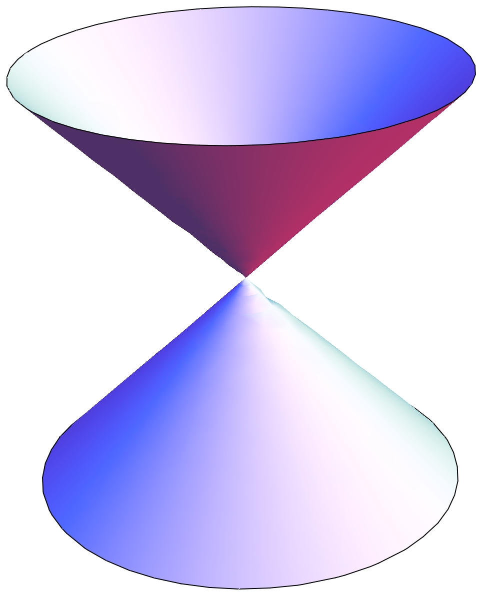
Gravitational physics

D.D.Kodwani

Light cone



Last updated on July 28, 2014

Contents

Preface

Introduction

Relativity, along with quantum mechanics, is one of the pillars on which all of modern physics is built upon. Some, like Pedro Ferreira, argue that general relativity is "the perfect theory" [**1**], however there certainly are points where it breaks down and the fact that it is incompatible with Quantum (field) theory means that even if it is the perfect theory, it requires some modifications.

General and special relativity provide mind blowing insights into our universe. Even today, almost a century on from when it was first discovered, general relativity continues to throw the cat among the pigeons from time to time, the most recent being the discovery of gravitational waves from the recent BICEP 2 experiment at the south pole, which provide an insight into the theory of inflation that we will come to in future chapters.

The notes are structured into three parts. The first part is aimed to provide a detailed look at the formulation of relativity theory, both general and special. The starting point for relativity was Maxwell's equations of electro-magnetism, which suggest a constant speed of propagation for electro-magnetic waves, which we know commonly refer to as the speed of light, c. It was Einstein's genius, that took this simple looking fact and transformed our intuition of space-time and everything in it. Part 2 goes into hardcore gravitational physics, much of which is used in present day research. Finally, part 3 gives a brief sketch of cosmology and it is this part that is most relevant for the summer project on inflation. Detailed calculations of the power spectrum and quantum field theory in curved space-time are provided. These go along with the summary report produced as part of this project.

> *"Physics is not a finished logical system. Rather, at any moment it spans a great confusion of ideas, some that survive like folk epics from the heroic periods of the past, and others that arise like utopian novels from our dim premonitions of a future grand synthesis"* Steven Weinberg.

Notation

- The operator ∇ is defined as:

$$\nabla = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z} \tag{0.1}$$

- Einstein summation convention is assumed, where all repeated indicies are summed over:

$$\sum_{\mu=0}^{4} x^\mu x_\mu \equiv x^\mu x_\mu = x^0 x_0 + x^1 x_1 + x^2 x_2 + x^3 x_3 \tag{0.2}$$

- Co-ordinates are generally defined as:

$$\begin{aligned}\mu &= 0 \equiv t \\ \mu &= 1 \equiv x \\ \mu &= 2 \equiv y \\ \mu &= 3 \equiv z \end{aligned} \tag{0.3}$$

- The *Kronecker* delta symbol, δ, is defined as:

$$\delta_{\mu\nu} = \begin{cases} 1 & \text{if } \mu = \nu \\ 0 & \text{if } \mu \neq \nu \end{cases} \tag{0.4}$$

- The *Levi-Civita* symbol, $\epsilon_{ijkl...n}$, in n dimensions is defined as:

$$\epsilon_{i,j,k,l...n} = \begin{cases} 1 & \text{if } (i,j,k,l...n) \text{ is an even permutation of (1,2,3,4...n)} \\ 0 & otherwise \\ -1 & \text{if } (i,j,k,l...n) \text{ is an odd permutation of (1,2,3,4...n)} \end{cases} \tag{0.5}$$

- The identity matrix is symbolized by $\mathbb{I}$, in the dimension that should be obvious from the expression it is in relation with.
- $\mathbb{R}^n$ stands for an n dimensional Euclidean space.
- If an object is infinitely differentiable on a subset of a given field, it is denoted by C^∞. In other words, if a function, f, is parametrised by λ and is C^∞:

$$\exists \frac{d^n}{d\lambda} f(\lambda) \; \forall n \in \{0, 1, ...\infty\} \tag{0.6}$$

In general, *Greek* indicies run over 0,1,2,3 and *Latin indices* run over 1,2,3, unless stated otherwise.

There is a slight change in notation for part 2. Firstly the metric has the signature +,-,-,- as there will be aspects of particle physics here and this is the convention used in particle physics. Moreover, we work in units:

$$\hbar \equiv c \equiv 1 \tag{0.7}$$

from part 2 onwards, although at certain points in part 1 these notations are also used.

Part 1

Maxwell to Einstein: Laying the foundations of relativity

CHAPTER 1

Special relativity

1. A close look at Maxwell's equations

Let's begin with James Clerk Maxwell's discoveries in 1860's. Maxwell's discovery was to unify the laws of electricity and magnetism, which at the time seemed like two different things. This discovery has proved to one of the most important of mankind, not just in physics, but also in technology. Almost every appliance we use today is linked in some way to Maxwell's equations.

They are simple looking differential equations that carry beauty and symmetry between them in equal measure, so lets start off by writing them down in their original form (before Maxwell's key discovery):

(1) The first equation is one that is attributed to *Carl Friedrich Gauss*:

$$\nabla.\vec{E} = \frac{\rho}{\epsilon_0} \tag{1.1}$$

where ρ is the charge density, ϵ_0 is the permittivity of free space in a vacuum, $\vec{E}$ is the electric field. This equation states that the divergence of the electric field, is proportional to the density of the electric charge, with the constant of proportionality being $\frac{1}{\epsilon_0 \times \text{Volume}}$. Or in other words, the electric charge is the source of the electric field.

(2) This equation is not attributed to any person, it is simply an experimental fact that has been observed for years:

$$\nabla.\vec{B} = 0 \tag{1.2}$$

This equation actually carries a very puzzling fact of our universe. It states that the divergence of a magnetic field is *always* zero, which is a statement of the fact that no magnetic charges (magnetic monopoles) are ever observed in nature. There is no fundamental reason for this that we know of, thus it remains a big puzzle in physics today.

(3) Faraday is attributed with this equation, even though he never wrote it down, he was the first one to discover its effect in his famous experiments:

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{1.3}$$

This equation states that if there is a time varying magnetic field, it can also act as a source for an $\vec{E}$ field, or more specifically, the curl of the $\vec{E}$ field i.e the $\vec{E}$ field rotates around the $\vec{B}$ field.

(4) The final equation was first outlined by Ampere:

$$\nabla \times \vec{B} = -\mu_0 \vec{J} \tag{1.4}$$

where μ_0 is the permeability of free space. $\vec{J}$ is a current density and it acts as a source of the magnetic field.

Maxwell's genius was to realise that all these equations were describing the same phenomena, but not in the form they are currently in. He realised that these equations seemed to contradict the fact that was well established (by Benjamin Franklin) that electric charge is always conserved. More specifically, the equations did *not* obey:

$$\frac{d\rho}{dt} = -\nabla.\vec{J} \tag{1.5}$$

Which of course means if a electric charge is depleted it must move away from its position, so the fact that this is not obeyed is really a big problem! The problem can be spotted in Eq 1.4, to see how, lets take the divergence of both sides:

$$\nabla.(\nabla \times \vec{B}) = -\mu_0 \nabla.\vec{J} \tag{1.6}$$

The curl can be written as:

$$(\nabla \times \vec{B})_i = \epsilon_{ijk} \frac{\partial}{\partial x^j} B_k \tag{1.7}$$

So we can re-write the R.H.S of Eq 1.6 as:

$$\begin{aligned}
\nabla.(\nabla \times \vec{B}) &= \frac{\partial}{\partial x^i} \epsilon_{ijk} \frac{\partial B_k}{\partial x^j} \\
&= \epsilon_{ijk} \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} B_k \\
&= \epsilon_{1jk} \frac{\partial}{\partial x^1} \frac{\partial}{\partial x^j} B_k + \epsilon_{2jk} \frac{\partial}{\partial x^2} \frac{\partial B_k}{\partial x^j} + \epsilon_{3jk} \frac{\partial}{\partial x^3} \frac{\partial B_k}{\partial x^j} \\
&= \epsilon_{123} \frac{\partial}{\partial x^1} \frac{\partial}{\partial x^2} B_3 + \epsilon_{132} \frac{\partial}{\partial x^1} \frac{\partial}{\partial x^3} B_2 + \epsilon_{213} \frac{\partial}{\partial x^2} \frac{\partial}{\partial x^1} B_3 + \epsilon_{231} \frac{\partial}{\partial x^2} \frac{\partial}{\partial x^3} B_1 \\
&+ \epsilon_{312} \frac{\partial}{\partial x^3} \frac{\partial}{\partial x^1} B_2 + \epsilon_{321} \frac{\partial}{\partial x^3} \frac{\partial}{\partial x^2} B_1
\end{aligned} \tag{1.8}$$

Since ϵ_{ijk} is totally anti-symmetric and the fact that partial derivatives are commutative, it is obvious that the terms on the R.H.S of Eq 1.8 cancel. Therefore Eq 1.6 gives:

$$\nabla.\vec{J} = 0 \tag{1.9}$$

Which contradicts Eq 1.5. Maxwell then wanted to remove this inconsistency, and the obvious way to do this is to work backwards and assume Eq 1.5 holds and try to find what the $\frac{d\rho}{dt}$ term is from Eq 1.1, which is:

$$\epsilon_0 \frac{\partial}{\partial t} \nabla.\vec{E} \tag{1.10}$$

Now it is not hard to see that adding:

$$\epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t} \tag{1.11}$$

to Eq 1.4, will remove this inconsistency. This was Maxwell's great isnsight and this term is often referred to as Maxwell's displacement current. The equations are now ready to be written in their world renown form:

$$\nabla.\vec{E} = \frac{\rho}{\epsilon_0} \tag{1.12}$$

$$\nabla.\vec{B} = 0 \tag{1.13}$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \tag{1.14}$$

$$\nabla \times \vec{B} = \mu_0 \vec{J} + \epsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t} \tag{1.15}$$

It is easy for us to see how this inconsistency was removed with a few simple steps now, however it took Maxwell (who was probably the greatest mind of his generation) years to work out (in the process of developing these equations he had to develop the entire field of vector calculus and also introduce partial differential equations). Hindsight is a wonderful thing!

It was this last term that Maxwell added that leads to the fact that these equations predict a solution that takes on the form of a wave. To see how, let's take a simple system, in which there are no charges or currents:

$$\rho = \vec{J} = 0 \tag{1.16}$$

Under these conditions, take the curl of Eq 1.15:

$$\begin{aligned}
\nabla \times (\nabla \times \vec{B}) &= \epsilon_0 \mu_0 \left(\nabla \times \frac{\partial \vec{E}}{\partial t} \right) \\
\nabla \times \left(\epsilon_{klm} \frac{\partial}{\partial x^l} B_m \right) &= \epsilon_0 \mu_0 \left(\epsilon_{klm} \frac{\partial}{\partial x^l} E_m \right) \\
\epsilon_{ijk} \frac{\partial}{\partial x^j} \epsilon_{klm} \frac{\partial}{\partial x^l} B_m &= \epsilon_0 \mu_0 \epsilon_{klm} \frac{\partial}{\partial x^l} \left(\frac{\partial E_m}{\partial t} \right) \\
\frac{\partial}{\partial x^j} \frac{\partial}{\partial x^l} \epsilon_{ijk} \epsilon_{klm} B_m &= \epsilon_0 \mu_0 \epsilon_{klm} \frac{\partial}{\partial x^l} \left(\frac{\partial E_m}{\partial t} \right)
\end{aligned} \tag{1.17}$$

There is an identity that needs to be used now:

$$\epsilon_{ijk} \epsilon_{klm} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} \tag{1.18}$$

Using this, Eq 1.17 can be written as:

$$-\nabla^2 B = \epsilon_0 \mu_0 \frac{\partial}{\partial t} \left(\nabla \times \vec{E} \right) \tag{1.19}$$

Now substitute from Eq 1.14 for $\left(\nabla \times \vec{E} \right)$:

$$\nabla^2 \vec{B} = \epsilon_0 \mu_0 \frac{\partial^2 \vec{B}}{\partial t^2} \tag{1.20}$$

Similarly, taking the curl of Eq 1.14 and following the same procedure of finding $\nabla \times \vec{B}$ from Eq 1.15 to get:

$$\nabla^2 \vec{E} = \epsilon_0 \mu_0 \frac{\partial^2 \vec{E}}{\partial t^2} \tag{1.21}$$

These are both *wave equations*. As an example suppose the magnetic field is a function of time and only the x spatial dimension:

$$\frac{\partial^2 \vec{B}}{\partial t^2} = \frac{1}{\epsilon_0 \mu_0} \frac{\partial^2 \vec{B}}{\partial x^2} \tag{1.22}$$

The solution to this equation is any function of $(x - ct)$, where c takes the form of a velocity, and since the equations are linear we can also form superposition functions of the form:

$$\vec{B} = F_1(x - ct) + F_2(x + ct) \tag{1.23}$$

Substituting this solution into Eq 1.22 yields the relation:

$$c^2 = \frac{1}{\epsilon_0 \mu_0} \tag{1.24}$$

This function, is a function whose shape is fixed, but its center moves, which are just waves! So Maxwell found that these were waves of electro-magnetism that traveled at a fixed speed that depends on two of the fundamental constants of nature ϵ_0, μ_0. These had already been measured in experiments and Maxwell put the experimentally measured values in to obtain a value for c as $3 \times 10^8\ ms^{-1}$. Which is just the speed of light(which had also been measured in experiment before). This was Maxwell's true genius; using only pen and paper (and his brain!), he followed the mathematics to find that light, which nobody knew what it really was, is infact an electromagnetic(EM) wave. This solution also predicts that there should be waves of any wavelength that move with this speed, so infact we had discovered the whole spectrum of EM waves.

If that was not enough, the solution to Maxwell's equations tells us something more about the waves aswell, and that is seen by Eq 1.12 and Eq 1.13. For example, if we put the solution in Eq 1.23 into Eq 1.13, it is seen that the wave cannot change in the x direction, it is only the y and z components that can vary. which in essence tells us that the waves are *transverse*.

2. Symmetry of Maxwell's equations

2.1. Galilean transformations. As with most of our great theories, there is a lot more to them then first meets the eye(and it usually takes more people, other than the inventor, to fully appreciate its consequences) and the same is true for Maxwell's equations. The equations posses a lot of symmetries, some are more obvious then others. An obvious one is the dependence (or independence) of *time translations*.

Making a transformation of the form:

$$t \to t + t' \qquad \text{where } \ t' \in \{\text{Constant}\} \tag{1.25}$$

leaves Maxwell's equations invariant, as there is no explicit time dependence in them, just derivatives of time, so any time independent quantity will not affect the equations of motion. Another symmetry that is immediately obvious is that of *spatial translations*, consider the transformation:

$$\vec{x} \to \vec{x} + \vec{x}' \qquad \text{where } \ x' \in \{\text{Constant}\} \tag{1.26}$$

this also does not affect the EM equations as they only involve derivatives of $\vec{x}$. So there are four (x, y, z, t) independent translations under which Maxwell's equations are invariant. These are expected from Galilean transformations as we do not expect the physics to change depending on what coordinates are being used.

2.2. End of Newtonian mechanics. Maxwell's equations do something very strange that no other theory upto its time (or even since) have done. It predicts a speed. This seems totally counter-intuitive and created a big problem for physics at that time (although physicists did not realise this at that time), as from Newtonian mechanics, speed is always relative. The speed of an object never determines the physics of its system, it is always relative to something. Yet Maxwell's equations do not show any relative motion, or any sight of a frame of reference, it is a the same speed in all frames of reference. Maxwell realised this of course, but it was something he did not take further. It was not until Einstein, did the true consequences of this bizarre fact came to the surface.

In Newtonian mechanics, there is implicitly a symmetry built into the equations, that if you change a system into a coordinate system in an inertial (non-accelerating) frame of reference, then the laws of physics don't change; i.e making a transformation:

$$x \to x + \dot{x}t \qquad \text{where } \ \dot{x} \in \{\text{Constant}\} \tag{1.27}$$

doesnt change anything. The reason for this is obvious, and it is the fact that Newton's equation:

$$F = m\ddot{x} \tag{1.28}$$

only involves second derivatives of time.

The whole idea behind symmetry and laws of physics (Noether's theorem etc), begins from the symmetry hidden in Maxwell's equations. To see these symmetries we have to re-write the equations in tensor formalism, which unifies Maxwell's equations further.

2.3. Tensor formalism. Begin by defining four vectors:

$$x^\mu \equiv (ct, \vec{x}) = (ct, x, y, z) \tag{1.29}$$

where x^μ is the four vector and $\vec{x}$ is a three vector consisting of the x, y, z components. We can do this, not just for convenience but from the fact that c is a fundamental constant of nature and so time and space become interchangeable by simply multiplying/dividing by c. We often call x^μ, a space-time coordinate. This is actually a far more profound consequence of Maxwell's equations then it is often given credit for, as it truly unifies space-time.

Before we go any further, there a few more items that need to be defined:

$$\begin{aligned} \partial_\mu \equiv \frac{\partial}{\partial x^\mu} &= \left(\frac{\partial}{\partial x^0}, \frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3}\right) \\ &= \left(\frac{\partial}{c\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \end{aligned} \tag{1.30}$$

Next, we define a matrix, known as the Minkowski metric that is crucial for all of our physical theories:

$$\eta^{\mu\nu} \equiv \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.31}$$

It sometimes also has the signature of (+ - - -), but the important part is the difference in sign between temporal and spatial components. This change in sign accounts for all the differences between space and time, for example, the fact that we cannot move backward in time comes from the difference in sign. In fact the main difference between EM and gravity is that gravity is always attractive, as supposed to repulsive and attractive forces in EM. This also comes about due to the difference in sign.

Using $\eta^{\mu\nu}$, we can rewrite the wave equation as:

$$\eta^{\mu\nu} \frac{\partial}{\partial x^\mu} \frac{\partial}{\partial x^\nu} B = 0 \tag{1.32}$$

the B is the magnetic field vector, $\vec{B}$, unless stated otherwise. To see it explicitly, we can simply expand the indicies:

$$\begin{aligned} \eta^{\mu\nu} \frac{\partial}{\partial x^\mu} \frac{\partial}{\partial x^\nu} B &= \eta^{00} \frac{\partial^2}{\partial {x^0}^2} B + \eta^{11} \frac{\partial^2}{\partial {x^1}^2} B + \eta^{22} \frac{\partial^2}{\partial {x^2}^2} B + \eta^{33} \frac{\partial^2}{\partial {x^3}^2} B \\ &= -\frac{1}{c^2} \frac{\partial^2 B}{\partial t^2} + \frac{\partial^2 B}{\partial x^2} + \frac{\partial^2 B}{\partial y^2} + \frac{\partial^2 B}{\partial z^2} = 0 \end{aligned} \tag{1.33}$$

Which can be re-written as:

$$\frac{1}{c^2} \frac{\partial^2 B}{\partial t^2} = \nabla^2 B \tag{1.34}$$

Once space-time are unifies as four-vectors, we can use a similar trick for charge density and current density:

$$j^\mu \equiv (c\rho, \vec{J}) \tag{1.35}$$

This is known as the four vector electric current. The obvious question to ask now is how to unify the electric field E and the magnetic field B in this same way. It is not so obvious however, as in total there are six independent components for them (i.e E_x, E_y, E_z and same for B), so we can't use a four vector. The next thing to try is a matrix (a rank two tensor). But a rank two tensor with two indicies, $F_{\mu\nu}$, has 16 components, and we only need 6, so it is far too many. To restrict the number of components, we can impose a condition on the matrix that it is *antisymmetric*:

$$F_{\mu\nu} = -F_{\nu\mu} \tag{1.36}$$

This means that it only has 6 *independent* components as:

$$F_{\mu\mu} = -F_{\mu\mu} \equiv 0 \ \ \forall\mu \tag{1.37}$$

This implies that all diagonal elements are zero, which leaves 12 components. Now if one imposes the condition:

$$F_{\mu\nu} = -F_{\nu\mu} \ \ \forall\mu \neq \nu \tag{1.38}$$

then there are only 6 independent components as half of them are related to the other half by a minus sign. So we can define the antisymmetric tensor for the electric and magnetic fields, also known as the field strength tensor, is defined as:

$$F_{\mu\nu} \equiv \begin{pmatrix} 0 & -\frac{E_x}{c} & -\frac{E_y}{c} & -\frac{E_z}{c} \\ \frac{E_x}{c} & 0 & B_z & -B_y \\ \frac{E_y}{c} & -B_z & 0 & B_x \\ \frac{E_z}{c} & B_y & -B_x & 0 \end{pmatrix} \tag{1.39}$$

So the components are:

$$\begin{aligned} F_{0i} &= -\frac{E_i}{c} \\ F_{ij} &= \epsilon_{ijk} B_k \end{aligned} \tag{1.40}$$

One has to be careful with indicies that are covariant (at the bottom) or contravariant (at the top). To convert from covariant to contravariant and vice-verse we can define:

$$\eta_{\mu\nu} \equiv (\eta^{\mu\nu})^{-1} \tag{1.41}$$

So when we contract them:

$$\eta^{\mu\alpha}\eta_{\alpha\beta} = \delta^{\mu}_{\beta} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.42}$$

which means we construct the wave operator as:

$$\begin{aligned} \eta^{\mu\nu}\partial_\mu\partial_\nu &= \partial^\nu\partial_\nu \\ \partial^\nu = \eta^{\mu\nu}\partial_\mu &\equiv \left(-\frac{1}{c}\frac{\partial}{\partial t} + \nabla\right) \end{aligned} \tag{1.43}$$

In the same way:

$$\begin{aligned} j^\mu &= \eta_{\mu\nu} j^\mu \\ &= j_\nu \equiv \left(-\rho c, \vec{J}\right) \end{aligned} \tag{1.44}$$

$$F^{\mu}_{\alpha} = \eta^{\mu\beta} F_{\beta\alpha} \tag{1.45}$$

$$F^{\mu\nu} = \eta^{\mu\alpha}\eta^{\nu\beta} F_{\alpha\beta} \tag{1.46}$$

So the raising and lowering the indices simply changes the sign of the time components (depending on ths signature of the η of course).

Using all the tools, we can write Maxwell's equations as follows:

CLAIM 1. Equations 1.12 and 1.15 can be combined into:

$$\partial_\mu F^{\mu\nu} = -\mu_0 j^\nu \tag{1.47}$$

PROOF 1. Begin with the $\mu = 0$, temporal index:

$$\partial_0 F^{0i} = -\mu_0 j^i \qquad \text{as } F^{00} = 0 \text{ and } i \in \{1, 2, 3\} \tag{1.48}$$

Combining this with Eq 1.35 gives:

$$\frac{\dot{E}_i}{c} = \mu_0 \vec{J} \tag{1.49}$$

And a similar procedure has to be followed to obtain the remaining terms.[1]

A similar long winded procedure can be used to prove:

$$\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_{F\alpha\beta} = 0 \tag{1.50}$$

incorporates equations 1.13 and 1.14 of Maxwell's equations. This can be written as:

$$\partial_{[\alpha} F_{\beta\gamma]} = 0 \tag{1.51}$$

The square bracket means the indicies are anti-symmetrised, which means:

$$\partial_{[\alpha} F_{\beta\gamma]} = \frac{1}{6} \left(\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} - \partial_\alpha F_{\gamma\beta} - \partial_\beta F_{\alpha\gamma} - \partial_\gamma F_{\beta\alpha} \right) \tag{1.52}$$

But F is anti-symmetric, therefore:

$$\partial_{[\alpha} F_{\beta\gamma]} = \frac{1}{3} \left(\partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} \right) = 0 \tag{1.53}$$

Which is equivalent to the statement in Eq 1.50. Sometimes Eq 1.50 is also written in yet another way:

$$\epsilon^{\mu\nu\rho\alpha} \partial_\nu F_{\rho\alpha} = 0 \tag{1.54}$$

2.4. Electric and magnetic potentials. The electric field, ϕ, is a scalar and is defined as:

$$\vec{E} \equiv -\nabla\phi \qquad \text{Time-independent case} \tag{1.55}$$

The magnetic potential, $\vec{A}$, is a vector and is defined as:

$$\vec{B} \equiv \nabla \times \vec{A} \tag{1.56}$$

They are combined together to form a four vector potential:

$$A^\mu \equiv \left(\frac{\phi}{c}, \vec{A} \right) = \left(\frac{\phi}{c}, A_x, A_y, A_z \right) \tag{1.57}$$

Then the EM field strength tensor, $F_{\mu\nu}$ can be written as:

CLAIM 2.

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{1.58}$$

[1] One of the problems with any tensor formulation of equations is that they are highly tangled up and unpacking them usually takes a long time. This is why Einstein's field equations are so difficult to solve.

Proof 2. First consider the temporal component of Eq 1.58:

$$F_{00} = \partial_0 A_0 = \partial_0 A_0 = 0 \tag{1.59}$$

The spatial components:

$$F_{0i} = \partial_0 A_i - \partial_i A_0 \tag{1.60}$$

Which can be written in component form as:

$$\frac{\vec{E}}{c} = \frac{\partial A}{\partial t} - \nabla\phi\frac{1}{c} \tag{1.61}$$

Which implies:

$$\vec{E} = -\nabla\phi \tag{1.62}$$

and

$$\frac{\partial \vec{A}}{\partial t} = 0 \qquad \text{Time independent case} \tag{1.63}$$

Similarly, one can prove the remaining components aswell.

Eq 1.58 is of fundamental importance, not just in EM, but also in all of physics. This is because it is the first equation that shows how the concept of gauge invariance comes about. To see how this works, consider the transformation:

$$A_\mu \to A_\mu + \partial_\mu \chi \tag{1.64}$$

This will leave $F_{\mu\nu}$ invariant (as the partial derivatives commute and both terms in Eq 1.58 will give equal and opposite terms that will cancel) and hence the equations of motion invariant. In fact this also leads on to gauge symmetries of the Standard Model of particle physics.

Finally, the Lorentz force:

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}) \tag{1.65}$$

To express this in terms of four vectors lets write down the equation of force:

Claim 3. Using Newton's second law:

$$m\frac{\partial^2 x^\mu}{\partial \tau^2} = qF^\mu_\nu \frac{\partial^\nu}{\partial \tau} \tag{1.66}$$

Proof 3. Lets start with the temporal components:

$$\begin{aligned} m\underbrace{\frac{\partial^2 x^0}{\partial \tau^2}}_{\equiv 0} &= qF^0_\nu \frac{\partial x^\nu}{\partial \tau} \\ 0 &= qF^0_i \frac{\partial x^i}{\partial \tau} \\ 0 &= qF^0_i \vec{v} \Rightarrow \vec{v} \equiv 0 \end{aligned} \tag{1.67}$$

The spatial components are:

$$\begin{aligned} m\frac{\partial^2 x^i}{\partial \tau^2} &= qF^i_\nu \frac{\partial x^\nu}{\partial \tau} \\ m\vec{a} &= qF^i_\nu \frac{\partial x^\nu}{\partial \tau} \\ &= qF^i_0 \underbrace{\frac{\partial x^0}{\partial \tau}}_{\equiv 1} + qF^i_j \frac{\partial x^j}{\partial \tau} \\ &= q\vec{E} + q(\vec{B} \times \vec{v}) \end{aligned} \tag{1.68}$$

3. Lorentz Symmetry

Lorentz symmetry was first discovered by Lorentz(hence bears his name), and was first seen as a mathematical property of Maxwell's equations. However, it was Einstein who realised that this was not just a property, but also a fundamental feature of our universe. The symmetry is defined by a a change of space-time coordinates:

$$x^\mu(ct, \vec{x}) :\rightarrow x^{\mu\prime} = (ct', \vec{x}') \equiv \Lambda^\mu_\nu x^\nu \tag{1.69}$$

This is a linear transformation, where Λ^μ_ν is a transformation matrix. Under this transformation, the Maxwell equations are invariant. Similarly one can also have transformations in the other EM quantities:

$$\begin{aligned} A'^\mu &= \Lambda^\mu_\nu A^\nu \\ j'^\mu &= \Lambda^\mu_\nu j^\nu \\ F'^{\mu\nu} &= \Lambda^\mu_\alpha \Lambda^\nu_\beta F^{\alpha\beta} \end{aligned} \tag{1.70}$$

These equations can be solved for the primed reference frame, or the unprimed reference frame if the transformation matrices Λ^μ_ν is invertible:

$$x^\nu = \tilde{\Lambda}^\nu_\mu x'^\mu \qquad \text{Such that } \tilde{\Lambda}^\nu_\mu \equiv (\Lambda^\mu_\nu)^{-1} \tag{1.71}$$

This can also be used to write down the transformations for derivatives:

$$\begin{aligned} \frac{\partial}{\partial x'^\mu} &= \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial}{\partial x^\alpha} \\ &= \tilde{\Lambda}^\alpha_\mu \frac{\partial}{\partial x^\alpha} \end{aligned} \tag{1.72}$$

So we see that covariant indicies transform with $\tilde{\Lambda}$ and vice-verse. It is immediately obvious that contracted indicies do not transform at all (as the covariant index will transform under $\tilde{\Lambda}$ and indicies with Λ, thus the expression has $\Lambda\tilde{\Lambda} \equiv \mathbb{I}$).

Now recall:

$$F^{\alpha\beta} = \eta^{\alpha\delta}\eta^{\beta\gamma}F_{\delta,\gamma} \tag{1.73}$$

To check that the Lorentz transformation:

$$F'^{\alpha\beta} = \Lambda^\alpha_\gamma \Lambda^\beta_\delta F^{\gamma\delta} \tag{1.74}$$

leaves Maxwell's equations invariant, the transformation of the η's, under the Lorentz transformations, needs to be worked out. The condition on η is obviously that they need to be Lorentz invariant:

$$\begin{aligned}\eta_{\alpha\beta} &= \eta_{\gamma\delta}\tilde{\Lambda}^{\gamma}_{\alpha}\tilde{\Lambda}^{\delta}_{\beta} \\ \eta^{\alpha\beta} &= \Lambda^{\alpha}_{\gamma}\Lambda^{\beta}_{\delta}\eta^{\gamma\delta} \end{aligned} \tag{1.75}$$

These conditions *define* Lorentz transformations, i.e matrices that satisfy Eq 1.75 are Lorentz transformations and the set of all matrices that satisfy this condition form the *Lorentz group*.

An example is of a Lorentz transformation is via an orthogonal matrix, O_{ij} (say).

Claim 4. The length of any vector in 3-D is invariant under an orthogonal transformation matrix, O_{ij}:

$$x'_i = O_{ij}x_j \Rightarrow |\vec{x}| \equiv |\vec{x'}| \tag{1.76}$$

Proof 4. Take the starting point as the length of the primed vector:

$$\begin{aligned}|\vec{x}'|^2 &\equiv \vec{x}'^{\dagger}\vec{x} \\ &= \vec{x}^{\dagger}\underbrace{O^{\dagger}O}_{\equiv 1}\vec{x} \\ &= \vec{x}^{\dagger}\vec{x} \\ &\equiv |\vec{x}|^2 \end{aligned} \tag{1.77}$$

Lorentz transformations are analogous to the one above, the only difference is the η which has a minus sign in the temporal components. We can re-write Eq 1.75 as:

$$\Lambda^T\eta\Lambda = \eta \tag{1.78}$$

As an example for this Λ matrix, consider an orthogonal matrix:

$$\Lambda^{\alpha}_{\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.79}$$

This is a rotation matrix for the x, y coordinates and we can see that by defining a 2×2 matrix:

$$O \equiv \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{1.80}$$

and $OO^T \equiv \mathbb{I}$, which implies that:

$$(x', y') = O\begin{pmatrix} x \\ y \end{pmatrix} \tag{1.81}$$

Which of course give the equations:

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \end{aligned} \tag{1.82}$$

Geometrically the rotation is given by Figure 1. This is the simplest example of a Lorentz transformation. More interesting Lorentz transformations take space into time, such as:

$$\Lambda^{\mu}_{\nu} = \begin{pmatrix} \cosh\theta & -\sinh\theta & 0 & 0 \\ \sinh\theta & \cosh\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.83}$$

This gives the transformation equations:

$$ct' = ct\cosh\theta - x\sinh\theta \tag{1.84}$$

This is also a linear transformation on the coordinates, but it is not a rotation due to the hyperbolic functions. Geometry is shown in Figure 2

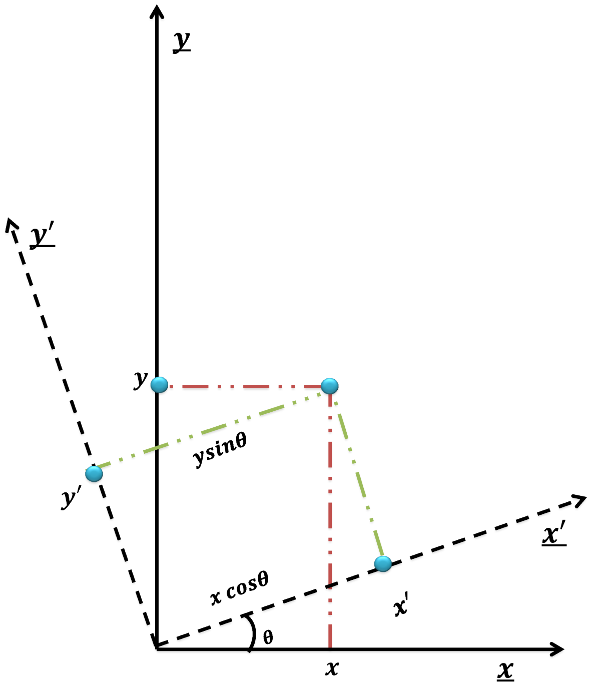


FIGURE 1. Geometry of transformation for Λ^{α}_{β} matrix

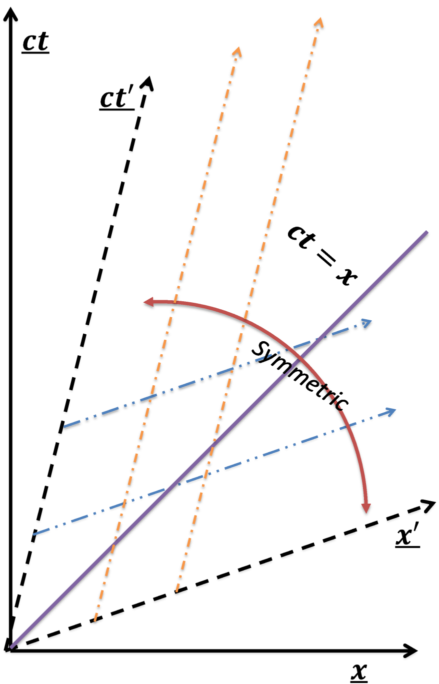


FIGURE 2. Geometry of transformation for $\Lambda^{\mu}\nu$ matrix. The dashed orange lines are lines of constant x', the dashed blue lines are lines of constant t'.

Now, the x' axis, is by definition where $t' = 0$. Therefore x' axis is defined by the condition:

$$ct \cosh\theta = x \sinh\theta \Rightarrow ct = x \tanh\theta \tag{1.85}$$

This represents a straight line with a constant gradient that is less than 1 (as $\tanh\theta < 1$), likewise ct' axis is where $x' = 0$, so we get:

$$ct = x \coth\theta \tag{1.86}$$

which is a straight line with gradient that is larger than 1, in fact the slopes of both lines are the inverse of each other, i.e if the gradient of ct' axis 6, then the gradient of the x' axis is $\frac{1}{6}$ and so on. This kind of Lorentz transformation is called a *boost*.

The physical meaning of the new coordinates is that the ct' axis is the locus of an observer who is in the primed reference frame. In the original coordinates, this observer is located at:

$$x = \tanh\theta ct \Rightarrow \tanh\theta = \frac{v}{c} \tag{1.87}$$

one often defines:

$$-\infty < \theta < \infty \tag{1.88}$$

as the *rapidity*. So if θ is known, the relation to all other hyperbolic functions is also known:

$$\begin{aligned} \cosh\theta &= \frac{1}{\sqrt{1 - \tanh^2\theta}} \\ &= \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \end{aligned} \tag{1.89}$$

$$\sinh\theta = \frac{\frac{v}{c}}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{1.90}$$

We also define the usual γ factor as:

$$\gamma \equiv \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{1.91}$$

Therefore, the Lorentz transformations are given by:

$$\begin{aligned} ct' &= \gamma\left(ct - \frac{vx}{c}\right) \\ x' &= \gamma(x - vt) \\ y' &= y \\ z' &= z \end{aligned} \tag{1.92}$$

Since:

$$\eta_{\mu\nu} x'^{\mu} x'^{\nu} = \eta_{\mu\nu} x^{\mu} x^{\nu} \tag{1.93}$$

The length of a space-time vector is invariant under the Lorentz transformations:

$$x^2 = x'^2 \tag{1.94}$$

This is often called the *proper distance* (also called a space-time interval) and this is invariant under Lorentz transformations (but ordinary time and spatial intervals are not). Any observer will also agree on what is in the past and the future, this is known as the concept of *causality*. Is it preserves the fact any action follows a cause, for example no observer will ever see a ball being struck in a cricket match before the ball has been delivered. To summarise, the concept of simultaneity is no longer valid for different reference frames, as observes will see events happening at a different time to each other, however the *order* of events will always be the same and this conservers *causality*.

3.1. Clocks and meter sticks. Now lets work out two of the most famous consequences in special relativity, which are time dilation and length contraction.

Time Dilation

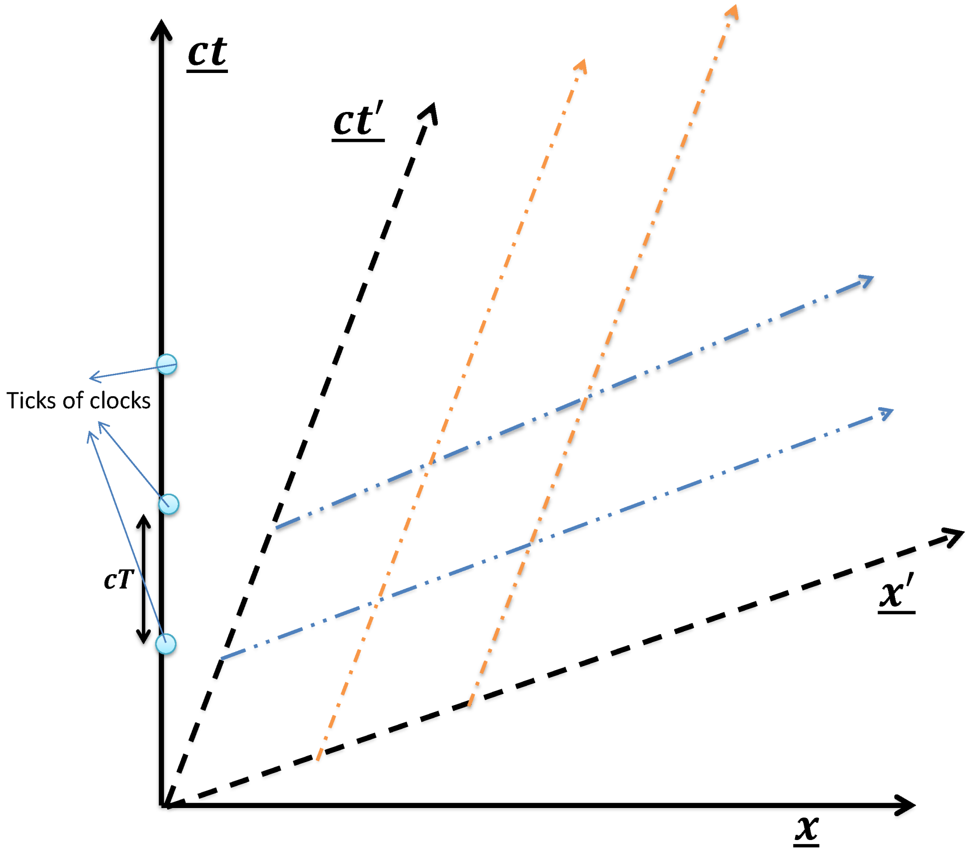


FIGURE 3. Space-time diagram showing ticks on a clock

Consider a clock at $x = 0$. We want to see how the clock ticks in a different, primed coordinate system. In the primed coordinates we will see that the value of the ticks of the clocks is becoming more negative in the x axis. This is intuitively obvious; as the observer moves away from the clock, the clock is at a more negative distance from the observer. The interesting question is weather the time *interval* measured on the clock is the same in both reference frames. Suppose the time interval in the unprimed frame is T, then the ticks in the unprimed frame are:

Tick	ct	x
1	0	0
2	cT	0
3	2cT	0

TABLE 1. Ticks in the unprimed frame

In the primed reference frame:

Tick	ct'	x'
1	0	0
2	γcT	-γvT
3	$2\gamma cT$	- 2 γvT

TABLE 2. Ticks in the primed frame

So the space coordinate for the first tick is at $-\gamma vT$, which is negative as expected, but looking at it carefully, $\gamma T \equiv T'$, therefore x' for the first tick is at $-vT$ and similarly the temporal coordinate for the first tick is at cT', the fact that:

$$\gamma T = T' \tag{1.95}$$

is called *time dilation*. In other words, the time interval observed in a moving frame will be longer than in a stationary frame. This follows straight from the fact that the speed of light is a constant in a stationary reference frame. So the fact that an observer is moving relative to the clock, does not mean the velocities of the light ray and the moving reference frame add according to Galilean transformations, instead they follow relativistic addition of velocities.

Meter sticks

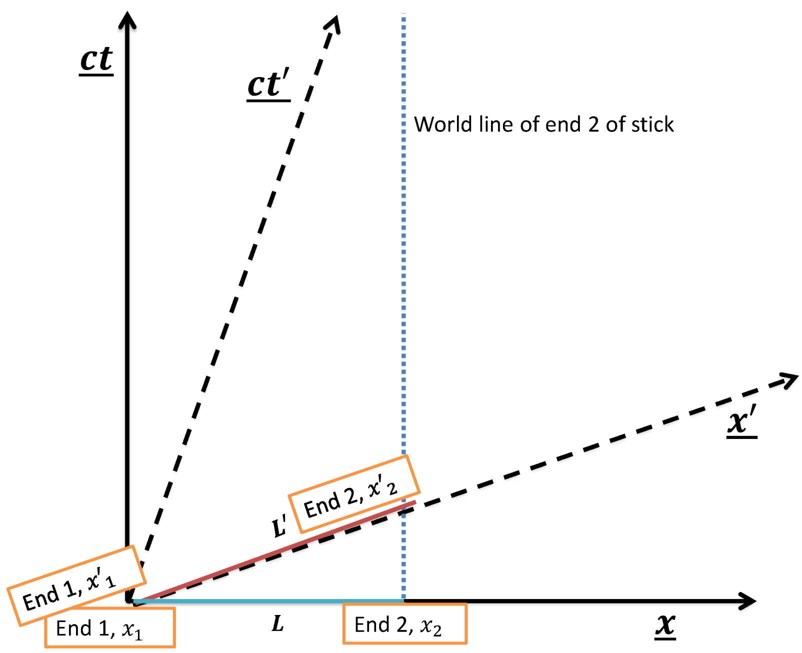


FIGURE 4. Space-time diagram of a meter stick. Red line shows the length of rod in the primed reference frame, L'. Blue line shows the length of the rod in the unprimed reference frame, L

As was done for the clock, one must now ask what is the length of the stick as measured by an observer in a moving primed reference frame. Lets look at the two ends of the stick both frames:

	Unprimed frame	Primed frame
Left end (End 1)	$x_1 = 0, t_1 = ct$	$x'_1 = -\gamma vt \equiv -vt, t'_1 = \gamma t$
Right end (End 2)	$x_2 = L, t_2 = ct$	$x'_2 = \gamma(L - vt), t'_2 = \gamma\left(t - \frac{vL}{c^2}\right)$

Table 3. Coordinates of meter stick in both reference frames.

let $t'_2 = 0$, to see the length of the stick in the primed frame, we simply solve for t, under this condition:

$$t = \frac{vL}{c^2} \tag{1.96}$$

Now we can use this relation to find out how the lengths of the stick changes in both reference frames. If we take End 1 of the stick to be at the origin, then the coordinate x'_2 represents its length in the primed frame:

$$\begin{aligned} x'_2 &\equiv L' = \gamma\left(L - vt\right) \\ &= \gamma\left(L - v\left(\frac{vL}{c^2}\right)\right) \\ &= \gamma L\left(1 - \frac{v^2}{c^2}\right) \\ &= \frac{L}{\gamma} \end{aligned} \tag{1.97}$$

which is smaller than L, which means that in the primed frame, sticks appear smaller than in the unprimed reference frame. At first there appears to be a contradiction between the two observers. The observer in the stationary frame, O1, will say that the observer in the moving frame, O2, has a watch that ticks slower. On the other hand O2, will say that his/her watch is ticking at the rate of 1 second per 1 second, and that the watch on O1 is ticking slower.

