C}$ occurred al its origin O at $t = 0$. All the particles, being free, will move uniformly and radially away from O, with all possible speeds $u < c$. The observer S, at rest with respect to O, sees a ball of dust particles whose unattained boundary expands at the speed of light.

At first glance it seems that this model does not satisfy the Cosmological Principle, because O is a privileged point. But the "boundary" of the universe behaves, kinematically, as a spherical front of light emitted in O at time$t = 0$ (creation event, $\mathcal{C}$). And each particle, having been present in $\mathcal{C}$, consider itself at the center of the wavefront.



The motion is radial from O, with $r = u \cdot t$ for each particle. Then $u = r/t = H \cdot r$ with $H = 1/t$ and Hubble's law holds. But Hubble's law holds *for every single* moving "grain": in fact, for particle 2, as for any other, $(\overline{r}_1 = \overline{u}_1 t,\, \overline{r}_2 = \overline{u}_2 t)\; \overline{r}_1 - \overline{r}_2 = (\overline{u}_1 - \overline{u}_2)t \quad \Rightarrow \quad d = v \cdot t$. Each particie will consider the whole motion pattern to be radially away from itself, and of course uniform.

There remains the question whether we can have an isotropic density distribution around each particle. To study this, let τ denote the proper time elapsed at each particle since creation. Then n_0, the proper particle density measured at any given particle P, is of the form

$$n_0 = \frac{N}{\tau^3} \qquad (N = const) \tag{89}$$

because a small sphere around P, containing a fixed number of particles, expands with the constant, relative velocity δu of the farthest particles, and thus has radius $\delta u \cdot \tau$ and volume $\frac{4}{3}\pi \delta u^3 \tau^3$.

For the particle P at a distance r from the origin of S in O it is:

$$\tau = \frac{t}{\gamma(u)} \qquad u = \frac{r}{t} \qquad \gamma(u) = \frac{1}{\sqrt{1 - u^2/c^2}} \tag{90}$$

and the number density of particles in P, *relative to* O, is (recall that volumes are contracted by a factor$1/\gamma$ in the direction of motion)

$$n = \frac{\gamma(u) \cdot N}{\tau^3} = \frac{\gamma^4 N}{t^3} = \frac{Nt}{(t^2 - r^2/c^2)^2} \tag{91}$$

Notice that (91) near the origin O (for $r \to 0$) gives $n \simeq N/t^3$. It is clear that, conversely, a density defined by (91) relative to the origin particle in O reduces to (89) at each particle, and thus to (91) relative to any *other* particle taken as origin. This is therefore the density distribution we must require to hold around *any* particle.

Observe how this density approaches infinity at the "edge" $r = ct$; this is due to the fact that when we look at points near the edge, we look back to times near the creation, when the density tended tominfinity. Note also that τ is the cosmic time in Milne's model, linked to the local number density via Eq.(89).

Although Milne's model satisfies the cosmological principle, is not satisfactory since it accepts that there is an "outside" of the universe of galaxies (over the edge $r = ct$) which can, however, interact with these (for example by sending light signals that can be seen). The expansion of dust grains occurs *in space*, while in a satisfactory cosmological model, without an "outside", the space between galaxies expands.

Consider now Milne's model from another point of view. We take now cosmic time as time coordinate and consider co-moving coordinates (e.g., u, θ, ϕ) with respect to the inertial frame S. The speed u is a co-moving coordinate since it doesn't change in time and is linked to the proper distance r. The metric in spherical polar coordinates will be, relative to S,

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - \left[\mathrm{d}r^2 + r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right] \tag{92}$$

and suppose the origin $r = t = 0$ coincides with Milne's creation-event $\mathcal{C}$. We move now to cosmic time τ, (see Eq. 90): $\tau = t \cdot (1 - u^2/c^2)^{\frac{1}{2}}$ and define a new co-moving coordinate $\tilde{r}$, more suitable than u, as $c\tilde{r} = u/(1 - u^2/c^2)^{\frac{1}{2}}$. With the help of the relation $\tilde{r} = \mathrm{sinh}\psi$ (remember that $\cosh^2\psi - \sinh^2\psi = 1$) we obtain:

$$r = ut = \frac{u\tau}{\sqrt{1 - u^2/c^2}} = c\tau\tilde{r} = c\tau\mathrm{sinh}\psi$$

$$\mathrm{sinh}\psi = \frac{u/c}{\sqrt{1 - u^2/c^2}} \qquad \to \qquad \mathrm{sinh}^2\psi = \frac{u^2/c^2}{1 - u^2/c^2} = \mathrm{cosh}^2\psi - 1$$

$$\mathrm{cosh}^2\psi = \frac{1 - u^2/c^2 + u^2/c^2}{1 - u^2/c^2} = \frac{1}{1 - u^2/c^2} \qquad \to \qquad \mathrm{cosh}\psi = \frac{1}{\sqrt{1 - u^2/c^2}}$$

This allows us to write $t = \tau\mathrm{cosh}\psi$. Note also that $\mathrm{cosh}^2\psi = 1 + \mathrm{sinh}^2\psi = 1 + \tilde{r}^2$, and, from $\tilde{r} = \mathrm{sinh}\psi$, by differentiation, we have $d\tilde{r} = \mathrm{cosh}\psi\, d\psi$, that is $d\psi = d\tilde{r}/\sqrt{1 + \tilde{r}^2}$.

If now, in (92), we go from r, t to $\tilde{r}, \tau$:

$$\begin{aligned}
\mathrm{d}s^2 &= c^2\left[\tau\mathrm{sinh}\psi\mathrm{d}\psi + \mathrm{cosh}\psi\mathrm{d}\tau\right]^2 - \left[c\tau\mathrm{cosh}\psi\mathrm{d}\psi + c\mathrm{sinh}\psi\mathrm{d}\tau\right]^2 - c^2\tau^2\mathrm{sinh}^2\psi\mathrm{d}\Omega^2 \\
&= c^2\mathrm{d}\tau^2 - c^2\tau^2\left[\mathrm{d}\psi^2 + \tilde{r}^2\mathrm{d}\Omega^2\right] \\
&= c^2\mathrm{d}\tau^2 + c^2\tau^2\left[\frac{\mathrm{d}\tilde{r}^2}{1 + \tilde{r}^2} + \tilde{r}^2\mathrm{d}\Omega^2\right]
\end{aligned}$$

This corresponds to a *R&W* metric with $k = -1$ and with $a(\tau) = c\tau$ (remember that time t in *R&W* metric is the proper time here named τ)!

The definition of cosmic time as the proper time of co-moving observers, the finite value of the speed of light, and a negligible (actually vanishing) mass density produce a negatively curved space section ($k = -1$), but the space-time curvature is zero, since we just changed reference frame from the flat Minkowski metric of Eq. 92. If we gradually increase mass density, k becomes first equal to zero and then equal to $+1$.

Moreover, in this model (and *only* in this model!) based solely on Special Relativity, it is correct to apply the the relativistic Doppler effect formula linking redshifts and recession velocity:

$$1 + z = \frac{\sqrt{1 + (v/c)}}{\sqrt{1 - (v/c)}} \tag{93}$$

3 Cosmological Models

3.1 Friedmann equations

We are now ready to derive the equations ruling the behaviour of the scale factor $a(t)$ in a universe described by a R&W metric, and by the energy-momentum tensor of a perfect fluid.

We start with the metric tensor; coordinates are $(x^0, x^1, x^2, x^3) = (ct, r, \theta, \phi)$:

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{a^2}{1-kr^2} & 0 & 0 \\ 0 & 0 & -a^2r^2 & 0 \\ 0 & 0 & 0 & -a^2r^2\sin^2\theta \end{pmatrix} \qquad {}^4g = -\frac{a^6r^4\sin^2\theta}{1-kr^2} \qquad g^{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1-kr^2}{a^2} & 0 & 0 \\ 0 & 0 & -\frac{1}{a^2r^2} & 0 \\ 0 & 0 & 0 & -\frac{1}{a^2r^2\sin^2\theta} \end{pmatrix}$$

We than evaluate the *affine connections.* Most of them vanish; those different from zero are

$$\Gamma^0_{ij} = -\frac{\dot{a}}{a}\frac{g_{ij}}{c}$$

$$\Gamma^i_{0j} = \frac{\dot{a}}{a}\frac{\delta^i_j}{c}$$

where $i, j = 1, 2, 3$. Connections like Γ^i_{jk} are the same already estimated when we derived R&W metric (the -1 factors present in g_{ij} and g^{ij} are simplified): just substitute, in place of the unknown function $f(r)$, the expression $\frac{1}{1-kr^2}$; for instance: $\Gamma^1_{11} = \frac{kr}{1-kr^2}$.

Then we move from $\Gamma^\alpha_{\beta\gamma}$ to Ricci tensor $R_{\alpha\beta} = R^\gamma_{\ \alpha\gamma\beta}$; with a bit of patience, the components different from zero are:

$$R_{00} = -\frac{3}{c^2}\frac{\ddot{a}}{a}$$

$$R_{ij} = -\frac{g_{ij}}{c^2}\left[\frac{\ddot{a}}{a} + \frac{2\dot{a}^2}{a^2} + \frac{2kc^2}{a^2}\right] \qquad (3 \text{ components } \neq 0: \ R_{11}, \ R_{22}, \ R_{33})$$

and the Ricci scalar is:

$$R = g^{\alpha\beta}R_{\alpha\beta} = -\frac{6}{c^2}\left[\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{kc^2}{a^2}\right]$$

We calculate the components of the stress-energy tensor $T_{\alpha\beta} = (p + \rho c^2)u_\alpha u_\beta - pg_{\alpha\beta}$

In the co-moving frame $u^\alpha = (1, 0, 0, 0)$; you can easily check that $u_\alpha = g_{\alpha\beta}u^\beta = g_{\alpha 0}u^0 = g_{\alpha 0} = (1, 0, 0, 0)$. So

$$T_{00} = (p + \rho c^2) - p = \rho c^2 \qquad \text{e} \qquad T_{ij} = -pg_{ij}$$

So we can finally write

$$R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} = \frac{8\pi G}{c^4}T_{\alpha\beta} + \Lambda g_{\alpha\beta}$$

The 00 component gives

$$\dot{a}^2 + kc^2 = \frac{8\pi G}{3}\rho a^2 + \frac{1}{3}a^2c^2\Lambda \qquad (F1) \qquad (94)$$

From any of the three components (11, 22, 33) we obtain:

$$\ddot{a} + \frac{1}{2a}\left(\dot{a}^2 + kc^2\right) = -\frac{4\pi G}{c^2}p\,a + \frac{1}{2}\Lambda c^2 a$$

and, by using $(F1)$:

$$\ddot{a} = -\frac{4\pi G}{3}(\rho + \frac{3p}{c^2})a + \frac{1}{3}\Lambda c^2 a \qquad (F2) \qquad (95)$$

Equations $(F1)$ and $(F2)$ aren't really independent: if we make explicit ρ from $(F1)$ and derive with respect to time we get:

$$\dot{\rho} = \frac{3\dot{a}}{4\pi G a^2}\left[\ddot{a} - \frac{1}{a}\left(\dot{a}^2 + kc^2\right)\right]$$

By using $(F2)$ for $\ddot{a}$ and reusing $(F1)$ for the term in parentheses, we get:

$$\dot{\rho} + 3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) = 0 \qquad (F3) \tag{96}$$

We have already seen a similar expression (Eq. 40) when we were speaking about the energy-momentum tensor (we used the edge L of a cubic cell instead of the scale factor a). It is just a way to express energy and momentum conservation through the First Principle of Thermodynamics, and is linked to the four-divergence of $T^{\alpha\beta}$: $T^{\alpha\beta}{}_{;\beta} = 0$. The above relation can be again derived from $\mathrm{d}Q = \mathrm{d}U + \mathrm{d}L = 0$, which represent an adiabatic expansion:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\rho c^2 a^3\right) = -p\frac{\mathrm{d}}{\mathrm{d}t}\left(a^3\right) \qquad \rightarrow \qquad \mathrm{d}\left(\rho c^2 a^3\right) + p\,\mathrm{d}\left(a^3\right) = 0 \tag{97}$$

3.2 The density of the Universe

One of the key parameters in Friedmann's equations is the density of the Universe; we will try to estimate its value. First of all we define the so-called **critical density** ρ_{cr}:

$$\rho_{cr} = \frac{3H^2}{8\pi G} \tag{98}$$

which, like the Hubble parameter H, is a function of time. Its present value, from $H_0 = h \cdot 100\, km\, s^{-1} Mpc^{-1} = h \cdot 3.241 \cdot 10^{-18} s^{-1} = (h/3.086 \cdot 10^{17})\, s^{-1}$ is $\rho_{cr} \simeq 1.879 \cdot 10^{-29} h^2 g\, cm^{-3}$.

Usually the density ρ is referred to ρ_{cr} by means of the **density parameter** Ω :

$$\Omega \equiv \frac{\rho}{\rho_{cr}}$$

Since there are many contributions to the overall density of the Universe, there is a particular value of Ω for each of them. Let's see now these different contributions.

