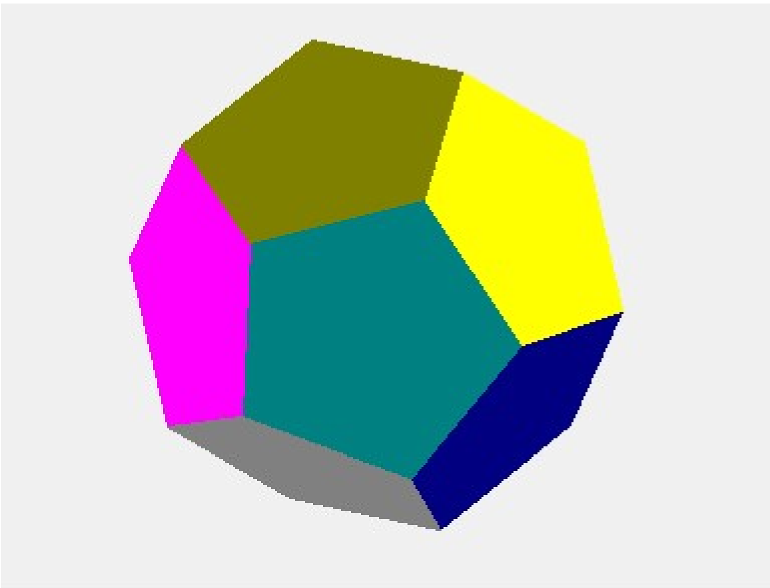
General Relativity and Cosmology

A personal outline

This is an article from my home page: www.olewitthansen.dk



Ole Witt-Hansen 2016

contents

Acknowledgement

I have, among other things been teaching in the Danish high school (gymnasium) for 40 years.
I was teaching the students Physics, Mathematics and Computer Science. Until 1988, and from 1988 -2005, (in an amputated form), this 3 year education had a highly professional level. It was together with the German and French among the best in the world.
But regarding the theoretical level, it deteriorated completely after 2005.
Instead of teaching elementary concepts, it became a broad IT and picture supported informative education of the achievements of natural science.
There was however, one new invention in the third year. This was "the large subject", where the students, were supposed to go into more depths. In 2011 the subjects in Physics was cosmology.
Having taken a minor university course in Astronomy in 1965, I knew nothing of of Cosmology, except from popular (non mathematical) books e.g. Kip Thorne: Kip S. Thorne: Black Holes and time warps, and I had never taken a course in General Relativity, since I wrote my thesis in particle physics.
The physics department published a book on cosmology, which however was totally inadequate and incomprehensible, even for the teacher, so I decided to make my own educational material, as many times before.
I had two books on my shelf: C. Møller: The theory of Relativity (1951) , and Landau and Lifschits: The classical Theory of Fields. (1975), but neither were really any help at that stage.
Then I ordered (most by chance) from Amazon: Ta-Pei-Cheng: Relativity, Gravitation and Cosmology, and that was a challenge I daresay. I had passed a university course in Gaussian differential geometry in 1967, but not in generalized Riemannian geometry.
However, I succeeded to write some quite applicable notes to the students in 2011, but I never got rid of an urge to understand more deeply the General theory of Relativity.
I pensioned myself in November 2012 because I felt, for academic reasons that I could no longer substantiate the professional collapse in the Danish high school.
The next couple of years I spent building my Website, with a comprehensive production of books and article on Physics, Mathematics and (obsolete) Computer science.
In 2015 I took up again General Relativity. I decided to write in English, since the Danish audience for these subjects is limited.
First I wrote 30 pages of notes about Gaussian Differential Geometry. Then I translated from Danish to English a chapter on The Special Theory of Relativity from a textbook, I had written in 1978.
Then I added a paper on Tensor Analysis, as an inlet to the present exposition: General Relativity and Cosmology.
The mathematical level of tensor analysis and General Relativity is certainly demanding, but I hope that my readers will share my own enlargement in understanding Einstein's magnificent theory.

Ole Witt-Hansen
June 2016

1. The foundation of The General Theory of Relativity

Newton's law of gravitation was until the beginning of the twentieth century considered being the ultimate physical description of the phenomena of gravitation. The obvious reason was that his law of gravitation had been able to explain fully the gravitation on the earth, and by Newton's revolutionary intellect, it was the same simple law that determined the orbital movement of the moon, as well as the planets orbital motion around the sun.

The appearance of Einstein's Special theory of Relativity in 1905 did not significantly influence the belief of Newton's theory of gravitation, and in fact it still hasn't, as long as one is confined to non relativistic motion, although neither Newton's three laws nor his theory of gravitation did comply with Special Relativity. Newton's laws of dynamics were rather easy to modify to comply with the Lorentz transformation, but that was not the case with his law of gravitation.

Also the velocities of the planets are of order most 3,0 10^4 m/s, so that v/c is of order 10^{-4}, and $(v/c)^2$ is of order 10^{-8} , and for this reason, it is highly unlikely that anyone had ever observed any deviation from Newton's law of gravitation, due to Special Relativity. With one exception, which concerns the precession of the perihelion of Mercury, being a discrepancy of 43'' per century of the observed and calculated precession – even when taking into accounts the influence of the other nearest planets.

There was however one more thing. Some observations indicated that a star could be observed in two separate positions. This phenomena could, however be explained if the light rays from the star were bended by a massive object situated in front of the star in the line of sight.
But the phenomena can actually be explained in the framework of Newtonian mechanics, although the predicted bending is only half of what is derived from Einstein's General theory of Relativity.

1.1 The equivalence principle

The equivalence principle bluntly formulated is that all frames of reference, including accelerated frames are inertial frames, in which the laws of physics are valid.

The accelerated frame must however be such that all entities have the same acceleration, which mean that the acceleration is due to gravity.

The equivalence principle has due to the formulation of Einstein two parts:

The *weak* equivalence principle claims the equality between the heavy mass (that is the mass m_G that enters in Newton's law of gravitation), and the inertial mass m_I (that enters Newton's second law of dynamics).

(1.1) $$F = G\frac{m_G M}{r^2} \qquad \text{and} \qquad F = m_I a$$

The *strong* equivalence principle states that motion due to gravitation is equivalent to acceleration in a "free fall" in a curved space time. An observer inside a box e.g. in a spaceship, will feel no gravity. So gravity is transformed away by acceleration. Inside the box the laws of physics will be the same as in a Galileo inertial frame.

In the Special Theory of Relativity the laws of physics are the same in every **inertial rest frame.** The coordinates and time measurements for two inertial frames that move relative to each other with velocity v, are given by the Lorentz transformations:

(1.2) $$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}} \quad \text{and} \quad t' = \frac{t - \frac{xv}{c^2}}{\sqrt{1 - v^2/c^2}}$$

On the other hand, if you are placed in a accelerated coordinate system, e.g. a train or an airplane which enters a curve, Newton's laws of dynamics are no longer valid, since the passengers will follow the law of inertia, and therefore will feel a push in the opposite direction of the turn, even if the are no visible force.
Such forces are called fictitious forces. They have no tangible origin, but they are due to the fact that the system is accelerated, because of the external forces, while you follow the law of inertia.

If you move in a rotating system e.g. sitting on a carousel, you will experience two fictitious forces, the centrifugal force, which stems from the fact, that it requires a centripetal force to uphold a circular motion, and the Coriolis force, which is more subtle, but comes about, because the velocity in a rotating system increases linearly with the distance from the centre. $v = \omega r$, whereas ω the angular velocity is the same.

So obviously not all accelerated coordinate system are inertial systems.
Newton's laws of dynamics are not in general valid in an accelerated system. The law of inertia is not valid in an accelerated system, that is, unless all entities in the frame have exactly the same acceleration, which is the case of a gravitational "free fall". But the crucial condition is that the heavy mass is exactly identical to the inertial mass. Otherwise Newton's second law and his law of gravitation would not "treat" the mass on equal footing.

In 1915 the principle of equivalence was a more abstract supposition made by Einstein, but since the sixties with launching of manned spaceships, it has been irrevocably confirmed that the spaceship capsule is in fact an inertial system.
The men in the capsule feel neither gravity nor acceleration. Gravity is transformed away by acceleration. In fact gravity is equivalent to acceleration, and by this token the principle of equivalence is formulated.

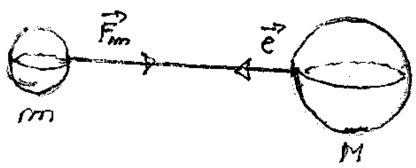
The principle of equivalence implies the existence of local inertial frames at every space-time point. In a sufficiently small region a local observer will sense no effects of gravity.
These considerations among others motivated Einstein to propose a geometrical curved space-time description of the gravitational field.

1.2 Review of the Newtonian theory of gravitation

It is one of the most fundamental principles in theoretical physics that a new theory of the physical world must comply with an established theory in some limit.
For example The Special Theory of Relativity becomes classical Newtonian mechanics when the velocities involved are much smaller than the speed of light, that is, when v/c goes to zero.
Likewise in quantum physics one regains the formulas of classical mechanics, in the limit where the Planck constant h goes to zero.

For the same reason, an extension of a theory of gravitation must comply with the Newtonian theory of gravitation in the area, where the theory has successfully been applied for more than two hundred years.
In fact we expect some formal resemblance between Einstein's General theory of Relativity and the Newtonian theory of gravitation.
We shall therefore proceed giving a recapitulation of Newton's theory in a mathematical compact form.

Newton's law of gravitation is illustrated below. F_m is the force on the mass m from the gravitational attraction from the mass M. The distance between the centres of the two spherical bodies is r. $\vec{e}$ is a unit vector pointing from M to m. G = 6,67 10^{-11} Nm^2/kg^2 is the gravitational constant. Newton's law of gravitation, may then be written:



(1.1) $$\vec{F}_m = -G\frac{mM}{r^2}\vec{e}$$

From Newton's second law we find the acceleration:

$$\vec{g}(r) = \frac{\vec{F}_m}{m}$$

(1.2) $$\vec{g}(r) = -\frac{GM}{r^2}\vec{e}$$

We say that the mass M gives rise to a gravitational field, as it is the case in electrostatics with Coulomb's law and we can likewise apply Gauss' law on the field:

If we take the surface integral over a closed surface it equals $4\pi G$ times the source of the field inside the surface. Let $d\vec{A}$ be an outward normal vector to the surface at the infinitesimal area dA.

(1.3) $$\int_S \vec{g}\cdot d\vec{A} = -4\pi GM$$

For at sphere with a point source in the centre this becomes obvious.
Foe symmetry reasons a field from a point source (or just a spherical symmetric source) must have a radial direction, and must have the same numerical value at a constant radius.

$$\int_{Sphere} \vec{g}(r)\cdot d\vec{A} = -\frac{GM}{r^2}\int_{Sphere} \vec{e}\cdot d\vec{A} = -\frac{GM}{r^2}\int_{Sphere} dA = -\frac{GM}{r^2}4\pi r^2 = -4\pi GM$$

Conversely, Stokes law can be used to derive the formula (1.2) for a point source.

By applying Stokes law, we can transform the surface integral into an integral over the volume using the divergence of the field, and at the same time turn the right hand side of (1.3) into a volume integral, where ρ is the mass density.

(1.4) $$\int_S \vec{g}\cdot d\vec{A} = \int_V \vec{\nabla}\cdot\vec{g}\cdot dV = -4\pi G\int_V \rho dV$$

Since the equation must hold for all volume elements, it must also hold for the differential volume elements, so we must have.

(1.5) $$\vec{\nabla} \cdot \vec{g} = -4\pi G\rho$$

This is Newton's field equation written in differential form. Mathematically it is identical to Gauss' law for electrostatics. Since the gravitational field is free from rotation:

$$\oint_c \vec{\nabla} \times \vec{g} \cdot d\vec{s} = 0 \ ,$$

the gravitational field can be expressed as the gradient of a potential.

(1.6) $$\vec{g} = -\vec{\nabla}\phi$$

From Newton's second law $\vec{F} = m\vec{a} = m\ddot{\vec{r}}$ we have $\vec{a} = \vec{g} = -\vec{\nabla}\phi$, and the Newtonian equation of motion becomes.

(1.7) $$\ddot{\vec{r}} = -\vec{\nabla}\phi$$

One may notice, however that Newton's theory of gravitation is not compatible with The Special Theory of Relativity. One reason is that the gravitational force is "acting at a distance" (propagates with infinite speed), and the other main reason is that space and time coordinates are not treated on equal footing, as they must be in Special Relativity.

The Special Theory of Relativity was published in 1905, and soon after completing it, Einstein began working on a relativistic theory of gravitation. In the period 1907 – 1914 he relied heavily on the equivalence principle to extract some general results. (Much in the same manner as Bohr used the correspondence principle to extract result from quantum mechanics).
Not until 1915 he realized that gravitation could be described within the framework of Riemannian differential geometry, being a generalisation from the Gaussian theory from ordinary 3-dimensional space to 4-dimensional space-time.

The heart of the General theory of Relativity is to describe accelerated motion – not as a consequence of gravitation – but following the shortest path (in time) in a warped space time.

On a surface in Euclidian space the **geodesics** are the "straight lines on a warped surface".
For example the geodesics on a sphere are great circles. Geodesics are quite fundamental in differential geometry, and it is closely related to the concept of **covariant differentiation**, since a geodesic is a curve along a path, where the covariant derivative is zero.
"Covariant" refers to objects that transform as a tensor under covariant coordinate transformations. (For details see Differential Geometry 2).

The concept of a geodesic follows one of the deepest principle in physics: **The principle of least time**, which also is the basis of the Lagrange approach of analytical mechanics.
It states that a particle (including light), will always follow the path that takes the minimum time.
For example the law of refraction can rather easily be deduced from the principle of least time.

The **curvature** of the 4-dimensional Minkowski space-time originate from massive bodies, but far away from such masses, we have a flat space-time that can be described by The Special Theory of Relativity.

We will show that the General theory of Relativity can in fact make predictions, which can be verified experimentally, predictions which differ from that of Newtonian mechanics.
Also General Relativity complies with the classical theory of gravitation in the non relativistic limit.

1.3 Implications of the principle of equivalence

We start out with Einstein's precision of the equivalence principle, a formulation which he used, when making deductions.

(1.8)	Physics in a frame, freely falling in a gravitational field, is equivalent to physics in an inertial frame without gravity. Physics in a non accelerating frame with gravity ***g*,** is equivalent to Physics in a frame without gravity, but accelerating with ***a* = -*g***

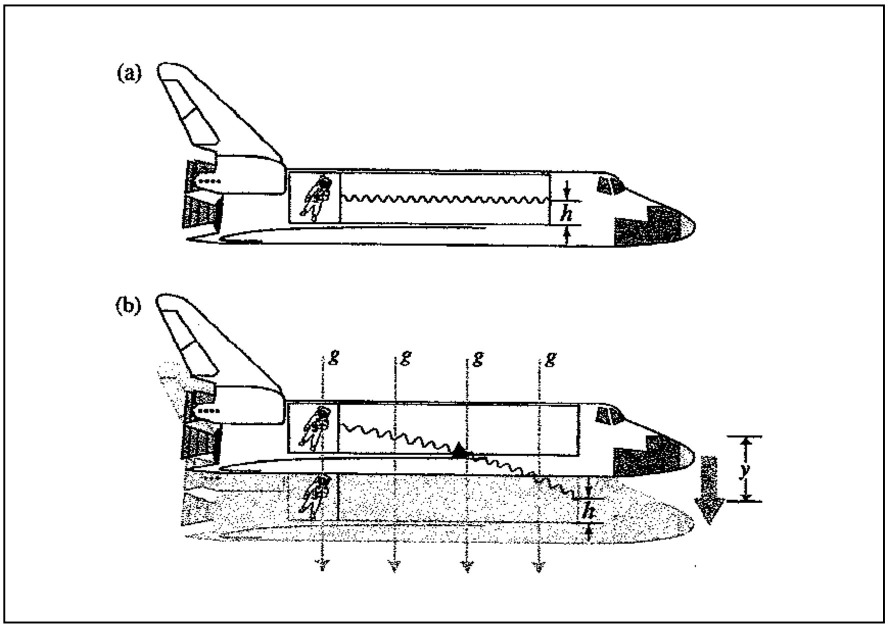
Thus according to the equivalence principle, accelerating frames of reference can be treated in exactly the same way as inertial frames. They are simply frames with gravity.
From this we also obtain a definition of an inertial frame (without reference to any external environment such as fixed stars), as a frame where there is no gravity.

In this way we also avoid the problem of incorporating Special Relativity into gravity but rather to use gravitation as a means to broaden the principle of relativity from inertial frames to all coordinate systems including accelerating frames.

We shall now study some consequence of applying the equivalence principle to various situations.
We shall subsequently show that gravity can bend a light ray, shift the frequency of an electromagnetic wave and cause clocks to run slower.

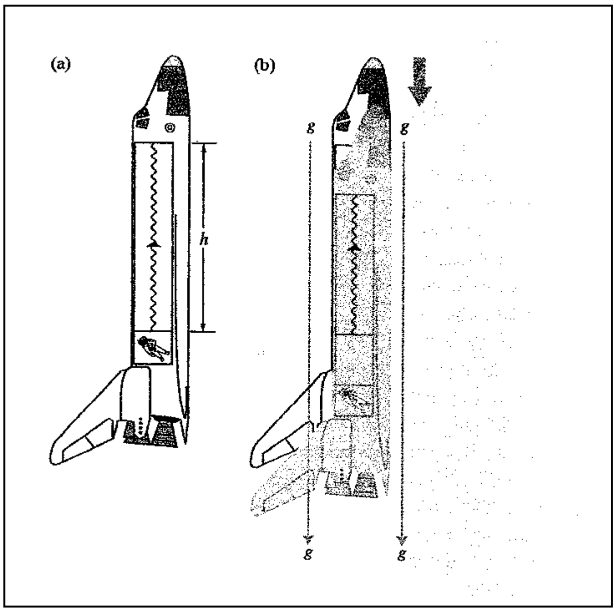
1.4 A spaceship in free fall

We first consider the situation of an observer inside a spaceship in a free fall. According to the equivalence principle there is no gravitational effects in this inertial frame and Special Relativity applies.
We shall then consider the events from the point of view of an observer watching the spaceship from outside. He sees a gravitational field, and the freely falling observer in the spaceship is seen to be accelerating in this gravitational field.
According to the equivalence principle, the combined effect of acceleration and gravity as seen from the observer outside the spaceship must reproduce the Special Relativity description as observed by the observer in free fall.
Physics should be independent of the choice of coordinate system.



First we look at a light ray travelling horizontally from one end of a spaceship to the other. The spaceship is falling vertically in a constant gravitational field ***g***. From the viewpoint of an observer in the spaceship (and according to the equivalence principle), he finds himself in an inertial frame so there are no effects associated with gravity or acceleration. The light moves straight across the spaceship from one side to the other.
Seen from the observer in the spaceship the light is received in the same height as it is emitted.

However, to an observer outside the spaceship there is both gravity ***g*** and acceleration.
From his point of view the light ray will follow the spaceship, and it will therefore appear to bend in the gravitational field, (since we know that the speed of light is independent of the speed of the emitter). (See figure). Since two observes agree that the signal was emitted and received in the same height, the equivalence principle suggest that a light ray is bended in the presence of a gravitational field. This is of course a thought experiment. With realistic scales the light will only be bended a fraction of the size of an atom.



We shall then consider a situation, where the light ray is emitted along the direction of the free fall.
From the equivalence principle, the observer in the spaceship is in an inertial frame and he cannot detect any effects, neither from gravitation nor from acceleration.
Specifically the emitted and received frequency of the light will be the same.
From the observer outside the spaceship, there will be a Doppler shift corresponding to a receiver moving away from the source.
If the receiver moves with a velocity *u*, then the observed wavelength λ' will be prolonged by a distance *uT*,' where *T'* is the

period of the signal, as measured from the receiver.
If λ is the wavelength of the emitted signal, we therefore have the relation: λ' = λ+ *uT'* .

Using the relation: λ = *cT*, we find: *cT' = cT + uT'* or (*c-u*)*T' = cT*, and inserting $T = 2\pi/\omega$, then dividing by *c* we get a formula for the (non relativistic) Doppler shift:
(It is not a normal Doppler shift, however since the emitter and the receiver they are not in relative motion)

(1.9) $$\omega' = \omega(1-\frac{u}{c}) \quad \Leftrightarrow \quad \frac{\omega-\omega'}{\omega}=\frac{u}{c} \quad \Leftrightarrow \quad \frac{\Delta\omega}{\omega}=\frac{u}{c}$$

What we see is that the received frequency is less than the emitted frequency, which means that the light is shifted towards the blue.
But according to the equivalence principle the two observes must agree that the frequency has not shifted.
The only possible explanation (apart from discarding the equivalence principle) is that the shift in frequency must be cancelled by gravity. We therefore postulate a shift due to gravity:

(1.10) $$\left(\frac{\Delta\omega}{\omega}\right)_g = -\frac{u}{c}$$

We shall now analyze this a bit more closely. We express the right hand side of (1.10) in terms of the gravitational potential ϕ. In the time Δ*t* (as seen from the outside), it takes the light ray to travel the distance *h*, it has had an increase in velocity: Δ*u* = *g*Δ*t* = *gh/c*. where.

$$g = \frac{GM}{r^2} \quad \text{And where the gravitational potential is} \quad \Phi = -\frac{GM}{r}$$

For the difference in the two positions separated by *h* we find:

$$\Delta\Phi \approx \frac{\partial\Phi}{\partial r}h = \frac{GM}{r^2}h = gh \quad \Rightarrow \quad \Delta u = \frac{gh}{c} = \frac{GM}{cr^2}h \quad \Rightarrow \quad \left(\frac{\Delta\omega}{\omega}\right)_g = -\frac{\Delta u}{c} = -\frac{\Delta\Phi}{c^2}$$

We thus find:

(1.11) $$\left(\frac{\Delta\omega}{\omega}\right)_g = -\frac{\Delta\Phi}{c^2} = -\frac{\Phi_2-\Phi_1}{c^2}$$

This means that a light ray emitted at a lower gravitational potential will be received at a higher gravitational potential at a lower frequency, that is, the signal is red shifted. This is in accordance with the outside observer, who would expect a blue shift, associated with the movement of the spaceship. The two shifts in frequency however exactly cancel each other to uphold the equivalence principle.
This gravitational shift in frequency is however impossible to observe directly, because it is of order 10^{-6} for the sun, and is masked by the Doppler shift due to the thermal motion of the atoms.

It is therefore surprising that the first conclusive data came from a terrestrial experiment performed by Pound, Rebka and Sneider in 1960 and 1964.

They succeeded in measuring a extremely small shift in radiation travelling up 22.5 *m*, which was the height of an elevator shaft in the building housing the Harvard physics department.
With these dimensions the relative shift in frequency is: $(\Delta\omega/\omega)_g = gh/c^2 = O(10^{-15})$
Normally it would be impossible to fix the frequency of an emitter or an absorber to this accuracy due to the thermal recoils of the atoms, but with the Mössbauer effect for crystals, the recoil is not from a single atom, but from the crystal as a whole, so with the bandwidth only limited by the Heisenberg uncertainty principle, it was possible to obtain this high accuracy.

Pound and collaborators worked with the excited atom Fe^*-57, which can be obtained through the nuclear beta-decay of Cobalt–57. It makes the transitions to the ground state by emitting a gamma ray.
Because of the extreme narrow line width, gamma rays emitted at the bottom of the elevator shaft after climbing up 22.5 *m* up could no longer be resonantly absorbed in sheets of Fe-57 in its ground state. Pound and Rebka then moved the elevator slowly towards the bottom, so that the Doppler blue shift had just the right amount to compensate for the gravitational red shift.
A calculation shows that to obtain this, the elevator should move with the speed 7.35 10^{-7} m/s.

$$\left(\frac{\Delta\omega}{\omega}\right)_g = \frac{gh}{c^2} = \frac{9.8\cdot 22.5}{3\cdot 10^8} = 7.35\,10^{-7}\,m$$

1.4 Gravitational time dilation

When you think about it, it becomes incomprehensible that an observer in a gravitational field stationary with respect to the emitter should receive radiation at a different rate, than it is emitted. And it appears certainly strange, but as Einstein was the first to explain the reason is not a change in frequency, but a **gravitational time dilation** $d\tau$, where $\omega = 2\pi/d\tau$ and $d\tau$ is the proper time rate of the period of the wave crest. From the gravitational frequency shift formula (1.11):

$$\frac{\Delta\omega}{\omega} = \frac{\omega_2 - \omega_1}{\omega_1} = -\frac{\Phi_2 - \Phi_1}{c^2}$$

we then get, when inserting $\omega = 2\pi/d\tau$.

$$\frac{\omega_2 - \omega_1}{\omega_1} = \frac{\dfrac{2\pi}{d\tau_2} - \dfrac{2\pi}{d\tau_1}}{\dfrac{2\pi}{d\tau_1}} = \frac{\dfrac{d\tau_1 - d\tau_2}{d\tau_1 d\tau_2}}{\dfrac{1}{d\tau_1}} = \frac{d\tau_1 - d\tau_2}{d\tau_2},$$

so we find

(1.12) $$\frac{d\tau_1 - d\tau_2}{d\tau_2} = \frac{\Phi_1 - \Phi_2}{c^2} \quad\Leftrightarrow\quad d\tau_1 = \left(1 + \frac{\Phi_1 - \Phi_2}{c^2}\right) d\tau_2$$

For a static gravitational field, this can be integrated to

(1.13) $$\tau_1 = \left(1 + \frac{\Phi_1 - \Phi_2}{c^2}\right)\tau_2$$

Namely, a clock situated at a higher gravitational potential will run faster. This is in contrast to Special Relativity, where two observers are moving relative to each other. In that case both observers will claim, that the counterparts clock run slower, while with gravitational dilation, the observer at a higher gravitational point claims that the clock at a lower gravitational potential run slower, and the observer at a lower gravitational potential will claim that the clocks at a higher gravitational potential run faster.

As an example we shall calculate the difference in clocks rates on the surface of the sun and on the earth. We must then calculate:

$$\frac{\tau_{earth} - \tau_{sun}}{\tau_{earth}} = \frac{\Phi_{earth} - \Phi_{sun}}{c^2} \qquad \text{Where} \quad \Phi = -G\frac{M_{sun}}{r}$$

Inserting M_{sun}=2.0 10^{30} *kg*, r_{sun} = 6.96 10^{8} *m* , r_{earth} = 1,50 10^{11} m and G = 6.67 10^{-11} *J m/kg*.

$$\frac{\tau_{earth} - \tau_{sun}}{\tau_{earth}} = \frac{\Phi_{earth} - \Phi_{sun}}{c^2} = 2.1\ 10^{-6}$$

This result together with other predictions from General Relativity, shows that Einstein's theory (albeit its vast epistemological consequences), will probably never have influence on human life on earth in a way that mechanics, electrodynamics, quantum physics or even Special Relativity has had.

Although the gravitational time dilation is extremely small, it has been possible to verify it in terrestrial circumstances using two identical Caesium atomic clocks, one in an airplane at an altitude of about 10 *km* and holding it there for a long time. After correcting for Special Relativity, the high altitude clock was found to gain over the grounds clock by an amount corresponding to $\Delta t = (gh/c^2)t$, where h was the altitude of the plane and g is the gravitational acceleration and gh is the gravitational potential difference.

The gravitational time dilation should preferably be in accordance with Special Relativity.
To establish that this is in fact the case, we shall analyze the free fall of a spaceship within the framework of Special Relativity. The spaceship is according to the equivalence principle an inertial system without gravity, and Special Relativity should apply.

We consider the free fall at two positions, position (1), where the speed is v_1 and position (2) where the speed is v_2.
Let t denote the time as recorded in the free fall frame and τ the time recorded in the two fixed positions. According to Special theory

$$\Delta t = \frac{\Delta t' + \frac{\Delta x' v}{c^2}}{\sqrt{1 - v^2/c^2}}$$

For the observer in the spaceship $\Delta x'$ = 0, so we can write to special theory relations.

$$\Delta t_1 = \frac{\Delta\tau_1}{\sqrt{1 - {v_1}^2 / c^2}} \qquad and \qquad \Delta t_2 = \frac{\Delta\tau_2}{\sqrt{1 - {v_2}^2 / c^2}}$$

We are interested in finding the clock rates $\Delta\tau_1$ and $\Delta\tau_2$ when $\Delta t_1 = \Delta t_2$.
By division of the two equations above we get:

$$\frac{\Delta\tau_1}{\Delta\tau_2} = \frac{\sqrt{1 - {v_1}^2 / c^2}}{\sqrt{1 - {v_2}^2 / c^2}}$$

And by first applying the formula: $\frac{1}{1+h} \approx 1 - h$ to the denominator, and afterwards using the formula: $\sqrt{1+h} \approx 1 + \frac{1}{2}h$: we can simplify the expression above.

$$\frac{\Delta\tau_1}{\Delta\tau_2} = \frac{\sqrt{1 - {v_1}^2 / c^2}}{\sqrt{1 - {v_2}^2 / c^2}} \approx \sqrt{(1 - {v_1}^2 / c^2)(1 + {v_2}^2 / c^2)} \approx \sqrt{(1 - {v_1}^2 / c^2)(1 + {v_2}^2 / c^2)}$$

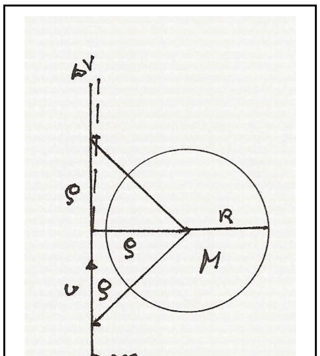
$$\frac{\Delta\tau_1}{\Delta\tau_2} =\approx \sqrt{1 + {v_2}^2 / c^2 - {v_1}^2 / c^2 - {v_1}^2 / c^2 \cdot {v_2}^2 / c^2)} \approx \sqrt{1 + {v_2}^2 / c^2 - {v_1}^2 / c^2} \approx 1 + \frac{1}{2}\frac{{v_2}^2 - {v_1}^2}{c^2}$$

$$\tfrac{1}{2}{v_2}^2 - \tfrac{1}{2}{v_1}^2 = \frac{\Delta E_{kin}}{m} = -\frac{\Delta\Phi}{m},$$

so again we find:

$$\frac{\Delta\tau_1}{\Delta\tau_2} = 1 + \frac{\Phi_1 - \Phi_2}{c^2} \text{ as in (1.12).}$$

1.5 Bending of a light ray. A classical approach



We shall initiate to make a calculation of the deflection, when a particle with mass *m* and velocity *v* passes a much lager body with mass *M* (the sun) at the distance ρ between the centres of the two bodies.
ρ is called the parameter of impact.
The calculation of the deflecting angle is identical with Rutherford's calculation of the deflection of the Alfa-particles when passing by an Au-nuclei.
An exact calculation is mildly complicated, but one can show that you get the same result with a simplified geometry.
We assume therefore that the force only acts on the distance 2ρ, and that the force is constant on this distance and equal the force, when the particle passes the mssive body at a distance ρ.
We get the expression for force from the law of gravitation.

$F = G\frac{mM}{\rho^2}$, if we put $\Delta s = 2\rho$, the time for traversing Δs is: $\Delta t = \frac{\Delta s}{v} = \frac{2\rho}{v}$

The (small) deflection (in radian) of the particle can then be calculated from: $\alpha = \frac{\Delta v}{v}$, and the acceleration of the particle is:

$$a = \frac{\Delta v}{\Delta t} = \frac{F}{m} = \frac{GM}{\rho^2} \quad \Rightarrow \quad \Delta v = \frac{GM}{\rho^2}\Delta t = \frac{GM}{\rho^2}\frac{2\rho}{v} = \frac{2GM}{\rho v}$$

In this way we find an expression for the deflection:

$$\alpha = \frac{\Delta v}{v} = \frac{2GM}{\rho v^2}$$

If you substitute for ρ, the radius of the sun r_{sun},, we arrive at the formula:

(1.14) $$\alpha = \frac{2GM}{r_{sun} v^2}$$

One should notice however, that the deflection is independent of the mass of the particle.
This is actually no surprise, since the heavy masse (from Newton's 2. law) is equal to the gravitational mass, according to the weak equivalence principle.