So what is the real answer...? The answer is, that they are both correct. It is easier to see how the contradiction disappears from the geometry of the situation. For the information that follows, refer to Figure 5:

	Observer 1 measures	Observer 2 measures
Time interval for O1	Yellow line without arrow	Green line with arrows
Time interval for O2	Yellow line with arrows	Green line without arrows

Table 4. Time intervals measured by both observer's for each other in their respective frames of reference

One has to be careful here, it is not the distance on the axis that gives rise to the Lorentz contraction and time dilation. Instead it is the space-time intervals that need to be used, as it is these that are invariant under Lorentz transformations:

$$c^2t^2 - x^2 = c^2t'^2 - x'^2 \tag{1.98}$$

The space-time intervals are hyperbolic curves. The same procedure can be followed for the length in both reference frames to look at the effect of length contraction(i.e draw the lines of constant x, x', t, t' and find the intersect to the world line of End 2 in the Figure 5[2]). Following the same logic will show that both the observers will be correct when they saw that the length is

[2]This is left to the reader as drawing more lines in Figure 5 would make it look more messy!

increasing in the reference frame of the other. This shows the confusing yet profound consequences of the speed of light being a constant in each reference frame and is at the heart of the theory of special relativity.

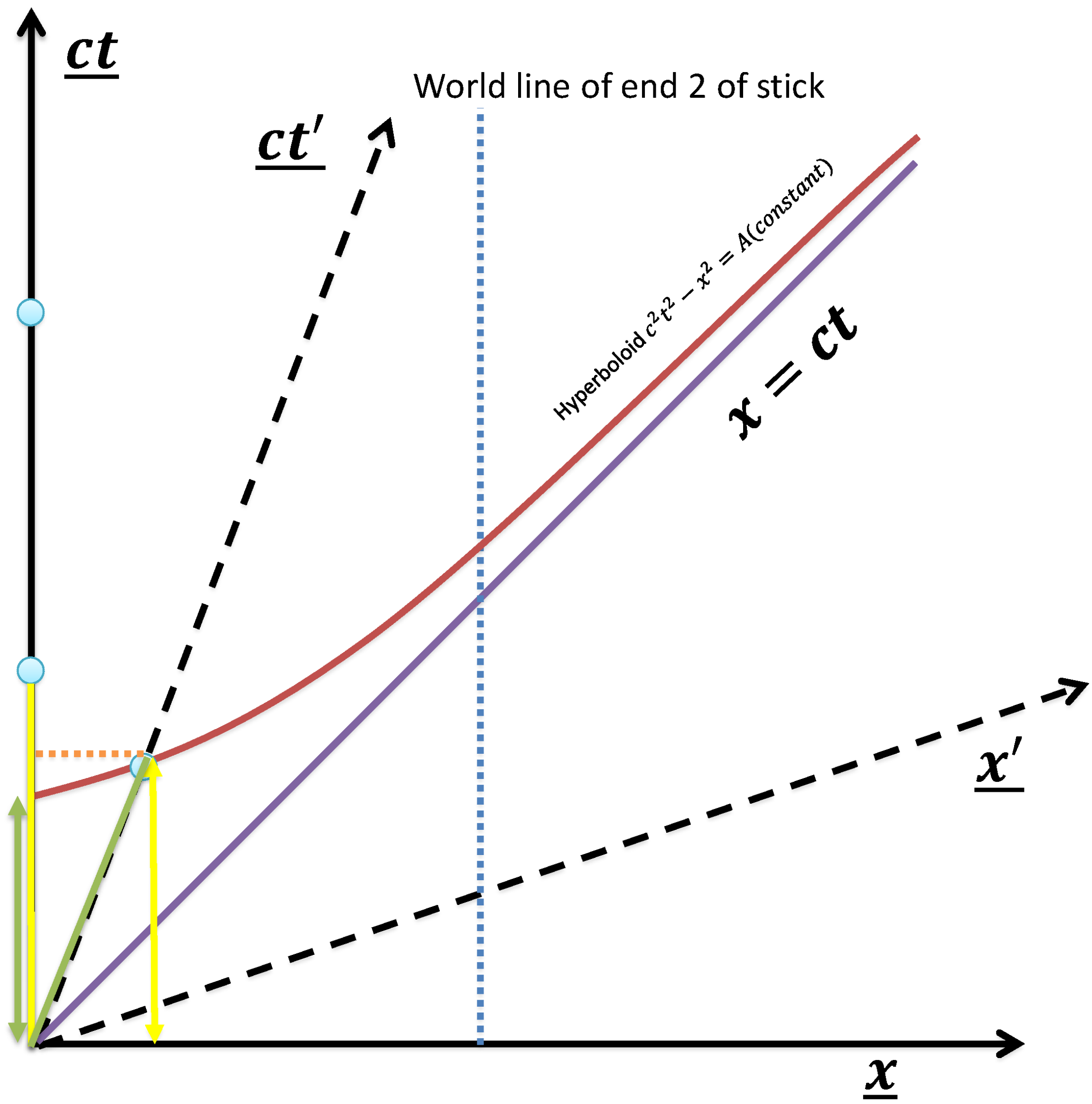


FIGURE 5. The geometry of time dilation and length contraction

4. Relativistic mechanics

Mechanics has been about describing the trajectory of particles given a force and some initial conditions. This has been the case for Newtonian mechanics and is no different in relativistic mechanics. Suppose T is some trajectory of a particle in space-time, and it can be parametrised by our coordinates x^μ and some parameter λ. In general, λ must be a Lorentz invariant quantity, and a good quantity to use the *proper time* that describes the flow of the trajectory.

At each point one can define a parametric, 4-velocity:

$$\text{Parametric 4-velocity} = \frac{\partial x^\mu}{\partial \lambda} \tag{1.99}$$

this simply shows how the coordinates change with λ. So now we can define a length element as:

$$ds^2 = \eta_{\mu\nu} \frac{\partial x^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda} d\lambda^2 \tag{1.100}$$

There are certain classifications that ds falls under and they are as follows:

- If $ds^2 > 0$, we call the trajectory *space-like*. There are no particles in nature that have space-like trajectories since that would involve moving faster than the speed of light (there exist hypothetical particles in some theories, like string theory for example, that travel faster than the speed of light).
- If $ds^2 = 0$, it is called a *light-like* trajectory. These are all particles that are mass-less and hence travel at the speed of light.
- If $ds^2 < 0$, the trajectory is called *time-like*. These are all massive particles.

The type of trajectory is always the same in any coordinate system. For time-like trajectories, we define:

$$c\partial\tau \equiv = \sqrt{-\eta_{\mu\nu}\partial x^\mu \partial x^\nu} \tag{1.101}$$

This quantity is Lorentz invariant and the expression above can be re-written to get it in a more familiar form:

$$\begin{aligned} \partial\tau &= \frac{1}{c}\sqrt{-\eta_{\mu\nu}\frac{\partial x^\mu}{\partial t}\frac{\partial x^\nu}{\partial t}}dt \\ &= dt\sqrt{1-\frac{1}{c^2}\left(\frac{\partial x^i}{\partial t}\right)^2} \\ &= dt\sqrt{1-\frac{v^2}{c^2}} = \frac{dt}{\gamma} \end{aligned} \tag{1.102}$$

4.1. 4-velocity. Define:

$$\begin{aligned} u^\mu &= \frac{\partial x^\mu}{\partial \tau} \\ &= \frac{\partial x^\mu}{\partial t}\frac{\partial t}{\partial \tau} \\ &= \gamma(c, \vec{v}) \end{aligned} \tag{1.103}$$

Therefore:

$$u^\mu u_\mu = -c^2 \tag{1.104}$$

4.2. 4-momentum. Particle mass depends on velocity as follows:

$$m = \gamma m_0 \tag{1.105}$$

The proof is quite long and is worth doing once; it is found in many resources like [**2**][3].
The four momentum, p^μ, is given by:

[3]Having just worked through it, it is not worth typing up!

$$\begin{aligned} p^\mu &= m_0 u^\mu \\ &= (\gamma m_0 c, \gamma m_0 \vec{v}) \\ &= (mc, m\vec{v}) \\ &= \left(\frac{E}{c}, \vec{p}\right) \end{aligned} \tag{1.106}$$

Therefore:

$$p^2 = p^\mu p_\mu = mc^2 \tag{1.107}$$

This is a Lorentz invariant quantity and proves to be very useful in all areas from physics, especially in particle physics, when the kinematics of colliding particles in considered as m, also known as the invariant mass of a particle, is the same in all frames and hence be used in working out unknown quantities in different reference frames. For example, in a the rest frame, where $\gamma \equiv 1$:

$$p^2 = p^\mu_{(rest)} p_{\mu(rest)} = -m_0^2 c^2 \tag{1.108}$$

In other frames:

$$p^2 = p^\mu_{(other)} p_{\mu(other)} = -\frac{E^2}{c^2} + |\vec{p}|^2 \tag{1.109}$$

Since p^2 is a Lorentz invariant quantity, Eq 1.108 and 1.109 and be equated to give:

$$E = c\sqrt{\vec{p}^2 + m_0^2 c^2} \tag{1.110}$$

Which can be further simplified:

$$\begin{aligned} \vec{E} &= m_0 c^2 \sqrt{1 - \left(\frac{p}{mc}\right)^2} \\ &= mc^2 \left(1 + \frac{1}{2}\frac{p^2}{mc^2}\right) \\ &= \underbrace{m_0 c^2}_{Restmass} + \underbrace{\frac{1}{2}\frac{p^2}{m_0}}_{KE} \end{aligned} \tag{1.111}$$

This works even for $m_0 = 0$ as then:

$$E = c|\vec{p}| \tag{1.112}$$

which is just the energy of a photon.

4.3. Dynamics. Newton's laws are the starting point for the dynamics in relativity aswell. The difference comes in the formalism of four vectors; the four vector force can now be defined in terms of the four vector momentum as:

$$F^\mu = \frac{\partial p^\mu}{\partial \tau} \tag{1.113}$$

An example is the one looked at while discussing EM and its the Lorentz force. The Lorentz force, in terms of four vectors can be written as:

$$F^\mu = qF^{\mu\nu} u_\nu \tag{1.114}$$

where $F^{\mu\nu}$ is the usual field strength tensor. Let's explicitly do a calculation to see how this formalism works for forces. Consider:

$$\begin{aligned} F.u &= \eta_{\mu\nu} F^{\mu\nu} \\ &= \nu_{\mu\nu} u^{\mu} \underbrace{\frac{\partial p^{\mu}}{\partial \tau}}_{F^{\mu}} \end{aligned} \tag{1.115}$$

But $p^{\mu} = m_o u^{\mu}$, therefore:

$$F.u = \eta_{\mu\nu} u^{\nu} \frac{\partial (m_0 u^{\mu})}{\partial \tau} \tag{1.116}$$

Use chain rule:

$$\begin{aligned} F.u &= \eta_{\mu\nu} u^{\nu} \frac{\partial m_0}{\partial \tau} u^{\mu} + \eta_{\mu\nu} u^{\nu} \frac{\partial u^{\mu}}{\partial \tau} m_0 \\ &= \eta_{\mu\nu} u^{\nu} \frac{\partial m_0}{\partial \tau} u^{\mu} + (\eta_{\mu\nu} u^{\nu}) \frac{\partial u^{\mu}}{\partial \tau} m_0 \\ &= \eta_{\mu\nu} u^{\nu} \frac{\partial m_0}{\partial \tau} u^{\mu} - u_{\mu} \frac{\partial u^{\mu}}{\partial \tau} m_0 \end{aligned} \tag{1.117}$$

Since derivatives compute both terms of the R.H.S can be simplified:

$$\begin{aligned} F.u &= \eta_{\mu\nu} u^{\nu} \frac{\partial m_0}{\partial \tau} u^{\mu} - \frac{\partial}{\partial \tau} \underbrace{(u_{\mu} u^{\mu})}_{c^2} m_0 \\ &= \eta_{\mu\nu} u^{\nu} \frac{\partial m_0}{\partial \tau} u^{\mu} \\ &= \eta_{\mu\nu} u^{\nu} u^{\mu} \frac{\partial m_0}{\partial \tau} \\ &= -c^2 \frac{\partial m_0}{\partial \tau} \end{aligned} \tag{1.118}$$

This means that if the motion is along the line of the force (or any components of the motion are along the line of force) then the rest mass varies with time. When though about carefully, this makes sense as the rest mass is a source of energy and we expect energy to change (work is done) when a force acts along the direction of motion. This is not the case for electromagnetic forces. If the Lorentz force is multiplied by u:

$$Fu = q u_{\mu} u_{\nu} F^{\mu\nu} \tag{1.119}$$

But since $u_{\mu} u_{\nu}$ is symmetric in $\mu\nu$ and $F^{\mu\nu}$ is anti-symmetric in $\mu\nu$, $Fu \equiv 0$. This means EM forces do not change the energy of particles they act on, which is another peculiarity of Maxwell's equations.

CHAPTER 2

Introducing gravity

1. Motivating general relativity

Let's go back to the time at which Einstein was thinking about how to generalise his theory. He started to think about the geometry of space-time and gravity and how they might be linked, immediately after completing his work on special relativity.

Einstein wanted to find a field theory for gravity that superseded Newton's theory, as the classical gravity was inconsistent with special reativity. This is immediately obvious from the equation of Newtonian gravitation force:

$$\vec{F}_{(gravity)} = \frac{Gm_1 m_2}{\vec{r}^3}\hat{\vec{r}} \tag{2.1}$$

There is no time dependence in the equation, which implies that the force was transmitted *instantaneously*. This obviously contradicts special relativity and its causal structure. Einstein wanted to eliminate this "spooky action at a distance". But gravity was not the only thing at the time that was described by such an equation. Coulomb's law for electro-static forces had exactly the same form:

$$\vec{F}_{(Coulomb)} = k\frac{q_1 q_2}{\vec{r}^3}\hat{\vec{r}} \tag{2.2}$$

Again this violated causality. But this was just static limit of the EM fields that were described fully by Maxwell's equations. This is what motivated Einstein into thinking that maybe the Newtonian gravity is also a static limit for a more general field theory for gravity.

1.1. Relativity of space-time: Newton's Bucket. Another thing he did not agree with was the absolute nature of space and time, a view that had been promoted by Newton and had been accepted for over three centuries. This was not the accepted view before Newton however, Aristotle(arguably the greatest philosopher of all time) and Descartes held a view of relativity himself. Their views were much more local and specific than the view that Einstein would eventually come up with. They believed that it was not sensible to quantify things in an absolute way. The only reality of a physical object was in its relation to other surrounding physical objects. The same arguments can be applied to motion as well as space. To describe his view on absolute space-time, Newton came up with a thought experiment.

- Imagine that you have a bucket full of water.
- The bucket is then attached to a chord and then the chord is twisted around itself several times.
- If the chord is then released it will rotate and along with it so would the bucket.
- The water in the bucket would not start rotating immediately but it would start rotating after a few seconds due to the friction between the bucket and the water.
- When the water starts to rotate it forms a concave shape. In other words the concave shape of the water implies a rotation, but what is the rotation in reference to? i.e rotation with respect to what?

If Descartes views were correct and the motion had to be relative to some nearby objects, then the obvious object to use a reference would be the bucket as it is what contains the water. However this cannot be true as in the beginning when the bucket was rotating and the water was not rotating, their is still some relative motion, yet the water does not form that concave shape.

Moreover, after a certain amount of time, one would expect the water to be rotating at the same speed as the bucket, so their would be no relative motion, yet the concave shape would remain. Therefore the water cannot be relative to the bucket.

Newton argued that this implied the existence of absolute space to which the water was in relative motion with. This argument of Newton held strong for over 300 years.

In the 19^{th} century, Mach made an argument that the bucket and the water rotated with respect to all of the other matter in the universe. In other words its the motion relative to the distance stars that caused the concave shape. This is obviously not true, as it implies that if there was just a universe with the bucket and the water and the bucket was rotated the water would not rotate. Moreover, it also violates causality, as it implies that the presence of matter billions of light years away can cause the water to rotate!

Einstein obviously did not like the concept of causality in Mach's argument; his, correct, view was that the motion was relative to the local gravitational field. This would prove to be of fundamental important to his theory of gravity, because Einstein had stumbled upon a very profound fact that the classical Newtonian (in this case centripetal) acceleration was linked to gravity.

2. Equivalence principle

This principle is probably the most fundamental insight into our universe that general relativity gives us (and it also solves the problem of Newton's bucket!). In comes from an observation that all bodies in a given gravitational field experience the same acceleration. The connection between acceleration and gravity is as obvious(well it is now, as I said, hindsight is a wonderful thing!) as it is profound. Newtonian force is given by:

$$F = m_I \vec{a} \tag{2.3}$$

where m_I is the inertial mass, it is simply the resistance of acceleration to the force. In a uniform gravitational field, we know:

$$F = m_g \vec{g} \tag{2.4}$$

where m_g is the gravitational mass, which couples gravitational pull to the force. In principle m_I and m_g can be two different quantities, however experiments show that:

$$m_I = m_g \tag{2.5}$$

to a very high precision, which implies:

$$\vec{a} = \vec{g} \tag{2.6}$$

This is known as the *equivalence principle*. To explain this Einstein created a thought experiment. Imagine a person in an elevator, if the person feels a force downwards there is no way to determine weather the force is caused by acceleration upwards of the elevator or by a gravitational force pulling on the elevator. A consequence of this is known as red-shift.

2.1. Newtonian gravitational red-shift. This is a more ad-hoc derivation of the red-shift. A more rigorous derivation will follow in upcoming sections. Consider again an elevator that is initially on the ground and is being accelerated upwards (same as being pulled down by gravity).

Suppose a torch light emits a light ray from the bottom of the elevator, by the time it reaches the top of the elevator, the elevator itself has moved up a distance $\frac{1}{2}gt^2$ and the final velocity is gt. The distance traveled by the light is:

$$ct = \Delta h + \frac{1}{2}gt^2 \tag{2.7}$$

and the final velocity is:

$$v = gt \tag{2.8}$$

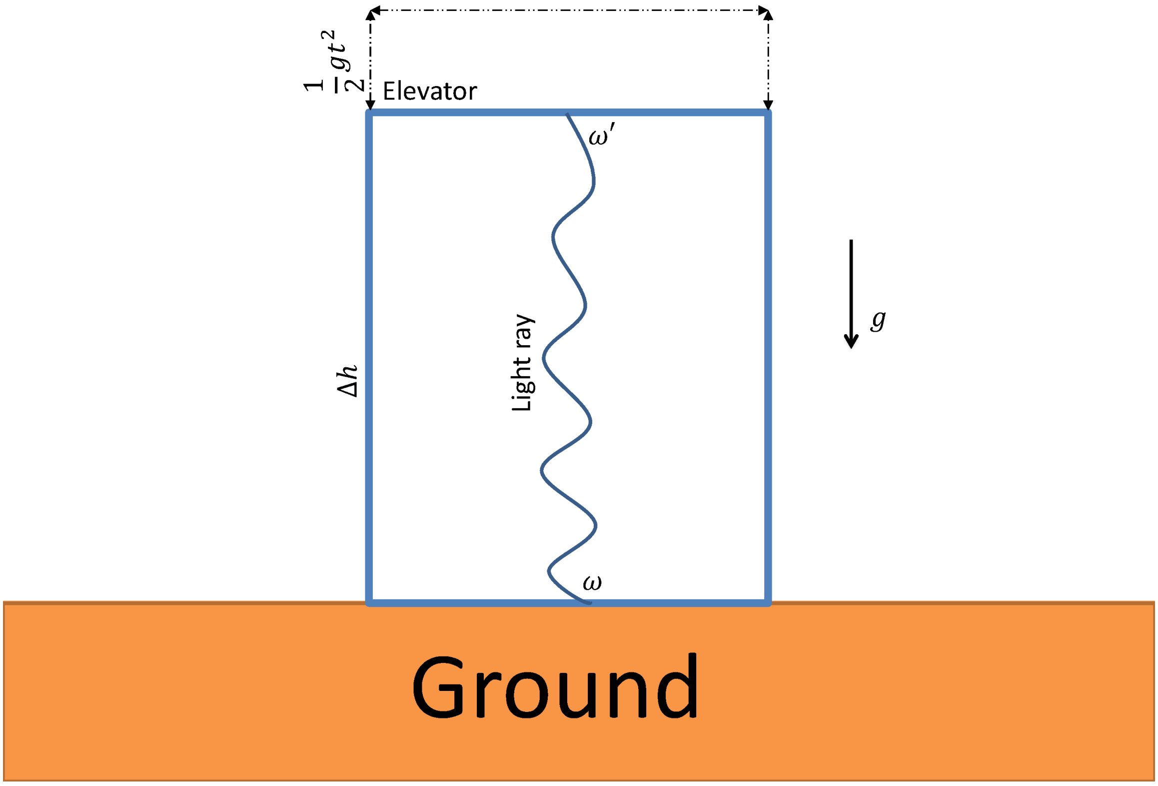


FIGURE 6. Elevator motion

we expect $t << 1$ and g is relatively small at the surface of the earth, so we can neglect the second term in Eq 2.7 and solve for it for t and substitute this into Eq 2.8:

$$v = g\frac{c}{\Delta h} \tag{2.9}$$

So the four wave-vector vector of the light at the bottom of the elevator is:

$$k^{\mu} = \left(\frac{\omega}{c}, 0, 0, \frac{\omega}{c}\right) \tag{2.10}$$

here it is assumed that the motion is only in the z direction. At the top it is:

$$\begin{aligned} k^{\mu\prime} &= \left(\frac{\omega'}{c}, 0, 0, \frac{\omega'}{c}\right) \\ &= \gamma\frac{\omega}{c}(1-\beta)(1, 0, 0, 1) \end{aligned} \tag{2.11}$$

Therefore:

$$\omega' = \gamma\omega(1-\beta) \tag{2.12}$$

if $\beta << 1$:

$$\omega' \approx \omega(1-\beta) \tag{2.13}$$

Substitute Eq 2.9 into 2.13:

$$\omega' = \omega\left(1 - \frac{\Delta h g}{c^2}\right) \tag{2.14}$$

So now a quantity can be defined as the gravitational redshift, z:

$$\begin{aligned} z &= \frac{\omega' - \omega}{\omega} \\ &= \frac{\omega'}{\omega} - 1 \\ &= -\frac{\Delta h g}{c^2} \end{aligned} \tag{2.15}$$

Note that this is negative, so the frequency of light at the top of the elevator is smaller than at the bottom. Physically this makes sense as a light photon will have to climb up a gravitational potential well in order to move up and therefore it must loose energy and since energy is proportional to ω (another thing Einstein showed with the photo-electric effect) ω must also decrease. Hence photons are shifted towards the red end of the spectrum, which is why the effect is called red-shift.

The time period of the wave is inversely proportional to its frequency:

$$\Delta\tau \propto \frac{1}{\omega} \tag{2.16}$$

This means that if one is in a strong gravitational field, the time, τ, will go slower (as ω will increase).

2.2. Bending of light: The Geodesic equation. Suppose the same elevator is moving upwards and a light ray is emitted from one of the walls of the elevator towards the other wall such that the light intersects the other wall at right angles. If the elevator is stationary, then the light ray will intersect the other wall at right angles as viewed by an observer in the elevator and by any external observer outside the elevator.

Once the elevator is in motion, the light ray will still appear to intersect the other wall at right angles as viewed by the observer in the elevator as it has no relative motion with respect to the light ray. However, to an external observer, the light will no longer appear to intersect the other wall at right angles. Instead, the light ray will appear to take a curved path towards the other wall. More specifically, the light ray is curved downwards. This is a general principle, the light rays are bent towards the gravitational field. The consequence is actually more profound that this; the photons path is not the thing that actually changes, it is infact the space-time itself that is changing. The space-time is bent by gravitating objects (massive, energetic objects) and particles follow geodesics (straight lines) in these new geometries.

The effect of gravity can be removed by moving into a freely falling reference frame, i.e the elevator accelerates at the same rate as $\vec{g}$ is pulling down. In this case one would not observe the bending of light as this would be an inertial frame.

To analyse how the space time is bent around by gravitating objects, the geodesic equation needs to be introduced. For simplicity lets set $c = 1$, as is the convention in general relativity theory. First let's formalise the phenomena of light curving in non-inertial (accelerating frames).

- Define:

$$\text{Local coordinates} = \xi^{\mu} \tag{2.17}$$

These coordinates are freely falling, i.e an inertial frame

- Define:

$$\text{Global coordinates} = x^\alpha \tag{2.18}$$

These are coordinates for an external observer.

In general, x^α and ξ^μ are related by a coordinate transformations:

$$x^\alpha = x^\alpha(\xi^\mu) \tag{2.19}$$

An example is a static observer on the surface of the earth observes x^α coordinates. In an inertial frame, we expect the rules of special relativity to apply, hence the equation of motion in the absence of any forces would be:

$$\frac{\partial^2 \xi^\mu}{\partial \tau^2} = 0 \tag{2.20}$$

Thus the solution for the trajectory would be a straight line:

$$\xi^\mu = m\vec{x} + c \qquad m, c \in \{\mathbb{R}\} \tag{2.21}$$

The next step is to find the trajectory in the global coordinate system; first let's use the chain rule:

$$\begin{aligned} \frac{\partial \xi^\mu}{\partial \tau^2} &= \frac{\partial}{\partial \tau}\left(\frac{\partial \xi^\mu}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial \tau}\right) \\ &= \frac{\partial \xi^\mu}{\partial x^\alpha}\frac{\partial^2 x^\alpha}{\partial \tau^2} + \frac{\partial^2 \xi^\mu}{\partial x^\beta \partial x^\alpha}\frac{\partial x^\beta}{\partial \tau}\frac{\partial x^\alpha}{\partial \tau} = 0 \end{aligned} \tag{2.22}$$

Recall:

$$\frac{\partial x^\nu}{\partial x^\alpha} = \delta^\nu_\alpha = \frac{\partial x^\nu}{\partial \xi^\mu}\frac{\partial \xi^\nu}{\partial x^\alpha} \tag{2.23}$$

Multiply Eq 2.22 by $\frac{\partial x^\nu}{\partial \xi^\mu}$ and use Eq 2.23 to get and δ in the first term of the R.H.S on Eq 2.22 that can then contracted:

$$\underbrace{\frac{\partial^2 x^\nu}{\partial \tau^2}}_{\frac{\partial u^\nu}{\partial \tau}} + \underbrace{\frac{\partial x^\nu}{\partial \xi^\mu}\frac{\partial^2 \xi^\mu}{\partial x^\beta \partial x^\alpha}}_{\text{Term a}}\underbrace{\frac{\partial x^\beta}{\partial \tau}}_{u^\beta}\underbrace{\frac{\partial x^\alpha}{\partial \tau}}_{u^\alpha} = 0 \tag{2.24}$$

Term a has three free indicies as μ is contracted, so it can be defined as an object with three independent indicies as, $\Gamma^\nu_{\beta\alpha}$. which is symmetric in α and β (as swapping the two 4-velocities does not have any effect). This gives the final form of the equation of motion in any arbitrary curved space-time:

$$\frac{\partial u^\nu}{\partial \tau} + \Gamma^\nu_{\alpha\beta}\frac{\partial x^\alpha}{\partial \tau}\frac{\partial x^\beta}{\partial \tau} = 0 \tag{2.25}$$

This is the geodesic equation. To understand what the $\Gamma's$ actually mean, the metric also needs to be transformed into these curved coordinates. The metric in general is given by:

$$\begin{aligned} ds^2 &= -d\tau^2 = \eta_{\mu\nu} d\xi^\mu d\xi^\nu \\ &= \underbrace{\eta_{\mu\nu}\frac{\partial \xi^\mu}{\partial x^\alpha}\frac{\partial \xi^\nu}{\partial x^\beta}}_{g_{\alpha\beta}} dx^\alpha dx^\beta \end{aligned} \tag{2.26}$$

This $g_{\alpha\beta}$ is known as the *metric tensor*. Since it is a symmetric 4×4 matrix, it has 10 independent parameters or ten components. It is not hard to see now, that if we take a derivative of $g_{\alpha\beta}$, the structure will be the same as it is for $\Gamma^\nu_{\alpha\beta}$. Let's define an object $\Gamma_{\mu\alpha\beta}$, such that:

$$\Gamma^{\nu}_{\alpha\beta} \equiv g^{\nu\mu}\Gamma_{\mu\alpha\beta} \tag{2.27}$$

These Γ's are called *Christoffel* symbols;

$$\begin{aligned} \Gamma_{\mu\alpha\beta} &= \text{Christoffel symbol of 1st kind} \\ \Gamma^{\nu}_{\alpha\beta} &= \text{Christoffel symbol of 2nd kind} \end{aligned} \tag{2.28}$$

Both the Christoffel symbols are symmetric in α and β, so any expansion for it must also be symmetric in α and β:

CLAIM 5.

$$\Gamma_{\mu\alpha\beta} \equiv \frac{1}{2}\left(\frac{\partial g_{\mu\alpha}}{\partial x^{\beta}} + \frac{\partial g_{\mu\beta}}{\partial x^{\alpha}} - \frac{\partial g_{\alpha\beta}}{\partial x^{\mu}}\right) \tag{2.29}$$

PROOF 5. The way to verify this claim is to simply compute the individual components on the R.H.S and see that they equal the L.H.S:

$$\begin{aligned} \frac{\partial g_{\mu\nu}}{\partial x^{\beta}} &= \frac{\partial}{\partial x^{\beta}}\left(\eta_{\kappa\sigma}\frac{\partial \xi^{\kappa}}{\partial x^{\mu}}\frac{\partial x^{\sigma}}{\partial x^{\alpha}}\right) \\ &= \eta_{\kappa\sigma}\frac{\partial}{\partial x^{\beta}}\left(\frac{\xi^{\kappa}}{\partial x^{\mu}}\frac{\partial \xi^{\sigma}}{\partial x^{\alpha}}\right) \\ &= \eta_{\kappa\sigma}\left(\frac{\partial^2 \xi^{\kappa}}{\partial x^{\beta}\partial x^{\mu}}\frac{\partial \xi^{\sigma}}{\partial x^{\alpha}} + \frac{\partial \xi^{\kappa}}{\partial x^{\mu}}\frac{\partial^2 \xi^{\sigma}}{\partial x^{\beta}\partial x^{\alpha}}\right) \end{aligned} \tag{2.30}$$

But notice:

$$\begin{aligned} \frac{\partial^2 \xi^{\kappa}}{\partial x^{\beta}\partial x^{\mu}} &= \overbrace{\frac{\partial \xi^{\kappa}}{\partial x^{\gamma}}\frac{\partial x^{\gamma}}{\partial \xi^{\delta}}}^{\delta^{\kappa}_{\delta}}\frac{\partial^2 \xi^{\delta}}{\partial x^{\beta}\partial x^{\mu}} \\ &= \frac{\partial \xi^{\kappa}}{\partial x^{\gamma}}\underbrace{\frac{\partial x^{\gamma}}{\partial \xi^{\delta}}\frac{\partial^2 \xi^{\delta}}{\partial x^{\beta}\partial x^{\mu}}}_{\Gamma^{\gamma}_{\beta\mu}} \end{aligned} \tag{2.31}$$

So infact:

$$\frac{\partial^2 \xi^{\kappa}}{\partial x^{\beta}\partial x^{\mu}} = \frac{\partial \xi^{\kappa}}{\partial x^{\gamma}}\Gamma^{\gamma}_{\beta\mu} \tag{2.32}$$

Substituting Eq 2.32 into 2.30:

$$\begin{aligned} \frac{\partial g_{\mu\alpha}}{\partial x^{\beta}} &= \eta_{\kappa\sigma}\left(\frac{\partial \xi^{\kappa}}{\partial x^{\gamma}}\Gamma^{\gamma}_{\beta\mu}\frac{\partial \xi^{\sigma}}{\partial x^{\alpha}} + \frac{\partial \xi^{\kappa}}{\partial x^{\mu}}\frac{\partial^2 \xi^{\sigma}}{\partial\partial x^{\beta}\partial x^{\alpha}}\right) \\ &= \eta_{\kappa\sigma}\frac{\partial \eta^{\kappa}}{\partial x^{\gamma}}\frac{\partial \xi^{\sigma}}{\partial x^{\alpha}}\Gamma^{\gamma}_{\beta\mu} + \underbrace{\eta_{\kappa\sigma}\frac{\partial \xi^{\kappa}}{\partial x^{\mu}}\frac{\partial^2 \xi^{\sigma}}{\partial x^{\beta}\partial x^{\alpha}}}_{\text{Term b}} \end{aligned} \tag{2.33}$$

One has to use the same trick as the one used in Eq 2.32 for Term b in the equation above to obtain:

$$\frac{\partial g_{\mu\alpha}}{\partial x^{\beta}} = g_{\gamma\alpha}\Gamma^{\gamma}_{\beta\mu} + g_{\mu\gamma}\Gamma^{\gamma}_{\alpha\beta} \tag{2.34}$$

The same method needs to be used for the second and third terms in Eq 2.29 to complete the proof.

So we see that the geodesic equation for any space-time geometry can be derived using the equivalence principle. Note that in general the gravitational field is not homogeneous, therefore the geodesics are not the same for all particles. This means that gravity cannot mimic acceleration, except at a specific point in space. This is infact the key in working with gravitational forces, one looks for tidal forces on objects as they cannot be produced by an accelerating reference frame.

In general a gravitational force is described by a general metric, $g_{\mu\nu}$, which is not necessarily related via the same transformation to the flat metric. So in general the line element in space-time is given by:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \tag{2.35}$$

This gives the true insight into general relativity; gravity is related to the geometry of space-time. Moreover, it is also a gauge theory, as the object that defined the curvature is a metric and a metric is not a physical observable[1]. Hence any changes made to the metric that leave the equations of motion unchanged are also equally valid solutions. To be more specific, the physical observables for a given geometry(metric) are encoded in a *Riemann tensor*, which contains second derivatives of the metric.

3. Newtonian limit

In Newtonian gravity, the defining object of a system is its gravitational potential energy, which is a scalar. In this transition from Newtonian gravity to Einstein's gravity, we have introduced 10 components (the components of the metric tensor) to define a system! So at first sight these two theories seem completely incompatible.

However, Newtonian gravity certainly works upto a certain limit. It has got us to the moon and most space travel only needs Newtonian theory, so how does general relativity approach the Newtonian limit? To understand this, lets look at the assumptions of Newtonian gravity:

- Slow (non-relativistic) motion:

$$\left| \frac{\partial x^\mu}{\partial \tau} \right| << \frac{\partial t}{\partial \tau} \tag{2.36}$$

  This basically means that time does not change much from proper time. If this is true, all the velocity terms in the geodesic equations can be neglected:

$$\frac{\partial^2 x^\mu}{\partial \tau^2} + \Gamma^\mu_{00} \left(\frac{\partial t}{\partial \tau} \right)^2 = 0 \tag{2.37}$$

- Stationary field:

$$\frac{\partial g_{\mu\nu}}{\partial t} \equiv 0 \tag{2.38}$$

  Therefore:

$$\Gamma^\mu_{00} = -\frac{1}{2} \frac{\partial g_{00}}{\partial x^\mu} g^{\nu\mu} \tag{2.39}$$

- Weak field:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} + \mathcal{O}(h^2) \tag{2.40}$$

  This is basically saying that the metric is just the Minkowski (flat) metric with a small perturbation $h_{\mu\nu}$ added to it, i.e the field strength is small.

[1]In the same that in quantum mechanics the object that carries the information of a system is its wavefunction, which is not a physical observable hence can be changed by an arbitrary phase factor.

To make further progress, the inverse of $g_{\mu\nu}$ is needed. If the metric is diagonal, the inverse is simply:

$$g_{\mu\nu} = \frac{1}{g^{\mu\nu}} \tag{2.41}$$

CLAIM 6. In general:

$$g^{\mu\nu} = \eta^{\mu\nu} - \eta^{\mu\nu} \equiv (g_{\mu\nu})^{-1} \tag{2.42}$$

PROOF 6. Since $g^{\mu\nu}$ is claimed to be the inverse, its product with $g_{\mu\nu}$ must be the unit matrix, this can be shown schematically as follows (one can also do the calculation using the contraction of indices and this will yield the same solution):

$$\begin{aligned} g_{\mu\nu}g^{\mu\nu} &= (\eta_{\mu\nu} + h_{\mu\nu})(\eta^{\mu\nu} - h^{\mu\nu}) \\ &= \eta_{\mu\nu}\eta^{\mu\nu} + h_{\mu\nu}\eta^{\mu\nu} - \eta_{\mu\nu}h^{\mu\nu} - h_{\mu\nu}h^{\mu\nu} \\ &= \delta + h - h - (h^2) \\ &= \delta \end{aligned} \tag{2.43}$$

Look back to Eq 2.39. There are two terms, one contains a derivative and one without; Since h is small, we keep terms only that are linear in h. So under the weak field approximation, the derivative of η is obviously zero since it is a constant, hence the only term that contributes in the derivative will be h:

$$\Gamma^{\mu}_{00} = -\frac{1}{2}\frac{\partial h_{00}}{\partial x^{\mu}} g^{\nu\mu} \tag{2.44}$$

and since the h^2 terms are being ignored:

$$\begin{aligned} \Gamma^{\mu}_{00} &= -\frac{1}{2}\frac{\partial h_{00}}{\partial x^{\mu}}(-h^{\nu\mu} + \eta_{\mu\nu}) \\ &= -\frac{1}{2}\frac{\partial h_{00}}{\partial x^{\mu}}\eta_{\mu\nu} + \frac{1}{2}\frac{\partial h_{00}}{\partial x^{\mu}}h^{\nu\mu} \\ &= -\frac{1}{2}\frac{\partial h_{00}}{\partial x^{\mu}}\eta_{\mu\nu} \end{aligned} \tag{2.45}$$

Now Eq 2.37 can be written in terms of its separate components as:

$$\mu = 0 \Rightarrow \frac{\partial^2 x^0}{\partial \tau^2} + \Gamma^0_{00}\left(\frac{\partial t}{\partial \tau}\right)^2 \tag{2.46}$$

But:

$$\begin{aligned} \Gamma^{0}_{00} &= -\frac{1}{2}\frac{\partial h_{00}}{\partial x^{\nu}}\eta^{\nu 0} \\ &= 0 \quad \forall \nu \end{aligned} \tag{2.47}$$

The last line holds as $\eta^{\nu 0} \neq 0$ only for $\nu = 0$ and $\partial_\nu \equiv 0$ for $\nu = 0$. Therefore:

$$\frac{\partial^2 x^0}{\partial \tau^2} = \frac{\partial^2 t}{\partial \tau^2} = 0 \tag{2.48}$$

This means:

$$\frac{dt}{d\tau} = C \qquad C \in \{\mathbb{R}\} \tag{2.49}$$

So time flows at some constant rate, which once can always absorb into the definition of τ to give the proportionality constant as 1. Now for the spatial components we have:

$$\mu = i(\in \{1,2,3\}) \Rightarrow \frac{\partial^2 x^i}{\partial \tau^2} + \Gamma^i_{00}\left(\frac{\partial x^0}{\partial \tau}\right)^2 = 0 \tag{2.50}$$

Substitute Eq 2.49 into 2.50:

$$\frac{\partial^2 x^i}{\partial \tau^2} + \Gamma^i_{00}C^2 = 0 \tag{2.51}$$

The first term can be expanded using the chain rule:

$$\begin{aligned} \frac{\partial^2 x^i}{\partial t^2}\frac{\partial t^2}{\partial \tau^2} + \Gamma^i_{00}c^2 &= 0 \\ \frac{\partial^2 x^i}{\partial t^2}C^2 + \Gamma^i_{00}c^2 &= 0 \end{aligned} \tag{2.52}$$

Substituting Eq 2.45 into Eq 2.52:

$$\frac{\partial^2 x^i}{\partial t^2} = \frac{1}{2}\frac{\partial h_{00}}{\partial x^i} \tag{2.53}$$

In Newtonian gravity, the corresponding relation is:

$$\frac{\partial^2 x^i}{\partial t^2} = -\frac{\partial \phi}{\partial x^i} \tag{2.54}$$

where ϕ is the gravitation potential. By comparing Eq 2.54 to Eq 2.53:

$$\phi \equiv -\frac{1}{2}h_{00} \tag{2.55}$$

Now recall that the gravitational potential follows the Laplace equation:

$$\nabla^2 \phi = 4\pi G\rho \tag{2.56}$$

and by comparing Eq 2.56 and Eq 2.55:

$$\nabla^2 h_{00} = -\frac{8\pi G\rho}{c^2} \tag{2.57}$$

Therefore:

$$g_{00} = -\left(1 + \frac{2\phi}{c^2}\right) \tag{2.58}$$

where c^2 has been put back to get the correct units. Note that $\frac{\phi}{c^2} = \frac{GM}{c^2 r} << 1$, for the Newtonian approximation to hold. In essence this is the region in which one makes a transition from the Newtonian to the general relativity regime. As an example, one can look at the value of $\frac{\phi}{c^2}$ at the surface of several objects.

Object	$\frac{\phi}{c^2}$	Newton or Einstein?
Proton	10^{-39}	$<< 1$ therefore Newton
Earth	10^{-9}	$<< 1$ therefore Newton
Sun	10^{-6}	$<< 1$ therefore Newton
Neutron star	$10^{-2} - 10^{-1}$	Transition regime
Black holes	$10^{-1} - 1$	Einstein

TABLE 5. The gravitational approximation for a range of physical objects

4. Einstein gravitational red-shift

The derivation presented in section 2.1 is for a weak gravitational field, hence we can assume a homogenous gravitational field. This is a very specific derivation for the red-shift. For non-homogenous fields, this needs to be modified. In a strong gravitational field, time slows down. To see how, look at the proper time:

$$d\tau = \sqrt{-g_{\mu\nu}\frac{\partial x^\mu}{\partial t}\frac{\partial x^\mu}{\partial t}}dt \tag{2.59}$$

In units where $c \equiv \hbar \equiv 1$, the frequency of light in a primed (moving observer) frame and in an unprimed frame (stationary observer) is given by:

$$\begin{aligned} d\tau &= \frac{1}{\nu} \\ d\tau' &= \frac{1}{\nu'} \end{aligned} \tag{2.60}$$

The ratio of frequencies is:

$$\frac{\nu'}{\nu} = \frac{d\tau}{d\tau'} = \frac{\sqrt{-g_{\mu\nu}\frac{\partial x^\mu}{\partial t}\frac{\partial x^\nu}{\partial t}}dt}{\sqrt{-g_{\mu\nu}\frac{\partial x'^\mu}{\partial t'}\frac{\partial x'^\nu}{\partial t'}}dt'} \tag{2.61}$$

Suppose that the observers in both frames are stationary in their respective frames, i.e all velocities will be 0. The only non-trivial solution will be when $\mu = 0$:

$$\frac{\nu'}{\nu} = \frac{\partial \tau}{\partial \tau'} = \sqrt{\frac{g_{00}(x)}{g'_{00}(x')}}\frac{\partial t}{\partial t'} \tag{2.62}$$

Next, we assume that the space-time is stationary (i.e geometry not changing with time). Then it is possible to choose a global coordinate time, so $t \equiv t'$ and the expression above simplifies further:

$$\frac{\nu'}{\nu} = \sqrt{\frac{g_{00}(x)}{g_{00}(x)}} \tag{2.63}$$

This is the gravitational red-shift, which is a more general result than the expression in Eq 2.15. To recover the result of Eq 2.15, first substitute for:

$$g_{00}(x) = -(1 + 2\phi') \tag{2.64}$$

Therefore:

$$\frac{\nu'}{\nu} = \sqrt{\frac{1+2\phi}{1+2\phi'}} \tag{2.65}$$

For small ϕ, we can Taylor expand:

$$\frac{\nu'}{\nu} \approx (1+\phi)(1-\phi) \tag{2.66}$$

The red-shift is defined as usual:

$$\begin{aligned} z &= \frac{\Delta\nu}{\nu} \\ &= \frac{\nu' - \nu}{\nu} \\ &= \frac{\nu'}{\nu} - 1 \\ &= \phi - \phi' \end{aligned} \tag{2.67}$$

If the gravitational field is homogenous, the $\phi = gh$:

$$z = g(h - h') \equiv \frac{g(h - h')}{c^2} \tag{2.68}$$

Which is the same as Eq 2.15.

5. Gravitational field theory

Returning to Einstein's goal of forming a field theory of gravity. The first thing to ask is, what are the ingredients of a field theory. The very first field theory has already been discussed, it was Maxwell's theory of EM. The first thing one needs is obviously *fields.* In Maxwell's theory, this was A^{μ}.

The next thing is to find out the dynamics of the field, which is described by an *equation of motion.* The equations of motion are generally found by using the Euler-Lagrange equation, once a Lagrangian for the theory is found. In EM the equation of motion is:

$$\frac{\partial p^{\mu}}{\partial \tau} = qF^{\mu\nu}U_{\nu} \tag{2.69}$$

Finally, one needs a field equation; this provides information about how a source of the fields actually produces the fields. In EM these are just Maxwell's equations as seen in Eq 1.47 and Eq 1.53. In Newtonian theory, the field is ϕ, the equation of motion for gravity is:

$$\frac{\partial \vec{p}}{\partial t} = -\nabla\phi \tag{2.70}$$

and the field equation is:

$$\nabla^2\phi = \rho \tag{2.71}$$

where ρ is the mass density.

For Einstein's theory, the field is $g_{\mu\nu}$ and the equations of motion is:

$$\frac{\partial^2 x^{\mu}}{\partial t^2} + \Gamma^{\mu}_{\alpha\beta}\frac{\partial x^{\alpha}}{\partial t}\frac{\partial x^{\beta}}{\partial t} = 0 \tag{2.72}$$

The field equation is actually a set of equations called Einstein's field equations:

$$G_{\mu\nu} = 8\pi T_{\mu\nu} \tag{2.73}$$

which will now be derived. $T_{\mu\nu}$ is the source of the field and is known as the stress energy tensor, and since the field in this case is the metric tensor and the metric tensor represents the geometry; we see that the source of the field gives rise to curvature of space-time and this is the corner stone statement of general relativity.