3.2.1 Luminous Matter

The density ρ_{lum} of luminous matter, basically *stars*, can be derived from the luminosity density ρ_L of the Universe ($\rho_L \sim 2 \cdot 10^8 h\ L_\odot Mpc^{-3}$), by assuming a *mass-to-light ratio* $\langle M/L\rangle \sim 1\, M_\odot/L_\odot$. We obtain $\Omega_{lum} \equiv \rho_{lum}/\rho_{cr}$:

$$\Omega_{lum} h \simeq 0.002\ -\ 0.006$$

3.2.2 Galaxies

The presence of massive, dark halos around the luminous part of galaxies, revealed by flat rotation curves, increases the mass-to-light ratio to a value $\langle M/L\rangle \sim 30\, h\, M_\odot/L_\odot$[16]. This gives an estimate of Ω_{gal} which is an order of magnitude larger than Ω_{lum}:

$$\Omega_{gal} \geq 0.03\ -\ 0.05$$

3.2.3 Galaxy clusters

From the virial theorem applied to groups and clusters of galaxies (Zwicky, 1933), as well as from mass estimates by means of gravitational lensing or from the X-ray emission of the *intra-cluster medium* (*ICM*), we obtain $\langle M/L\rangle \sim 100 - 400\, h\ M_\odot/L_\odot$, i. e. a value about 10 times larger than that for single galaxies. So we get the estimate

$$\Omega_{cl} \sim 0.1 - 0.3$$

[16]Recall that, for a galaxy or a galaxy cluster with mass M, $M \sim V^2 R \sim V^2 \theta D$, where V is the velocity dispersion , D is the distance, θ the subtended angle, and R is the size of the system. The mean recession velocity of the system $\langle v\rangle$ is, according to Hubble's law, $\langle v\rangle = H_0 D \propto h\, D$, that is $D \propto h^{-1}$. So mass scales as $M \propto h^{-1}$. The measured flux is $F \simeq L/D^2$ and so $L \propto D^2 \propto h^{-2}$. It turns out that $M/L \propto h^{-1}/h^{-2} \propto h$. The luminosity density will then scale as $L/\mathrm{Volume} \propto L/\mathrm{length}^3 \propto h^{-2}/h^{-3} \propto h$.

3.2.4 Primordial (Big Bang) Nucleosynthesis

The *Hot Big Bang* involves the synthesis (called primordial, or *Big Bang*, nucleosynthesis, *BBN*) of ^3He, ^4He, D, ^7Li when the universe was about three minutes old. As we shall see, from theoretical calculations, compared with the observations, we have that the density of *baryons* (in astrophysics this term is referred to protons, neutrons, nuclei of helium, etc., i.e. the ordinary matter we and thing around us are made of) gives a contribution

$$\Omega_b h^2 \simeq 0.005 - 0.024$$

with a preference for the higher values of this range, as suggested by observations of the *Cosmic Microwave Background* (*CMB*).

3.2.5 The Baryon Catastrophe

The hot, diffused plasma (*ICM*) present in the space among galaxies of galaxy clusters, detected by means of its X-ray emission, contributes by about $6\,h^{-3/2}\%$ to the total mass of the cluster. Stars, in galaxies, contribute with a further 2%.

From numerical simulations, the ratio between the mass M_b in baryons and total mass M_{tot} of a cluster is representative of the relation, on a cosmic scale, between the density of baryons and the total matter density, or between the corresponding density parameters Ω_b and Ω_M. By taking into account that some of the baryons may be dark, we get (by assuming $\Omega_b h^2 \simeq 0.02$)

$$\Omega_b/\Omega_M \geq 0.06\,h^{-3/2} + 0.02 \;\rightarrow\; \Omega_M \leq \frac{0.02\,h^{-2}}{0.06\,h^{-3/2} + 0.02} \leq 0.33$$

if we use $h \simeq 0.7$. The term *baryon catastrophe* dates back to the early nineties, when it was believed that $\Omega_M \simeq 1$; the catastrophe was represented by the impossibility to have a value of Ω_M close to one.

3.2.6 Radiation and (massless) neutrinos

The Universe is also filled by the cosmic microwave radiation (*Cosmic Microwave Background*, *CMB*), discovered in 1965; it has a thermal, black body spectrum with a temperature, today, $T_\gamma \simeq 2.73K$; this corresponds to a mass-energy density

$$\rho_\gamma = \frac{aT_\gamma^4}{c^2} \simeq 4.7 \cdot 10^{-34} g/cm^3 \qquad \Rightarrow \qquad \Omega_\gamma \simeq 2.5 \cdot 10^{-5} h^{-2}$$

There is also a contribution from a cosmic background of neutrinos. If they have no or negligible mass, they behave like photons, i.e. as relativistic matter, providing

$$\rho_\nu \sim N_\nu \cdot 10^{-34} g/cm^3$$

where N_ν is the number of lepton generations; if $N_\nu = 3$, then $\Omega_\nu \simeq 1.7 \cdot 10^{-5}$. The total contribution Ω_R in the form of relativistic matter, i.e. photons and (massless) neutrinos is then

$$\Omega_R h^2 = (\Omega_\gamma + \Omega_\nu) h^2 \simeq 4.2 \cdot 10^{-5}$$

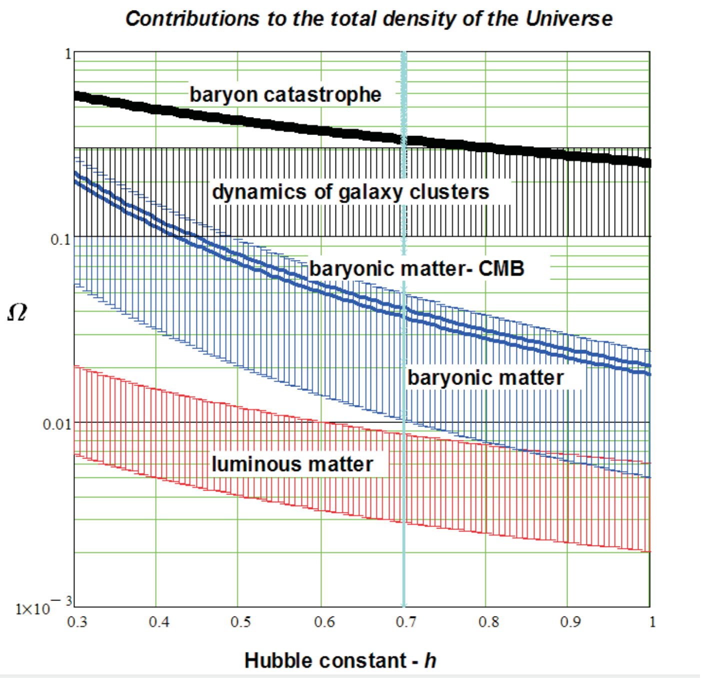
If neutrinos, as we shall see later, have a mass different from zero, their contribution may be larger than this, but then it should be counted among the non-relativistic matter.

We see that, at present, the dominant contribution to the density of mass-energy is provided by matter, so $\rho \sim \rho_M$. As this matter does not possess relativistic motions, pressure $p_0 = p(t_0)$ will be negligible.

3.2.7 Baryonic and non-baryonic dark matter

The following Figure summarizes the conditions described above for the various components of matter in the Universe.

If we take $H_0 \sim 70\, km \cdot s^{-1} \cdot Mpc^{-1}$ we see that, compared to the luminous matter, the total contribution by baryonic matter is larger by at least an order of magnitude. The dynamics of clusters of galaxies and the baryonic catastrophe further increase by another order of magnitude the density.



This shows that in the Universe the vast majority of matter is dark, i.e. not luminous. Moreover, there are dark baryons, but the main contribution to the density of the Universe is due to some form of non-baryonic dark matter.

3.2.8 The cosmological constant

In the last ten years the observations of *SNIa* up to $z \sim 1$, and the observed properties of the *CMB*, both by satellites (*COBE, WMAP*, *Planck*) and by stratospheric balloons, have suggested that the geometry of the space part (at $t = const.$) of the *R&W* metric is consistent with an Euclidean metric ($k = 0$). This is due to the contribution of a cosmological constant Λ different from zero.

It is useful to include the cosmological constant into the energy-momentum tensor by the definition of an effective energy-momentum tensor:

$$\tilde{T}_{\alpha\beta} \equiv T_{\alpha\beta} + \frac{\Lambda c^4}{8\pi G} g_{\alpha\beta} = \left(\tilde{p} + \tilde{\rho} c^2\right) u_\alpha u_\beta - \tilde{p} g_{\alpha\beta} \tag{99}$$

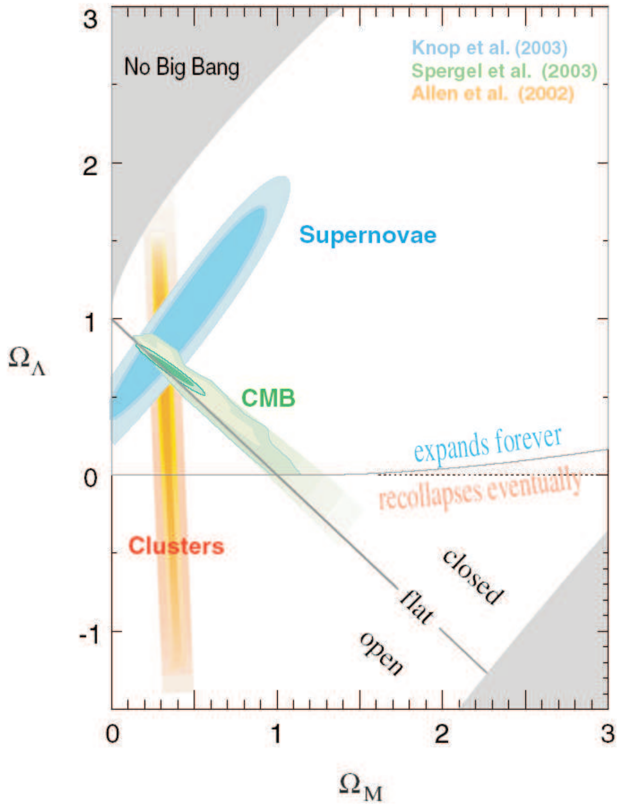
where effective pressure $\tilde{p}$ and effective density $\tilde{\rho}$ are defined as

$$\tilde{p} = p - \frac{\Lambda c^4}{8\pi G} \qquad \text{and} \qquad \tilde{\rho} = \rho + \frac{\Lambda c^2}{8\pi G} = \rho + \rho_\Lambda \tag{100}$$

So Λ can be associated to a density $\rho_\Lambda = \Lambda c^2 / 8\pi G$ and to a density parameter

$$\Omega_\Lambda = \frac{\rho_\Lambda}{\rho_{cr}} = \frac{\Lambda c^2}{8\pi G} \cdot \frac{8\pi G}{3H^2} = \frac{\Lambda c^2}{3H^2} \tag{101}$$

The following Figure shows the observational constraints on the values of Ω_Λ and Ω_M derived from distant *SNIA*, from the *CMB* and from galaxy clusters. The meaning of the various curves will be clarified later [Adapted from: Knop et al., 2003, *The Astrophysical Journal* (*ApJ*) **598**, 102].



As one can see, the confidence regions related to *SNIA* and to the *CMB* intersect almost at right angles, allowing a good determination of $\Omega_\Lambda \simeq 0.7$ and $\Omega_M \simeq 0.3$.

3.3 Peculiar motions

Before we go on with the study of cosmological models, we examine the so-called *peculiar velocites* by using the already computed affine connections. The *peculiar velocity* u^α (4-velocity) is the velocity of a particle with respect to the local co-moving frame. The equation of the geodesic motion is, as usual,

$$\frac{\mathrm{d}u^\alpha}{\mathrm{d}s} + \Gamma^\alpha_{\beta\gamma} u^\beta u^\gamma = 0 \qquad \text{(remember that } u^\alpha \equiv \frac{\mathrm{d}x^\alpha}{\mathrm{d}s})$$

For the $\alpha = 0$ component:

$$\frac{\mathrm{d}u^0}{\mathrm{d}s} + \Gamma^0_{\beta\gamma} u^\beta u^\gamma = 0$$

For the *R&W* metric (see above) the only non-vanishing component of $\Gamma^0_{\beta\gamma}$ is $\Gamma^0_{ij} = -\frac{\dot{a}}{a}\frac{g_{ij}}{c}$, $(i,j = 1,2,3)$, and using the fact that $1 = g_{\alpha\beta}u^\alpha u^\beta = (u^0)^2 + g_{ij}u^i u^j = (u^0)^2 - |\overline{u}|^2$, with $\overline{u}$ the space part of the 4-vector, the geodesic equation becomes

$$\frac{\mathrm{d}u^0}{\mathrm{d}s} = +\frac{\dot{a}}{a}\frac{g_{ij}}{c}u^i u^j = -\frac{\dot{a}}{a}\frac{|\overline{u}|^2}{c}$$

If we differentiate $1 = (u^0)^2 - |\overline{u}|^2$ we have $u^0 \mathrm{d}u^0 = |\overline{u}|\mathrm{d}|\overline{u}|$, and since $u^0 = \frac{\mathrm{d}x^o}{\mathrm{d}s} = c\frac{\mathrm{d}t}{\mathrm{d}s}$ we get

$$\frac{|\overline{u}|}{u^0}\frac{\mathrm{d}|\overline{u}|}{\mathrm{d}s} = -\frac{\dot{a}}{a}\frac{|\overline{u}|^2}{c} \quad \Rightarrow \quad \frac{\mathrm{d}|\overline{u}|}{\mathrm{d}t} = -\frac{\dot{a}}{a}|\overline{u}| \quad \Rightarrow \quad \frac{|\dot{\overline{u}}|}{|\overline{u}|} = -\frac{\dot{a}}{a} \tag{102}$$

This implies that $|\overline{u}| \propto 1/a$, and recalling thet $p^\alpha = m_0 u^\alpha$, $|\overline{p}| \propto 1/a$: the magnitude of the 3-momentum of a freely-propagating particle "red shifts" as a^{-1}.