If we then "carelessly" substitute the velocity *v* of the particle with *c*, the speed of light, we could get an (non relativistic) indication of the deflection angle for a light ray passing near to the sun.
The deflection is

$$\alpha = \frac{2GM}{r_{sol} c^2} \qquad \alpha = \frac{2 \cdot 6{,}67\,10^{-11} \cdot 2{,}0\,10^{30}}{6{,}95\,10^{8} \cdot (3{,}0\,10^{8})^2} = 4{,}27\,10^{-6} = 2{,}45\,10^{-4}\ \text{deg}$$

An arc second = $1/3600^0$ = $2{,}78\ 10^{-4}$ degree, so Newton's non relativistic theory predicts a deflection of about 1 arc sec.

It is remarkably that the deflection is of the same magnitude as is predicted from General Relativity, however General Relativity predicts (fortunately) exactly the double deflection.

It is obvious that if a deflection of this magnitude should be measured, it certainly requires a telescope of extreme accuracy. And this became the challenge for the prominent physicist and astronomer A. S. Eddington.
In the early days following Einstein's publication in 1915, Eddington had the reputation of being among the *three* persons in the world who had a full understanding of Einstein's theory (although he wondered whom the third one might be)

The only indirect confirmation of the General Theory was hitherto a General Relativity calculation of the precession of the perihelion of Mercury, but when there occurred a total solar eclipse in may 1919. Eddington decided to go to the Caribbean island of Principe, where he made observations at the total solar eclipse on the 29'th of may 1919.

As stated above it required rather sensitive telescope observations to measure deflection of 2 arc sec, which was the deflection from the normal line of sight to the observed star but nevertheless Eddington concluded that's what he did.
Actually a similar measurements performed in Brazil seemed to point towards the Newtonian deflection, but it totally drowned in international noise of Eddington's alleged success.

1.6 Gravity induced index of refraction

At a given position r with gravitational potential $\phi(r)$ a determination of the speed of light involves the measurement of a displacement dr for a time interval $d\tau$, measured by a clock at rest at this position. The result is the light speed, according to a local proper time.

(1.15) $$\frac{dr}{d\tau} = c$$

Because of the gravitational time dilation as stated in (1.12)

(1.12) $$d\tau_1 = \left(1 + \frac{\Phi_1 - \Phi_2}{c^2}\right) d\tau_2$$

An observer in another position with a different gravitational potential, will obtain a different value for the speed of light, when using a clock located at his position.
We choose a reference point far away from the gravitational source, where $\phi(\infty) = 0$.
Here the time is denoted t, whereas at a distance r from the source the local proper time is τ
According to (1.12) we have

(1.16) $$d\tau = \left(1 + \frac{\Phi(r) - \Phi(\infty)}{c^2}\right) dt \qquad\qquad d\tau = \left(1 + \frac{\Phi(r)}{c^2}\right) dt$$

This implies that the speed of light is measured by a remote observer as

(1.17) $$c(r) = \frac{dr}{dt} = \left(1 + \frac{\Phi(r)}{c^2}\right)\frac{dr}{d\tau} = \left(1 + \frac{\Phi(r)}{c^2}\right) c \qquad \text{(where } \Phi(r) < 0\text{)}$$

Namely the speed of light will be seen by an observer, with his coordinate system to vary from position to position, as the gravitational potential varies from position to position.
For such an observer the effect of the gravitational field is viewed as a gravitational index of refraction in space.

(1.18) $$n(r) = \frac{c}{c(r)} = \left(1 + \frac{\Phi(r)}{c^2}\right)^{-1} \approx 1 - \frac{\Phi(r)}{c^2}$$

One should stress that the deviation of $c(r)$ from c does not mean, that the speed of light has changed, or is no longer a universal constant, but merely signifies that clocks at different positions in a gravitational field measures a different time.
The term gravitational index of refraction is actually misleading because it is an entirely different phenomena than the normal index of refraction, when light changes speed, when passing from one transparent material to another.

2. Differential geometry and curved space

Einstein's law of gravitation is formulated in the geometric framework of curved space-time, so we shall now depart from the Newtonian theory of gravitation as an attraction acting on a distance.

The mathematical framework in which the General Theory is formulated is Riemannian differential geometry, which is a (non trivial) generalization of the Gaussian theory of intrinsic theory of two dimensional surfaces, but founded in Euclidian geometry.

Differential geometry is a description of surfaces applying generalized coordinates, which are intrinsic to that space. While a two dimensional surface can be viewed as embedded in a three dimensional surface this is not the case when we consider four dimensional space-time.

We may elaborate a bit on this.
A surface in 3D space has a parametric description: $(x(u_1, u_2), y(u_1, u_2), z(u_1, u_2))$, where the parameters u_1, u_2 belong to some interval.
The main achievement of Gauss is that he was to be able to describe the geometric properties of a surface (including curvature), in terms of the intrinsic coordinates (u_1, u_2), and that the calculated curvature is independent of the choice of parameters.
Gaussian differential geometry is treated in detail in: "Differential geometry 1", on my physics Website.

The necessity of applying intrinsic coordinates to obtain the Gaussian curvature of surfaces also explains why differential geometry is necessary for describing warped 4D space-time.
However in 4D space-time the geometric must be treated in an highly abstract and complex form of Riemannian generalized differential geometry.

Riemannian geometry and Tensor Analysis are treated in some detail in the paper: "Differential geometry 2". However, while Gaussian differential geometry is mathematically demanding, it is still conceptually palatable, Riemannian geometry and tensor analysis, on the other hand is highly abstract and conceptually complex.

The foundation of any geometric description requires a metric. In differential geometry it is called a metric form.

The metric in ordinary Euclidian space is position independent, (it is the same at every point of space, and the square of the distance element is:

$$ds^2 = dx^2 + dy^2 + dz^2$$

The square of the distance element is an invariant when changing the coordinate system.
In differential geometry the metric is a quadratic form, which in general is position dependent.

(2.1) $$ds^2 = g_{i\,j} dx^i dx^j$$

$g_{i\,j}$ is called the metric form or just the metric. It is a function of the generalized coordinates x^i.
That the coordinates have upper indices, reflects the tensor notation, since the coordinate differentials transform as a contravariant tensor.
Furthermore we have used the summation convention (due to Einstein), that whenever an index appears twice summation is implied. For more details see: "Differential geometry 2"

The (Euclidian) 4-dimensional Minkowski space of Special Relativity has the coordinates:

(*ct, x, y, z)*, an the distance element is written

(2.2) $$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$$

This distance element is position independent and positive definite, since dividing by *dt* gives:

$$\left(\frac{ds}{dt}\right)^2 = c^2 - \left(\frac{dx}{dt}\right)^2 - \left(\frac{dy}{dt}\right)^2 - \left(\frac{dz}{dt}\right)^2 = c^2 - v_x{}^2 - v_y{}^2 - v_z{}^2 = c^2 - v^2 > 0$$

For various reasons, the "distance element" is often used with the opposite sign:

(2.3) $$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$$

This has the advantage, that when projecting into the usual three dimensional Euclidian subspace, we regain the usual distance element.
Here we shall follow most books on General Relativity and use (2.3).
For the "flat" Minkowski space the metric form is therefore:

(2.4) $$g_{ij} = diag\ (-1, 1, 1, 1)$$

Where "*diag*" signifies that all non diagonal elements are zero. For details of the metric form see "Differential Geometry 1" and "Differential geometry 2".

In his search for a relativistic description of gravity and to overcome the Newtonian theory as gravity acting of a distance, Einstein realized that the Riemannian geometric description with a curved Minkowski space-time might be the solution.
Firstly, such a theory was born to comply with Special Relativity, which had shown to be incompatible with Newton's law of gravitation.

Far away from massive bodies the geometry is the flat Minkowski space of Special Relativity, but near massive bodies the space is warped, and the metric form is no longer diagonal.

In differential geometry a **geodesic** in a surface (or in a higher dimensional space) is the shortest path between to points, if you are compelled to stay on the surface.
The geodesics in Euclidian space are straight lines, and the geodesics on a sphere are great circles i.e. circles that have the same diameters as the sphere

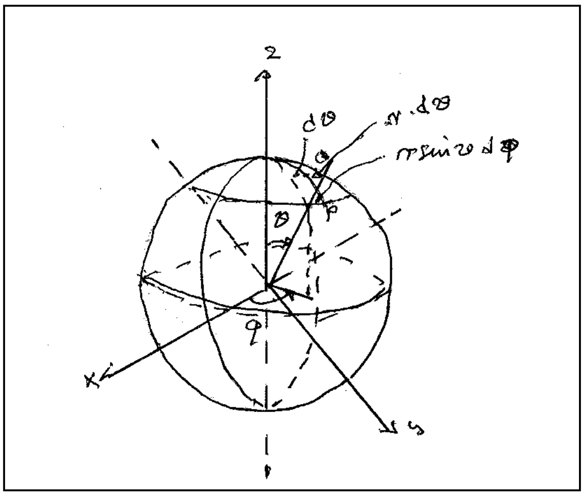
How can this be combined with movements in a gravitational field?
Well one of the most fundamental principle in theoretical physics is the "Principle of Least Time", which states that the path followed by a particle always is the path of least time.
For a light ray one can rather easily show that the law of refraction is a consequence of the Principle of Least Time. This principle is also the foundation of Lagrangian analytic mechanics.

In the description of the field of gravitation as a warped space, a particle will follow a geodesic according to the Principle of Least Time. Surely this is a description that differs largely from the description of a particle moving in a gravitational field. But fortunately General Relativity converges to the Newtonian description in the non relativistic limit.

In the presence of massive bodies 4D space-time is warped, the metric form is position dependent, and each point has a curvature given by the Riemannian curvature tensor.
However, mathematically it is very complex. For details see “Differential geometry 2”.

2.1 Metric on a sphere

The simplest example of a curved surface is the sphere. A point *P*(x, y, z) has the polar coordinates (θ,φ) See figure below. θ is the polar angle and φ is the azimuth angle. The coordinates are when expressed by the polar angles.



$x = r\cos\theta\cos\varphi,\;\; y = r\cos\theta\sin\varphi,\; z = r\sin\theta$

To come from a point *P* (θ,φ) to $Q(\theta + d\theta,\ \varphi + d\varphi)$ you first go “vertically” the distance $r\,sin\theta d\theta$ and horizontally the distance $rd\varphi$. Since the two steps are perpendicular to each other, the square of the distance is:

(2.5) $ds^2 = r^2 d\theta^2 + r^2\sin^2\theta d\theta^2$

This means that the metric tensor is.

$$g_{\theta\varphi} = \begin{pmatrix} r^2 & 0 \\ 0 & r^2\sin^2\theta \end{pmatrix}$$

As mentioned before, the geometry on a surface can be visualized, but this is certainly not the case for a curved 4D-space. Here we must resort to two dimensional analogies and projections on a two dimensional subspace.

2.2 The geodesic equation as the equation of motion.

In the differential geometry of surfaces, as well as in the Riemann generalization to higher dimensions the geometry is defined by the metric form, which is generally a position dependent function of generalized coordinates. The distance element is defined by the quadratic form in generalized coordinates:

(2.5) $$ds^2 = g_{i\,j}dx^i dx^j$$

The derivation of the geodesic equation is done for 2D surfaces in “Differential geometry 1”, and for higher dimensions in “Differential geometry 2”. The method is the Lagragnian approach, with the Lagrange function:

$$L = \frac{ds}{dt} = \sqrt{g_{i\,j}(x)\dot{x}^i\dot{x}^j} \quad where \qquad x = (x^1, x^2, x^3, \ldots x^n)\,,$$

and solving the Lagrange equation:

(2.6) $$\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial \dot{y}} = 0$$

The result is

(2.8) $$\frac{d^2x^j}{ds^2} + \Gamma^j_{ik}\frac{dx^k}{ds}\frac{dx^i}{ds} = 0$$

Here *s* is a natural parameter i.e. the length of the curve from a reference point, and Γ^j_{ik} is the so called Christoffel symbol.

(2.9) $$\Gamma^k_{ij} = \tfrac{1}{2} g^{kl}\left(\frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l}\right)$$

It makes little sense to try to solve the equation (2.8), since it consists of four second order coupled differential equation, and even if one had an analytic expression for the 4x4 components of the metric, the task seem unattainable. Even to show that the geodesics on a sphere are great circles is controversial. See "Differential Geometry 1" for details.

The differential equation for the geodesic has a direct connection to the concept of **covariant differentiation**. It turns out that the components of the derivative of a vector do not transform covariantly, that is, they are not tensors, even when the differential operator and a vector both are good tensors.. Instead it is possible, with the help of the so called **Christoffel** symbols to form a differential operator, which transforms as a tensor under coordinate transformation.

$$D_i v^j = \frac{dv^j}{ds} + \Gamma^j_{ik} v^k \frac{dx^i}{ds}$$

D denotes covariant differentiation. If we replace the vector v^j with $\frac{dx^i}{ds}$ we get:

(2.10) $$D_i\left(\frac{dx^j}{ds}\right) = \frac{d^2x^j}{ds^2} + \Gamma^j_{ik}\frac{dx^k}{ds}\frac{dx^i}{ds}$$

If we put the covariant derivative to zero, it marks "a straight line" in a curved space.
For example the covariant derivative on a sphere is zero along a great circle.
Not surprisingly the "straight lines" in a curved space is also the shortest path between two points, and if we put the covariant derivative to zero, we find again the geodesic equation.

(2.11) $$D_i\left(\frac{dx^j}{ds}\right) = 0 \quad \Leftrightarrow \quad \frac{d^2x^j}{ds^2} + \Gamma^j_{ik}\frac{dx^k}{ds}\frac{dx^i}{ds} = 0$$

It is mandatory for an enlargement of a physical theory that it complies with the existing theory in some limit. Concerning General Relativity, it must comply with Newtonian gravitation and mechanics in the non relativistic limit. To confirm this is however far more circumstantial than is the case of Special Relativity.

2.3 The Newtonian limit of General Relativity

The Newtonian limit means:

$$\frac{dx^i}{dt} << c \Leftrightarrow dx^i << cdt \qquad \text{it follows}$$

$$\frac{dx^i}{ds} << c\frac{dt}{ds} \Leftrightarrow \frac{dx^i}{ds} << \frac{dx^0}{ds} \qquad \text{since } dx^0 = cdt$$

Keeping only the dominant term in the double sum over *i* and *j* the geodesic equation:

(2.12) $$\frac{d^2x^i}{ds^2} + \Gamma^i_{jk}\frac{dx^j}{ds}\frac{dx^k}{ds} = 0$$

Reduces to:

(2.13) $$\frac{d^2x^i}{ds^2} + \Gamma^j_{00}\frac{dx^0}{ds}\frac{dx^0}{ds} = 0$$

In a static field, $\frac{\partial g_{ij}}{\partial x^0} = 0$, so because all the time derivatives vanish, the Christoffel symbol takes the simple form

(2.14) $$\Gamma^i_{00} = -\tfrac{1}{2}g^{ij}\frac{\partial g_{00}}{\partial x^j}$$

If we furthermore take the limit of a weak field, and we assume that the metric form is not too different from the flat space time metric: $d_{ij} = diag(-1,1,1,1)$, so we write

(2.15) $$g_{ij} = d_{ij} + \varepsilon_{ij}$$

Where $\varepsilon_{ij}(x)$ is a small correction to the field: $\varepsilon_{ij}(x) << d_{ij}$.

What we get is $\frac{\partial g_{ij}}{\partial x_k} = \frac{\partial \varepsilon_{ij}}{\partial x_k}$, so the Christoffel symbol is of the same order as ε_{ij}.

To the leading order we have since d_{ij} constant.

(2.16) $$\Gamma^i_{00} = -\tfrac{1}{2}d^{ij}\frac{\partial \varepsilon_{00}}{\partial x^j}$$

Which, because d^{ij} is diagonal, (2.16) has for a static ε_{00} has the following components:

(2.17) $$-\Gamma^0_{00} = -\tfrac{1}{2}\frac{\partial \varepsilon_{00}}{\partial x^0} = 0 \qquad \text{and} \qquad -\Gamma^i_{00} = -\tfrac{1}{2}\frac{\partial \varepsilon_{00}}{\partial x^i} = 0$$

If we put (2.17) into the non relativistic geodesic equation (2.13) the $i = 0$ part leads to:

(2.18) $$\frac{\partial x^0}{ds} = \text{constant}$$

And the $i > 0$ part is:

(2.19)
$$\frac{d^2x^i}{ds^2}+\Gamma^i_{00}\frac{dx^0}{ds}\frac{dx^0}{ds}=\left(\frac{d^2x^i}{c^2dt^2}+\Gamma^i_{00}\right)\left(\frac{dx^0}{ds}\right)^2=0$$

To obtain the last equation, we have used:

$$\frac{d^2x^i}{ds^2}=\frac{d^2x^i}{dx^{0\,2}}\left(\frac{dx^0}{ds}\right)^2 \quad and \quad dx^0=cdt$$

The equation above together with (2.17) implies

(2.20)
$$\frac{d^2x^i}{c^2dt^2}-\tfrac{1}{2}\frac{\partial\varepsilon_{00}}{\partial x^i}=0$$

This we should compare with the Newtonian equation of motion

(2.21)
$$\frac{\partial^2 r}{dt^2}=-\nabla\varphi$$

Thus $\varepsilon_{00}=-2\Phi/c^2$, and using $g_{ij}=d_{ij}+\varepsilon_{ij}$, where d_{ij} =*diag*(-1, 1, 1, 1),

(2.22)
$$g_{00}=-\left(1+\frac{2\Phi(x)}{c^2}\right)$$

From (2.22) we may regard the metric tensor as a relativistic generalization of the gravitational potential. At the same time it also provides us with a criterion of the gravitational potential to be week namely $|\Phi|/c^2<<1$.

Let us for example calculate the gravitational potential at the surface of the earth divided by c^2. We know that

$$g=G\frac{M}{R^2}$$

so the potential $\Phi(r)=-G\dfrac{M}{R}=-gR=-9.82\cdot 6.370\,10^6\ m^2/s^2=6.26\,10^7\ m^2/s^2$

It then follows: $|\Phi|/c^2=6.95\,10^{-10}$.

So even in a gravitational field about a billion times larger than that from the earth will still be “a weak field”. A fact which overwhelmingly justify our approximations above.

2.4 Space-time is locally flat

In the Gaussian differential geometry every point P on the surface has a tangent plane, which in the vicinity of P approximates the surface. That is, in the vicinity of P, the surface is Euclidian. Assume that the surface has a parameter representation (u^1, u^2)

If $\boldsymbol{D_1P}(u^1, u^2)$ and $\boldsymbol{D_2 P}(u^1, u^2)$ are the tangent vectors along u^1 and u^2 , we can make a coordinate transformation to a system determined by $\boldsymbol{D_1P}$, $\boldsymbol{D_2 P}$ and $\boldsymbol{n} = \boldsymbol{D_1PxD_2 P}$.
If we denote the coordinates in this system (x^1, x^2, x^3), and make a Taylor expansion of the surface given by $x^3 = f(x^1, x^2)$ around P in these coordinates, the partial derivatives $\frac{\partial f}{dx^1} = 0$ and $\frac{\partial f}{dx^2} = 0$ so a differentiable surface will always be flat in the vicinity of P. If we expand to the second derivative, it can be shown, that we get an expression:

(2.23) $$x^3 = \kappa_1 {x_1}^2 + \kappa_2 {x_2}^2$$

where κ_1 and κ_2 are the main curvatures, that is, the curvatures along u^1 and u^2.
Depending of the sign of the Taylor expansion will be a hyperboloid or an elliptic parabola.
Although it is simple to visualise the "flatness theorem" in a Euclidian space, it is a bit more complicated, when you look at the 4 dimensional space-time, but the method is generally the same.

The flatness theorem for space-time states that it is always possible to find a coordinate transformation, so that the metric form is "flat" at some point. That is $x^i \to \bar{x}^i$ and $g_{ij} \to \bar{g}_{ij}$ and such that $\bar{g}_{ij}(0) = \delta_{ij}$ and that $\bar{g}_{ij}(\bar{x})$ is constant up to the second correction, that is, the first order derivatives vanishes.

We prove this by an explicit construction of the transformation. To determine the transformation, we take advantage of the definition of the covariant derivative:

$$D_i v^j = \frac{dv^j}{ds} + \Gamma^j_{ik} v^k \frac{dx^i}{ds}$$

and the fact that the "straight lines" in space-time is given by the geodesic equation

$$D_i\left(\frac{dx^j}{ds}\right) = \frac{d^2x^j}{ds^2} + \Gamma^j_{ik} \frac{dx^k}{ds}\frac{dx^i}{ds} = 0$$

and we choose (being encouraged by the geodesic equation) the transformation.

(2.24) $$\frac{\partial x^i}{\partial \bar{x}^j} = \delta^i_j - \Gamma^i_{jk} \bar{x}^k$$

Making a power expansion of the metric form up to the first order:

(2.25) $$g_{kl}(x) = g_{kl}(0) + \frac{\partial g_{kl}}{\partial x^m} x^m$$

The transformation of a contravariant v^i vector from the coordinates x to the coordinates $\bar{x}$ is given by:

$$v^i = \frac{\partial x^i}{\partial \overline{x}^j}\overline{v}^j$$

Using this for x^i gives:

(2.26) $$x^i = (\delta^i_j - \Gamma^i_{jk}\overline{x}^k)\overline{x}^j \quad \Leftrightarrow \quad x^i = \overline{x}^i - \Gamma^i_{jk}\overline{x}^k\overline{x}^j$$

The transformation equation for the metric form is:

(2.27) $$\overline{g}_{ij}(x) = \frac{\partial x^k}{\partial \overline{x}^i}\frac{\partial x^l}{\partial \overline{x}^j} g_{kl}(x)$$

Inserting (2.24) and (2.25) w have

(2.28) $$\overline{g}_{ij}(x) = \frac{\partial x^i}{\partial \overline{x}^j}\frac{\partial x^k}{\partial \overline{x}^i} g_{kl}(x) = (\delta^i_j - \Gamma^i_{jk}\overline{x}^k)(\delta^k_i - \Gamma^k_{lm}\overline{x}^m)(g_{kl}(0) + \frac{\partial g_{kl}}{\partial x^m}x^m + ...)$$

Keeping only up to first order terms, using (2.26) on the last tem, this can be reduced to:

(2.29) $$\overline{g}_{ij}(x) = g_{ij}(0) - (\Gamma^k_{il}g_{kj}(0) + \Gamma^k_{lj}g_{ik}(0) - \frac{\partial}{\partial x^l}g_{ij})x^l$$

The term in the bracket vanishes because of the identity: (See "Differential Geometry 2" for details)

(2.30) $$\frac{\partial}{\partial x^k}g_{ij} - \Gamma^l_{ki}g_{lj} - \Gamma^l_{kj}g_{il} = 0$$

So $\overline{g}_{ij}(x) = g_{ij}(0)$ to the first order of a Taylor expansion, which proves that there always exists a coordinate transformation, so that the space is locally (Euclidian) flat in the transformed system.

3. General Relativity as a geometric theory of gravity

The mathematical realization of the equivalence principle is the meaning of general covariance. This means that the equations of General Relativity must be covariant tensors to generalized coordinate transformations.
The Equivalence Principle means that for a body in a free fall, you can transform to an accelerated frame, such that acceleration and gravity vanishes, and you are left with an inertial frame.

General Relativity is a generalized differential geometric theory, where the geodesics are the trajectory of motion.

In this space ordinary differentiation is replaced by covariant differentiation, symbolized with a capital *D*.

(3.1) $$\frac{\partial v^j}{\partial s} \quad \text{is replaced by} \quad D_i v^j = \frac{dv^j}{ds} + \Gamma^j_{ik}v^k\frac{dx^i}{ds}$$

In Special Relativity, the equation of motion of a free particle moving in Minkowski space, may be written as the derivative of the four velocity u^i:

(3.2) $$\frac{du^i}{d\tau} = 0,$$

where τ is the proper time.

In General Relativity this corresponds to a free fall in a gravitational field, and the "only" difference is that the ordinary time derivative is replaced by the covariant derivative.

(3.3) $$\frac{Du^i}{d\tau} = 0$$

We can then demonstrate, that (3.3) is in fact the geodesic for the equation of motion. In (3.1) we replace the vector v^i with the velocity u^i to get:

(3.4) $$\frac{Du^j}{d\tau} = \frac{du^j}{d\tau} + \Gamma^j_{ik} u^k \frac{dx^i}{ds} = 0$$

If we insert $u^j = \frac{dx^j}{d\tau}$ we obtain.

(3.6) $$\frac{D^2 x^j}{d\tau^2} = \frac{d^2 x^j}{d\tau^2} + \Gamma^j_{ik} \frac{dx^k}{d\tau} \frac{dx^i}{d\tau} = 0$$

Which, we recognize as the geodesic (the path followed by a particle or a light ray) for a particle moving in a gravitational field. The geometry of the fields is hidden behind the Christoffel symbols.

3.1 Einstein's equation for the gravitational field

Einstein's theory of gravitation is a geometrical theory, where the gravitational field is identified with the warped space-time described by the metric form $g_{ij}(x)$, and x are generalized coordinates in 4-space.
Following the acceptance that space-time is curved and the Riemannian generalized differential geometry is the appropriate mathematical framework to describe it, we shall try to substantiate the Einstein Equation for the gravitational field.
The underlying mathematics is described in some detail in "Differential Geometry 2".
There we derive the Einstein equation, which is a covariant equation satisfying The Principle of Relativity.

We have already demonstrated that the metric tensor $g_{ij}(x)$ is the relativistic generalization of the gravitational potential.
In our analyze of the gravitational time dilation, we found in (1.16) the connection between the proper time τ and the time t measured by someone far away from the gravitational source.

(3.7) $$d\tau = \left(1 + \frac{\Phi(r)}{c^2}\right) dt$$

Where $\Phi(r)$ is the gravitational potential in the distance r from the source.
Since g_{ij} is a quadratic form in the generalized coordinates, and g_{00} is the time component, we shall make the identification.

(3.8) $$g_{00} = -\left(1+\frac{\Phi(x)}{c^2}\right)^2 \approx -\left(1+\frac{2\Phi(x)}{c^2}\right) \qquad \text{since } \frac{\Phi(x)}{c^2} << 1$$

This comes about because $ds^2 = g_{00}dx^0dx^0$ when $d\vec{x} = 0$
We have introduced the energy momentum tensor T_{ij} , as the (General Relativity) generalization of the mass density. In the same limit as above, we may write: $\rho(x) \to T_{00}$.
The General Relativity equation of gravity must have the proper non relativistic limit.

(3.9) $$\nabla^2\varphi = 4\pi G\rho$$

This equation is the familiar Newton law of gravitation written in differential form.
φ is the gravitational potential, ρ is the mass density and G is the gravitational constant.
From this one can infer that the General Relativity equation of gravity must have the form:

(3.10) $$G_{op}g = \kappa T$$

Where G_{op} is some differential operator, which must be a symmetric tensor of second rank, that is, it must be covariantly constant, implying that the covariant derivative = 0.
G_{op} is both of second order in the derivative of the metric form g and quadratic in g.

In “Differential Geometry 2”, it is substantiated that there is in fact only one such tensor, within the framework of Riemannian geometry, namely the so called Einstein tensor, symbolically written as G_{ij}.
Einstein therefore proposed the following fundamental tensor equation to replace the Newtonian non relativistic scalar theory.

(3.11) $$G_{ij} = \kappa T_{ij}$$

where the constant κ is to be determined by taking the Newtonian limit of this equation.

G_{ij} can be expressed by means of the Ricci tensor R_{ij} , and the Ricci scalar R. See “Differential Geometry 2” for more details.

(3.11) $$G_{ij} = R_{ij} - \tfrac{1}{2}Rg_{ij} = \kappa T_{ij}$$

If we take the trace of the equation we find: $-R = \kappa T$. Since $T = g^{ij}T_{ij}$. *w*e can rewrite the field equation in an equivalent form, by replacing Rg_{ij} by $-\kappa Tg_{ij}$. So we find the following form of Einstein’s field equation

(3.12) $$R_{ij} = \kappa(T_{ij} - \tfrac{1}{2}Tg_{ij})$$

3.1.1 Newtonian limit of the field equation

We shall then proceed to show that the Newtonian field equation for the gravitational field is simply the leading approximation to (3.12), for a non relativistic source producing a weak field, where weak field as usual means $\Phi/c^2 << 1$.

In the non relativistic limit of $v/c << 1$ and the rest energy density term T_{00} is dominant.
We will focus only at the (0, 0) components of (3.12), since all other terms are of order v/c.

(3.13) $$R_{00} = \kappa(T_{00} - \tfrac{1}{2}Tg_{00})$$

with

(3.14) $$T = g^{ij}T_{ij} \approx g^{00}T_{00} = \frac{1}{g_{00}}T^{00}$$

Thus (3.12) becomes

(3.15) $$R_{00} = \tfrac{1}{2}\kappa T_{00}$$

To recover the Newtonian limit, we need to show that $R_{00} \to \nabla^2 g_{00}$
From the definition of the Ricci tensor in terms of the Riemann curvature tensor R_{ikjl}, we have:

(3.16) $$R_{00} = g^{ij}R_{i0j0}$$

Here the indices only go from 1..3, since the components with 0 index vanishes, because of the symmetry properties of the Riemann tensor.

Since the Newtonian limit also correspond to the weak field limit, we shall only keep the second derivatives of the metric form in favor of the square of the first derivative components of the metric form.

(3.17) $$R_{ijkl} = \tfrac{1}{2}\left(\frac{\partial^2 g_{jl}}{\partial x^i \partial x^k} - \frac{\partial^2 g_{il}}{\partial x^j \partial x^k} - \frac{\partial^2 g_{ik}}{\partial x^j \partial x^l} - \frac{\partial^2 g_{jk}}{\partial x^i \partial x^l}\right)$$

(3.18) $$R_{00} = g^{ik}R_{i0k0} = \tfrac{1}{2}g^{ij}\left(\frac{\partial^2 g_{00}}{\partial x^i \partial x^j} - \frac{\partial^2 g_{i0}}{\partial x^j \partial x^k} - \frac{\partial^2 g_{i0}}{\partial x^k \partial x^l} - \frac{\partial^2 g_{k0}}{\partial x^i \partial x^l}\right)$$

Since the Newtonian limit also correspond to a static situation, we can drop all terms with a derivatives with respect to $x^0 = ct$. So (3.18) reduces to:

(3.19) $$R_{00} = \tfrac{1}{2}\nabla^2 g_{00}$$

Using (3.8) $g_{00} \approx -\left(1 + \frac{2\Phi(x)}{c^2}\right)$ and $T_{00} = \rho c^2$, then (3.15) $R_{00} = \tfrac{1}{2}\kappa T_{00}$ becomes.