	Maxwell	Newton	Einstein
Field	A^μ	ϕ	$g_{\mu\nu}$
Equation of motion	$\frac{\partial p^\mu}{\partial \tau} = qF^{\mu\nu}U_\nu$	$\frac{\partial \vec{p}}{\partial t} = -\nabla\phi$	$\frac{\partial^2 x^\mu}{\partial \tau}^2 + \Gamma^\mu_{\alpha\beta}\frac{\partial x^\alpha}{\partial \tau}\frac{\partial x^\beta}{\partial \tau} = 0$
Field equations	$\partial_\mu F^{\mu\nu} = -\mu_0 j^\nu$ and $\partial_{[\mu}F_{\nu\alpha]} = 0$	$\nabla^2\phi = \rho$	$G_{\mu\nu} = 8\pi T_{\mu\nu}$

Table 6. Summary of classical field theories

CHAPTER 3

Curved space-time: Riemannian geometry

Everything that has been described so far has been in terms of special cases or in the limit of weak fields etc. The true power of Einstein's theory comes from its geometry and some of the previous results will be re-derived in a more general way in this chapter, as well as providing the mathematical tools for working with curved space-time.

1. Manifolds

A manifold looks like flat space "locally", i.e it has the differential structure of $\mathbb{R}^n$ locally, but is does not, in general have its global properties. Its formal definition is:

DEFINITION 1. A manifold is a set of points together with a collection of subsets, $\{O_\alpha\}$ such that:

(1) Each point, $p \in M(Manifold)$ lies in at least one of the subsets O_α; which implies $\{O_\alpha\}$ covers the entire manifold.
(2) For each α, there is a one to one, *onto* map, ϕ_α, which takes O_α to another subset U_α, where U_α is an open subset of $\mathbb{R}^n$
(3) If any two sets O_α overlap; $O_\alpha \cap O_\beta \neq 0$, then the map $\psi_\beta \cdot \psi_\alpha^{-1}$ is infinitely differentiable, C^∞.

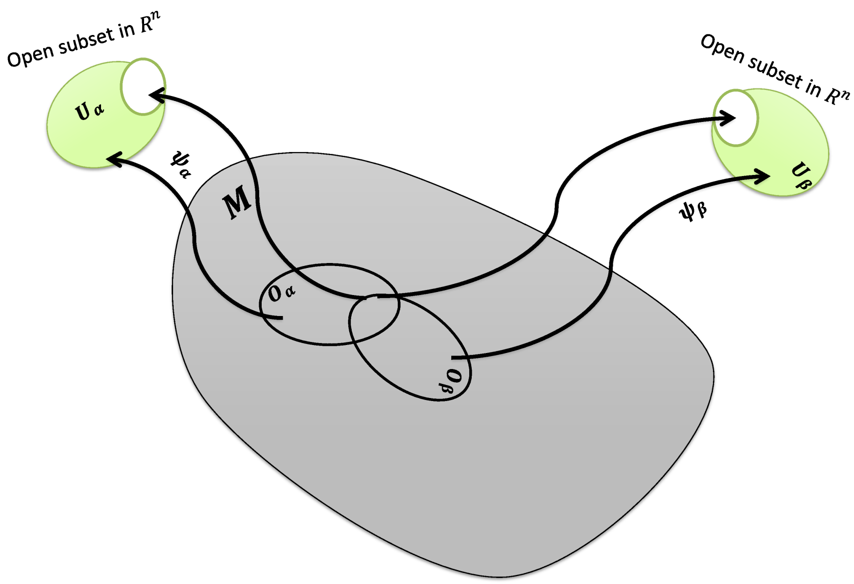


FIGURE 7. Manifold with subsets mapping to $\mathbb{R}^n$

From O_β, there is a map ψ_β which goes to $\mathbb{R}^n$ ($n = Dim(M)$) into an open subset U_β in $\mathbb{R}^n$ space. There is a similar map for U_α in $\mathbb{R}^n$ from O_α in M. The overlap region could be separated into the U_α and U_β regions so to combine it, one would have to do the transformation:

$$\psi_\alpha \cdot \psi_\alpha^{-1} \quad \text{or} \quad \psi_\beta \cdot \psi_\alpha^{-1} \tag{3.1}$$

So $\psi_\beta \cdot \psi_\alpha^{-1}$ takes a point in $\mathbb{R}^n$ and moves it to another point in $\mathbb{R}^n$, so it must just be an ordinary (multidimensional in general) function that is infinitely differentiable. To summarise, Manifolds are made of pieces that look like open subsets of $\mathbb{R}^n$, which are smoothly sewn together.

DEFINITION 2. ψ_α is called a *chart* (usually in mathematical literature) or *coordinate system* (usually in physics literature). Sometimes one would just write x_μ, instead of ψ_α as it represents the coordinate system.

DEFINITION 3. A collection of maps, $\{O_\alpha, \psi_\alpha\}$, is usually called an *atlas*

EXAMPLE 1. A simple example of a manifold is $\mathbb{R}^n$. It satisfies all the properties in the definition of a manifold and the proof is trivial.

EXAMPLE 2. A sphere in n dimensions, $\mathbb{S}^n$; defined by the equation:

$$\mathbb{S}^n = \{(x_1, x_2, ..., x_n) \in \mathbb{R}^n | x_1^2 + x_2^2 + ...x_n^2 = 1\} \tag{3.2}$$

EXAMPLE 3. As a more concrete example, consider the 2-sphere embedded in a 3 dimensional Euclidean space. The general Eq 3.2 now becomes:

$$\mathbb{S}^3 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_1^2 + x_2^2 + x_3^2 = 1\} \tag{3.3}$$

We can define a set of maps to form an atlas. There are many ways to choose the set of maps, as an example lets consider the following six maps, $\{\mathcal{M}_1, \mathcal{M}_2 ... \mathcal{M}_6\}$ that form an atlas for this 2-sphere:

- This covers the entire *northern hemisphere*:

$$\mathcal{M}_1 : \{(x_1, x_2, x_3) \in \mathbb{S}^2 | x_3 > 0\} \tag{3.4}$$

- This covers the entire *southern hemisphere*:

$$\mathcal{M}_2 : \{(x_1, x_2, x_3) \in \mathbb{S}^2 | x_3 < 0\} \tag{3.5}$$

- This covers the entire *west hemisphere*:

$$\mathcal{M}_3 : \{(x_1, x_2, x_3) \in \mathbb{S}^2 | x_1 > 0\} \tag{3.6}$$

- This covers the entire *northern hemisphere*:

$$\mathcal{M}_4 : \{(x_1, x_2, x_3) \in \mathbb{S}^2 | x_1 < 0\} \tag{3.7}$$

  Note that this still does not cover the entire sphere. There are two points (0,1,0) and (0,-1,0) on the equator that are not covered by these maps. Thus these two maps need to be specified separately:

- Covers the point (0,1,0):

$$\mathcal{M}_5 : \{(x_1, x_2, x_3) \in \mathbb{S}^2 | x_1 = x_3 \equiv 0, x_2 = 1\} \tag{3.8}$$

- Covers the point (0,-1,0):

$$\mathcal{M}_6 : \{(x_1, x_2, x_3) \in \mathbb{S}^2 | x_1 = x_3 \equiv 0, x_2 = -1\} \tag{3.9}$$

There is infact a much better way to parametrize the maps, which is in terms of stereographic coordinates, which only requires two maps.

EXAMPLE 4. A sheet in 2 dimensions with a 1-D chord bisecting the plane of the sheet perpendicularly is not a manifold. To see why, lets define a subset, $\{\mathcal{A}_{chord}\}$:

$$\{\mathcal{A}_{chord}\} = \text{Set of all points on the chord} \tag{3.10}$$

and a subset $\{\mathcal{A}_{plane}\}$:

$$\{\mathcal{A}_{plane}\} = \text{Set of all points on the plane} \tag{3.11}$$

Now, it is not possible to construct a 1 to 1 mapping; $\mathcal{A}_{chord} \mapsto \mathcal{A}_{plane}$, since the dimensions of the chord and the plane are not the same.

EXAMPLE 5. A Riemann surface of genus g, is essentially a two-torus with g holes defining its topology (as supposed to one hole for a "normal" torus). A torus with genus 1 is also a manifold.

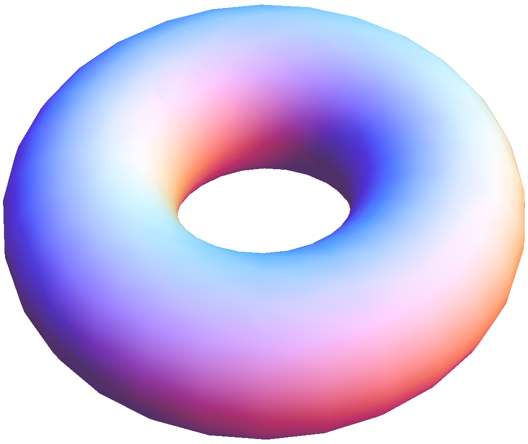


FIGURE 8. Torus with genus 1 manifold

EXAMPLE 6. If a cone is connected to another cone as shown in Figure 9, it is also not a manifold due to the apex point where the two cones intersect. This is because at the point, the mapping from the manifold to R^n is not C^∞.

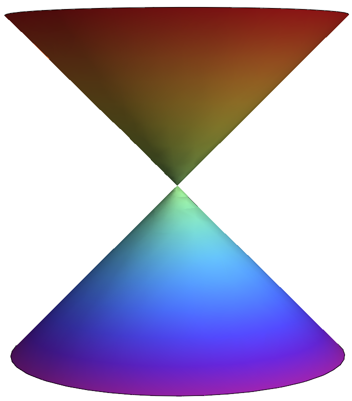


FIGURE 9. This structure is not a manifold due to the point at the apex.

General relativity only makes sense on smooth manifolds, as one does not expect or observe any massive spikes in the curvature of space time from one point to the next. Other theories like string theory, do not require this condition. This is because strings are 1-D objects, whereas in general relativity the fundamental particles are taken to be points of 0 dimension. Since a string is an extended object it can smooth out singularities. The fundamental objects in string theory are called "Orbifolds", for example if one considers an orbifold which is a one dimensional surface, with a point of origin, say 0. The orbifold can be parametrized by a variable, say x, and every point x will have a one to one mapping to every other point $-x$ expect the origin point 0. So this point becomes singular and is a problem for point like particles in this manifold. However, a string can simply avoiding that point by going over that point and joining the manifold at two points x_1 and $-x_1$.

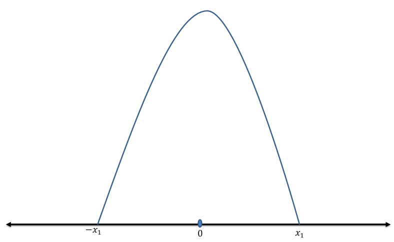


FIGURE 10. String on orbifold

Returning to manifolds, if there is a map that takes one manifold M to another manifold M'; $f : M \mapsto M'$, where each manifold has a map to $\mathbb{R}^n$ via:

$$\psi : M \mapsto \mathbb{R}^n$$
$$\psi' : M' \mapsto \mathbb{R}^n \tag{3.12}$$

This means there is another way to get from M to M' via the application of the three maps:

$$\psi \cdot f \cdot \psi'^{-1} \text{ Takes } M \mapsto M' \tag{3.13}$$

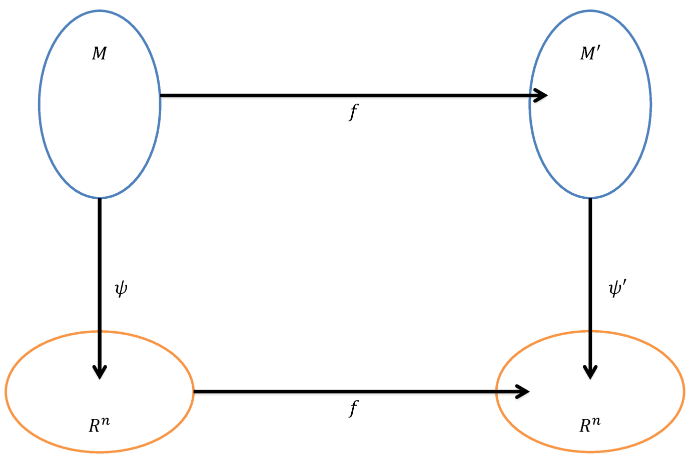


FIGURE 11. Schematic of mappings to illustrate Diffeomorphism

DEFINITION 4. If a condition is imposed on Eq 3.13 such that the mapping is C^∞, 1 to 1, onto (i.e all points of the image are covered by the mapping), and the inverse of the derivative also exists, then the mapping $M \mapsto M'$ is called a *Diffeomorphism.*

2. Tensors

Tensors can be thought of as actors on the stage of manifolds. If there exists a scalar function that maps a manifold M to $\mathbb{R}^n$, one can require that the function is independent of the parametrisation (independent of which coordinate system is chosen):

$$f(x(p)) \equiv f'(x'(p)) \quad \text{or} \quad f'(x') \equiv f(x) \tag{3.14}$$

This is the key idea behind tensors.

2.1. Tangent vectors. The simplest thing to think about when thinking about vectors is a directional derivative at a point.

DEFINITION 5. In $\mathbb{R}^n$, a vector v^μ has some components in a specific coordinates:

$$v^\mu = (v^1, v^2, ..., v^n) \tag{3.15}$$

For any vector of this form, one can define in a 1 to 1 mapping a directional derivative operator:

$$\hat{v} \equiv v^\mu \frac{\partial f}{\partial x^\mu} = R \in \mathbb{Z} \tag{3.16}$$

For example, if one considers just the 1 dimensional case, with $x^\mu \equiv x$, the $\hat{v}f(x)$ will give an indication of how the function $f(x)$ is changing in the x-direction.

DEFINITION 6. Let F be the collection of C^∞ on a manifold M. Then the tangent vector, V, at a point $p \in M$ is a map:

$$VF \mapsto \mathbb{R} \tag{3.17}$$

and it has the following properties:

$$\text{Linearity} \Rightarrow V(af + bg) = aV(f) + bV(a) \tag{3.18}$$

$$\text{Leibniz} \Rightarrow V(fg) = V(f)g(p) + V(g)f(p) \tag{3.19}$$

THEOREM 1. The set of tangent vectors at $p \in M$ forms a *tangent vector space*, $T_p(m)$. The vector space has the same dimensionality as M, with the basis $\frac{\partial}{\partial x^\mu}$, so any vector V can be expressed as a linear combination of the basis vectors:

$$V = V^\mu \frac{\partial}{\partial x^\mu} \tag{3.20}$$

where V^μ represents the components of the vector in the coordinate system that corresponds to the basis $\frac{\partial}{\partial x^\mu}$.

This vector can now be written a different coordinate system:

$$V = V^\mu(x) \frac{\partial x'^\nu}{\partial x^\mu} \left(\frac{\partial}{x^{\nu\prime}} \right) \tag{3.21}$$

This gives the transformation for components of the vector into a different coordinate system is given by:

$$V'^\nu(x') = \frac{\partial x'^\nu}{\partial x^\mu} V^\mu(x) \tag{3.22}$$

Until now the vectors and vectors spaces are defined at a given point, p, in the manifold. This construction can be extended to the whole manifold by defining a vector field, $\mathbb{V}$.

Definition 7. Define a collection of all the tangent vectors V at each point $p \in M$, by a tangent vector field, $\mathbb{V}$:

$$\mathbb{V} \equiv \{V_p \in T_pM,\ \forall p \in M,\ V(f) = \text{Smooth}\} \tag{3.23}$$

where V_p is a tangent vector at point p and T_p is a tangent vector space at point p in the manifold. This collection of vectors on the whole manifold is known as a *field*. The last condition in the definition is that the vectors must transform smoothly on the manifold, meaning that if a vector is pointing $\uparrow$, say, at point p_1, then at point $p_2 = p_1 + \delta p_1$, it cannot point in $\downarrow$, say, as this would involve a jump, which would not be smooth and hence not satisfy the definition of a manifold.

Definition 8. The set of all T_pM, $\{T_pM\}, \forall p$, is known as a *tangent bundle*, $T(M)$.

Definition 9. A *cotangent vector*, $\vec{\omega}$, at a point p is a map which, for any vector in the tangent space, assigns a number:

$$\begin{aligned} \vec{v}_1 &\to a_1 \\ \vec{v}_2 &\to a_2 \\ \vec{v}_n &\to a_n \end{aligned} \tag{3.24}$$

where $v_n \in T_pM$ and $n = dim(T_pM)$ and $a_n \in \mathbb{R}$. These cotangent vectors, form a *cotangent vector space*, denoted by T_p^*M, which is a dual spcae to the tangent space, with a basis dx^μ. The basis is defined by:

$$x^\mu \left(\frac{\partial}{\partial x^\nu}\right) \equiv \delta^\mu_\nu \left(\equiv \frac{\partial x^\mu}{\partial x^\nu}\right) \tag{3.25}$$

Now, one can write:

$$\omega = \omega_\mu dx^\mu \tag{3.26}$$

i.e a linear combination (same as was done for the tangent vectors).

To change coordinate systems:

$$\begin{aligned} \omega &= \omega_\mu dx^\mu \\ &= \omega_\mu \frac{\partial x^\mu}{\partial x'^\nu} \partial x'^{\mu} \end{aligned} \tag{3.27}$$

Therefore the transformation of the components is given by:

$$\omega'_\nu(x') = \frac{\partial x^\mu}{\partial x'^\nu} \omega_\mu(x) \tag{3.28}$$

This shows that the tangent vectors and the cotangent vectors transform in the same way, except that the transformation matrix is now the inverse of each other.

In general, a tensor of type (k, l) is a multi-linear map:

$$T : \underbrace{T_p^* \times T_p^* ... \times T^*}_{k \text{ times}} \times \underbrace{T_p \times T_p \times ... \times T_p}_{l \text{ times}} \to \mathbb{R} \tag{3.29}$$

Which can be re-written as:

$$T'^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \beta_l}(x') = \frac{\partial x'^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial x'^{\alpha_k}}{\partial x^{\mu_k}} \frac{\partial x_1^\beta}{\partial x'^{\nu_1}} \cdots \frac{\partial x^{\alpha_k}}{\partial x'^{\nu_l}} T^{\mu_1 \ldots \mu_k}_{\nu_1 \ldots \nu_k}(x) \tag{3.30}$$

The k indicies transform like (tangent) vectors and the l indicies transform like co(tangent)-vectors.

In general, co-vectors are different to vectors, however, once a metric has been introduced, the vectors can be transformed into co-vectors by contracting with the metric tensor and vice-verse. A metric acts as a 1 to 1 map between vectors and co-vectors.

2.2. Tensor algebra. Let T and S be tensors, then they can be added since they are linear (as long as the indicies run over the sum number of values, i.e the dimensions are the same).

DEFINITION 10. Once can also define a product of tensors:

$$T^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \beta_l} S^{\mu_1 \ldots \mu_{k'}}_{\nu_1 \ldots \nu_{l'}} \tag{3.31}$$

sometimes simply referred to as a *tensor product.* The result of the product of tensors is also a tensor of rank $(k+k'\ l+l')$, where T is a tensor of rank (k,l) and S is a tensor of rank (k',l').

DEFINITION 11. Tensors can be *contracted* when two of the indicies of the two tensors are the same and hence summed over (i.e they are just dummy indicies):

$$T^{\alpha_1 \ldots \alpha_k}_{\beta_1 \ldots \beta_k} S^{\mu_1 \ldots \mu_{k'}}_{\nu_1 (\equiv \alpha_1) \ldots \nu_{l'}} = \text{Rank}(k+k'-1, l+l'-1) \tag{3.32}$$

where α_1 is summed over

If a tensor is zero in one coordinate system, then it is zero in all coordinate system. In fact a tensor is the *same* in all coordinate systems by definition. This is what makes them very useful, as once the laws of physics are written in terms of tensors they can be applied and compared in any coordinate system.

2.3. Connection. Laws of physics are generally defined by differential equations, so it is important to formulate a derivative of a tensor, i.e tensor calculus. As an introductory example, the derivative of a scalar, ϕ, is a tensor of Rank(0,1)

CLAIM 7.

$$\partial_\mu \phi(x) \equiv \text{Tensor}(0,1) \tag{3.33}$$

PROOF 7.

$$\begin{aligned} \frac{\partial \phi(x)}{\partial x^\nu} &= \frac{\partial \phi(x)}{\partial x'^\nu} \frac{\partial x'^\nu}{\partial x^\mu} \\ &= \frac{\partial x'^\nu}{\partial x^\mu} \frac{\partial}{\partial x'^\nu} \phi(x) \end{aligned} \tag{3.34}$$

So the components transform as:

$$\frac{\partial x'^\nu}{\partial x^\mu} \phi(x) \tag{3.35}$$

which is the transformation of rank(0,1) tensor.

Now consider the derivative of a vector, V'^α:

$$\begin{aligned} \frac{\partial}{\partial x'^\mu} V'^\alpha(x') &= \frac{\partial}{\partial x'^\mu} \left(\frac{\partial x'^\alpha}{\partial x^\nu} V^\nu \right) \\ &= \underbrace{\frac{\partial x'^\alpha}{\partial x^\nu} \frac{\partial x^\beta}{\partial x'^\mu} V^\nu_\beta}_{\text{Term 1}} + \underbrace{\frac{\partial x^\beta}{\partial x'^\mu} \frac{\partial^2 x'^\alpha}{\partial x^\beta \partial x^\nu} V^\nu}_{\text{Term 2}} \end{aligned} \tag{3.36}$$

Term 1 shows that the first part transforms as a vector, the second part transforms as a co-vector, which is expected for a tensor of rank(1,1). Term 2 has a part with a second derivative and this creates a problem as taking its derivative, in say, the Cartesian coordinate system would make it zero. However, one can choose another coordinate system, in which this will not be zero, hence it is not a tensor!.

This was a problem that Riemann set about solving in his PHD. The solution he came up with was the idea of a *covariant* derivative.

DEFINITION 12. The covariant derivative is defined as:

$$\nabla_\mu V^\mu \equiv \partial_\mu V^\nu + \Gamma^\nu_{\alpha\mu} V^\alpha \tag{3.37}$$

where $\Gamma^\nu_{\alpha\mu}$ is called a *connection*. Note that this is not the same as the Christoffel symbols, it is more general. We require that $\nabla_\mu V^\nu$ is a tensor and the $\partial_\mu V^\nu$ term is already known. This provides a condition on the $\Gamma^\nu_{\alpha\mu}$, such that $\nabla_\mu V^\nu$ is a tensor.

Similarly one can define a covariant derivative for a covariant object (i.e indicies at the bottom):

$$\nabla_\alpha \omega_\alpha \equiv \partial_\alpha \omega_\beta - \Gamma^\gamma_{\beta\alpha}\omega_\gamma \tag{3.38}$$

More generally, for any tensor:

$$\nabla_\alpha T^{\gamma_1\ldots\gamma_k}_{\beta_1\ldots\beta_l} \equiv \partial_\alpha T^{\gamma_1\ldots\gamma_k}_{\beta_1\ldots\beta_l} + \Gamma^{\gamma_1}_{\delta_1\alpha} T^{\delta_1\ldots\delta_k}_{\beta_1\ldots\beta_l} - \Gamma^{\rho_1}_{\beta_1\alpha} T^{\gamma_1\ldots\gamma_k}_{\rho_1\ldots\rho_l} \tag{3.39}$$

DEFINITION 13. It turns out, the condition on Γ is:

$$\Gamma'^{\alpha'}_{\mu'\beta'} \equiv \frac{\partial x'^\alpha}{\partial x^\alpha}\frac{\partial x^\beta}{\partial x'^\beta}\frac{\partial x^\mu}{\partial x'^\mu}\Gamma^\alpha_{\mu\beta} - \underbrace{\frac{\partial^2 x'^\alpha}{\partial x^\alpha \partial x^\beta}\frac{\partial x^\beta}{\partial x'^\beta}\frac{\partial x^\mu}{\partial x'^\mu}}_{\text{Term 2}} \tag{3.40}$$

This is also not a tensor due to the second derivative in Term 2.

DEFINITION 14. The anti-symmetrization of the Γ indicies is defined as:

$$\Gamma^\alpha_{[\beta,\gamma]} \equiv \frac{1}{2}(\Gamma^\alpha_{\beta\gamma} - \Gamma^\alpha_{\alpha\beta}) = -\frac{1}{2}T^\alpha_{\beta\alpha} \tag{3.41}$$

This object is known as the *torsion*.

Infact, this object is a tensor as the Term 2 in Eq 3.40 will be equal and opposite due to the symmetrization of the γ and β indices and hence cancel, leaving the transformation property of tensor. In general relativity, torsion is assumed to be zero. In general, particles with spin will create torsion.

Note that throughout this formalism, the metric has never been mentioned. This means that all these objects will exist independently of a metric. Once a metric is introduced, these expressions can be re-derived with the metric and they are somewhat easier.

3. The metric

Recall that in Minkowski space-time, the Lorentz invariant distance is given by:

$$\eta_{\mu\nu} dx^\mu dx^\nu = d\vec{x}^2 - c^2 dt^2 \tag{3.42}$$

This distance will be observed to be the same for every observer that is related by a Lorentz transformation:

$$x^\mu \to x'^\mu = \Lambda^\mu_\nu x^\nu \tag{3.43}$$

The linear transformations Λ^μ_ν, are defined by:

$$\Lambda^\mu_\nu \Lambda^\alpha_\beta \eta_{\mu\alpha} = \eta_{\nu\beta} \tag{3.44}$$

This equation allows 6 independent parameters (as η is a diagonal metric); 3 boosts and 3 rotations. The Lorentz group is a 6 parameter group. Einstein found that all the Laws of physics are written in tensor form and are invariant under Lorentz transformations.

But Lorentz transformations are just linear transformations as stated above. It is natural to think about non-linear transformations, i.e frames that are not moving at a constant speed relative to each other, but frames that are accelerating. Consider the transformation:

$$x' = x + \frac{1}{2}at^2 \tag{3.45}$$

and a particle is moving in the un-primed frame with the equation of motion:

$$\frac{d^2x_p}{dt^2} = 0 \tag{3.46}$$

i.e no acceleration of the particle. Then in the primed reference frame, it would look like:

$$\frac{\partial^2 x'_p}{\partial t^2} = a \tag{3.47}$$

i.e there seems to be a force. This is what lead Einstein to his Equivalence principle, that was discussed in depth in Chapter 2.

3.1. Curved space-time. A curved space-time must locally look like flat space-time. The trick with gravity is to combine all the flat space-time regions present locally, into a curved space-time that shows gravity emerging. In curved space-time distances are measured using a metric, in the same way as the Minkowski mertic:

$$g_{\mu\nu}(x)dx^\mu dx^\nu \tag{3.48}$$

where $g_{\mu\nu}$ is a symmetric metric tensor, which has 10 independent parameters and follows the properties:

- $$\forall\ g_{\mu\nu}\ \exists\ (g_{\mu\nu})^{-1} \tag{3.49}$$
  Which implies that the determinant of the metric tensor is not zero:
  $$|g_{\mu\nu}| \neq 0 \tag{3.50}$$
- It has three positive and one negative eigenvalues (difference between space and time), this type of metric is called *pseudo - Riemannian* (As Riemann invented the mathematics for curved space. He was working with higher dimensions of Euclidean space, so the components of the metric did not have a change in sign).
- We say that a space possessing such a metric is locally Minkowski, because one can always choose coordinates such that the metric at one point is equal to the Minkowski metric.

To see why the metric can be set to the flat (Minkowski) metric locally, consider the metric at a particular point, x_p:

$$g_{\mu\nu}(x_p) \tag{3.51}$$

since $g_{\mu\nu}$ is symmetric, $\exists\ \mathbb{O}^\mu_\alpha$ such that:

$$\mathbb{O}^\mu_\alpha \mathbb{O}^\nu_\beta g_{\mu\nu}(x_p) = D_{\alpha\beta} \tag{3.52}$$

where $D_{\alpha\beta}$ is a diagonal matrix. In other words, this condition basically states that the metric is diagonalizable via:

$$\mathbb{O} g \mathbb{O}^T = D \tag{3.53}$$

This implies:

$$g = \mathbb{O}^T D \mathbb{O} \tag{3.54}$$

The diagonal matrix takes the form:

$$D = \begin{pmatrix} -l_0^2 & 0 & 0 & 0 \\ 0 & l_1^2 & 0 & 0 \\ 0 & 0 & l_2^2 & 0 \\ 0 & 0 & 0 & l_3 \end{pmatrix} \tag{3.55}$$

where l_μ are the eigenvalues. Defining:

$$L \equiv \begin{pmatrix} l_0 & 0 & 0 & \\ 0 & l_1 & 0 & 0 \\ 0 & 0 & l_2 & 0 \\ 0 & 0 & 0 & l_3 \end{pmatrix} \tag{3.56}$$

Therefore the metric can be written as:

$$g = \mathbb{O}^T L \eta L^T \mathbb{O} \tag{3.57}$$

or equivalently:

$$dx^\mu dx^\nu g_{\mu\nu} = dy^\mu dy^\nu \eta_{\mu\nu} \tag{3.58}$$

with:

$$y = L\mathbb{O}x \tag{3.59}$$

where we have used:

$$dx^T g dx = \underbrace{dx^T \mathbb{O}^T L}_{dy^T \eta dy} \eta L \mathbb{O} dx \tag{3.60}$$

Therefore we can transform $g_{\mu\nu}$ into $\eta_{\mu\nu}$ using orthogonal matrices $\mathbb{O}$ and L, which have 6 and 4 parameters respectively. This means that at every point in space-time one can fix the components such that the metric tensor $g_{\mu\nu}$ is a flat metric as $g_{\mu\nu}$ has 10 independent parameters aswell. In other words, it is possibly to choose a locally inertial coordinate system that can mimic and hence remove gravity.

Since the metric must yield the same line element in every coordinate system, the condition imposed is:

$$g_{\mu\nu} dx^\mu dx^\nu \equiv g'_{\mu\nu} dx'^\mu dx'^\nu \tag{3.61}$$

This is the definition of the metric and thus the coordinate transformation for the metric is:

$$\begin{aligned} g'_{\mu\nu}(x') &= \frac{\partial x^\alpha}{\partial x'^\mu} \frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta}(x) \\ &= \eta_{\mu\nu} \qquad \text{At a particular point} \end{aligned} \tag{3.62}$$

Claim 8. In a "sufficiently" small neighborhood of the point:

$$g_{\mu\nu} = \eta_{\mu\nu} \quad \text{and} \quad \frac{\partial}{\partial x^\alpha} g_{\mu\nu} \approx 0 \tag{3.63}$$

when these conditions are applied, the laws of physics take the same special relativistic form as they are in flat space-time.

Proof 8. Start with Eq 3.62 and assume that the coordinate transformation has Taylor series about an arbitrary point a^α, now one can expand x^μ about that point:

$$x^\mu(x'^\alpha) = a^\mu + \frac{\partial x^\mu}{\partial x'^\nu}_a (x'^\nu - a'^\nu) + \frac{\partial^2 x^\mu}{\partial x'^\alpha \partial x'^\beta} (x' - a')^\alpha (x' - a')^\beta + ... \tag{3.64}$$

Substitute Eq 3.64 into Eq 3.62:

$$g'_{\mu\nu}(x') = \underbrace{\left(\frac{\partial x}{\partial x'}\right)^2 |g|_a}_{\eta} + \overbrace{\left(\frac{\partial x}{\partial x'} \left|\frac{\partial^2 x}{\partial x \partial x'}\right|_a g + \frac{\partial x}{\partial x'} \left|\frac{\partial x}{\partial x'}\right|_a\right) \frac{\partial g}{\partial x} (x' - a')^\alpha}^{\text{Term 2}} + ... \tag{3.65}$$

The indices have been deliberately left out here, as the large number of indices makes the expression very confusing, whereas this equation gives a schematic structure to the solution. The $\frac{\partial x}{\partial x'}$ term in the first term on the L.H.S can be used to rotate and rescale the eigenvalues of g such that it becomes η.

The $\left|\frac{\partial^2 x^\alpha}{\partial x'^\beta \partial x'^\gamma}\right|_a$ part from Term 2, in Eq 3.65, is an arbitrary matrix. The matrix is symmetric in γ and β, therefore it has 10 components from these two indicies. α is independent of γ and β so it gives rise to 4 more independent parameters meaning that in total, therefore are $4 \times 10 = 40$ free parameters. This means that there are 40 free parameters that can be arbitrarily set. The coefficients of term two are the derivatives of the metric evaluated at the given point:

$$\left|\partial g'_{\mu\nu} \partial x'^\beta\right|_a \tag{3.66}$$

This term is also symmetric in μ and ν and has one independent index β, therefore in total this also has 40 free parameters. Using the 40 parameters from the first part of term 2 and the 40 parameters from this part, one can solve a set of 40 linear equations to set Term 2 to be zero. This the there is no change in curvature in to first order and the metric in the first term can be set to the flat Minkowski metric.

At first sight this seems strange. But this is where Riemann conducted his pioneering work. He showed that at the second order things become more interesting, as one will have more free parameters then equations to constrain them, so infact it is not possible to choose a coordinate transformation from a curved metric to a flat metric at a given point. The second order will have terms like:

$$g'_{\mu\nu}(x') = \text{R.H.S of Eq 3.65} + \underbrace{\frac{\partial x^3}{\partial x' \partial x' \partial x'}}_{\text{Term a}} + \text{Terms } \alpha \tag{3.67}$$

But Terms α have already been determined as they will have terms from the R.H.S of Eq 3.65, so the only undetermined Term is Term a. These are numbers that are coefficients of the $(x' - a')^2$. Now the question is, is it possible to set the parameters of :

$$\left|\frac{\partial x^3}{\partial x' \partial x' \partial x'}\right|_a \tag{3.68}$$

such that the second derivative of the curvature is 0:

$$\frac{\partial g_{\mu\nu}}{\partial x^\alpha \partial x^\beta} = 0 \tag{3.69}$$

Eq 3.68 is symmetric in α, β, γ as partial derivatives commute, μ is a free index. A symmetric three index tensor with indicies running over 4 values. Let's count the possible number of parameters:

$$\begin{aligned} \text{3 indicies are the same:} \quad &\Rightarrow \quad \alpha = \beta = \gamma = 0, 1, 2, 3 \qquad \text{4 values} \\ \text{2 indicies are the same:} \quad &\Rightarrow \quad \alpha = \beta = 0, \gamma = 1 \text{ etc} \qquad \text{12 values} \\ \text{no indicies are the same:} \quad &\Rightarrow \quad \text{4 values} \end{aligned} \tag{3.70}$$

Therefore the number of parameters here is 20. There is also the free μ index in the top which is independent from these indicies and takes on 4 values itself, so in total there are 80 free parameters from Eq 3.68.

On the other hand Eq 3.69, is symmetric in μ, ν and α, β, hence they each contribute 10 parameters each, giving a total of 100 free parameters. Thus there are only 80 equations that can be solved and set to zero. This means there are 20 non zero second derivatives of the curvature and it is precisely these 20 parameters that describe the curvature of space-time, that cannot be simply removed by choosing a coordinate system. These 20 free parameters form the Riemann tensor. It took Einstein over 10 years to understand this mathematics of Riemann and incorporate it into his theory and from the above calculations, you can see why!.

The metric can not only measure distances, it can also measure volumes. This is important, as the action involves an integration over the volume of space time. It has already been shown that the metric transforms as:

$$g'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu} g_{\alpha\beta} \tag{3.71}$$

The determinant of the metric is:

$$|g'_{\mu\nu}| \equiv g' = \left|\frac{\partial x}{\partial x'}\right|^2 g(\equiv |g_\alpha \beta|) \tag{3.72}$$

Therefore the determinant of the metric transforms with 2 powers of the Jacobin matrix. Recall that the metric has signature of (-,+,+,+), thus its determinant will necessarily be *negative* and to compensate for this negative sign in the square root, one has to include a minus sign. The determinany therefore transforms as:

$$\sqrt{-g} = \left|\frac{\partial x}{\partial x'}\right| \sqrt{-g} \tag{3.73}$$

and in general, integrating over d^4x', will transform as:

$$d^4x' = \left|\frac{\partial x'}{\partial x}\right| d^4x \tag{3.74}$$

Therefore:

$$d^4x'\sqrt{-g'} \equiv d^4x\sqrt{-g} \tag{3.75}$$

This is known as the invariant volume component and is completely independent of the coordinate system. One more important function will be the inverse metric, $g^{\mu\nu}$ and is defined as:

$$g^{\mu\nu} g_{\nu\alpha} = \delta^\mu_\alpha \tag{3.76}$$

This transforms in the opposite way to $g_{\mu\nu}$:

$$g'^{\mu\nu}(x') = \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x'^\nu}{\partial x^\beta} g^{\alpha\beta} \tag{3.77}$$

This is used to raise and lower indicies in the same way that η is used in flat space.

EXAMPLE 7. Consider a 2 dimensional sphere in a 3 dimensional Euclidean space, $S^2 \in \mathbb{R}^3$. The sphere satisfies:

$$x^2 + y^2 + z^2 = 1 \tag{3.78}$$

The simplest metric is the metric inherited by the Euclidean space:

$$dx^2 + dy^2 + x^2 \tag{3.79}$$

Define a metric parametrised by spherical coordinates over the sphere with unit radius, this gives the transformation from Cartesian to Spherical coordinates:

$$\begin{aligned} x &= \sin\theta\cos\phi \\ y &= \sin\theta\sin\phi \\ z &= \cos\theta \end{aligned} \tag{3.80}$$

Thus, the metric embedded into the spherical coordinates is:

$$d\phi^2 + \sin^2\theta d\phi^2 \tag{3.81}$$

when $\theta \approx \frac{\pi}{2}$, the metric reduces to:

$$d\theta^2 + d\phi^2 \tag{3.82}$$

so if the spherical coordinates are constrained to a point, the metric looks flat. In fact it does'nt matter what the point is, as even $\sin^2\theta \neq 1$ will just be a number < 1, which is still another flat matter scaled slightly differently. On the other hand if $\sin\theta = 0$ i.e at the poles, the metric is not flat, as the the phi coordinate becomes singular (i.e all ϕ coordinates will specify the same point!).

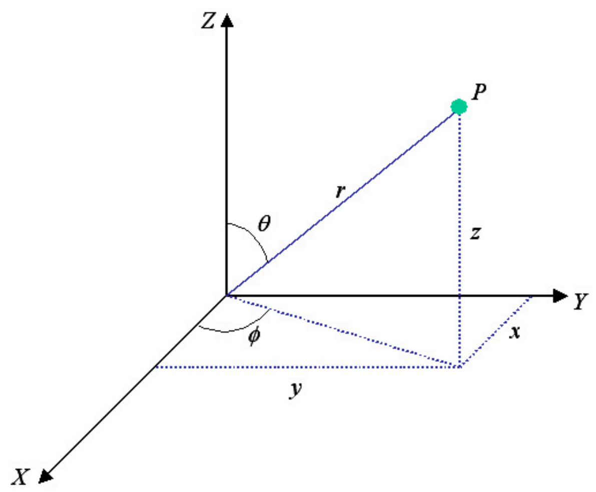


FIGURE 12. Spherical coordinates

The metric tensor in these coordinates is given by:

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix} \tag{3.83}$$

volume element in these coordinates is given by:

$$\text{Volume} = \sqrt{g}\, d^2x = \sin\theta d\theta d\phi \tag{3.84}$$

Note that there is no minus sign in the square root here. This is because the metric being used here is Riemannian, i.e it does not have a time component that changes the sign. As mentioned before. the coordinates defined for the sphere do not appear flat at the poles of the sphere, as the ϕ coordinates become singular at two points. One can do better by parametrising the sphere by stereographic coordinates.

The way to define stereographic coordinates is:

- Take a plane tangent to the north pole, P.
- Take every point on the sphere and connect it to the south pole in a straight line with the line intersecting the plane P.
- This project every point on the sphere to a unique point on this plane.

It is obvious that the singular point will be the south pole as it would have a line that is tangent to itself and hence parallel to P. In flat Euclidean space, parallel line will never intersect. It can be shown[1] that the metric in the stereographic coordinates is:

$$\frac{4(dR^2 + R^2 d\theta)}{(1+R^2)^2} \tag{3.85}$$

where the geometry of in Figure 13 defines:

$$\begin{aligned} \frac{z_1}{2} &= \frac{x_1}{1+x_3} \\ \frac{z_2}{2} &= \frac{x_2}{1+x_3} \\ z_1 &= R\cos\theta \\ z_2 &= R\sin\theta \end{aligned} \tag{3.86}$$

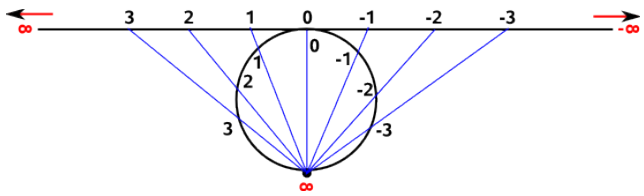


FIGURE 13. Stereographic coordinates

3.2. de-Sitter space-time. John Nash proved the following theorem:

THEOREM 2. Every Riemannian manifold can be isometrically$^{2 embeddedinaEuclideanspace,ofsomedimension.}$

It turns out that any 4D manifold can be locally embedded in 10 Euclidean dimensions. This theorem can be applied to the simplest 4D (1+3) space-time known as the de-Sitter space-time. It seems that our universe is evolving towards this metric. The de-Sitter space is basically a 4-D sphere, i.e a space with maximum symmetry.

Imagine a 5D Minkowski space, with the metric:

$$\eta_{AB} = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.87}$$

where $A, B \in \{0,1,2,3,4\}$. The line element is simply the Pythagorean theorem in 5D:

$$\eta_{AB}dx^A dx^B = d(x^0)^2 + d(x^1)^2 + d(x^2)^2 + d(x^3)^2 + d(x^4)^2 \tag{3.88}$$

Consider a 4D hyperboloid embedded in this space (Recall that a hyperboloid is just like a sphere in that it is described by a quadratic equation, the only difference being a minus sign in the quadratic as supposed to all positive signs):

$$\text{4D hyperboloid} \Rightarrow L^2 = -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2 \tag{3.89}$$

This geometry represents a bouncing universe, as x^0 time $= -\infty$, the universe is a big sphere. At $x = 0$, the universe is small sphere, and at x^0 the universe blows up in size again. Eq 3.89 can be solved by setting:

[1]The calculation is rather lengthy and would diverge too far away from the main points of this example. The calculation is very common and can be found at many places on the web.

$$\begin{aligned} x^0 &= L \sinh t \\ x^I &= L \cosh t \Omega^I \qquad I \in \{1,2,3,4\} \end{aligned} \tag{3.90}$$

Such that:

$$\Omega^I \Omega_I = 1 \tag{3.91}$$

The Ω parametrizes a 3-sphere in this space:

$$\Omega^I \equiv (\sin\chi \sin\theta \cos\phi, \sin\chi \sin\theta \sin\phi, \sin\chi \cos\theta \cos\chi) \tag{3.92}$$

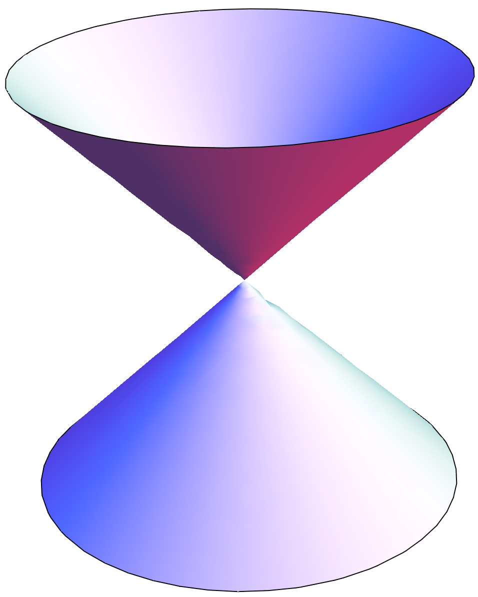


FIGURE 14. de-Sitter space

The line element in this metric becomes:

$$L^2 = L^2 \left(-dt^2 + \cosh^2 t d\Omega_3^2\right) \tag{3.93}$$

where:

$$d\Omega_3^2 = d\chi^2 + \sin^2\chi (d\theta^2 + \sin^2\theta d\phi^2) \tag{3.94}$$

At a given time, the size of the universe is given by the $\cosh^2 t$ term. At large times it becomes exponentially large and yet the dark energy remains constant. This is a bizarre property of dark energy and makes the universe we appear to live in very strange indeed.

So far the de-Sitter space has been parametrised by a closed universe. But it can be represented in many other ways. Since the space is embedded in a Minkowski space; there are three natural slicing's. The closed slicing is where the normal to the plane of the surface is space-like. The other two are obviously time-like or light-like. The flat spacing is parametrised by:

$$\begin{aligned} x^0 &= L(\sinh t + \frac{1}{2}r^2 e^t) \\ x^i &= Le^t x^i, \qquad i \in \{1,2,3\} \\ x^4 &= L(\cosh t - \frac{1}{2}r^2 e^t) \end{aligned} \tag{3.95}$$

Then:

$$L(-dt^2 + e^{2t}dx^i dx^i) = d(x^i)^2 + d(x^4)^2) - d(x^0)^2 \tag{3.96}$$

This expands exponentially at large t. Note that:

$$-\infty < t < \infty \tag{3.97}$$

and at $-\infty$, $e^{2t}dx^{i2}$ terms are zero, hence the metric is singular at $t = -\infty$ (as the determinant is 0, hence not invertible). At $t = -\infty, \rightarrow x^0 + x^4 = 0$. Thus flat slicing becomes singular on the surface represented by the line $x^0 + x^4 = 0$. It only covers half of the space time as there is an equal amount above the plane and below it.