We see again that the co-moving frame is the most natural one. In fact, in an expanding universe the peculiar velocity (3-velocity) decreases with increasing expansion: the particles tend to move at rest with respect to co-moving observers. We can understand this just thinking that if a particle is moving away from a point at a certain speed it will cross the positons of observers, in motion with respect to the original one, for which the peculiar velocity of the particle will be smaller:

$$D = v_p \cdot \Delta t \qquad v_0' = H_0 D = H_0 v_p \Delta t \qquad v_p' = v_p - v_0'$$

$$\Rightarrow \Delta v_p = v_p' - v_p = -v_0' = -H_0 v_p \Delta t \qquad \Rightarrow \qquad \frac{1}{v_p}\frac{\mathrm{d}v_p}{\mathrm{d}t} = -H_0 = -\frac{\dot{a}_0}{a_0}$$

Since the temperature of an ideal gas is proportional to the square of the mean velocitiy of particles: $T_{gas} \propto |\overline{u}|^2 \propto 1/a^2$ ($T_{rad} \propto 1/a$), consistent with the adiabatic expansion of a perfect gas: $pV^\gamma = const \Rightarrow TV^{\gamma-1} = const \quad \gamma = 5/3, \quad V \propto a^3 \quad \Rightarrow Ta^2 = const.$

3.4 The equation of state

Friedmann equations contain, in addition to density and cosmological constant, another relevant parameter, the *pressure* p of the cosmic fluid. In cosmology the pressure is linked to the density by means of a *barotropic* equation of state such as $p = w\rho c^2$, with $w = const$, $0 \leq w \leq 1$.

As we have seen, the case $w = 0$ corresponds to non relativistic matter for which, even if $p = w(T)\rho c^2$ is non-zero, $p \ll \rho c^2$ (note that pressure is always associated to density in Friedmann equations) , and then $w \simeq 0$. For a non-degenerate, ultrarelativistic fluid in thermal equilibrium the equation of state is $p = \frac{1}{3}\rho c^2$ with , also valid for a photon gas.

The quantity w is also linked to the adiabatic speed of sound (at constant entropy):

$$c_s = \left(\frac{\partial p}{\partial \rho}\right)_S^{1/2} \qquad \Rightarrow \qquad c_s = c\sqrt{w}$$

If $w = 0$, then $c_s = 0$; if $w = 1/3$ then $c_s = c/\sqrt{3}$. We have already seen that, with this equation of state, $\rho_w V^{1+w} = const$, but $V \propto a^3 \quad \rightarrow \quad \rho_w a^{3(1+w)} = const = \rho_{0w} a_0^{3(1+w)}$ (the suffix 0 implies $t = t_0$):

$$\rho_w = \rho_{0w} (\frac{a_0}{a})^{3(1+w)} \tag{103}$$

- $w = 0 \quad \rightarrow \quad \rho_M a^3 = \rho_{0M} a_0^3; \qquad \frac{a_0}{a} = 1 + z \qquad \Rightarrow \quad \rho_M = \rho_{0M}(1+z)^3$
- $w = 1/3 \quad \rightarrow \quad \rho_R a^4 = \rho_{0R} a_0^4 \qquad \Rightarrow \quad \rho_R = \rho_{0R}(1+z)^4$

Regarding the behavior of the cosmological constant, we refer to the effectitive quantities defined in Eq. 100. If we imagine that the pressure and density of matter and radiation are negligible, we obtain

$$\tilde{p} \equiv p_\Lambda = -\frac{\Lambda c^4}{8\pi G} \qquad \text{e} \qquad \tilde{\rho} \equiv \rho_\Lambda = +\frac{\Lambda c^2}{8\pi G}$$

which gives

$$p_\Lambda = -\rho_\Lambda c^2 \qquad \Longrightarrow \qquad w_\Lambda = -1$$

Thus we see that the cosmological constant is characterized by an equation of state with $w = w_\Lambda = -1$. A similar case, as we shall see, is present during the phase of *inflation* that occurs in the early Universe. If $w = -1$, Eq. 103 tells that ρ_w (in this case ρ_Λ) is constant in time: $\rho_\Lambda = const.$

Eq. 103 can be derived also from ($F3$) in the form given by Eq. 97

$$\mathrm{d}\big(\rho c^2 a^3\big) + p\,\mathrm{d}\big(a^3\big) = 0$$

In fact, by writing $p = w\rho c^2$,

$$\mathrm{d}\big(\rho c^2 a^3\big) + w\rho c^2\,\mathrm{d}\big(a^3\big) = 0$$

$$a^3 \mathrm{d}(\rho c^2) + \rho c^2 (1+w) \cdot 3a^2 \mathrm{d}a = 0$$

$$\frac{\mathrm{d}\rho}{\rho} = -3(1+w)\frac{\mathrm{d}a}{a}$$

which, by integration with $w = cost$, gives

$$\int_{\rho}^{\rho_0} \mathrm{d}\ \ln\rho = -3(1+w)\int_{a}^{a_0} \mathrm{d}\ \ln a$$

and again

$$\rho_w = \rho_{0w}(\frac{a_0}{a})^{3(1+w)}$$

If w is not constant in time (as occurs in some theories that, in place of the cosmological constant, consider a scalar field variable in time to explain the origin of so-called *dark energy*) tje factor $1+w$ cannot be taken out of the above integral and the formal solution is

$$\rho(a) = \rho_0 \exp\left\{3\int_{a}^{a_0}[1+w(a)]\ \mathrm{d}\ \ln a\right\}$$

If we want to use the redshift z, remember that $a = a_0/(1+z)$, so $\mathrm{d}a = -a_0/(1+z)^2\mathrm{d}z$ and

$$\rho(z) = \rho_0 \exp\left\{3\int_{z}^{0}[1+w(z)]\ \frac{-a_0}{(1+z)^2}\ \frac{1+z}{a_0}\ \mathrm{d}z\right\}$$

$$\rho(z) = \rho_0 \exp\left\{3\int_{0}^{z}\frac{[1+w(z)]}{1+z}\ \mathrm{d}z\right\} \tag{104}$$

3.5 A useful relation among cosmological parameters

We start from $(F1)$ and divide it by a^2:

$$\begin{aligned}
\dot{a}^2 + kc^2 &= \frac{8\pi G}{3}\rho a^2 + \frac{1}{3}a^2c^2\Lambda \qquad /\cdot\frac{1}{a^2}\\
\frac{\dot{a}^2}{a^2} + \frac{kc^2}{a^2} &= \frac{8\pi G}{3}\rho\frac{H^2}{H^2} + \frac{1}{3}c^2\Lambda\frac{H^2}{H^2}\\
H^2 + \frac{kc^2}{a^2} &= H^2\cdot\frac{\rho}{\rho_{cr}} + H^2\cdot\frac{c^2\Lambda}{3H^2}\\
\frac{kc^2}{a^2} &= H^2\left[\frac{\rho}{\rho_{cr}} + \frac{c^2\Lambda}{3H^2} - 1\right]
\end{aligned}$$

Remember that $\Lambda c^2/3H^2 = \Omega_\Lambda$, $\rho = \rho_M + \rho_R$ and then

$$\frac{kc^2}{a^2} = H^2\left[\Omega_M + \Omega_R + \Omega_\Lambda - 1\right]$$

We name $\Omega \equiv \Omega_M + \Omega_R + \Omega_\Lambda$, and get

$$\frac{kc^2}{a^2} = H^2\left[\Omega - 1\right] \tag{105}$$

This relation holds at any time; in particular, today (at $t = t_0$, with $H = H_0$ and $\Omega = \Omega_0$) we have

$$\frac{kc^2}{a_0^2} = H_0^2\left[\Omega_0 - 1\right] \tag{106}$$

This relation shows the link between the density parameter Ω_0 and the curvature k of the space section at constant cosmic time:

- $\Omega_0 > 1 \quad \Rightarrow \quad k = +1$
- $\Omega_0 = 1 \quad \Rightarrow \quad k = 0$
- $\Omega_0 < 1 \quad \Rightarrow \quad k = -1$

Eq. 106 express also (for $k \neq 0$) the present value of the scale factor:

$$a_0 = \frac{c}{H_0}\sqrt{\frac{k}{\Omega_0 - 1}}$$

3.6 The Hubble parameter

Let's consider again $(F1)$, divide it by a_0^2, and rimember that $kc^2/a_0^2 = H_0^2\,[\Omega_0 - 1]$, $\Omega_0 \equiv \sum_w \Omega_{0w}$, $\Omega_{0w} \equiv \rho_{0w}/\rho_{0cr}$; we include the cosmological constant into ρ through the relation $\rho_\Lambda = \Lambda c^2/8\pi G$:

$$\dot{a}^2 + kc^2 = \frac{8\pi G}{3}\rho a^2 \qquad / \cdot \frac{1}{a_0^2}$$

$$\frac{\dot{a}^2}{a_0^2} - \frac{8\pi G}{3H_0^2}\,H_0^2\,\rho\left(\frac{a}{a_0}\right)^2 = -\frac{kc^2}{a_0^2}$$

where $\rho \equiv \sum_w \rho_w = \sum_w \rho_{0w}(\frac{a_0}{a})^{3(1+w)} = \rho_{0\,cr}\sum_w \Omega_{0w}(\frac{a_0}{a})^{3(1+w)}$. Then

$$\frac{\dot{a}^2}{a_0^2} - H_0^2\sum_w \Omega_{0w}(\frac{a_0}{a})^{3(1+w)}\left(\frac{a}{a_0}\right)^2 = -H_0^2\left[\sum_w \Omega_{0w} - 1\right]$$

$$\frac{\dot{a}^2}{a_0^2} = H_0^2\left[\sum_w \Omega_{0w}(\frac{a_0}{a})^{1+3w} + \left(1 - \sum_w \Omega_{0w}\right)\right] \qquad (107)$$

If we remember that $H(t) \equiv \dot{a}/a_0$, and we multiply by $(a_0/a)^2$, the previous relation gives

$$H^2(t) = H_0^2\left(\frac{a_0}{a}\right)^2\left[\sum_w \Omega_{0w}\left(\frac{a_0}{a}\right)^{1+3w} + (1 - \sum_w \Omega_{0w})\right] \qquad (108)$$

Since $a_0/a = 1+z$, this gives the dependence on redshift $H(z)$

$$H^2(z) = H_0^2(1+z)^2\left[\sum_w \Omega_{0w}(1+z)^{1+3w} + (1 - \sum_w \Omega_{0w})\right] \qquad (109)$$

Making explicit the different components we finally get

$$H^2(z) = H_0^2(1+z)^2\Big[\Omega_R(1+z)^2 + \Omega_M(1+z) + \Omega_\Lambda(1+z)^{-2} + 1 - (\Omega_R + \Omega_M + \Omega_\Lambda)\Big] \qquad (110)$$

3.7 The three eras of the Universe

In Eq. 110, describing the evolution of $H(z)$, and also the evolution of the scale factor a, there are three different contributions depending on Ω_R, Ω_M, Ω_Λ. These contributions evolve differently in redshift.