(3.20) $$-\tfrac{1}{2}\nabla^2\left(1+2\frac{\Phi}{c^2}\right)=\tfrac{1}{2}\kappa\rho c^2$$

Or

(3.21) $$\nabla^2\Phi=-\tfrac{1}{2}\kappa\rho c^4$$

Thus we see that Einstein's equation indeed has the correct Newtonian limit

(3.22) $$\nabla^2\Phi=-4\pi G\rho$$

if we make the identification

(3.23) $$4\pi G\rho=\tfrac{1}{2}\kappa c^4 \quad \Leftrightarrow \quad \kappa=\frac{8\pi G}{c^4}$$

Putting this back into the Einstein equation we get:

(3.24) $$R_{i\,j}-\tfrac{1}{2}Rg_{i\,j}=\frac{8\pi G}{c^4}T_{i\,j}$$

Or when written in the equivalent form (3.12) $R_{i\,j}=\kappa(T_{i\,j}-\tfrac{1}{2}Tg_{i\,j})$

(3.25) $$R_{i\,j}=\frac{8\pi G}{c^4}(T_{i\,j}-\tfrac{1}{2}Tg_{i\,j})$$

This is a set of ten coupled nonlinear partial differential equations of which a more general solution does not exist. Only for a spherical symmetric source Karl Schwarzschild managed as early as in 1915, (when he was stationed at the eastern front in the First world war), to obtain a solution, which has been named after him.

3.2 Spherically symmetrical metric of space-time

We shall now demonstrate how Schwarzschild succeeded in finding an exterior solution to the Einstein equation valid for a spherically symmetric source with mass M.
The solution is the metric form $g_{i\,j}(x)$ for the space-time geometry lying outside the source, which has become known as the **Schwarzschild exterior solution**.

First we shall demonstrate that a spherical symmetric tensor has only two unknown functions to be determined, namely $g_{00}(t,r)$ and $g_{rr}(t,r)$.
For a spherical source the metric function must also be spherical symmetric, at least in the spatial part. Therefore we need the spherical coordinates on a 2D sphere derived in (2.5)

$$ds^2=r^2d\theta^2+r^2\sin^2\theta d\theta^2$$

This is easily to extend to 3D space, by adding dr^2.

(326) $$ds^2=dr^2+r^2d\theta^2+r^2\sin^2\theta d\theta^2\,.$$

With the two unknown functions mentioned above, the metric function becomes.

(3.27) $$g_{ij} = diag(g_{00}, g_{rr}, r^2, r^2 \sin^2\theta)$$

The solution should far away from a gravitational source approach a flat space-time.

Then we shall try to demonstrate the assertion above: That a spherically symmetric tensor of rank 2 has only two unknown scalar functions. The square of the distance element is.

$$ds^2 = g_{ij}(x)dx^i dx^j$$

It must be diagonal in the spherical symmetric case, and the spatial part must have the usual form

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\theta^2 .$$

One may argue that for the spherical symmetric case, the most general (diagonal) form of the metric function must result in the following expression for the square of the distance element.

(3.28) $$ds^2 = \alpha(dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\theta^2) + \beta r dr^2 + \gamma dr dt + \delta dt^2$$

Where α, β, γ and δ, are scalar functions of (*r, t*), which are to be determined.
From tensor analysis, we know that we can perform a general coordinate transformation, chosen to give the metric function its simplest form.
First we want to eliminate the product term *drdt*, and we do so by introducing a new time coordinate *t*'.

$$t' = t + f(r) \quad \Rightarrow \quad dt' = dt + \frac{df}{dr}dr \quad and \quad dt^2 = dt'^2 + (\frac{df}{dr})^2 dr^2 - 2\frac{df}{dr}drdt'$$

The product term $drdt'$ has now (from 3.28) the coefficient $(\gamma - 2\delta\frac{df}{dr})$, which can be eliminated by choosing *f* to satisfy the differential equation:

$$\frac{df}{dr} = \frac{\gamma}{2\delta}$$

Incidentally the absence of any linear *dt* term insures that the metric is time-reversal invariant.
The next step is to choose a new radial coordinate, so that the angular dependence becomes trivial.

$$r'^2 = \alpha(r,t)r^2$$

Then the right hand side of (3.28) reduces to $r^2(d\theta^2 + \sin^2\theta d\theta^2)$

With these new choices of coordinates the expression (3.28) is left with only two unknown scalar functions in the metric. Renaming *r'* and *t'* to *r* and *t, t*he metric then takes the diagonal form.

(3.29) $$ds^2 = g_{00}(r,t)c^2dt^2 + g_{rr}(r,t)dr^2 + r^2(d\theta^2 + \sin^2\theta d\theta^2)$$

In particular $g_{i\,0} = g_{0\,i}$ where $i \neq 0$ (i is the spatial index), which means that we may treat the spatial subspace separately.

For each value of t, we may visualize the spherical symmetrical space as a series of spheres having different radial coordinates, but with its centre at the origin of the symmetrical source.
On each sphere the clocks are synchronized. As a consequence of the spherical symmetry the proper radial distance ρ is defined as:

(3.30) $$dr^2 = ds_r^{\ 2} = d\rho^2$$

ds_r^2 is the invariant distance element with $d\theta = d\phi = 0$, and the coordinate time t is the proper time τ for an observer in a fixed location.

(3.31) $$-c^2dt^2 = ds_t^{\ 2} = -c^2d\tau^2$$

Thus the coordinates (r, t) have the physical interpretation as the radial distance and time measured by an observer far away from the spherical gravitational source.
In the presence of a gravitational source we have a warped space-time. In particular $g_{r\,r} \neq 1$, so there is a curvature in the spatial radial direction, leaving the proper radial distance $\rho \neq r$ as

(3.32) $$d\rho = \sqrt{g_{rr}}dr$$

Similarly the proper time differs from the coordinate time because $g_{00} \neq -1$ signifying the warping of the space-time in the time direction

(3.33) $$d\tau = \sqrt{-g_{00}}dt$$

3.3 The Schwarzschild solution for a spherical source

In classical dynamics we perceive the 3D space as the background, where physical processes take place. On the contrary, in the framework of General Relativity, space-time is the cause of physical processes. Space-time is determined by the distribution of mass and energy, (i.e. the energy momentum tensor), and generates the physical events.
Given the distribution of mass and energy and solve the Einstein equation to find the metric $g_{i\,j}(x)$.

In a following chapter, we shall derive the Schwarzschild exterior solution for a spherically symmetric source directly from Einstein's equation, but fearing that you might give up in the process, we merely state the solution here, elaborating on some important consequences.

(3.34) $$g_{00}(t,r) = -\frac{1}{g_{rr}(t,r)} = -1 + \frac{r_S}{r}$$

Here r_S is the constant scale of length, called the Schwarzschild radius. We shall see that the deviation from flat space-time of (2.3) $ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2$ is determined by the ratio r_S/r, and the resulting metric is called the Schwarzschild metric.

(3.35) $$g_{ij} = diag((-1+\frac{r_S}{r}),(1-\frac{r_S}{r})^{-1}, r^2, r^2\sin^2\theta)$$

We can then relate the Schwarzschild radius to already known quantities by considering the Newtonian limit (3.38)

(3.36) $$g_{00} = -\left(1+\frac{\Phi(x)}{c^2}\right)^2 \approx -\left(1+\frac{2\Phi(x)}{c^2}\right)$$

Comparing to (3.34)

(3.37) $$g_{00}(t,r) = -1+\frac{r_S}{r} = -1-\frac{2\Phi(x)}{c^2} \quad and \quad \Phi(x) = -G\frac{M}{r}$$

Leads to

(3.38) $$-1+\frac{r_S}{r} = -1-\frac{2GM}{c^2r} \quad \Rightarrow \quad r_S = \frac{2GM}{c^2}$$

The Schwarzschild radius is generally a very small distance, at least in our solar system e.g.

(3.39) $$r_{sun} = 3\ km \qquad r_{earth} = 9\ mm$$

So in general the ratio r_S/r, which signify the space-time deviation from a flat Minkowski metric, is a very small quantity.
For the exterior solution to be applicable the smallest value that r can take is the radius of the spherical source. For the sun and the earth the above relation translate into:

(3.39) $$\frac{r_{S\,(sun)}}{R_{sun}} = O(10^{-6}) \quad \text{and} \quad \frac{r_{S\,(earth)}}{R_{earth}} = O(10^{-10})$$

Clearly the second term in the metric

$$g_{ij} = diag((-1+\frac{r_S}{r}),(-1+\frac{r_S}{r})^{-1}, r^2, r^2\sin^2\theta)$$

becomes singular at $r_S = r$.

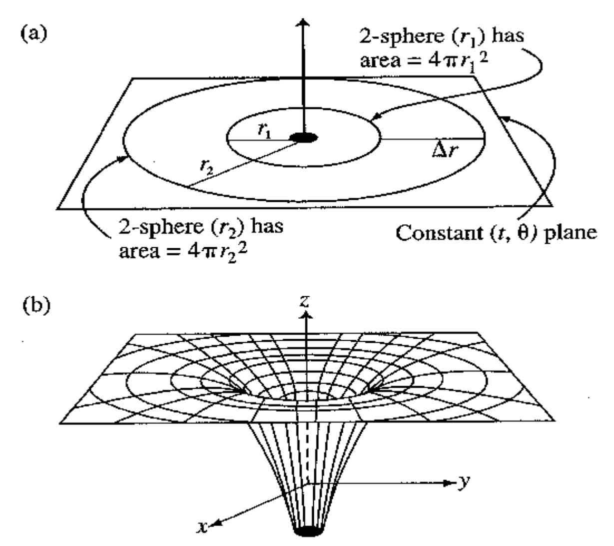
However this singularity did no draw much attention in the physical society in the early days of General Relativity, mostly because the physicists occupied in this field did not consider $r_S = r$ to be physically realizable for the simple reason that r_S is so extremely small, and $r_S << r$ for the exterior solution.
Another objection was that in the case where $r_S > R_{source}$ the density of the source should be unrealistically high, actually comparable to the density of nuclei.
Only 30 years later this possibility was recognized in what is now known as Black Holes.

For a stellar object being a Black Hole having a mass comparable to the sun, the density would be:

(3.40) $$\rho_S = \frac{M}{\frac{4}{3}\pi r_S^{\;3}} = \frac{2\,10^{30}}{\frac{4}{3}\pi(3.0\,10^3)^3} = 1.8\,10^{19}\ kg/m^3$$

A number, which is close to the nuclear density.

Although it is easy to visualize a 2D curved surface from a 3D Euclidian space, it is clearly not possible to visualize a 3D curved space from a higher dimension space. To do so it is necessary resorting to look into projections of the space (ct, r, θ, ϕ) onto a 2D subspace.
The most obvious choice is a fixed t, and $\theta = \pi/2$. This will reveal a 2D surface in the familiar coordinates (r, ϕ). Figure (3.41).



(a) Shows the $\theta = \pi/2$ plane (r, ϕ) cutting across the spherical source. (b) Shows a fictitious 3D embedding space, where the physical 2D subspace (a) is shown as a curved surface.
In this example, the drawings have used the Schwarzschild solution.

$$g_{rr} = (1 - \frac{r_S}{r})^{-1}$$

The singular nature of the space at $r_S \approx r$ is reflected by the steep slope of the curved surface near the Schwarzschild circle.
In the absence of gravity there is just a flat space as depicted in fig. (3.41) (a), but in the presence of gravity this becomes a curved 2D surface. If we want to visualize this warped 2D surface, one way to do so, is to embed it into a fictitious 3D Euclidian space as shown in figure (3.41) (b).
A particle moving in this $\theta = \pi/2$ plane will trace out a bended trajectory, as it follows a geodesic of the warped surface. To illustrate this, we have used the Schwarzschild solution for a compact spherical source. Here ρ is the proper radial distance. $\rho \neq r$, and $d\rho = \sqrt{g_{rr}}\,dr$

(3.42) $$\Delta\rho = \Delta r\sqrt{(1 - \frac{r_S}{r})}$$

In the embedded diagram the distance from the centre in the curved surface is ρ, while in the horizontal plane it is the coordinate distance r. Thus a small change in r correspond to a large change in ρ when r approaches r_S , the Schwarzschild radius.

3.4 Deflection of a light ray due to gravity

As mentioned earlier, Einstein deduced from the equivalence principle that a light ray will be a deflected when passing near the sun. And he concluded that this was due to the gravitational time dilation, expressed as the deviation from the flat space metric. See section (1.6).

(3.43) $$g_{00}(t,r) = -\left(1 - \frac{2\Phi(x)}{c^2}\right) = -1 - \frac{2GM}{c^2 r} = -1 + \frac{r_S}{r}$$

In General Relativity the warping of space-time takes place both in the radial and time direction: $g_{00} \neq -1$ and $g_{rr} \neq 1$.
We shall then proceed to calculate the deflection of a light ray (as a consequence of gravitational warping) arriving at the double angle from what we found earlier in the Newtonian approach.

Measuring this deflection was actually the first attempt of an experimentally verification of Einstein's theory. The experiment was conducted by the astronomer A. S. Eddington in a total solar eclipse on the island of Principe in the Caribbean in the year of 1919.

The bending of a light ray passing close to a massive body is similar to a light ray passing a lens. Let us consider a light ray in the fixed direction $d\theta = d\phi = 0$

(3.44) $$ds^2 = g_{00} c^2 dt^2 + g_{rr} dr^2 = 0$$

To an observer far from the source, using the coordinate time and radial distance (t, r), the effective speed of light is according to (1.17)

(3.45) $$c(r) = \frac{dr}{dt} = \left(1 + \frac{\Phi(r)}{c^2}\right)\frac{dr}{d\tau} = c\sqrt{-\frac{g_{00}(r)}{g_{rr}(r)}}$$

A slightly different way of arriving at the same results is by noting that the speed of light is universal.
Thus $c = d\rho / d\tau$ and using $d\rho = \sqrt{g_{rr}}\,dr$ and $d\tau = \sqrt{-g_{00}}\,dt$ this relates the proper distance and the proper time, corresponding to the coordinate time and distance (t, r).

From (3.45) it is then straightforward to find the "index of refraction".

(3.46) $$n(r) = \frac{c}{c(r)} = \sqrt{-\frac{g_{rr}(r)}{g_{00}(r)}} = \frac{1}{-g_{00}(r)} = \left(1 + 2\frac{\Phi(r)}{c^2}\right)^{-1}$$

The relation $$ds^2 = g_{00} c^2 dt^2 + g_{rr} dr^2 = 0$$

shows us that the retardation of the light signal is twice as large as given by the Newtonian

calculation(1.17): $$c(r) = \frac{dr}{dt} = \left(1 + \frac{\Phi(r)}{c^2}\right)\frac{dr}{d\tau} = \left(1 + \frac{\Phi(r)}{c^2}\right)c$$

From (3.46) we then find:

(3.47) $$c(r) == \left(1+2\frac{\Phi(r)}{c^2}\right)c$$

Inserting

$$\Phi(r) = -\frac{GM}{r} \quad \text{in} \quad n(r) = 1-2\frac{\Phi(r)}{c^2} \quad \text{gives} \quad n(r) = 1-2\frac{\Phi(r)}{c^2} = 1+2\frac{GM}{c^2 r}$$

To calculate the deflection we use the same method as the Newtonian calculation of the angular deflection, when a body at high speed passes the sun at a small distance.
The method is an approximation, but it is known to give the correct (not approximated) result.

First we differentiate (3.47) $c(r) == \left(1+2\frac{\Phi(r)}{c^2}\right)c = \left(1-2\frac{GM}{rc^2}\right)c$ with respect to *r*.

(3.48)

$$\frac{dc(r)}{dr} = 2\frac{GM}{r^2c^2} \quad \Leftrightarrow \quad dc(r) = 2\frac{GM}{r^2c^2}dr$$

The trick is to assume that the potential is constant equal to the value it has at the minimum radial distance $r_{min} = \rho$, and let it be effective only at a tangential distance 2 ρ. Namely ρ before it passes the sun, and ρ after. In this way $dr = \Delta r = 2\rho$, and we get for the deviation δ.

(3.49) $$\delta = 2\frac{GM}{\rho^2c^2}2\rho = \frac{4GM}{\rho c^2}$$

It is seen, that it is twice the deviation, of the Newtonian that we found earlier in (1.18).
The three figures below illustrate the phenomena of bending a light ray passing a massive star.
The first figure shows a (vastly exaggerated) illustration of how the line of sight changes, when a light ray passes near a massive star, or better a Black Hole

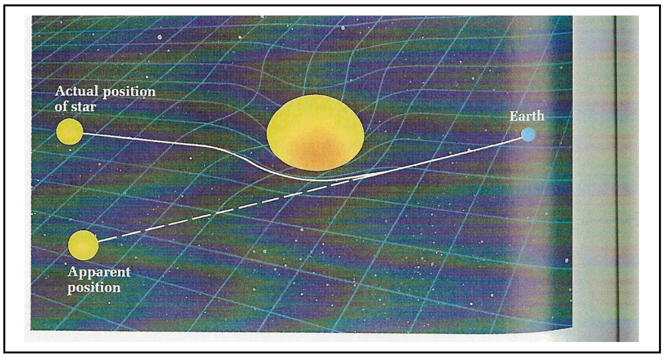


Fig (3.50)
Graphical illustration of the change in the line of sight, when a light ray is passing a massive star.

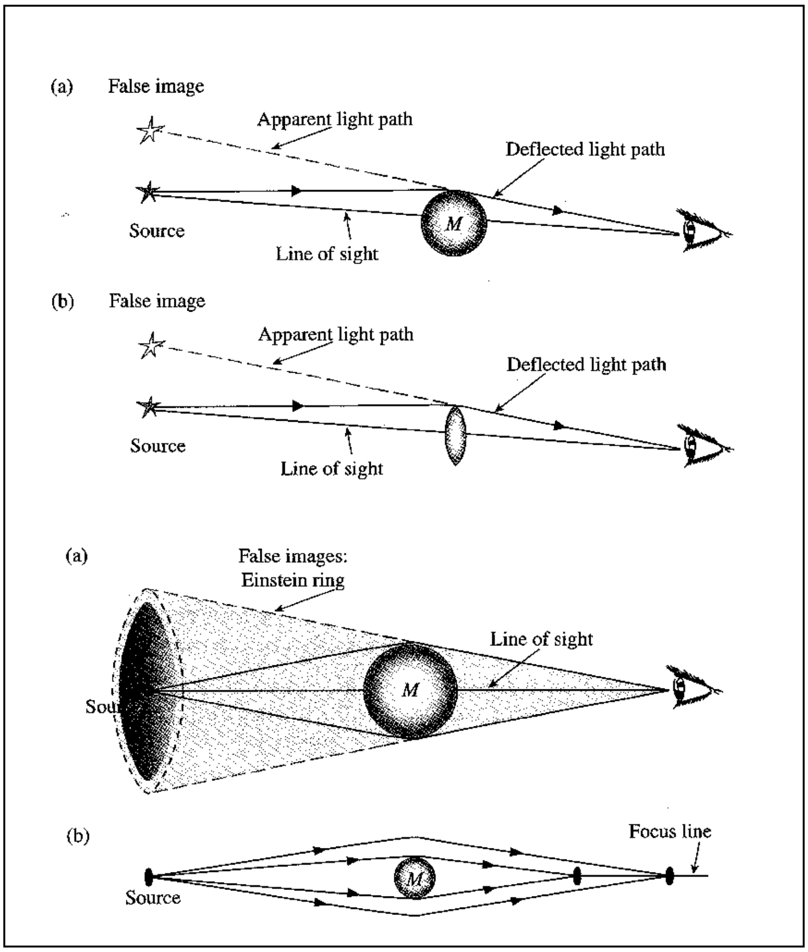


Fig (3.51) A graphical illustration showing the bending a light ray, when passing a massive star

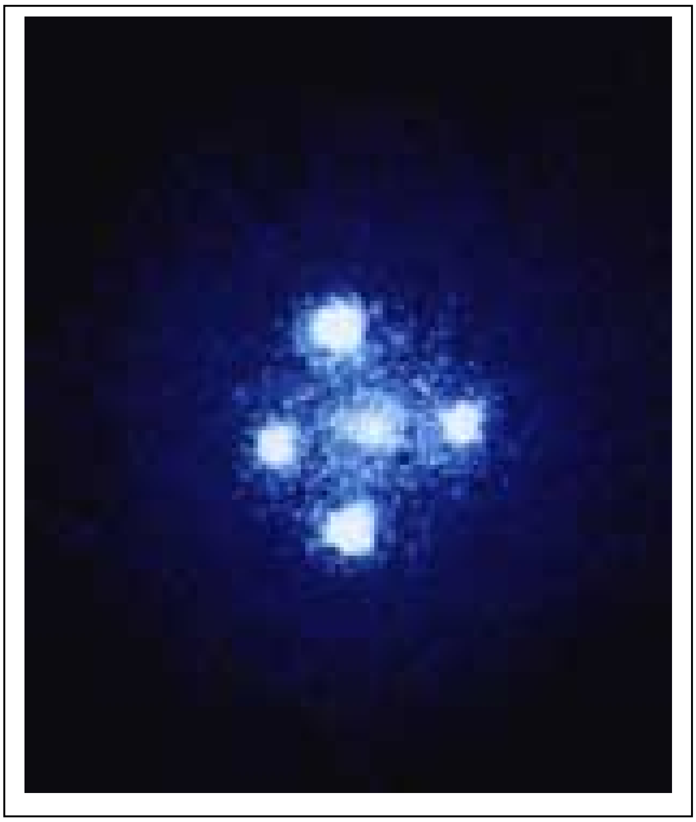


Figure (3.52). This is probably the most famous example of a gravitational lens. It shows a galaxy, which apparently has four nuclei.

A better explanation is that the four "nuclei" are really light from a very luminous stellar object (a Quasar) which is bended by a gravitational lens, probably a Black Hole, since it is "invisible"

3.4 Precession of Mercury's perihelion

The first and for several years the only prediction of General Relativity which deviated from Newtonian theory of gravitation, (and turned out to be correct), was Einstein's calculation of the rate of precession from the planet Mercury's perihelion

So this section deals with the motion a particle, using the Schwarzschild solution for a spherical symmetric field, and specifically we shall calculate the (very small) deviation from the Newtonian case

Newton's celestial mechanics has since its introduction in the eighteen century been remarkably successful. However, about 1850 astronomers had observed a very small discrepancy between theory and the observed precession of the planet Mercury's perihelion.

The pure r^{-2} force from Newton's law of gravitation predicted a closed elliptical orbit, which is fixed in space. However the perturbations due to the presence of other planets lead to a trajectory that was no longer closed. Since the perturbations are very small, the deviation from a closed orbit can be described as an elliptical orbit having a perihelion with precession.

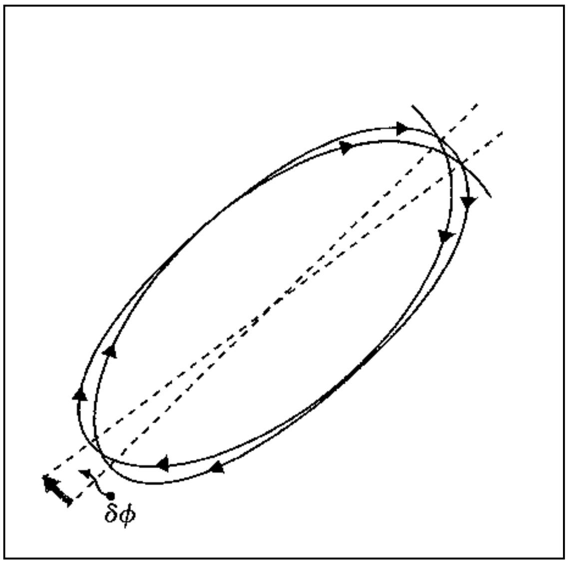


Fig. 3.53 An elliptical orbit, based on Newton's law of gravitation which is subject to small perturbations can be described by an elliptical orbit with a precession perihelion.

In the case of Mercury, the perturbation caused by the nearby planets can account for most of the perihelion advance, namely 1.556 ^{0} per century. However, there was still a discrepancy on 43 arc sec/century left to be accounted for.

Following a similar situation involving Uranus that eventually led to the prediction and discovery of the outer planet Neptune in 1846, a hitherto not discovered planet named Vulcan was predicted to lie inside the Mercury orbit. But it was never found. This was the perihelion precession problem that Einstein solved by applying his new theory of gravitation.

We have seen that corrections from General Relativity to masses corresponding to the mass of the Sun are of the order of 10^{-6}, and as the calculation shows, the correction is indeed small.

The calculations are far from simple, but since it is one of the only tangible confirmations of General Relativity, we shall present it in some detail.

The problem to be solved is as follows: Given the gravitational filed, that is, the Schwarzschild space-time due to the Sun, we are going to find the motion of the planet Mercury.

We do this by solving the geodesic equation by means of the Euler Lagrange equation

The well known Euler Lagrange equation is

(3.54) $$\frac{\partial L}{\partial x^i} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{x}^i} = 0$$

Using the Lagrangian. (For justification of this choice, see "Differential Geometry 2"

(3.55) $$L = \left(\frac{ds}{d\tau}\right)^2 = g_{ij}(x)\dot{x}^i\dot{x}^j$$

Where τ is the proper time, g_{ij} is the Schwarzschild metric, and a bullet over a variable as usual means differentiating with respect to time (τ).

Because of the conservation of angular momentum the trajectory will always remain in a plane spanned by the initial position vector and the momentum vector.
By choosing $\theta = \frac{1}{2}\pi$ in the Euclidian *x*- *y* plane, the Lagrangian takes the form.

(3.56) $$L = -\left(1 - \frac{r_S}{r}\right)c^2\dot{t}^2 + \left(1 - \frac{r_S}{r}\right)^{-1}\dot{r}^2 + r^2\dot{\varphi}^2 = -c^2$$

r_S is the Schwarzschild radius, and the last equation follows from $L = \left(\frac{ds}{d\tau}\right)^2 = -c^2$, since $ds^2 = -c^2 d\tau^2$ because $dx = dy = dz = 0$ when $t = \tau$.

If the Lagrangian $L = L(\dot{q}_i, q_i)$ does not explicitly depend on q_i then

(3.56) $$\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = 0 \quad \text{reduces to} \quad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = 0 \quad \text{or} \quad \frac{\partial L}{\partial \dot{q}^i} = const$$

Our *L* does not explicitly depend on ϕ and *t*. The two corresponding constants represent essentially the angular momentum *L* and the energy *E*. We invent two new integrating constants λ and ε.

(3.57) Angular momentum: $$\frac{\partial L}{\partial \dot{\varphi}} = const \quad \Rightarrow \quad 2r^2\dot{\varphi}^2 = \lambda$$

(3.58) Energy: $$\frac{\partial L}{\partial \dot{t}} = -2\left(1 - \frac{r_S}{r}\right)c^2\dot{t} = -2c^2\varepsilon \quad \Rightarrow \quad \left(1 - \frac{r_S}{r}\right)\dot{t} = \varepsilon$$

Multiplying (3.58) by $\frac{1}{2}m\left(1 - \frac{r_S}{r}\right)$ and inserting the two constants (3.57) and (3.58) we find in two steps:

$$L = -\tfrac{1}{2}m\left(1 - \frac{r_S}{r}\right)^2 c^2\dot{t}^2 + \tfrac{1}{2}m\dot{r}^2 + \tfrac{1}{2}mr^2\dot{\varphi}^2\left(1 - \frac{r_S}{r}\right) = -\tfrac{1}{2}mc^2\left(1 - \frac{r_S}{r}\right)$$

(3.59) $$L = -\tfrac{1}{2}m\,c^2\varepsilon^2 + \tfrac{1}{2}m\dot{r}^2 + \tfrac{1}{8}m\left(1 - \frac{r_S}{r}\right)\frac{\lambda^2}{r^2} = -\tfrac{1}{2}mc^2\left(1 - \frac{2GM}{c^2 r}\right)$$

Where we have used (3.38): $r_S = \frac{2GM}{c^2}$, or when rearranging the terms.

(3.60) $$\tfrac{1}{2}m\dot{r}^2 + \tfrac{1}{8}m\left(1 - \frac{r_S}{r}\right)\frac{\lambda^2}{r^2} - \frac{GmM}{r} = \tfrac{1}{2}m\varepsilon^2c^2 - \tfrac{1}{2}mc^2$$

Renaming the variables $\frac{\lambda^2}{4} = \frac{L^2}{m^2}$ *and* $\frac{\varepsilon^2 - 1}{2} = \frac{E}{mc^2}$ we obtain

(3.61) $$\tfrac{1}{2}m\dot{r}^2 + \left(1 - \frac{r_S}{r}\right)\frac{L^2}{2mr^2} - \frac{GmM}{r} = E$$

The choice of the variable E indicates of course, that it is the total energy, and apart from the factor $(1 - r_S/r)$, the equation is identical to that of Newtonian mechanics. Remember r_S/r is O(10^{-6})
If we define

(3.62) $$\Phi_{eff}(r) = \left(1 - \frac{r_S}{r}\right)\frac{L^2}{2m^2r^2} - \frac{GM}{r}$$

(3.61) can be written:

(3.63) $$\tfrac{1}{2}m\dot{r}^2 + m\Phi_{eff}(r) = E$$

The Newtonian equation

(3.64) $$\tfrac{1}{2}m\dot{r}^2 + \frac{L^2}{2mr^2} - \frac{GmM}{r} = E$$

is known to have the solution

(3.65) $$r = \frac{p}{1 + e\cos\varphi} \quad where \quad p = \frac{L^2}{m^2} \quad and \quad e = \sqrt{1 + \frac{2EL^2}{m\alpha^2}}$$

Where we have set: $\alpha = \frac{L^2}{GMm^2}$.

Since the factor $(1 - r_S/r)$ differs only from 1 by a small amount, it may be treated as a perturbation to the Newtonian solution in classic perturbation theory. So we proceed to solve the relativistic energy equation (3.61) as a standard central force problem, even if it becomes a bit complex.
To obtain a the solution $r = r(\phi)$, we first use the conservation of momentum:

(3.66) $$L = mr^2\dot{\phi} \quad \Rightarrow \quad \dot{\phi} = \frac{d\phi}{d\tau} = \frac{L}{mr^2}$$

Subsequently we make the substitution: $\rho = \frac{1}{r} \quad \Rightarrow \quad \frac{d\rho}{d\varphi} = -\frac{1}{r^2}\frac{dr}{d\varphi}$.

By the same token: $\dot{r} = \dfrac{dr}{d\tau} = \dfrac{dr}{d\phi}\dfrac{d\phi}{d\tau} = \dfrac{dr}{d\phi}\dfrac{L}{mr^2}$, according to (3.66).

Inserting in (3.61):

$$\tfrac{1}{2}m\dot{r}^2 + \left(1 - \frac{r_S}{r}\right)\frac{L^2}{2mr^2} - \frac{GmM}{r} = E$$

and reducing one finds:

(3.67) $$\left(\frac{d\rho}{d\varphi}\right)^2 + \rho^2 - \frac{2}{\alpha}\rho - r_S\rho^3 = C$$

Firstly, we solve the equation without the gravitational correction.

(3.68) $$\left(\frac{d\rho}{d\varphi}\right)^2 + \rho^2 - \frac{2}{\alpha}\rho = C$$

Where we have set $\alpha = \dfrac{L^2}{GMm^2}$, and C is some constant.

The equation (3.68) can be solved by differentiating with respect to ϕ.

(3.69) $$2\frac{d\rho}{d\varphi}\frac{d^2\rho}{d\varphi^2} + 2\rho\frac{d\rho}{d\varphi} - \frac{2}{\alpha}\frac{d\rho}{d\varphi} = 0$$

When reducing with $\dfrac{d\rho}{d\varphi}$ we obtain a more simple equation.

(3.70) $$\frac{d^2\rho}{d\varphi^2} + \rho = \frac{1}{\alpha}$$

which can be transferred to

(3.71) $$\frac{d^2(\rho - \frac{1}{\alpha})}{d\varphi^2} + (\rho - \frac{1}{\alpha}) = 0$$

Which is simply the differential equation for the harmonic oscillator.
The equation for the harmonic oscillator is usually written:

(3.72) $$m\ddot{x} + kx = 0 \quad \Leftrightarrow \quad \ddot{x} + \frac{k}{m}x = 0$$

And it has the solution:

(3.73) $$x = A\cos(\omega t + \varphi_0)$$

Where A is the amplitude, ω is the cyclical frequency, and ϕ_0 is the initial phase.