This is precisely the same as the idea of *inflation*. Inflation requires the very early universe looked flat and hence had the same form as the metric. The length scale during inflation would have been very small (i.e the damping term is exponentially small), therefore after a small amount of time, the universe would expand exponentially. Going back to the metric:

$$L^2(-dt^2 + \cosh^2 td\,\Omega_3^2) \tag{3.98}$$

Define a proper time:

$$t_p \equiv Lt \tag{3.99}$$

Substitute Eq 3.99 into Eq 3.98:

$$\cosh^2\left(\frac{t_p}{L}\right)d\Omega_3^2 - dt_p^2 \tag{3.100}$$

It has already been shown that at a given point, one can choose coordinates such that locally the metric looks flat. The $\cosh^2\left(\frac{t_p}{L}\right)d\Omega_3^2$ is some scale, i.e the radius of the 3-sphere, so it looks like flat Minkowski (as it has a time component) space-time.

A curious fact about de-Sitter space is that it has a *temperature*. Suppose the proper time is related to some imaginary time, τ:

$$t_p \equiv i(\tau - \frac{\pi}{2}) \tag{3.101}$$

Substitute Eq 3.101 into Eq 3.100:

$$d\tau^2 + L^2 \sin^{\left(\frac{\tau}{L}\right)} d\Omega_3^2 \tag{3.102}$$

$\sin^2$ is obviously periodic, hence the metric is periodic in imaginary time. In fact the metric is that of a 4-sphere. The period is just $2\pi L$. From the Boltzmann distribution:

$$e^{-\frac{H}{T}} = e^{-iHt} \tag{3.103}$$

where:

$$t \equiv -\frac{i}{T(temp)} \qquad H = \text{Hamiltonian} \tag{3.104}$$

Then in statistical mechanics, one sums over all states and for a diagonalised metric, it is just the trace:

$$Tr(e^{\frac{-H}{T}}) \tag{3.105}$$

From the e^{-iHT} equivalence, the function $Tr(e^{-\frac{H}{T}})$, must be periodic (as $e^{i\theta}$ is periodic) with period $\Delta\tau$:

$$\Delta\tau = \frac{1}{T} \tag{3.106}$$

By equating the space-time periodicity in imaginary time and the period in a thermal system; it is easy to see that:

$$\frac{1}{T} = 2\pi L \Rightarrow T = \frac{1}{2\pi L} \equiv \frac{H}{2\pi} \tag{3.107}$$

The de-Sitter space time is a very useful example of embedding a metric into a higher dimension space-time (i.e Nash's theorem).

3.3. Curvature in the metric. Recall that previously, the curvature of arbitrary curved spaces has been discussed in terms of the connection and the co-variant derivative as shown in Eq 3.192, these are discussed without the metric. The Γ's can be fixed in general relativity in terms of the metric by making two simplifying assumptions:

- Assume there is no torsion:

$$T^{\nu}_{\mu\alpha} \equiv 0 \Rightarrow \Gamma^{\nu}_{\mu\alpha} \equiv \Gamma^{\nu}_{\alpha\mu} \tag{3.108}$$

- The metric tensor is a constant with respect to the covariant derivative:

$$\nabla_{\alpha} g_{\mu\nu} \equiv 0 \tag{3.109}$$

This quantity is called the *metricity*.

Claim 9. If:

$$\nabla_{\alpha} g_{\mu\nu} = 0 \tag{3.110}$$

And the torsion is zero, then the connection is given by:

$$\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2} g^{\mu\nu}(g_{\alpha\lambda,\beta} + g_{\beta\lambda,\alpha} - g_{\alpha\beta,\lambda}) \tag{3.111}$$

In terms of the metric, these are called the *Christoffel* symbols.

Proof 9. Start with the the definition of the metricity being zero:

$$\begin{aligned} \Delta_{\alpha} g_{\mu\nu} &= 0 \\ &= \partial_{\alpha} g_{\beta\alpha} - \Gamma^{\delta}_{\beta\alpha} g_{\delta\gamma} - \Gamma^{\delta}_{\gamma\alpha} g_{\beta\delta} \end{aligned} \tag{3.112}$$

One can include the cycles of the indicies:

$$\partial_{\alpha} g_{\alpha\beta} - \Gamma^{\delta}_{\alpha\gamma} g_{\alpha\beta} - \Gamma^{\delta}_{\beta\gamma} g_{\alpha\delta} = 0 \tag{3.113}$$

$$\partial_{\beta} g_{\gamma\alpha} - \Gamma^{\delta}_{\gamma\beta} g_{\delta\alpha} - \Gamma^{\delta}_{\alpha\beta} g_{\gamma\delta} = 0 \tag{3.114}$$

$$\partial_{\alpha} g_{\beta\gamma} - \Gamma^{\delta}_{\beta\alpha} g_{\delta\gamma} - \Gamma^{\delta}_{\gamma\alpha} g_{\beta\delta} = 0 \tag{3.115}$$

Adding the equations above gives:

$$\partial_{\alpha} g_{\beta\gamma} + \partial_{\beta} g_{\gamma\alpha} - \partial_{\gamma} g_{\alpha\beta} - 2\Gamma^{\delta}_{\gamma\alpha} g_{\beta\delta} = 0 \tag{3.116}$$

Multiply through by $g^{\gamma\nu}$:

$$\partial_{\alpha} g_{\beta\gamma} g^{\beta\nu} + \partial_{\beta} g_{\gamma\alpha} g^{\beta\nu} - \partial_{\gamma} g_{\alpha\beta} g^{\beta\nu} - 2\Gamma^{\delta}_{\beta\alpha} g^{\gamma}_{\delta} g^{\nu}_{\gamma} = 0 \tag{3.117}$$

Which can just be re-arranged to give the required form:

$$\Gamma^{\nu}_{\gamma\alpha} = \frac{1}{2} g^{\nu\gamma} \left(\partial_\alpha g_{\beta\gamma} + \partial_\beta g_{\gamma\alpha} - \partial_\gamma g_{\alpha\beta} \right) \tag{3.118}$$

This can also be proved in the other order, i.e if Eq 3.111 is true, then Eq 3.110 must also be true:

$$\begin{aligned} g_{\mu\nu;\gamma} &= \frac{\partial g_{\mu\nu}}{\partial x^\gamma} - \Gamma^{\alpha}_{\mu\gamma} g_{\nu\alpha} - \Gamma^{\alpha}_{\nu\gamma} g_{\mu\alpha} \\ &= g_{\mu\nu,\gamma} - \Gamma^{\alpha}_{\mu\gamma} g_{\alpha_\nu} - \Gamma^{\alpha}_{\nu\gamma} g_{\mu\alpha} \end{aligned} \tag{3.119}$$

The Christoffel symbols are:

$$\Gamma^{\alpha}_{\mu\gamma} = \frac{1}{2} g^{\alpha\sigma} \left(\frac{\partial g_{\mu\sigma}}{\partial x^\gamma} + \frac{\partial g_{\gamma\sigma}}{\partial x^\mu} - \frac{\partial g_{\mu\gamma}}{\partial x^\sigma} \right) \tag{3.120}$$

$$\Gamma^{\alpha}_{\nu\gamma} = \frac{1}{2} g^{\alpha\sigma} \left(\frac{\partial g_{\nu\sigma}}{\partial x^\gamma} + \frac{\partial g_{\gamma\sigma}}{\partial x^\nu} - \frac{\partial g_{\nu\gamma}}{\partial x^\sigma} \right) \tag{3.121}$$

Substituting Eq 3.121 and 3.120 into Eq 3.119:

$$\begin{aligned} g_{\mu\nu;\gamma} &= g_{\mu\nu,\gamma} - \frac{1}{2} \left(\underbrace{g_{\alpha\nu} g^{\alpha\sigma}}_{Contract} \left(\frac{\partial g_{\mu\sigma}}{\partial x^\gamma} + \frac{\partial g_{\gamma\sigma}}{\partial x^\mu} - \frac{\partial g_{\mu\gamma}}{\partial x^\sigma} \right) + \underbrace{g_{\alpha\mu} g^{\alpha\sigma}}_{Contract} \left(\frac{\partial g_{\nu\sigma}}{\partial x^\gamma} + \frac{\partial g_{\gamma\sigma}}{\partial x^\nu} - \frac{\partial g_{\nu\gamma}}{\partial x^\sigma} \right) \right) \\ &= g_{\mu\nu,\gamma} - \frac{1}{2} \left(g^{\sigma}_{\nu} \left(\frac{\partial g_{\mu\sigma}}{\partial x^\gamma} + \frac{\partial g_{\gamma\sigma}}{\partial x^\mu} - \frac{\partial g_{\mu\gamma}}{\partial x^\sigma} \right) + g^{\sigma}_{\mu} \left(\frac{\partial g_{\nu\sigma}}{\partial x^\gamma} + \frac{\partial g_{\gamma\sigma}}{\partial x^\nu} - \frac{\partial g_{\nu\gamma}}{\partial x^\sigma} \right) \right) \end{aligned} \tag{3.122}$$

Now the $g's$ can be contracted with the $g's$ inside the derivative to obtain:

$$g_{\mu\nu;\gamma} = g_{\mu\nu,\gamma} - \frac{1}{2} (\frac{\partial g_{\mu\nu}}{\partial x^\gamma} + \frac{\partial g_{\gamma\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\gamma}}{\partial x^\nu} + \frac{\partial g_{\nu\mu}}{\partial x^\gamma} + \frac{\partial g_{\gamma\mu}}{\partial x^\nu} - \frac{\partial g_{\nu\gamma}}{\partial x^\mu}) \tag{3.123}$$

Now we have to remember that the metric tensors are symmetric, i.e:

$$g_{\mu\nu} = g_{\nu\mu} \tag{3.124}$$

Therefore we are left with:

$$\begin{aligned} g_{\mu\nu;\gamma} &= g_{\mu\nu,\gamma} - \frac{1}{2} \left(2 \frac{\partial g_{\mu\nu}}{\partial x^\gamma} \right) \\ &= g_{\mu\nu,\gamma} - \frac{\partial g_{\mu} \nu}{\partial x^\gamma} \\ &= g_{\mu\nu,\gamma} - g_{\mu\nu,\gamma} \\ &= 0 \end{aligned} \tag{3.125}$$

Because under these assumptions, the connection only depends on the first derivative of the metric, and can be set to zero at a particular point, through a coordinate transformation. Thus covariant derivatives just become ordinary derivatives at a given point.

If metricity is not zero, then a vector V_μ:

$$\nabla_\alpha V^\mu = 0 \tag{3.126}$$

i.e it is a constant vector on a curved space. The length square of this vector is:

$$g_{\mu\nu} V^{\mu\nu} \tag{3.127}$$

This means:

$$\begin{aligned}\nabla_\alpha \phi &= \frac{\partial}{\partial x^\alpha}(g_{\mu\nu}V^\mu V^\nu) = \nabla_\alpha(g_{\mu\nu}V^\mu V^\alpha) \\ &= \underbrace{(\nabla_\alpha g_{\mu\nu})V^\mu V^\nu}_{\text{Term 1}} + \underbrace{g_{\mu\nu}\nabla_\alpha V^\mu V^\nu}_{\text{Term a}} + \underbrace{g_{\mu\nu}V^\mu \nabla^\mu \nabla_\alpha V^\nu}_{\text{Term b}}\end{aligned} \tag{3.128}$$

Where ϕ is the length, thus a scalar quantity. By assumption Terms a and b are both zero. The fist term involves the metricity and if this is not zero, i.e if one of the assumptions of GR are not true, then the length squared of a constant vector is not constant under the covariant derivative. This is very strange and it actually changes what the constant vector actually is.

Now the metric is embedded into the manifold with a formal description. In general, the coordinate system that corresponds to the x mapping and the x' mapping may have different regions in which they are valid, except for an overlapping regions in which they must necessarily agree. The way this is done, is by writing equations that are invariant under coordinate transformations, i.e in terms of tensors.

3.4. Parallel transport. The most fundamental object on a manifold is a curve. The curve in general will be parametrised by some parameter, λ, that runs along the curve in a given coordinate system. The coordinates can be chosen with an arbitrary transformation to different coordinate systems:

$$x'^\mu(\lambda) = x'^\mu(x^\mu(\lambda)) \tag{3.129}$$

The transformation from one coordinate system to another, is always a function of the old coordinates, related in general by a transformation matrix. The actual transformation is not a vector, its just some function (like $\sin\theta$ etc) But the tangent vector $\frac{\partial x'^\mu}{\partial \lambda}$, is a vector:

$$\frac{\partial x'^\mu}{\partial \lambda} = \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial \lambda} \tag{3.130}$$

If we define:

$$\xi^\mu \equiv \frac{\partial x^\mu}{\partial \lambda} \tag{3.131}$$

then:

$$\xi'^\mu = \frac{\partial x'^\mu}{\partial x^\alpha}\xi^\alpha \tag{3.132}$$

One can now insist that this equation for the vector is invariant under coordinate transformations. Consider a vector $V^\mu(\lambda)$, defined at every point on the curve. A vector is said to be *parallel transported* along a curve, if:

$$\frac{DV^\mu}{D\lambda} \equiv \frac{\partial V^\mu}{\partial \lambda} + \Gamma^\mu_{\alpha\beta}V^\alpha(\lambda)\frac{\partial x^\beta}{\partial \lambda} \equiv 0 \tag{3.133}$$

where D represents the *absolute derivative* and is defined by the equation above. $\Gamma^\mu_{\alpha\beta}$ is calculated at the point of the curve:

$$T^\mu_{\alpha\beta} = \Gamma^\mu_{\alpha\beta}(x^\mu(\lambda)) \tag{3.134}$$

which is equivalent to:

$$\frac{\partial V^\mu}{\partial \lambda} = -\Gamma^\mu_{\alpha\beta}V^\alpha(\lambda)\frac{\partial x^\beta}{\partial \lambda} \tag{3.135}$$

This is the closest one gets to saying that a vector is constant on a curve.

EXAMPLE 8. Suppose a contravariant vector:

$$V^i = (1, 2, 3) \tag{3.136}$$

in the above co-ordinate system is parallel transported from point (r, θ, ϕ) to the point $(r + \alpha, \theta + \beta, \phi + \gamma)$, where $\alpha << 1, \beta << 1, \gamma << 1$, under the metric:

$$ds^2 = dr^2 + r^2 d\theta^2 + \sin^2 r d\phi^2 \tag{3.137}$$

The change in V^i under parallel transport is given by Eq 3.135. The derivative terms in Eq 3.135 represents how much a coordinate changes with respect to the parameter that parametrizes the curve, which in this case is simply:

$$\frac{\partial x^1}{\partial \lambda} = \alpha \qquad \frac{\partial x^2}{\partial \lambda} = \beta \qquad \frac{\partial x^3}{\partial \lambda} = \gamma \tag{3.138}$$

Therefore the change in V^i is:

$$\begin{aligned} \frac{\partial V^1}{\partial \lambda} &= -\Gamma^1_{\alpha\beta} V^\alpha \frac{\partial x^\beta}{\partial \lambda} \\ &= -\Gamma^1_{22} V^2 \frac{\partial x^2}{\partial \lambda} - \Gamma^1_{33} V^3 \frac{\partial x^3}{\partial \lambda} \\ &= \frac{2\beta}{r} + \frac{\sin 2r}{2} 3\gamma \end{aligned} \tag{3.139}$$

Similarly:

$$\begin{aligned} \frac{\partial V^2}{\partial \lambda} &= -\Gamma^2_{\alpha\beta} V^\alpha \frac{\partial x^\beta}{\partial \lambda} \\ &= -\Gamma^2_{12} V^2 \frac{\partial x^1}{\partial \lambda} \\ &= -\frac{2\alpha}{r} - \frac{\beta}{r} \end{aligned} \tag{3.140}$$

$$\begin{aligned} \frac{\partial V^3}{\partial \lambda} &= -\Gamma^3_{\alpha\beta} V^\alpha \frac{\partial x^\beta}{\partial \lambda} \\ &= -\Gamma^3_{13} V^3 \frac{\partial x^1}{\partial \lambda} - \Gamma^3_{31} V^1 \frac{\partial x^3}{\partial \lambda} \\ &= -3\alpha \cot r - \gamma \cot r \end{aligned} \tag{3.141}$$

The simplest curves are straight line in flat space, which correspond to *geodesics* in curved space-time. The definition of a geodesic is that the absolute derivative of the tangent vector is proportional to the tangent vector:

$$\frac{D\xi^\mu}{D\lambda} = f(\lambda)\xi^\mu \tag{3.142}$$

for some function $f(\lambda)$ which is introduced to remove the proportionality sign and ξ is the tangent vector defined by:

$$\xi^\mu \equiv \frac{\partial x^\mu}{\partial \lambda} \tag{3.143}$$

If a curve is a geodesic, the curve can be can be re-parametrised in a way that $f(\lambda) = 0$. So change λ to $\sigma(\lambda)$ to re-parametrize the geodesic. The $\sigma(\lambda)$ is monotonic in λ, i.e for each value of λ, there is a value for $\sigma(\lambda)$. Let's rewrite Eq 3.142 in terms of σ:

$$x^\mu(\lambda) \to x^\mu(\lambda(\sigma)) \equiv x^\mu(\sigma) \tag{3.144}$$

Using the definition of the absolute derivative in Eq 3.133, Eq 3.142 can be written under this parametrization as:

$$\frac{\partial}{\partial\lambda}\left(\frac{\partial\sigma}{\partial\lambda}\frac{\partial x^\mu}{\partial x^\mu}\partial\sigma\right)+\Gamma^\mu_{\beta\alpha}\frac{\partial^\beta}{\partial\sigma}\frac{\partial x^\alpha}{\partial\sigma}\left(\frac{\partial\sigma}{\partial\lambda}\right)^2=f(\lambda)\frac{\partial x^\mu}{\partial\sigma}\frac{\partial\sigma}{\partial\lambda} \tag{3.145}$$

Using the chain rule

$$\underbrace{\frac{\partial^2\sigma}{\partial\lambda^2}\frac{\partial x^\mu}{\partial\sigma}}_{\text{Term a}}+\underbrace{\left(\frac{\partial\sigma}{\partial\lambda}\right)^2\left(\frac{\partial^2 x^\mu}{\partial\sigma^2}+\Gamma^\mu_{\beta\alpha}\frac{\partial x^\beta}{\partial\sigma}\frac{\partial x^\alpha}{\partial\sigma}\right)}_{\text{Term }\alpha}=\underbrace{f(\lambda)\frac{\partial\sigma}{\partial\lambda}\frac{\partial x^\mu}{\partial\sigma}}_{\text{Term b}} \tag{3.146}$$

By equating Term a and Term b:

$$\frac{\partial^2\sigma}{\partial\lambda^2}\equiv f(\lambda)\frac{\partial\sigma}{\partial\lambda} \tag{3.147}$$

which is s simple second-order partial differential equation, which has the solution:

$$\frac{\partial\sigma}{\partial\lambda}=Ae^{\int f(\lambda)d\lambda}\qquad\forall A\in\mathbb{C} \tag{3.148}$$

This leaves Term α in Eq 3.146. Since $\frac{\partial\sigma}{\partial\lambda}$ is non-zero (as by definition $\sigma\equiv\sigma(\lambda)$), the only way Eq 3.146 can be true is if:

$$\frac{\partial^2 x^\mu}{\partial\sigma^2}+\Gamma^\mu_{\beta\alpha}\frac{\partial x^\beta}{\partial\sigma}\frac{\partial x^\alpha}{\partial\sigma}=0 \tag{3.149}$$

which is simply the geodesic equation. σ is called the *affine* parameter, affine means the σ is determined up to transformations of the form:

$$\sigma'=A\sigma+B\qquad\forall A,B\in\mathbb{C} \tag{3.150}$$

In our universe, all particles follow geodesics in curved space-time.

4. Principle of least action

The principle of least action is the corner-stone of all physics. In this section, the geodesic equations will be derived using the action principle and show that it yields the same result as before.

4.1. Introducing the action. The action in general relativity takes the form:

$$S=-mc\int d\tau \tag{3.151}$$

where τ is the proper time that parametrizes the world line is space-time. The particles follow a curve is space-time $x^\mu(\tau)$, it turns out that τ is an affine parameter and the equation of motion for the particle is the geodesic equation with this parameter. Re-write:

$$\begin{aligned}S&=-mc\int\sqrt{-\eta_{\mu\nu}dx^\mu dx^\nu}\\&=-mc\int\sqrt{c\partial t^2-\partial\vec{x}^2}\\&=-mc\int c\partial t\sqrt{1-\left(\frac{\partial\vec{x}}{\partial t}\right)^2\frac{1}{c^2}}\end{aligned} \tag{3.152}$$

The equivalence principle states that the laws of physics must be exactly the same in curved space-time, as the laws of physics in flat space-time within a small neighborhood, so $\eta_{\mu\nu}\to g_{\mu\nu}$:

$$S=-mc\int\sqrt{-\eta_{\mu\nu}dx^\mu dx^\nu} \tag{3.153}$$

This can be parametrised using a parameter, say λ:

$$S = -mc \int \sqrt{-g_{\mu\nu} \frac{\partial x^\mu}{\partial \lambda} \frac{\partial x^\nu}{\partial \lambda}} d\lambda \tag{3.154}$$

4.2. Aside: Why does E=mc²?. If we take a general action:

$$S = \int_{x_i(t_i)}^{x_f(t_f)} L dt \tag{3.155}$$

for some Lagrangian:

$$L = L(\vec{x}, \dot{\vec{x}}) \tag{3.156}$$

Minimising the action yields the Euler-Lagrange equations, which must be satisfied and are the equations of motion of the particles (i.e the solution to these equations gives the trajectory of the particle).

Evaluated on the classical trajectory, this action is a function of the trajectory at $x(t_i)$ and $x(t_f)$ (or a function to be more precise). This is called Hamiltons principle function. Consider varying the boundary conditions:

$$\delta V \equiv \{x_i, t_i, x_f, t_f\} \tag{3.157}$$

This will cause variations in the classical trajectories $\vec{x}_c(t)$. Now we want to know, what is the variation in the classical action, under δV:

$$\delta S_{classical} = \int_{x_i(t_i+\delta t_i)+\delta x_i}^{x_f(t_f+\delta t_f)+\delta x_f} L dt \tag{3.158}$$

Using Leibniz's integral rule:

$$\delta S_{classical} = \delta t_f L(t_f) - \delta t_i L(t_i) + \int_{t_i}^{t_f} \frac{\partial L}{\partial \vec{x}} \delta \vec{x} dt + \int_{t_i}^{t_f} \frac{\partial L}{\partial \dot{\vec{x}}} \delta \dot{\vec{x}} dt \tag{3.159}$$

Using the Euler-Lagrange equation:

$$\frac{\partial L}{\partial x} = \frac{\partial}{\partial t} \left(\frac{\partial L}{\partial \dot{x}} \right) \tag{3.160}$$

Eq 3.159 can be re-written as:

$$\delta S_{classical} = \left[\delta t L + \frac{\partial L}{\partial \dot{\vec{x}}} \delta \vec{x} \right]_{t_f}^{t_i} \tag{3.161}$$

Now one has to realise a subtlety:

$$\delta \vec{x_f} = \delta t_f + \dot{\vec{x}}_f \delta t_f \tag{3.162}$$

Hence Eq 3.161 can be re-written as:

$$\delta S = \left[\delta L + (\delta \vec{x} - \dot{\vec{x}} \delta t) \frac{\partial L}{\partial \dot{x}} \right]_{t_i}^{t_f} \tag{3.163}$$

Using the definitions:

$$\begin{aligned} p &\equiv \frac{\partial L}{\partial \dot{\vec{x}}} \quad \text{Canonical momentum} \\ H &\equiv p\dot{\vec{x}} - L \quad \text{Hamiltonian} \end{aligned} \tag{3.164}$$

Thus the Hamiltonian variation function gives the canonical momentum and the Hamiltonian as the coefficients of the variations in the final and initial positions and times. Therefore:

$$\begin{aligned} \frac{\partial S'}{\partial x_f} &= \vec{p}_f \quad \text{(-ve for initial)} \\ \frac{\partial S'}{\partial t_f} &= -H_f \quad \text{(+ve for initial)} \end{aligned} \tag{3.165}$$

Noether's theorem: Imagine that the action has a symmetry, for example translation:

$$\vec{x} \to \vec{x} + c(\text{constant}) \tag{3.166}$$

regard this is a variation, therefore:

$$\delta x_f = \delta x_i \equiv c \tag{3.167}$$

We know that:

$$\delta S_{classical} = 0 = [\delta x_i p]_i^f \tag{3.168}$$

since $x_f \neq x_i$, this implies:

$$\vec{p}_f = \vec{p}_i \tag{3.169}$$

Thus *momentum is conserved.* If then, there is a time translation symmetry:

$$t \to t + c \tag{3.170}$$

now regard this as a variation:

$$\delta S_{classical} = [\delta t H]_i^f = 0 \tag{3.171}$$

since $t_f \neq t_i$:

$$H_f \equiv H_i \tag{3.172}$$

therefore *energy is conserved.* Lets go back to the action defined in Eq 3.152; this shows that the Lagrangian is:

$$L = mc^2 \sqrt{1 - \left(\frac{\partial x}{\partial t}\right)^2 \frac{1}{c^2}} \tag{3.173}$$

The momentum is given by:

$$\begin{aligned} p &= \frac{\partial L}{\partial \dot{x}} \\ &= mc^2 \frac{\left(\frac{\partial x}{\partial t}\right)}{\sqrt{1 - \left(\frac{\partial x}{\partial t}\right)^2 \frac{1}{c}}} \\ &= mv\gamma \end{aligned} \tag{3.174}$$

where:

$$\frac{\partial x}{\partial t} \equiv v \tag{3.175}$$

For the Hamiltonian, i.e energy:

$$\begin{aligned} H &= p\dot{x} - L \\ &= mv^2\gamma - mc^2\sqrt{1-\frac{v^2}{c^2}} \\ &= mv^2\gamma - \left(mc^2\left(1-\frac{v^2}{c^2}\right)\gamma\right) \\ &= -mv^2\gamma - mc^2\gamma + \frac{mc^2}{c^2}v^2\gamma \\ &= mc^2\gamma \end{aligned} \tag{3.176}$$

4.3. Geodesic equation from the action. To get the geodesic equations, lets vary the action in Eq 3.154:

$$\begin{aligned} \delta S &= \frac{\partial S}{\partial \lambda}d\lambda \\ &= -\frac{mc}{2}\int \frac{d\lambda}{\sqrt{-g_{\mu\nu}x'^{\mu}x'^{\nu}}}(-g_{\mu\nu,\lambda}\delta x^{\lambda}x'^{\nu}x'^{\mu} + \underbrace{2g_{\mu\nu}\delta x'^{\mu}x'\nu}_{Term1}) \end{aligned} \tag{3.177}$$

Term 1 has a factor of two in front, as the same terms come from x'^{μ} and $x^{'\nu}$, Notice that:

$$\begin{aligned} \dot{x}^{\mu} &= \frac{1}{c}\frac{\partial x^{\mu}}{\partial \tau} \\ &= \frac{\partial x^{\mu}}{\partial \lambda}\frac{1}{\sqrt{-g_{\mu\nu}x'^{\mu}x'^{\nu}}} \end{aligned} \tag{3.178}$$

Thus Eq 3.177 simplifies to:

$$\delta S = -mc^2\int \frac{1}{2}\partial\tau\left(-g_{\mu\nu,\lambda}\delta x^{\lambda}\dot{x}^{\mu}\dot{x}^{\nu} - 2g_{\lambda\nu}\delta\dot{x}^{\lambda}\dot{x}^{\nu}\right) \tag{3.179}$$

The next thing to do, as in any variational problem, is to integrate by parts. Notice that this integration by parts was not done with respect to λ. First the λ is converted into τ; this is motivated by the fact that λ is an arbitrary parametrisation imposed on a curve. Integrating by parts with respect to λ, will give equations of motion with λ, which will then give the trajectories with respect to λ, which is not very useful when λ is not known!. This is why the curve is re-parametrised by:

$$c\partial\tau = \partial\lambda\sqrt{-g_{\mu\nu}x'^{\mu}x'^{\nu}} \tag{3.180}$$

So it is seen from the expression above that the λ terms on the L.H.S are completely redundant, as they just cancel out, so the trajectory obtained from the solutions of the equations of motion will be as functions of τ. Integrating by parts:

$$\delta S \propto \int(-g_{\mu\nu,\lambda}\dot{x}^{\mu}\dot{x}^{\nu} + 2g_{\lambda\nu}\dot{x}^{\nu})\cdot\delta x^{\lambda} = 0 \tag{3.181}$$

This implies that:

$$\frac{\partial}{\partial\tau}\left(g_{\lambda\nu}\dot{x}^{\nu}\right) = \frac{1}{2}g_{\mu\nu,\lambda}\dot{x}^{\mu}\dot{x}^{\nu} \tag{3.182}$$

Which can be further simplified by expanding out the product of derivatives first:

$$\begin{aligned}\frac{\partial g_{\lambda\nu}}{\partial t}\dot{x}^\nu + g_{\lambda\nu}\ddot{x}^\nu &= \frac{1}{2}g_{\mu\nu,\lambda}\dot{x}^\mu\dot{x}^\nu \\ \frac{g_{\lambda\nu}}{\partial x^\mu}\frac{\partial x^\mu}{\partial t}\dot{x}^\nu + g_{\lambda\nu}\ddot{x}^\nu &= \frac{1}{2}g_{\mu\nu,\lambda}\dot{x}^\mu\dot{x}^\nu \\ g_{\lambda\nu,\mu}\dot{x}^\mu\dot{x}^\nu + g_{\mu\nu}\ddot{x}^\nu &= \frac{1}{2}g_{\mu\nu,\lambda}\dot{x}^\mu\dot{x}^\nu \\ \ddot{x}^\mu = -\Gamma^\mu_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda & \end{aligned} \tag{3.183}$$

Which is the usual geodesic equation.

5. Riemann Tensor

Until now, a curve on a manifold is considered and a condition is imposed on it, such that the tangent vectors to the curves are parallel to the curve, which gives rise to the geodesic equation. The next thing to do, is to consider two curves that are close by on a manifold shown in figure 15.

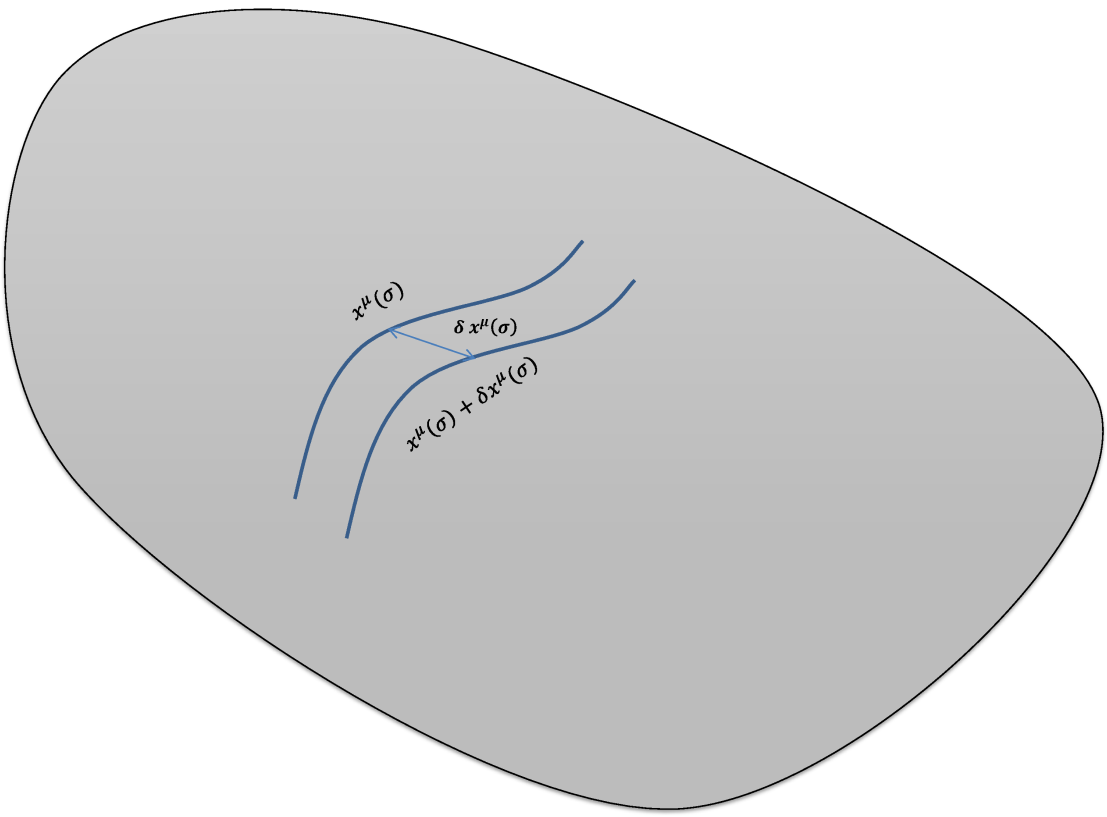


FIGURE 15. Manifold with two geodesics

The two geodesics are displaced by $\delta x^\mu(\sigma)$. The two geodesic equations are defined by the derivatives of the tangent vectors for each curve being zero:

$$\frac{D^2x^\mu(\sigma)}{D\sigma^2} = \frac{\partial^2 x^\mu(\sigma)}{\partial\sigma^2} + \Gamma^\mu_{\nu\lambda}(x(\sigma))\frac{\partial x^\nu}{\partial\sigma}\frac{\partial x^\lambda}{\partial\sigma} = 0 \tag{3.184}$$

and

$$\frac{D^2(x^\mu+\delta x^\mu)}{D\sigma^2}=\frac{\partial^2}{\partial\sigma^2}(x^\mu(\sigma)+\delta x^\mu(\sigma))+\Gamma^\mu_{\nu\lambda}(x(\sigma))\frac{\partial x^\nu}{\partial\sigma}\frac{\partial x^\lambda}{\partial\sigma}=0 \tag{3.185}$$

Subtract the two equations and ignore second order terms in δx^λ:

$$\frac{\partial^2\delta x^\mu}{\partial\sigma^2}+\partial_\alpha\Gamma^\mu_{\nu\lambda}(x)\delta x^\alpha\frac{\partial x^\nu}{\partial\sigma}\frac{\partial x^\lambda}{\partial\sigma}+2\Gamma^\mu_{\nu\lambda}\frac{\partial\delta x^\mu}{\partial\sigma}\frac{\partial\delta x^\mu}{\partial\sigma}=0 \tag{3.186}$$

But now:

$$\frac{\partial^2\delta x^\mu}{\partial\sigma^2}\neq \text{Vector} \tag{3.187}$$

Thus it is not a useful quantity as it will vary from coordinate system to coordinate system, so one needs to re-write this with more connections:

$$\begin{aligned}\frac{D^2\delta x^\mu}{D\sigma} &= \frac{D}{D\sigma}\left(\frac{D\delta x^\mu}{D\sigma}\right)\\ &= \frac{D}{D\sigma}\left(\frac{\partial\delta x^\mu}{\partial\sigma}+\Gamma^\mu_{\nu\lambda}\delta x^\nu\frac{\partial x^\lambda}{\partial\sigma}\right)\\ &= \underbrace{\frac{\partial^2\delta x^\mu}{\partial\sigma^2}}_{\text{Term a}}+\Gamma^\mu_{\nu\lambda,\alpha}\frac{\partial x^\alpha}{\partial\sigma}\delta x^\nu\frac{\partial x^\lambda}{\partial\sigma}+\Gamma^\mu_{\nu\lambda}\frac{\partial\delta x^\nu}{\partial\sigma}\frac{\partial x^\lambda}{\partial\sigma}+\Gamma^\mu_{\nu\lambda}\delta x^\nu\underbrace{\frac{\partial^2 x^\lambda}{\partial\sigma^2}}_{\text{Term b}}+\Gamma^\mu_{\alpha\beta}\frac{\partial\delta x^\alpha}{\partial\sigma}\frac{\partial x^\beta}{\partial x^\sigma}+\Gamma^\mu_{\alpha\beta}\Gamma^\alpha_{\nu\lambda}\delta x^\nu\frac{\partial x^\lambda}{\partial\sigma}\frac{\partial x^\beta}{\partial\sigma}\end{aligned} \tag{3.188}$$

Substitute Eq 3.185 into Term b and Eq 3.186 into Term a. The equation above then simplifies to:

$$\begin{aligned}\frac{D^2\delta x^\mu}{D\sigma} &= -\Gamma^\mu_{\nu\lambda,\rho}\delta x^\rho\dot{x}^\mu\dot{x}^\lambda+\Gamma^\mu_{\rho\lambda,\nu}\dot{x}^\nu\delta x^\rho\dot{x}^\lambda-\Gamma^\mu_{\nu\lambda}\Gamma^\lambda_{\alpha\beta}\dot{x}^\alpha\dot{x}^\beta dx^\nu+\Gamma^\mu_{\alpha\beta}\Gamma^\alpha_{\nu\lambda}\delta x^\nu\dot{x}^\lambda\dot{\beta}\\ &= -\left(\Gamma^\mu_{\nu\lambda,\rho}-\Gamma^\mu_{\rho\nu,\lambda}+\Gamma^\mu_{\rho\gamma}\Gamma^\gamma_{\nu\lambda}-\Gamma^\mu_{\gamma\lambda}\Gamma^\gamma_{\rho\nu}\right)\delta x^p\dot{x}^\nu\dot{x}^\lambda\\ &\equiv R^\mu_{\nu\lambda\rho}\dot{x}^\nu\dot{x}^\lambda\delta x^\rho\end{aligned} \tag{3.189}$$

where we have defined:

$$R^\mu_{\nu\lambda\rho}=-\left(\Gamma^\mu_{\nu\lambda,\rho}-\Gamma^\mu_{\rho\nu,\lambda}+\Gamma^\mu_{\rho\gamma}\Gamma^\gamma_{\nu\lambda}-\Gamma^\mu_{\gamma\lambda}\Gamma^\gamma_{\rho\nu}\right) \tag{3.190}$$

This is known as the *Riemann tensor* and it measures geodesic deviation (also called the *intrinsic curvature of the manifold. This is different to extrinsic curvature, which corresponds to the curvature of an external embedding space.*)

Claim 10. The commutator of the covariant derivative acting on an arbitrary vector gives the Riemann tensor times the vector:

$$(\nabla_\mu\nabla_\nu-\nabla_\nu\nabla_\mu)V^\alpha=R^\alpha_{\gamma\mu\nu}V^\gamma \tag{3.191}$$

Proof 10. The covariant derivative is defined as:

$$\nabla_\mu V^\alpha\equiv\frac{\partial V^\alpha}{\partial x^\mu}+V^\mu\Gamma^\alpha_{\mu\nu} \tag{3.192}$$

For a mixed tensor:

$$\nabla_\gamma W^\mu_\nu\equiv\partial_\alpha W^\mu_\nu+\Gamma^\mu_{\gamma\sigma}W^\sigma_\nu-\Gamma^\sigma_{\gamma\nu}W^\mu_\sigma \tag{3.193}$$

Using these two definitions we can write out the second covariant derivatives as:

$$\nabla_\mu(\nabla_\nu V^\alpha)=\partial_\mu\partial_\nu V^\alpha+\Gamma^\alpha_{\nu\gamma}\partial_\mu V^\gamma+V^\gamma\partial_\mu\Gamma^\alpha_{\nu\gamma}-\Gamma^\sigma_{\nu\mu}(\partial_\sigma V^\alpha+V^\gamma\Gamma^\alpha_{\sigma\gamma})+\Gamma^\alpha_{\mu\sigma}(\partial_\nu V^\sigma+V^\gamma\Gamma^\sigma_{\nu\gamma}) \tag{3.194}$$

$$\nabla_\nu(\nabla_\mu V^\alpha) = \partial_\mu\partial_\nu A^\alpha + \Gamma^\alpha_{\mu\gamma}\partial_\nu V^\alpha + V^\gamma\partial_\nu\Gamma^\alpha_{\mu\gamma} - \Gamma^\sigma_{\nu\mu}(\partial_\sigma V^\alpha + V^\gamma\Gamma^\alpha_{\sigma\gamma}) + \Gamma^\alpha_{\nu\sigma}(\partial_\mu V^\sigma + V^\gamma\Gamma^\sigma_{\mu\gamma}) \quad (3.195)$$

Subtract these two equations:

$$(\nabla_\mu\nabla_\nu - \nabla_\nu\nabla_\mu)V^\alpha = (\partial_\mu\Gamma^\alpha_{\nu\gamma} - \partial_\nu\Gamma^\alpha_{\mu\gamma} + \Gamma^\alpha_{\mu\sigma}\Gamma^\sigma_{\nu\gamma} - \Gamma^\alpha_{\nu\sigma}\Gamma^\sigma_{\mu\gamma})V^\gamma \quad (3.196)$$

The term in the brackets on the R.H.S is simply the Riemann tensor.

So the Riemann tensor gives the degree of non-commutativity between the covariant derivative, which is to be expected, as if there is no curvature, then the covariant derivative is the same as the partial derivative(as the Γ is zero) and partial derivatives commute, which will lead to the Riemann tensor being zero. In general for $[\nabla_\mu, \nabla_\nu]$ on a vector give a $+R$ for upper indicies and $-R$ for lower indicies.

5.1. Symmetries of Riemann tensor. The first thing to do is to lower indicies.

Claim 11.

$$R_{\mu\nu\lambda\rho} = g_{\mu\alpha}R^\alpha_{\nu\lambda\rho} \quad (3.197)$$

We have already shown that at a given point on a manifold, the Christoffel symbols can be set to zero by an appropriate coordinate transformation. In this approximation:

$$R_{\mu\nu\lambda\rho} = \frac{1}{2}(g_{\mu\rho,\nu\lambda} - g_{\mu\lambda,\nu\rho} - g_{\nu\rho,\mu\lambda} + g_{\lambda\nu,\rho\mu}) \quad (3.198)$$

Proof 11. In coordinates that have $\Gamma = 0$, the Riemann tensor is:

$$R^\mu_{\nu\lambda\rho} = \partial_\lambda\Gamma^\mu_{\rho\nu} - \partial_\rho\Gamma^\mu_{\lambda\nu} \quad (3.199)$$

To lower the index:

$$\begin{aligned} R_{\mu\nu\lambda\rho} &\equiv g_{\mu\kappa}R^\kappa_{\nu\lambda\rho} \\ &= g_{\mu\kappa}(\partial_\lambda\Gamma^\kappa_{\rho\nu} - \partial_\rho\Gamma^\kappa_{\lambda\nu}) \end{aligned} \quad (3.200)$$

In this local neighborhood on the manifold we are assuming the first derivative of g to be zero, therefore this equation simplifies to:

$$\partial_\lambda\Gamma^\kappa_{\beta\nu} = \frac{1}{2}g^{\kappa\sigma}(\partial_\lambda\partial_\nu g_{\rho\sigma} + \partial_\lambda\partial_\rho g_{\sigma\nu} - \partial_\lambda\partial_\sigma g_{\nu\rho}) \quad (3.201)$$

Multiply this by $g_{\kappa\mu}$:

$$\begin{aligned} g_{\kappa\mu}\partial_\lambda\Gamma^\kappa_{\rho\nu} &= \frac{1}{2}g_{\kappa\mu}g^{\kappa\sigma}(\partial_\lambda\partial_\nu g_{\rho\sigma} + \partial_\lambda\partial_\rho g_{\sigma\nu} - \partial_\lambda\partial_\sigma g_{\nu\rho}) \\ &= \frac{1}{2}(\partial_\lambda\partial_\nu g_{\rho\mu} + \partial_\lambda\partial_\rho g_{\mu\nu} - \partial_\lambda\partial_\mu g_{\nu\rho}) \end{aligned} \quad (3.202)$$

Similarly:

$$g_{\mu\nu}\partial_\rho\Gamma^\kappa_{\lambda\nu} = \frac{1}{2}(\partial_\rho\partial_\nu g_{\lambda\mu} + \partial_\rho\partial_\lambda g_{\mu\nu} - \partial_\rho\partial_\mu g_{\nu\lambda}) \quad (3.203)$$

Substitute Eq 3.203 from Eq 3.202 and put into Eq 3.200:

$$R_{\mu\nu\lambda\rho} = \frac{1}{2}(\partial_\lambda\partial_\nu g_{\rho\mu} + \partial_\rho\partial_\mu g_{\nu\lambda} - \partial_\lambda\partial_\mu g_{\nu\rho} - \partial_\rho\partial_\nu g_{\lambda\mu}) \quad (3.204)$$

Now one can analyse the symmetries of the Riemann tensor at a given point. Eq 3.204 has two obvious symmetries and one that it is not so obvious.