At high z the term containing Ω_Λ, as well as the term $1 - (\Omega_R + \Omega_M + \Omega_\Lambda)$ which is on the order of unity, are negligible if compared to the terms containing Ω_R and Ω_M. Moreover, the term depending on Ω_R grows more rapidly and, even if today $\Omega_R \ll \Omega_M$, relativistic matter dominates the dynamics of the Universe before the epoch of *equivalence* (or *matter/radiation equality*), corresponding to

$$\Omega_R(1+z_{eq})^2 = \Omega_M(1+z_{eq})$$

$$1 + z_{eq} = \frac{\Omega_M}{\Omega_R} \simeq 23800\,\Omega_M h^2 \qquad (111)$$

which, for $\Omega_M \simeq 0.3$ and $h \simeq 0.7$, gives $z_{eq} \simeq 3700$. We have first, starting with the birth of the Universe, an era during which the dynamics is dominated by relativistic matter. Then, after the equivalence, non-relativistic matter dominates. This era lasts until the cosmological constant enters the game, that is when

$$\Omega_M(1+z) = \Omega_\Lambda(1+z)^{-2}$$

i.e. at a redshift z_Λ given by

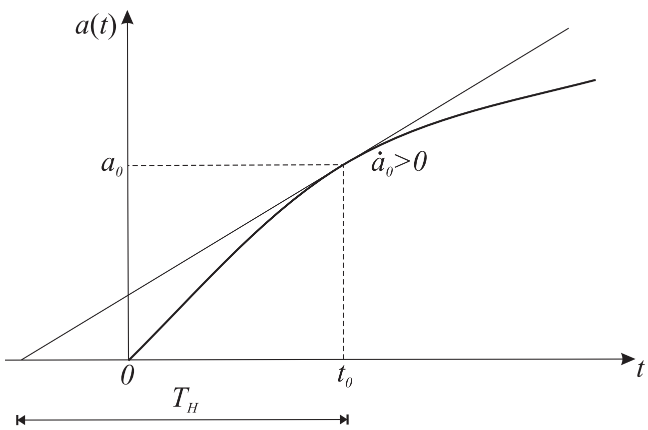
$$1 + z_\Lambda = \left(\frac{\Omega_\Lambda}{\Omega_M}\right)^{\frac{1}{3}} \tag{112}$$

which, for $\Omega_\Lambda = 0.7$ and $\Omega_M = 0.3$, gives $z_\Lambda = 0.33$.

Summarizing, we have three *phases* (*eras*): a first era dynamically dominated by relativistic matter (radiation) (*radiation dominated - RD - era*); a second phase dominated by non-relativistic matter (*matter dominated - MD - era*); a third phase dominated by the cosmological constant (or by some form of dark energy) named *dark-energy dominated era* or *vacuum dominated* (*VD*) *era*. During the *RD* era, as we shall see, an inflationary phase occurs, dynamically similar to the *VD* era.

3.8 Hubble time

Suppose that, at a time (for instance at $t = t_0$), $\dot{a} > 0$ (expansion); Eq.($F2$) tells us that if $(\rho + 3p/c^2) > 0$ (i.e. if $(1 + 3w)\rho > 0$, $w > -1/3$) $\ddot{a}$ is always < 0, the graph of $a(t)$ has the concavity facing downwards, and $a(t)$ must be zero at a certain instant, that we can take as $t = 0$. At $t = 0$ ρ and H diverge, and we have a singularity, the so-called *Big Bang*.



We also see that $a_0/T_H = \dot{a}_0 \quad \Rightarrow \quad 1/T_H = H_0 \quad$ e $\quad T_H > t_0 \quad$ that is $\quad H_0 t_0 \leq 1$: the inverse of H_0 gives an upper limit to the age of the Universe (T_H is called *Hubble time* $= 1/H_0$). This is no more true if the expansion is dominated by the cosmological constant, as we shall see.

We also note that, if at any moment is $\dot{a} < 0$, the concavity of $a(t)$ implies that in the future there will be an unstoppable collapse: the *Big Crunch*. We finally note that the effect of expansion is not due to the pressure, which always acts in the direction to decelerate the expansion, if $w > -1/3$.

3.9 Evolution of the density parameter Ω

If we divide Eq. 105, referred to a generic time,

$$\frac{kc^2}{a^2} = H^2(\Omega - 1)$$

by the corresponding Eq. 106, referred to time $t = t_0$, we get

$$\frac{a_0^2}{a^2} = \frac{H^2}{H_0^2}\frac{\Omega - 1}{\Omega_0 - 1} \qquad \Rightarrow \qquad \Omega - 1 = (\Omega_0 - 1)\frac{H_0^2 a_0^2}{H^2 a^2}\,.$$

The time evolution of H (Eq. 108)gives:

$$\Omega - 1 = \frac{\Omega_0 - 1}{1 - \Omega_0 + \sum_w \Omega_{0w}(1+z)^{1+3w}}$$

Making explicit the three components R, M, Λ:

$$\Omega - 1 = \frac{\Omega_0 - 1}{\Omega_R(1+z)^2 + \Omega_M(1+z) + \Omega_\Lambda(1+z)^{-2} + 1 - \Omega_0} \tag{113}$$

which gives the time evolution of $\Omega(z)$. We first see that, since the denominator of the right hand side is always positive (see Eq. 110) the sign of $\Omega(z) - 1$ does not change during the evolution. So if $\Omega_0 > 1$, $\Omega(z)$ is always greater than one through cosmic history. Similarly if $\Omega_0 < 1$; if $\Omega_0 = 1$ it keeps that value at all times. This effect is strictly linked to the fact that the curvature$k$ cannot change during cosmic evolution.

If we go back in time, for $z \to \infty$, $\Omega - 1 \to 0$, i.e. $\Omega \to 1$: going back into the past, the Universe increasingly resembles that with $k = 0$ and the effects of the curvature are negligible in the early stages of cosmic evolution.

The fact that Ω tends to diverge from 1 as time goes on, while today it seems to be very close to 1, requires that in the distant past Ω was actually extremely close to 1, with considerable *"fine tuning"* between density and expansion rate. This is the so-called **flatness problem**, which is solved by the paradigm of inflation. The existence of a phase of inflation, dominated by the energy density of a false vacuum that mimics the effects of a cosmological constant, provides the mechanism through which Ω is so forced towards unity in the early Universe, that it stays up to date very near to 1. To understand the reason for the cosmological constant forces Ω to one, let's look at Eq. 113. Remember that $a_0/a = 1 + z$, so that in the future, when $a \to \infty$, $z \to -1$. Now let $z \to -1$ in Eq. 113: the numerator of the left hand side diverges and $\Omega - 1 \to 0$, i.e. $\Omega \to 1$. Something similar happened during inflation.

3.10 Evolution of the deceleration parameter $q(z)$

The deceleration parameter, defined by Eq. 77, by using the definition of the Hubble parameter, can be written

$$q(t) \equiv -\frac{\ddot{a}(t)\,a(t)}{\dot{a}(t)^2} = -\frac{\ddot{a}(t)\,a(t)^2}{a(t)\,\dot{a}(t)^2} = -\frac{\ddot{a}(t)}{a(t)\,H(t)^2} \tag{114}$$

We use again Eq. 107; time t_0 corresponds to a generic reference time, not necessarily to the presen time. We rewrite Eq. 107 as

$$\dot{a}^2 = H_0^2 a_0^2 \left[\sum_{w_i} \Omega_{0w_i} (\frac{a_0}{a})^{1+3w_i} + \left(1 - \sum_{w_i} \Omega_{0w_i}\right)\right] \tag{115}$$

and derive it with respect to time. We obtain

$$2\dot{a}\ddot{a} = H_0^2 a_0^2 \left[\sum_{w_i} \Omega_{0w_i} \,(1+3w_i)\,(\frac{a_0}{a})^{3w_i} \cdot \frac{-a_0\dot{a}}{a^2}\right]$$

$$\ddot{a} = -\frac{H_0^2 a_0^3}{2a^2}\left[\sum_{w_i} \Omega_{0w_i}\,(1+3w_i)\,(\frac{a_0}{a})^{3w_i}\right] = -\frac{H_0^2 a_0^3}{a^2}\left[\frac{1}{2}\sum_{w_i} \Omega_{0w_i} (\frac{a_0}{a})^{3w_i} + \frac{3}{2}\sum_{w_i} w_i \Omega_{0w_i} (\frac{a_0}{a})^{3w_i}\right] \tag{116}$$

We insert this in Eq. 114 and we have

$$q = -\frac{\ddot{a}}{a\,H^2} = \frac{H_0^2 a_0^3}{H^2 a^3}\left[\frac{1}{2}\sum_{w_i} \Omega_{0w_i} (\frac{a_0}{a})^{3w_i} + \frac{3}{2}\sum_{w_i} w_i \Omega_{0w_i} (\frac{a_0}{a})^{3w_i}\right].$$

If we now take $t = t_0$ this relation simplifies and we get

$$q_0 = \frac{1}{2}\sum_{w_i} \Omega_{0w_i} + \frac{3}{2}\sum_{w_i} w_i \Omega_{0w_i}. \tag{117}$$

But, as we already said, time t_0 is just a suitable reference time, and the above relation holds at any time or redshift, provided we use the density parameters corresponding to that reference time. We put $\Omega_{tot}(z) \equiv \sum_{w_i} \Omega_{w_i}(z)$ and finally get

$$q(z) = \frac{1}{2}\Omega_{tot}(z) + \frac{3}{2}\sum_{w_i} w_i \Omega_{w_i}(z). \tag{118}$$

We apply now this formula to a particular Universe. We neglect Ω_R (i.e. $z \ll z_{eq}$), and assume $k = 0$. So $\Omega_{tot}(z) \equiv 1$ at any time and $\Omega_M(z) + \Omega_\Lambda(z) = 1$. Eq. 118 gives

$$q(z) = \frac{1}{2} - \frac{3}{2}\Omega_\Lambda(z). \tag{119}$$

How do we estimate $\Omega_\Lambda(z)$? In our particular case Eq. 110 becomes

$$H^2(z) = H_0^2(1+z)^2\left[\Omega_M(1+z) + \Omega_\Lambda(1+z)^{-2}\right] = H_0^2(1+z)^3\left[\Omega_M + \frac{\Omega_\Lambda}{(1+z)^3}\right].$$

Now $\Omega_\Lambda(z) = 1 - \Omega_M(z)$ and, by definition,

$$\Omega_M(z) = \frac{\rho_M(z)}{\rho_{cr}(z)} = \frac{\rho_{0M}(1+z)^3}{\rho_{0,cr}\cdot {}^{H^2}/_{H_0^2}}, \tag{120}$$

where we used the relation

$$\rho_{cr} \equiv \frac{3H^2}{8\pi G} = \frac{3H_0^2}{8\pi G}\cdot\frac{H^2}{H_0^2} = \rho_{0,cr}\cdot\frac{H^2}{H_0^2}.$$

So Eq. 120 can be written

$$\Omega_M(z) = \Omega_M \frac{(1+z)^3}{(1+z)^3\left[\Omega_M + \frac{\Omega_\Lambda}{(1+z)^3}\right]} \tag{121}$$

and

$$\Omega_\Lambda(z) = 1 - \frac{\Omega_M}{\left[\Omega_M + \frac{\Omega_\Lambda}{(1+z)^3}\right]} = 1 - \frac{\Omega_M(1+z)^3}{\Omega_M(1+z)^3 + \Omega_\Lambda} = \frac{\Omega_\Lambda}{\Omega_\Lambda + (1-\Omega_\Lambda)(1+z)^3}. \tag{122}$$

Finally, Eq. 119 can be written

$$q(z) = \frac{1}{2} - \frac{3}{2}\left[\frac{\Omega_\Lambda}{\Omega_\Lambda + (1-\Omega_\Lambda)(1+z)^3}\right]. \tag{123}$$

We see that, for $z \gg 1$, $q \to 1/2$, i.e. $\ddot{a} < 0$, the expansion slows dows. For $z = 0$, since $\Omega_\Lambda \sim 0.7$, $q < 0$, the expansion is accelerated. Eq. 123 tells us that $q = 0$ for a particular value of $z = \widetilde{z}$ given by

$$\widetilde{z} = \left(\frac{2\Omega_\Lambda}{1-\Omega_\Lambda}\right)^{1/3} - 1 \tag{124}$$

which, for $\Omega_\Lambda = 0.7$, gives $\widetilde{z} = 0.67$. We can compare this result with Eq. 112, another way to define the beginning of the era dominated by the cosmological constant (or dark energy).

3.11 Cosmological models

Friedmann equations $(F1)$ and $(F2)$

$$\dot{a}^2 + kc^2 = \frac{8\pi G}{3}\rho a^2 + \frac{1}{3}a^2c^2\Lambda \tag{F1}$$

$$\ddot{a} = -\frac{4\pi G}{3}(\rho + \frac{3p}{c^2})a + \frac{1}{3}\Lambda c^2 a \tag{F2}$$

allow a qualitative classification of cosmological models with $\Lambda \neq 0$, for different values of the curvature parameter k. We assume for simplicity that $\rho \equiv \rho_M = \varrho_{0M}(a_0/a)^3$, i.e. we are in the *MD* era (a more general treatment leads to the same qualitative conclusions).

Let's first consider eq. $(F2)$ and assume that $\ddot{a} = 0$; we get (pressure is negligible in *MD* era)

$$\frac{1}{3}\Lambda c^2 a = \frac{4\pi G}{3}(\rho + \frac{3p}{c^2})a \qquad /\cdot\frac{3}{c^2 a} \tag{125}$$

$$\Lambda(a) = \frac{4\pi G}{c^2}\rho_{0M}\frac{a_0^3}{a^3} \equiv \frac{B}{2a^3} \tag{126}$$

We draw this critical curve in the Λ versus a plane, as in the following figure (the dashed line). This curve is contained in the region $\Lambda > 0$. It marks the border between the region where $\ddot{a} < 0$ and that where $\ddot{a} > 0$.