Copying this expression, we find the solution to (3.71)

(3.74) $$\rho = A\cos(\varphi) + \frac{1}{\alpha}$$

If we set $A = \frac{e}{\alpha}$ and substituting back $\rho = \frac{1}{r}$ we get:

(3.75) $$\frac{1}{r} = \frac{e}{\alpha}\cos\varphi + \frac{1}{\alpha} \quad \Leftrightarrow \quad r = \frac{\alpha}{1 + e\cos\varphi}$$

as already stated in (3.65).

The minimum distance is $r_{\min} = \frac{\alpha}{1+e}$, where e is the eccentricity, and $r(\frac{\pi}{2}) = \alpha = (1+e)r_{\min}$.

Now we are ready to make the relativistic correction, plugging in $\rho + \rho_1$, where ρ_1 is the correction to (3.71):

$$(\frac{d\rho}{d\varphi})^2 + \rho^2 - \frac{2}{\alpha}\rho - r_S\rho^3 = C.$$

(3.76) $$\frac{d(\rho+\rho_1)}{d\varphi}^2 + (\rho+\rho_1)^2 - \frac{2}{\alpha}(\rho+\rho_1) - r_S(\rho+\rho_1)^3 = C$$

$$\left((\frac{d\rho}{d\varphi})^2 + (\frac{d\rho_1}{d\varphi})^2 + 2\frac{d\rho}{d\varphi}\frac{d\rho_1}{d\varphi}\right) + (\rho^2 + \rho_1{}^2 + 2\rho\rho_1) - \frac{2}{\alpha}(\rho+\rho_1) - r_S(\rho^3 + 3\rho^2\rho_1 + 3\rho\rho_1{}^2 + \rho_1{}^3) = C$$

Then using (3.68): $$(\frac{d\rho}{d\varphi})^2 + \rho^2 - \frac{2}{\alpha}\rho = C$$

it reduces to:

(3.77) $$\left((\frac{d\rho_1}{d\varphi})^2 + 2\frac{d\rho}{d\varphi}\frac{d\rho_1}{d\varphi}\right) + (\rho_1{}^2 + 2\rho\rho_1) - \frac{2}{\alpha}\rho_1 - r_S(\rho^3 + 3\rho^2\rho_1 + 3\rho\rho_1{}^2 + \rho_1{}^3) = 0$$

Keeping only the leading and next leading terms of $\rho_1 = O(r_S)$, (3.77) reduces to:

(3.78) $$2\frac{d\rho}{d\varphi}\frac{d\rho_1}{d\varphi} + 2\rho\rho_1 - \frac{2}{\alpha}\rho_1 - r_S\rho^3 = 0$$

Inserting this solution to the unperturbed equation $\rho = A\cos(\varphi) + \frac{1}{\alpha}$ (3.78) becomes:

(3.79) $$-e\sin\varphi\frac{d\rho_1}{d\varphi} + e\rho_1\cos\varphi = \frac{r_S(1+e\cos\varphi)^3}{2\alpha^2}$$

You may then verify (if you have this kind of inclinations) that this equation has the solution.

(3.80) $$\rho_1 = \frac{r_S}{2\alpha^2}\left((3+2e^2) + \frac{1+3e^2}{e}\cos\varphi - e^2\cos^2\varphi + 3e\varphi\sin\varphi\right)$$

The first two terms have the form of a zero-order solution: $A+B\cos\phi$.
The third term, being periodic in ϕ as the zero-order terms are corrections that follow the non relativistic solution, and thus represent unobservable small corrections. So we shall only focus on the fourth non periodic term, which is ever increasing with ϕ. So we are left with:

(3.81) $$\rho_1 = \frac{r_S}{2\alpha^2}(3e\varphi\sin\varphi)$$

Replacing ρ with $\rho+\rho_1$ in the solution (3.74) $\rho = A\cos(\varphi) + \frac{1}{\alpha}$ gives:

(3.82) $$\rho + \rho_{1+} = A\cos(\varphi) + \frac{1}{\alpha} + \frac{r_S}{2\alpha^2}(3e\varphi\sin\varphi)$$

Using $\rho = \frac{1}{r}$ this can be cast into

(3.83) $$r = \frac{\alpha}{1 + e\cos\varphi + \delta e\varphi\sin\varphi}$$

where we have set $\delta = \frac{3r_S}{2\alpha}$ being a very small quantity.
The angular terms in the denominator can be cast into the same form as the denominator of the zero order solution, after approximating $\cos(\delta\phi) = 1$ and $\sin(\delta\phi) = 0$, using:

$$e\cos(\phi - \delta\varphi) = e\cos\varphi\cos(\delta\varphi) + e\sin\varphi\sin(\delta\varphi) \approx e\cos\varphi + e\delta\varphi\sin\varphi,$$

Finally getting:

(3.84) $$r = \frac{\alpha}{1 + e\cos((1-\delta)\varphi)}$$

According to (3.84) the planet returns to its perihelion r_{min}, not at $\phi = 2\pi$, but at $\varphi = 2\pi/(1-\delta) \approx 2\pi + 3\pi r_S/\alpha$, where

$$\alpha = \frac{L^2}{GMm^2} = (1+e)r_{\min},$$

namely the amount the perihelion advances (i.e. the whole orbit rotates) per revolution.

(3.85) $$\delta\varphi = 3\pi r_S\frac{GMm^2}{L^2} = \frac{3\pi r_S}{(1+e)r_{\min}}$$

Using the Schwarzschild radius for the Sun : r_S =2.95 *km*, Mercury's eccentricity e = 0.206 and its perihelion r_{min} = 4.6 10^7 *km* we can calculate the numerical value of the advance.

(3.86) $$\delta\varphi = 5 \cdot 10^{-7} \quad radian/revolution$$

Which equals 5 10^{-7} x 180/π x 60 x 60 = 0.103 ´´ (arc sec) per revolution.
In terms of the advance per century.

(3.87) $$0{,}103'' \text{ x } 0.103 \frac{100 \; years}{Mecury's \;\; period \; of \; 0.241 \; years} = 43'' \quad per \; century$$

This was in beautiful accordance with observations, being an outstanding triumph for Einstein's theory of gravitation, and it made the physical society aware, that the General Theory of Relativity was not just a mathematical abstraction, which had no relevance to the real world, but described realities of our physical world. (Almost the same situation, as when Einstein first presented his Special Theory of Relativity)

3.5 Black Holes

We shall now look into the space-time structure of an object with the mass so compressed, that its radius is smaller than the Schwarzschild radius $r_S = 2GM/c^2$. Such objects have been given the name "Black Holes", because it is impossible to transmit outwardly any signal, any light from the region inside the Schwarzschild radius.

One simple argument for such a strange behaviour is that the escape velocity from the "surface" of a Black Hole is greater than the velocity of light.
The escape velocity from a massive body is the velocity needed to escape its gravitational field.
But this explanation does not immediately apply in General Relativity. The correct explanation is that a Black Hole gives rise to an infinite gravitational time dilation, that is, a vanishing small effective velocity of light, as seen for an observer far away from the gravitational source.

3.51 A non relativistic approach to Black Holes

If this is difficult to comprehend, it is a curiosum that this result also can be achieved from non relativistic Newtonian mechanics. The escape velocity e.g. from the Earth can easily be found from energy conservation.
Let *M* be the mass of the Earth, *R* be the radius of the Earth and v_0 the escape velocity, with which an object with mass *m* is launched from the surface of the Earth. If the body is at rest at infinity, it follows from conservation of mechanical energy:

(3.88) $$\tfrac{1}{2} m {v_0}^2 - G \frac{mM}{R} = 0$$

Solving for v_0 gives:

(3.89) $$v_0 = \sqrt{\frac{2GmM}{R}}$$

For the Earth the escape velocity is 11.2 *km/s*.
Notice that the result is independent of the mass *m* of the body, so if tentatively put $v_0 = c$, and solve for *R*, we find:

(3.90) $$R = \frac{2GM}{c^2}$$

which is recognized as the Schwarzschild radius!

3.5.2 Singularities in the Schwarzschild metric

In General Relativity Black Holes appear as singularities to the Schwarzschild metric.

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

The metric has singularities at $r = 0$, at $r = r_S$, and at $\theta = 0$ and π. The last singularity is a coordinate singularity that can be transformed away by a different choice of coordinate.
It will not show up in physical measurements.
However the $r = 0$ is real. This is not surprising since the Newtonian gravitational potential has this singularity for a point mass.

The singularity at $r = r_S$ is technically also a coordinate singularity, that is, it can be transformed away, by a coordinate transformation. Meaning that physical measurements are in principle possible at $r = r_S$.
But it certainly does not mean that this singularity is not very special. $r = r_S$ constitutes an event horizon, where no signal emitted from within can reach an observer outside the event horizon.

A Black Hole is made possible by the collapse of a massive star, where a physical singularity is formed at $r = 0$, and a spherical collapse proceeds onward, until the sufficient mass and density is enough to form a event horizon.

3.5.3 Time measurements in the Schwarzschild space-time

As mentioned the Schwarzschild radius

$$r_S = \frac{2GM}{c}$$

is technically a coordinate singularity, and is therefore due to physical measurements.

Let us therefore consider a spaceship falling feely from the outside and traveling across the Schwarzschild surface of a Black Hole, and describe the time measurements made from the spaceship and from an observer far away (the coordinate time).
The general central force energy equation reads:

(3.92) $$\tfrac{1}{2} m\dot{r}^2 + \left(1 - \frac{r_S}{r}\right)\frac{L^2}{2mr^2} - \frac{GMm}{r} = E$$

The equation is simplified, if we specialize having no radial motion, i.e. the angular momentum $L = 0$, and if the motion starts from where the energy $E = 0$. It then reduces to:

(3.93) $$\tfrac{1}{2}m\dot{r}^2 - \frac{GMm}{r} = 0$$

or

(3.94) $$\left(\frac{dr}{d\tau}\right)^2 = \frac{2GM}{r} \quad \Rightarrow \quad \frac{1}{c^2}\left(\frac{dr}{d\tau}\right)^2 = \frac{2GM}{c^2 r}$$

Inserting $r_S = \frac{2GM}{c}$, and solving with respect to $d\tau$ gives:

(3.95) $$cd\tau = \pm\sqrt{\frac{r}{r_S}}dr$$

The plus sign corresponds to an exploding star, or an outward bound spaceship, while the minus sign corresponds to a collapsing star or an inward-bound spaceship. (3.95) is integrated to

(3.96) $$\tau(r) = \tau_0 - \frac{2r_S}{3c}\left[\left(\frac{r}{r_S}\right)^{\frac{3}{2}} - \left(\frac{r_0}{r_S}\right)^{\frac{3}{2}}\right]$$

where τ_0 is the time, where the probe is at some reference point r_0.

Thus the proper time τ is perfectly smooth over the Schwarzschild surface. The time for the star to collapse from $r_0 = r_S$ to the singularity at $r = 0$ is found by putting this into (3.96), giving:

$$\Delta\tau = \tau(0) - \tau(r_S) = \ \tau_0 - \frac{2r_S}{3c}\left[\left(\frac{0}{r_S}\right)^{\frac{3}{2}} - \left(\frac{r_0}{r_S}\right)^{\frac{3}{2}}\right] - (\tau_0 - \frac{2r_S}{3c}\left[\left(\frac{r_S}{r_S}\right)^{\frac{3}{2}} - \left(\frac{r_0}{r_S}\right)^{\frac{3}{2}}\right]) = \frac{2r_S}{3c}$$

Which is of order 10^{-4} s for a star with mass 10 times the solar mass, where r_S is about 30 *km.*

3.5.4 The Schwarzschild coordinate time

While the time measurement done by an observer travelling across the Schwarzschild surface is perfectly smooth, this is not the case for the time *t,* measured by an observer far away.

From (3.43) we have $g_{00} = -1 + \frac{r_S}{r}$ and $d\tau = \sqrt{-g_{00}}dt \quad \Rightarrow \quad d\tau = \sqrt{1 - \frac{r_S}{r}}dt$

From (3.92)

$$cd\tau = \pm\sqrt{\frac{r}{r_S}}dr$$

we then get:

(3.97)
$$cdt = -\sqrt{\frac{r}{r_S}}\frac{dr}{c\sqrt{1-\frac{r_S}{r}}},$$

which can be integrated by parts.

$$t = -\frac{1}{c}\int_{r_0}^{r}\sqrt{\frac{r}{r_S}}\frac{dr}{c\sqrt{1-\frac{r_S}{r}}} = k\int_{r_0}^{r}\frac{\sqrt{r}dr}{\sqrt{1-\frac{r_S}{r}}}, \quad \text{where } k = -\frac{1}{c\sqrt{r_S}}$$

$$t = k\int_{r_0}^{r}\frac{\sqrt{r^2}dr}{\sqrt{r-r_S}} = k\int_{r_0}^{r}\frac{rdr}{\sqrt{r-r_S}} = 2\int_{r_0}^{r} rd\sqrt{r-r_S}$$

$$t - t_0 = 2k\left(r\sqrt{r-r_S} - \int_{\infty}^{r}\sqrt{r-r_S}\,dr\right) = 2k\left[r\sqrt{r-r_S} - \tfrac{2}{3}(r-r_S)^{\frac{3}{2}}\right]_{r_0}^{r}$$

(3.98)
$$t - t_0 = -\frac{2}{c\sqrt{r_S}}\left(r\sqrt{r-r_S} - \tfrac{2}{3}(r-r_S)^{\frac{3}{2}} - r\sqrt{r_0-r_S} + \tfrac{2}{3}(r_0-r_S)^{\frac{3}{2}}\right)$$

The result however, can not be evaluated for $r \le r_S$ because we hit the pole at r_S in $\frac{1}{c}\left(1-\frac{r_S}{r}\right)^{-\frac{1}{2}}$ of the integrand. Looking at (3.97) it is however clear that *dt* diverges, when $r \to r_S$, and we can therefore conclude, as seen from an observer far away, it will take a probe an infinite time to reach the Schwarzschild surface.

There is however another way of viewing the event horizon and that is by studying the gravitational red-shift. We shall see that within the Schwarzschild surface the gravitational red-shift becomes infinite.
The relation between the coordinate time *dt* and the proper time $d\tau$, is

(3.99)
$$d\tau = \sqrt{g_{00}}dt = \left(1-\frac{r_S}{r}\right)^{\frac{1}{2}}dt \qquad \Rightarrow \qquad dt = \left(1-\frac{r_S}{r}\right)^{-\frac{1}{2}}d\tau$$

Which shows, that the coordinate time interval becomes infinite when $r \to r_S$.

3.5.5 Infinite gravitational red-shift

In terms of emitted (em) and recorded (rec) frequencies: we have earlier seen

(3.100)
$$\frac{\omega_{rec}}{\omega_{em}} = \sqrt{\frac{(g_{00})_{em}}{(g_{00})_{rec}}} = \sqrt{\frac{1-r_S/r_{em}}{1-r_S/r_{rec}}}$$

The received frequency approaches zero, when $r_{em} \to r_S$ so it will take an infinite time to receive the next wave crest. Thus no transmission is possible from within a Black Hole.

3.5.6 Orbital motion around a Black Hole

The equation for the potential that was discussed in (3.62)

$$\Phi_{eff}(r) = -\frac{GM}{r} + \frac{L^2}{2m^2r^2} - \frac{r_S L^2}{2m^2r^3},$$

This equation can also be used to study the motion of a massive body around a Black Hole.

The first term is the Newtonian potential, the second term is the classical centrifugal barrier, where L is the angular momentum of the body with mass m, with respect to the centre of the source, and the last term is the General Relativity contribution.
This term is extremely small, as we saw in the calculation of the precession of the planet Mercury. But it becomes important, when the radial distance is comparable to the Schwarzschild radius r_S, as it is the case of neutron stars or black holes.
We can find the extremes of the potential by setting the derivative of the potential to zero.

(3.101) $$\frac{d\Phi_{eff}(r)}{dr} = 0 \quad \Leftrightarrow \quad \frac{GM}{r^2} - \frac{L^2}{m^2r^3} + \frac{3r_S L^2}{2m^2r^4} = 0$$

Multiplying by $m^2\, r^4$ gives:

(3.102) $$GMm^2r^2 - rL^2 + \frac{3}{2}r_S L^2 = 0$$

Which is an ordinary second order algebraic equation, with the discriminant d.

$$d = r^2L^4 - 6\,GMm^2r^2r_S L^2 \qquad \Rightarrow \quad r = \frac{rL^2 \pm \sqrt{r^2L^4 - 6\,GMm^2r^2r_S L^2}}{2GMm^2r^2},$$

or

(3.103) $$r = \frac{L^2}{2GMm^2}\left(1 \pm \sqrt{1 - \frac{6\,GMm^2}{L^2}r_S}\right)$$

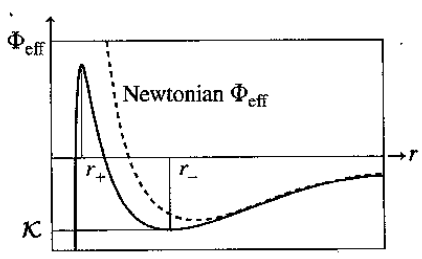


Fig. 6.12 Schwarzschild vs. Newtonian effective potential.

We note the difference between the effective potential in the Newtonian limit ($r_S = 0$) where the centrifugal barrier always dominates: $\Phi_{eff}(r) \to \infty \quad for \quad r \to 0$
In this case there is no r_+ solution. A particle cannot fall into to the $r = 0$ centre, as long as $L \neq 0$.

In the relativistic case however for very small r the r_S term becomes dominant, and

$$\Phi_{eff}(r) \to -\infty \quad for \quad r \to 0$$

When $E \geq m\Phi_{eff}(r_+)$ a particle can plunge into the gravity centre even if $L \neq 0$.
If $E = m\Phi_{eff}(r_-)$, it corresponds to a stable orbit at $r = r_-$, as in the Newtonian case.
However the radius of the orbit cannot be arbitrary small. As is seen from (3.103)

$$1 - \frac{6\,GMm^2}{L^2} r_S = 1 \ ,$$

we have that the smallest radius r_{min} at

$$r_{\min} = \frac{L^2}{2GMm^2} = 3r_S$$

Because of the extraordinary feature of strongly warped space-time near the Schwarzschild surface, it took a long time for the physical community to accept the reality of the Black Hole prediction, presented by the Schwarzschild solution.

One reason was the evident, since black holes do not emit light they cannot be observed, and physics has had a long tradition of rejection theories containing unobservable objects.

However no other solution was applicable, when considering the collapse of a very massive star, running out of nuclear fuel. And by the end of the 1950'ties, Black Holes was generally accepted by the physical society, as the evidence from observations became manifest.

However, although black holes cannot be observed directly, there are several indirect indications of there existence. One is a double star having an invisible partner.

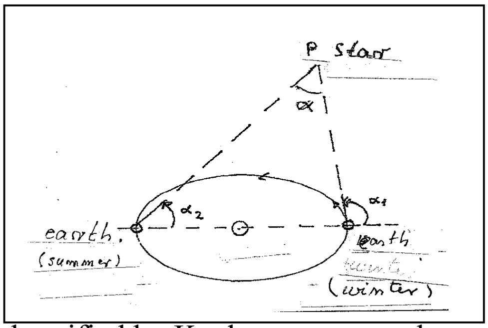
The other is the intense gamma rays produced from an invisible "star". Gamma rays are produced when a Black Hole sucks gasses from its exterior. Near the Schwarzschild surface they become so energetic that they produce gamma rays, showing a spectrum in accordance with what one might expect if the source is a Black Hole.

4. General Relativity and Cosmology

- The framework required to study the whole universe as a physical system is General Relativity.
- The universe, when observed at a distance scales over 100 *Mpc* is homogeneous and isotropic.
- Hubble's discovery that the universe is expanding suggests strongly that it had a beginning, when all objects were concentrated in a point of infinite density. The estimate of the age of the universe taken from the astrophysicists observed data is at least 12.5 Giga-years.
- There is a considerable amount of evidence showing that most of the mass in the universe does not shine. The mass density in the universe, including both luminous and dark matter, has been estimated to around one third of the "critical density".
- The space-time satisfying the cosmological principle is described by the Robertson–Walker metric in co-moving coordinates, called the cosmic rest frame.
- In an expanding universe with a space that may be curved, any treatment of length and time must be carried out with care. We investigate the relations between cosmic red-shift and proper distances as well as luminosity distances.

4.1 Astronomical observations and astronomical distances

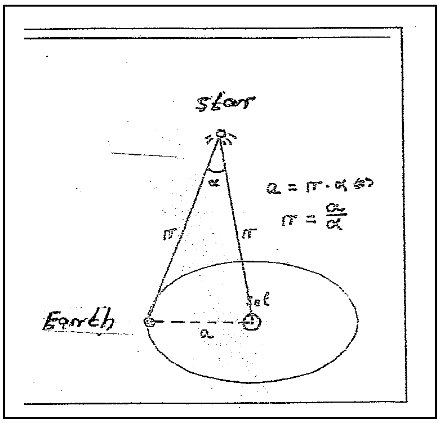
Before we dive into General Relativity concepts, we shall look a bit into the astronomical part of cosmology.
From ancient times people have observed that there were two kinds of “stars”, the fixed stars and the wandering stars. The fixed stars were seen at the same fixed positions in the sky, but following the rotation of the earth. (“Were fixed on the glass sky spheres that rotated around the earth fixed in space”).
The wandering stars followed the motion of the fixed stars, but were superimposed by an independent motion that differed from one wandering star to another.
Since Copernicus and Galileo in the late 1400 abandoned the geocentric view of the world in favour of the heliocentric view, we have known that the “fixed stars” are stars and galaxies, whereas the wandering stars are the planets.
However one of the most resistant arguments against the heliocentric view of the World, was that the fixed stars in fact seem to be fixed, even in a period over 200 years.
The simple argument was, if the earth rotated around the sun, the line of sight to the fixed stars would be different from summer to winter. fig.(4.1)



The change of the line of sight to a star, as a consequence of the earth’s orbital motion around the sun, is called the **parallax** of a star.

In figure (4.1) the parallax α is the upper angle in the triangle. The figure shows that $\alpha = \alpha_1 - \alpha_2$.
Following the establishment of the Heliocentric view of the world, established by Copernicus and Galileo and verified by Kepler, encouraged several astronomers during the eighteenth and nineteenth century to pursue the measurement of the parallax for the nearest stars – but without success.
This failure could have two reasons. Either the heliocentric view of the world was wrong or the stars were much father away in space than hitherto assumed.

Not until 1838 the German astronomer (and constructor of precision binoculars) succeeded to measure a parallax for the star Sirius B, which is situated 4.26 *ly* (light years) away.



The measurement of distance by parallax has lead to a definition of a astronomical distances called 1 *parsec*, which is still heavily used in astronomy. Fig (4.2).
The arc *a*, corresponding to the angle α (in radians) is equal to the radius of the arc *r* times α

$$a = \alpha \cdot r$$

If one measures the parallax α, and the arc *a* is well known, the distance *r* can then be calculated.

1 *parsec* is then defined as the distance, under which the half great axis (*a*) of the Earth is seen under an angle of 1 arc sec = 1/3600 ^{0}.

(4.3) $$1\, arc\sec = \frac{1}{3600}\frac{\pi}{180} = 4.848\,10^{-6} \quad radian$$

If we put *a = 149.6 mill km,* we find.

(4.4) $$1 \;\; par\sec = 1\; pc = \frac{a}{\alpha} = \frac{149.6\,10^{9}}{4.848\,10^{-6}} = 3.086\,10^{16}\; m$$

This can be compared to

(4.5) 1 *ly* (light year) = *c· 1* year = 3.0 10^8 365·24·3600 =9.46 10^{15} *m*.
so
(4.6) $1\; pc = 3.26\; ly$

The nearest star Alfa-Centauri has a parallax corresponding to a distance of 1.3 *parsec*.

For stars, which are located more than 100 *ly* away, it is no longer possible to determine the distance by measuring their parallax. (There exist galaxies that are billions of light years away)

For distant objects, it is necessary to use alternative methods. **The Cepheid method** has been used since the 1920'ties and after 1940 it has been supplemented by the so called **photometric method.**

Some stars have a pulsating brightness and the frequency is dependent of the luminosity of the star. These stars are called Cepheid's. (Named after the first observed "Cepheid").
For Cepheid's, where the distance can be measured by parallax, one has discovered the notable fact that the pulsating frequency has an almost on to one correspondence with the luminosity of the Cepheid.

Equipped with the knowledge of the relation between luminosity and pulsating frequency, one has been able to determine the distance to Cepheid's, where the parallax method no longer is available.

From measuring the frequency, one can namely determine the absolute luminosity L_{abs}, by measuring the apparent luminosity L_{ob}. The distance may then be found by the quadratic-distance-relation, i.e. the flux through a unit area of a sphere emitted from a source at its centre decreases with r^{-2}. (Since without absorption the total flux is constant, and the area of a sphere is $4\pi r^2$).
We then have the relation:

(4.7) $$L_{observed} = \frac{L_{absolute}}{4\pi r^2}$$

From which the distance *r* to the star can be calculated.

This was exactly the method used by Edwin Hubble when he in the 1920'ties determined corresponding values of the distances to and velocities of about 18 galaxies.
The velocities of the galaxies were then measured by the red-shift in their radiation.

The procedure of the **photometric method** is to establish a catalog of stars at known distances together with a detailed registration of their luminous spectrum. Then to apply this to distant stars which match the spectrum, measure the apparent luminosity, and use the quadratic-distance-relation.

If two stars at distances r_1 and r_2 have the same absolute luminosity L_{abs} , the apparent luminosities L_1 and L_2 are

(4.8) $$L_1 = \frac{L_{abs}}{4\pi r_1^2} \quad and \quad L_2 = \frac{L_{abs}}{4\pi r_2^2}$$

And they will satisfy the equation:

(4.9) $$L_1 4\pi r_1^2 = L_2 4\pi r_2^2 \quad \Rightarrow \quad r_2 = r_1\sqrt{\frac{L_1}{L_2}}$$

4.2 The universe is not infinite and static

Until the mid 1950'ties it was the most common view that the outer space "always" had existed, that it was infinite, homogenous and static in a galactic scale. (*The steady state model)*.
Also, before Hubble's observations, there existed no observations to the contrary.
There were however, two serious objections to *the steady state model.*

Firstly, as Einstein had pointed out.
If Newton's law of gravitation had unrestricted validity, also at cosmological distances, then the universe would necessarily over time begin to contract due to gravitation.

And this would tend to be in conflict by the notion of a static universe.

For this reason Einstein invented the so called **cosmological constant** in his theory of General Relativity, which implied a correction to the law of gravitation, but inferred a repulsive force to compensate for the gravitational attraction on a galactic scale.

But, as it has been the case several times in the history of theoretical physics, such an invention is highly problematic, if it is done only to explain discrepancies between experimental observations and existing theory. Especially when the new "theoretically concept" has had no other manifestations in earlier observations or experiments.

Einstein later denounced it, as "The greatest mistake in his life". (Actually there have been others in the same weight class).
Nevertheless the *cosmological constant* is still a "physical ghost" in scientific literature together with "dark matter", "dark energy" and other non observable objects.

But already in the 1930'ties some were on the track of a flagrant discrepancy between the rotational speed and the masses of the galaxies.
It was a indisputable fact that the rotational speed of the galaxies were to big to be explained by the observable amount of matter.
From this comes the notion of "dark matter". Dark matter has however become an accepted ingredient of the universe, even if it never has been subject to a just reasonably good theoretical explanation, whereas to its nature or its origin.

There is another and more specific reason why the universe cannot be infinite and homogenous (meaning an all over constant density of galaxies).

It is rather straightforward to show, that this notion implies that the heaven would be predominantly white and not black.

Let us therefore assume that the density, that is, the number of stars corresponding to say 1 cubic *parsec* is n_s. And let us assume that their average luminosity is L_{av} . The luminosity of the sun is: $L_{sun} = 3.827\ 10^{26}\ W$.

According to the quadratic-distance-relation the intensity in the distance r from a star will be

$$I_{av} = \frac{L_{av}}{4\pi r r^2}$$

We now consider an arbitrary point P in space. At this point we shall calculate the intensity due to the infinity of stars in space. We do this by slicing the space around P up in spherical shells with a thickness dr and radius r. The number of stars in each shell is dN, *and* the contribution from each shell is $dI(r)$. Finally we find the intensity at P by summing the contributions from all the shells by integrating.
The volume of a sphere shell with radius r and thickness dr, is its surface $4\pi r^2$ times dr.

(4.10) $$dI = I(r)dN = I(r)n_s 4\pi r^2 dr = \frac{L_{av}}{4\pi r^2} n_s 4\pi r^2 dr = L_{av} n_s dr$$

Since L_{av} and n_s were assumed to be the same throughout the universe, it is clear the an integration necessarily must give infinity.

(4.11) $$I(P) = \int_{r_{\min}}^{\infty} L_{av} n_s dr = \infty$$

So according to *the steady state model*, the sky ought to be white and not black.

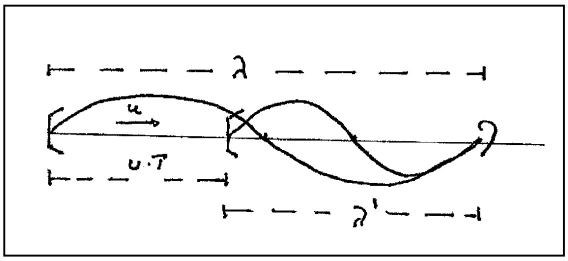
Objections to this argument, e.g. that light is absorbed in galactic nebulae, may lead to a more complex calculation, but it does not alter the conclusion.

So we may conclude that either the universe is not finite or the density of stars decrease with the distance from the earth. The last speculation has in fact much in common with the geocentric view of the world, and can hardly be substantiated by other than metaphysical arguments.

4.3 Red-shift

The red-shift is the designation of the phenomena that the wavelength (and frequency) of light is shifted, when it is emitted from a source, moving away from an observer.
The corresponding situation, when a source is moving towards an observer is called blue-shift, and the generic designation for the phenomena is the **Doppler shift**.