- The Riemann tensor is symmetric under the first and third indicies and the second and fourth indicies:

$$R_{\mu\nu} \equiv R_{\lambda\rho\mu\nu} \quad (3.205)$$

- The Riemann tensor is anti-symmetric under the first and second indicies and the third and fourth indicies:

$$R_{\mu\nu\lambda\rho} = -R_{\mu\nu\rho\lambda} \equiv R_{\nu\mu\rho\lambda} \tag{3.206}$$

- This is not the obvious symmetry and its to do with the cyclicity of the Riemann tensor:

Claim 12. $R_{\mu\nu\lambda\rho} + R_{\mu\rho\nu\lambda} + R_{\mu\lambda\rho\nu} = 0$
This is the statement as:

$$R_{\mu[\nu\lambda\rho]} = 0 \tag{3.207}$$

Proof 12. Writing out the Riemann tensor explicitly from Eq 3.204:

$$\begin{aligned} R_{\mu\nu\lambda\rho} + R_{\mu\rho\nu\lambda} + R_{\mu\lambda\rho\nu} &= \frac{1}{2}(\partial_\lambda\partial_\nu g_{\rho\mu} + \partial_\rho\partial_\mu g_{\nu\lambda} - \partial_\lambda\partial_\mu g_{\nu\rho} - \partial_\rho\partial_\nu g_{\lambda\mu}) \\ &+ \frac{1}{2}(\partial_\rho\partial_\lambda g_{\nu\mu} + \partial_\nu\partial_\mu g_{\lambda\rho} - \partial_\rho\partial_\mu g_{\lambda\nu} - \partial_\nu\partial_\lambda g_{\rho\mu}) \\ &+ \frac{1}{2}(\partial_\nu\partial_\rho g_{\lambda\mu} + \partial_\lambda\partial_\mu g_{\rho\nu} - \partial_\nu\partial_\mu g_{\rho\lambda} - \partial_\lambda\partial_\rho g_{\nu\mu}) \end{aligned} \tag{3.208}$$

Using the symmetry:

$$g_{\mu\nu} \equiv g_{\nu\mu} \tag{3.209}$$

and the commutativity of the derivatives:

$$R_{\mu\nu\lambda\rho} + R_{\mu\rho\nu\lambda} + R_{\mu\lambda\rho\nu} \equiv 0 \tag{3.210}$$

The number of independent components of Riemann tensor, are reduced by the symmetries. The Riemann tensor can be thought of as a matrix with two lower indicies where each index is a pair of anti-symmetric indicies i.e:

$$R_{\psi\phi} = R_{\mu\nu\rho\lambda} \tag{3.211}$$

So think about a symmetric matrix, whose indicies takes on values from 1...D, where D is the dimension of the matrix, and each ψ and ϕ is anti-symmetric matrix, so it will take on values:

$$\frac{D(D-1)}{2} \tag{3.212}$$

The factor of half come as the two values can be flipped, i.e:

$$\mu, \nu \to \nu, \mu \tag{3.213}$$

Therefore the overall symmetric matrix, $R_{\psi\phi}$ will have $\frac{D(D-1)}{2}$ dimensions, and a symmetric matrix the values:

$$\frac{1}{2}D'(D'+1) \tag{3.214}$$

where:

$$D' \equiv \frac{D(D-1)}{2} \tag{3.215}$$

therefore:

$$\begin{aligned} \#\text{ of values for Riemann} &= \frac{1}{2}\left(\frac{D(D-1)}{2}\right)\left(\frac{D(D-1)}{2}+1\right) \\ &= \frac{1}{8}D(D-1)(D^2-D+2) \end{aligned} \tag{3.216}$$

However, there is a still the symmetry of the cyclicity that needs to be taken into account. This does not provide an obvious condition on the number of values, as some of the symmetries of cyclicity may have already been imposed by the two previous conditions, and one would not want to count any symmetries twice!.

Claim 13. The cyclicity condition implies that the Riemann tensor is totally anti-symmetric on the 4 indicies, if the previous two conditions are imposed.

Proof 13. The only way to see this, is by explicitly writing out the indicies:

$$\begin{aligned} \mu\nu\lambda\rho &\rightarrow \nu\mu\lambda\rho \rightarrow -\mu\nu\rho\lambda \\ \mu\rho\nu\lambda &\rightarrow \nu\rho\mu\lambda \rightarrow -\mu\lambda\rho\nu \\ \mu\lambda\rho\nu &\rightarrow \nu\lambda\rho\mu \rightarrow -\mu\rho\nu\lambda \end{aligned} \tag{3.217}$$

Therefore it is totally anti-symmetric on all 4 indicies.

This gives the number of equations constraining the parameters as:

$$\frac{D(D-1)(D-2)(D-3)}{4!} \tag{3.218}$$

Therefore the number of independent components of Riemann is:

$$\begin{aligned} \text{Total \# of components} &= \frac{1}{8}D(D-1)(D^2-D+2) - \frac{D(D-1)(D-2)(D-3)}{4!} \\ &= \frac{D^2}{12}(D^2-1) \end{aligned} \tag{3.219}$$

So for 4 dimensions the number of parameters for the Riemann tensor is 20 and this is what is expected as previously it was found that there are 20 unidentified parameters when trying to set the second order derivatives of the metric to zero at a point. The Riemann tensor fixes these quantities by specifying a value for each of component.

Theorem 3. If:

$$R^{\lambda}_{\mu\nu\rho}(x) = 0 \qquad \forall x \in M \tag{3.220}$$

then there exists a coordinate system in which:

$$g_{\mu\nu} = \eta_{\mu\nu} \tag{3.221}$$

i.e it is just Minkowski space (in some funny coordinate system). Thus a space-time is flat, iff $R^{\mu}_{\nu\lambda\rho} = 0$.

5.2. Ricci Tensor. This is constructed from the Riemann tensor. One can contract the first two indicies and the last two indicies of the Riemann tensor as they are anti-symmetric, therefore:

$$R_{\lambda\rho\mu\nu}g^{\mu\nu} = 0 \tag{3.222}$$

To get non-zero values one has to contract the first and third or second and fourth indicies. The resulting tensor is known as the Ricci tensor. It is a symmetric 4×4 matrix, hence it has 10 components:

$$R_{\alpha\beta} = R^{\lambda}_{\alpha\lambda\beta} \tag{3.223}$$

$$\equiv R_{\beta\gamma} \tag{3.224}$$

There is another quantity, known as the Ricci scalar, that is defined as:

$$R = g^{\alpha\beta}R_{\alpha\beta} \tag{3.225}$$

It is also called the intrinsic curvature scalar, i.e it only depends on quantities in the metric (in the space itself) and does not depend on how the space is embedded into a manifold.

5.3. Weyl tensor. This is related to the conformal symmetry of the space-time, i.e the symmetry of multiplying the metric by an arbitrary function of position. It is defined as:

$$C_{\alpha\beta\mu\nu} \equiv R_{\alpha\beta\mu\nu} - \frac{1}{2}g_{\alpha_\mu}R_{\nu\beta} + \frac{1}{2}g_{\alpha\nu}R_{\mu\beta} + \frac{1}{2}g_{\beta\mu}R_{\nu\alpha} - \frac{1}{2}g_{\beta\nu}R_{\mu\alpha} - \frac{1}{6}g_{\alpha\nu}g_{\mu\beta}R + \frac{1}{6}g_{\alpha\mu}g_{\nu\beta}R \quad (3.226)$$

and it is constructed by removing the trace of the Riemann tensor. The Riemann tensor is anti-symmetric in α, β and μ, ν, so the components on the diagonal are zero. However, in general $\alpha\mu$ and $\rho\nu$ components are not and so need to be removed for the Weyl tensor. Therefore:

$$g^{\alpha\mu}C_{\alpha\beta\mu\nu} = 0 \quad (3.227)$$

C is traceless on any two indicies.

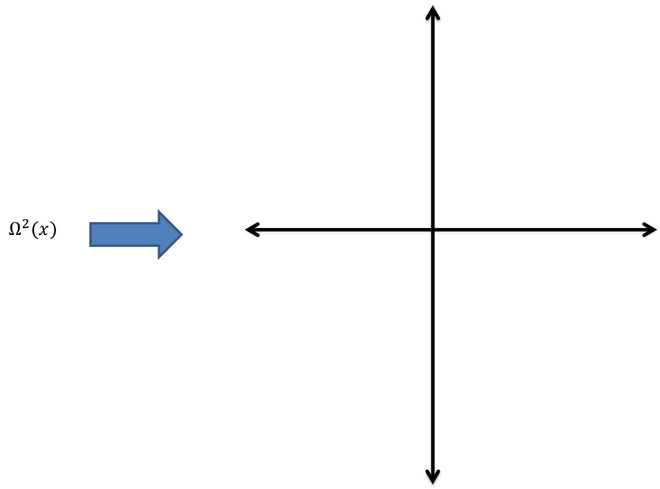


FIGURE 16. Conformal scaling function

It has 10 independent components. Therefore Riemann can be thought of as Ricci + Weyl. The Ricci components are fixed by the matter and the Weyl components come from the gravitational waves.

THEOREM 4. If:

$$C_{\alpha\beta\gamma\delta} = 0 \quad (3.228)$$

then there exists coordinates in which:

$$g_{\mu\nu} = \Omega^2(x)\eta_{\mu\nu} \quad (3.229)$$

This $\Omega^2(x)$ is known as a conformal factor, so the space-time is conformally flat.

The function $\Omega^2(x)$ is a local scaling, that does not change any angles, as shown in figure 16.

5.4. Bianchi identities. The Bianchi identity is an identity for the derivative of the Riemann tensor and is defined via:

$$R_{\alpha\beta\mu\nu;\kappa} + R_{\alpha\beta\kappa\mu;\nu} + R_{\alpha\beta\nu\kappa;\mu} = 0 \quad (3.230)$$

Because the Riemann tensor is anti-symmetric on the indicies μ, ν, then just as before, when a tensor is constructed by cycling three indicies, two of which are anti-symmetric, its the same as anti-symmetrizing. Therefore the Bianchi identity is equivalent to:

$$R_{\alpha\beta[\mu\nu;k]} = 0 \tag{3.231}$$

The property of the Riemann tensor in Eq 3.207, follows in general from the Jacobi identity:

$$([\nabla_\mu[\nabla_\nu, \nabla_\lambda]] + \text{cycles})\,\psi = 0 \tag{3.232}$$

where ψ is some scalar. Then the Bianchi identity follows from a similar identity except with the operator acting on a vector, V^α

$$([\nabla_\mu[\nabla_\nu, \nabla_\lambda]] + \text{cycles})\,V^\alpha = 0 \tag{3.233}$$

The full formal derivation is quite lengthy and would take us too far of course at this point.[3]

6. Summary of differential geometry

There is an easy way to outline the concepts of differential geometry, especially in general relativity. One can divide the concepts into three layers:

- The most fundamental object in differential geometry is a manifold. A manifold is formally defined in definition 1, however, it can just be thought of as an object that locally looks flat.

  A manifold does not have any intrinsic way to define a coordinate system on it. Therefore one is at liberty to choose which ever coordinates that are appropriate. The idea is to describe certain geometric objects on the manifold, such that the coordinates being used make no difference to the description. An obvious object is a tensor:

  $$T^{\alpha_1\ldots\alpha_m}_{\beta_1\ldots\beta_n} \tag{3.234}$$

  This is a (m, n) rank tensor. Mathematically, a tensor is defined by a transformation law, given by Eq 2.1. A manifold also contains a tangent vector space at each point on the manifold.
- It is not possible compare different points on a manifold, as taking a derivative of a tensor at one point does not transform as a tensor, instead one has an extra term, as is shown in Term 2 in Eq 3.36. To cancel of this term, the covariant derivative is introduced in Eq 3.192. The covariant derivative gives a natural concept of parallel transport, which is the idea that vectors at one point in the manifold can now be compared with vectors in another point in the manifold. This also induces a special curve on the manifold, *the geodesic*. The geodesic curves parallel transport their own tangent vector:

  $$U^\mu \nabla_\mu U^\nu = 0 \tag{3.235}$$

  The next thing to define is Riemann curvature. It is defined by the non-commutativity of the covariant derivative, given in Eq 3.191. Once this has been introduced, the indices of the Riemann tensor can be contracted to give the Ricci tensor and the Ricci scalar.
- Finally, one adds metric to the manifold, which gives a measure of distance of the manifold. With a metric, the line element between two points is given by:

  $$ds^2 \equiv g_{\mu\nu} dx^\mu dx^\nu \tag{3.236}$$

  In general relativity the metric is defined in a way that its covariant derivative is zero, which actually follows from the fact that the theory is torsion free.

[3] Full proof can be found on many websites, like http://www-personal.umich.edu/ jbourj/gr/homework%205.pdf

CHAPTER 4

Einstein's field equations

This chapter will form the key result in the general theory of relativity; the Einstein field equations. The chapter is deliberately kept very short as to highlight the importance of Einstein's equations.

1. Einstein tensor

Contract the first and third index of Eq 3.230:

$$g^{\alpha\mu}g^{\beta\kappa}(R_{\alpha\beta\mu\nu;\kappa} + R_{\alpha\beta\kappa\mu;\nu} + R_{\alpha\beta\nu\kappa;\mu}) = 0 \tag{4.1}$$

Contracting the indices explicitly:

$$\begin{aligned} g^{\alpha\mu}g^{\beta\kappa}R_{\alpha\beta\mu\nu;\kappa} + g^{\alpha\mu}g^{\beta\kappa}R_{\alpha\beta\kappa\mu;\nu} + g^{\alpha\mu}g^{\beta\kappa}R_{\alpha\beta\nu\kappa;\mu} &= g^{\beta\kappa}R_{\beta\nu;\kappa} - g^{\alpha\mu}g^{\beta\kappa}R_{\alpha\beta\mu\kappa;\nu} + g^{\alpha\mu}R_{\alpha\nu;\mu} \\ &= \nabla^{\beta}R_{\beta\nu} - \nabla_{\nu}R + \nabla^{\alpha}R_{\alpha\nu} = 0 \end{aligned} \tag{4.2}$$

where $R_{\beta\nu}$ and $R_{\alpha\nu}$ are the Ricci tensors and R is the Ricci scalar. The first and third terms are the same as α and β are dummy indicies, therefore:

$$\begin{aligned} \nabla^{\beta} &= \frac{1}{2}\nabla_{\mu}R \\ &= \nabla^{\beta}(R_{\beta\nu} - \frac{1}{2}g_{\beta\nu}R) = 0 \end{aligned} \tag{4.3}$$

Einstein realised[1] that this quantity was of fundamental important and would lead to being a term in his field equations.

Definition 15. Another tensor is defined as:

$$G_{\beta\nu} \equiv R_{\beta\nu} - \frac{1}{2}g_{\beta\nu}R \tag{4.4}$$

This is known as the *Einstein tensor.*

Under this definition Eq 4.3 becomes:

$$\nabla^{\beta}G_{\beta\nu} = 0 \tag{4.5}$$

2. Stress-energy tensor

In relativity, the density of matter (or energy) is described by the stress-energy tensor, $T^{\mu\nu}$.

Imagine a collection of particles, moving with a four velocity $U^{\mu}(x)$. Since the metric is Minkowski like (i.e with the time component with the opposite sign), the four velocity must obey:

$$g_{\mu\nu}U^{\mu}U^{\nu} = -1 \tag{4.6}$$

In the rest frame of a fluid, the four velocity is just $(1, \vec{0})$ and the energy is given by $\rho(\vec{x})c^2$, which is the mass density and the pressure, $P(\vec{x})$:

[1] In fact, in his first attempt, he got this term wrong!

$$\begin{aligned} T^{00} &= \rho(\vec{x})c^2 \\ T^{0i} &= 0 \\ T^{ij} &= P(\vec{x})\delta^{ij} \end{aligned} \tag{4.7}$$

Thus, for an isotropic fluid:

$$T^{\mu\nu} = (P + \rho c^2)U^\mu U^\nu + P\eta^{\mu\nu} \tag{4.8}$$

This equation has been motivated by the physics in the rest frame, however it is a tensor, hence by definition it is the same in all coordinate systems, i.e all frames. Once can that this tensor is correct in the rest frame explicitly:

$$\begin{aligned} T^{00} &= (P + \rho c^2)U^0 U^0 + \rho\eta^{00} \\ &= (P + \rho c^2) - P \\ &\equiv \rho c^2 \end{aligned} \tag{4.9}$$

$$\begin{aligned} T^{0i} &= (P + \rho c^2)U^0 \underbrace{U^1}_{\equiv 0} + P \underbrace{\eta^{0i}}_{\equiv 0} \\ &= 0 \end{aligned} \tag{4.10}$$

$$\begin{aligned} T^{ij} &= (P + \rho c^2) \underbrace{U^i U^j}_{\equiv 0} + P \underbrace{\eta^{ij}}_{\delta^{ij}} \\ &= P\delta^{ij} \end{aligned} \tag{4.11}$$

The basic property of $T^{\mu\nu}$ is that it is conserved, i.e:

$$\partial_\mu T^{\mu\nu} = 0 \tag{4.12}$$

In fact, for an isotropic fluid, this condition gives rise to the Navier-Stokes equation.

Einstein said, $T^{\mu\nu}$, is the same in general relativity by simply replacing the flat metric, η, with a general curvature metric g. Conservation of stress-energy then become a covariant derivative from a regular derivative:

$$\nabla_\mu T^{\mu\nu} = 0 \tag{4.13}$$

This equation implies the conservation of energy and momentum. To see how, lets write out Eq 4.13 at a given point (i.e setting the Γ's to zero) and for the components in which $\nu = 0$:

$$\partial_\mu T^{\mu\nu} = \partial_0 T^{00} + \partial_i(T^{i0}) = 0 \tag{4.14}$$

If the equation is now integrated over space:

$$\int d^3\vec{x}\left(\partial_0(T^{00} + \partial_i(T^{i0}))\right) = 0 \tag{4.15}$$

Since the anti-derivative and derivative commute:

$$\partial_0 \int d^3\vec{x} T^{00} + \int d^3\vec{x}\partial_i T^{i0} = 0 \tag{4.16}$$

But:

$$\int d^3\vec{x} T^{00} \equiv \text{Total energy}, E \tag{4.17}$$

Thus Eq 4.16 becomes:

$$\partial_0 E + \int d^3x(\partial_i T^{i0}) = 0 \tag{4.18}$$

The second term is an integral over the divergence, so one can use Gauss's divergence theorem, that states that the total integral of the divergence of a vector is simply the flux integrated over the closed surface. Therefore Eq 4.18 becomes:

$$\frac{\partial E}{\partial t} = -\int_{surface} T^{i0} dS^i \tag{4.19}$$

This is the statement that the rate of change of energy is equal to the flow of momentum over a given surface. Now turn to the i components of Eq 4.13, with $\nu = 0$:

$$\partial_0 T^{0i} + \partial_i T^{ii} = 0 \tag{4.20}$$

Once again integrate over space:

$$\int d^3\vec{x}_0 T^{0i} + \int d^3\vec{x}\partial_i T^{ii} = 0 \tag{4.21}$$

Once again use the commutativity between the derivative and anti-derivative:

$$\partial_0 \underbrace{\int d^3\vec{x} T^{0i}}_{\text{Term a}} + \int d^3\vec{x}\partial_i T^{ii} = 0 \tag{4.22}$$

Term a is the total momentum:

$$\frac{\partial \vec{p}_i}{\partial t} + \int d^3\vec{x}\partial_i T^{ii} = 0 \tag{4.23}$$

Once again use the divergence theorem for the second term:

$$\frac{\partial \vec{p}_i}{\partial t} = -\int T^{ii} dS \tag{4.24}$$

But T^{ii} is the pressure and hence integrating over an area will yield a force. So infact Eq 4.24 is simply the statement of Newton's second law:

$$\frac{\partial \vec{p}_i}{\partial t} = -F \tag{4.25}$$

Therefore Eq 4.13 encompasses the conservation of energy and momentum and it comes fundamentally from Noether's theorem under the time and spatial translation invariance.

3. The field equations

By looking at Eq 4.13 and Eq 4.5, it is tempting to formulate the field equations as:

$$G^{\mu\beta} \propto T^{\mu\beta} \Rightarrow G^{\mu\beta} \equiv \kappa T^{\mu\beta} \tag{4.26}$$

The trick now is to realise that this equation must reduce to the Newtonian limit for sufficiently small curvatures and thus a direct comparison to the equation in the Newtonian limit can be used to calculate the coefficient κ.

The first thing to do is calculate $G^{\alpha\beta}$ for the metric, that deviates at a very small scale from the flat Minkowski metric:

$$\begin{aligned} g_{\alpha\beta} &= \eta_{\alpha\beta} + h_{\alpha\beta} \\ g^{\alpha\beta} &= \eta^{\alpha\beta} - h_{\alpha\beta} \end{aligned} \tag{4.27}$$

Now we can calculate the Riemann tensor:

$$R_{\mu\nu\rho\lambda} = \frac{1}{2}(h_{\mu\rho,\nu\lambda} - h_{\nu\rho,\mu\lambda} - h_{\mu\lambda,\nu\rho} + h_{\lambda\nu,\rho\mu}) + \mathcal{O}(h^2) \tag{4.28}$$

Its convenient to express Eq 4.26 in terms of the Ricci tensor and Ricci scalar:

$$R^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}R = \kappa T^{\alpha\beta} \tag{4.29}$$

If one takes the trace of this equation:

$$\begin{aligned} g_{\alpha\beta}R^{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}g^{\alpha\beta}R &= \kappa g_{\alpha\beta}T^{\alpha\beta} \\ R - \frac{1}{2}(4)R &= \kappa g_{\alpha\beta}T^{\alpha\beta} \\ -R = \kappa T \end{aligned} \tag{4.30}$$

So we can substitute Eq 4.30 into Eq 4.29:

$$R^{\alpha\beta} = \kappa(T^{\alpha\beta} - g^{\alpha\beta}T) \tag{4.31}$$

This equation is slightly easier to work with as all the T dependence on the R.H.S and the only thing needed is R.

Next the R^{00} component needs to be extracted from the Riemann tensor (as the eventual aim is to get the weak field approximation, which has an h^{00}). The clever trick one has to realise now, is that R^{00} is equivalent to R_{00}, since the raising and lowering of the indicies will involve applying $g^{\mu\nu}, g_{\mu\nu}$ and since:

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu} \tag{4.32}$$

and the Riemann tensor is already of $\mathcal{O}(h)$, then applying g will give terms with ηh and $\mathcal{O}(h^2)$, but $\mathcal{O}(h^2)$ terms are ignore and ηh will simply give 1, since η is flat. So lets proceed and calculate R_{00}:

$$R_{00} = \eta^{\mu\rho}R_{\mu 0\rho 0} \tag{4.33}$$

Here again the g will have $\eta - h$ term that will give rise to $\mathcal{O}(h^2)$ terms which are just ignored for reasons given above, leaving only the η. Let's look at $R_{\mu 0\rho 0}$:

$$R_{\mu 0\rho 0} = \frac{1}{2}(h_{\mu\rho,00} - h_{0\rho,\mu 0} - h_{\mu 0,0\rho} + h_{00,\rho\mu}) \tag{4.34}$$

But recall that the field is also in a static limit, i.e all the time derivatives are zero. Which means the expression simplifies to:

$$R_{00} = \eta^{\mu\rho}\frac{1}{2}h_{00,\mu\rho} \tag{4.35}$$

since we are ignoring time derivatives, μ and ρ are restricted to spatial components:

$$R_{00} = n^{ij}\frac{1}{2}h_{00,ij} = -\frac{1}{2}\nabla^2 h_{00} \tag{4.36}$$

Therefore:

$$R_{00} = -\frac{1}{2}\nabla^2 h_{00} = \kappa(T_{00} - \frac{1}{2}\eta^{00}T) \tag{4.37}$$

For T, assume that $U^\mu \approx (1, \vec{0})$ (fluid at rest) and $P << \rho c^2$ (e.g ideal gas $P \approx \rho c^2, v << c$):

$$T^{\mu\nu} \approx \begin{pmatrix} \rho c^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.38}$$

Therefore:

$$T = \eta_{\mu\nu} T^{\mu\nu} = -\rho c^2 \Rightarrow T_{00} = \rho c^2 \tag{4.39}$$

Substitute Eq 4.36 and Eq 4.39 into Eq 4.37:

$$\begin{aligned} -\frac{1}{2}\nabla^2 h_{00} &= \kappa(\rho c^2 + \frac{1}{2}(-\rho c^2)) \\ &= \frac{\kappa}{2}\rho c^2 \end{aligned} \tag{4.40}$$

Therefore:

$$-\nabla^2 h_{00} = \kappa \rho c^2 \tag{4.41}$$

Now this can be compared to the previously calculated weak field limit in Eq 2.57, which gives:

$$\kappa \equiv \frac{8\pi G}{c^4} \tag{4.42}$$

Finally, this gives the Einstein field equations:

$$G_{\alpha\beta} = \frac{8\pi G}{c^4} T_{\alpha\beta} \tag{4.43}$$

CHAPTER 5

Solutions to Einstein's equations

This chapter could also be called *cosmology*, as solutions to Einstein's equations give rise to (almost) the entire field of cosmology. At the time that Einstein worked out this equation, he did not imagine that the entire Universe would follow it. Of course, in the derivation of this equation, the only things he used were some clever comparison between curved space and Newtonian gravity. But it turns out that all gravitational object follow this equation to extreme precision.

Einstein started to play around with his equation to see what they predicted about the universe. These equations are highly tangled up and are very difficult to solve without making many simplifying assumptions. Einstein used his equations to try to produce a universe that was static, as nobody knew about the expansion of the universe at that time. Many of the first models that he proposed were incorrect, however there was one thing that was thought to be incorrect at first (infact Einstein called it his "Biggest blunder") and now turns out to be true and that is the cosmological constant.

To see how this comes in, recall the motivation of writing out the equation:

$$G_{\alpha\beta} \propto T_{\alpha\beta} \tag{5.1}$$

It was known that:

$$\nabla_{\mu} T^{\mu\nu} = 0 \tag{5.2}$$

Therefore one required a tensor that also satisfied this condition, to be able to equate it to $T^{\mu\nu}$ and it turns out that adding a term of the form:

$$G_{\alpha\beta} = \frac{8\pi G}{c^4} T_{\alpha\beta} + c g_{\alpha\beta} \tag{5.3}$$

will still satisfy the fact that the covariant derivative of both sides would be zero. This is easy to see, as one of the assumptions of general relativity is that the metricity is zero:

$$\nabla_{\alpha} g^{\alpha\beta} = 0 \tag{5.4}$$

Using this additional term, Einstein shows that it is possible to make a static universe model, however we shall see that this model turns out to be unstable.

1. Friedmann-Robertson-Walker(FRW) equation

To simplify the field equations solutions, one imposes some symmetries. These were first discovered by Friedmann, Robertson and Walker and the symmetries are:

- **Isotropy**: Uniformity in all directions, i.e if one looks at any direction in space it looks the same at large scales.
- **Homogeneity**: Symmetry under positional translation i.e i.e every point in space is the same.

Together, these symmetries are sometimes called the *maximal symmetry* in space[1]. Together, these two assumptions are called the "cosmological principle".

[1] This is infact a humbling assumption to make; the universe we live in has no special points, so there is nothing special about our position in the universe. This is the Copernican view applied to the whole universe. We

In d dimensions, the greatest number of possible symmetries of a manifold with a metric is:

$$\frac{d(d+1)}{2} = \#\text{ of possible symmetries} \tag{5.5}$$

The proof for this is quite long and can be found, for example, in [**3**]. In 3 dimensions, there are 6 possible symmetries; defined by three different curved surfaces.

EXAMPLE 9. One of the curved surfaces is a 3 sphere, $\mathbb{S}^3$. This has a positive curvature and leads to a closed surface. The three sphere can be embedded in a 4 dimensional Euclidean space, and is characterised by the following equation:

$$x^2 + y^2 + z^2 + u^2 + 1 \tag{5.6}$$

This has 6 rotational symmetries.

EXAMPLE 10. A 3 dimensional Euclidean space, $\mathbb{E}^3$, is an example of a *flat* space and is characterized by:

$$x^2 + y^2 + z^2 = 1 \tag{5.7}$$

This appears to be the actual space that we live in and has 3 rotations and 3 translations as part of its symmetry.

EXAMPLE 11. The final surface is a 3 dimensional hyperbolic surface, $\mathbb{H}^3$, which has a negative curvature. It can be embedded in a 4 dimensional Minkowski space and thus follows:

$$t^2 - x^2 - y^2 - z^2 = 1 \tag{5.8}$$

Here one can have three ordinary rotations around $(x, y), (x, z), (y, z)$ axis. Or one can do have translations between $(t, x), (t, y), (t, z)$ since these take t to spatial components and vice-verse, these are called *boosts*. This appears to be the *space-time* we live in.

The metric for $\mathbb{S}^3$ is given by solving Eq 5.6:

$$\begin{aligned} x &= \sin\chi \sin\theta \cos\phi \\ y &= \sin\chi \sin\theta \sin\phi \\ z &= \sin\chi \cos\theta \\ u &= \cos\chi \end{aligned} \tag{5.9}$$

The metric imposed by the embedding space is:

$$ds^2 = d\chi^2 + \sin^2\chi(d\phi^2 + \sin^2\theta d\phi^2) \tag{5.10}$$

This is the metric of $\mathbb{S}^3$ with maximal symmetry. One can ask what is the 2 dimensional space at a fixed χ, the term in the bracket is obviously a 2 sphere, and $\sin^2\chi$ acts as the radius squared of the two sphere, so it is often defined as:

$$r \equiv \sin\chi \tag{5.11}$$

This gives the line element:

$$ds^2 = \frac{dr^2}{1 - r^2} + r^2(d\phi^2 + \sin^2\theta d\phi^2) \equiv \frac{dr^2}{1 - r^2} + r^2 d\Omega_2^2 \tag{5.12}$$

where Ω_2 is the metric on $\mathbb{S}^2$. Here one has assumed that $\mathbb{S}^3$ had unit radius, if the radius was r_0, then:

---

are simply sitting on a rock around an average star and the universe would look the same to any other observer, positioned somewhere else in the universe)

$$ds^2 = \frac{dr^2}{1 - \frac{r^2}{r_0^2}} + r^2 d\Omega_2^2 \tag{5.13}$$

for $r << r_0$, the metric looks flat. Similarly for 3 dimensional hyperbolic space, $\mathbb{H}^3$:

$$\begin{aligned} t &= \cosh\chi \\ x &= \sinh\chi \sin\theta \cos\theta \\ y &= \sinh\chi \sin\theta \sin\phi \\ z &= \sinh\chi \cos\theta \end{aligned} \tag{5.14}$$

The metric induced from Minkowski space:

$$\begin{aligned} ds^2 &= -dt^2 + d^2x + dy^2 + dz^2 \\ &= d\chi^2 + \sinh^2\chi d\Omega_2^2 \end{aligned} \tag{5.15}$$

again defining:

$$r \equiv \sinh\chi \tag{5.16}$$

This gives the metric with radius r_0:

$$ds^2 = \frac{dr^2}{1 + \frac{r^2}{r_0^2}} + r^2 d\Omega_2^2 \tag{5.17}$$

These are the spatial metrics for the simplest most symmetrical universes. Now lets look at a space-time metric:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \tag{5.18}$$

Lets also work in units with $c \equiv 1$ for simplicity:

$$ds^2 = g_{00}dt^2 + 2g_{0i}dtdx^i + g_{ij}dx^i dx^j \tag{5.19}$$

Homogeneity and isotropy leads to g_{00} being constant, as a scalar function of space and time that is translation invariant and rotational invariant has to be constant with respect to space, therefore:

$$g_{00} \equiv g_{00}(t) \neq g_{00}(\vec{x}) \tag{5.20}$$

This can be absorbed in dt^2 by simply rescaling the time coordinate accordingly:

$$-g_{00}dt^2 \equiv dt'^2 \Rightarrow t' = \int^t \sqrt{-g_{00}(t)}dt \tag{5.21}$$

Therefore the metric becomes:

$$ds^2 = -dt'^2 + 2g'_{0i}dt'dx^i + g_{ij}dx^i dx^j \tag{5.22}$$

The primes can now be dropped by simply defining new coordinates:

$$ds^2 \equiv -dt^2 + 2g_{0i}dtdx^i + g_{ij}dx^i dx^j \tag{5.23}$$

Now lets look at the g_{0i} term, it is a vector and a vector, by definition, has a direction. Since the space is assumed to be isotropic, there cannot be any preferred direction, therefore g_{0i} must be zero.

The final term is the spatial part of the metric, g_{ij}. It is, in general, a function of (x, t). Under maximal symmetry, the spatial components, $\vec{x}$, must be in $\mathbb{S}^3, \mathbb{E}^3, \mathbb{H}^3$. Thus it follows that g_{ij} is:

$$g_{ij}(t,\vec{x}) = a^2(t)\gamma_{ij}(\vec{x}) \tag{5.24}$$

where $\gamma_{ij}(\vec{x})$ is the metric on $\mathbb{S}^3, \mathbb{E}^3, \mathbb{H}^3$. The result of isotropy and homogeneity is that we can always choose coordinates, such that the line element:

$$ds^2 = -dt^2 + a^2(t)\gamma_{ij}dx^i dx^j \tag{5.25}$$

This is known as the FRW line element, corresponding to the FRW metric. The quantity a is the scale factor of the universe.

1.1. Introducing matter. Assume the simplest kind of matter; a perfect fluid:

$$T^{\mu\nu} = (\rho + P)U^\mu U^\nu + Pg^{\mu\nu} \tag{5.26}$$

Isotropy implies all vectors are zero for reasons explained above, thus the four velocity must reduce to:

$$U^\mu = (U^0, \vec{0}) \tag{5.27}$$

and from $d\tau^2 = -g_{\mu\nu}dx^\mu dx^\nu$:

$$g_{\mu\nu}U^\mu U^\nu = -1 \tag{5.28}$$

Since $g_{00} = -1$, then $U^0 \equiv 1$, this is a future pointing time-like vector. This means $T_{\mu\nu}$ has the form:

$$T^{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & \frac{P\gamma^{11}}{a^2} & 0 & 0 \\ 0 & 0 & \frac{P\gamma^{22}}{a^2} & 0 \\ 0 & 0 & 0 & \frac{P\gamma^{33}}{a^2} \end{pmatrix} \tag{5.29}$$

and the metric tensor is:

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \gamma^{11}a^2 & 0 & 0 \\ 0 & 0 & \gamma^{22}a^2 & 0 \\ 0 & 0 & 0 & \gamma^{33}a^2 \end{pmatrix} \tag{5.30}$$

another useful quantity is the inverse of the metric tensor:

$$g^{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \frac{1}{\gamma^{11}a^2} & 0 & 0 \\ 0 & 0 & \frac{1}{\gamma^{22}a^2} & 0 \\ 0 & 0 & 0 & \frac{1}{\gamma^{33}a^2} \end{pmatrix} \tag{5.31}$$

Actually, isotropy and homogeneity force $T^{\mu\nu}$ to take the form of a perfect fluid. When thought about carefully this fact is coming from Einstein's great insight that geometry and matter are related, hence the type of matter in the universe will be constrained by the geometry of the universe.

1.2. Einstein field equations for FRW metric. Let's work out the Field equations for the FRW metric. The equations are found by computing the Ricci tensor and scalar:

$$\begin{aligned} R_{00} &= -\frac{3\ddot{a}}{a} \\ R_{0i} &= 0 \\ R_{ij} &= R^{(3)}_{ij}(\gamma) + (a\ddot{a} + 2\dot{a}^2)\gamma_{ij} \\ R &= 6\left(\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{2}\right)^2 + \frac{k}{a^2}\right) \end{aligned} \tag{5.32}$$

where k is some constant that represents the curvature in the maximal symmetric solutions. Thus the 00 component for the Einstein tensor is:

$$\begin{aligned} G_{00} &= R_{00} - \frac{1}{2} g_{00} R \\ &= R_{00} + \frac{1}{2} R \\ &= 3\left(\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right) \end{aligned} \tag{5.33}$$

The other ingredient for the 00 field equation is the 00 component of the stress-energy tensor:

$$T_{00} = 8\pi k \rho \tag{5.34}$$

in fact:

$$\begin{aligned} \mathbb{S}^3 \Rightarrow k &= \frac{1}{r_0^2} \\ \mathbb{H}^3 \Rightarrow k &= -\frac{1}{r_0^2} \\ \mathbb{E}^3 \Rightarrow k &= 0 \end{aligned} \tag{5.35}$$

Thus the 00 field equation is:

$$3\left(\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2}\right) = 8\pi k \rho \tag{5.36}$$

This is known as the *Friedmann* equation. As the r_0 is a scale factor radius, it can be absorbed into $a(t)$, one can define r_0 such that it is 1 for the present value, since it can simply be absorbed into the scale factor. In this case the values of k are just $\pm 1, 0$ and this is how it is represented in most literature. Another thing to note about this equation is that it only has first time derivatives of a, thus it is known as a *constraint* equation (an equation involving $\ddot{a}$ is called an *evolution* equation).

Now lets turn to the spatial equation. The Einstein tensor is:

$$\begin{aligned} G_{ij} &= R_{ij} - \frac{1}{2} a^2 \gamma_{ij} R \\ &= \dot{a}^2 + 2a\ddot{a} + k \end{aligned} \tag{5.37}$$

and the stress and energy tensor:

$$T_{ij} = \frac{P\gamma ij}{a^2} \tag{5.38}$$

Therefore the spatial field equation is:

$$\dot{a}^2 + 2a\ddot{a} + k = 8\pi G \rho a^2 \tag{5.39}$$

As a final equation, it is useful to use the fact that the stress energy tensor is a constant with respect to the covariant derivative:

$$\begin{aligned} \nabla_\mu T^{\mu\nu} &= \partial_\mu T^{\mu\nu} + \Gamma^\mu_{\alpha\nu} T^{\alpha\nu} + \Gamma^\nu_{\alpha\mu} T^{\mu\alpha} \\ &= 0 \end{aligned} \tag{5.40}$$

This implies:

$$\dot{\rho} = -\frac{3\dot{a}}{a}(P+\rho) \tag{5.41}$$

this is also known as the *continuity* equation.

These three equations describe the evolution of a universe with maximal symmetry, also known as the FRW universe. At this level of the discussion the pressure, $P(t)$, is a completely arbitrary function, ρ on the other hand must satisfy Eq 5.41. Infact, Eq 5.41 is not independent of the other two as the Bianchi identity gives the continuity equation. There is one more equation that turns out to be very useful, that is obtained by adding together Eq 5.39 and 5.36:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) \tag{5.42}$$

This is known as *Raychaudhuri* equation[**4**] and this is an evolution equation as it involves the second time derivative of a. In general, the nature of Einstein's field equations is such that the 00 equation is usually a constraint equation and the G_{ij} equations are evolution equations.

1.3. Different types of matter. The evolution of the universe can be described once the matter inside the universe has been defined. The type of matter is described by an equation of state which, in general, relates the pressure to the volume.

Example 12. The simplest possibility is that a matter dominated universe, that has an equation of state:

$$P = 0 \tag{5.43}$$

This is sometimes referred to as *dust* and could possibly be a candidate for cold dark matter. Putting this into the continuity equation gives:

$$\dot{\rho}(t) = -\frac{3\dot{a}}{a}\rho \tag{5.44}$$

which has a solution:

$$\rho = \frac{C_m}{a^3} \tag{5.45}$$

C_m is some constant related to the mass of the particles. This makes sense as for a certain amount of particles in a volume, the density decreases as the volume increases.

Example 13. A universe dominated by radiation, i.e photons etc, has and equation of state:

$$\dot{\rho} = -4\frac{\dot{a}}{a}\rho \tag{5.46}$$

Putting this into the continuity equation:

$$\frac{C_r}{a^4} = \rho \tag{5.47}$$

This has an extra $\frac{1}{a}$ factor compared to the matter particles, as the energy of photons is:

$$\begin{aligned} E_\gamma &= h\nu \\ &= \frac{hc}{\lambda} \\ &\propto \tfrac{1}{\lambda} \equiv \tfrac{1}{a} \end{aligned} \tag{5.48}$$

i.e the wavelength of the photons is simply stretched out by the expansion of the universe.

EXAMPLE 14. A universe dominated by the cosmological constant has the equation of state:

$$P = -\rho \tag{5.49}$$

Therefore the continuity equation gives:

$$\dot{\rho} = 0 \tag{5.50}$$

So the density of the dark energy is constant:

$$\rho_\Lambda = C_\Lambda \tag{5.51}$$

This is extremely strange, as the energy from this type of matter will just grow with the universe and yet it is still compatible with the conservation of the stress energy tensor with respect to the covariant derivative[2].

In flat space, the Friedmann equation reduces to:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G C_\Lambda}{3} \tag{5.52}$$

This gives the scale factor evolution:

$$a \propto e^{H_\Lambda t} \tag{5.53}$$

where:

$$H_\Lambda^2 \equiv \frac{8\pi G}{3} C_\Lambda \tag{5.54}$$

So we see that the universe expands exponentially. As a side note; it is now understood that the dark energy receives many contributions, such as energy from vacuum fluctuations in space from all fields, energy from gluon condensates, the Higgs field with potential $V(H)$, etc.

EXAMPLE 15. Now lets consider a more general solution which contains ordinary matter, radiation and a cosmological constant in the universe. The Friedmann equation is:

$$\begin{aligned} \dot{a}^2 - \frac{8\pi G}{3}(\rho_m + \rho_r + \rho_\Lambda)C^2 &= -k \\ \dot{a}^2 - \frac{8\pi G}{3}\left(\frac{C_m}{a} + \frac{C_r}{a} + \rho_\Lambda a^2\right) &= -k \end{aligned} \tag{5.55}$$

Compare this to:

$$E = \frac{mv^2}{2} + V(x) \tag{5.56}$$

If $m = 2$, $E = -k$, then Eq 5.55 is just the equation of a particle moving in an effective potential:

$$V_{eff}(a) = -\frac{8\pi G}{3}\left(\frac{C_m}{a} + \frac{C_n}{a^2} + C_\Lambda a^2\right) \tag{5.57}$$

This type of universe has a few properties that are listed below.

- Positive Λ leads to de-Sitter space-time and negative Λ leads to anti-de Sitter space-time.
- The energy of a particle is constant. The energy, as stated above, is equal to $-k$. Hence for different values of k, one gets different types of universes.

[2]To me, this phenomena seems to show that the universe is truly mathematical in its nature. The fact that Einstein's field equations allow for this arbitrary constant and nature obeys this equation and also has this constant is an example that nature really does follow the rules of mathematics. At no point in deriving the Einstein equations was it considered that there should be a parameter, Λ, that is also satisfied by the field equations.

- $k = 0$. In this case, the scale factor starts out at 0 from the big bang and then simply goes over the potential if Λ is positive. If Λ is negative, then there will come a point, where the $V_{eff} = 0$ line, will intercept the potential of the Λ line and hence the potential energy will become greater than the kinetic energy and the universe will collapse in a so called *big crunch*.
- $k > 0$. The universe is a sphere and will have a negative energy (as $E = -k$). Thus the universe will expand upto a point, then it won't have enough energy to climb up the potential energy and hence roll back down to $a = 0$, again the big crunch scenario.

  Another possibility is that the universe came in from infinity (i.e a is contracting in on itself) and will again be unable to climb up the potential barrier and hence bounce back to infinity, a so called "bouncing universe".
- $k < 0$. k is negative means the potential energy is positive and the solution is similar to the $k = 0$ universe, except the evolution will be faster due to the longer kinetic energy.

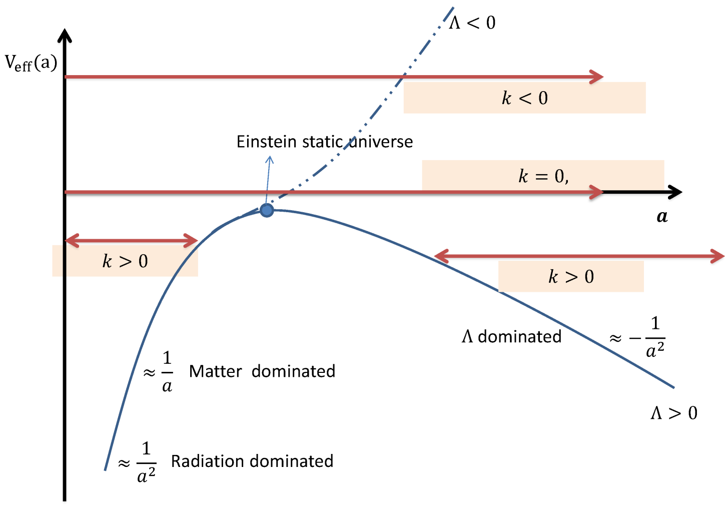


FIGURE 17. Diagram showing the evolution of a FRW universe with matter, radiation and a cosmological constant

1.4. Einstein static universe. Einstein's first model of the universe was this static model as shown by the point on the Figure 17. For a static universe one would let:

$$\dot{a} = \ddot{a} = 0 \tag{5.58}$$

and from the Raychaudhri equation:

$$\rho + 3P = 0 \tag{5.59}$$

and if the universe is dominated by matter and Λ, this is:

$$\rho_\Lambda + \rho_m + 3(P_\Lambda + P_m) = 0 \tag{5.60}$$

But $P_\Lambda = -\rho_\Lambda$ and $P_m = 0$, therefore:

$$\rho_m = 2\rho_\Lambda \tag{5.61}$$

for Einstein's static universe. From Friedmann's equation:

$$k = 4\pi G \rho_m a^2 \tag{5.62}$$

and recall that k is defined in terms of the radius of curvature and the scale factor, a:

$$k = \frac{a^2}{R_c^2} \tag{5.63}$$

Therefore:

$$R_c = \frac{1}{\sqrt{4\pi G\rho_0}} \tag{5.64}$$

for the Einstein static universe. So the universe is finely balanced between matter pulling it in and Λ pushing it out. But obviously this type of universe is unstable and was disregarded very quickly.

1.5. Comparing the models with observations. The first step is to re-write the Friedmann equation by dividing by a^2:

$$\left(\underbrace{\frac{\dot{a}}{a}}_{\equiv H}\right)^2 - \frac{8\pi G}{3}(\rho_m + \rho_r + \rho_\Lambda) + \frac{k}{a^2} = 0 \tag{5.65}$$

where H is the Hubble parameter. The different matter densities are re-defined by a matter density parameter as:

$$\frac{8\pi G\rho_m}{H^2} \equiv \Omega_m \tag{5.66}$$

and the same for $\rho_r, \rho_\Lambda, \rho_k$ etc. Thus Eq 5.65 can be re-written as:

$$1 - (\Omega_m + \Omega_r + \Omega_\Lambda + \Omega_k) = 0 \tag{5.67}$$

Observations indicate:

$$\begin{aligned} \Omega_m &= 0.28 \\ \Omega_r &= 0.0003 \\ \Omega_\Lambda &\approx 0.72 \\ \Omega_k &= 0 \end{aligned} \tag{5.68}$$

Ω_m includes cold dark matter, baryogenic matter and massive neutrinos.