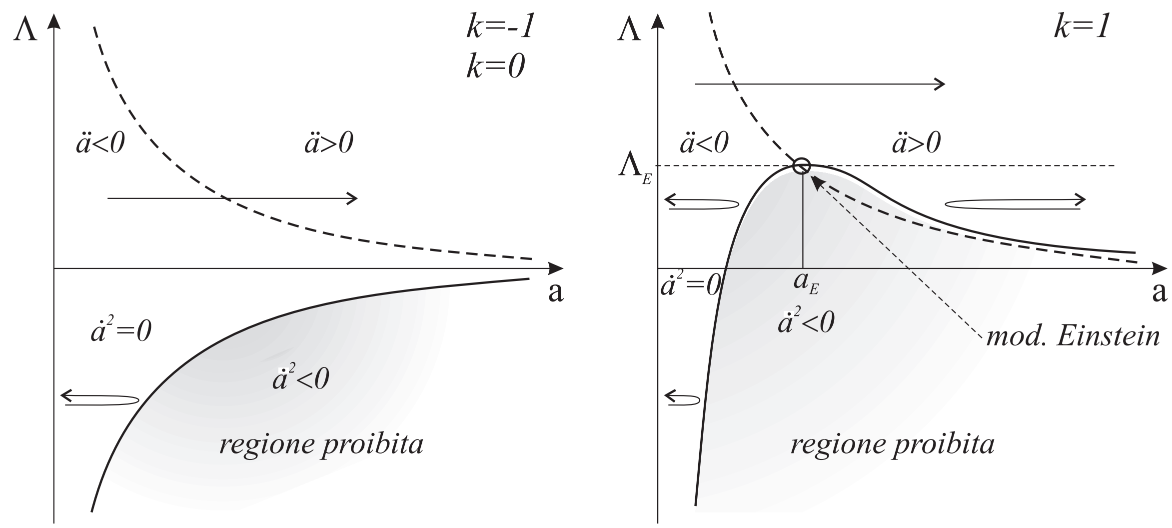
Let's now consider eq. $(F1)$ and assume that, in this case, $\dot{a}^2 = 0$. We act as for eq. $(F2)$ and we get

$$\frac{1}{3}\Lambda c^2 a^2 = kc^2 - \frac{8\pi G}{3}\rho a^2 \qquad / \cdot \frac{3}{c^2 a^2} \tag{127}$$

$$\Lambda(a) = \frac{3k}{a^2} - \frac{8\pi G}{c^2}\rho_{0M}\frac{a_0^3}{a^3} \equiv \frac{A}{a^2} - \frac{B}{a^3}. \tag{128}$$

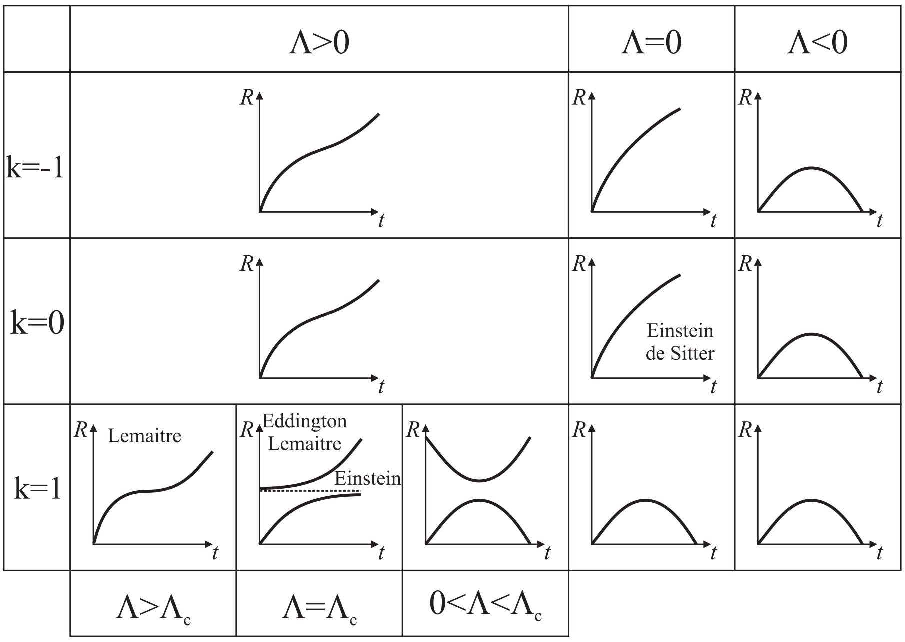
Here we must distinguish two cases: $k = 0$ or -1, and $k = +1$. In the first case the right hand side of eq. 128 is always negative, as shown on the left side of the following figure. The curve corresponding to $\dot{a}^2 = 0$ marks the border between the region of the (Λ, a) plane where $\dot{a}^2 > 0$ and that where $\dot{a}^2 < 0$, forbidden by dynamics. So, for a given vale of $\Lambda < 0$, a starts from zero and grows (it moves horizontally in the plane) until the crossing of the $\dot{a}^2 = 0$ line. Since it cannot enter the forbidden region, the scale factor must go back to zero: the Universe recollapses. On the contrary, for $\Lambda > 0$, the scale factor can grow indefinitely; it starts in the $\ddot{a} < 0$ region and self-gravitation slows down expansion, but after the crossing of the $\ddot{a} = 0$ line, in the $\ddot{a} > 0$ region, the expansion accelerates. If $\Lambda \equiv 0$, $\ddot{a}$ is always negative and expansion is slowed forever.

The other case, $k = +1$, is more complex. According to eq. 128, for small values of a, the $-B/a^3$ term prevails and $\Lambda(a) < 0$, while for large values of a the A/a^2 term, positive, prevails. So the $\dot{a}^2 = 0$ line crosses the $\Lambda = 0$ line and has a maximum at a particular value of a: $a = a_E = 3B/2A$, corresponding to $\Lambda = \Lambda_E$. Moreover, the two curves defined by eq. 126 and eq. 128 cross at the point ($a = a_E$, $\Lambda = \Lambda_E$), as can be easily checked. So at this point we have both $\dot{a}^2 = 0$ and $\ddot{a} = 0$: this correspond to the static **Einstein model** (see below for more details). This situation is described on the right side of the following figure. For $\Lambda < 0$, and for $\Lambda \equiv 0$, the Universe has to recollapse. For $\Lambda > 0$ we have three possibilities. For $\Lambda > \Lambda_E$ we have again an initial deceleration followed by acceleration and expansion to infinity. Among the cases with $\Lambda > \Lambda_E$ there is the so-called **Lemaitre model**, in which, if $\Lambda = \Lambda_E(1+\epsilon)$ with $\epsilon \ll 1$, the evolution of the scale factor can have an arbitrarily long stationary phase (the smaller is ϵ, the longer this almost-static phase). This model was invoked in 1967 to explain an observed excess of quasars[17] at $z \sim 2$ (this observational evidence is today ascribed to the evolution of quasars). In this model $T_H \equiv 1/H > t_0$ is violated, as it is in many models with $\Lambda > 0$. If $0 < \Lambda < \Lambda_E$, depending on the initial condition, we may have a contraction phase, followed by an expansion era **(bouncing universe)**, or still a model that recolapses, such as those with $\Lambda < 0$.



The following figure shows, in summary, all the cases above discussed.

[17]Quasars are active galactic nuclei, i.e. supermassive black holes hosted in the center of galaxies and emitting a huge amount of energy



Now we will start to see in detail some models that are interesting for historical reasons or because they may be useful approximations in certain stages of cosmic evolution.

3.11.1 Einstein model

If we include the cosmological constant into the energy-momentum tensor, as done in eq. 100, eq. $(F1)$ and eq. $(F2)$ become:

$$\dot{a}^2 + kc^2 = \frac{8\pi G}{3}\tilde{\rho}a^2 \qquad (F1)$$

$$\ddot{a} = -\frac{4\pi G}{3}(\tilde{\rho} + \frac{3\tilde{p}}{c^2})a \qquad (F2)$$

These equations have a static solution with $\ddot{a} = 0$ and $\dot{a} = 0$ if:

$$\tilde{\rho} = -\frac{3\tilde{p}}{c^2} = \frac{3kc^2}{8\pi Ga^2}$$

If this model corresponds to our Universe as it is today, dominated by non relativistic matter, then matter pressure is negligible and $p \simeq 0$, so that:

$$-\frac{3\tilde{p}}{c^2} = -\frac{3}{c^2}\Big(-\frac{\Lambda c^4}{8\pi G}\Big) = \frac{3kc^2}{8\pi Ga^2} \qquad \Rightarrow \qquad \Lambda = \frac{k}{a^2}$$

$$\tilde{\rho} = \rho + \frac{\Lambda c^2}{8\pi G} = \frac{3kc^2}{8\pi Ga^2} \qquad \Rightarrow \qquad \rho = \frac{kc^2}{4\pi Ga^2}$$

Since $\rho > 0$ we have $k = +1$ and $\Lambda > 0$. The value of Λ which makes the Universe static is

$$\Lambda_E = \frac{k}{a^2} = \frac{4\pi G\rho}{c^2} \qquad \text{and} \quad a = a_E = \frac{c}{\sqrt{4\pi G\rho}}$$

This model is unstable because even small fluctuations in density may lead both to local collapse or local expansion. There are two models, the so-called ***Eddington-Lemaitre models***, with $\Lambda = \Lambda_E$ and $k = +1$, whose asymptote is the Einstein model. The first starts from $a = 0$ and tends asymptotically to a_E, with $a < a_E$. The second one starts from $a = a_E$ and, after a long time (tending to infinity) slowly diverges with $a > a_E$.

3.11.2 de Sitter model

This model is empty ($p = 0$ and $\rho = 0$) and flat ($k = 0$). In this case eq. $(F1)$ is simply

$$\dot{a}^2 = \frac{\Lambda}{3}c^2a^2 \qquad \Rightarrow \qquad a(t) = A\, e^{\sqrt{\frac{\Lambda}{3}}ct} \qquad \text{with} \quad H = \frac{\dot{a}}{a} = \sqrt{\frac{\Lambda}{3}}c = const$$

This model, characterized by an equation of state $\tilde{p} = -\tilde{\rho}c^2$, describes the *inflation* phase in the early universe. It also represents the asymptotic behavior of the models with $\Lambda > 0$, as can be seen by examining eq. (F1) and by letting a grow to infinity. Therefore, this model also represents the asymptotic behavior of the cosmological model which currently has more credit.

3.12 Einstein-de Sitter model

In this model (***EdS***) we neglect the cosmological constant and it is assumed that the dynamics is dominated by a single component (radiation or matter), with $\Omega_{0w} \equiv 1$ (this, as seen above, implies that $\Omega_w(z) \equiv 1$ always); i.e. $k = 0$. To be more precise, the *EdS* model has $k = 0$ and $w = 0$, but in general we call by that name even models with $w \neq 0$. Then also $\Lambda = 0$. We will have, from eq. $(F1)$, in the phase dominated by the w component:

$$\frac{\dot{a}^2}{a_0^2} = H_0^2\left(\frac{a_0}{a}\right)^{1+3w} = H_0^2(1+z)^{1+3w} \tag{129}$$

that is

$$a^{\frac{1+3w}{2}}\mathrm{d}a = C\mathrm{d}t \quad \Rightarrow \quad a^{\frac{3(1+w)}{2}} = C' \cdot t$$

and, referring to a_0 and t_0:

$$a(t) = a_0\left(\frac{t}{t_0}\right)^{\frac{2}{3(1+w)}}. \tag{130}$$

The Universe expands forever. Eq. 130 can be written as a function of z:

$$t = t_0(1+z)^{-\frac{3(1+w)}{2}} \tag{131}$$

For this model we also have:

$$\begin{aligned}
&H \equiv \frac{\dot{a}}{a} = \frac{2}{3(1+w)t} \quad \rightarrow \quad Ht = const \quad \rightarrow \quad H = \frac{H_0 t_0}{t} = H_0(1+z)^{\frac{3(1+w)}{2}}\\
&q \equiv -\frac{\ddot{a}a}{\dot{a}^2} = -\frac{2}{3(1+w)}\frac{\dot{a}t - a}{t^2}\frac{a}{\dot{a}^2} = \frac{1+3w}{2} = const = q_0\\
&t_{0cr,w} \equiv t_0 = \frac{2}{3(1+w)H_0} \qquad \text{from the first relation}\\
&\rho_w a^{3(1+w)} = const \quad \rightarrow \quad \frac{\rho_w}{\rho_{0w}} = \left(\frac{a}{a_0}\right)^{-3(1+w)} = \left(\frac{t}{t_0}\right)^{-2} \quad \rightarrow \quad \rho_w(t) = \frac{\rho_{0w}t_0^2}{t^2}.
\end{aligned}$$

Since $\Omega_{0w} = \Omega_w = 1$, it follows that $\rho = \rho_{cr} = \frac{3H^2}{8\pi G}$, and

$$\rho_w(t) = \frac{3H_0^2}{8\pi G}\left(\frac{2}{3(1+w)H_0}\right)^2\frac{1}{t^2} = \frac{1}{6\pi G(1+w)^2t^2} \tag{132}$$

It is useful to explicitly write these relations in the two cases:

- $w = 0$ "dust", matter-dominated Universe (= *non relativistic fluid*)

$$a(t) = a_0\left(\frac{t}{t_0}\right)^{\frac{2}{3}} \qquad t = t_0(1+z)^{-\frac{3}{2}} \qquad H = \frac{2}{3t} = H_0(1+z)^{\frac{3}{2}}$$

$$q_0 = \frac{1}{2} \qquad t_{0cr,m} = t_0 = \frac{2}{3}\frac{1}{H_0} \qquad \rho_m(t) = \frac{1}{6\pi G t^2}$$

- $w = 1/3$ relativistic matter, radiation dominated Universe

$$a(t) = a_0\left(\frac{t}{t_0}\right)^{\frac{1}{2}} \qquad t = t_0(1+z)^{-2} \qquad H = \frac{1}{2t} = H_0(1+z)^2$$

$$q_0 = 1 \qquad t_{0cr,r} = t_0 = \frac{1}{2H_0} \qquad \rho_r = \frac{3}{32\pi G t^2}$$

Note that when the pressure increases (from $w = 0$ to $w = 1/3$) the deceleration parameter q_0 grows.