Although the galaxies move at very high speed, they are not relativistic and one may safely apply a non relativistic approach.
The Doppler shift is well known from sound waves and the figure, which is shown below actually refer to sound waves. Fig (4.12)



There are several ways to derive a formula for the Doppler shift, but it is most easily done by looking at the figure to the left.
The source emits a periodic signal travelling at a speed *v,* and the source itself is moving towards the observer with a speed *u.* The emitted wavelength is λ and the observed wavelength is λ′.
If *T* is the period of the signal, it is clear from the figure that the following relation between λ and λ′ must hold:

(4.13) $\lambda = \lambda' + uT$ but since, $T = \dfrac{\lambda}{v}$, we find: $\lambda' = \lambda(1 - \dfrac{u}{v})$

This means that the wavelength is shortened, as seen from the observer, and it is called blue-shift, because it means a displacement of the spectral lines towards the blue end of the spectrum.
If the source moves away from the spectator, all we have to do is to replace *u* by *–u*. And if we are talking about light and not sound waves, we replace *u* by *v*, and *v* by the speed of light *c*.

$$\lambda' = \lambda(1 + \frac{v}{c}) \qquad (4.14)$$

Since here the wavelength is widened, it is called **red-shift**
The relativistic formula when the source moves away from the observer looks a bit different.
(See e.g. "The Special Theory of Relativity", on the current Web-side)

$$\lambda' = \lambda\sqrt{\frac{c+v}{c-v}} \quad or \quad \lambda' = \lambda\sqrt{\frac{1+v/c}{1-v/c}} \qquad (4.15)$$

A Taylor expansion of (4.15) after *v/c* to the first order leads to the non relativistic formula (4.14).
If λ' is measured and compared to the actual λ, the velocity *v* of the source can be calculated from

$$v = \frac{\left(\frac{\lambda'}{\lambda}\right)^2 - 1}{\left(\frac{\lambda'}{\lambda}\right)^2 + 1}c \qquad (4.16)$$

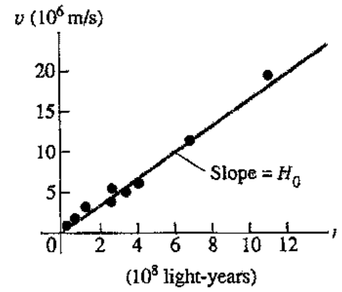
Although the galaxies move fast some 100 *km/s*, it is still considered a non relativistic motion, so in most cases the non relativistic formula (4.14) is used, but written a bit different.

$$\lambda' = \lambda(1 + \frac{v}{c}) \Leftrightarrow \lambda' - \lambda = \lambda\frac{v}{c} \Leftrightarrow \frac{\Delta\lambda}{\lambda} = \frac{v}{c} \ (= z) \qquad (4.17)$$

The relative change in the wavelength is equal to the fraction of the speed of light, at which the source is moving. Using the relation $\lambda = \dfrac{2\pi v}{\omega}$ we get almost the same expression for the frequencies as for the wavelengths, apart from a minus sign.

(4.18) $$\frac{\Delta\omega}{\omega} = -\frac{v}{c} \ .$$

4.4 The universe is expanding



46—15 Graph of recession speed as a function of distance for several galaxies. The graph is nearly a straight line, illustrating Hubble's law. The slope of the line is Hubble's constant, H_0.

A comparison between the spectre from galaxies with the familiar spectre from the atoms show that emitted spectre from distant galaxies are the same as we know from the 92 elements on the earth. But also that they are all red-shifted, indicating that the galaxies are moving away from us.

But likewise important, this applies only to distant galaxies, but not to the stars in our own milky way, where some are red-shifted and some are blue-shifted.

The American astronomer Edwin Hubble conducted in the 1920'ties a comparison between the speed and the distances for a dozen of galaxies. He calculated their velocity from the measured red-shift, and the distances from the Cepheid method. As mentioned above, he concluded that the light from all the galaxies were red-shifted, indicating that they were moving away from us, and furthermore that there seemed to be a simple proportionality between their velocities and their distances. See the figure.

In Hubble's original graphical publication however, the proportionality was much less obvious, since many of the measured distances were off with about 25%.
Later and more accurate measurements have confirmed Hubble's famous conjecture which is now known as Hubble's law. If r is the distance to the galaxy and v is the outward speed, Hubble's law is usually written.

(4.19) $$v = H_0 \cdot r$$

H_0 is designated Hubble's constant. The accepted value is today $H_0 = 33\ 10^{-18}\ s^{-1}$ with an uncertainty of about 20%. The zero suffix refers to the present period of the universe.
Often Hubble's law is written with the unit *km/s* per *light-year.*

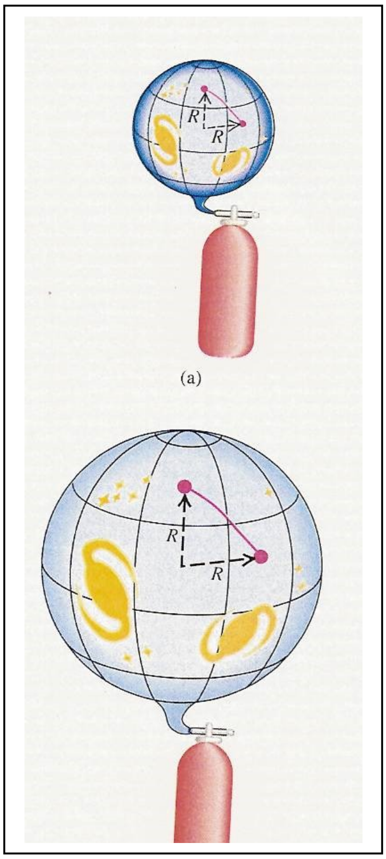
(4.20) $$H_0 = 1.9\ 10^{-5}\ (km/s)/ly$$

The initial value, which Hubble obtained was only about a fifth of the present value, due to the very delicate job of determine the distances to remote galaxies.

For very remote objects a direct determination of distances are impossible, but in those cases Hubble's law is applied to find the distances from measuring the red-shift, using the relation $\Delta\lambda / \lambda = v / c$ to determine the speed v, and finally using Hubble's law to determine the distance r.

4.5 The cosmological principle

Since all galaxies are moving away from us one could from a common line of thought conclude that the earth is the centre of the universe, but such a conjecture can only be substantiated metaphysically. From physical rationality it is incomprehensible to conceive the earth as being the centre of the universe. Fig (4.21).



However with the appearance of Einstein's General Theory of Relativity, the possibility arose of a curved finite space-time, which might explain why all galaxies are receding from us.
Since the beginning of the1960'ties the vast majority of the scientific world have agreed on the validity of the **cosmological principle**, which can be formulated as follows:

The universe is homogenous and isotropic.
On the largest scale of *Mpc* (mega parsec) this has not been contradicted by observations until now.
Another evidence is the1960'ties discovery of the Cosmic Microwave Background (CMB), which are the remnants of the radiation created in the big bang.

Homogenous: Means that on the largest scale the distribution of galaxies are the same overall in the universe.

Isotropic: Means that the universe looks the same from every point in the universe. The universe is not confined, but it is neither infinite.
It is however not possible to comprehend this model within the framework of Euclidian geometry, in which mathematical objects are either confined or infinite. On the contrary, the fact that all galaxies tend to move away from each other indicates, that they are not moving away in space, but it is space itself that expands.

To comprehend this mathematically we have to resort to non Euclidian geometries i.e. Einstein's General Relativity. Both in General Relativity and in Special Relativity time is treated on equal footing with the spatial coordinates, which certainly does not make the understanding of either more transparent from a common view of the physical world.

In an attempt to visualize the curved expanding space, it is usual to use a two dimensional curved sphere. Often a two dimensional analogy is illustrated as an inflating balloon, inhabited with "2D spices", which cannot move or see outside the two dimensions. But we can, because the 2D sphere is embedded in a 3D space.
The space is curved, so if you walk along a "straight line" (a great circle), you will eventually reach the position, where you started. In principle, if you could look straight ahead "infinitely", you should be able to see your back, since the line of sight is along a geodesic.

When the balloon is inflated, it is obvious that no matter where you are placed on the sphere the "galaxies" will move away from each other. Furthermore if the radius increases with time, any arc b of a great circle will increase at the same rate, this because $b = \alpha \cdot r$, where α is the constant radian of the arc.
If $r = kt$, the speed of separation between to galaxies is: $v=b/t = \alpha \cdot r/(r/k)= \alpha k$, is seen to be directly proportional to their angular separation α. The father away from each other, the faster they move away from each other.

Although the expansion of space-time is beyond common human imagination, this 2D analogy might give an explanation to the notion, that the velocity in which the galaxies are receding from each other is directly proportional to their separation (Hubble's law) .

4.6 The Big Bang

While the universe is expanding, one could contemplate to "run the movie backwards", and then one can hardly avoid the conclusion that the universe must originate from a very small size, rather a point, or more precisely, a singularity of infinite density.

The Big Bang is a short formulation of the conjecture that the universe was created at $t = 0,$ with a gigantic outlet of energy caused by an inflation of vacuum.

If one supposes that the galaxies, which were created 30 million years later, "have moved" (actually they do not move in space, rather the space is expanding) with the same velocity v (and has not been substantially slowed down, as a consequence of mutual gravitational attraction), then they will "now" have reached a distance r, given by:

(4.22)
$$r = v \cdot t_{universe}$$

Where $t_{universe}$ is the age of the universe. If one compares this with Hubble's law:

$$v = H_0 r \quad \Leftrightarrow \quad r = \frac{v}{H_0},$$

Then it follows.

(4.23)
$$t_{universe} = \frac{1}{H_0}$$

It is somewhat surprising that the age of the universe has this simple relation to Hubble's constant. Presently we shall demonstrate that Hubble's law is in fact a consequence of General Relativity.

Inserting the present value for Hubble's constant $H_0 = 2.33\ 10^{-18}\ s^{-1}$ we find.

(4.24)
$$t_{universe} = \frac{1}{H_0} = 1.36\,10^{10}\ yr = 13.6\ \ billion\ \ years$$

4.7 Critical density

Until the 1950'ties (and in spite of General Relativity and Hubble's law), the universe was still considered static and Euclidian flat.
But in an 2D analogy to General Relativity it is depicted as an infinite plane covered by valleys, corresponding to stars and galaxies supplemented by some narrow and deep holes representing white dwarfs and neutron stars, and finally some infinite deep black holes, which were a point of no return.

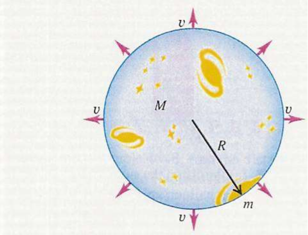
Even if stellar objects moved relative to each other and new stars were born and deceased, the universe was infinite and had always existed. The most common opinion was that the universe was static and flat and neither expanded or the opposite.

The obvious problem with a static universe, (as mentioned previously), was that a static universe eventually would draw together as a consequence of gravitational attraction and inevitably collapse in a "Big Crunch".
Observations indicating that the galaxies approach each other do not exist. After Hubble had shown that the galaxies on the contrary moved away from each other, Einstein invented the so called cosmological constant, (a correction to Newton's law of gravitation), to compensate for the gravitational attraction in a cosmic scale, and in this way provide an explanation for the static universe.

We are however left with the question of the curvature of the universe, which is rather closely connected with the average density of the universe.
With a positive curvature (closed universe) the expansion will eventually stop, and the universe will begin a contraction, ending with a Big Crunch.

A negative curvature means that the universe will continue to expand and with increased speed.

A flat universe means that the curvature of the universe is zero. The universe is Euclidian for cosmological distances, and the expansion will continue ad infinite until it stops.
The relevant physical concept to decide which of these pictures is correct is called the **critical density**.
It is possible to make a (non relativistic) calculation resulting in finding the average density in the universe from Newton's gravitational law, supplied by Hubble's law.
It is done in the same manner in which you find the escape velocity from the earth.



46–17 The gravitational force on our galaxy (mass m) is the force exerted by the mass M within the sphere with radius R. Because of the spherical symmetry, the mass outside this sphere has no effect.

If there is a homogenous distribution of mass in a cosmic scale, then one may calculate the force on a galaxy on the surface of a sphere with radius R, from the mass inside the sphere.
The sphere is supposed to have an extension far greater than the size of the galaxy. One may show, (with a homogenous distribution of mass within the sphere), that the force on the surface depends only upon the mass within the sphere, and as if the whole mass was placed in the centre. (Gauss' law).The force must (through symmetry considerations) be radial directed.
If the density is denoted ρ_c, we obtain for the mass M in the sphere.

(4.26) $$M = \rho_c V = \rho_c \frac{4}{3}\pi R^3$$

The escape velocity v, that is, the velocity a galaxy must have at the surface of the sphere to escape its gravitational field, can be calculated from conservation of energy.
At infinity both the kinetic and potential energy is zero. Putting the kinetic and potential energy at the surface to zero the escape velocity can be calculated.
It should be clear that the escape velocity correspond to the critical density, namely a universe, where the expansion only will stop at infinite. ($v = 0$)

(4.27) $$½mv^2 - G\frac{mM}{R} = 0 \quad \Leftrightarrow \quad ½v^2 - G\frac{M}{R} = 0 \quad \Leftrightarrow \quad M = \frac{v^2 R}{2G}$$

Now inserting (4.26) and Hubble's law $v = H_0 R$, we get:

(4.28) $$\rho_c \frac{4}{3}\pi R^3 = \frac{(H_0 R)^2 R}{2G} \quad \Rightarrow \quad \rho_c = \frac{3H_0^{\ 2}}{8\pi G}$$

Using known values for the constants, we achieve the following value for the critical density.

(4.29) $$\rho_c = 7.2\ 10^{-27}\ kg/m^3$$

The mass of the hydrogen atom is close to 1 u = 1.67 10^{-27} kg, so the critical density correspond to about four hydrogen atoms per cubic meter.
Attempts to "weigh" the universe, including luminous objects, interstellar gasses, neutrinos and black holes gives an estimate of only 1/100 of ρ_c.

Many theorists however, have the strong conviction that the universe is closed and flat, and that has risen many speculations on the nature of "**dark matter**".

Dark matter plays a role in several other connections. If the rotation velocities of the stars in a galaxy are measured (as well as it seem possible), and when they are compared to measurements of the mass of the galaxy (as well as it seem possible) within the orbit of the star, there is a discrepancy which cannot be subscribed to uncertainty in measurements. This is to be understood as the estimated mass is far to small to keep the stars in rotation when compared to their measured angular velocity.
However the notion of "dark matter" is for the present just as speculative, as the notion of the cosmological constant.

4.8 Luminous and dark matter

During the 90'ties most astrophysicists had more or less been convinced of the existence of the "dark" matter in the universe. "Dark" refers to the fact that it does not shine and it does not emit electromagnetic radiation. Neither as Planck radiation caused by high temperatures (as is the case for the stars), nor as secondary radiation, as is the case of interstellar gasses.
On the contrary observations of the tangential velocity of the galaxy halo strongly indicate that only a fraction of the matter in the universe consist of "ordinary" matter, the major part is dark matter although its constituents and origin is unknown.

If the observations are reliable (and they are for the present), the dark matter do not emit electromagnetic radiation, so either (if it consists of ordinary matter) must be neutral particles or have a low temperature. Obvious candidates for the dark matter could therefore be gas-planets like the outer planets in our solar system Jupiter, Neptune, Uranus or it might be neutron stars or Black Holes.
However for various theoretical reasons these candidates for the dark matter have been ruled out. Also massive neutrinos have been suggested as a source of dark matter but it has later been discarded.

Another more vivacious presumption is the existence of an unknown type of elementary particles. This suggestion has not created much happiness among elementary particle physicists, especially after the Standard Model in the late 90'ties obtained a consistency by the discovery of the W and Z bosens, and notably the τ bosons, which more or less closed the model.

But personally I cannot come up with a quick answer why the dark matter could not just be accumulations of proto-stars like Jupiter, with masses less than 0.07 M_{sun} (which is the lower limit for initiating the fusion of hydrogen and forming a luminous star).

That there must be more matter having gravitational properties in the galaxies then the luminous matter, has been manifest since the observations of the rotational frequency of the galaxies, made by Fritz Zwicky already in the 1930'ties.
All measurements of the distribution of mass in the universe are inflicted with uncertainty, but this does not change the fact that the luminous mass cannot account for the mass necessary to explain the rotational motion of the galaxies.
Previously it has been proposed that the galaxies might have a gigantic Black Hole (3 -4 mill. sun masses) at its centre, but this explanation does not comply with the observations. For an orbital motion in a central field, Kepler's 3. law $a^3/T^2 = const.$ applies, where $2a$ is the great axe in the elliptic orbit, and T is the period.

For such motions it is the central gravitational force which delivers the centripetal force. If the central mass of the galaxy is M, and the rotating star has mass m, then we have for a circular orbit.

(4.30)
$$F_c = F_G \quad \Leftrightarrow$$
$$m\omega^2 r = G\frac{mM}{r^2} \quad \Leftrightarrow$$
$$\left(\frac{2\pi}{T}\right)^2 r = G\frac{M}{r^2} \quad \Leftrightarrow$$
$$\frac{r^3}{T^2} = \frac{GM}{4\pi^2} = const$$

Using $T = 2\pi/\omega$ (where ω is the angular velocity) it follows:

(4.31)
$$r^3\omega^2 = k \quad or \quad \omega = \frac{k}{r^{\frac{3}{2}}}$$

If the whole galaxy mass is placed near its centre, the angular velocity of the stars will decrease with the distance r to exponent of 3/2, but neither this, is in accordance with observations.

Fig (4.32) The Andromeda galaxy

That the angular velocity of the stars in a galaxy decreases with the distance from the centre of the galaxy can be understood, even without any knowledge of physics or astronomy.

Take for example a look at a photo from the Andromeda galaxy, where the backward turned spiral arms indicate decreasing angular velocity with increasing radius.
If we, on the other hand assume that the mass of the galaxy is evenly distributed in a disc, it is only the mass which lies within the stars orbit that contributes to the gravitational force on the star.

If M_0 is the total mass of the disc, r is the radius in the orbit of the star, and r_0 is the radius in the galactic disc, then the mass within the disc is given by.

(4.33) $$M(r) = \left(\frac{r}{r_0}\right)^2 M_0$$

The mass density of the disc is namely $\rho = M_0/\pi r_0^{\ 2}$, so $M(r) = \rho\pi r^2$, which gives the expression above. Making the same derivation of Kepler's 3. law as above, we find:

(4.34) $$F_c = F_G \quad \Leftrightarrow \quad m\omega^2 r = G\frac{mM(r)}{r^2} \quad \Leftrightarrow$$

$$\omega^2 r = G\frac{(r/r_0)^2 M_0}{r^2} \quad \Rightarrow$$

$$\omega^2 r = \frac{GM_0}{r_0^2} = k \quad \Leftrightarrow$$

$$\omega = \frac{k_1}{\sqrt{r}}$$

This is in somewhat better accordance with the observations, but still the stars move with angular velocities corresponding to galactic masses which are some 60% larger than the observed.

The dark matter is still subject to speculation. It could actually be accounted for, by a large "cosmological constant", but this does not fit into numerous other observations.
A Black Hole having a mass of some million sun masses could also compensate for the mass deficiency, but lacks to explain the dependency of the angular velocities on their distances from the centre of the galaxy.

4.9 Quasars

In the mid sixties some stellar objects were observed having an extreme high red-shift $z > 0.5$, which means that they are receding with a speed comparable with the speed of light. They have been named Quasars. According to the Hubble equation, we have namely.

(4.35) $v = H_0 \cdot r$ with $H_0 = 1.9\ 10^{-5}$ *(km/s)/ ly (lightyears)*

Plugging in $v = 0.5c$, where c is $3.0\ 10^5$ *m/s* one finds:

(4.36) $$r = \frac{v}{H_0} = \frac{1{,}5 \cdot 10^5}{1{,}9 \cdot 10^{-5}}\, ly = 7{,}8\, 10^9 \quad lightyears$$

Quasars with a distance of 10 billion light years have been observed.
Quasars have another remarkable quality. They are visible! A quasar can emit more light than 1000 ordinary galaxies from an extension less than one light year. It is the supposition that the galaxies get their colossus amount of energy from matter falling into a Black Hole in the centre of the quasar. To supply the sufficient energy the Black Hole must have a mass at 100 to 1000 million times the mass of the sun.

The conjecture is that the enormous quantity of energy comes from a so called active core, which is also a central part of other stellar systems, in which some of the most energetic processes in the universe take place. Roughly 1% of the galaxies have an active core. By an active core means that there is emitted far more energy from its centre, than one should expect from the abundance of stars in the vicinity.
To explain the enormous outburst of energy from quasars, we can estimate that the mass that falls into the Black Hole at the centre at a yearly basis must be comparable to the mass of the sun. (The lifetime of the sun is about 9 billion years).
The quasars are extreme examples of galaxies having an active core. Our own galaxy The Milky Way is supposed to have a core, with a mass which is "only" three million times the mass of the sun.
While studying the quasars one should remember however, that we are looking perhaps 10 billion years back in time and that we view the universe as it appeared at that time. What it looks like to day is of course inaccessible.

4.10 Neutron stars and Black Holes

When a star has exhausted its ability to produce energy by fusion of lighter atoms to heavier, converting binding energy to radiation, starting from hydrogen to helium and further from helium to carbon until iron (which has the largest binding energy per nucleus) it is no longer possible for the star to resist the pressure from the outer layers of the star. Following a rapid contraction the star will eject the outer layers in space, leaving a most massive remnant of the star.

How massive it becomes depends on the size of the star. If the mass is less than 1.4 M_{Sun}, the star will become a so called *white dwarf* , which is a small and very hot star with the size of the earth. The density is typically $10^{10} – 10^{11}$ kg/m^3.

If the mass of the star is above 2 M_{Sun} , the star will end its life in a so called Supernova Explosion. The contraction is rapid and violent, releasing an enormous outburst of energy, causing the outer layers of the star to be thrown out into space.

The pressure in what is left over from the explosion is so high that the electrons are pressed back into the protons to form neutrons. The star is now a gigantic nucleus, consisting only of neutrons. The star has become a *neutron star*. The density is of the same magnitude as that of a nucleus. The order of 10^{18} kg/m^3, and its radius is only about 10 *km*.

If the mass is larger than 3 M_{Sun} , the life of the star will also end with a supernova explosion, but the contraction will continue past the neutron star state, onto a gravitational collapse, which we have described earlier as a singularity in space using the Schwarzschild solution to the Einstein equation of General Relativity.

Back in the 1960'ties the situation was this, that even if Black Holes and neutron stars were predicted theoretically no one had any supposition on how to make observations, since they were by nature invisible.
For this reason, even in the 1960'ties neutron stars were discarded by some members of the physical society for epistemologies reason, since the stronghold of the positivism by Ernst Mach was still manifest.
But that changed with the discovery of the "pulsars". A pulsar is an invisible object emitting signal at a frequency some *ms* in the radio band. The presumption was that the radio waves come from

the magnetic field of the neutron stars, emitting electromagnetic radiation, because of the extreme high rotational frequency. The radiated frequency is the same as the rotational frequency of the star.

The first pulsar was discovered at the centre of the Crab nebular, which is the remnant of a supernova explosion (only recorded in China) in 1054. This pulsar was found in 1967 by the research student Jocelyn Bell and her tutor Anthony Hewish, (of which only the latter received the Nobel price). Since then more than 1000 pulsars have been detected, and the emitted radiation suits well with the theoretical framework.

Until the 1970'ties several members of the physical society, had an epistemological problem with black holes for the simple reason that they could not be observed, and therefore could be no part of the physical science.
But since then, there have been several indirect testimonies of their existence. The main reason is that satellites have made it possible to obtain observation data outside the atmosphere of the earth. Mainly the measurement of X-rays, which otherwise would be absorbed when passing though the atmosphere.

When a Black Hole sucks interstellar matter or matter from a nearby star, the accelerations of ionized atoms and in particular the electrons, are so violent, that the emitted radiation are X-rays. Some galactic cores are called active, because they exhaust a disproportionately large amount of energy as radiation, including X-rays, in comparison to their extension. The presumption is that these galaxy cores are gigantic black holes with a mass of several million sun masses.

Another indirect evidence of a black hole is Cygnus X-1, a star which orbits around an invisible companion. As the mass of the companion must necessarily be larger than Cygnus X-1, it seems plausible to assume that we are talking about a Black Hole. Observation of X-rays presumably coming from the invisible companion tend to confirm this assumption.

A third kind of evidence are the gravitational lenses, which cause the same stellar object to be seen in different places. This comes about because the rays are bended when passing a very massive object.

Below is illustrated (schematically) the principle of a gravitational lens where the rays from a star in the line of sight are deflected when passing near to a massive Black Hole.
Further below is an authentic image of that phenomena, called Einstein's cross.
Apparently it shows a galaxy, with no less than 4 cores.
The more probable explanation is that the 4 cores are in reality light from a very luminous object (a quasar), lying behind a massive Black Hole in the line of sight.
Fig (4.37) Gravitational lens.

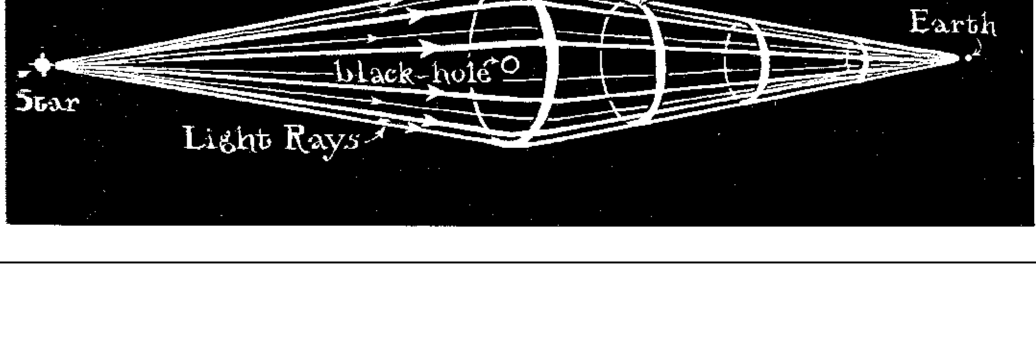
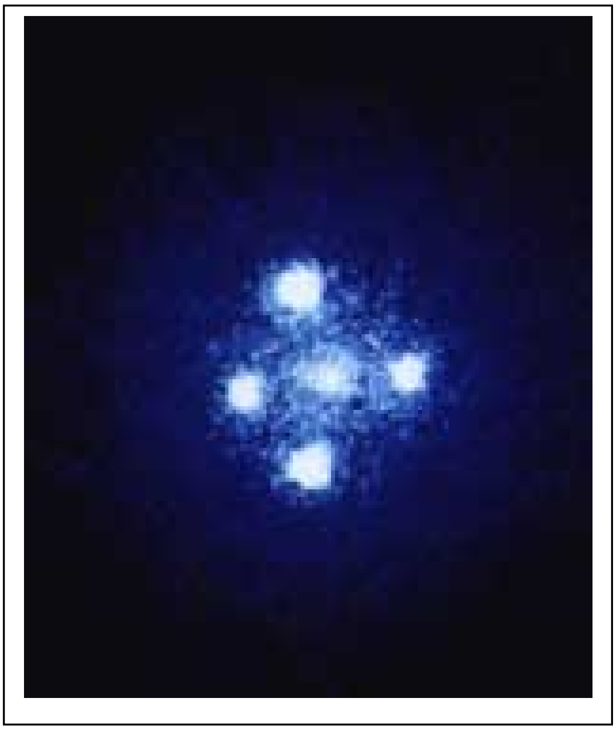


Fig (4.38) Einstein's cross.



4.11 The distribution of matter on a cosmic scale

Distances in our planetary system are sometimes measured in astronomical units.
1 *AU* is the mean distance from the earth to the sun. 1 *AU* = 150 million kilometres = $1.50\ 10^{11}$ *m*.
The most common units for distances in the universe are 1 *parsec* =1 *pc* = $3.086\ 10^{16}$ *m* , and
1 *light year* = 1 *ly* = $9.46\ 10^{15}$ *m*. 1 *pc* = 3.26 *ly*.

The nearest star Alpha Centauri is found 1.3 *pc* = 4.22 *ly* away.
Our own Milky Way is a typical spiral galaxy that holds about 10^{11} stars in a disc with a diameter of 30 *Kpc,* and a disc thickness of about 2 *kpc*.
Galaxies in turn arrange themselves in clusters, which are arranged in even larger super-clusters.
Our galaxy is part of a small group called the Local Group comprised of 30 galaxies in a volume measuring 1 *Mpc* across, and where the distance to the (mythical) Andromeda Galaxy is 0.7 *Mpc*
Galactic distances are thus measured in *Mpc* or even in millions of *Mpc.*

We have previously calculated the *critical density* of the universe. This is the density that would result in a flat universe and from a Newtonian calculation turns out to be:

(4.39) $$\rho_c = \frac{3H_0^{\ 2}}{8\pi G} \qquad \rho_c = 9.7\ 10^{-27}\ kg/m^3$$

As mentioned above there is a considerably amount of evidence that most of the matter in the universe does not shine. This is referred to as *dark matter.* To do an estimate of the density of the universe, we therefore split the matter (M) in the universe up in luminous matter (LM) and dark matter (DM).

(4.40) $$\rho_M = \rho_{LM} + \rho_{DM}$$

It is practical to express the results in terms of the relative densities. $D = \dfrac{\rho}{\rho_c}$.

(4.41) $$D_M = \frac{\rho_{LM}}{\rho_c} + \frac{\rho_{DM}}{\rho_c} = D_{LM} + D_{DM}$$

All entities are related to the present epoch, of course.

In the resent years the parameters in (4.41) have been determined rather accurately by somewhat indirect methods, involving a detailed statistical analysis of the temperature fluctuations in the cosmic microwave background radiation (CMB). The details however go far beyond the scope of this chapter.
The idea applied to measure the density for a luminous matter is through the relation to the luminosity density, denoted here by λ . L is the luminosity, and M is the mass.

(4.42) $$\rho_{LM} = \lambda \frac{M}{L}$$

We have separated the mass density into the luminosity density and the mass luminosity ratio.
The luminosity density can be obtained by counting the galaxies per unit volume, multiplied by the average galaxy luminosity. This method has resulted in an estimate.

(4.43) $$\lambda = Luminosity\ density = 0.2\,10^9 L_{sun}/(Mpc)^3$$

The ratio *M/L* is the amount of mass associated on the average with a given amount of light.
The latter quantity is much more difficult to ascertain. Depending of the methods, one finds results, of which the average is.

(4.44) $$\frac{M}{L} = 4\frac{M_{sun}}{L_{sun}}$$

Plugging this into (4.43), we find:

(4.45) $$\rho_{LM} = \lambda \frac{M}{L} = 8\,10^8 M_{sun}/(Mpc)^3 = 5.0\,10^{-29}\ \ kg/m^3 ,$$

this in turn gives the relative density of luminous matter (LM)

(4.46) $$D_{LM} = \frac{\rho_{LM}}{\rho_c} = \frac{5.0\,10^{-29}}{7.2\,10^{-27}} = 0.007 \text{ (Other calculations give 0.005)}$$

This estimate asserts, rather surprisingly that only 0.5 - 0.7 % of the matter in the universe is luminous matter, the rest is dark matter! Non luminous matter could be planets, brown dwarfs (gas planets like Jupiter), black holes as well as interstellar gasses.
This we shall denote atomic matter (AM), since it eventually is built from the elements in the periodic system.

For various reasons we shall not go farther into what the astrophysicists claim, namely that the dark matter cannot be subscribed to atomic matter. Most of the dark matter is called exotic matter. There have been proposed various candidates for the exotic particles, but until now it is pure speculation. Also the exotic particles can not be observed, since they have no electromagnetic interaction.
Alternatively a revision of Newton's law of gravitation on galactic distances has been proposed, of which Einstein's cosmological constant is an example, but neither of the theories have prevailed.

However there exists observational evidence on the amount of "exotic matter".

The fraction of atomic matter (AM) is the sum of luminous matter (LM) and atomic dark matter (ADM)

(4.47) $$D_{AM} = D_{LM} + D_{ADM}$$

The dark matter (DM) can again be written as the sum of atomic dark matter and exotic matter (EXO)

(4.48) $$D_{DM} = D_{EXO} + D_{ADM}$$

So the content of matter (M) in the universe may be split up into 3 parts.

(4.49) $$D_M = D_{LM} + D_{ADM} + D_{EXO}$$

In an attempt to estimate the amount of non exotic matter one must try to measure $< v^2 >$ the mean square velocity of the galaxy haloes, and then apply the *virial theorem* of statistical mechanics.

The virial theorem states that for a closed system in equilibrium the mean kinetic energy is numerically equal to half the mean potential energy per particle.