1.6. Age of the universe.

Example 16. For a flat universe that only contains matter, the Friedmann equation becomes:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\frac{C_m}{a^3} \tag{5.69}$$

From Eq 5.45:

$$a\dot{a}^2 = c(constant) \tag{5.70}$$

Which has the solution:

$$a \propto t^{\frac{2}{3}} \tag{5.71}$$

Thus the Hubble parameter is:

$$H = \frac{\dot{a}}{a} = \frac{2}{3t} \tag{5.72}$$

So the age of the universe can be found by substituting the current value of the Hubble parameter into this equation:

$$t_{age} = \frac{2}{3}H^{-1} \tag{5.73}$$

This is the age of the matter dominated universe and has a value of $\approx 10^{10}$ years. However, stars have been observed that are 12 billion years old! which immediately implies that the universe cannot contain matter by itself.

EXAMPLE 17. For a universe with matter and Λ, the Friedmann equation gives:

$$\begin{aligned} t &= \int_0^{a_{\text{today}}} \frac{da}{\sqrt{\frac{8\pi G}{3}\left(\frac{C_m}{a} + C_\Lambda a^2\right)}} \\ &= \frac{1}{H}\int_0^1 \frac{da}{\sqrt{\frac{\Omega_m}{a} + \Omega_\Lambda a^2}} \end{aligned} \tag{5.74}$$

where a_{today} has been scaled to one. Now define:

$$a \equiv x^{\frac{2}{3}} \tag{5.75}$$

Thus Eq 5.74 becomes:

$$\begin{aligned} t &= \frac{2}{3H}\int_0^1 \frac{dx}{(\Omega_m + \Omega_\Lambda x^2)^{\frac{1}{2}}} \\ &= \frac{2}{3}\frac{1}{\Omega_\Lambda^{\frac{1}{2}}}\sinh^{-1}\left(\frac{\Omega_\Lambda}{\Omega_m}\right)^{\frac{1}{2}} \approx \frac{1}{H} \end{aligned} \tag{5.76}$$

for observed values of Ω_Λ and Ω_m, this turns out to be about 14 billion years. One can also include the radiation term:

$$\frac{1}{H}\int_0^1 \frac{da}{\sqrt{\frac{\Omega_m}{a} + \Omega_\Lambda a^2 + \frac{\Omega_r}{a^2}}} \tag{5.77}$$

But note that radiation only dominates at early times and since it is in the denominator, the integrand will not be affected too much by it and hence in general the radiation term can be ignored.

2. Schwarzschild solution

Black holes are the most interesting and paradoxical objects in physics. They continue to be of enormous interest, both theoretically, because they raise a lot of paradoxes, and observationally, because in the last decade or so, its become clear that black holes are everywhere in the universe, not just in the center of galaxies.

In many ways, black holes are the simplest solutions to Einstein's field equations. At first, spherically symmetric solutions are studied. Realistic black holes are not spherically symmetric as they are rotating and have an angular momentum, meaning they are *axially* symmetric but not spherically symmetric. For these calculations, the cosmological constant is neglected, and the field equations are considered in a vacuum:

$$G_{\alpha\beta} = 0 \tag{5.78}$$

Since the space-time is assumed to be spherically symmetric, the solutions to the field equations i.e the metric, is also spherically symmetric:

$$\vec{x}' = O\vec{x} \tag{5.79}$$

where:

$$O^T O = 1 \tag{5.80}$$

O is a rotation matrix:

$$O \in \{O(3)\} \tag{5.81}$$

A rotation matrix is generated by 3 generators, which represent rotations around the $\vec{x}, \vec{y}, \vec{z}$ axis. This means the line element should be built from only rotationally invariant quantities, which in this case are:

$$\text{Lorentz invariant quantities} = \{dt, t, r = \sqrt{x^2}, \vec{x}\cdot d\vec{x} = rdr, d\vec{x}^2 = dr^2 + r^2(\theta^2 + \sin^2\theta d\phi^2)\} \tag{5.82}$$

where $\vec{x}$ is:

$$\vec{x} = \begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix} \tag{5.83}$$

The most general line element formed from these parameters is:

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -C(r,t)dt^2 + D(r,t)dr^2 + 2E(r,t)drdt + F(r,t)r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{5.84}$$

This equation can be simplified by a choice of coordinates. First define:

$$F(r,t)r^2 \equiv r'^2 \tag{5.85}$$

This would mean that the area of a surface that is parametrised by θ and ϕ is just $4\pi r'^2$ and then we can just re-label r' to r:

$$ds^2 = -C(r,t)dt^2 + D(r,t)dr^2 + 2E(r,t)drdt + r^2 d\Omega_2^2 \tag{5.86}$$

The next simplification is made to remove the cross terms in the line element i.e E(r,t), as it is much easier to deal with a diagonal metric. To do this, we re-define the time as:

$$dt' = \eta(r,t)\left(C(r,t)dt - E(r,t)dr\right) \tag{5.87}$$

η is called the integration factor and is designed to ensure that the coordinate transformations exists with dt', i.e we expect:

$$dt' = \frac{\partial t'}{\partial t}dt + \frac{\partial t'}{\partial r}dr \tag{5.88}$$

Thus by comparing Eq 5.87 to 5.88, we get:

$$\begin{aligned} \frac{\partial t'}{\partial t} &= \eta(r,t)C(r,t) \\ \frac{\partial t'}{\partial r} &= -\eta(r,t)E(r,t) \end{aligned} \tag{5.89}$$

The trick now is to use the fact that partial derivatives commute:

$$\frac{\partial^2 t'}{\partial r\partial t} = \frac{\partial^2 t'}{\partial t\partial r} \tag{5.90}$$

Substitute Eq 5.89 for the partial derivatives:

$$\frac{\partial}{\partial r}\left(\eta(r,t)C(r,t)\right) = -\frac{\partial}{\partial t}\left(\eta(r,t)E(r,t)\right) \tag{5.91}$$

These are simply the conditions for t' to exist. Given $C(r,t)$ and $E(r,t)$, this is a differential equation for η. In fact it can be viewed as the evolution equation for η. So for any $\eta(r,t_0)$, the equation determines $\eta(r,t)$, for $t > t_0$ (assuming $E \neq 0$). All of this has been done to show that t' must exist, and now we can put this into the line element:

Claim 14. The metric with t' can be written as:

$$ds^2 = -\frac{\partial t'^2}{\eta^2 C} + (D + C^{-1}E^2)dr^2 + r^2 d\Omega_2^2 \tag{5.92}$$

Proof 14. Substitute Eq 5.91 into 5.92:

$$\begin{aligned} ds^2 &= -\frac{1}{\eta^2 C}\eta^2(Cdt - Edr)^2 + \left(D + \frac{E^2}{C}\right)dr^2 + r^2 d\Omega_2^2 \\ &= -\frac{1}{C}(C^2dt^2 - 2ECdtdt + E^2dr^2) + Ddr^2 + \frac{E^2dr^2}{C} + r^2 d\Omega_2^2 \\ &= -cdt^2 + 2Edtdr - \frac{E^2dr^2}{C} + Ddr^2 + \frac{E^2dr^2}{C} + r^2 d\Omega_2^2 \\ &= -cdt^2 + 2Edtdr + Ddr^2 + r^2 d\Omega_2^2 \\ &= \text{Eq } 5.86 \end{aligned} \tag{5.93}$$

To further simplify, define:

$$\begin{aligned} ds^2 &\equiv -B(r,t)dt^2 + A(r,t)dr^2 + r^2 d\Omega_2^2 \\ B(r,t) &\equiv \frac{1}{\eta^2 C} \\ A(r,t) &\equiv D + C^{-1}E^2 \end{aligned} \tag{5.94}$$

This is the most general spherically symmetric solution to the Einstein field equations, in coordinates that make the metric diagonal. Now one can calculate the Einstein field equations:

$$G_{\alpha\beta} = 0 \Rightarrow R_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}R = 0 \tag{5.95}$$

Taking the trace of this equation:

$$\begin{aligned} g^{\alpha\beta}R_{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}g_{\alpha\beta}R &= 0 \\ R - 2R &= 0 \\ R &= 0 \end{aligned} \tag{5.96}$$

Thus the Einstein equations in vacuum imply that:

$$R_{\alpha\beta} = R = 0 \tag{5.97}$$

Computing the Christoffel symbols with $x^0 = t, x^1 = r, x^2 = \theta, x^3 = \phi$:

$$\begin{aligned} &\Gamma^0_{00} = \frac{1}{2}\frac{\dot{B}}{B}, \Gamma^0_{01} = \Gamma^0_{10} = \frac{1}{2}\frac{B'}{B}, \Gamma^0_{11} = \frac{1}{2}\frac{\dot{A}}{B} \\ &\Gamma^1_{10} = \Gamma^1_{01} = \frac{\dot{A}}{2A}, \Gamma^1_{00} = \frac{B'}{2A}, \Gamma^1_{11} = \frac{A'}{2A}, \Gamma^1_{22} = -\frac{r}{A}, \Gamma^1_{33} = -r\frac{\sin^2\theta}{A} \\ &\Gamma^2_{21} = \Gamma^2_{12} = \frac{1}{r}, \Gamma^2_{33} - \sin\theta\cos\theta, \Gamma^3_{31} = \Gamma^3_{13} = \frac{1}{r}, \Gamma^3_{32} = \Gamma^3_{23} = \cot\theta \end{aligned} \tag{5.98}$$

The ' represents derivatives w.r.t r and the $\dot{}$ represents a derivative w.r.t t. The Ricci tensor is:

$$R_{\mu\kappa} = R^\lambda_{\mu\lambda\kappa} = -(\Gamma^\lambda_{\mu\nu,\kappa} - \Gamma^\lambda_{\mu\kappa,\lambda} + \Gamma^\eta_{\mu\nu}\Gamma^\lambda_{\kappa\eta} - \Gamma^\eta_{\mu\kappa}\Gamma^\lambda_{\eta\lambda}) \tag{5.99}$$

Computing each of the components:

$$
\begin{aligned}
R_{00} &= \frac{B''}{2A} - \frac{B'A'}{4A^2} + \frac{B'}{Ar} - \frac{B'^2}{4AB} - \frac{\ddot{A}}{2A} + \frac{\dot{A}^2}{4A^2} + \frac{\dot{B}\dot{A}}{4AB} \\
R_{11} &= -\frac{B''}{2B} + \frac{B'^2}{4B^2} + \frac{A'B'}{4AB} + \frac{A'}{Ar} - \frac{\ddot{A}}{2B} + \frac{\dot{A}\dot{B}}{4B^2} + \frac{\dot{A}^2}{4AB} \\
R_{10} &= \frac{\dot{A}}{Ar} \\
R_{22} &= 1 - \frac{1}{A} + \frac{rA'}{2A^2} - \frac{rB'}{2AB}
\end{aligned} \tag{5.100}
$$

There is another component, R_{33} but that is related to the R_{22} component:

$$R_{33} = \sin^2\theta R_{22} \tag{5.101}$$

All other components of the Ricci tensor are 0 by spherical symmetry. Now one has to solve these 4 coupled equations for A and B and these turn out to be surprisingly simple. First lets look at the R_{10} equation:

$$\frac{\partial A(r,t)}{\partial t} = 0 \tag{5.102}$$

i.e A is independent of t, $A(r) = A$. This also means that all $\dot{A}$ terms will be zero. To make further progress, compute:

$$\frac{R_{00}}{B} + \frac{R_{11}}{A} \tag{5.103}$$

The reason is to get rid of the second derivatives:

$$
\begin{aligned}
\frac{R_{00}}{B} + \frac{R_{11}}{A} &= \frac{A'}{A^2} + \frac{B'}{ABr} = 0 \\
&= (AB') = 0
\end{aligned} \tag{5.104}
$$

Therefore:

$$AB = f(t) \tag{5.105}$$

Now we can impose some boundary conditions. The space-time is required to be asymptotically Minkowski at large r. This implies that $A \to 1$ and $B \to 1$ as $r \to \infty$. But now it is obvious that under these BC's $f(t)$ must be 1 as it is independent of r. Therefore:

$$B \equiv \frac{1}{A} \tag{5.106}$$

But $A = A(r)$ therefore $B = B(r)$, i.e R must be time-independent. The only equation left to use now is R_{22}:

$$R_{22} = 1 - B - rB' = 0 \tag{5.107}$$

This implies:

$$(rB)' = 1 \tag{5.108}$$

Therefore:

$$rB = r + C(constant) \tag{5.109}$$

Which can be re-arranged for B:

$$B = 1 + \frac{C}{r} \tag{5.110}$$

Substitute Eq 5.110 into 5.106:

$$A = \frac{1}{1 + \frac{C}{r}} \tag{5.111}$$

The fact that all the time dependence disappears is a consequence of "Birkhoff's theorem".

Consider a mass, M, at large r, the gravitational potential must take on the Newtonian form:

$$\phi = -\frac{GM}{r} \tag{5.112}$$

and recall from the correspondence with Newtonian gravity that h_{00} (the deviation from flat space), is:

$$h_{00} = -\frac{2\phi}{c^2} \tag{5.113}$$

This implies that:

$$g_{00} \approx \eta_{00} + h_{00} = -1 - \frac{2\phi}{c^2} \tag{5.114}$$

This fixes the constant, C:

$$\begin{aligned} B &= 1 - \frac{2GM}{rc^2} \\ A &= B^{-1} \end{aligned} \tag{5.115}$$

So the line element becomes:

$$ds^2 = -\left(1 - \frac{2MG}{rc^2}\right) dt^2 + \frac{dr^2}{\left(1 - \frac{2GM}{rc^2}\right)} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{5.116}$$

This is called the Schwarzschild metric and was discovered in 1915. The Schwarzschild metric appears to be singular at:

$$r \equiv r_s = \frac{2GM}{c^2} \tag{5.117}$$

and this is known as the *Schwarzschild radius*. In the 1960's it was realised that this is actually just a coordinate singularity. In other words, it is possible to remove this singularity by a change of coordinates. An easy explanation to the apparent singularity is to consider the mass being considered to be an extended object, like the sun. The value of r_s is approximately 3 km and this radius is obviously inside the sun. The solution obtained above is only valid in the absence of matter, i.e outside the surface of the sun. Thus any results from these solutions cannot be extrapolated to results in the sun. The field equations need to be solved separately inside the sun and these will give solutions that do not contain this singularity. But if the sun collapses to a radius of 3 km (assuming no mass loss), then there would be a singularity at r_s.

2.1. Particle trajectories. Now we will think about falling into a black hole and thinking about what will be observed. Let's begin with the action, but if there are symmetries that leave the equations of motion unchanged, then these symmetries can be used in the action to simplify it. The action was previously:

$$S = -m \int d\tau \tag{5.118}$$

where $cd\tau = -g_{\mu\nu}dx^\mu dx^\nu$:

$$S = -mc^2 \int \sqrt{-g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu} d\lambda \tag{5.119}$$

where the $\dot{}$ represents a derivative w.r.t λ. This action is useful as it is invariant under re-parametrising . However, the square root can prove to be a problem sometime, hence one uses a different action.

Claim 15. The action in Eq 5.119, can be written as:

$$S = S(x^\mu(\lambda), e)\frac{m}{2}\int \left(g_{\mu\nu}\frac{\dot{x}^\mu\dot{x}^\nu}{e} - c^2 e\right)d\lambda \tag{5.120}$$

where e is called "einbein", and is the square root of the 1-D metric on the world line:

$$\begin{aligned}\lambda &\to \tilde{\lambda}(\lambda)\\ e &\to \frac{d\lambda}{d\tilde{\lambda}}e\end{aligned} \tag{5.121}$$

therefore $ed\lambda$ is invariant under re-parametrization of the world line.

Proof 15. The Euler-Lagrange equation for e gives:

$$\frac{-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}{e^2} - c^2 = 0 \tag{5.122}$$

This gives:

$$e = \frac{1}{c}\sqrt{-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu} \tag{5.123}$$

substitute back into the action:

$$S = -mc\int \sqrt{-g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu}d\lambda \tag{5.124}$$

which retains Eq 5.119.

The improved action gives both the constraint on the 4-velocity and the geodesic equation. From Eq 5.122:

$$g_{\mu\nu}U^\mu U^\nu = -c^2 \tag{5.125}$$

where:

$$U^\mu = \frac{\dot{x}^\mu}{e} \tag{5.126}$$

Thus the Euler-Lagrange (EL) equation for e gives the constraint on the four velocity and similarly the EL equation for x^μ gives the geodesic equation:

$$\frac{\partial^2 x^\mu}{\partial\tau} + \Gamma^\mu_{\nu\lambda}\frac{\partial x^\nu}{\partial\tau}\frac{\partial x^\alpha}{\partial\tau} = 0 \tag{5.127}$$

where:

$$d\tau = ed\lambda \tag{5.128}$$

For a photon, $m = 0$, thus the action is 0 and it cannot be used to get any meaningful results. Instead, for photons, a different action is used:

$$S = \frac{C}{2}\int \frac{\dot{x}^\mu\dot{x}^\nu}{e}g_{\mu\nu}d\lambda \tag{5.129}$$

where C is identified with the magnitude of the momentum of the photon:

$$p_\mu = \frac{\partial L}{\partial\dot{x}^\mu} = \frac{Cg_{\mu\nu}\dot{x}}{e} \tag{5.130}$$

C gives a representation of what the energy of the photon is. The EL equation for e is:

$$\dot{x}^\mu \dot{x}^\nu g_{\mu\nu} = 0 \tag{5.131}$$

i.e the photons follow a null trajectory in space-time (as expected).

From Noether's theorem, it has already been shown that if there is translational symmetry in $g_{\mu\nu}$ w.r.t a particular coordinate, its conjugate momentum will be conserved. For example, for the Schwarzschild metric, the metric is independent of t and ϕ, thus the corresponding conjugate momentum for these objects is conserved.

2.2. Action of the Schwarzschild metric. It is convenient to work in a gauge (coordinate choice):

$$e \equiv 1 \tag{5.132}$$

Which means that from:

$$d\tau = ed \Rightarrow d\tau = d\lambda \tag{5.133}$$

So the geodesics are being parametrised by the proper time. The action is then:

$$S = \frac{m}{2}\int d\tau \left(-\left(1 - 2\frac{GM}{rc^2}c^2\right)\right)\dot{t}^2 + \frac{\dot{r}^2}{\left(1 - \frac{2GM}{rc^2}\right)} + r^2(\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2) \tag{5.134}$$

where the , represents a derivative w.r.t τ now. Now one can calculate the equations of motion:

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{x}^\mu}\right) = \frac{\partial L}{\partial x^\mu} \tag{5.135}$$

For θ, the equations of motion are:

$$\frac{d}{d\tau}(r^2\dot{\theta}) = r^2\sin\theta\cos\theta\dot{\phi}^2 \tag{5.136}$$

Note that we are studying the metric of a spherically symmetric black-hole (or any other massive object), thus we can choose to consider $\theta = \frac{\pi}{2}$ to simplfy the equations, without any loss of generality. The particle will be moving on an equatorial orbit, and Eq 5.136 gives:

$$\frac{d}{d\tau}(r\theta) = 0 \tag{5.137}$$

Now the EL equation for ϕ:

$$\begin{aligned}\frac{d}{d\tau}(r^2\sin^2\theta\dot{\phi}) &= 0\\ r^2\dot{\phi} &= L_z(constant)\end{aligned} \tag{5.138}$$

where the m has been dropped as it is simply a scaling factor. The EL equation for t:

$$\frac{d}{d\tau}\left(c\dot{t}\left(1 - \frac{2GM}{rc^2}\right)\right) = 0 \tag{5.139}$$

therefore:

$$c\dot{t}\left(1 - \frac{2GM}{rc^2}\right) \equiv E(constant) \tag{5.140}$$

This is interpreted as the energy as it comes from the symmetry under time translations. To proceed, use the constraint:

$$g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = -c^2 \tag{5.141}$$

Which gives:

$$-\left(1-\frac{2GM}{rc^2}\right)E^2+\left(1-\frac{2GM}{rc^2}\right)^{-1}\dot{r}^2+\frac{L_z^2}{r^2}=-c^2 \tag{5.142}$$

This is the equation for a massive particle; if there was a photon, the only difference would be that the R.H.S would be 0 as the photon has 0 proper time. Eq 5.142 can be re-written as:

$$\dot{r}^2+V(r)_m^{(eff)}=E^2 \tag{5.143}$$

where:

$$V(r)_m^{(eff)}\equiv c^2\left(1+\frac{L_2^2}{r^2c^2}\right)\left(1-\frac{2GM}{rc^2}\right) \tag{5.144}$$

Thus the material particle moves with a kinetic energy in an effective potential, $V(r)_m^{(eff)}$. This potential shows that the singularity at the Schwarzschild radius is not really a singularity. As at $r=r_s$ the effective potential is just 0 i.e it does not diverge. The solution to Eq 5.143 can be found by looking for $r(\phi)$ as supposed to $r(t)$. In other words, the radius is now parametrised by ϕ:

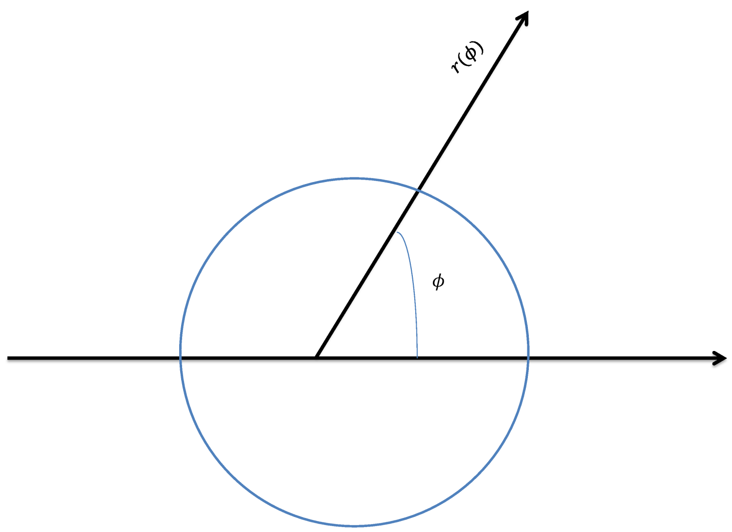


FIGURE 18. Parametrising r by ϕ

So for a given value of ϕ, one obtains a value for $r(\phi)$. We have:

$$L_z=r^2\dot{\phi} \tag{5.145}$$

Therefore:

$$\dot{r}^2=r'^2\dot{\phi}^2=\frac{r'^2L_z^2}{r^4} \tag{5.146}$$

where the ' represents a derivative w.r.t ϕ. Now define:

$$u\equiv\frac{1}{r} \tag{5.147}$$

It turns out that the Newtonian problem of $\frac{1}{r}$ potentials become a simple harmonic oscillation (SHO) under this transformation. Therefore one gets:

$$\dot{r}^2-L_z^2u'^2 \tag{5.148}$$

Putting these ingredients into Eq 5.143:

$$u'^2 = \underbrace{\frac{E^2 - c^2}{L_z^2}}_{T_1} - \underbrace{u^2}_{T_2} + \underbrace{\frac{2GMu}{L_z^2}}_{T_3} + \underbrace{\frac{2GMu^3}{c^2}}_{T_4} \tag{5.149}$$

In this form, T_2 and T_3 are independent of C, and are the usual Newtonian terms. T_1 is just a constant and T_4 is the relativistic correction.

To solve this equation, differentiate w.r.t ϕ, to eliminate the constant term T_1:

$$2u'u'' = -2u'u + \frac{2GMu'}{L_z^2} + \frac{6GMu^2u'}{c^2} \tag{5.150}$$

Canceling u':

$$u'' = \underbrace{-u}_{T_\alpha} + \underbrace{\frac{GM}{L_z^2}}_{T_\beta} + \underbrace{\frac{3GMu^2}{c^2}}_{T_\gamma} \tag{5.151}$$

This equation can be solved by an expansion in c^2. When $v << c^2$, the T_γ term can be ignored:

$$u'' = -u + \frac{GM}{L_z^2} \tag{5.152}$$

The idea now is to calculate the solution for this linear differential equation and compute the T_γ as a perturbation, using perturbation theory. The total solution will then be:

$$u = u_0 + \frac{u_1}{c^2} + ... \tag{5.153}$$

where u_0 is the solution to Eq 5.152 and u_1 will be the perturbation term. u_0 is simply the solution to a Harmonic oscillator, displaced by $\frac{GM}{L_z^2}$. The minimum of the harmonic oscillator is when:

$$u = \frac{GM}{L_z^2} \tag{5.154}$$

This is the in-homogenous solution. The homogenous solution δu, needs to be added to this, which is:

$$\delta U = A_1 \cos\phi \tag{5.155}$$

where A_1 is an arbitrary constant. Combining the two terms, gives the general solution:

$$u = \frac{GM}{L_z^2} + A_1 \cos\phi \tag{5.156}$$

which can be re-defined as:

$$u \equiv \frac{GM}{L_z^2}(1 + e\cos\phi) \tag{5.157}$$

where e is defined as the ellipcity of the orbit. It parametrizes how much the orbit is deviating from a circular orbit. Since:

$$r = \frac{1}{u} \equiv \frac{r_0}{1 + e\cos\phi} \tag{5.158}$$

this is an equation of an ellipse, with r_0 as the average radius. Now lets calculate the perturbing correction to the orbit. Substitute Eq 5.153 into 5.151 and equate the coefficients of $\frac{1}{c^2}$:

$$u_1'' = -u_1 + \frac{3GM}{c^2}\left(u_0 + \frac{1}{c^2}u_1\right) \equiv -u_1 + \frac{3GM}{c^2}u \tag{5.159}$$

Substitute Eq 5.157 into 5.159:

$$\begin{aligned} u_1'' &= -u_1 + \frac{3GM}{c^2}\frac{(GM)^2}{L_z^4}(1+e\cos\phi)^2 \\ &= -u_1 + \frac{3GM}{c^2}\frac{(GM)^2}{L_z^4}(1+2e\cos\phi+e^2\cos^2\phi) \end{aligned} \tag{5.160}$$

where the $\frac{1}{c^4}$ terms have been neglected. This equation looks like a harmonic oscillator equation for u, which is driven by a force in the second term. This force is oscillating (with a constant force aswell but that is not important as it will displace an object slightly, but the displacement will not grow).

There is a term with $e^2\cos^2\phi$, which has twice the frequency compared to the oscillation of u, and the term with $e\cos\phi$, which has the *same* frequency will cause a resonance affect in the oscillator and will cause a growth in the displacement. Einstein realised that this relativistic correction to the Newtonian orbit, actually builds over every orbit, even though the term is considered a perturbation.

This lead to one of the most dramatic early test of GR. Let's re-write the solution with just the resonance term, as that is the one that will cause the major deviation:

$$u'' = -u_1 + \epsilon\cos\phi \tag{5.161}$$

where:

$$\epsilon \equiv \frac{2(GM)^2 e}{L_z^2 c^2} \tag{5.162}$$

The solution to this equation is:

$$u_1 = \frac{\epsilon}{2}\phi\sin\phi \tag{5.163}$$

So it is clear that this oscillation has an amplitude that grows with ϕ. Therefore the total solution is:

$$u \approx \frac{GM}{L_z^2}\left(1+e\cos\phi+3\frac{(GM)^2}{L_z^2c^2}\phi\sin\phi\right) \tag{5.164}$$

To interpret this growing perturbation, imagine that the coefficients of $\sin\phi$ are very small and hence one can use the trig identity:

$$\cos A - B = \cos A\cos B + \sin A\sin B \qquad B << 1 \Rightarrow \cos B \approx 1, \sin B \approx B \tag{5.165}$$

Therefore the solution can be re-written as:

$$u \approx \frac{GM}{L_z^2}\left(1+e\cos\left(\phi\left(1-\frac{3(GM)^2}{L_z^2c^2}\right)\right)\right) \tag{5.166}$$

So we see that the offset of this resonance is to actually change the orbit of the object around the black hole.

As an example, lets work out the precession (change in ϕ of maximum radius per orbit):

$$\Delta\phi = G\pi\left(\frac{GM}{L_z}\right)^2\frac{1}{c^2} = \frac{6\pi GM}{r_0c^2} \tag{5.167}$$

where r_0 is the average radius of the orbit and:

$$\frac{GM}{r_0} \approx v_{\text{escape}} \tag{5.168}$$

Therefore:

$$\Delta\phi \approx \frac{v_{\text{escape}}^2}{c^2} \tag{5.169}$$

Its largest for smallest r_0 For mercury, one gets:

$$\Delta\phi = 42.98''/\text{century} \tag{5.170}$$

The observations of the 19th century had discovered an anomaly of 43″/century. This is a remarkable prediction of general relativity and was the first confidence boost for Einstein that he was on the right track.

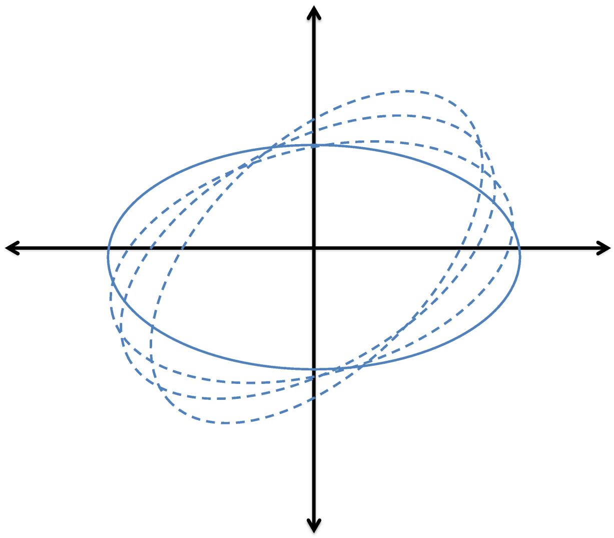


FIGURE 19. Precession of orbits due to relativistic corrections

2.3. Motion of photons. Eq 5.144 simplifies for photons to:

$$V(r) = \frac{L_z^2}{r^2}\left(1 - \frac{2GM}{rc^2}\right) \tag{5.171}$$

Let's re-write the potential in terms of the Schwarzschild radius:

$$V_\gamma(r) = \frac{L_z^2}{r^2}\left(1 - \frac{r_s}{r}\right) = L_z^2\left(\frac{1}{r^2} - \frac{r_s}{r^3}\right) \tag{5.172}$$

The maximum of the potential is:

$$\frac{\partial V_\gamma(r)}{\partial r} = 0 \Rightarrow r_{max} = \frac{3}{2}r_s \tag{5.173}$$

Therefore:

$$r_{max} = \frac{3GM}{c^2} \tag{5.174}$$

At this point the photon orbits the mass at a constant radius. Again setting:

$$u \equiv \frac{1}{r} \tag{5.175}$$

and follow the same procedure as was done for the photon, yields the equation:

$$u'' = -\underbrace{u}_{T_1} + \underbrace{\frac{3GMu^2}{c^2}}_{T_2} \tag{5.176}$$

where T_1 is the Newtonian result and T_2 is the general relativity correction. The term that is missing here, is the one that shifted the center of the harmonic oscillator. Here the geometry shows:

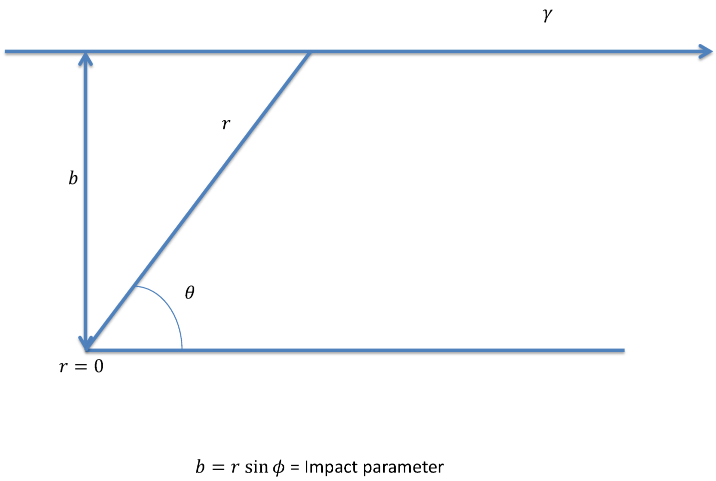


FIGURE 20. Impact parameter geometry

Therefore:

$$u_0 = \frac{1}{r} = \frac{\sin}{b} \tag{5.177}$$

This is the zeroth order solution. Putting this into Eq 5.176:

$$u_1'' + u_1 = \frac{3GM}{2b^2c^2}\sin^2\phi = \frac{3GM}{2b^2c^2}(1 - \cos 2) \tag{5.178}$$

Which has the solution:

$$u_1 = \frac{3GM}{2b^2c^2}\left(1 + \frac{1}{3}\cos 2\phi\right) + \underbrace{A\sin\phi}_{T_\alpha} \tag{5.179}$$

where A is a constant. T_α is not important as it is simply another harmonic oscillator solution and by choosing appropriate boundary/initial conditions, it can be absorbed into the $\sin\phi$ term from before. So the overall solution is:

$$u = \frac{1}{b}\sin\phi + u_1 \tag{5.180}$$

as $r \to \infty$, $u \to 0$ and thus $\phi = 0$ without u_1 term. If the u_1 terms is not neglected, then it is slightly different. Firstly, one expects $\phi << 1$, i.e weak field approximation, therefore one can define:

$$\phi \equiv \epsilon << 1 \tag{5.181}$$

So Eq 5.180 becomes:

$$u \approx \frac{\epsilon}{b} + \frac{3GM}{2b^2c^2}\left(1 + \frac{1}{3}\right) = 0 \tag{5.182}$$

Therefore:

$$\epsilon = -\frac{2GM}{c^2 b} \tag{5.183}$$

This is the deflection in ϕ. In other words, if the photon is followed to $+\infty$, instead of $\phi \to 0$, it becomes slightly negative, $\phi \to \epsilon$, this is the correction from general relativity. So the total deflection angle is:

$$\Delta\phi = \frac{4GM}{c^2 b} \tag{5.184}$$

if we take the values of the sun, $M = M_{sun}$ and $b = r_{sun}$, then:

$$\delta\phi_{sun} = 1.75'' \tag{5.185}$$

i.e photons just grazing the surface of the sun are deflected by this angle. This prediction of general relativity was measured by Eddington and was also proved to be correct.

2.4. Interior of black-holes. Recall the Schwarzschild metric:

$$ds^2 = -\left(1 - \frac{r_s}{r}\right)dt^2 + \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 + r^2 d\Omega_2^2 \tag{5.186}$$

$r = r_s$, the metric is singular, but this is just a coordinate singularity. For $r < r_s$, the coordinate r is *time-like*, t is *space-like*. This means that if all other coordinates are fixed and r is varied, then the line element is negative, meaning its a time-like direction. If t is varied on the other hand, then the line element increases and hence is space-like.

Consider radial, light rays, i.e null geodesics. In this case $d\Omega \equiv 0$, thus the metric simplifies to:

$$ds^2 = -\left(1 - \frac{r_s}{r}\right)dt^2 + \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 \tag{5.187}$$

For photons $ds^2 = 0$:

$$\left(1 - \frac{r_s}{r}\right)dt^2 = \left(1 - \frac{r_s}{r}\right)^{-1} dr^2 \tag{5.188}$$

Therefore:

$$dt = \pm\frac{dr}{(1 - \frac{r_s}{r})} \equiv \pm dr\left(1 + \frac{r_s}{r - r_s}\right) \tag{5.189}$$

Integrating:

$$t = \pm\left(r + r_s \ln\left(\frac{r_s}{r} - 1\right)\right) + \text{constant} \qquad (r > r_s) \tag{5.190}$$

For $r < r_s$:

$$t = \pm\left(r + r_s \ln\left(1 - \frac{r}{r_s}\right)\right) \tag{5.191}$$

As $r \to 0$:

$$t \approx \pm\left(r + r_s\left(-\frac{r}{r_s}\right)\right) = \frac{r^2}{3r_s^2} + ... \tag{5.192}$$

Therefore:

$$t = \pm\left(-\frac{r^2}{2r_s}\right) + \text{constant} \tag{5.193}$$

The signs are now flipped! the positive solution gives the in going light ray and the negative solution gives the outgoing light ray. In other words, it just shows what was previously stated, that spatial and time dimensions are flipped.

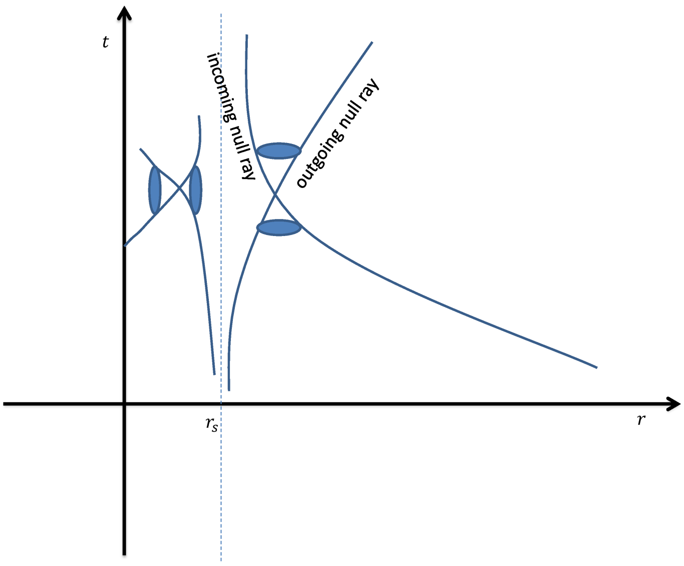


FIGURE 21. Light cones approaching the event horizon; light cones shrink in width as they approach the event horizon. After the horizon, the space and time components are flipped and hence the light cone is rotated by 90°

Now lets change the coordinate to ingoing "Eddington-Finkelstein" coordinates. The idea behind this is to say that along a null ray, the solutions are given by Eq 5.190 and 5.193, where the constant does not change. This means if one defines a coordinate in which the entire terms in these equations is the coordinate, then the null ray will be simply constant or zero:

$$\text{ingoing null ray} \equiv u \equiv \text{constant} \tag{5.194}$$

where:

$$u \equiv t + r + \ln\left|\frac{r}{r_s} - 1\right| \tag{5.195}$$

notice the modulus sign has been included so that the coordinate is defined even when $r < r_s$. The coordinate transformation is singular if $r \equiv r_s$. This is expected, as the original coordinate system had a singularity aswell, and to remove that singularity one must also include another singularity to cancel it:

$$du = dt + \frac{dr}{1 - \frac{r_s}{r}} \tag{5.196}$$

Therefore:

$$dt = du - \frac{dr}{1 - \frac{r_s}{r}} \tag{5.197}$$

Substitute this into the line element in Eq 5.187:

$$ds^2 = -\left(1 - \frac{r_s}{r}\right) du^2 + 2dudr + r^2 d\Omega_2^2 \tag{5.198}$$

If $r = r_s$:

$$ds^2 = 2dudr + r^2 d\Omega_2^2 \tag{5.199}$$

which gives the metric tensor:

$$g_{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2 \theta \end{pmatrix} \tag{5.200}$$

In these coordinates the metric is non-singular which confirms that the singularity at $r = r_s$, is a coordinate singularity. Now one can follow an ingoing light ray in these new-coordinates, which are just curves of constant u. Inside the black hole ($r < r_s$):

$$u = \text{constant} \Rightarrow t = -r - r_s \ln \left| 1 - \frac{r}{r_s} \right| + u(\equiv \text{constant}) \tag{5.201}$$

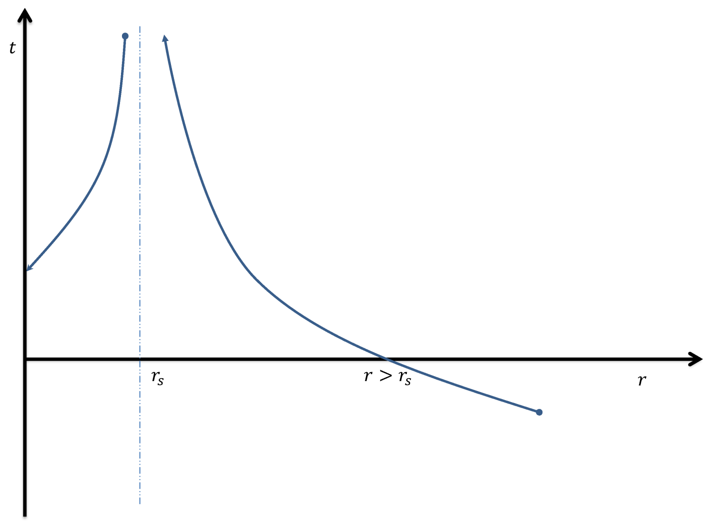


FIGURE 22. Track of ingoing light ray in $u(r)$ coordinates then converted into the singular coordinates

Notice the direction of time here. As time is increasing, r is decreasing. This is because the null ray being followed had an increasing time at $r = \infty$ (the null ray was on the future part of the light cone). Following the null ray beyond the event horizon leads to the null ray decreasing in r after r_s. Therefore one can infer that in the region of space:

$$0 < r < r_s \tag{5.202}$$

the decreasing r must lead to an increasing time. Which in essence tells us that the null ray has to hit the singularity (i.e since the time keeps moving forward, the r will continue to decrease until it hits the singularity). Therefore there seems to be an "end to time" at $r = 0$.

The trajectory of radial light rays can be obtained from the metric in Eq 5.198, where $ds = d\Omega = 0$:

$$\left(1 - \frac{r_s}{r}\right) du^2 = 2dudr \tag{5.203}$$

Which gives:

$$dr = \frac{1}{2}\left(1 - \frac{r_s}{r}\right) du \tag{5.204}$$

or:

$$du = 0 \tag{5.205}$$

Define a new time, t', such that $u = t' + r$; this implies that the ingoing null rays have:

$$r = u - t' \tag{5.206}$$

i.e they are lines at 45°. The outgoing null rays in these new coordinates can be re-written as:

$$du = \frac{2dr}{1 - \frac{r_s}{r}} = dt' + dr \tag{5.207}$$

This gives:

$$dt' = \frac{r + r_s}{r - r_s} \tag{5.208}$$

The ingoing rays are shown below:

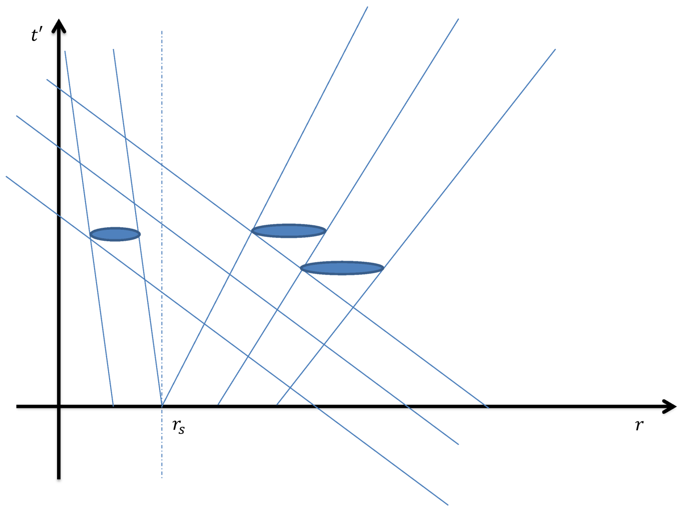


FIGURE 23. Outgoing lines $r > r_s$ increase in steepness as r decreases.

at $r = r_0$, the outgoing lines have an infinite radius, i.e they become straight lines at $r = r_s$. This affect is known as the tipping of the light cones. When $r < r_s$, the gradient is negative and the light cone is forced into the singularity. Note that the coordinates have been chosed to correspond to ingoing rays. Now lets use coordinates corresponding to outgoing rays:

$$v \equiv t - r_0 \ln\left|\frac{r}{r_s} - 1\right| \tag{5.209}$$

Which gives the metric:

$$ds^2 = -\left(1 - \frac{r_s}{r}\right) dv^2 - 2dvdr + r^2 d\Omega_2^2 \tag{5.210}$$

This is also regular at $r = r_s$. For outgoing null geodesic:

$$t = r + r_s + \ln\left|\frac{r}{r_s} - 1\right| + v(\equiv \text{constant}) \tag{5.211}$$

However, this implies the direction of increasing r,corresponds to increasing time! This is direct contradiction to the result obtained from the $u(r)$ coordinates, as there are two *regular* coordinate systems, which disagree about what is the direction for the part of the light curve. The fact that r increases with time in these coordinates inside the horizon means that a particle will fly out of the black hole!. This is not thought to be possible, as this would mean the black hole would radiate particles from the singularity and an external observed would be able to see the singularity directly.

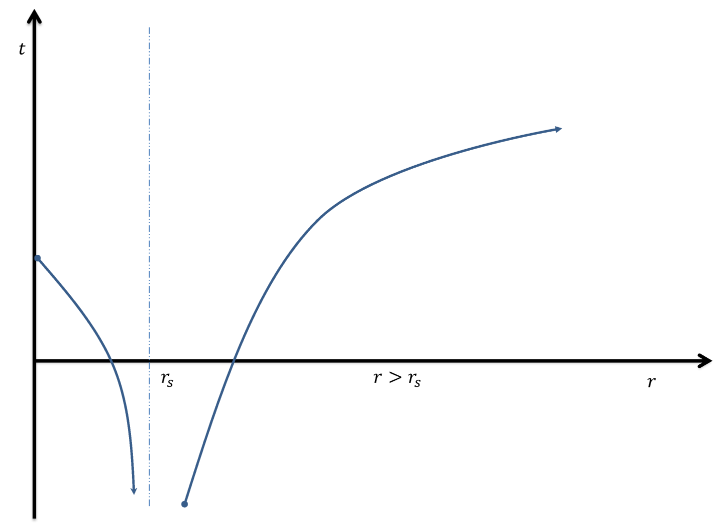


FIGURE 24. Track of ingoing light ray in $v(r)$ coordinates then converted into the singular coordinates

Actually each of these descriptions, describe a different interior. So there are actually two interiors to a black hole! Both versions of the interior have an opposite direction of time, both connected to the same exterior horizon.