This model can also be extended to more general cases. Indeed, in the phases dominated by radiation or matter, when the contribution of Ω_Λ is negligible as well as the term of curvature, the evolution of $a(t)$ is given by

$$\frac{\dot{a}^2}{a_0^2} = H_0^2 \Omega_{0w} \left(\frac{a_0}{a}\right)^{1+3w} \tag{133}$$

similar to eq. 129, but with an effective Hubble constant $H_{0,eff}$ given by

$$H_{0,eff} = H_0 \sqrt{\Omega_{0w}}$$

So, at high z, we can use the relations derived for *EdS* model just by putting $H_0 \to H_{0,eff}$; for instance:

$$t(z) = \frac{2}{3(1+w)H_0\sqrt{\Omega_{0w}}}(1+z)^{-\frac{3(1+w)}{2}} \tag{134}$$

$$H(z) = H_0\sqrt{\Omega_{0w}}(1+z)^{\frac{3(1+w)}{2}} \tag{135}$$

The relations giving $\rho_w(t)$ (in which the terms with H_0 are simplified) and q (which does not depend on H_0).

These relations are useful to have some *rough estimates* of the correct values of cosmological quantities. Moreover, at high z the *EdS* model is an excellent approximation of the real model of Universe, whatever its curvature.

For instance, we can use eq. 134 to estimate the age of the Universe at *matter/radiation equality*, at $z = z_{eq}$ given by eq. 111. Since z_{eq} is at the edge of the *MD* era we use eq. 134 with $w = 0$:

$$t(z_{eq}) =\simeq \frac{2}{3H_0\sqrt{\Omega_M}(1+z_{eq})^{\frac{3}{2}}} \simeq 1.8 \cdot 10^3 (\Omega_M h^2)^{-2} \; years, \tag{136}$$

about $8 \cdot 10^4$ years for $\Omega_M = 0.3$ and $h = 0.7$. A more precise calculation gives an age of about $5 \cdot 10^4$ years.

3.13 Matter dominated models

As we have just seen, the epoch of matter/radiation equality corresponds to an age of the Universe of about 50000 years, much less then about 13.5 billion years (see below), the present age of the cosmos in which we live. If we neglect this "small" amount of time (if compared to the total age), and suppose that the cosmological constant is zero or negligible, we obtain the classic cosmological models listed in all the texts, popular and not, more than ten years old. So let's see in detail these models, pointing out that the *EdS* model dominated by matter is already one of these cases, the one with $\Omega \equiv \Omega_M = 1$.

For matter dominated models eq. 107 gives

$$\left(\frac{\dot{a}}{a_0}\right)^2 \simeq H_0^2 \left[\Omega_M \frac{a_0}{a} + 1 - \Omega_M\right] \tag{137}$$

Let's see the two cases $\Omega_M > 1$ and $\Omega_M < 1$:

- $\Omega_M > 1$: Since $1 - \Omega_M < 0$, while $\Omega_M(a_0/a)$ becomes smaller and smaller as a grows, there is a value of the scale factor which makes $\dot{a} = 0$; larger values of a would make $\dot{a}$ an imaginary number. This means that the scale factor has a maximum value a_m at the time $t = t_m$, given by

$$a(t_m) = a_m = a_0 \frac{\Omega_M}{\Omega_M - 1} \tag{138}$$

The evolution of the scale factor can be derived by using an auxiliary parameter, the *development angle* θ, defined by the relation $(\frac{a}{a_0}) \cdot \frac{2(\Omega_M - 1)}{\Omega_M} \equiv 1 - \cos\theta$, $\theta \in [0, 2\pi]$. The parametric solution is:

$$H_0 t = \frac{\Omega_M}{2(\Omega_M - 1)^{3/2}}(\theta - \sin\theta) \qquad a(t) = a_0 \frac{\Omega_M}{2(\Omega_M - 1)}(1 - \cos\theta) \tag{139}$$

This is the parametric equation of a cycloid. The maximum is obtained for $\theta = \theta_m = \pi$:

$$a(t_m) = a_m = a_0 \frac{\Omega_M}{\Omega_M - 1} \qquad H_0 t_m = \frac{\Omega_M}{2(\Omega_M - 1)^{3/2}}(\theta_m - \sin\theta_m) = \frac{\Omega_M \pi}{2(\Omega_M - 1)^{3/2}}$$

For $t = 2t_m$ ($\theta = 2\pi$) the scale factor collapses (*Big Crunch*). We obtain an expression for t_0 by putting $a(t_0) = a_0$:

$$1 - \cos\theta_0 \equiv \frac{2(\Omega_M - 1)}{\Omega_M} \qquad \Rightarrow \qquad \cos\theta_0 = \frac{2 - \Omega_M}{\Omega_M} = \left(\frac{2}{\Omega_M} - 1\right).$$

By transforming $\cos\theta_0 \to \sin\theta_0$ we easily obtain, from eq. 139,

$$H_0 t_0 = \frac{\Omega_M}{2(\Omega_M - 1)^{3/2}}\left[\arccos\left(\frac{2}{\Omega_M} - 1\right) - \frac{2}{\Omega_M}\sqrt{\Omega_M - 1}\right] < \frac{2}{3},$$

i.e. less than for the *EdS* model, in which $H_0 t_0 = \frac{2}{3}$.

- $\Omega_M < 1$: In this case it is useful to write $(\frac{a}{a_0}) \cdot \frac{2(1-\Omega_M)}{\Omega_M} \equiv \cosh\psi - 1$, and the parametric solution is

$$H_0 t = \frac{\Omega_M}{2(1 - \Omega_M)^{3/2}}(\sinh\psi - \psi) \qquad a(t) = a_0 \frac{\Omega_M}{2(1 - \Omega_M)}(\cosh\psi - 1) \tag{140}$$

In a way similar to what we did for $\Omega_M > 1$

$$\cosh\psi_0 = 1 + \frac{2(1 - \Omega_M)}{\Omega_M} = \frac{2}{\Omega_M} - 1 \quad \text{and} \quad Ht_0 = \frac{\Omega_M}{2(1 - \Omega_M)^{3/2}}\left[\frac{2}{\Omega_M}\sqrt{1 - \Omega_M} - \cosh^{-1}\left(\frac{2}{\Omega_M} - 1\right)\right] > \frac{2}{3}$$

By means of the relation $\cosh^{-1}(x) = \ln[x + \sqrt{x^2 - 1}]$ and expanding for $\Omega_M \to 0$ we get

$$H_0 t_0 \simeq 1 + \frac{\Omega_M \ln\Omega_M}{2} \to 1.$$

The asymptotic behaviour for the scale factor can be inferred from eq. 137 for $a \to \infty$

$$\left(\frac{\dot{a}}{a_0}\right)^2 \simeq H_0^2 (1 - \Omega_M) = const \implies \dot{a} = const \implies a(t) \propto t \tag{141}$$

and a grows linearly with time, while in *EdS* $a \propto t^{\frac{2}{3}}$.

Why do we study models that are clearly not representative of the currently accepted cosmological model? Regarding the model with $\Omega_M > 1$, it may represent the evolution (simplified) of a density fluctuation in excess of the average density: if we suppose that a spherical region of the universe has a density greater than the mean density and also greater than the critical one, the spherical region evolves (its radius evolves) as the scale factor of a Universe with $\Omega_{M,local} > 1$, reaching a maximum and then recollpsing, while the rest of the Universe continues to expand. The model with $\Omega_M < 1$ shows us that, as the density tends to zero due to the expansion, the universe approaches the *Milne* model, with $a \propto t$, $H = 1/t$, $k = -1$.

3.14 Models with $\Lambda \neq 0$

We have already seen a qualitative classification of these models. Let's now try to be more quantitative. We have also seen that the models that begin with a Big Bang recollapse or expand indefinitely. But in other cases the universe never had a Big Bang: the universe collapsed in the past, but the repulsive effect of a $\Lambda > 0$ slowed the collapse to turn it into expansion.

To see in detail the various cases it is necessary to integrate numerically Friedmann equations; if we neglect the contribution of radiation, one can also proceed analytically. The Friedmann equation ($F1$) with $\Omega_R = 0$ (see eq. 107)becomes

$$\left(\frac{\dot{a}}{a_0}\right)^2 = H_0^2\left[\Omega_M\left(\frac{a_0}{a}\right) + \Omega_\Lambda\left(\frac{a_0}{a}\right)^{-2} + 1 - \Omega_M - \Omega_\Lambda\right] \tag{142}$$

If we put $a/a_0 = R$ and $\tau = H_0 \cdot t$ we get:

$$\left(\frac{dR}{d\tau}\right)^2 = 1 + \Omega_M\left(\frac{1}{R} - 1\right) + \Omega_\Lambda(R^2 - 1) \tag{143}$$

The present epoch corresponds to $R = 1$ and the slope of the curve, at present, is equal to one.

Let's explore the future (and the past) of cosmological models as a function of the present values of the density parameters Ω_M and Ω_Λ.

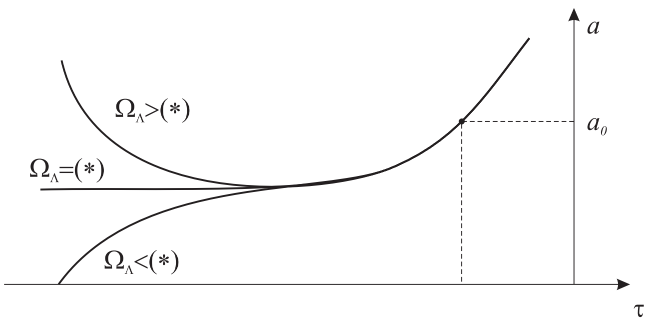
We have seen that, if $\Lambda < 0$ ($\Omega_\Lambda < 0$), the Universe finally recollapses in all cases. We have also seen that, if $k \leq 0$ and $\Lambda > 0$ ($\Omega_\Lambda > 0$), the Universe finally expands to infinity like in *de Sitter* model.

The case with $k = +1$ and $\Lambda > 0$ ($\Omega_\Lambda > 0$) is more complicate. If we look at the past, it is possible, at least in principle, that our Universe is a bouncing universe. Since it is now expanding, the bounce already happened and $a(t)$ had a minimum in the past. So we must find the conditions for $\mathrm{d}R/\mathrm{d}\tau = 0$. But eq. 142 leads in this case to a cubic equation for R, and cubic equations are not so easy to solve.