In this context the kinetic energy is denoted T and the potential energy is denoted U.
First we illustrate this by looking at a body in a circular orbit in a central gravitational field. Here we know:

(4.50) $$U = -G\frac{mM}{r} \quad and \quad F = m\frac{v^2}{r} = G\frac{mM}{r^2}$$

From which immediately follows:

(4.51) $$T = \tfrac{1}{2}mv^2 = \tfrac{1}{2}G\frac{mM}{r} = -\tfrac{1}{2}U \quad \Rightarrow \quad 2T + U = 0$$

Note that $v^2 = GM\frac{1}{r}$, so that the factor of proportionality between v^2 and $\frac{1}{r}$ is GM, when a particle moves in a central field.

.

This can now be extended an arbitrary number n of particles, moving in a gravitational field.

$$< T >= \tfrac{1}{n}\sum \tfrac{1}{2}m\vec{v}_i \cdot \vec{v}_i \quad \Rightarrow \quad \frac{d}{dt} < T > = \tfrac{1}{n}\frac{d}{dt}\sum \tfrac{1}{2}m\vec{v}_i \cdot \vec{v}_i = \tfrac{1}{n}\sum m\frac{d\vec{v}_i}{dt}\cdot \vec{v}_i =$$

$$\tfrac{1}{n}\sum_i\sum_j \vec{v}_i\cdot\vec{F}_{ij} = -\tfrac{1}{n}\sum_i \tfrac{1}{2}\sum_j \vec{v}_i\cdot\nabla U_{ij} = -\tfrac{1}{2}\tfrac{1}{n}\sum_i\sum_j (v_{ix}, v_{iy}, v_{iz})\cdot(\frac{\partial U_{ij}}{\partial x}, \frac{\partial U_{ij}}{\partial x}, \frac{\partial U_{ij}}{\partial x})$$

The factor $\tfrac{1}{2}$ enters because $\vec{v}_i\cdot\nabla U_{ij}$ and $\vec{v}_j\cdot\nabla U_{ji}$ only gives one contribution to the potential energy. Noticing that:

$$v_x\frac{\partial U}{\partial x}+v_y\frac{\partial U}{\partial y}+v_z\frac{\partial U}{\partial z}=\frac{\partial U}{\partial x}\frac{dx}{dt}+\frac{\partial U}{\partial y}\frac{dy}{dt}+\frac{\partial U}{\partial z}\frac{dz}{dt}=\frac{dU}{dt}$$

We find

$$\frac{d}{dt}<T>=-\tfrac{1}{2}\frac{d}{dt}<U>$$

Or

(4.52) $$2<T>+<U>=0$$

The mean kinetic energy is half the loss in the mean potential energy and the other half goes to radiation.

For a central force: $U=-G\frac{m_i M}{r}$ and after inserting this in the virial theorem, we get:

$$2<\tfrac{1}{2}m\vec{v}\cdot\vec{v}>-<G\frac{mM}{r}>=0$$

And after reducing by *m*

(4.53) $$<v^2>=GM<\frac{1}{r}>$$

From such considerations and measuring the mean of the square of the velocity, applying gravitational lenses and several other means, it has been estimated that the relative mass density is only 30% of the critical density.

(4.54) $$D_M\approx 0.30$$

Further studies have revealed that the atomic matter (AM) is only about 4% of the critical mass density.

Since D_{AM} = 0.04 is greater than D_{LU} = 0.007, it means that only a fraction of atomic matter is visible to us. Furthermore since $D_{AM} = 0.04$ is much less than D_M , the intriguing conclusion is that most of the matter in the universe is exotic of which neither its composition nor its origin is known.

If your university education in physics goes back to the 1960'ties the notion of non atomic dark matter is very difficult to comprehend. Year back is was rather fundamental that the world consisted of elementary particles of which the missing ones that sustain the Standard Model have recently been found.

The notion that about 90% of the matter in the universe is made up of yet unknown not atomic matter is simply incomprehensible. If it should be so, then why have we never observed non atomic matter in our planetary system.

But rather than claim that the observations must be "wrong" it tends to a revision of the law of gravitation on galactic distances. But on the other hand the cosmological constant has also led to a dead end.

4.12 The cosmological principle and the Robertson–Walker metric

The Cosmological Principle has been stated in section (4.7). The main content is that the universe is homogenous and isotropic in a galactic scale of e.g. 1 *Mpc*.
Homogenous means that there is roughly the same number of galaxies within a volume of perhaps 1 cubic *Mpc*, and isotropic means that the universe looks the same no matter from where you observe it. Independently of your position in the universe you will observe that the galaxies are receding from you in accordance with Hubble's law.

This behaviour of the universe introduces a model where the universe is perceived as a cosmic fluid. Thus the universe is a physical system where the fluid particles are the galaxies.
Each fluid element has its proper time t, and its spatial coordinate x_i. A co-moving observer flows with a cosmic fluid element.

The co-moving coordinate time can be synchronized over the whole system. For example the proper time is inversely proportional to cosmic microwave background temperature.
We shall now present a model of the universe based on General Relativity conceived as a cosmic fluid. The corresponding metric is called the Robertson–Walker metric.
First, however we shall take a review on the concept of curvature.

4.13 Curvature in General Relativity

In the next section, we shall occupy ourselves with the overall structure of the universe.
When calculating the critical density we considered the three possibilities:

An open universe: $\rho_{universe} < \rho_{critical}$

A closed universe: $\rho_{universe} > \rho_{critical}$

A flat universe: $\rho_{universe} = \rho_{critical}$

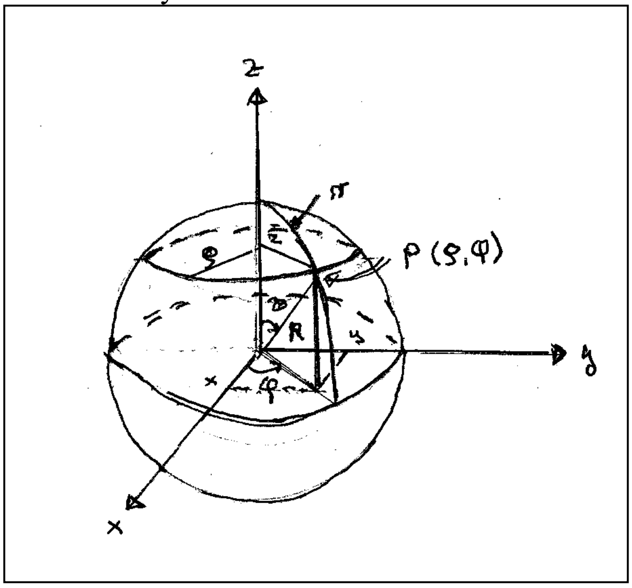
For the open universe the density is too low to stop the expansion and the expansion will continue.

For a closed universe the velocities of the galaxies are lower than the escape velocity, so a contraction will eventually be initiated.

In a flat space, the velocities of the galaxies are equal to the escape velocities, and the expansion will stop at a static universe within an infinite time.

4.13.1 Gaussian curvature

Gaussian differential geometry describes a surface in generalized coordinates or parameters (x^1, x^2) For a spherical symmetric surface, one can use either the polar coordinate system or the cylindrical coordinate system. The notation is shown in the figure (4.55)



A point P on the sphere have the **cylindrical coordinates** (ρ, φ), so that:

$$x = \rho\cos\varphi\,,\ y = \rho\sin\varphi\,,\ z = \pm\sqrt{R^2 - \rho^2}$$

Since $\rho = R\sin\theta$ and $z = R\cos\theta$, we get the **spherical coordinates**. (4.56)

$$x = R\sin\theta\cos\varphi\,,\ y = R\sin\theta\sin\varphi\,,\ z = R\cos\theta$$

The variable r denotes the latitude measured from the North Pole, and ϕ is the longitude. $\theta = \frac{r}{R}$, and this allows us to write (4.56) in a manner called the **polar coordinates**:

(4.57)
$$x = R\sin\frac{r}{R}\cos\varphi\,,\ y = R\sin\frac{r}{R}\sin\varphi\,,\ z = R\cos\frac{r}{R}$$

Gauss' Teorema Egregium (The excellent theorem) is a formula for the Gaussian curvature expressed by the metric and its first and second derivatives. (See "Differential Geometry 1"). Without loss of generality we shall quote the result for a diagonalized metric:

$$g_{ij} = \begin{pmatrix} g_{11} & 0 \\ 0 & g_{22} \end{pmatrix} \quad \text{or} \quad g_{ij} = diag(g_{11}, g_{22})$$

The formula for the Gaussian curvature K is:

(4.58)
$$K = \frac{1}{2g_{11}g_{22}}\left\{-\frac{\partial^2 g_{11}}{\partial x_1^{\,2}} - \frac{\partial^2 g_{22}}{\partial x_1^{\,2}} + \frac{1}{2g_{11}}\left[\frac{\partial g_{11}}{\partial x_1}\frac{\partial g_{22}}{\partial x_1} + \left(\frac{\partial g_{11}}{\partial x_2}\right)^2\right]\right\} + \frac{1}{2g_{22}}\left[\frac{\partial g_{11}}{\partial x_2}\frac{\partial g_{22}}{\partial x_2} + \left(\frac{\partial g_{22}}{\partial x_1}\right)^2\right]$$

It is important to note that the curvature is expressed only in terms of the metric and its derivatives, thus the curvature K is an intrinsic geometrical object.

If you view a surface from outside i.e. from a 3D space, you may wonder if there should not be two curvatures to describe the surface in a point. In fact there are, the normal curvature and the Gaussian curvature, but the first refers to how the surface looks from 3D space.

To illustrate this, consider a 1-dimensional curve. From an intrinsic point of view it is described by only one parameter, the length. This parameter can measure no intrinsic curvature whatsoever.
So the number of curvatures are zero. Correspondingly the number of intrinsic curvatures for a surface is one.
For a position independent metric as the Euclidian space, the curvature $K = 0$, simply because all the derivatives in (4.58) vanishes.
If we change our coordinate system to a position dependent metric, e.g. for a plane

$$(x_1\, x_2) \rightarrow (r,\varphi) \text{ where } x_1 = r\cos\varphi \;\; and \;\; x_2 = r\sin\varphi,$$

We find for the square of the distance element: $ds^2 = dr^2 + r^2 d\varphi^2$ and therefore $g_{11} = 1$ and $g_{22} = r^2$, so the metric is position dependent.
This should however certainly not change the curvature and it doesn't, as the following calculation shows. $x_1 = r \;\; and \;\; x_2 = \varphi$

(4.59) $$g_{11} = 1 \;\Rightarrow \frac{\partial g_{11}}{\partial x_1} = \frac{\partial g_{11}}{\partial x_2} = 0 \;\; and \;\; \frac{\partial g_{22}}{\partial x_1} = 2x_1 \;\; and \;\; \frac{\partial^2 g_{22}}{\partial {x_1}^2} = 2$$

$$K = \frac{1}{2(x_1)^2}\left\{-2 + \frac{1}{2(x_1)^2}\left[4(x_1)^2\right]\right\} = 0$$

For a **spherical surface**, where we use polar coordinates: $(x_1\, x_2) = (r,\varphi)$, we have the distance element

$$ds^2 = dr^2 + R^2\sin^2\theta d\varphi^2 = dr^2 + R^2\sin^2\left(\frac{r}{R}\right)d\varphi^2 \qquad \text{See fig. (4.55)}$$

Thus $g_{12} = g_{21} = 0$, $g_{11} = 1$ and $g_{22} = R^2\sin^2\left(\frac{r}{R}\right)$ with

$$\frac{\partial g_{11}}{\partial x_1} = \frac{\partial g_{11}}{\partial x_2} = 0,$$

$$\frac{\partial g_{22}}{\partial x_1} = 2R\sin(x_1/R)\cos(x_1/R) = R\sin(2x_1/R)$$

$$\frac{\partial^2 g_{22}}{\partial {x_1}^2} = 2\cos(2x_1/R)$$

This leads to

(4.60) $$K = \frac{1}{2R^2\sin^2(x_1/R)}\left\{2\sin^2(x_1/R) - 2\cos^2(x_1/R) + \frac{4R^2\sin^2(x_1/R)\cos^2(x_1/R)}{2R^2\sin^2(x_1/R)}\right\} = \frac{1}{R^2}$$

So the curvature of a sphere is $1/R^2$.
You will naturally get the same result (with a little effort) using cylindrical coordinates.

4.13.2 Spaces with a constant curvature

The cosmological principle states that the universe is isotropic and homogenous. As a consequence of General Relativity the space must have constant curvature. Since it is not possible to visualize a 3D space in the same manner as a 2D surface, we shall first study 2D surfaces, and subsequently make a plausible ad hoc generalization to three dimensions.

We have already seen that a sphere has a constant curvature $K = 1/R^2$ where R is the radius.
A plane has the curvature $K = 0$.

One can prove that there exists only 3 surfaces having a constant curvature: $K = k/R^2$, where $k = \{-1, 0, 1\}$. $K = 1$, is the sphere, $k = 0$ is the plane, whereas $k = -1$ is called a 2-**pseudosphere**. If we merely want to try to change the metric to get a negative curvature $K = -k/R^2$ it would (from a mathematical point of view) be worth trying to change the spherical parameters from trigonometric ones to hyperbolic ones.

$$\sin \to \sinh \quad and \quad \cos \to \cosh$$

Indicating a change of the square of the distance element:

$$ds^2 = dr^2 + R^2 \sin^2\left(\frac{r}{R}\right) d\varphi^2 \to dr^2 + R^2 \sinh^2\left(\frac{r}{R}\right) d\varphi^2 \qquad (4.61)$$

So that

$$g_{11} = 1 \quad and \quad g_{22} = R^2 \sinh^2(r/R),$$

From which we find:

$$\frac{\partial g_{22}}{\partial x_1} = \frac{\partial g_{22}}{\partial r} = 2R\sinh(r/R)\cosh(r/R) = R\sinh(2r/R) = R\sinh(2x_1/R)$$

and

$$\frac{\partial^2 g_{22}}{\partial x_1^2} = \frac{\partial^2 g_{22}}{\partial r^2} = 2\cosh(2r/R) = 2\cosh(2x_1/R)$$

Plugging this into (4.58)

$$K = \frac{1}{2g_{11}g_{22}}\left\{-\frac{\partial^2 g_{11}}{\partial x_1^2} - \frac{\partial^2 g_{22}}{\partial x_1^2} + \frac{1}{2g_{11}}\left[\frac{\partial g_{11}}{\partial x_1}\frac{\partial g_{22}}{\partial x_1} + \left(\frac{\partial g_{11}}{\partial x_2}\right)^2\right]\right\} + \frac{1}{2g_{22}}\left[\frac{\partial g_{11}}{\partial x_2}\frac{\partial g_{22}}{\partial x_2} + \left(\frac{\partial g_{22}}{\partial x_1}\right)^2\right]$$

To give:

$$K = \frac{1}{R^2\sinh^2(r/R)}\left\{-2\cosh(2r/R) + \frac{(R\sinh(2r/R))^2}{R^2\sinh^2(r/R)}\right\}$$

$$K = \frac{1}{R^2\sinh^2(r/R)}\left\{-2(\cosh^2(r/R) + \sinh^2(r/R)) + \frac{(R\sinh(2r/R))^2}{R^2\sinh^2(r/R)}\right\}$$

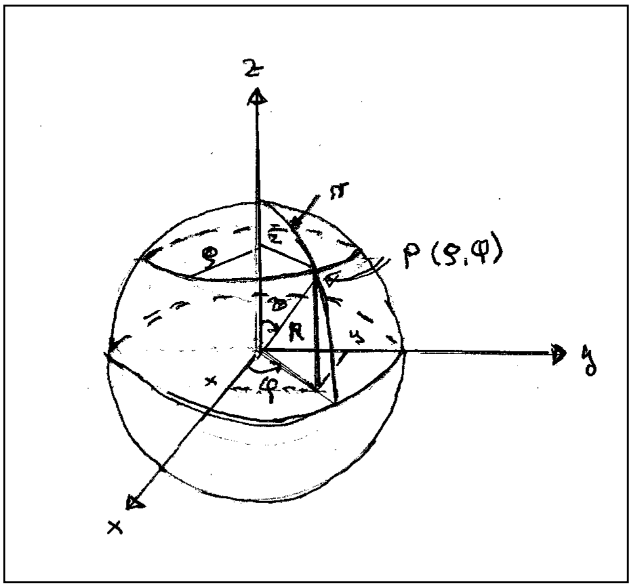
$$K = \frac{1}{R^2 \sinh^2(r/R)}\left\{-2\cosh^2(r/R) - 2\sinh^2(r/R) + \frac{\sinh^2(2r/R)}{\sinh^2(r/R)}\right\}$$

$$K = \frac{1}{R^2 \sinh^2(r/R)}\left\{-2\cosh^2(r/R) - 2\sinh^2(r/R) + \frac{2\sinh^2(r/R)\cosh^2(r/R)}{\sinh^2(r/R)}\right\}$$

$$K = \frac{1}{R^2 \sinh^2(r/R)}\left\{-2\cosh^2(r/R) - 2\sinh^2(r/R) + 2\cosh^2(r/R)\right\}$$

(4.62) $$K = -\frac{1}{R^2}$$

A surface (space) with a negative curvature is referred to as the **hyperbolic case**.



If we want to express the metric in **cylindrical coordinates** (ρ, φ), where:

$$x = \rho\cos\varphi \,,\ y = \rho\sin\varphi \,,\ z = \pm\sqrt{R^2 - \rho^2} = R\cos\vartheta$$

The square of the distance element is:

$$ds^2 = {ds_\rho}^2 + {ds_\theta}^2 = \rho^2 d\varphi^2 + R^2 d\theta^2$$

To change the variable from θ to ρ we use

$\sin\theta = \rho / R$ and $\cos\theta = \sqrt{1 - \sin^2\theta}$

$$d\theta = \frac{d\theta}{d\rho}d\rho = \left(\frac{d\rho}{d\theta}\right)^{-1} d\rho = \frac{1}{R\cos\theta}d\rho$$

$$d\theta = \frac{1}{R\sqrt{1 - \rho^2/R^2}}d\rho = \frac{R}{\sqrt{R^2 - \rho^2}}d\rho$$

And the distance element becomes

(4.63) $$ds^2 = R^2 d\theta^2 + \rho^2 d\varphi^2 \qquad ds^2 = \frac{R^2}{R^2 - \rho^2}d\rho^2 + \rho^2 d\varphi^2$$

Which yields the metric for the cylindrical coordinates $(x_1, x_2) = (\rho, \phi)$.

(4.64) $$g_{12} = g_{21} = 0\,;\ g_{11} = \frac{R^2}{R^2 - \rho^2}\,,\ g_{22} = \rho^2$$

The (only) three surfaces having a constant curvature can be collectively expressed in polar coordinates.

Using $\eta = \frac{r}{R}$ as a dimensionless radial coordinate, the distance elements can be written in a compact form:

(4.65) $$\left[ds^2\right]^k_{2D,\theta} = \begin{cases} R^2(d\eta^2 + \sin^2 \eta d\phi^2) & for \quad k=1 \\ R^2(d\eta^2 + \eta^2 d\phi^2) & for \quad k=0 \\ R^2(d\eta^2 + \sinh^2 \eta d\phi^2) & for \quad k=-1 \end{cases}$$

Introducing $\sqrt{-1} = i$ (the imaginary unit $i^2 = -1$), and since $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$, it follows $\sin ix = \frac{1}{2i}(e^{-x} - e^{x}) = i \sinh x$, and using $\lim_{k \to 0} \frac{\sin k\eta}{k} = \eta$, (4.6) can be further compressed to

(4.66) $$\left[ds^2\right]^k_{2D,\theta} = R^2(d\eta^2 + \frac{1}{k}\sin^2(\sqrt{k}\eta)d\phi^2)$$

We shall then argue (but not prove) how the geometry of a surface with constant curvature can be generalized to a 3D space.

Compared to the 2D coordinates (r, ϕ) or (ρ, ϕ), the 3D spherical (r, θ, ϕ) or (ρ, θ, ϕ) requires an additional (polar) angle coordinate. Specifically for the $k = 0$ case, we have the polar coordinates for the flat 2D surface, and the spherical coordinates for the Euclidian 3D space.
Their respective metric relations are:

(4.67) and
$$\left[ds^2\right]^{k=0}_{2D,\theta} = dr^2 + r^2 d\phi^2$$
$$\left[ds^2\right]^{k=0}_{3D,\theta} = dr^2 + r^2 d\Omega^2, \quad \text{where} \quad d\Omega^2 = d\theta^2 + + \sin^2 \theta d\phi^2$$

Now the idea is, that we do the same thing in (4.66), but also in the case where $k \neq 0$, replacing $d\phi^2$ with $d\Omega^2 = d\theta^2 + + \sin^2 \theta d\phi^2$, which brings us to have three (spherical polar) coordinates η, θ *and* ϕ. Making this alteration in (4.66) we have.

(4.68) $$\left[ds^2\right]^k_{3D,\eta} = \begin{cases} R^2(d\eta^2 + \sin^2 \eta d\Omega^2) & for \quad k=1 \\ R^2(d\eta^2 + \eta^2 d\Omega^2) & for \quad k=0 \\ R^2(d\eta^2 + \sinh^2 \eta d\Omega^2) & for \quad k=-1 \end{cases}$$

However, it is no longer possible to visualize this 3D metric as embedded in a space of higher dimension. Similarly if we replace the radial coordinate ρ in $ds^2 = \frac{R^2}{R^2 - \rho^2} d\rho^2 + \rho^2 d\varphi^2$, by the dimensionless distance: $\xi = \rho / R$ we get using (4.63)

(4.69) $$\left[ds^2\right]^k_{3D,\xi} = R^2\left(\frac{d\xi^2}{1-kd\xi^2} + \xi^2 d\Omega^2\right)$$

Equation (4.68) and (4.69) reduces to 2D spaces if we make a projection by setting either $d\theta = 0$ or $d\phi = 0$. This means that all 2D spaces are appropriately curved.
A rigorous mathematical derivation of this result is more demanding, but the results are good.

The metric in (4.68) and (4.69) with $k = +1$ describes a 3-sphere, and for $k = -1$ a 3-pseudosphere, and the overall distance scale R is identified as the radius of these spheres.

4.13 The Robertson – Walker metric

We shall now present a General Relativity model of the universe as a cosmic fluid.
The corresponding metric is called the Robertson–Walker metric.
Because each "fluid element" carries its own position coordinate, the fluid element is also a cosmic rest frame, as each fluid elements position coordinate are unchanged with time.

As we shall see below the expanding universe with all the galaxies rushing away from each other, is not described by changing coordinate positions, but rather by an ever increasing metric.
This emphasizes that the physics of the universe is not a universe expanding in space, but an expansion of space itself.

We have earlier seen the spherical symmetry restricts the metric to have only two scalar functions.

(4.70) $$g_{ij} = diag(g_{00}(r,t), g_{rr}(r,t), r^2, r^2 \sin^2\theta)$$

Here we shall dwell into the geometry, outlined from the cosmological principle, which turns out to have the Robertson-Walker metric, when expressed in co-moving coordinates.
Because the fluid elements coordinate time is its proper time, we must have $g_{00} = -1$, $g_{0i} = 0$ and $g_{i0} = 0$, where i is only the spatial index.

Assume that g_{ij} is a metric that satisfies the cosmological principle, we have:

(4.71) $$ds^2 = -c^2dt^2 + g_{ij}(r,t)dx^i dx^j \qquad (i, j) = 1..3$$

If we put $g_{ij}(r,t)dx^i dx^j \equiv dl^2$, we get:

$$ds^2 = -c^2dt^2 + dl^2$$

Because of the cosmological principle, (the universe is isotropic, and homogenous), dl^2 can have no dependence on the spatial coordinates, and therefore it must be an overall time dependent scale factor $R(t)$.

(4.71) $$dl^2 = R(t)^2 d\hat{l}^2$$

Where the reduced length element $d\hat{l}$ is both time independent and dimensionless.
It use useful to define a dimensionless scale factor.

(4.72) $$a(t) = \frac{R(t)}{R_0}$$

$a(t)$ is normalized so that $a(t_0) = 1$. R_0 is often referred to, as the radius of the universe now.
In this way the universe is pictured as a 3D map with a label fixed by the co-moving coordinates $\hat{x}_i$. The evolution of the fluid elements enters entirely through the map scale $R(t) = a(t)R_0$.

(4.73) $$x_i(t) = R(t)\hat{x}_i$$

With $\hat{x}_i$ being the fixed (time independent) dimensionless map coordinate, while $a(t)$ is the size of the grids, and is independent of map coordinates.
As the universe expands the relative distance relations are unchanged i.e. the shapes are not changed.
The Robertson–Walker metric is designed for a space-time, which is homogenous and isotropic, and therefore it has a constant curvature. We have in (4.68) and (4.69) written down the metric in two spherical coordinate systems.

(4.74) $$\left[ds^2\right]^k_{3D,\eta} = \begin{cases} R^2(d\eta^2 + \sin^2\eta d\Omega^2) & for \quad k=1 \\ R^2(d\eta^2 + \eta^2 d\Omega^2) & for \quad k=0 \\ R^2(d\eta^2 + \sinh^2\eta d\Omega^2) & for \quad k=-1 \end{cases}$$

(4.74) $$\left[ds^2\right]^k_{3D,\varsigma} = R^2\left(\frac{d\varsigma^2}{1-kd\varsigma^2} + \varsigma^2 d\Omega^2\right)$$

k takes the values: 0 for a 3D Euclidian space, +1 for a 3-sphere, and -1 for a 3-pseudospere.
In the context of cosmology, the structure of the universe is, for

$k = +1$, a positively curved space, (parabolic-elliptic), and is referred to as a "**closed universe**",

$k = -1$, a negatively curved space (hyperbolic), and is referred to as a "**open universe**", and

$k = 0$ a "**flat universe**".

4.14 Proper distances in the Robertson–Walker metric

In practice, one can use either of the geometries (4.73) or (4.74), since they are equivalent.
In the following, however we shall use the (4.74) with the coordinates $(\varsigma, \theta, \phi)$.
In an expanding universe that may be curved, one should be careful in any treatment of distances.
We shall lay out with the most simple, the proper distances.

The proper distance $d_p(\varsigma, t)$ to a point at the co-moving radial distance ν, and cosmic time t, can be calculated from the metric (5.77) with $d\Omega = 0$, and $dt = 0$.

(4.75)
$$d_p(\varsigma,t) = a(t)R_0\int_0^{\nu}\frac{d\varsigma}{\sqrt{1-kd\varsigma^2}}$$
$$= a(t)\left(\frac{R_0}{\sqrt{k}}\right)\sin^{-1}\left(\sqrt{k}\varsigma\right)$$

From which we have for a positive curvature $k = +1$

(4.76)
$$d_p(\varsigma,t) = a(t)R_0\sin^{-1}(\varsigma)$$

For a negative curvature $k = -1$

(4.77)
$$d_p(\varsigma,t) = a(t)R_0\sinh^{-1}(\varsigma)$$

And for a flat space

(4.78)
$$d_p(\varsigma,t) = a(t)R_0\varsigma$$

The relation

(4.79)
$$d_p(\varsigma,t) = a(t)d_p(\varsigma,t_0)$$

Implies a proper velocity

(4.80)
$$v_p(t) = \frac{d(d_p)}{dt} = \dot{a}(t)d_p(\varsigma,t_0) \quad \Rightarrow$$

(4.80)
$$v_p(t) = \frac{\dot{a}(t)}{a(t)}d_p(t)$$

What is recognized as Hubble's law with:

(4.81)
$$v = v_p(t) \quad , \quad H(t) = \frac{\dot{a}(t)}{a(t)} \quad r = d_p(t) \quad and \quad H_0 = \dot{a}(t_0), \quad \text{since } a(t_0) = 1$$

Recall that the appearance of the overall scale factor in the special part of the Robertson–Walker metric follows from our imposition of the homogeneity and isotropy condition.
Equation (4.80) then confirms the assertion that the cosmological principle automatically leads to Hubble's law.
Again we emphasize that the drift of galaxies away from any point in the universe, is not a drift in space but an expansion of space itself.

5. The expanding universe

In the last section, we studied the "kinematics" of the standard model of cosmology.
The requirement of homogeneity and isotropy fixes the space-time to have the Robertson-Walker metric in co-moving coordinates. This geometry is specified by a curvature signature $k = 0, +1, -1$, and a time dependent scale factor $a(t)$.

We shall now turn to the dynamics of the homogeneous and isotropic universe, and determine the unknown quantities *k* and *a*(*t*) by the matter and energy sources, using Einstein field equations.

The fact that all galaxies are moving away from each other is a strong indication that the universe has been smaller and hotter, and ultimately has been concentrated in a point of infinite density. Considering the expanding universe, there are notably two evidences for the Big Bang theory.

The first one is the discovery of an all pervasive cosmic microwave background (CMB).

The idea is that when the universe became transparent, that is, the photon age, some 350 mill years after the big bag, it was inhabited with energetic photons, say some Gev. But when space expanded, so did the wavelength. So when one estimates the factor with which the universe has grown until now, and compare it to the factor of enlargement of the wavelengths of the early universe, the two numbers are in fact comparable.
A photon of 1 GeV corresponds to a wavelength

$$\lambda = \frac{hc}{E_\gamma} = \frac{4.15\,10^{-15}eVs\;3.0\,10^{8}m/s}{10^{9}eV} = 1.25 10^{-15}m$$

The wavelength of the CMB radiation is peaked at $\lambda = 2.4\ 10^{-3}\ m$. This means that the wavelength is increased by a factor $O(10^{12})$. However we neither know precisely the size of the universe at the photon age or now, but since the universe has an age of $13.6\ 10^{9}\ y$, the two factors of expansion are comparable, although a direct calculation requires a much more detailed knowledge of the early universe.
Another indirect confirmation of the Big Bang theory of the expanding universe was delivered by the discovery of the CMB radiation, which is still here and everywhere.
In an infinite universe the CMB radiation would long gone have disappeared at the speed of light.

The CMB radiation in the universe can likewise, and more correctly be viewed as the Planck "black body" radiation of a closed universe, where the temperature has decrease from extreme temperatures to the present Planck temperature of 2.7 *K*.
The second evidence is the observed abundance of the light nuclear elements, ^{4}He, $^{2}_{1}H$, Li in agreement with values predicted by particle physics.

5.1 The Friedmann equations

The Einstein equation relates space-time geometry to the mass/energy distribution.

$$G_{i\,j} = \kappa T_{i\,j}$$

We have already established the Robertson-Walker geometry with co-moving coordinates as the appropriate geometry for a homogenous and isotropic universe.
The simplest choice of the energy momentum tensor is an ideal (a non viscous) fluid.
The proper description of the ideal fluid tensor will be given later, but it is specified by two parameters mass density $\rho(t)$ and pressure $p(t)$. General Relativity relates the geometric parameters of curvature signature *k* and the scale factor $a(t) = R(t)/R_0$ to the cosmic fluid parameters.
The resulting equations are called the Friedmann equations. The first and second Friedmann equation are

(5.1) $$\frac{\dot{a}(t)^2}{a(t)^2}+\frac{kc^2}{R_0{}^2a(t)^2}=\frac{8\pi G}{3}\rho$$

(5.2) $$\frac{\ddot{a}(t)}{a(t)}=-\frac{4\pi G}{c^2}\left(p+\tfrac{1}{3}\rho c^2\right)$$

Because pressure and density are positive quantities, the second derivative of $a(t)$ is negative. This is in accordance with the increase in separation, as the gravitational attraction decreases.
A linear combination of the two Friedmann equations gives:

(5.3) $$\frac{d}{dt}(\rho c^2a^3)=-p\frac{da^3}{dt}$$

This resembles the first law of thermodynamics on differential form, which is the most general statement of energy conservation.

$$dE=dQ-PdV \quad and \quad dQ=0 \quad \Rightarrow \quad dE=-PdV$$

The more simple equation (5.3) is often used instead of (5.1) and (5.2), and is then referred to as the Friedmann equation.
However since there are three unknown functions $a(t)$, $\rho(t)$, $p(t)$, and only two equations, we shall need one more equation.
The equation relates the pressure to the density, which we choose in its most simple form.