2.5. Kruskal coordinates. Martin Kruskal was the first person to realise this duality of interiors of black holes (while he was an undergraduate!) and decided to combine the coordinate systems for outgoing and ingoing null geodesics:

$$r > r_s : \begin{cases} u = t + r + r_s \ln\left(\frac{r}{r_s} - 1\right) \equiv \text{constant} & \text{outgoing photons} \\ v = t - r - r_s \ln\left(\frac{r}{r_s} - 1\right) \equiv \text{constant} & \text{incoming photons} \end{cases} \tag{5.212}$$

The derivatives are:

$$\begin{aligned} du &= dt + \frac{dr}{1 - \frac{r_s}{r}} \\ dv &= dt - \frac{dr}{1 - \frac{r_s}{r}} \end{aligned} \tag{5.213}$$

The line element for Schwarzschild is:

$$ds^2 = -\left(1 - r_s r\right) du dv + r^2 d\Omega_2^2 \tag{5.214}$$

Combining the equations for u and v:

$$r + r_s \ln\left(\frac{r}{r_s} - 1\right) = \frac{u - v}{2} \tag{5.215}$$

From this:

$$\frac{r}{r_s} - 1 = e^{-\frac{r}{r_s}} e^{\frac{u-v}{2r_s}} \tag{5.216}$$

Therefore:

$$1 - \frac{r_s}{r} = \frac{r_s}{r} e^{-\frac{r}{r_s}} e^{\frac{u-v}{2r_s}} \tag{5.217}$$

Which gives the metric:

$$ds^2 = -\frac{r_s}{r} e^{-\frac{r}{r_s}} e^{\frac{u-v}{2r_s}} du dv + r^2 d\Omega_2^2 \tag{5.218}$$

This metric looks good as it is not singular anywhere except at $r = 0$, which really is a singularity and thus cannot be removed. Now we define a new set of null coordinates:

$$\begin{aligned} U &\equiv r_s e^{-\frac{u}{2r_s}} \\ V &\equiv r_s e^{-\frac{v}{2r_s}} \end{aligned} \tag{5.219}$$

So the metric is:

$$ds^2 = \frac{4r_s}{r} e^{-\frac{r}{r_s}} dU dV + r^2 d\Omega_2^2 \tag{5.220}$$

Finally, set:

$$U \equiv X + TV \equiv X - T$$

X is a new radius. The metric then becomes:

$$ds^2 = \frac{4r_3}{r} e^{-rr_s} \left(-dT^2 + dX^2\right) + r^2 d\Omega_2^2 \tag{5.220}$$

where:

$$X^2 - T^2 = UV = r_s^2 e^{\frac{U-V}{2r_s}} \tag{5.221}$$

Substitute for $e^{\frac{U-V}{2r_s}}$ from Eq 5.217:

$$X^2 - T^2 = r_s^2 \left(\frac{r}{r_s} - 1\right) e^r r_s \equiv f(X^2 - T^2) \tag{5.222}$$

This gives r as a function of $X^2 - T^2$, which is how the metric should be interpreted.

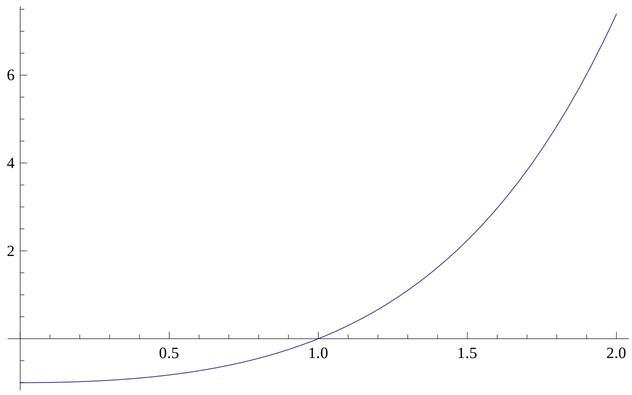


FIGURE 25. This shows $f(X^2 - T^2)$ has a 1 to 1 mapping to r.

So we see that the function $f(X^2 - T^2)$ corresponds to one value of r as long as r is positive. Even when $f(X^2 - T^2)$ is negative, there is a 1 to 1 correspondece to r except at $-r_s^2$. In other words, this function only exists on $r > 0$, for $X^2 - T^2 > -r_s^2$.

Converting back to the t coordinate:

$$t = \frac{u+v}{2} = r_s \ln\left(\frac{U}{V}\right) = r_s \ln\left|\frac{X+T}{X-T}\right| \tag{5.223}$$

At $X = T$, $t = r_s \ln \infty = \infty$, and similarly for $-r_s$; thus the straight lines at 45° correspond to lines with $t = \pm\infty$ in Figure 26

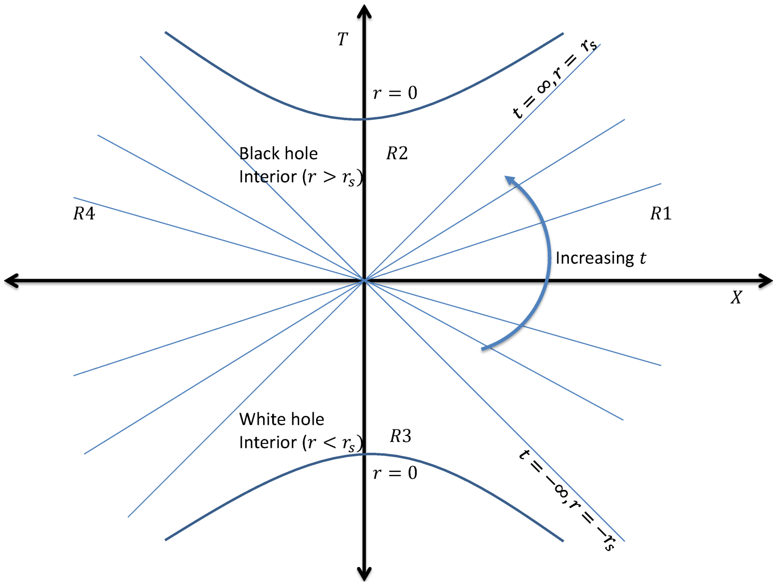


FIGURE 26. Dual interior of black holes

So we see that there are two interiors. In R2 and R3 a particle from R3 following a null geodesic can actually come out of the interior (these objects are called "white holes"), i.e out of the black hole, as r increases with t. And similarly a particle following a null geodesic in R2 will not be able to exit the interior as r decreases with t.

Kruskal found four regions to this solution. This is called the maximal geodesic extension of Schwarzschild and gives a deep insight into what the Schwarzschild solution really is. There are only two possibilities for particles; either a particle in the Schwarzschild metric will hit the singularity or it will continue to infinity. So there are no special points in space that contain infinities.

But this is not all, this solution also shows that one can go to infinity in R1 or to infinity in R3 (i.e flat space-time in two regions). So there are two asymptotic regions of space-time (or two universes!), that are connected by a so called Einstein-Rosen bridge.

The light from R4 can never reach R1 (which is where we are) as the light will travel along the line $t = \infty$ and this will asymptotically approach the singularity, hence will never be able to reach R1 and the same holds from light going from R1 to R4.

In all physical process that give rise to black holes, i.e stellar death, R3 and R4 do not exist as far as what has been observed so far. This is because at early time, there is just a star in space and the Schwarzschild metric is only valued outside the star. As the star collapses, at a certain time, it falls across its Schwarschild radius, and then all the particles enclosed in this radius must collapse to this singularity.

In fact this solution gave rise to a whole new field in general relativity of wormholes, which have the same concept of black holes in different metrics, such as the Kerr metric, connecting two causally disconnected regions of space-time with a bridge, also called a wormhole. These types of objects are thought to be unstable however, under certain conditions can be stabilized, [**5**], but have never been observed in nature. However, as they are correct solutions to Einstein's field equations, they have to be given some consideration (in fact science fiction movies and books have also given these solutions a lot of consideration!).

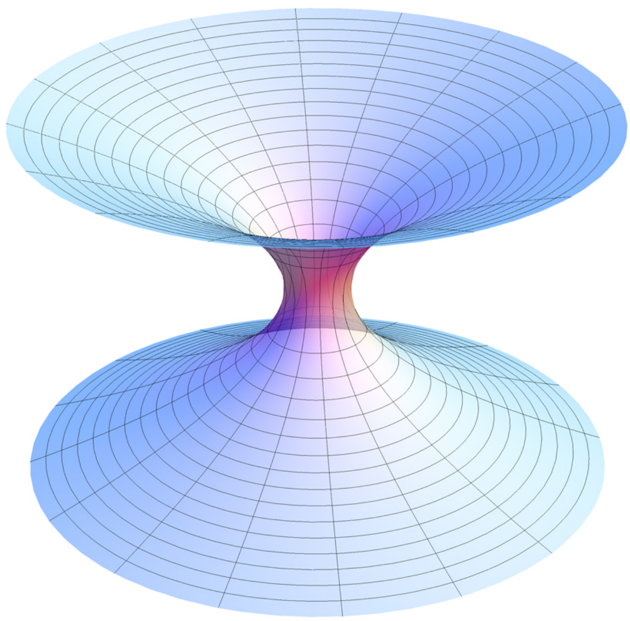


FIGURE 27. Einstein-Rosen bridge connecting two asymptotically flat space-times

3. Kerr black holes

Black holes are formed by collapsing stars and generally, stars will carry some angular momentum thus the black holes formed from these stars will have some rotation. It took a long time for the solution to Einstein's field equations, for a rotating black hole to be found. Kerr was the first one to find it in 1963 and it was a surprising result that this problem had an analytic solution. The metric corresponding to a rotating black hole is:

$$ds^2 = -\left(1 - \frac{r_s r}{\rho^2}\right) dt^2 - \frac{r_s a r \sin^2\phi}{\rho^2}\left(dt d\phi + d\phi dt\right) + \frac{\rho^2 dr^2}{\Delta} + \rho^2 d\theta^2 + \frac{\sin^2\theta}{\rho^2}\left((r^2+a^2)^2 - a^2\Delta\sin^2\theta\right) d\phi^2 \tag{5.224}$$

where:

$$\begin{aligned} \Delta^2 &\equiv r^2 - r_s r + a^2 \\ \rho^2 &\equiv r^2 + a^2\cos^2\theta \\ r_s &\equiv \frac{2GM}{c^2} \\ a &\equiv \frac{J}{Mc} \end{aligned} \tag{5.225}$$

where J is the angular momentum. These coordinates are called *Bayer-Lindquist* coordinates. The coordinates that Kerr use were actually more clever and they shall be used later, however, these coordinates allow for easy comparison to the Schwarzschild metric.

CLAIM 16. In the limit $a \to 0$, and fixed M, the Kerr metric reduces to the Schwarzschild metric

Proof 16. If $a = 0$:

$$\begin{aligned} \Delta^2 &= r^2 - r_s r \\ \rho^2 &= r^2 \end{aligned} \tag{5.226}$$

Therefore the line element is:

$$\begin{aligned} ds^2 &= -\left(1 - \frac{r_s r}{r^2}\right) dt^2 + \frac{r^2}{r^2 - r_s r} dr^2 + r^2 d\theta^2 + \frac{\sin^2\theta}{r^2}(r^2)^2 d\phi^2 \\ &= -\left(1 - \frac{r_s}{r}\right) dt^2 + \frac{r}{r - r_s} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \\ &= -\left(1 - \frac{r_s}{r}\right) dt^2 + \left(\frac{1}{1 - \frac{r_s}{r}}\right) dr^2 + r^2 d\theta_2^2 \\ &= \text{Eq 5.116} \end{aligned} \tag{5.227}$$

In instead, one takes the limit that $M \to 0$ at fixed a, one would expect the metric to be Minkowski, but the metric gives:

$$ds^2 = -dt^2 + \frac{(r^2 + a^2\cos^2\theta)}{r^2 + a^2} dr^2 + (r^2 + a^2\cos^2\theta) d\theta^2 + (r^2 + a^2)\sin^2\theta d\phi^2 \tag{5.228}$$

This does not look like flat space at first sight. To check it, one has to compute the Riemann tensor and it turns out that it is zero and therefore the space is indeed flat. Therefore there must be a coordinate system in which the metric looks Minkowski. To motivate the correct coordinate transformation, lets look at the spherical polar coordinates in Minkowski space:

$$dx^2 + dy^2 + dz^2 = dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2 \tag{5.229}$$

The spatial part of Eq 5.228 looks similar to the spherical polars in Minkowski space, in fact the radius seems to be transformed into $(r^2 + a^2)^{\frac{1}{2}}$, so lets try the coordinates:

$$\begin{aligned} x &= (r^2 + a^2)^{\frac{1}{2}} \sin\theta\cos\phi \\ y &= (r^2 + a^2)^{\frac{1}{2}} \sin\theta\sin\phi \\ z &= r\cos\theta \\ t &= t \end{aligned} \tag{5.230}$$

Claim 17. The coordinate transformations in Eq 5.230, will lead to a Minkowski metric

Proof 17. Start of by computing the derivatives of Eq 5.230:

$$\begin{aligned} dx &= \frac{rdr}{(r^2 + a^2)^{\frac{1}{2}}} \sin\theta\cos\phi + d\theta (r^2 + a^2)^{\frac{1}{2}} \cos\theta\cos\phi - (r^2 + a^2)^{\frac{1}{2}} \sin\theta\sin\phi d\phi \\ dy &= \frac{rdr}{(r^2 + a^2)^{\frac{1}{2}}} \sin\theta\sin\phi + d\theta (r^2 + a^2)^{\frac{1}{2}} \cos\theta\sin\phi + (r^2 + a^2)^{\frac{1}{2}} \cos\theta\cos\phi d\phi \\ dz &= dr\cos\theta - d\theta r\sin\theta \end{aligned} \tag{5.231}$$

therefore:

$$dx^2 + dy^2 + dz^2 = \frac{r^2 + a^2\cos\theta}{r^2 + a^2} dr^2 + (r^2 + a^2\cos^2\theta) d\theta^2 + (r^2 + a^2)\sin^2\theta d\phi^2 \tag{5.232}$$

which is the spatial part of Eq 5.228. These coordinates are called *ellipsoidal* coordinates.

when r is a constant:

$$\begin{aligned} y^2 + z^2 &= (r^2 + a^2)\sin^2\theta + r^2\cos^2\theta \\ &= r^2 + a^2\sin^2\theta \end{aligned} \tag{5.233}$$

when $r = 0 \Rightarrow z = 0$ and:

$$y = a\sin\theta \tag{5.234}$$

therefore y simply oscillates with θ. This metric has appears to have to singularities; one at $\Delta = 0$ and one at $\rho = 0$. We will see that the singularity at $\Delta = 0$ is actually a coordinate singularity and therefore it can be removed by an appropriate coordinate transformation. The singularity at $\rho = 0$ implies:

$$\rho^2 = r^2 + a^2\cos^2\theta = 0 \tag{5.235}$$

The only way this equation can be true is if both $r = 0$ and $\cos\theta = 0 \Rightarrow \theta = \pm\frac{\pi}{2}$, as a is non zero. So infact the singularity in this case is not actually a point, it is a ring. This can be seen by the definition of these coordinates:

$$\begin{aligned} &\frac{x^2 + y^2}{r^2 + a^2} + \frac{z^2}{r^2} \equiv 1 \\ &z \equiv r\cos\theta \end{aligned} \tag{5.236}$$

Therefore for this singularity:

$$x^2 + y^2 = 1 \tag{5.237}$$

which is a ring (circle) in two dimensions.

3.1. Event horizons.

Definition 16. An event horizon is a surface beyond which it is not possible to communicate with observers at infinity.

These surfaces have a property, that is their area, which is actually associated with the entropy of a black hole. It can be shown that any space-time surface lying inside the event horizon has the same area.

Example 18. For the Schwarzschild metric, Eq 5.116, if we choose a surface of constant t at the event horizon, which in this case is the Schwarzschild radius, r_s, the metric reduces to:

$$ds^2 = r_s^2 d\Omega_2^2 \tag{5.238}$$

Integrating over the solid angle, simply gives:

$$\text{Area} = 4\pi r_s^2 \tag{5.239}$$

In the Schwarzschild/Kerr metric, the condition for an event horizon is that:

$$g^{\mu\nu}\partial_\mu r \partial_\nu r = 0 \tag{5.240}$$

where:

$$\begin{aligned} \text{Schwarzschild} \quad &\Rightarrow \quad g^{rr} = 1 - \frac{r_s}{r} \\ \text{Kerr} \quad &\Rightarrow \quad g^{rr} = \frac{\Delta}{\rho^2} \end{aligned} \tag{5.241}$$

For large r, the metric is expected to asymptotically become flat. The quantity $\partial_\mu r$ becomes null at the event horizon, which means a light ray will simply orbit the surface at the same radius. For the Kerr metric, $g^{rr} = 0$ when $\Delta = 0$:

$$\Delta = r^2 - r_s r + a^2 = 0$$
$$r = \frac{r_s}{2} \pm \sqrt{\frac{r_s^2}{4} - a^2} \equiv r_\pm \tag{5.242}$$

where we have assumed $\frac{r_s^2}{4} > a^2$. When:

$$J = \frac{GM}{c^2} \tag{5.243}$$

the black hole is dais to be in the *extremal* limit, as the black hole cannot rotate any faster than this.

When Kerr first derived his solution, the coordinates he used were different and given by:

$$\begin{aligned} dT &= dr + dr\frac{r_s r}{\Delta} \\ R &= r \\ \Theta &= \theta \\ \Phi &= \phi - \int^r \frac{dr}{\Delta} \\ d\Phi &= d\phi - \frac{dr}{\Delta} \end{aligned} \tag{5.244}$$

This gives the metric:

$$ds^2 = -dT^2 + dR^2 + 2a\sin^2\theta dr d\Phi + \rho^2 d\Theta^2 + (R^2 + a^2)\sin^2\Theta d\Phi^2 - \frac{r_s R}{\rho^2}\left(dR + a\sin^2\Theta d\Phi + dT\right)^2 \tag{5.245}$$

where:

$$\rho = R^2 + a^2\cos^2\Theta \equiv r^2 + a^2\cos^2\theta \tag{5.246}$$

The reason to used these coordinates is that there is no Δ in this metric, which means there is no singularity at $\Delta = 0$. These coordinates are called *Kerr-Eddington-Finkelstein*, as they are analogous to the Eddington-Finkelstein coordinates for the Schwarzschild metric. However even in these coordinates, notice that $\rho = 0$ is singular.

3.2. Stationary limit surface. Beyond the surface of the stationary limit, it is not possible to have a constant r, θ, ϕ. This surface is defined by:

$$g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu = -1 \tag{5.247}$$

for a massive particle. In the stationary limit, by definition:

$$\dot{r} \equiv \dot{\phi} \equiv \dot{\theta} \equiv 0 \tag{5.248}$$

Therefore the only way Eq 5.247 can be true is if $g_{00} < 0$. Thus the limiting case for particles being stationary is:

$$g_{00} = 0 \tag{5.249}$$

Therefore:

$$-\frac{1}{\rho^2}(r^2 + a^2\cos^2\theta - r_s r) = 0 \tag{5.250}$$

Which has the solution:

$$r_\pm^{(s)} = \frac{r_s}{2} \pm \sqrt{\frac{r_s^2}{4} - a^2\cos^2\theta} \tag{5.251}$$

Therefore:

$$r_+^s > r_+ \qquad \text{when } \theta \neq 0 \tag{5.252}$$

when $\theta = 0 \Rightarrow r_+^{(s)} = r_+$. The region:

$$r_+ < r < r_+^{(s)} \tag{5.253}$$

is called the *Ergosphere*. In this region, the space-time is rotating so fast that a particle cannot stay stationary in it.

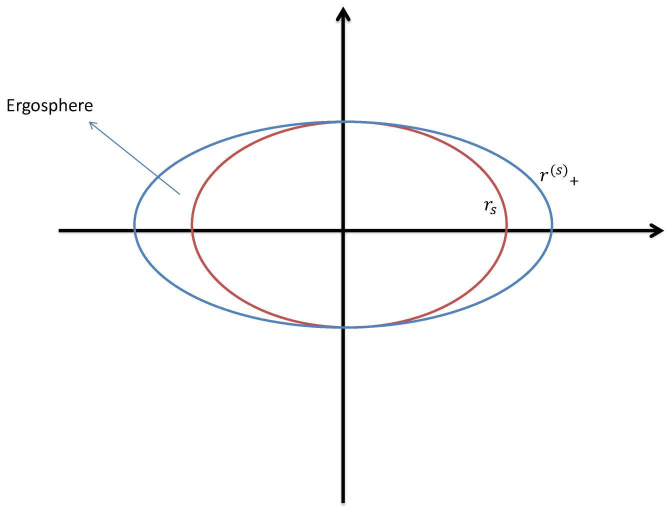


FIGURE 28. Ergosphere of a rotating black hole

3.3. Penrose process. Because the ergosphere is outside the event horizon, it is still possible to communicate with a particle in this region of space. Penrose showed that it is possible to obtain energy from the black hole from the ergosphere.

The idea is to put an object into the ergosphere and then have the object throw a particle into the black hole and then the particle comes out of the ergosphere with more energy then it had in the beginning. This is because the conservation of energy is related to the coordinate t; it is the conjugate momentum of t. Now, by the definition of the ergosphere, g_{00} is zero, which means there is not time component in the metric and hence the Lagrangian. Since E is the canonical momentum conjugate to t, it is given by:

$$E = \frac{\partial L}{\partial t} = 0 \tag{5.254}$$

Therefore any variation in time must leave the energy unchanged, i.e energy is conserved.

If the trajectory of a particle is parametrised by a parameter λ, then:

$$E \propto \frac{\partial t}{\partial \lambda} \tag{5.255}$$

In general t will increase with λ, i.e a future pointing particle will have a positive value of $\frac{\partial t}{\partial \lambda}$, thus the energy will be positive. This means that the future moving particles (which are the only allowed particles) can only have access to these positive energy states. The negative energy states are in a causally disconnected region of space-time[3]

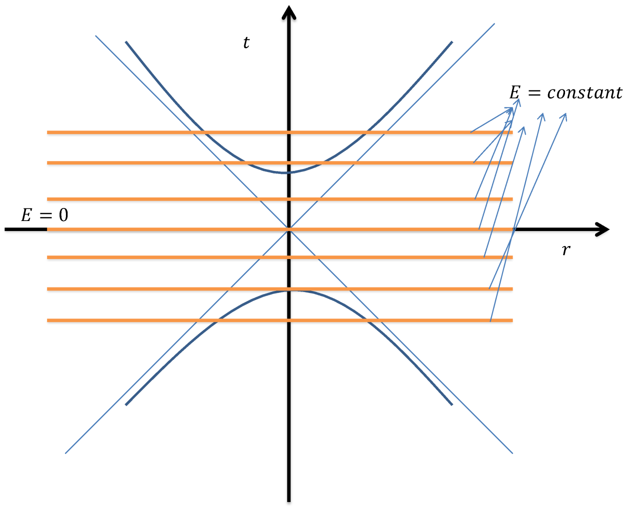


FIGURE 29. Lines of constant energy outside Ergosphere

On the other hand, inside the ergosphere the space and time coordinates are flipped, which means that any future moving particle can exist in both positive and negative energy states! This bizarre behavior is still not fully understood, however this is what leads to a particle being able to come out of the ergosphere with more energy than it had before. As if an object throws a particle in the negative r direction, i.e towards the black hole from a positive r position, then it will have to gain the energy that the particle moving in the negative r direction will be loosing as the energy is conserved.

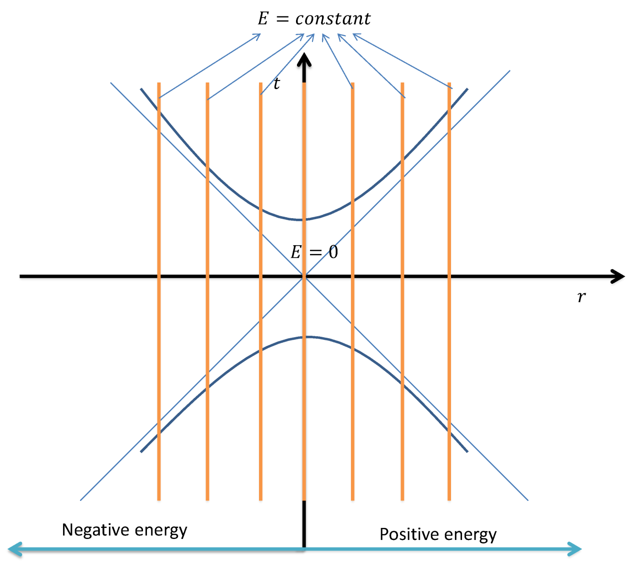


FIGURE 30. Lines of constant energy inside Ergosphere

[3]This shows another strange property of the hypothetical particles, *tachyons*, that can travel faster than the speed of light. Not only would they violate the natural speed limit, but they would also not conserve energy!

Part 2

Advanced gravitational theory

CHAPTER 6

Mathematical framework

1. Differentiating of a manifold

Manifolds have been discussed in Chapter 3. The main type of derivative discussed in there was the covariant derivative, Eq 3.192. However this is not the most general type of derivative. There are other forms of differentiation that can be defined on a manifold and they will be the discussed in detail in this section. The starting point to thinking about differentiation on a manifold, is to recall the familiar 3D vector calculus that Maxwell invented to describe his unification of electricity and magnetism, as shown in Chapter 1. Firstly, we can define:

$$\vec{\nabla} \equiv \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{6.1}$$

a differential operator, which actually behaves like a vector (or co-vector, it does't make any difference as the $\vec{\nabla}$, is defined in 3D flat Euclidean space, therefore the indicies can be raised or lowered as required by the Minkowski metric). Derivatives have a direction and a magnitude. So let's recap from Chapter 3, about what goes wrong on a manifold. The derivatives $\frac{\partial}{\partial x^i}$, gives components of gradients. So one can think of taking $\frac{\partial}{\partial x^i}$ of an object on a manifold; consider a vector field on the manifold, $\vec{V}$. Now take the derivative:

$$\frac{\partial V^\mu}{\partial x^\nu} \tag{6.2}$$

So these are the components, of the derivatives, of V^μ. Then under a change of coordinates:

$$\frac{\partial V^\mu}{\partial x^\nu} = \frac{\partial}{\partial x^\nu}\left(V'^\mu \frac{\partial x'^\mu}{\partial y'^\mu}\right) \tag{6.3}$$

where the bracket term is the definition of how a vector transforms under coordinate transformations. Using the product rule:

$$\frac{\partial V^\mu}{\partial x^\nu} = \underbrace{\frac{\partial y\nu'}{\partial y^\nu}\frac{\partial V^{\mu\prime}}{\partial y'^\mu}\frac{\partial x^\mu}{\partial y'^\mu}}_{T_\alpha} + \underbrace{V'^\mu \frac{\partial^2 x^\mu}{\partial y'^\mu \partial y'^\mu}\frac{\partial y'^\nu}{\partial x^\nu}}_{T_\beta} \tag{6.4}$$

T_α is the expected term and T_β which is not a tensor, but it is on object that is symmetric in ν', μ'. This was already shown in Eq 3.36 and shows why the concept of simply taking partial derivatives of a vector does not work. This might even be visible from the fact that $\frac{\partial}{\partial x^\mu}$, is not a geometric object, it depends on the coordinates one is working with. Therefore one has to go back to the very beginning and the fundamental theorem of calculus:

$$f'(t) = \lim_{\delta t \to 0} \frac{f(t+\delta t) - f(t)}{\delta t} \tag{6.5}$$

to rethink about how to define a derivative to get a geometric definition of a derivative. In the usual real analysis, the derivative at point is defined by a limiting process as shown in the equation above. The "$t + \delta t$" in this expression is saying, move a little bit away from a point t. In terms of a manifold, one would need to move from one point to another in the manifold:

$$p_1 \to p_2\{|p_2 = p_1 + \delta p; \delta p << 1\} \qquad p_1, p_2 \in M \tag{6.6}$$

The problem is, that the manifold is not the same at every point. Each point on the manifold corresponds to a different target space. Thus it is difficult to define the notion of δt. One has to specify how to move across this distance. To do this, one has to define a way to move to neighboring points that is independent of coordinates (i.e geometrically). Note that the disappearing in partial differentiation is symmetric.

The covariant derivative is one way of getting rid of that extra term as discussed before and provides a *connection*, which is a way of linking tangent spaces together. As any two tangent spaces at different points, are, by definition, subsets at $\mathbb{R}^n$, means there should exist a map linking them together, which is given by the connection. Moving to a nearby point, to compare a geometric object, is also done with the help of a connection, but this leads to concept of Lie derivative. The Lie derivative is taken by moving along vector fields (in fact once there is a vector field, the tangent space can also be moved along the vector field).

1.1. Forms & Exterior derivatives. This is a first look at a geometric derivative that does not depend on a metric or a coordinate. The exterior derivative, $\vec{d}$, acting on a function, f, maps the function onto a co-vector of 1 form:

$$\vec{d}f \mapsto T*(M) \qquad \forall f \in \{C^\infty(M)\} \tag{6.7}$$

such that:

$$\langle \vec{d}f | \vec{T} \rangle = \vec{T}f \ \ \forall \, \vec{T}_p \in T(M) \text{ at } p \tag{6.8}$$

This is saying that $\vec{d}f$, is a vector (as it is a linear map from $T_p(M) \mapsto \mathbb{R}$), acting on a point in the tangent spaces $\vec{T}$, to give a number $\vec{T}f$, that's true for every element in $T_p(M)$, therefore $d\vec{f}$ is an independent geometric object and in the coordinate basis:

$$\langle \vec{d}f | \frac{\partial}{\partial x^\mu} \rangle = \frac{\partial f}{\partial x^\mu} \tag{6.9}$$

Then $\vec{d}$ looks like the gradient operator in vector calculus. So a vector is defined as a 1 form, the next obvious question to ask, what a p-form is. A p-form is an anti-symmetric rank p covariant tensor. It is an element of:

$$\Lambda^{(p)}_Q(M) = T*_Q(M) \wedge T*_Q(M) \otimes ... \tag{6.10}$$

where $\otimes$ is known as the exterior/wedge product. It is defined to an be anti-symmetric product.

Example 19. As an example of an exterior product, lets consider the exterior product between 1 forms:

$$A \wedge B \equiv A \otimes B - B \otimes A \tag{6.11}$$

Definition 17. Generally for a p form and a q form the exterior product is defined as:

$$\left(A^{(p)} \wedge B^{(q)}\right)_{a_1,...a_{p+q}} \equiv \frac{(p+q)!}{p!q!} A_{[a_1,...a_p} B_{a_{p+1},...a_{p+q}]} \tag{6.12}$$

An exterior product is linear, but not commutative:

$$A^{(p)} \wedge B^{(q)} \equiv (-)^{pq} B^{(a)} \wedge A^{(p)} \tag{6.13}$$

The maximum value for a p form is the dimensionality of the manifold. A rank n form (where $n = dim(M)$), is proportional to the alternating symbol, ϵ (not the same as the permutation symbol), specifically defined for this part as:

$$\epsilon_{a_1,...a_n} = \begin{cases} +1 & \text{odd permutations of } a_1,...a_n \\ -1 & \text{even permutations of } a_1,...a_n \end{cases} \tag{6.14}$$

ϵ can be thought of as a tensor density, as in doing a coordinate transformations, one picks up the terms of the determinant of the Jacobian matrix:

$$\begin{aligned}\epsilon_{\mu'\nu'\lambda'\rho'} &= \frac{\partial x^{\mu}}{\partial y'^{\mu}}\frac{\partial x^{\rho}}{\partial y^{\rho}}\epsilon_{\mu\nu\lambda\rho} \\ &\equiv det\left(\frac{\partial x}{\partial y}\right)\epsilon_{\mu'\nu'\lambda'\rho'} \end{aligned} \tag{6.15}$$

If one has a metric defined as:

$$\begin{aligned} g_{\mu'\nu'} &= \frac{\partial x^{\mu}}{\partial y'^{\mu}}\frac{\partial x^{\nu}}{\partial y'^{\nu}}g_{\mu\nu} \\ &= det(g_{\mu'\nu'}) \\ &= \left|\frac{\partial x}{\partial y}\right|^2 det(g_{\mu\nu}) \end{aligned} \tag{6.16}$$

thus one can define:

$$\epsilon_{\mu\nu\lambda\rho} \equiv \sqrt{(g)}\epsilon_{\mu\nu\lambda\rho} \tag{6.17}$$

This quantity transforms as a tensor. Therefore once a metric has been introduced, we can define a map $*$:

$$* : \Lambda^p(M) \mapsto \Lambda^{n-p}(M) \tag{6.18}$$

$*$ is called the *Hodge star operator*. So we can go from a p form to an $n-p$ form, by taking the ϵ tensor and contracting it with the p form:

$$(*A)_{a_1,\dots a_{n-p}} = \frac{1}{p!}\epsilon^{b_1,\dots b_n}_{a_1,\dots a_{n+p}}A_{b_1,\dots b_n} \tag{6.19}$$

DEFINITION 18. Now one can defined an exterior derivative that maps p forms to $p+1$ forms:

$$(\vec{d}A)_{a_1,\dots a_{p+1}} \equiv \frac{|(p+1)|}{p!}\partial_{[a_1}A_{a_2,\dots a_{n+1}]} \tag{6.20}$$

This is the definition of the exterior derivative. In words it states, take the partial derivative of p form A, and then anti-symmetrise in all possible permutations of the indicies. It also follows:

$$\vec{d}(A^{(p)} \wedge B^{(q)}) = \vec{d}A^{(p)} \wedge B^{(q)}(-)^p A^{(p)} \wedge \vec{d}B^{(q)} \tag{6.21}$$

One can also use $*$ to define:

$$\delta = *\vec{d}* \tag{6.22}$$

Which is a map from Λ^p to Λ'.

EXAMPLE 20. As an example of how these exterior derivatives are used, consider the electromagnetic potential:

$$A_\mu = (\phi_1 - \vec{A}) \tag{6.23}$$

electro-magnetism is naturally described as the gauge theory and the gauge potential, A_μ, is made from the vector and scalar potential. A_μ is a 1 form, the exterior derivative of it is:

$$\begin{aligned}(\vec{d}A)_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \\ &= \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix}\end{aligned} \tag{6.24}$$

Which is just the stress-energy tensor of electro-magnetism. Since $\vec{d}^2$, as $\vec{d}$ takes a p form and $p+1$ form and the p form is limited by the dimensionality of the manifold, $\vec{d}F = \vec{d}(\vec{d}A) = 0$, gives half of Maxwell's equations. The other half come from $\vec{d}F = 0$. It also shows the gauge invariance of these equations:

$$A \to A + \vec{d}\Lambda \tag{6.25}$$

$$\begin{aligned} \vec{d}A \to \quad &= \quad \vec{d}A + \vec{d}(\vec{d}\Lambda) \\ &= \quad \vec{d}A \end{aligned} \tag{6.26}$$

EXAMPLE 21. Now lets consider a two form:

$$B_{\mu\nu} = 2 \text{ form} \tag{6.27}$$

Then the gauge transformations:

$$B^{(2)} \to B^{(2)} + \vec{d}A^{(1)} \tag{6.28}$$

H can be defined as the field strength:

$$H \equiv \vec{d}B \tag{6.29}$$

In components:

$$H_{\mu\nu\lambda} = \partial_\mu B_{\nu\lambda} + \partial_\nu B_{\lambda\mu} + \partial_\lambda B_{\mu\nu} \tag{6.30}$$

This is used in string theory and super-gravity.

This can be generalised to p forms. But the exterior derivative is just one way of taking a geometric derivative. Another way is a *Lie* derivative, which will be discussed next.

1.2. Lie derivative. A Lie derivative takes a derivative along a vector field in a manifold.

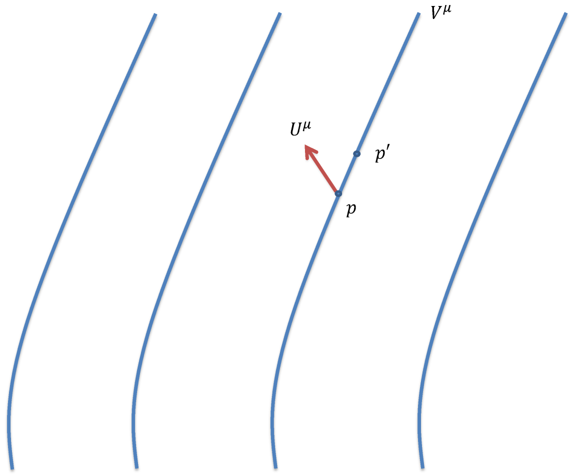


FIGURE 31. Schematic of vector field of manifold

V^μ is a vector field. To take the derivative, one has to look at the values of the field at a local point, say p, and the compare it to the value of the field at a neighboring point. This is not a general exterior derivative or the covariant derivative, as one is fixing the direction along a vector field and then comparing how it changes along two neighboring points in that direction.

The vector field $\vec{V}$, defines a way of pushing forward the tangent space at p to p'. Move a small amount along a curve:

$$x_t^\mu = x_0^\mu + \delta t V^\mu + \mathcal{O}(\delta t^2) \tag{6.31}$$

This states that we start at a point p, that's x_0^μ. Move a small amount δt along this curve, then to leading order, the coordinate transformation is:

$$x_t^\mu = x_0^\mu + \delta t V^\mu \tag{6.32}$$

Therefore:

$$U_t^\mu = \frac{\partial X_t^\mu}{\partial x_0^\mu} U_0^\nu \tag{6.33}$$

If the coordinate transformation of U under $X_0^\mu \to X_t^\mu$, and its a contravariant coordinate transformation. Substitute Eq 6.32 into Eq 6.33:

$$U_t^\mu = U_0^\mu + \delta t V_{0,\nu}^\mu U_0^\nu \tag{6.34}$$

This can be used to define the derivative, as now we have the framework to transport U from p to p'.

Definition 19. We define:

$$\begin{aligned}
(\mathcal{L}_\nu U)^\mu &\equiv \lim_{\delta t \to 0} \frac{[U(x_t^\mu) - U_t]}{\delta t} \\
&= \frac{1}{\delta t}\left[U^\mu(x^\lambda + \delta t V^\lambda) - (U_0^\mu + \delta t V_{0,\nu}^\mu U^\nu)\right] \\
&= \frac{1}{\delta t}\left[U_0^\mu + \delta t V^\lambda U_{,\nu}^\mu - U_0^\mu - \delta t V_{0,\nu}^\mu U^\nu\right] \\
&= \frac{1}{\delta t}\left[\delta t V^\lambda U_{,\nu}^\mu - \delta t V_{0,\nu}^\mu U^\nu\right] \\
&= V^\lambda U_{,\lambda}^\lambda U^\lambda
\end{aligned} \tag{6.35}$$

This is the Lie derivative for a co-vector:

$$(\mathcal{L}_\nu, \omega)_\mu = \omega_{\mu\sigma} V^\sigma + \omega_\sigma V_{,\mu}^\sigma \tag{6.36}$$

For vectors $\mathcal{L}_\nu U$ is also defined as a commutator:

$$\mathcal{L}_\nu U = [V, U] \tag{6.37}$$

where:

$$[V, U]\, f \equiv \vec{v}(\vec{u}\vec{p}) - \vec{u}(\vec{v}, \vec{p}) \qquad \forall f \in C^*(M) \tag{6.38}$$

To illustrate the significance of this Lie bracket consider Figure 32, where U and V are both vectors fields, this is showing that, in general, going from one point to another in a curved space is *not independent* of the path taken. In other words, taking a path $V + U$ is not some as the path $U + V$. The lie bracket is important as it is related to the difference between the two paths.

Claim 18. The difference between the two paths shown in Figure 32, the dotted line, is proportional to the lie bracket:

$$[U, V] \tag{6.39}$$

Proof 18. This will be shown by sticking to the underlying coordinate chart. Look at the coordinates at point A on the diagram, to get to A:

$$X^\mu_A = X^\mu_0 + tV^\mu_0 + \frac{1}{2}t^2 V^\mu_{0,\nu}V^\nu_0 \tag{6.40}$$

This basically says, start at point O, then move distance, t, along V^μ_0, (i.e the direction of the vector field, V^μ, at point O) and to second order one just expands about small quantity, t. This is just a Taylor expansion. Similarly for point B:

$$\begin{aligned} X^\mu_B &= X^\mu_A + sU^\mu_A + \frac{1}{2}s^2 U^\mu_{\lambda,\nu}U^\nu_A \\ &= X^\mu_0 + tV^\mu_0 + \frac{1}{2}t^2 V^\mu_{0,\nu}V^\nu_0 + sU^\mu_A + \frac{1}{2}s^2 U^\mu_A \end{aligned} \tag{6.41}$$

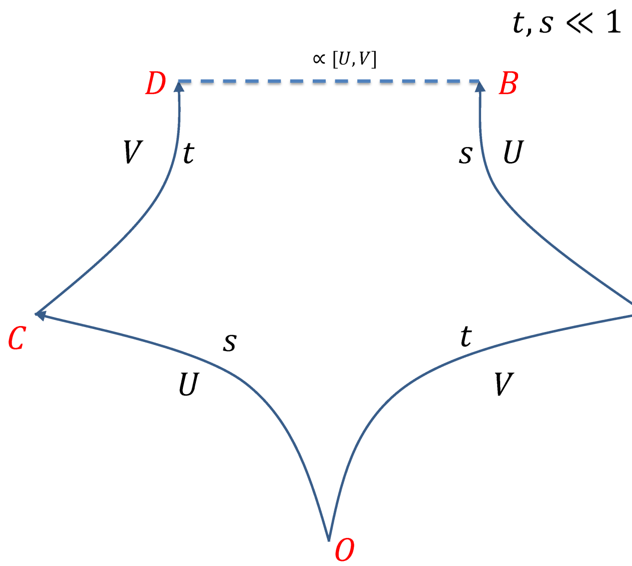


Figure 32. Schematic showing that the Lie derivative is related to the path difference along a manifold

Now we have to expand the terms U^μ_A in a similar way but since all terms of order higher than 2 are ignored, we get:

$$X^\mu_B = X^\mu_0 + tV^\mu_0 + \frac{1}{2}t^2 V^\mu_{0,\nu}V^\nu_0 + s(U^\mu_0 + sU^\mu_0) + \frac{1}{2}s^2 U^\mu_0 \tag{6.42}$$

Now use the transformation property:

$$U^\mu_0 = \frac{\partial U^\mu}{\partial x^\nu}U^\nu_0 \equiv U^\mu_{0,\nu}U^\nu_0 \tag{6.43}$$

and the fact that at a point the vector fields have the same value:

$$U^\nu_0 s = tV^\nu_0 \tag{6.44}$$

Therefore:

$$X^\mu_B = X^\mu_0 + tV^\mu_0 + \frac{1}{2}t^2 V^\mu_{0,\nu}V^\nu_0 + s(U^\mu_0 + tU^\mu_{0,\nu}V^\nu_0) + \frac{1}{2}ts^2 U^\mu_{0,\nu}U^\nu_0 \tag{6.45}$$

X_D can be found in the same way, by swapping U and V fields and s and t parameters:

$$X^\mu_D = X^\mu_0 + sU^\mu_0 + \frac{1}{2}S^2 U^\mu_{0,\nu}U^\nu_0 + t(V^\mu_0 + sV^\mu_{0,\nu}U^\nu_0) + \frac{1}{2}t^2 V^\mu_{0,\nu}V^\nu_0 \tag{6.46}$$

So we see that $X_B^\mu = X_D^\mu$ up to linear orders in t and s. However there is a cross term which is not the same:

$$X_D^\mu - X_B^\mu = st\left[V_{0,\nu}^\mu U_0^\nu - U_{0,\nu}^\mu V_0^\nu\right] = st[U,V]^\mu \tag{6.47}$$

Therefore as claimed, the distance between X_D and X_B is proportional to the Lie bracket.

Thus, the Lie bracket gives an indication of how much the lines, or sometimes called *integral* curves(as they are curves obtained by integrating a vector field), fail to close. This means that if the Lie bracket of two vector fields vanishes, it means that the order in which a path is taken in a manifold is not important. Also, one can uniquely go along one end and then the other, to define a closed surface locally, which is uniquely labeled by how far one has gone along one curve and then the other. Which means U and V can define a coordinate system on that surface. This is one of the main uses of the Lie bracket. Note that there is a relation between the exterior product and the Lie derivative:

$$\langle \vec{d}\vec{w}|\vec{u},\vec{v}\rangle \equiv \vec{u}\left(\langle\vec{w}|\vec{v}\rangle\right) - \vec{v}\left(\langle\vec{w}|\vec{y}\rangle\right) - \langle\vec{w}|\left[\vec{u},\vec{v}\right]\rangle \tag{6.48}$$

2. Killing vector

Definition 20. A killing vector is a vector field, $\vec{g}$, along which the metric is Lie invariant, i.e the Lie derivative of the vector field is zero:

$$\mathcal{L}_v\vec{g} = 0 \tag{6.49}$$

where $\vec{g}$ is the metric.

Example 22. As an example:

$$v = \frac{\partial}{\partial\phi} \tag{6.50}$$

is a killing vector if ϕ is some periodic coordinate. The statement is that if v is a killing vector, the metric does not change under $\frac{\partial}{\partial\phi}$. If ϕ is a coordinate defining a circle, then the fact that the metric is invariant under $\frac{\partial}{\partial\phi}$, is a statement of axial symmetry or symmetry of rotation around some axis. This is the key point; the killing vector gives an insight into the symmetries that are possessed by a metric.