We will do in another way. As we can see in the following figure, the limiting case between a bouncing universe and a model with a Big Bang in the past corresponds to the case in which our universe is described by an *Eddington-Lemaitre* model, i.e. a deviation from an originally static *Einstein* model. In this case, in the past, both $\dot{a}$ and $\ddot{a}$ had to be equal to zero. So we use eq. 107 and eq. 116 with $\Omega_R = 0$:

$$\Omega_M(1+z) + \Omega_\Lambda(1+z)^{-2} + 1 - \Omega_M - \Omega_\Lambda \equiv 0 \tag{144}$$

$$\Omega_M(1+z)^0 + \Omega_\Lambda(-2)(1+z)^{-3} = \Omega_M - 2\Omega_\Lambda(1+z)^{-3} \equiv 0 \tag{145}$$



Eq. 145 gives immediately

$$\Omega_\Lambda = \frac{\Omega_M}{2}(1+z)^3 \tag{146}$$

This means that, given Ω_M, if $\ddot{a} = 0$ at redshift z, Ω_Λ is given by eq. 146. Remember that Ω_M and Ω_Λ refer always to the present epoch. We can work in a similar way with eq. 144; we multiply it by $(1+z)^2$ and, after some algebra, we finally get

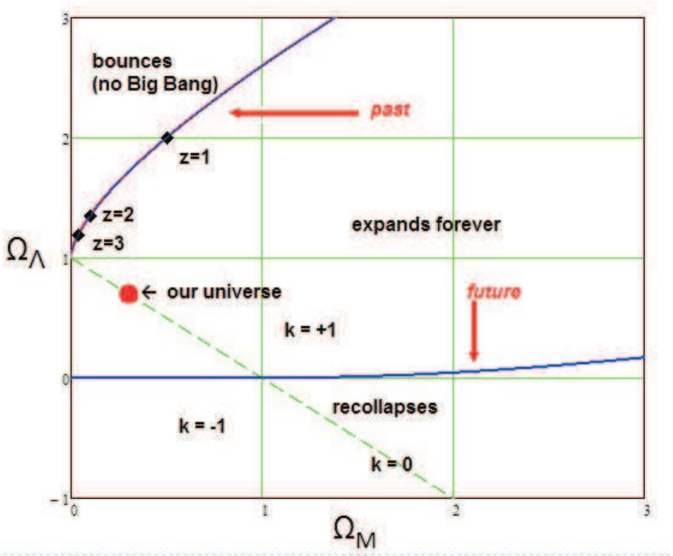
$$\Omega_\Lambda = \frac{(1+z)^2}{z\,(z+2)}\,(1 + \Omega_M z) \tag{147}$$

We want that both $\dot{a}$ and $\ddot{a}$ are zero, so eq. 146 and 147 have both to be fulfilled. If we solve both equations we get

$$\begin{cases} \Omega_M(z) = \frac{2}{z^2(3+z)} \\ \Omega_\Lambda(z) = \frac{(1+z)^3}{z^2(3+z)} \qquad (*) \end{cases} \tag{148}$$

If we want to explore the past, z must go from 0 to ∞. If we let z to move within this range we obtain the curve labelled *"past"* in the following Figure. Some values of z are shown. Note that in bouncing models the scale factor has a minimum value and, since $a_0/a = 1 + z$, this means that there is a maximum value for z. The minimum value of a and the maximum value for z correspond to those of the Eddington-Lemaitre model and are, also in this case, those shown in the following Figure on the *"past"* curve. We see that, in order to have a maximum redshift much larger than one, the value of Ω_M must be very low, in disagreement with with the observations. So bouncing models are excluded for our Universe.

But our Universe could tend to an *Eddington-Lemaitre* model not in the past but in the future. This would mark the border between a recollapsing universe and a universe expanding to infinity. We can again apply eqs. 148, but the trick is to let z to move now within the range between 0 and -1, since as a becomes larger than a_0 the relation $a_0/a = 1 + z$ shows that $z < 0$ and $z \to -1$ as $a \to \infty$.



The Figure above shows a synthesis of the discussion. The following Figure shows some models with the corresponding values of $(\Omega_M, \Omega_\Lambda)$.

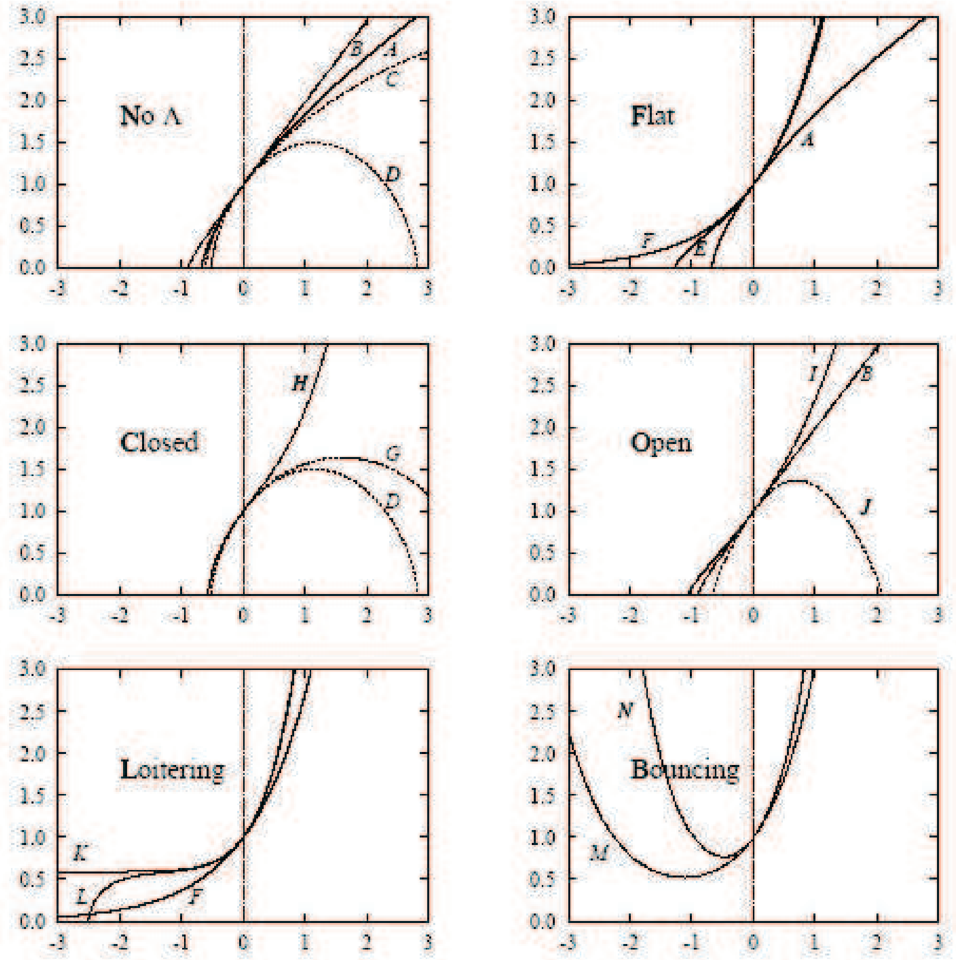


Fig. 1: Evolution of the scale parameter with respect to time for different values of matter density and cosmological parameter. The horizontal axis represents $\tau = H_0(t - t_0)$, while the vertical axis is $y = a/a_0$ in each case. The values of $(\Omega_M, \Omega_\Lambda)$ for different plots are: A=(1,0), B=(0.1,0), C=(1.5,0), D=(3,0), E=(0.1,0.9), F=(0,1), G=(3,.1), H=(3,1), I=(.1,.5), J=(.5,−1), K=(1.1,2.707), L=(1,2.59), M=(0.1,1.5), N=(0.1,2.5). From Ref. [7].

3.15 Our Universe?

We have seen that our Universe, after *matter/radiation equality*, is first *matter dominated* and then *vacuum dominated*. There exists an analytical solution for a Universe with matter and cosmological constant and spatially flat ($\Omega_M + \Omega_\Lambda = 1$).

Eq. 107 is, in this case,

$$\left(\frac{\dot{a}}{a_0}\right)^2 = H_0^2\left[\Omega_M\left(\frac{a_0}{a}\right) + \Omega_\Lambda\left(\frac{a_0}{a}\right)^{-2}\right], \tag{149}$$

and, by putting $R \equiv a/a_0$,

$$\frac{dR}{dt} = H_0\left[\frac{\Omega_M}{R} + \Omega_\Lambda R^2\right]^{1/2} \quad \rightarrow \quad H_0 dt = \frac{dR}{\sqrt{\Omega_M/R + \Omega_\Lambda R^2}}$$

$$H_0 t = \int_0^R \frac{dx}{\sqrt{\Omega_M/x + \Omega_\Lambda x^2}} = \int_0^R \frac{\sqrt{x}dx}{\sqrt{\Omega_M + \Omega_\Lambda x^3}}$$

We use the substitution $x^3 \equiv u^2$ and we get

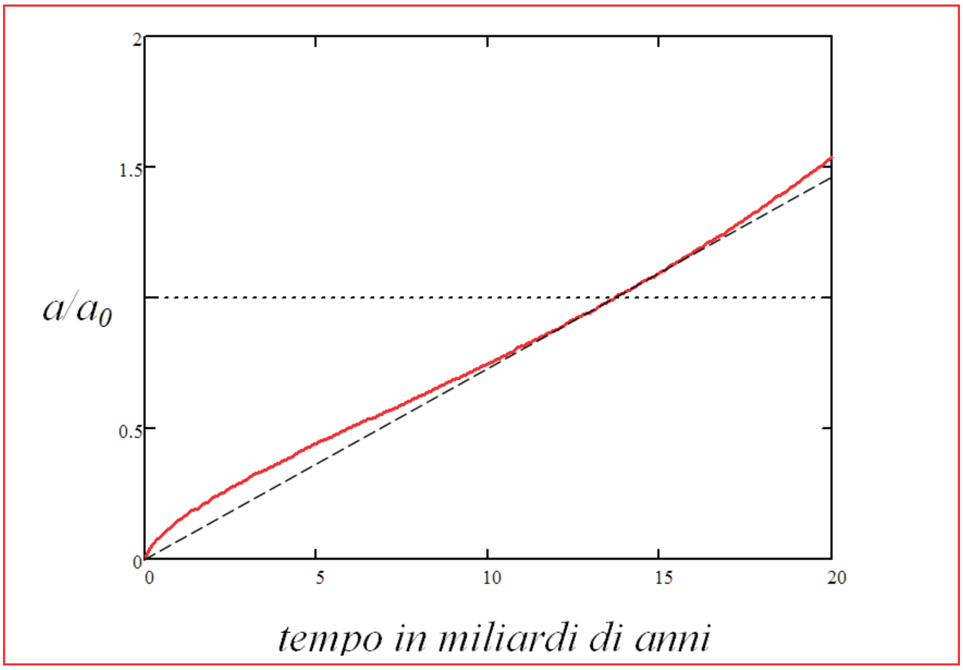
$$H_0 t = \frac{2}{3\sqrt{\Omega_\Lambda}} \int_0^{R^{3/2}} \frac{du}{\sqrt{\Omega_M/\Omega_\Lambda + u^2}}.$$

By solving the integral[18], we finally obtain

$$H_0 t = \frac{2}{3\sqrt{\Omega_\Lambda}} \sinh^{-1}\left[\left(\frac{a}{a_0}\right)^{3/2}\sqrt{\frac{\Omega_\Lambda}{\Omega_M}}\right] \tag{150}$$

$$a(t) = a_0\left(\frac{\Omega_M}{\Omega_\Lambda}\right)^{\frac{1}{3}}\left[\sinh\left(\frac{3\sqrt{\Omega_\Lambda}}{2}H_0 t\right)\right]^{\frac{2}{3}} \tag{151}$$

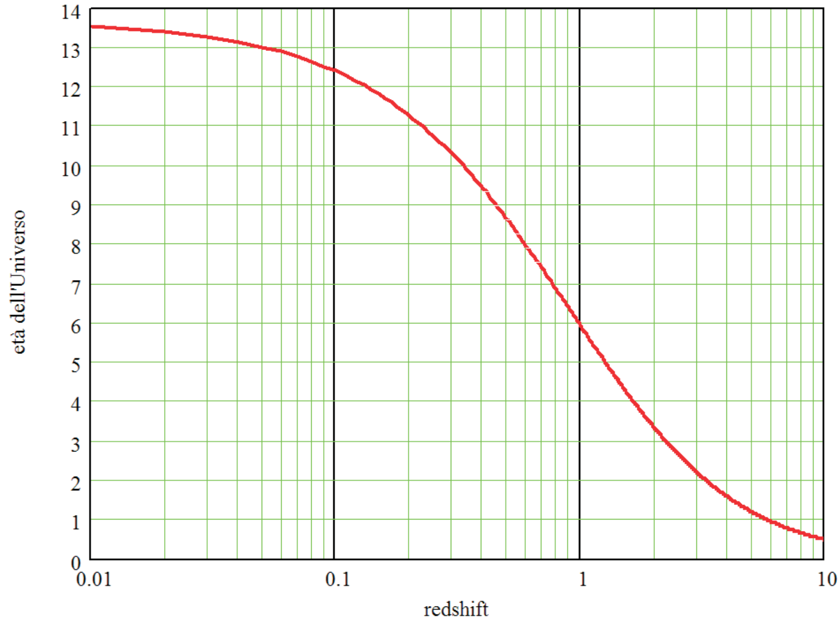
Since $a/a_0 = (1+z)^{-1}$, eq. 150 gives, for this model, the age of the Universe as a function of redshift. In the following Figure we draw $a(t)$ for $\Omega_M \sim 0.3$, $\Omega_\Lambda \sim 0.7$ and $h \sim 0.7$. Note that the dashed line, tangent to the curve at the present time, goes approximatly through the origin; $H_0 t_0 = 0.964$ and the present age of the Universe ($t_0 \simeq 13.5$ *Gyr*[19]) is almost equal to the Hubble time.