(5.4) $$p=w\rho c^2$$

Where w is a constant that determines the material content of the system. For non relativistic matter the pressure is negligible, so $w = 0$, and for radiation $w=\frac{1}{3}$, as we know it from electrodynamics.
A more rigorous derivation of the Friedmann equations will be given later, here we shall proceed with a quasi Newtonian approach, which is more transparent.

5.1 The critical density of the universe

If we divide the first Friedmann equation

$$\frac{\dot{a}(t)^2}{a(t)^2}+\frac{kc^2}{R_0{}^2a(t)^2}=\frac{8\pi G}{3}\rho \quad \text{with} \quad \rho_c=\frac{3}{8\pi G}\frac{\dot{a}^2}{a^2},$$

using ρ_c on the RHS and $\frac{3}{8\pi G}\frac{\dot{a}^2}{a^2}$ on the LHS, we get

$$\frac{8\pi G}{3}+\frac{c^2}{\dot{a}^2R_0{}^2}\frac{8\pi G}{3}k=\frac{8\pi G}{3}\frac{\rho}{\rho_c} \quad \Leftrightarrow \quad 1+\frac{c^2}{\dot{a}^2R_0{}^2}k=\frac{\rho}{\rho_c} \quad \Leftrightarrow$$

(5.5) $$-k = \left(\frac{\dot{a}R_0}{c}\right)^2 (1 - \frac{\rho}{\rho_c})$$

Using Hubble's constant

$$H(t) = \dot{a}(t)/a(t)\,,\ H_0 = \dot{a}(t_0)/a(t_0) = \dot{a}(t_0)\,,$$

$$\rho_c = \frac{3}{8\pi G}\frac{\dot{a}^2}{a^2} = \frac{3}{8\pi G}H(t)^2$$

and related to $t = t_0$:

$$\rho_{c,0} = \frac{3}{8\pi G}\dot{a}^2 = \frac{3}{8\pi G}H_0^{\ 2}$$

Denoting the relative density by D: $D = \frac{\rho}{\rho_c}$ we find:

$$1 + \frac{c^2}{\dot{a}^2 R_0^{\ 2}}k = \frac{\rho}{\rho_c} \quad \Leftrightarrow -\frac{c^2}{\dot{a}^2 R_0^{\ 2}}k = 1 - D \quad \Leftrightarrow$$

Related to $t = t_0$:

(5.6) $$-k\frac{c^2}{R_0^{\ 2}} = H_0^{\ 2}(1 - D_0)$$

If the density is less than the critical density, that is, if $D_0 < 1$, then $k = -1$, the universe has a negative curvature, and we have an open universe.

If the density is equal to the critical density, that is, if $D_0 = 1$, then $k = 0$, the universe is flat.

If the density is greater than the critical density, that is, if $D_0 > 1$, then $k = +1$, the universe has a positive curvature, and we have a closed universe.

From the available phenomenological data $D_0 = D_{M,0}$, and it therefore looks like that we have an negatively curved open universe.

However this conclusion may be altered, if we modify Einstein's equations by adding a "Cosmological Constant". Adding such a constant does not alter the consistency of Einstein's equation, the problem is (and this has always been the problem in theoretical physics), that such a constant has no other physical justification, than to explain a discrepancy between theory and data.

Using the cosmological constant and interpreting new observational data have caused several physicist/astronomers to be inclined to conceive the universe as flat

5.2 Particle physics and the Big Bang. The Cosmic microwave background

The Standard model of particle physics was commonly accepted in the beginning of the 90'ties. This model has enabled us theoretically to give a description of the development of the universe down to 10^{-43} *s* after the big bang (Vacuum inflation). It is impossible to give even a brief survey of the standard model, but its foundation is the knowledge of the four interaction of the fundamental particles of which we shall give a short review.

5.2.1 The four interactions

There are four fundamental interactions in nature, but two of them are only found on a subatomic level. Below they are listed after their historical appearance in physics.

1. **Gravitational forces**
   The gravitational force is responsible for holding the universe together. It holds together the planetary system, the stars, the galaxies etc. All particles including mass less particles like photons are influenced by gravity. The range of gravity is infinite. Gravity is the weakest of the 4 forces. If you compare the constant i Coulomb law $\frac{1}{4\pi\varepsilon_0} = 9.0\,10^9$ (SI-units) with the constant of gravitation $G = 6.67\,10^{-11}$ (SI-units) the factor of strength is about 10^{20}.

2. **Electromagnetic forces**
   The electromagnetic interactions are responsible for holding the atoms together. Furthermore they are responsible for the whole spectrum of electromagnetic radiation. Both the gravitational forces and the electromagnetic forces are present in the macroscopic world, but the electromagnetic forces are much stronger than the gravitational forces, with a factor that is roughly the inverse relation of the scale between an atom and the size of the planetary system.
   As is the case with the gravitational forces the range of the electromagnetic forces are infinite. Electromagnetic forces are active on all charged particles including particles with a magnetic moment.
   During the last 100 years we have gained a complete understanding of the macroscopic electromagnetic forces, given by the Maxwell Equations, and extended to quantum physics, by the Dirac equation and quantum electrodynamics.

3. **Strong interactions**
   These forces have no visibility in the macroscopic world, since they are the forces that hold the nuclei together. They are responsible for the α-decay of nuclei. Particles with strong interactions are called baryons and mesons and together for hadrons. Apart from the proton and neutron, there exist large families of unstable baryons and mesons. The range is very short some 10^{-15} *m*, about the size of the nuclei. Roughly speaking they are a million times stronger than the electromagnetic forces. With the introduction of the Standard Model, they are almost understood.

4. **Weak interactions**
   The weak interactions are much weaker than the electromagnetic. They are among other things responsible for the β-decay of nuclei and other hadrons. Their range is the same as the strong interactions. All hadrons have weak interactions. Particles with weak interactions, but not strong interactions are called leptons. The neutral neutrino is a mass less particle with weak interactions only.

Below is a table showing the properties of the four exiting forces of nature.

TABLE 46–1 The Four Interactions of Nature

Interaction	Strength	Range	Mediating particle			
			Name	Rest mass	Charge	Spin
Strong	1	Short (~ 1 fm)	Gluon	0	0	1
Electromagnetic	$\frac{1}{137}$	Long ($1/r^2$)	Photon	0	0	1
Weak	10^{-9}	Short (~ 0.001 fm)	$W^{\pm}, Z^0$	81, 91 GeV/c^2	$\pm e$, 0	1
Gravitational	10^{-38}	Long ($1/r^2$)	Graviton	0	0	2

In the end of the sixties the physicists succeeded in to accomplish a theory: The electro weak theory, which at very high energies (10^9 GeV) unified the electromagnetic and the weak interactions to one interaction: The electro-weak.
During the 70'ties, they finally succeeded to unify the electro-weak and the strong interactions to a grand unified theory (GUT), later to become the Standard Model, which unifies the 3 interactions at extreme high energies (10^{16} GeV). The level of energies where the three or just two interactions are unified are excluded from laboratory experiments. Actually the only place, where the unified theory can be "tested" (indirectly) is the first 10^{-36} s after the big bang, and for the electro-weak 10^{-36} s to 10^{-12} s.

5.2.2 Elementary particles and Quarks

Of the elementary particles, we know from the atomic nucleus, the proton and the neutron are hadrons. They have strong, electromagnetic and weak interactions.
The electron has electromagnetic and weak interactions. The β-decay (weak interaction) emits an electron and a neutrino. The neutrino is mass-less and it has only weak interactions, which makes it virtually impossible to detect in the laboratory. Neutrinos coming from the cosmic radiation were first detected in 1970.
Until 1950 these particles (and their antiparticles) were considered to be fundamental.
Later from accelerator experiments, there were discovered a large number of new, but unstable particles, especially the π-mesons and the myon, (which is a heavy lepton with mass 200 times that of the electron) not to mention some so called "strange particles".
All the new particles were unstable, with a lifetime ranging from 10^{-10} s to 10^{-6} s.

From a theoretical point of view it was not satisfactory with about 50 "elementary particles", and more to come, so the American Physicist Murray Gell-Mann suggested in 1964, that all hadrons were built of three truly fundamental particles, which he (with a highly literary reference to a novel by James Joyce) named quarks.
At that time he assumed, (encouraged by the group theory of SU2) that there were only three quarks, which he gave the rather picturesque names: *up, down, and strange.* These were sufficient the explain the building of hitherto known (SU3) multiplets of particles.
The quarks were "exotic" in the manner that, although they were spin ½ particles, they carried fractions of the charge of the electron, which for the up, down and strange quarks were $\frac{2}{3}e, -\frac{1}{3}e, and -\frac{1}{3}e$, where e is the (elementary) charge of the electron. And they had a baryon number $\frac{1}{3}$.

Particles carrying a fraction of either *e* (charge of the electron) or a fraction of the baryon number have never been observed in nature
For example the proton is built of two *up quarks* and one *down quark*, having the charge $\frac{2}{3}e+\frac{2}{3}e-\frac{1}{3}e=e$, and baryon number $\frac{1}{3}+\frac{1}{3}+\frac{1}{3}=1$

Free quarks have never been observed, and it is the general assumption, that they can not exist as free particles except under the extreme conditions that existed 10^{-36} *s* after the big bang.
Since then accelerator experiments and theoretical considerations, have made it necessary to add another triplet of quarks with equally picturesque names: *charm, top and Bottom.*

A baryon consists of 3 quarks, and a meson consists of a quark and an anti-quark.
The description of the physics of the quarks is called Quantum Cronodynamics, which is part of the Standard model, but the theoretical framework of particle physics is at least as complex as the General Theory of Relativity.
The theory has required the addition of yet another lepton, the tau-meson, the existence of which has been confirmed by accelerator experiments at CERN in 1995, so we now have a beautiful theoretical symmetry, with 6 quarks, their anti-quarks, 6 leptons (e, μ, τ) their neutrinos $(\nu_e, \nu_\mu, \nu_\tau)$ and their anti-particles. The properties of the leptons and the quarks are shown schematically below.
It should be mentioned, that the physicists at CERN in 2013 succeeded to detect the long searched for Higgs boson. This was important, since all the particles in the Standard Model had hitherto been mass-less. The Higgs boson has been predicted theoretically to be the source of mass through the vacuum field.

TABLE 46–2 The Leptons

Particle name	Symbol	Anti-particle	Rest mass (MeV/c^2)	L_e	L_μ	L_τ	Lifetime (s)	Principal Decay Modes
Electron	e^-	e^+	0.511	+1	0	0	Stable	
Neutrino (e)	ν_e	$\bar{\nu}_e$	0(?)	+1	0	0	Stable	
Muon	μ^-	μ^+	105.7	0	+1	0	2.20×10^{-6}	$e^-\bar{\nu}_e\nu_\mu$
Neutrino (μ)	ν_μ	$\bar{\nu}_\mu$	0(?)	0	+1	0	Stable	
Tau	τ^-	τ^+	1784	0	0	+1	$<4 \times 10^{-13}$	$\mu^-\bar{\nu}_\mu\nu_\tau$, $e^-\bar{\nu}_e\nu_\tau$
Neutrino (τ)	ν_τ	$\bar{\nu}_\tau$	0(?)	0	0	+1	Stable	

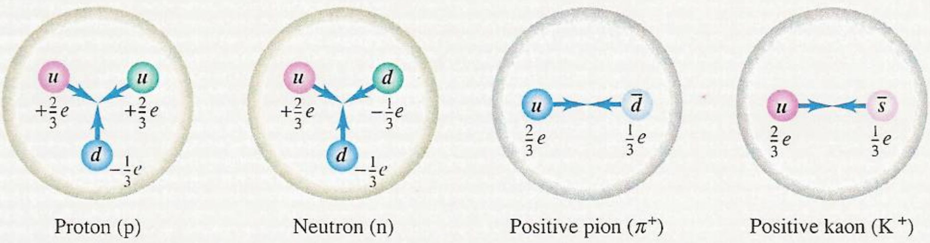
TABLE 46–5 Properties of Quarks

Symbol	Q/e	Spin	Baryon number, B	Strangeness, S	Charm, C	Bottomness, B'	Topness, T
u	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	0	0	0	0
d	$-\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	0	0	0	0
s	$-\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	−1	0	0	0
c	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	0	+1	0	0
b	$-\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	0	0	+1	0
t	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	0	0	0	+1

Following is a similar listing of the most well known hadrons.

TABLE 46–3 Some Known Hadrons and Their Properties

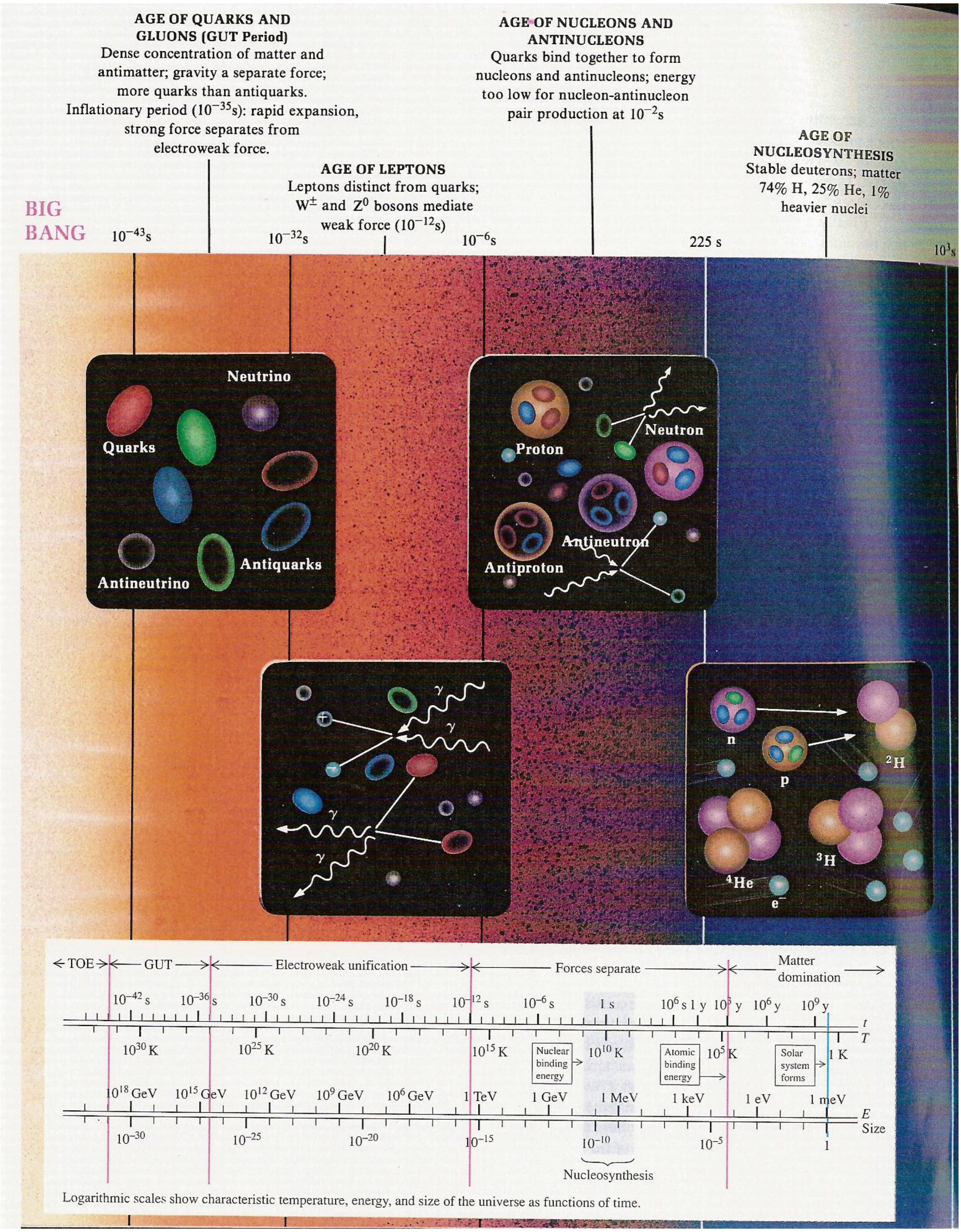
	Particle	Mass (MeV/c^2)	Charge ratio, Q/e	Spin	Baryon number, B	Strangeness, S	Mean lifetime (s)	Typical decay modes	Quark content
Mesons	π^0	135.0	0	0	0	0	0.87×10^{-16}	$\gamma\gamma$	$u\bar{u}$, $d\bar{d}$
	π^+	139.6	+1	0	0	0	2.6×10^{-8}	$\mu^+ \nu_\mu$	$u\bar{d}$
	π^-	139.6	−1	0	0	0	2.6×10^{-8}	$\mu^- \bar{\nu}_\mu$	$\bar{u}d$
	K^+	493.7	+1	0	0	+1	1.24×10^{-8}	$\mu^+ \nu_\mu$	$u\bar{s}$
	K^-	493.7	−1	0	0	−1	1.24×10^{-8}	$\mu^- \bar{\nu}_\mu$	$\bar{u}s$
	η^0	548.8	0	0	0	0	$\approx 10^{-18}$	$\gamma\gamma$	$u\bar{u}$, $d\bar{d}$, $s\bar{s}$
Baryons	p	938.3	+1	$\frac{1}{2}$	1	0	Stable	—	uud
	n	939.6	0	$\frac{1}{2}$	1	0	898	$pe^- \bar{\nu}_e$	udd
	Λ^0	1116	0	$\frac{1}{2}$	1	−1	2.63×10^{-10}	$p\pi^-$ or $n\pi^0$	uds
	Σ^+	1189	+1	$\frac{1}{2}$	1	−1	0.799×10^{-10}	$p\pi^0$ or $n\pi^+$	uus
	Σ^0	1193	0	$\frac{1}{2}$	1	−1	7.4×10^{-20}	$\Lambda^0 \gamma$	uds
	Σ^-	1197	−1	$\frac{1}{2}$	1	−1	1.48×10^{-10}	$\eta\pi^-$	dds
	Ξ^0	1315	0	$\frac{1}{2}$	1	−2	2.90×10^{-10}	$\Lambda^0 \pi^0$	uss
	Ξ^-	1321	−1	$\frac{1}{2}$	1	−2	1.64×10^{-10}	$\Lambda^0 \pi^-$	dss
	Δ^{++}	1232	+2	$\frac{3}{2}$	1	0	10^{-23}	$p\pi^+$	uuu
	Ω^-	1672	−1	$\frac{3}{2}$	1	−3	0.822×10^{-10}	$\Lambda^0 K^-$	sss
	Λ_c^+	2285	+1	$\frac{1}{2}$	1	0	1.91×10^{-13}	$\Sigma^+ \pi\pi\pi$	udc



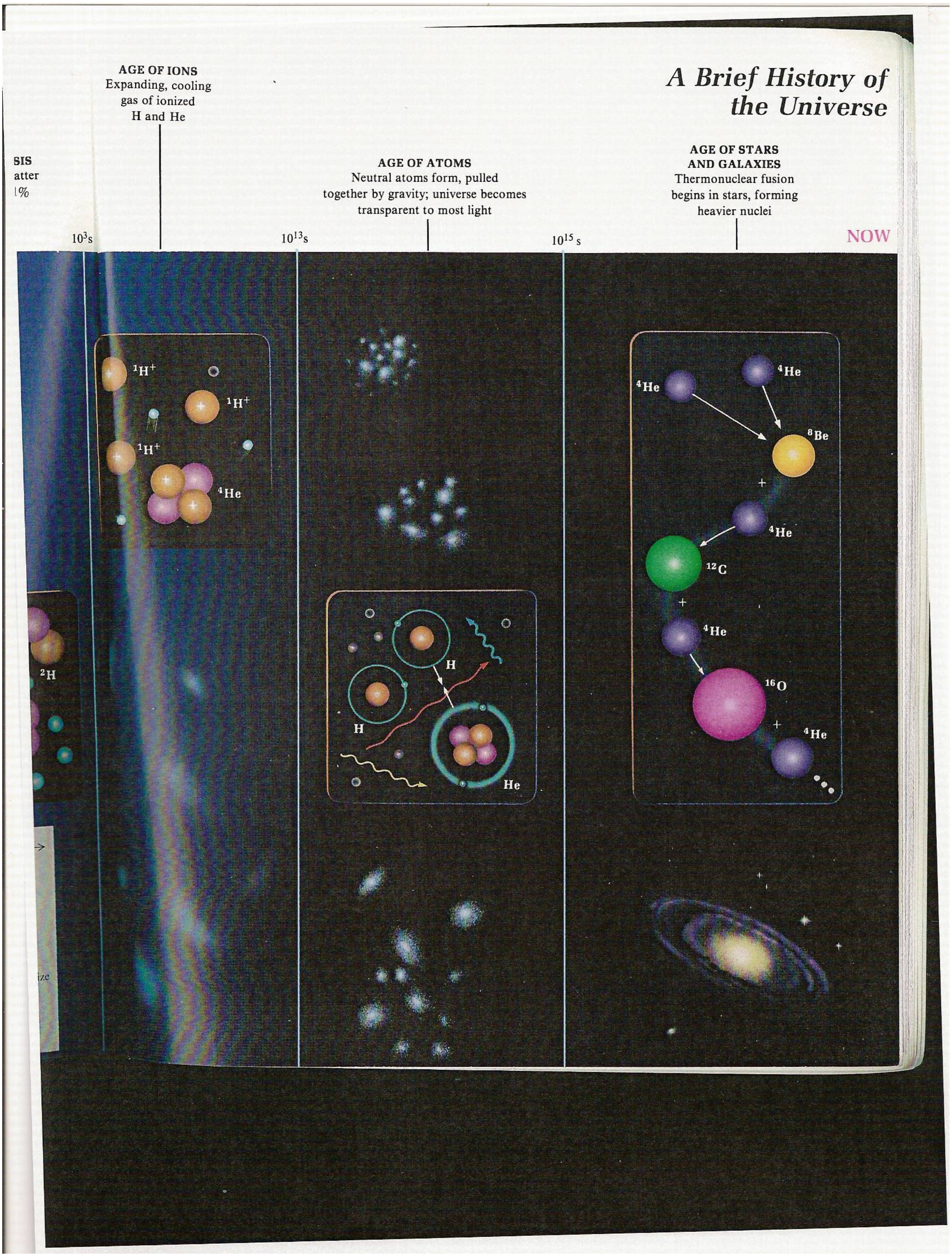
46–10 Quark content of four hadrons. The various color combinations needed for color neutrality are not shown.

5.2.4 The evolution of the universe in pictures

The following pages illustrate the early development of the universe from the point of view of particle physics.



Logarithmic scales show characteristic temperature, energy, and size of the universe as functions of time.


A Brief History of the Universe
AGE OF IONS
Expanding, cooling gas of ionized H and He
AGE OF ATOMS
Neutral atoms form, pulled together by gravity; universe becomes transparent to most light
AGE OF STARS AND GALAXIES
Thermonuclear fusion begins in stars, forming heavier nuclei
NOW
10³s
10¹³s
10¹⁵ s
¹H⁺
⁴He
²H
H
He
⁸Be
¹²C
¹⁶O

One may reasonably ask the question, why it is not possible to study on the universe before 10^{-43} *s* after the Big Bang.
For the particle energies which rule between 10^{-42} *s* to 10^{-36} *s* there is a theory (GUT) which unites the strong, the electromagnetic, and the weak interaction to one interaction, and one imagines that at even higher energies these three interaction will unify with gravitation.

However there is for the present no quantum field theory for the gravitational field. There is still no TOE (Theory of everything), and most physicist doubt that it will ever be found on a quantum field theoretical ground.

It is expected (for theoretical reasons) that the unification of the four interactions will take place at A distance 10^{-35} *m*. This length is called the **Planck length**, and it can be expressed by three constants of nature: The gravitational constant $G = 6.67\ 10^{-11}$ (SI-units), the speed of light $c = 3.0\ 10^{8}$ *m/s* and Planck's constant $h = 4.15\ 10^{-15}$ *eVs*.

(5.7) $$l_P = \sqrt{\frac{hG}{2\pi c^3}} = 1{,}616\,10^{-35}\,m$$ (The Planck length)

The time it takes the light to transverse the Planck length is:

(5.8) $$t_P = \frac{l_P}{c} = \sqrt{\frac{hG}{2\pi c}} = 0{,}593\,10^{-43}\,s$$

We have no possibility to draw any conclusions of what happened to the universe before the Planck time since the universe was smaller than the Planck length.

5.2.5 The cosmic microwave background

Even if the notion of a Big Bang has been known since 1930'ties, (following the establishment of Hubble's law), it had until the 1970'ties still a speculative nature, since no experimental verifications were anticipated, for obvious reasons.

This situation changed, when two electro engineers Arno Penzias and Robert Wilson discovered (in their effort to eliminate radio background noise) a microwave radiation of unknown origin, being the same in all directions.
Later when recording an analyzing the spectrum, it revealed that it coincided with the Planck *black body* radiation at a temperature 2.7 *K*.

When the cosmologists became aware of this discovery, they did not hesitate to conclude that is was a decisive evidence of the Big Bang, considering that the radiation had to be relic from the early seconds of Big Bang, and since no other explanation could possibly account for it.

From an everyday experience it might be difficult to accept that today, we can receive radiation, which was emitted 13.5 Gyr ago, but this explanation is not strictly correct.
Rather, the notion is that a finite universe is a Planck "black-body" in thermodynamic equilibrium and for such a body Planck's thermodynamic law of radiation applies.
For the microwave radiation to prevail the universe must be finite, such that emitted radiation will never be lost in "space", but one may also think of it, as when the space expands, so does the

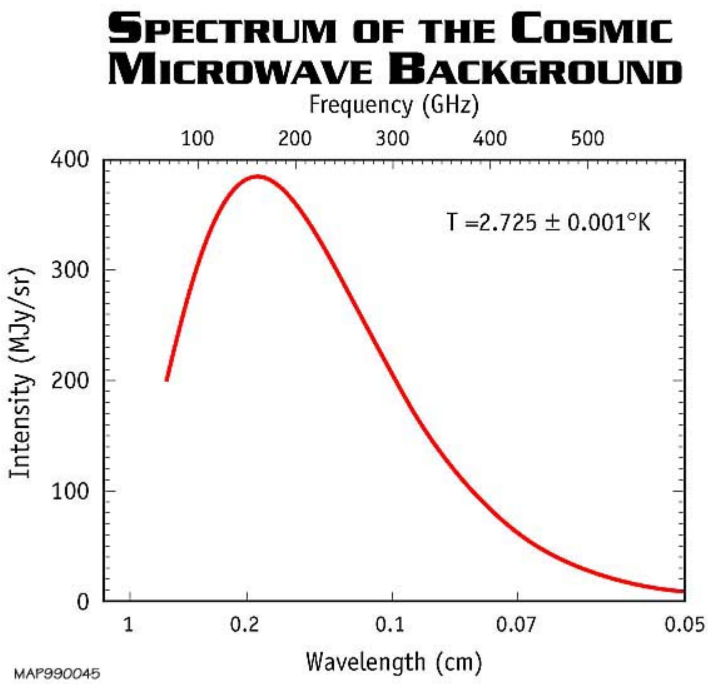
wavelength, and there seem to be a fairly good linear accordance between the expansion of space and the stretch of the photon wavelength from the early universe.

5.2.6 The Planck blackbody radiation formula

All bodies in thermodynamic equilibrium (at a temperature *T*) emit radiation given by Planck's law of radiation. This was actually the first quantum mechanical formula, where Planck introduced "Planck's constant" *h*, writing the energy of the photon as: $E = h\nu$, where ν is the frequency of the radiation.
The fraction $I(\nu)d\nu$ of radiation, which per unit volume of the material, is emitted in the frequency interval $d\nu$, is given by the formula (5.8), where *h* is Planck's constant $h = 4.15\ 10^{-15}\ eVs$, and *k* is Boltzmann's constant $k = 8{,}617\ 10^{-5}\ eV/K$.

(5.8)
$$I(\upsilon)d\upsilon = \frac{8\pi h\upsilon^3 d\upsilon}{c^3\left(e^{\frac{h\upsilon}{kT}} - 1\right)}$$

Below is shown the registered curve for the Cosmic Microwave Background radiation (CMB), which fits nicely into Planck's formula at a temperature of $T = 2.7\ K$. Fig (5.9)



6. The accelerating universe

As pointed out by Einstein, the universe could not be static and unchangeable, because the gravitational forces would eventually pull it together.

Having formulated his General Theory of Relativity, using Riemann's generalized tensor differential geometry, he realized that the gravitational contraction could be compensated for, by adding a constant, named the the *Cosmological Constant* multiplied by the metric form in his fundamental equation without changing the physical or mathematical (co-variance) consistency.

Later he declared that the cosmological constant, "was the greatest mistake in his life".
The cosmological constant represents the "vacuum-energy" of the universe. This constant energy density corresponds to a "negative pressure", giving rise to repulsive forces which increase with distance. A vacuum-energy dominated universe expands exponentially.

The inflationary theory of the cosmic origin, asserting that the universe has experimented a huge expansion at the earliest moment of the Big Bang, can provide the initial conditions for the standard model of cosmology, solving the flatness and horizon problems, and providing an explanation of the origin of energy/matter.

6.1 The cosmological constant

From the Newtonian theory, one would expect that expansion of the universe would slow down as a consequence of the gravitational attraction. However, there have never been observations that could corroborate that.
Before Hubble made his revolutionary discovery (At every point in space the galaxies move away, with a speed proportional to their distance. $v = H_0 r$) almost everyone believed that the universe was static.
You must remember that the discovery of the galaxies did not occur until the beginning of the 20'th century, and therefore the view of the universe consisted only of our own galaxy The Milky Way. The Milky Way was static in the same sense as the planetary system, but rotating around its centre because of gravitation.
But there exists no centre of the universe as a whole, about which the star and galaxies could rotate to maintain its stability, and therefore a static universe is (also from a Newtonian point of view) impossible.
The Friedman equation (5.3)

(6.1) $$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{c^2}\left(p + \tfrac{1}{3}\rho c^2\right) ,$$

is not compatible with the static condition of an unchanging scale factor, since $\dot{a} = \ddot{a} = 0$, would lead to an empty universe with $p = \rho = 0$.
The Einstein equation, which is the foundation of General Relativity

(6.2) $$G_{ij} = \kappa T_{ij} \qquad with \qquad \kappa = -\frac{8\pi G}{c^4}$$

Where G_{ij} is the Einstein curvature of space-time, and T_{ij} is the energy/momentum source of gravity.
In order to accomplish an equation which describe a static universe Einstein was led to alter his field equation such as it contained a repulsive component that became only potent at very large distances and increased with distance. The simplest choice that would preserve the consistency and covariance of Einstein's equation, could be a constant Λ (the cosmological constant) as a prefix to the metric form g_{ij} With this change the Einstein equation then reads.

(6.3) $$G_{ij} - \Lambda g_{ij} = \kappa T_{ij} \quad \textit{with} \quad \kappa = -\frac{8\pi G}{c^4}$$

Even if this formally can account for a static universe, there are several theoretical drawbacks.
Λ is invented only to compensate for the discrepancy between observation and experiment.
There exists no theory from which the existence or value of Λ can be predicted.
Even if there existed a quantum field description of gravity (and there doesn't) it is very unlikely that a cosmological constant would ever emerge from that theory.
Another problematic thing is that a cosmological constant will change the Newtonian limit, unless it is so small that it would be undetectable except at very large distances.

Although we have introduced the cosmological constant as an additional geometric term, we could as well move it to the other side of the equation and view it as an additional source of gravity.
In particular, when the regular energy-momentum is absent, that is, $T_{ij} = 0$, the Einstein equation becomes.

(6.4) $$G_{ij} = \Lambda g_{ij} \quad \text{where} \quad T_{ij}^{\Lambda} \equiv \frac{1}{\kappa}\Lambda g_{ij} = -\frac{\Lambda c^4}{8\pi G} g_{ij}$$

Then it can be interpreted as the energy–momentum tensor of the vacuum. In the same manner as T_{ij} for the cosmic fluid depends on two functions, the energy density ρ, and the pressure p, the vacuum energy–momentum tensor T_{ij}^{Λ}, can be similarly parametrized by the *vacuum energy density*, which is independent of time and the position in space, and the *vacuum pressure.*
We shall presently show, that these two quantities can be related to a positive cosmological constant Λ as follows:

(6.5) $$\rho_{\Lambda} = \frac{\Lambda c^2}{8\pi G}$$

And a negative pressure

(6.6) $$p_{\Lambda} = -\rho_{\Lambda} c^2$$

Although a negative pressure is impossible in classical mechanics, as well as in thermodynamics, it is not formally inconsistent with the adiabatic version of the first law of thermodynamics:

$$dE = -PdV.$$

Since when the energy density is constant

$$dE = \rho c^2 dV,$$

the pressure must necessarily be negative: $p = -\rho c^2$.

6.1.1 Gravitational repulsion caused by vacuum-energy

To understand why a negative pressure can give rise to a repulsive force, we shall first look into the Newtonian limit of the Einstein equation with a normal source composed of mass density ρ as well as pressure p.
We will show later, that in the Newtonian limit, the equation written in terms of the gravitational potential Φ, is

(6.7) $$\nabla^2\Phi = 4\pi G\left(\rho + 3\frac{p}{c^2}\right)$$

This equation informs us that not only mass, but also pressure can be the source of gravitational field.

If we also insert the vacuum contributions: $\rho = \rho_M + \rho_\Lambda$ and $p = p_M + p_\Lambda$
The Newton Poisson equation becomes.

(6.8) $$\nabla^2\Phi = 4\pi G\left(\rho_M + 3\frac{p_M}{c^2} + \rho_\Lambda + 3\frac{p_\Lambda}{c^2}\right)$$

$$\nabla^2\Phi = 4\pi G\left(\rho_M + 3\frac{p_M}{c^2} + \rho_\Lambda + 3\frac{-\rho_\Lambda c^2}{c^2}\right)$$

(6.9) $$\nabla^2\Phi = 4\pi G\rho_M - 8\pi G\rho_\Lambda = 4\pi G\rho_M - \Lambda c^2$$

Where we have used: $\rho_M >> 3\frac{p_M}{c^2} \approx 0$, and for the vacuum dominated case: $4\pi G\rho_M << \Lambda c^2$, so we end up with $\nabla^2\Phi = -\Lambda c^2$. The Laplace operator in spherical coordinates is

$$\nabla^2 \equiv \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin^2\phi}\frac{\partial^2}{\partial\theta^2} + \frac{1}{r^2\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial}{\partial\phi}\right).$$

Since $\Phi = \Phi(r)$ all the derivatives of θ and ϕ vanish, and we are left with.

$$\nabla^2\Phi = \frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\Lambda c^2 \quad \Leftrightarrow$$

$$\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\Lambda c^2 r^2$$

$$r^2\frac{d\Phi}{dr} = -\tfrac{1}{3}\Lambda c^2 r^3 \quad \Leftrightarrow \quad \frac{d\Phi}{dr} = -\tfrac{1}{3}\Lambda c^2 r \quad \Leftrightarrow$$

And by integrating:

(6.10) $$\Phi_\Lambda(r) = -\frac{\Lambda c^2}{6}r^2$$

Or simply and crude, by setting $x = y = z = r$, (because it really makes no difference in this context), so

$$\nabla^2\Phi = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = \frac{\partial^2\Phi}{\partial r^2} + \frac{\partial^2\Phi}{\partial r^2} + \frac{\partial^2\Phi}{\partial r^2} = 3\frac{\partial^2\Phi}{\partial r^2}$$

And integrating

$$3\frac{\partial^2\Phi}{\partial r^2} = -\Lambda c^2 \qquad \text{to give} \qquad \Phi_\Lambda(r) = -\frac{\Lambda c^2}{6}r^2$$

This potential gives rise to a repulsive force (per unit mass)

(6.11) $$g_\Lambda = -\nabla\Phi_\Lambda = \frac{\Lambda c^2}{3}\vec{r}$$

This should be compared to the Newtonian gravitational attraction $g = -G\frac{M}{r^3}\vec{r}$

6.1.2 The cosmological constant ensures a static universe

We shall then look back at the two Friedmann equations:.

(5.1) $$\frac{\dot{a}(t)^2}{a(t)^2} + \frac{kc^2}{R_0^{\ 2}a(t)^2} = \frac{8\pi G}{3}\rho$$

(5.2) $$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{c^2}\left(p + \tfrac{1}{3}\rho c^2\right)$$

And then by adding a non vanishing cosmological constant, they become

(6.12) $$\frac{\dot{a}(t)^2}{a(t)^2} + \frac{kc^2}{R_0^{\ 2}a(t)^2} = \frac{8\pi G}{3}(\rho_M + \rho_\Lambda)$$

(6.13) $$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{c^2}\left((p_M + p_\Lambda) + \tfrac{1}{3}(\rho_M + \rho_\Lambda)c^2\right)$$

The RHS of (6.13) does not necessarily need to be negative because of the negative pressure term $p_\Lambda = -\rho_\Lambda c^2$ Consequently a de-accelerating universe is no longer the ultimate outcome.
For non relativistic matter , and after setting $p_M = 0$ we have

(6.15) $$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left(\rho_M - 2\rho_\Lambda\right)$$

The static condition $\ddot{a} = 0$, leads to the constraint

(6.16) $$\rho_M = 2\rho_\Lambda = \frac{\Lambda c^2}{4\pi G}\ ,$$ (where we have used (6.5): $\rho_\Lambda = \frac{\Lambda c^2}{8\pi G}$)

The other static condition $\dot{a} = 0$ implies, using (6.12) the static condition $a = a_0 = 1$

(6.17) $$\frac{kc^2}{R_0^{\,2}} = 8\pi G \rho_\Lambda = \Lambda c^2$$

Since the RHS is positive, we must have $k = +1$, namely the static universe must have a positive curvature (a closed universe), with a scale factor (by solving (6.17) with respect to R_0.

(6.18) $$R_0 = \frac{1}{\sqrt{\Lambda}}$$

So, at the end of the day, our conclusion is, that the basic features of a static universe is, that the density and the radius are determined by an arbitrary input parameter.
Not only is this theoretically unsatisfactory, but worse the solution is unstable, so that even a small variation would cause it to deviate from its starting point.

7. Consequences of the Einstein equation

In section 3, we have derived and elaborated on some consequences of the Einstein equation

(7.1) $$G_{ij} = \kappa T_{ij} \quad \text{where } \kappa = -\frac{8\pi G}{c^4}$$

(7.2) $$R_{ij} - \tfrac{1}{2} R g_{ij} = \frac{8\pi G}{c^4} T_{ij}$$

Or when written in the equivalent form (3.12) $R_{ij} = \kappa (T_{ij} - \tfrac{1}{2} T g_{ij})$

(7.3) $$R_{ij} = \frac{8\pi G}{c^4}(T_{ij} - \tfrac{1}{2} T g_{ij})$$

7.1 The Schwarzschild solution

We shall now engage in solving the Einstein equation for a spherically symmetric source with total mass M.
The solution yields the metric function $g_{ij}(x)$ for the space-time geometry outside the source, and is called **the Schwarzschild exterior solution.** We have previously in (3.29) seen that a special spherical symmetric metric has only two unknown scalar functions.

(7.4) $$ds^2 = g_{00}(r,t) c^2 dt^2 + g_{rr}(r,t) dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\theta^2)$$

The objective is now to use Einstein's equation to solve it, to obtain expressions for $g_{00}(r,t)$ *and* $g_{rr}(r,t)$. The first step is to express the Ricci tensor elements R_{ij} in terms of the metric elements.We shall do this, by calculating the connections symbols based on the spherically symmetric form (7.4), but first we introduce some convenient notation:

(7.5) $$g_{00} = \frac{1}{g^{00}} \equiv -e^{\alpha} \quad , \quad g_{rr} = \frac{1}{g^{rr}} \equiv -e^{\beta}$$

So that the unknown functions are $\alpha(r,t)$ *and* $\beta(r,t)$. The mathematical procedure is a bit complex, so you may skip it.

We proceed to compute the Christoffel symbols via the Lagrange equation. In principle we may obtain the functions $\alpha(r,t)$ *and* $\beta(r,t)$ by differentiating the metric tensor, but a more efficient method is through interpreting the geodesic equation.

(7.6) $$\frac{d^2x^j}{ds^2} + \Gamma^j_{ik}\frac{dx^k}{ds}\frac{dx^i}{ds} = 0$$

The Euler Lagrange equation is:

(7.7) $$\frac{\partial L}{\partial x^i} - \frac{d}{d\tau}\frac{\partial L}{\partial \dot{x}^i} = 0$$

As demonstrated in “Differential Geometry 2”, we may choose the Lagrangian as:

(7.8) $$L = \tfrac{1}{2} g_{ij}\frac{dx^i}{d\tau}\frac{dx^j}{d\tau} = \tfrac{1}{2}\left[-e^{\alpha}\left(\frac{dx^0}{d\tau}\right)^2 + e^{\beta}\left(\frac{dr}{d\tau}\right)^2 + r^2\left(\frac{d\theta}{d\tau}\right)^2 + r^2\sin^2\theta\left(\frac{d\phi}{d\tau}\right)^2\right]$$

As long as the parameter is a natural parameter (e.g. the proper time τ.)

Once the geodesic equation is derived from (7.7), we can extract the value of Γ^i_{jk}, by comparing to the geodesic equation (7.6). Because we have:

$$\frac{\partial L}{\partial x^0} = \tfrac{1}{2}\left[-\dot{\alpha}e^{\alpha}\left(\frac{dx^0}{d\tau}\right)^2 + \dot{\beta}e^{\beta}\left(\frac{dr}{d\tau}\right)^2\right] \quad \text{and} \quad \frac{\partial L}{\partial \dot{x}^0} = -e^{\alpha}\left(\frac{dx^0}{d\tau}\right)$$

The $i = 0$ component of the Euler Lagrange equation (7.7) reads

$$\frac{\partial}{\partial \tau}\left[-e^{\alpha}\left(\frac{dx^0}{d\tau}\right)\right] - \tfrac{1}{2}\left[-\dot{\alpha}e^{\alpha}\left(\frac{dx^0}{d\tau}\right)^2 + \dot{\beta}e^{\beta}\left(\frac{dr}{d\tau}\right)^2\right] = 0$$

or

$$-e^{\alpha}\left[\frac{d^2x^0}{d\tau^2} + \frac{\partial\alpha}{\partial r}\frac{dr}{d\tau}\frac{dx^0}{d\tau} - \frac{\dot{\alpha}}{2}\left(\frac{dx^0}{d\tau}\right)^2 + \frac{\dot{\rho}}{2}e^{\beta-\alpha}\right] = 0$$

This is to be compared to the $i = 0$ component of the geodesic equation (7.6), where only the non vanishing $(dx^j/d\tau)(dx^k/d\tau)$ factors are displayed. It has the form:

$$\frac{d^2x^0}{d\tau^2}+2\Gamma^0_{r0}\frac{dr}{d\tau}\frac{dx^0}{d\tau}+\Gamma^0_{00}\left(\frac{dx^0}{d\tau}\right)^2+\Gamma^0_{rr}\left(\frac{dr}{d\tau}\right)^2=0$$

From which we may extract the result

(7.9) $$\Gamma^0_{r0}=\tfrac{1}{2}\frac{\partial\alpha}{\partial r}\ ,\quad \Gamma^0_{00}=-\tfrac{1}{2}\frac{\partial\alpha}{\partial\tau}=-\tfrac{1}{2}\dot{\alpha},,\quad \Gamma^0_{rr}=-\tfrac{1}{2}\dot{\beta}e^{\beta-\alpha}$$

Following a similar procedure, and using the symmetry properties of the Christoffel symbols.

(7.9) $$\Gamma^r_{00}=\tfrac{1}{2}\frac{\partial\alpha}{\partial r}e^{\alpha-\beta}\ ,\quad \Gamma^r_{rr}=-\tfrac{1}{2}\frac{\partial\beta}{\partial r}\ ,\quad \Gamma^r_{0r}=\tfrac{1}{2}\dot{\beta}$$

$$\Gamma^r_{00}=-re^{-\beta}\ ,\quad \Gamma^r_{\phi\phi}=-r\sin^2\theta e^{-\beta}\ ,\quad \Gamma^\theta_{\phi\phi}=-\sin\theta\cos\theta$$

$$\Gamma^\theta_{r\theta}=\frac{1}{r}\ ,\quad \Gamma^\varphi_{\phi\theta}=\cot\theta\ ,\quad \Gamma^\varphi_{r\theta}=\frac{1}{r}$$

In "Differential Geometry 2", the following expression for the Riemann tensor is derived.

(7.10) $$R^i_{jkl}=\frac{\partial\Gamma^i_{jl}}{\partial x^k}-\frac{\partial\Gamma^i_{jk}}{\partial x^l}+\Gamma^i_{mk}\Gamma^m_{jl}-\Gamma^i_{ml}\Gamma^m_{jk}$$

We can then use (7.10) to calculate the curvature tensor R^i_{jkl}, from which we can contract the indices of R^k_{jkl} to form the Ricci tensor.

$$R_{00}=-\left(\tfrac{1}{2}\frac{\partial^2\alpha}{\partial r^2}+\tfrac{1}{4}\left(\frac{\partial\alpha}{\partial r}\right)^2-\tfrac{1}{4}\frac{\partial\alpha}{\partial r}\frac{\partial\beta}{\partial r}+\frac{1}{r}\frac{\partial\alpha}{\partial r}\right)e^{\alpha-\beta}+\left(\tfrac{1}{2}\ddot{\beta}+\tfrac{1}{4}\dot{\beta}^2-\tfrac{1}{4}\dot{\alpha}\dot{\beta}\right)$$

(7.11) $$R_{rr}=\left(\tfrac{1}{2}\frac{\partial^2\alpha}{\partial r^2}+\tfrac{1}{4}\left(\frac{\partial\alpha}{\partial r}\right)^2-\tfrac{1}{4}\frac{\partial\alpha}{\partial r}\frac{\partial\beta}{\partial r}+\frac{1}{r}\frac{\partial\beta}{\partial r}\right)-\left(\tfrac{1}{2}\ddot{\beta}+\tfrac{1}{4}\dot{\beta}^2-\tfrac{1}{4}\dot{\alpha}\dot{\beta}\right)e^{\beta-\alpha}$$

$$R_{0r}=-\frac{\dot{\beta}}{r}$$

$$R_{\theta\theta}=\left(1+\tfrac{1}{2}r(\frac{\partial\alpha}{\partial r}-\frac{\partial\beta}{\partial r}\right)e^{-\beta}-1$$

$$R_{\phi\phi}=R_{\theta\theta}\sin^2\theta$$

Until now our calculations have been based only on what the spherical symmetry impose on the solution. Now we shall seek the actual solution to Einstein's field equations.

7.1.1 The Einstein equation for the space-time exterior to the source

Our aim is now to find the metric in the region outside a spherically symmetric source. Because the energy momentum tensor T_{ij} vanishes in the exterior the Einstein Field equation simply becomes.

(7.12) $$R_{ij} = 0$$

Although (7.12) looks simple, the Ricci tensor is actually a set of 16 second order non linear differential operators acting on the metric functions.

Before getting the solution for the two unknown functions

(7.13) $$g_{00}(r,t) = e^{-\alpha(r,t)} \quad and \quad g_{rr}(r,t) = e^{-\beta(r,t)}$$

We state that the solutions must necessarily be time independent, (as a consequence of the so called Birkhoff theorem, on which we relax the proof), and also since the alternative certainly would be strange.

(7.14) $$\alpha(r,t) = \alpha(r) \quad and \quad \beta(r,t) = \beta(r)$$

After substituting in (7.11) this condition, all the time derivatives terms vanish. $\dot{\alpha} = \dot{\beta} = \ddot{\beta} = 0$, and the Einstein vacuum equation (7.11), yield three component equations.

$$R_{00} = 0 \;\Rightarrow\; \tfrac{1}{2}\frac{\partial^2\alpha}{\partial r^2} + \tfrac{1}{4}\left(\frac{\partial\alpha}{\partial r}\right)^2 - \tfrac{1}{4}\frac{\partial\alpha}{\partial r}\frac{\partial\beta}{\partial r} + \frac{1}{r}\frac{\partial\alpha}{\partial r} = 0$$

(7.14) $$e^{\beta-\alpha}R_{00} + R_{rr} = 0 \;\Rightarrow\; \frac{\partial\alpha}{\partial r} + \frac{\partial\beta}{\partial r} = 0$$

(7.15) $$R_{\theta\theta} = 0 \;\Rightarrow\; \left(1 + \tfrac{1}{2}r(\frac{\partial\alpha}{\partial r} - \frac{\partial\beta}{\partial r}\right)e^{-\beta} - 1 = 0$$

Actually one of these equations is redundant. It can be shown that the solution from two of the equations automatically satisfies the third.

7.1.2 Solving the Einstein equations for a spherically symmetric source

We then carry out the solution to (7.14) and (7.15). After integrating (7.14) over *r*, and setting the integration constant to zero we find.

(7.16) $$\alpha(r) = -\beta(r)$$

Because α and β are exponent of the metric scalar function

$$g_{00} = -e^{\alpha} \quad , \quad g_{rr} = -e^{\beta} = -e^{-\alpha}, \quad \text{it follows}$$

(7.17) $$-g_{00} = \frac{1}{g_{rr}}$$

Inserting (7.14) into (7.15) we get:

(7.18) $$(1-r\frac{\partial\beta}{\partial r})e^{-\beta}-1=0$$

To solve this equation, we introduce a new variable λ.

(7.19) $$\lambda(r)=e^{-\beta(r)} \quad and \quad \frac{d\lambda}{dr}=-\frac{d\beta}{dr}e^{-\beta(r)}$$

So that (7.18) becomes:

(7.20) $$\frac{d\lambda}{dr}+\frac{\lambda}{r}=\frac{1}{r}$$

It has the solution $\lambda(r)=\lambda_0(r)+\lambda_1$, where λ_0 is the solution to the homogenous equation:

(7.21) $$\frac{d\lambda_0}{dr}=-\frac{\lambda_0}{r}$$

And (7.21) has the solution

(7.22) $$\ln\lambda_0=-\ln r+k_0 \quad \Leftrightarrow \quad \lambda_0 r=k_0$$

Thus the product $\lambda_0 r$ is a constant which we label $-r_S$ (*S* for Schwarzschild)

(7.23) $$\lambda_0 r=-r_S$$

Combining this with the particular solution of $\lambda_1 = 1$, we have, (remembering the definition of α and β:

$$g_{00}(r,t)=e^{-\alpha(r,t)} \quad and \quad g_{rr}(r,t)=e^{-\beta(r,t)}$$

and (3.34)

$$g_{00}(t,r)=-\frac{1}{g_{rr}(t,r)}=-1+\frac{r_S}{r},$$

and noticing that the λ function is just $1/g_{rr}$, yields the general solution.

(7.24) $$\lambda=1-\frac{r_S}{r}=\frac{1}{g_{rr}}=-g_{00}$$

The solution is called the Schwarzschild metric for the exterior of a spherical symmetric source.

(7.25) $$g_{ij}=diag((-1+\frac{r_S}{r}),(1-\frac{r_S}{r})^{-1},r^2,r^2\sin^2\theta)$$

The constant r_S is related (by taking the Newtonian limit) to the gravitational constant, and the mass of the source (3.38).

(7.26) $$g_{00} = -\left(1 + \frac{\Phi(x)}{c^2}\right)^2 \approx -\left(1 + \frac{2\Phi(x)}{c^2}\right)$$

(7.27) $$g_{00}(t,r) = -1 + \frac{r_S}{r} = -1 - \frac{2\Phi(x)}{c^2} \quad and \quad \Phi(x) = -G\frac{M}{r}$$

(7.28) $$-1 + \frac{r_S}{r} = -1 - \frac{2GM}{c^2 r} \quad \Rightarrow \quad r_S = \frac{2GM}{c^2}$$

In section 3, we have already discussed the consequences of the Schwarzschild metric concerning bending of light beams and Black Holes.

7.1.3 The Einstein equation for cosmology

After introducing his theory in 1915 the study of cosmology has been carried out within the framework of General Relativity. The foundation is Einstein's equation

(7.29) $$R_{ij} = \frac{8\pi G}{c^4}(T_{ij} - \tfrac{1}{2}Tg_{ij})$$

The solution for cosmology must comply with the cosmological principle of a homogenous and isotropic space, which has the solution of the Robertson-Walker metric.
The model of the universe is that of cosmic fluid. The 4D metric in this co-moving coordinate system, has the form as (4.71)

(7.30) $$ds^2 = -c^2dt^2 + g_{ij}(r,t)dx^i dx^j \qquad (i, j) = 1..3$$

We put $dl^2 = g_{ij}(r,t)dx^i dx^j$

$$ds^2 = -c^2dt^2 + dl^2$$

$$ds^2 = -c^2dt^2 + R(t)^2 d\hat{l}^2$$

Where the reduced length element $d\hat{l}$ is both time independent and dimensionless, and

(7.31) $$a(t) = \frac{R(t)}{R_0}$$

Normalized so that $a(t_0) = 1$. R_0 is often referred to as the radius of the universe.
We have previously argued that a universe complying with the cosmological principle must have a constant curvature. We extended in (4.74) (without a formal proof) the metric of a 2D constant curvature surface to 3D constant curvature space.

(7.32) $$d\hat{l}^2 = \left(\frac{d\varsigma^2}{1-kd\varsigma^2} + \varsigma^2 d\Omega^2\right) \quad \text{where} \quad d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$$

Where ς is the dimensionless radial distance. This is the Robertson Walker metric.
Here we shall make an alternative derivation, based directly on Einstein's equation, using the same steps, as in arriving at the Schwarzschild solution.

Homogeneity and isotropy mean that the space must be spherically symmetric, as seen from every point in space (7.4).

(7.33) $$ds^2 = g_{00}(r,t)c^2dt^2 + g_{rr}(r,t)dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)$$

This means that the metric for the 3D space $(\varsigma, \theta, \phi)$, should have the form as outlined in section (4.13.2). Keeping only the spatial part, we have

(7.34) $$d\hat{l}^2 = \hat{g}_{\varsigma\varsigma}d\varsigma^2 + \varsigma^2(d\theta^2 + \sin^2\theta d\phi^2)$$

As mentioned earlier the metric element $\hat{g}_{\varsigma\varsigma}$ is independent of time, and we shall follow the previous notation and put $\hat{g}_{\varsigma\varsigma} = e^{\beta(r)}$.
To elaborate further on homogeneity and isotropy, we demand that the Ricci scalar be constant.

(7.35) $$R^{(3D)} = -6k.$$

This does actually imply that the 3D space should be with constant curvature.
Looking up the Ricci tensor (7.11), and after setting:

$$\alpha = \dot{\alpha} = \frac{\partial\alpha}{\partial r} = \frac{\partial^2\alpha}{\partial r^2} = \dot{\beta} = \ddot{\beta} = 0 \ ,$$

we obtain the Ricci tensor elements for the 3D-space.

(7.36) $$R^{3D}_{\varsigma\varsigma} = -\frac{1}{\varsigma}\frac{d\beta}{d\varsigma}, \qquad R^{3D}_{00} = \left(1 - \frac{\varsigma}{2}\frac{d\beta}{d\varsigma}\right)e^{-\beta} - 1, \qquad R^{3D}_{\phi\phi} = \sin^2\theta R^{3D}_{\theta\theta}$$

Which is to be contracted with the inverse metric $\hat{g}^{ij}$ of (7.34).

(7.37) $$\hat{g}^{\varsigma\varsigma} = e^{-\beta(r)}, \quad \hat{g}^{\theta\theta} = \varsigma^{-2}, \quad \hat{g}^{\phi\phi} = \frac{1}{\varsigma^2\sin^2\theta}$$

to obtain the Ricci scalar:

(7.38) $$R^{3D} = \sum R^{3D}_{ii}\hat{g}^{ii} = R^{3D}_{\varsigma\varsigma}\hat{g}^{\varsigma\varsigma} + 2R^{3D}_{\theta\theta}\hat{g}^{\theta\theta}$$

$$R^{3D} = -\frac{e^{-\beta}}{\varsigma}\frac{d\beta}{d\varsigma} + \frac{2}{\varsigma^2}\left(\left(1 - \frac{\varsigma}{2}\frac{d\beta}{d\varsigma}\right)e^{-\beta} - 1\right)$$

Setting it to -6*k* as in (7.35), and rearranging:

(7.39)
$$\frac{2}{\varsigma^2}\frac{d}{d\varsigma}(\varsigma e^{-\beta} - \varsigma) = -6k$$

$$d(\varsigma e^{-\beta} - \varsigma) = -3k\varsigma^2 d\varsigma$$

It can be solved by integrating.

(7.40)
$$(1 - e^{-\beta})\varsigma = k\varsigma^3 + C$$

Where the constant C = 0, as can be seen in the $\varsigma = 0$ limit. We obtain the desired solution:

(7.41)
$$\hat{g}_{\varsigma\varsigma} = e^{\beta(\varsigma)} = \frac{1}{1 - k\varsigma^2}$$

In accordance with (7.32).

$$d\hat{l}^2 = \left(\frac{d\varsigma^2}{1 - kd\varsigma^2} + \varsigma^2 d\Omega^2\right),$$

This was obtained earlier by comparing with the two dimensional case.

7.1.4 The Friedmann equations

Her we shall explicate the exact relation between the Einstein and the Friedmann equations.
The Einstein equation:

(7.42)
$$G_{ij} = \kappa T_{ij} \quad \text{where} \quad \kappa = -\frac{8\pi G}{c^4}$$

For the homogenous and isotropic universe, the LHS is determined by the Robertson–Walker metric, with its two parameters: The curvature constant k , and the scale factor $R(t)$.

What is necessary is to specify the energy momentum tensor.
The simplest (and the only palatable) choice is to take the cosmic fluid as an ideal fluid, as discussed earlier in (5.1) to (5.3).

In Special Relativity the energy momentum tensor is given by:

(7.43)
$$T_{ij} = pg_{ij} + (\rho + \frac{p}{c^2})u_i u_j$$

Where p is the pressure, ρ is the mass density and u_i are the Minkowski 4-vectors. Since there are no derivatives, this expression must also hold in General Relativity.

In the cosmic rest frame, that is, the co-moving coordinates, in which each of the fluid elements (galaxies) carries it own position label.
All the fluid elements are at rest in such a frame, with the metric given by $g_{ij} = diag(-1, g_{ij})$, and (i,j) =1..3. , the energy momentum tensor takes on a particular simple form.

(7.44)
$$T_{ij} = \begin{pmatrix} \rho c^2 & 0 \\ 0 & p g_{ij} \end{pmatrix}$$

So the cosmological Friedmann equations are just the Einstein equation with the Robertson – Walker metric with the ideal fluid energy momentum tensor.

The $G_{00} = -8\pi G / c^2$ equation can then be written (after a lengthy calculation) in terms of the Robertson–Walker metric in terms of the curvature constant k and the scale factor $R(t)$.
The first Friedmann equation (5.1):

(7.45)
$$\frac{\dot{a}(t)^2}{a(t)^2} + \frac{kc^2}{R_0{}^2 a(t)^2} = \frac{8\pi G}{3}\rho \quad \text{or} \quad \frac{\dot{R}(t)^2 + kc^2}{R(t)^2} = \frac{8\pi G}{3}\rho, \quad \text{since } R(t) = a(t)R_0$$

From the equation: $G_{ij} = -8\pi G g_{ij} / c^4$, we get the second Friedmann equation:

(7.46)
$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{c^2}\left(p + \tfrac{1}{3}\rho c^2\right) \quad \text{or} \quad \frac{\ddot{R}(t)}{R(t)} = -\frac{4\pi G}{c^2}\left(p + \tfrac{1}{3}\rho c^2\right)$$

Although the Friedmann equations have a rather simple Newtonian interpretation, they must be understood in terms of General Relativity, as they still have the geometric concepts of 4D metric and curvature.

7.1.5 Einstein's equations with a cosmological term

As mentioned (several times), that until the mid 1950'ties the widespread view was that the universe was infinite and static, but as Einstein, among others, pointed out, this was not in accordance with universal gravitational attraction.
The universe would eventually begin to contract and end in a "Big Crunch".

Looking only at our own galaxy, the Milky Way, there was no problem with stability, because it is in rotation in a gravitational field like the planetary system.

But why the galaxies in 13.6 billion years had not come closer to each other was a mystery and it still is, by the way.
This caused Einstein to modify his equations for General Relativity to include a constant, called the cosmological constant (The greatest mistake in my life, he wrote some years later).
But nevertheless the cosmological constant is still spooking in scientific literature, although no one has ever been able to give a theoretical explanation of the cosmological constant or calculate its value from first principle.

What Einstein realized was that it was possible to modify his equation, without loss of consistency, by simply adding a constant Λ multiplied by the metric form to uphold its tensor nature (covariance).

(7.47) $$G_{ij} - \Lambda g_{ij} = \kappa T_{ij} \qquad with \qquad \kappa = -\frac{8\pi G}{c^4}$$

One serious problem is that the introduction of a cosmological constant will alter the Newtonian limit, but this can be avoided if it can be totally neglected, except at huge distances.
If we move the term with Λ to the other side:

(7.48) $$G_{ij} = \kappa(T_{ij} + \kappa^{-1}\Lambda g_{ij})$$

This allows for a new interpretation of $\kappa^{-1}\Lambda g_{ij}$, namely as an energy momentum tensor, which we call the "**vacuum energy tensor**". Since in the absence of ordinary mass/energy distribution $T_{ij} = 0$ (vacuum) The source term

$$T_{ij}^{\Lambda} \equiv \frac{1}{\kappa}\Lambda g_{ij} = -\frac{\Lambda c^4}{8\pi G} g_{ij}$$

can however still bring about a gravitational field.

This vacuum energy-momentum tensor can still be written in a form analogous to the conventional ideal fluid stress tensor.

(7.49) $$T_{ij}^{\Lambda} = \begin{pmatrix} \rho_\Lambda c^2 & 0 \\ 0 & p_\Lambda g_{ij} \end{pmatrix}$$

Comparing with (7.44)

$$T_{ij} = \begin{pmatrix} \rho c^2 & 0 \\ 0 & p g_{ij} \end{pmatrix},$$

It implies a constant vacuum energy density.

(7.50) $$\rho_\Lambda = -\frac{\Lambda}{\kappa c^2} = \frac{\Lambda c^2}{8\pi G}$$

If we take $\Lambda > 0$, so that $\rho_\Lambda > 0$, it further implies a negative vacuum pressure:

(7.51) $$p_\Lambda = -\rho_\Lambda c^2$$

Thus the cosmological constant corresponds to an energy density, which is constant in time and space. No matter how we change the volume the energy density, it is unchanged.

As we have discussed earlier *this pressure* is the source of gravitational repulsion on huge distances, which might haven driven the inflationary epoch of the Big Bang, and being sufficient to give the universe an accelerated expansion, as recent observations points toward.

References

Ta-Pei-Cheng: Relativity, Gravitation and Cosmolygy.
Young and Freedman: University Physics 12'th edition.
Landau and Liffshitz: The Classical Theory of Fields. 1975