[18] By defining $q \equiv \sqrt{x^2+a^2}$, $\int \frac{dx}{\sqrt{x^2+a^2}} = \ln(x+q)$ e $\ln\left(\frac{x+q}{a}\right) = \sinh^{-1}\left(\frac{x}{a}\right)$

[19] One $Gyr = 10^9 year$

The following Figure (with the same cosmological parameters) shows the link between redshift and age of the Universe (in billion year).



3.16 The age of the Universe

We have so far seen the age of the Universe as a function of redshift for some particular models. Now let's see a general formulation. From the very definition of the Hubble parameter, we have:

$$H \equiv \frac{\dot{a}}{a} = \frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}t} \qquad \Longrightarrow \qquad \mathrm{d}t = \frac{\mathrm{d}a}{a \cdot H} = \frac{\mathrm{d}(a/a_0)}{(a/a_0) \cdot H(a)}. \tag{152}$$

If we write $a/a_0 \equiv u$,

$$t(a/a_0) = \int_0^{a/a_0} \frac{du}{u \cdot H(u)} \tag{153}$$

which can be integrated, at least numerically, by using the known dependence of H on a/a_0 and on the cosmological parameters (see eq. 108). It is more useful to use the redshift; remember that $a/a_0 = 1/(1+z)$ and so

$$\mathrm{d}a = -\frac{a_0}{(1+z)^2}\,\mathrm{d}z \qquad \Longrightarrow \qquad dt = -\frac{a_0}{(1+z)^2\,a\,H}\,\mathrm{d}z = -\frac{\mathrm{d}z}{(1+z)\;H(z)} \tag{154}$$

where we used again $a_0/a = 1+z$. Finally, we get the useful, general relation

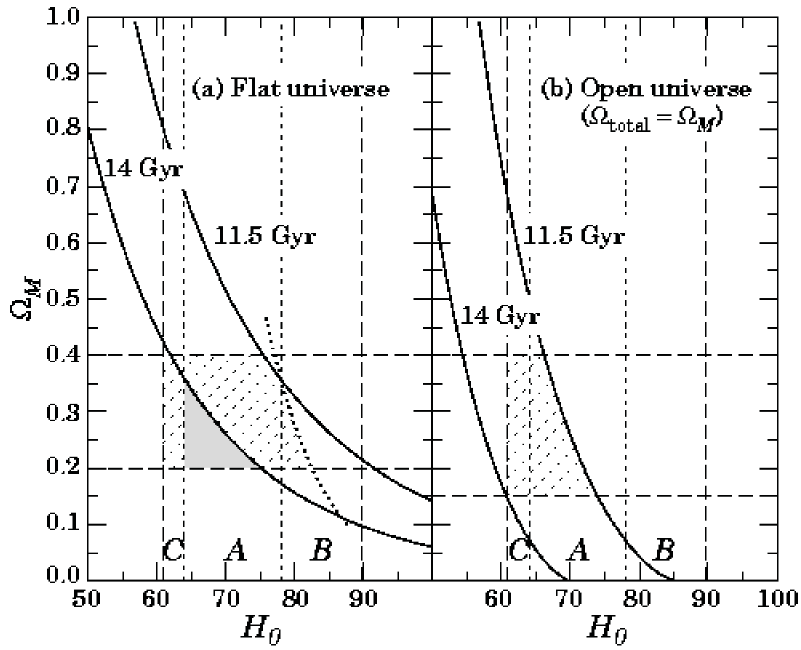
$$t(z) = \int_z^{\infty} \frac{\mathrm{d}z'}{(1+z')\,H(z')} \tag{155}$$

in which

$$H(z) = H_0(1+z)\left[\Omega_R(1+z)^2 + \Omega_M(1+z) + \Omega_\Lambda(1+z)^{-2} + 1 - \Omega_0\right]^{1/2}$$

where $\Omega_0 \equiv \Omega_R + \Omega_M + \Omega_\Lambda$.

The following Figure shows the effect produced by the cosmological constant on the age of the Universe. It compares, in the (H_0, Ω_M) plane, the "flat" case ($\Omega_M + \Omega_\Lambda = 1$) with the "open" case ($\Omega_0 = \Omega_M$). Note that the cosmological constant makes the age of the Universe longer, in better agreement with the estimated age of globular clusters.



A cosmological parameter often used is the so-called **look-back time**, $t_{lb} = t_0 - t(z)$, which is the time elapsed between redshift z and today. In particular, for *EdS* model:

$$t_{lb} = t_0 - t_0(1+z)^{-\frac{3(1+w)}{2}} = t_0\left[1 - (1+z)^{-\frac{3(1+w)}{2}}\right] \tag{156}$$

Example: Let's take $w = 0$, and assume $H_0 = 70km\,s^{-1}\,Mpc^{-1}$.

$$H_0 = \frac{1}{4.4 \cdot 10^{17}} s \qquad \Rightarrow \qquad t_0 = \frac{2}{3H_0} = \frac{2}{3} 4.4 \cdot 10^{17} s = 2.9 \cdot 10^{17} s = 9.3 \cdot 10^9 years$$

If we observe a quasar at $z = 2$, how many years ago was emitted the light we see today? Eq. 156 gives in this case

$$t_{lb} = t_0\left[1 - (1+z)^{-3/2}\right] = 0.8075\, t_0 = 7.5 \cdot 10^9 \text{ years ago.}$$

For the model described by eq. 150 $t_{lb}(z = 2) \simeq 10$ *Gyr*.

3.17 Horizons again

Let's estimate the size of the horizons for a particular,useful model, the *EdS* model.

$$d_H(t) = a(t)\int_0^t \frac{c\mathrm{d}t'}{a(t')} = a(t)\int_0^{t/t_0} t_0 \frac{c\mathrm{d}(\frac{t'}{t_0})}{a_0(\frac{t'}{t_0})^{\frac{2}{3(1+w)}}} = \frac{a(t)}{a_0} \cdot ct_0 \int_0^{t/t_0} \frac{\mathrm{d}x}{x^{\frac{2}{3(1+w)}}} =$$

$$= \frac{a(t)}{a_0} ct_0 \frac{3(1+w)}{1+3w}\left(\frac{t}{t_0}\right)^{\frac{1+3w}{3(1+w)}} = \frac{3(1+w)}{1+3w} ct$$

We see that $d_H(t) \propto ct$; if $w = 0$, $d_H(t) = 3ct$; if $w = 1/3$, $d_H(t) = 2ct$. If we calculate the event horizon for *EdS* model, we find that the integral involved diverges; this means that we don't have, in this model, an event horizon: if we are patient enough we will receive information from any future event.

Let's now estimate, for the same model, the value of the Hubble radius R_H, defined by eq. 76.

$$R(t) \equiv \frac{c}{H(t)} = \frac{c \cdot 3(1+w)t}{2} = \frac{1+3w}{2} d_H(t) \tag{157}$$

Although R_H and d_H appear comparable in the *EdS* model, they are actually two very different things: $R_H(t)$ is an instantaneous quantity, depending on the instantaneous value of H, while $d_H(t)$ is an integral quantity that depends on the entire past history of the Universe. For this reason, if an object has entered the particle horizon it will always remain inside, while it may happen that an object is, for example, first inside, then out, then back into the Hubble radius. If $H = const$ (as in *de Sitter* model), $R_H = const = c/H$.

4 Observational cosmology

4.1 Introduction

In the previous Section we have met many different theoretical cosmological models. Now it's time to see which of these models matches better the observational constraints. To do that we will need some useful tools and quantities. Then we will use them into the "classical" tests which have been used to derive estimates of the theoretical cosmological models.

4.2 $a_0\, r(z)$

As a first step we derive a quantity which is involved in almost every cosmological, observational test: $a_0\, r(z)$. Let's consider the radial motion of a photon travelling towar us (see eq. 70). We get

$$\frac{a\,\mathrm{d}r}{\sqrt{1-kr^2}} = -c\mathrm{d}t = -\frac{\mathrm{d}a}{\dot{a}} = -c\frac{\mathrm{d}a}{a\,H} \tag{158}$$

which can be rewritten (remember that $a = a_0/(1+z) \;\rightarrow\; \mathrm{d}a = -a_0/(1+z)^2\,\mathrm{d}z$) as

$$\frac{a_0}{1+z}\cdot\frac{dr}{\sqrt{1-kr^2}} = -\frac{c}{H}\cdot\frac{1+z}{a_0}\cdot\left[-\frac{a_0}{(1+z)^2}\right]\mathrm{d}z \tag{159}$$

and gives

$$\frac{a_0\,dr}{\sqrt{1-kr^2}} = \frac{c}{H(z)}\mathrm{d}z = \frac{c}{H_0\,E(z)}\mathrm{d}z \tag{160}$$

where we have defined the quantity $E(z) \equiv H(z)/H_0$. Going back to eq.110 we have ($\Omega_0 = \Omega_R + \Omega_M + \Omega_\Lambda$)

$$E(z) = \left[\Omega_R(1+z)^4 + \Omega_M(1+z)^3 + \Omega_\Lambda + (1-\Omega_0)(1+z)^2\right]^{1/2}. \tag{161}$$

If we use now eqs.71, 72 and 75 we have

$$f_k(r) = \int_0^r \frac{\mathrm{d}r'}{\sqrt{1-kr'^2}} = \frac{c}{a_0 H_0}\int_0^z \frac{\mathrm{d}z'}{E(z')} = \left\{\begin{array}{ll} \arcsin r & (k=+1) \\ r & (k=0) \\ \mathrm{arcsinh} r & (k=-1) \end{array}\right\}. \tag{162}$$

We use also eq.106:

$$\frac{kc^2}{a_0^2} = H_0^2\,[\Omega_0 - 1] \;\rightarrow\; \frac{c}{a_0 H_0} = \sqrt{|\Omega_0 - 1|} \;\rightarrow\; 1 = \frac{c}{a_0 H_0\sqrt{|\Omega_0-1|}}. \tag{163}$$

Finally, for $k=+1$,

$$\frac{c}{a_0 H_0}\int_0^z \frac{\mathrm{d}z'}{E(z')} = \arcsin r \;\rightarrow\; r = \sin\left(\frac{c}{a_0 H_0}\int_0^z \frac{\mathrm{d}z'}{E(z')}\right) = \frac{c}{a_0 H_0\sqrt{|\Omega_0-1|}}\cdot\sin\left(\sqrt{|\Omega_0-1|}\int_0^z\frac{\mathrm{d}z'}{E(z')}\right) \tag{164}$$

and:

$$a_0\,r(z) = \frac{c}{H_0\sqrt{|\Omega_0-1|}}\cdot\sin\left(\sqrt{|\Omega_0-1|}\int_0^z\frac{\mathrm{d}z'}{E(z')}\right) \qquad (k=+1). \tag{165}$$

For $k=-1$ we get a similar result, with $\sin\;\rightarrow\;\sinh$:

$$a_0\,r(z) = \frac{c}{H_0\sqrt{|\Omega_0-1|}}\cdot\sinh\left(\sqrt{|\Omega_0-1|}\int_0^z\frac{\mathrm{d}z'}{E(z')}\right) \qquad (k=-1). \tag{166}$$

For $k=0$ we have simply:

$$a_0\,r(z) = \frac{c}{H_0}\int_0^z\frac{\mathrm{d}z'}{E(z')} \qquad (k=0). \tag{167}$$

The function$E(z)$, eq. 161, can be approximated in different ways: for $z < z_{eq}$, and $\Omega_R = 0$,

$$E(z) = \Big[(1+z)^2(1+\Omega_M z) - z\,(2+z)\Omega_\Lambda\Big]^{1/2}; \tag{168}$$

when Ω_Λ is also negligible this gives

$$E(z) = (1+z)\,\sqrt{1+\Omega_M z}; \tag{169}$$

finally, when $\Omega_0 = 1$, i.e. $\Omega_M + \Omega_\Lambda = 1$, $\Omega_R = 0$,

$$E(z) = \Big[1 - \Omega_M + \Omega_M(1+z)^3\Big]^{1/2}. \tag{170}$$

There are no general. analytical expressions for $a_0 r(z)$; in the case $\Omega_\Lambda = 0$ the following expression can be obtained (the so called *Mattig formula*):

$$a_0 r(z) = \frac{2c}{H_0}\,\frac{\Omega_M z + (\Omega_M - 2)[(\Omega_M z + 1)^{1/2} - 1]}{\Omega_M^2(1+z)}, \tag{171}$$

which holds for both $\Omega_M > 1$ and $\Omega_M < 1$.

For $z \gg 1$ $(z \to \infty)$ Mattig formula gives

$$a_0 r(z) = \frac{2c}{H_0 \Omega_M} \qquad (\Omega_\Lambda = 0), \tag{172}$$

while, when $\Omega_M + \Omega_\Lambda = 1$, a useful *approximation* is

$$a_0 r(z) = \frac{2c}{H_0 \Omega_M^{0.4}} \qquad (\Omega_M + \Omega_\Lambda = 1). \tag{173}$$