3. Geometrical connection

The connection has already been discussed in Chapter 3, where it was introduced in relation with its use in the covariant derivative. Here we take a more geometric view and look at how to compute connection terms using exterior derivatives. So far we have seen how to take derivatives on a manifold. It was shown that the derivative must have a directive aswell as a magnitude and this was motivated by looking at the gradient in the usual, 3D vector calculus and then trying to simply use the partial derivative on manifold gives a term that the second derivative that did not transform as a tensor.

The connection provides a way of linking tangent spaces together.

Definition 21. Define:

$$\nabla\vec{e}_b \equiv \Gamma^a_{cb}\vec{e}_a \otimes \vec{w}^c \tag{6.51}$$

Recall that $\vec{e}$ is the basis for the tangent space and $\vec{w}$ is the basis for the $\vec{w}$ tangent space. This defines the action of the connection on a basis. This can be equivalently defined as:

$$\Gamma^a_{bc} \equiv \langle\vec{w}^a|\underbrace{\nabla_b}_{\vec{e}_b\cdot\nabla}\vec{e}_c\rangle \tag{6.52}$$

Note that the first lower index in Γ^c_{ab}, i.e a in this case, is always referring to a differentiating index.

The ∇ has the following properties:

- Commutes with contraction.
- Liebnizian.
- Reduces to $\vec{d}$ when acting on a function.

For a general vector:

$$\nabla\vec{V} = \nabla \underbrace{(\vec{V}^a\vec{e}^a)}_{\text{components}} \tag{6.53}$$

V expressed in a particular basis, $\vec{e}_a$ on the tangent space. From the property of commutativity and Leibnizian:

$$\begin{aligned} \nabla V &= (\nabla V^a)\vec{e}_a + V^a\nabla\vec{e}_a \\ &= \vec{d}V^a\vec{e}_a + V^a\Gamma^b_{ca}\vec{e}_b\vec{w}^c \end{aligned} \tag{6.54}$$

By simply re-labeling the indices:

$$T_1 \equiv \nabla V = (V^a_{,b} + V^c\Gamma^a_{bc})\vec{e}_a \otimes \vec{w}^b \tag{6.55}$$

T_1 is the usual definition of the covariant derivative. But $T_1 = \nabla_b V^a$. This is known as the *abstract* index rotation that is almost always used. However in general one has to include the basis terms at the end. In general relativity, we use a torsion-free metric connection:

$$\nabla\vec{g} = 0 \tag{6.56}$$

Torsion is defined as:

$$\vec{T}(\vec{U},\vec{V}) \equiv \nabla_\mu\vec{V} - \nabla_\nu\vec{U} - \left[\vec{U},\vec{V}\right] \tag{6.57}$$

it is a tensor as the difference of curvature of any 2 metrics is a tensor. This is very nearly the anti-symmetric part of the connection:

$$\Gamma^a_{bc} = \Gamma^a_{bc} - \Gamma^a_{cb} - C^a_{bc} \tag{6.58}$$

where:

$$C^a_{bc} \equiv \langle\vec{w}^a|\,[\vec{e}_a,\vec{e}_c]\rangle \tag{6.59}$$

this is a correction term and it allows for the fact that there can be vector fields that don't commute. These $C's$ are called the *structure* constants of the basis $\{\vec{e}_a\}$. If the structure constants are zero, the connection reduces to the Christoffel symbol, as shown in Eq 2.29.

Definition 22. One can also define the connection 1 form as:

$$\vec{\theta}^a_b = \Gamma^a_{cb}\vec{w}^c \tag{6.60}$$

Taking the connection components and contracting the differentiating index, from Eq 6.52:

$$\nabla\vec{e}_b \equiv \vec{\theta}^b_a \otimes \vec{e}_b \tag{6.61}$$

This defines the θ's.

Claim 19.

$$\vec{d}g_{ab} = \vec{\theta}_{ab} + \vec{\theta}_{ba} \tag{6.62}$$

Here $\vec{\theta}_{ab}$ is anti-symmetric in ortho-normal basis.

Proof 19. Start with:

$$\vec{d}(g_{ab}) = \vec{d}(g(\vec{e}_a, \vec{e}_b)) \tag{6.63}$$

g_{ab} are the components of the metric in the basis, $\vec{e}_a$. Which means we take 2 basis vectors, $\vec{e}_a$ and $\vec{e}_b$ and contract them using the metric tensor:

$$\vec{d}(g_{ab}) = \nabla g(\vec{e}_a, \vec{e}_b) \tag{6.64}$$

Because $\vec{d} = \nabla$ when acting on a function. Since the ∇ commutes with contraction and we have a metric connection(i.e Leibniz):

$$\begin{aligned} d(g_{ab}) &= \vec{g}(\nabla\vec{e}_a, \vec{e}_b) + \vec{g}(\vec{e}_a, \nabla\vec{e}_b) \\ &= \vec{\theta}^c_a \vec{g}(\vec{e}_c, \vec{e}_b) + \vec{\theta}^c_b \vec{g}(\vec{e}_b, \vec{e}_c) \end{aligned} \tag{6.65}$$

So the covariant derivative of each basis is given in terms of the connection 1 forms:

$$\begin{aligned} \vec{d}(g_{ab}) &= g_{cb}\vec{\theta}^c_a + g_{ac}\vec{\theta}^c_b \\ &= \vec{\theta}_{ba} + \vec{\theta}_{ab} \end{aligned} \tag{6.66}$$

Now we want to get an expression, of deriving these connection 1 forms. To do so, first lets re-arrange the wedge product:

$$\vec{\theta}^a_c \wedge \vec{w}^c = \Gamma^a_{bc}\vec{w}^b \wedge \vec{w}^c \tag{6.67}$$

$\vec{w}$ is anti-symmetric in b and c:

$$\vec{\theta}^a_c \wedge \vec{w}^c = \frac{1}{2}(\Gamma^a_{bc} - \Gamma^a_{cb})\vec{w}^b \wedge \vec{w}^c \tag{6.68}$$

But the anti-symmetric part of the connection is the torsion plus the structure constant:

$$\vec{\theta}^a_c \wedge \vec{w}^c = \frac{1}{2}(T^a_{bc} + C^a_{bc})\vec{w}^b \wedge \vec{w}^c \tag{6.69}$$

So the first part is geometric (with the torsion):

$$\vec{\theta}^a_c \wedge \vec{w}^c \equiv \vec{T}^a - \frac{1}{2}\langle \vec{w}^a | \, [\vec{e}_b, \vec{e}_c] \rangle \vec{w}^b \wedge \vec{w}^c \qquad \vec{T}^a \equiv \frac{1}{2}T^a_{bc}\vec{w}^b \wedge \vec{w}^c \tag{6.70}$$

But from Eq 6.48:

$$\langle \vec{w}^a | \, [\vec{e}_b, \vec{e}_c] \rangle = -\langle \vec{d}\vec{w}^a | \vec{e}_b \vec{e}_c \rangle - \vec{e}_b(\delta^a_c) + \vec{e}_c(\delta^a_b) = -\langle \vec{d}\vec{w}^a | \vec{e}_b \vec{e}_c \rangle \tag{6.71}$$

Combining Eq 6.70 and 6.71:

$$\vec{\theta}^a_c \wedge \vec{w}^c = \vec{T}^a - \vec{d}\vec{w}^a \tag{6.72}$$

This is known as Cartan's first equation.

4. Curvature

We define the curvature of the connection essentially as the commutator of derivatives, this is directly analogous to gauge theory, in which a commutator of gauge invariant derivatives is used to get the curvature or in this case the curvature of gauge connections or physical field strengths.

The Riemann curvature is defined as a map from three copies of the tangent space, into a copy of the tangent space. This is done by taking a commutator of covariant derivatives and then subtracting a Lie bracket part, which takes into account any inherent anti-symmetry that is already present because u and v don't commute. In components:

$$R^a_{bcd} = \Gamma^a_{bd,c} - \Gamma^a_{bc,d} + \Gamma^a_{ce}\Gamma^e_{db} - \Gamma^a_{de}\Gamma^e_{cb} - C^e_{cd}\Gamma^a_{eb} \tag{6.73}$$

It is very nearly the familiar expression in terms of the Christoffel symbols. The only extra bit is the last term in terms of the structure constants. Which is just saying that we are in a non-coordinate basis (in some awkward basis), and quite often this is the case and we may want to use an ortho-normal basis. The components of the Riemann curvature have to take into account the fact that the basis have this non-trivial behavior.

Cartan's second equation is a way of expressing the Riemann curvature using exterior derivatives and the connection 1 forms.

DEFINITION 23. First of all, define a curvature 2 form:

$$\vec{R}^a_b = \frac{1}{2} R^a_{bcd} \vec{w}^c \times \vec{w}^d \tag{6.74}$$

and then the curvature is given as:

$$\vec{R}^a_b = d\vec{\theta}^a_b + \vec{\theta}^a_c \wedge \vec{\theta}^c_b \tag{6.75}$$

This is Cartan's second equation.

Physically, curvature tells us about *tidal* forces. Suppose there is a vector field, T, which is geodesic:

$$\nabla_T T = 0 \tag{6.76}$$

This means there are a set of inertial observes on geodesics:

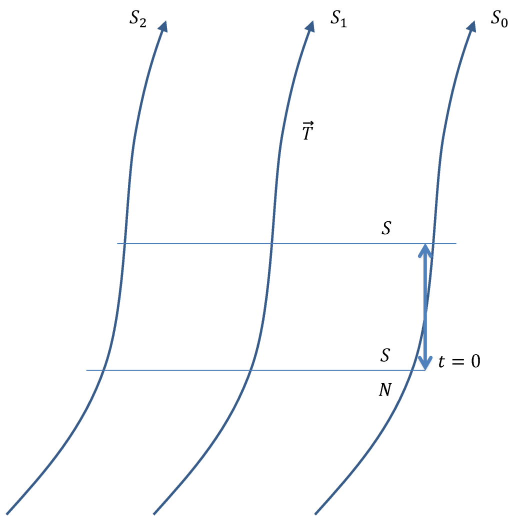


FIGURE 33. Geodesic of a vector field $\vec{T}$

The geodesics are labeled by parameter S. Suppose we go a small distance forward on the geodesics, and the geodesic is still labeled by S. In this way, we are setting up a local coordinate system. Let N be $\frac{\partial}{\partial S}$ connecting the geodesics. By construction the Lie bracket of T and N vanishes (as the two curves are closed). Thus the Riemann curvature is:

$$R^{\mu}_{\nu\lambda\rho} = -(\Gamma^{\mu}_{\nu\lambda,\rho} - \Gamma^{\mu}_{\rho\nu,\lambda} + \Gamma^{\mu}_{\rho\lambda}\Gamma^{\gamma}_{\nu\lambda} - \Gamma^{\mu}_{\gamma\lambda}\Gamma^{\gamma}_{\rho\nu}) \tag{6.77}$$

And if the curve is parametrised by time one gets the usual geodesic equations, as shown as Eq 3.149.

5. Using the Cartan formalism

To see how the Cartan structure equations are used, we will apply the formalism to a spherically symmetric static situation. The static part means that there is a killing vector:

$$\text{Static solution} \Rightarrow \text{Killing vector} \quad = \frac{\partial}{\partial t} \tag{6.78}$$

There is also a $t \to -t$ symmetry. Spherically symmetric means that there is an $SO(3)$ algebra of killing vectors, in other words:

$$[\xi_i, \xi_j] = \epsilon_{ijk}\xi_k \tag{6.79}$$

This says that there are surfaces in the geometry, which are invariant, and have a dimension ≤ 3. In this case we will be looking for two coordinates θ, ϕ on which the metric won't change. The metric being used is:

$$ds^2 = A(r)^2 dt^2 - B(r)^2 dr^2 - C(r)^2 d\Omega_2^2 \tag{6.80}$$

where the usual solid angle is defined as:

$$d\Omega_2^2 \equiv d\theta^2 + \sin^2\theta d\phi^2 \tag{6.81}$$

This metric is more general than is required because it allows for patterns to be seen in the solutions. Now we want to apply Cartan's equations, which means we have to go through the following steps.

- Identify an ortho-normal basis, so that the connection 1 forms are anti-symmetric. Since the metric is diagonal, the basis vectors are obvious:

$$\begin{aligned} \vec{w}^{\hat{t}} &\equiv Adt \\ \vec{w}^{\hat{\theta}} &\equiv Cd\vec{\theta} \\ \vec{w}^{\hat{r}} &\equiv Bd\vec{r} \\ \vec{w}^{\hat{\phi}} &\equiv C\sin\theta d\vec{\phi} \end{aligned} \tag{6.82}$$

  All we have to do is take the basic coordinate basis for the co-vector space, which is $dt, d\vec{r}, d\vec{\theta}, d\vec{\phi}$ and then multiply by an appropriate weighting factor, such that the metric is now simply the:

$$ds^2 = \vec{w}^{\hat{t}} + \vec{w}^{\hat{r}} + \vec{w}^{\hat{\theta}} + \vec{w}^{\hat{\phi}} \tag{6.83}$$

  the indices are hatted to emphasise the fact that we are referring to an co-tangent ortho-normal basis.
- Differentiate:

$$\vec{d}\vec{w}^t = A'(d\vec{r} \wedge dt) + A\vec{d}^2 t \tag{6.84}$$

  but recall that $d^2 = 0$, therefore:

$$d\vec{\omega}^t = A'\vec{dr} \wedge \vec{dt} \tag{6.85}$$

  But this needs to be expressed in terms of the ortho-normal basis:

$$d\vec{w}^{\hat{t}} = -\frac{A'}{AB}\vec{w}^{\hat{t}} \wedge \vec{w}^{\hat{r}} \tag{6.86}$$

  Where the minus sign comes from swapping r and t. Similarly:

$$\vec{d}\vec{w}^{\hat{r}} = B' \underbrace{\vec{dr} \wedge \vec{dr}}_{0} + B \underbrace{d^2 r}_{0} \equiv = 0 \tag{6.87}$$

$$\begin{aligned}\vec{d}\vec{w}^{\hat{\theta}} &= C'd\vec{\theta} \times d\vec{r} + C\vec{d}^2\theta \\ &= -\frac{C'}{CB}\vec{w}^{\hat{\theta}} \wedge \vec{w}^{\hat{r}}\end{aligned} \tag{6.88}$$

$$\begin{aligned}\vec{d}\vec{w}^{\hat{\phi}} &= C'\sin\theta d\vec{r} \wedge d\phi + C\cos\theta d\theta \wedge d\phi + C\sin\theta \vec{d}^2\phi \\ &= C'\sin\theta dr \wedge d\phi + C\cos\theta d\theta \wedge d\phi \\ &= -\frac{C'}{CB}\vec{w}^{\hat{\phi}} \wedge \vec{w}^{\hat{r}} - \frac{\cot\theta}{C}\vec{w}^{\hat{\phi}} \wedge \vec{w}^{\hat{\theta}}\end{aligned} \tag{6.89}$$

So now we have to take the derivative of the basis, the reason for doing that is because Cartan's first equation states that the connection 1 forms are related to exterior derivatives of the basis vectors.

- Now one can simply read off the connection: from Cartan's first equation:

$$\vec{d}\vec{w}^a + \vec{\theta}^a_b \wedge \vec{w}^b = 0 \tag{6.90}$$

So for the t component:

$$\vec{d}\vec{w}^{\hat{t}} = -\vec{\theta}^{\hat{t}}_{\hat{b}} \wedge \vec{w}^{\hat{b}} \tag{6.91}$$

Compare this to expression for $d\vec{w}^t$ just derived, Eq 6.86;

$$\begin{aligned}\vec{d}\vec{w}^{\hat{t}} &= -\frac{A'}{AB}\vec{w}^{\hat{t}} \wedge \vec{w}^{\hat{r}} \\ &= -\vec{\theta}^{\hat{t}}_{\hat{b}} \wedge \vec{w}^{\hat{b}}\end{aligned} \tag{6.92}$$

Therefore:

$$\vec{\theta}^{\hat{t}}_{\hat{r}} \equiv \frac{A'}{AB}\vec{w}^{\hat{t}} \tag{6.93}$$

Note that since are no off diagonal elements, this calculation is really straightforward. In fact most physically interesting systems are actually similar to this and do not diverge too much in there complexity. Applying the same method for the remaining components:

$$\begin{aligned}\vec{d}\vec{w}^{\hat{\theta}} &= -\vec{\theta}^{\hat{\theta}}_{\hat{b}} \wedge \vec{w}^{\hat{b}} \\ &= -\frac{C'}{CB}\vec{w}^{\hat{\theta}} \wedge \vec{w}^{\hat{r}} \\ &\Rightarrow \ \theta^{\hat{\theta}}_{\hat{r}} = \frac{C'}{CB}\vec{w}^{\hat{\theta}}\end{aligned} \tag{6.94}$$

$$\begin{aligned}\vec{d}\vec{w}^{\hat{\phi}} &= \vec{\theta^{\phi}}_{\hat{b}} \wedge \vec{w}^{\hat{b}} \\ &= -\frac{C'}{CB}\vec{w}^{\hat{\phi}} \wedge \vec{w}^{\hat{r}} - \frac{\cot\theta}{C}\vec{w}^{\hat{\phi}} \wedge \vec{w}^{\hat{\theta}}\end{aligned} \tag{6.95}$$

In this case the dummy index $\hat{b}$ will take two values, r and θ:

$$\begin{aligned}\vec{\theta}^{\hat{\phi}}_{\hat{r}} &= \frac{C'}{CB}\vec{w}^{\hat{\phi}} \\ \vec{\theta}^{\hat{\phi}}_{\hat{\theta}} &= \frac{\cot\theta}{C}\vec{w}^{\hat{\phi}}\end{aligned} \tag{6.96}$$

All others are zero.

- Now apply Cartan's second equation, Eq 6.75. We are looking for the curvature 2 forms and for Einstein metric connection, these are anti-symmetric on a and b and running over 4 values, one has 6 independent equations. First let's do the equations for $\hat{t}$ and $\hat{r}$:

$$R^{\hat{t}}_{\hat{r}} = \vec{d}\theta^{\hat{t}}_{\hat{r}} + \theta^{\hat{t}}_{\hat{a}} \wedge \theta^{\hat{a}}_{\hat{r}} \tag{6.97}$$

This is 0 $\forall \hat{a}$ as $\theta^{\hat{t}}_{\hat{r}}$ is the only non-zero component of the first part and $\theta^{\hat{r}}_{\hat{r}}$ on the second part is zero. Therefore Eq 6.97 simplifies to:

$$\begin{aligned} R^{\hat{t}}_{\hat{r}} &= \vec{d\theta}^{\hat{t}}_{\hat{r}} \\ &= \vec{d}\left(\frac{A'}{AB}\vec{w}^{\hat{t}}\right) \\ &= \vec{d}\left(\frac{A'}{B}dt\right) \\ &= \left(\frac{A'}{B}\right)' dr \wedge dt \\ &= \left(\frac{A''}{B} - \frac{A'B'}{B^2}\right) dr \wedge dt \\ &= \left(\frac{A''}{B} - \frac{A'B'}{B^2}\right) \frac{\vec{w}^{\hat{r}}}{B} \wedge \frac{\vec{w}^{\hat{t}}}{A} \\ &= \frac{1}{B^2}\left(\frac{A''}{A} - \frac{A'B'}{AB}\right) \vec{w}^{\hat{r}} \wedge \vec{w}^{\hat{t}} \end{aligned} \tag{6.98}$$

Similarly for the other components:

$$\begin{aligned} R^{\hat{\theta}}_{\hat{r}} &= \vec{d}\theta^{\hat{\theta}}_{\hat{r}} + \theta^{\hat{\theta}}_{\hat{a}} \wedge \theta^{\hat{a}}_{\hat{r}} \\ &= \vec{d}\theta^{\hat{\theta}}_{\hat{r}} \\ &= \vec{d}\left(\frac{C'}{CB}\right)\vec{w}^{\hat{\theta}} \\ &= \vec{d}\left(\left(\frac{C'}{CB}\right) C d\theta\right) \\ &= \left(\frac{C'}{B}\right)' dr \wedge d\theta \\ &= \frac{1}{B^2}\left(\frac{C''}{C} - \frac{C'B'}{CB}\right) \vec{w}^{\hat{r}} \wedge \vec{w}^{\hat{\theta}} \end{aligned} \tag{6.99}$$

Note that the ϕ equation actually has two equations coming from the $\theta^{\hat{\phi}}$ and $\theta^{\hat{\phi}}_{\hat{\theta}}$ terms:

$$\begin{aligned}
R^{\hat{\phi}}_{\hat{r}} &= \vec{d}\theta^{p\hat{h}i}_{\hat{r}} + \theta^{\hat{\phi}}_{\hat{a}} \wedge \theta^{\hat{a}}_{\hat{r}} \\
&= \vec{d}\theta^{\hat{\phi}}_{\hat{r}} + \theta^{\hat{\phi}}_{\hat{\theta}} \wedge \theta^{\hat{\theta}}_{\hat{r}} \\
&= \vec{d}\left(\frac{C'}{CB}\vec{w}^{\hat{\phi}}\right) + \frac{\cot\theta}{C} C \sin\theta d\phi \wedge \frac{C'}{CB} C d\theta \\
&= \vec{d}\left(\frac{C'}{CB}\vec{w}^{\hat{\phi}}\right) + \cos\theta d\phi \wedge \frac{C'}{B} d\theta \\
&= \vec{d}\left(\frac{C'}{B}\sin\theta d\phi\right) + \cos\theta d\phi \wedge \frac{C'}{B} d\theta \\
&= \left(\frac{C'}{B}\right)' \sin\theta dr \wedge d\phi + \frac{C'}{B}\cos\theta d\phi \wedge d\phi + C\cot\theta \sin\theta d\phi \wedge \frac{C'}{CB} C d\theta \\
&= \left(\frac{C'}{B}\right)' \sin\theta dr \wedge d\phi + \frac{C'}{B}\cos\theta d\theta \wedge d\phi + \cos\theta \frac{C'}{B} d\phi \wedge d\theta
\end{aligned} \tag{6.100}$$

But from the anti-symmetry of the wedge product:

$$\begin{aligned}
R^{\hat{\phi}}_{\hat{r}} &= \left(\frac{C'}{B'}\right)' \sin\theta dr \wedge d\phi + \frac{C'}{B}\cos\theta d\theta \wedge d\phi - \cos\theta \frac{C'}{B} d\theta \wedge d\phi \\
&= \left(\frac{C'}{B}\right)' \sin\theta dr \wedge d\phi \\
&= \frac{1}{B^2}\left(\frac{C''}{C} - \frac{C'B'}{CB}\right) \vec{w}^{\hat{r}} \wedge \vec{w}^{\hat{\phi}}
\end{aligned} \tag{6.101}$$

Similarly the other components are:

$$\begin{aligned}
R^{\hat{t}}_{\hat{\theta}} &= \theta^{\hat{t}}_{\hat{a}} \wedge \theta^{\hat{a}}_{\hat{\theta}} \\
&= \theta^{\hat{t}}_{\hat{r}} \wedge \theta^{\hat{r}}_{\hat{\theta}}
\end{aligned} \tag{6.102}$$

Now we need $\theta^{\hat{r}}_{\hat{\theta}}$ and this is found by:

$$\eta_{\hat{r}\hat{r}}\theta^{\hat{r}}_{\hat{\theta}} = -\theta_{\hat{r}\hat{\theta}} = \theta_{\hat{\theta}\hat{r}} \tag{6.103}$$

and

$$\eta^{\hat{\theta}\hat{\theta}}\theta_{\hat{\theta}\hat{r}} = -\theta^{\hat{\theta}}_{\hat{r}} \tag{6.104}$$

Therefore:

$$\theta^{\hat{r}}_{\hat{\theta}} = -\theta^{\hat{\theta}}_{\hat{r}} \tag{6.105}$$

Substitute Eq 6.105 into Eq 6.102:

$$R^{\hat{t}}_{\hat{\theta}} = -\frac{A'C'}{AB^2C}\vec{w}^{\hat{t}} \wedge \vec{w}^{\hat{\theta}} \tag{6.106}$$

and for $\hat{\phi}$:

$$R^{\hat{t}}_{\hat{\phi}} = -\frac{A'C'}{AB^2C}\vec{w}^{\hat{t}} \wedge \vec{w}^{\hat{\phi}} \tag{6.107}$$

The final one is:

$$\begin{aligned} R^{\hat{\phi}}_{\hat{\theta}} &= d\theta^{\hat{\phi}}_{\hat{\theta}} + \theta^{\hat{\phi}}_{\hat{a}} \wedge \theta^{\hat{a}}_{\hat{\theta}} \\ &= \vec{d}(\cos\theta d\phi) + \frac{C'}{CB}\vec{w}^{\phi} \wedge \left(-\frac{C'}{CB}\vec{w}^{\hat{\theta}}\right) \\ &= -\sin\theta d\theta \wedge d\phi - \frac{C'^2}{C^2B^2}\vec{w}^{\hat{\phi}} \wedge \vec{w}^{\hat{\theta}} \end{aligned} \tag{6.108}$$

putting all this together:

$$R^{\hat{\phi}}_{\hat{\theta}} = \frac{1}{C^2}\left(1 - \frac{C'^2}{B^2}\right)\vec{w}^{\hat{\phi}} \wedge \vec{w}^{\hat{\theta}} \tag{6.109}$$

- Last part is to find the Riemann tensor from this:

$$R^a_b = \frac{1}{2}R^a_{bcd}\vec{w}^c \wedge \vec{w}^d \tag{6.110}$$

Reading of the components from the previous equations:

$$R^{\hat{t}}_{\hat{r}\hat{t}\hat{r}} = -\frac{1}{B^2}\left(\frac{A''}{A} - \frac{A'B'}{AB}\right) \tag{6.111}$$

$$\begin{aligned} R^{\hat{\theta}}_{\hat{r}\hat{\theta}\hat{r}} &= -\frac{1}{B^2}\left(\frac{C''}{C} - \frac{C'B'}{CB}\right) \equiv R^{\hat{\phi}}_{\hat{r}\hat{\phi}\hat{r}} \\ R^{\hat{t}}_{\hat{\theta}\hat{t}\hat{\theta}} &= -\frac{A'C'}{AB^2C} \equiv R^{\hat{t}}_{\hat{\phi}\hat{t}\hat{\phi}} \\ R^{\hat{\phi}}_{\hat{\theta}\hat{\phi}\hat{\theta}} &= \frac{1}{C^2}\left(1 - \frac{C'^2}{B^2}\right) \end{aligned} \tag{6.112}$$

These are the components of the Riemann tensor in the ortho-normal basis. So the Riemann tensor, R, would be given contracting with this ortho-normal basis:

$$R = R^{\hat{a}}_{\hat{b}\hat{c}\hat{d}}\vec{e}_{\hat{a}}\vec{w}^{\hat{b}}\vec{w}^{\hat{c}}\vec{w}^{\hat{d}} \tag{6.113}$$

We usually want Riemann in a coordinate basis, for example:

$$R^t_{rtr} = R^{\hat{t}}_{\hat{r}\hat{t}\hat{r}}B^2 \tag{6.114}$$

where B^2 comes from the two r indicies, where the indices without the hat are coordinate bases (not ortho-normal basis). For the Einstein equations one can simply read of the Ricci tensor by keeping one index upstairs and one index downstairs, as then the weighting factors cancel out:

$$\begin{aligned} R^t_t &= R^{\hat{t}}_{\hat{t}} = \frac{1}{B^2}\left[\frac{A''}{A} - \frac{A'B'}{AB} + \frac{2A'C'}{AC}\right] \\ R^\theta_\theta &= \frac{1}{B^2}\left[\frac{C''}{C} - \frac{C'B'}{CB} + \frac{A'C'}{AC} + \frac{C'^2}{C^2}\right] - \frac{1}{C^2} = R^\phi_\phi \\ R^r_r &= \frac{1}{B^2}\left[\frac{A''}{A} + \frac{2C''}{C} - \frac{B'}{B}\left(\frac{A'}{A} + \frac{2C'}{C}\right)\right] \end{aligned} \tag{6.115}$$

Now these can be used to be build an Einstein tensor:

$$\begin{aligned} G^t_t &= \frac{1}{C^2} - \frac{1}{B^2}\left(\frac{2C''}{C} - \frac{2C'B'}{CB} + \frac{C'^2}{C^2}\right) \\ G^r_r &= \frac{1}{C^2} - \frac{1}{B^2}\left(\frac{2A'C'}{AC} + \frac{C'^2}{C^2}\right) \\ G^\theta_\theta &= -\frac{1}{B^2}\left(\frac{A''}{A} + \frac{C''}{C} + \frac{A'C'}{AC} - \frac{B'}{B}\left(\frac{A'}{A} + \frac{C'}{C}\right)\right) = G^\phi_\phi \end{aligned} \tag{6.116}$$

This example shows how to use the Carton formalism, as an alternative to the way of computing the Christoffel symbols and then computing the Einstein tensor. Both methods have their positive and negative aspects, and it is easier to use one or the other depending on the metric. An important concept in all of theoretical physics, is that of a gauge. In general relativity a gauge just means a change of coordinates. In this example, there are a few choices of which gauge to choose:

- Set $B \equiv 1$: In this case r equals proper radial distance.
- Set $A \equiv \frac{1}{B}$. This is useful in specific cases.
- Set $C \equiv R$: This is known as the *area* gauge, because the area of a 2 sphere is $4\pi r^2$. In this case R is not the proper distance, it is simply giving the area of the 2 sphere.

The idea of a gauge is that choosing a gauge should not affect the equations of motion and all three of the above satisfy this condition. As an example let's choose the area gauge: which gives the Einstein equations for time:

$$\frac{1}{r^2} + \frac{2B'B^{-3}}{r} - \frac{B^{-2}}{r^2} = 8\pi GT \tag{6.117}$$

This can be written as:

$$(rB^{-2})' = 1 - 8\pi GT^0_0 r^2 \tag{6.118}$$

$$B^{-2} = 1 - \frac{2GM(r)}{r} \tag{6.119}$$

This is the static spherically symmetry solution. Where:

$$M(r) \equiv \int 4\pi r^2 T^0_0 dr \tag{6.120}$$

Recall that T^0_0 is like an energy density, which means this equation is the same as the intuitive notion of the m mass inside the sphere. Now lets look at the rr equation which also does not contain a second derivative:

$$\frac{1}{r^2} - \left(1 - \frac{2GM(r)}{r}\right)\left(\frac{2A'}{Ar} + \frac{1}{r^2}\right) = -8\pi GP_r \tag{6.121}$$

where P_r is the radial pressure. This can be re-arranged as:

$$\frac{(A^2)'}{A^2} = \frac{2GM(r) + 8\pi Gr^2 P_r}{r(r - 2GM(r))} \tag{6.122}$$

At present this cannot be solves one needs an equation of state for P_r. Instead one can look at the conservation of energy-momentum:

$$\nabla_a T^{ab} = 0 \tag{6.123}$$

Now we use the ortho-normal basis and try to compute the connection in the ortho-normal basis:

$$\begin{aligned}
\Gamma^{\hat{t}}_{\hat{t}\hat{r}} &= \Gamma^{\hat{r}}_{\hat{t}\hat{t}} = \frac{A'}{AB} \\
\Gamma^{\hat{\theta}}_{\hat{\theta}\hat{r}} &= \frac{C'}{CB} = -\Gamma^{\hat{r}}_{\hat{\theta}\hat{\theta}}
\end{aligned} \tag{6.124}$$

The $\hat{r}$ component of Eq 6.123:

$$\begin{aligned}
\nabla_a T^{ar} &= \partial_{\hat{r}} P + \Gamma^{\hat{a}}_{\hat{a}\hat{r}} P_r + \Gamma^{\hat{r}}_{\hat{a}\hat{b}} T^{\hat{a}\hat{b}} \\
&= \frac{1}{B} P_r' + \left(\frac{A'}{AB} + \frac{2}{rB} \right) P_r + \left(\frac{A'}{AB} \rho - \frac{2}{rB} \right) P_r \\
&= \frac{1}{B} \left(P_r' + \frac{A'}{A} \left(P_r + \rho \right) \right)
\end{aligned} \tag{6.125}$$

Thus the conservation of energy and momentum gives an equation that links P_r and ρ.

CHAPTER 7

Space-time structures

1. Penrose diagrams

We have already seen the Kruskal coordinates in Chapter 5, these will now be used to describe the causal structure of space-time via the tool of Penrose diagrams. The space-time used will be the Schwarzschild solution to Einstein equations. The Schwarzschild solution has so far only been used without a cosmological constant. If we add a cosmological constant, Λg_{ab}, then in Eq 6.118, the T_0^0 is equal to a constant and therefore this equation can be inegrated easily:

$$rB^{-2} = 1 - Ar^2 \tag{7.1}$$

which implies:

$$B^{-2} = 1 \underbrace{- \frac{2GM}{r}}_{T_\alpha} \underbrace{- \frac{\Lambda r^2}{3}}_{T_\beta} = A^2 \tag{7.2}$$

If $\Lambda > 0$, the model universe is called the de-Sitter universe. T_α and T_β are both negative, therefore:

$$A^2 \to 0 \;\text{ as} \begin{cases} r \approx 2GM \\ r \approx \sqrt{\frac{3}{\Lambda}} \end{cases} \tag{7.3}$$

if M is sufficiently small. We already knew that when $g_{tt} \to 0$, one has a coordinate singularity, in the Scwarschild space-time. By a change of coordinate system, into the Kruskal coordinates, we have shown that there is a boundary in space-time between events that an asymptotic observer can see and those that they can't see. Therefore, if $g_{tt} = 0$, it is associated with a horizon.

In this case $r \approx 2GM$, will be an event horizon of a black hole and $r = \sqrt{\frac{3}{\Lambda}}$, is called a cosmological horizon, because in de-Sitter space, the space is expanding so rapidly that any individual observer see's that there is a boundary between events in the universe that they can monitor and those they cannot. The event horizon are clearly visible in this coordinate systems, but de-Sitter space actually has several coordinates systems that are commonly used to describe the space-time. If $M = 0, \Lambda = 3$:

$$ds^2 = (1 - r^2)dt^2 - (1 - r^2)^{-1}dr^2 - r^2 d\Omega_2^2 \tag{7.4}$$

This is known as the *static* patch as the metric is time-dependent. Now one change coordinates:

$$ds^2 = d\tau^2 - \cosh^2 \tau d\Omega_3^2 \tag{7.5}$$

In these coordinates the metric looks like a Lorentzian sphere (hyperboloid, shown by the cosh term). This called the global patch, as they cover the whole of the de-Sitter space. In the static patch, as $r \to 1$, there is a coordinate singularity. We have seen that this is removed in the Kruskal coordinates. Therefore the static patch does not cover the whole of de-Sitter space.

Another possible coordinate system used is:

$$ds^2 = dT^2 - e^{2T} \underbrace{d\vec{x}^2}_{\mathbb{R}^3} \tag{7.6}$$

The metric represents a flat universe, with exponential inflation. This is typically used in cosmology. If $\Lambda < 0$, the space is called anti de-Sitter (adS) space. The metric is now:

$$ds^2 = A^2 dt^2 - \frac{dr^2}{A^2} - r^2 d\Omega_2^2 \tag{7.7}$$

where:

$$A \equiv 1 - \frac{2GM}{r} + k^2 r^2 \qquad k^2 \equiv -\frac{\Lambda}{3} \tag{7.8}$$

the k has been put in to emphasise that the last term is now positive. This has only one horizon, which is at $\approx 2GM$. Black holes in adS are used very frequently in applications of string theory because of the famous adS-CFT correspondence (which I don't know anything about yet). Note that the 1 in Eq 7.8, comes from the fact that we were dealing with a 2 sphere. If one changes the $d\Omega_2^2$ to some general 2D space:

$$d\Omega_2^2 \to dx_\kappa^2 \tag{7.9}$$

where:

$$\kappa = \begin{cases} 1 \ \rightarrow \ \mathbb{S}^2 & \text{Spherical, closed} \\ 0 \ \rightarrow \ \mathbb{R}^2 & \text{flat} \\ -1 \ \rightarrow \ \mathbb{H}^2 & \text{Hyperboloid} \end{cases} \tag{7.10}$$

Then the 1 in the previous equation, becomes a κ, therefore $\kappa = 1$ for $\Lambda < 0$ or de-Sitter otherwise, the g_{tt} would be negative and that's not consistent with the metric signature. However in adS, κ can take on all three values $1, -1, 0$ as the $k^2 r^2$ term is positive (hence can compensate for the other two negative terms). Therefore in adS one has spherical spherical planar (flat) and hyperbolic black holes. Therefore black holes only have to be spherical in vacuum equations with positive Λ.

1.1. Causal structure of space-time. The Schwarzschild metric is:

$$ds^2 = \left(1 - \frac{2GM}{r}\right) dt^2 - \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 - r^2 d\Omega_2^2 \tag{7.11}$$

Here we follow the Kruskal coordinates using slightly different notation than previously done. To remove the coordinate singularities, Kruskal coordinates are used:

$$\begin{aligned} r* &= \int \frac{dr}{1 - \frac{2GM}{r}} \Rightarrow dr* = \frac{dr}{1 - \frac{2GM}{r}} \\ &= r + \ln(r - 2GM) \end{aligned} \tag{7.12}$$

The idea here is that including a factor of r^*, we get a single factor in front of the t and r components. This means that propagation of mass-less particles occurs along lines of $dt = \pm dr$. So the first step is to move to coordinates in which the characteristic curves are just straight lines. The coordinate system is still singular, but it hints at how to go beyond the singularity. Define:

$$\begin{aligned} U &\equiv -2GMe^{-\frac{(t-r^*)}{4GM}} \\ V &\equiv 2GMe^{\frac{t+r^*}{4GM}} \\ dUdV &\equiv UV(dt^2 - dr*^2) - \frac{1}{4(GM)^2} \end{aligned} \tag{7.13}$$

This also gives:

$$UV = -(2GM)^2 e^{\frac{r^*}{2GM}} \tag{7.14}$$

The dependence of the time coordinates in Eq 7.14 factors out:

$$e^{\frac{r^*}{2GM}} = e^{\frac{r}{2GM}} \frac{(r - 2GM)}{2GM} \tag{7.15}$$

therefore Eq 7.14 as $r \to 2GM$. The metric in these new coordinates is:

$$ds^2 = \frac{4}{2GMr} dUdV e^{-\frac{r}{2GM}} - r^2(U,V)d\Omega_2^2 \tag{7.16}$$

Notice that r is now written as a factor of U, V. Now as $r \to 2GM$:

$$ds^2 = C(\text{constant})dUdV - D(\text{constant})d\Omega_2^2 \tag{7.17}$$

Therefore the metric is perfectly regular. Since $r = 2GM$ is an horizon, lets try to identify where it is in these coordinates. When r is a constant, then UV has to be constant. Similarly when t is a constant $\frac{U}{V}$ is a constant. At $r = 2GM$, the r^* becomes negatively infinite, $UV \to 0$. Now to see what happens at $r = 0$ in these coordinates:

$$r = 0 \Rightarrow UV = (2GM)^2 \tag{7.18}$$

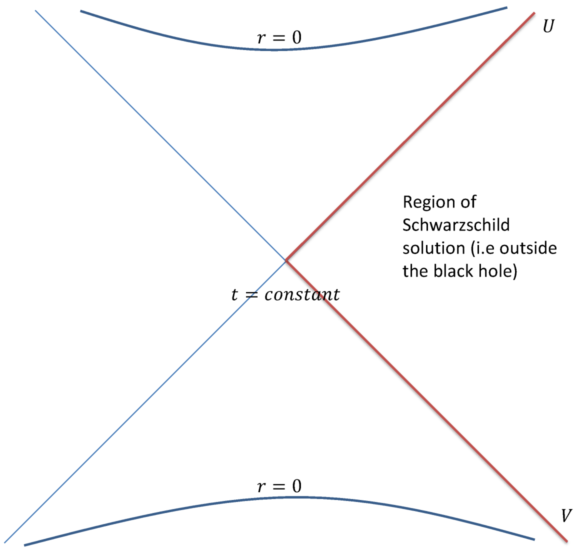


FIGURE 34. Black hole horizon shown by red line

This is what the maximally extended space-time looks like, $r = 0$ is a real singularity (i.e it cannot be removed by a coordinate transformation). The U and V coordinates run between $-\infty$ and $+\infty$.

The penrose diagram shrinks this into a single smaller picture, which is component (i.e can be fully drawn, of course nobody can draw an axis from $-\infty$ and $+\infty$). To do this, define new coordinates:

$$\begin{aligned} p &= \arctan\left(\frac{V}{2GM}\right) \\ q &= \arctan\left(\frac{U}{2GM}\right) \end{aligned} \tag{7.19}$$

This will bring the limits down from $-\infty$ to $+\infty$ to $[-\frac{\pi}{2}, \frac{\pi}{2}]$. The coordinates take on values are:

$$V = 0 \Rightarrow p = 0$$
$$V = 2GM \Rightarrow p = \frac{\pi}{4} \tag{7.20}$$

Note that combining the two definitions we get an identity:

$$\tan(p+q) = \frac{\tan p + \tan q}{1 - \tan p \tan q} = \frac{U+V}{(2GM)^2 - UV} \tag{7.21}$$

when $UV \to 2GM^2$, which is the singularity; therefore:

$$\tan(p+q) \to \infty \tag{7.22}$$

i.e $p + q = \frac{\pi}{2}$.

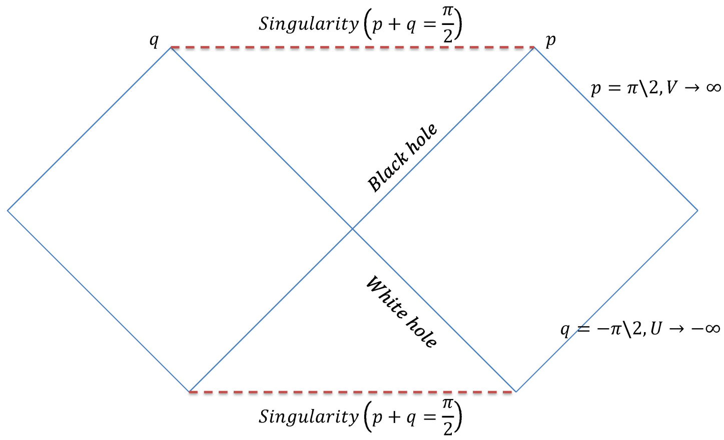


FIGURE 35. Penrose diagram for Schwarzschild solution

This sketch represents the causal structure of the Schwarzschild solution. These coordinates identify the horizon as a null surface, which we already know it is because it is a boundary between what can be seen and what cannot. Imagine a photon that always moves at $r = 2GM$ always trying to move out but being pulled back in by gravity at the same rate. Due to time-symmetry, one also gets a surface under time reflection. In the maximally extended Schwarzschild space-time, one gets both of these surfaces corresponding to a black hole and a white hole.

1.2. Penrose diagram of Minkowski space-time. The Minkowski metric in spherical coordinates is the usual:

$$ds^2 = dt^2 - dr^2 - r^2 d\Omega_2^2 \tag{7.23}$$

In flat space there are no singularities in the metric and one can simply define:

$$u \equiv t - r$$
$$v \equiv t + r \tag{7.24}$$

If r is positive (which it is as r is a radial coordinate), then:

$$U - V > 0 \tag{7.25}$$

The space-time is diagram is shown below:

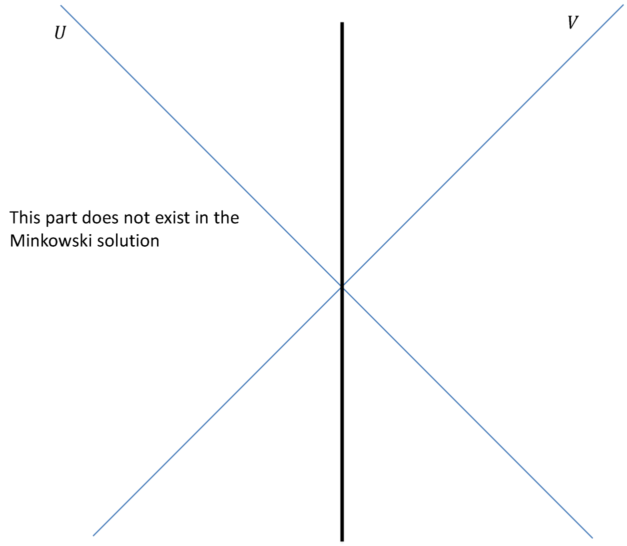


FIGURE 36. Space-time diagram for the Minkowski metric, showing that the left side of the black line does not exist.

Now using the the same definition for p, q in Eq 7.19, we get the causal structure:

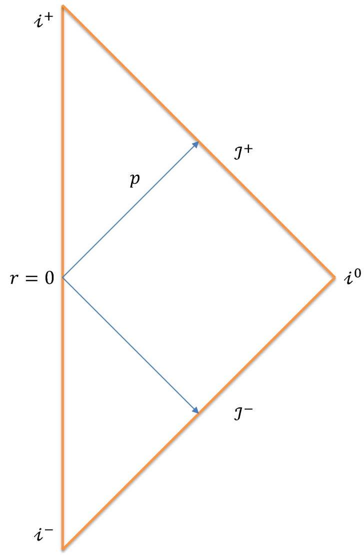


FIGURE 37. Causal structure of Minkowski metric

$r = 0$ is not a horizon or singularity, as a light ray going through $r = 0$. will simply come out the other side with same positive value for r, so in the diagram, it will appear to simply bounce of the $r = 0$ point. Points i^-, i^+, i^0 and lines $\mathscr{I}^+, \mathscr{I}^-$ are boundaries. Any observer has to eventually end up at i^+ and it is called a *future time like infinity*. Any observer must have come from i^-, and this is called *past time like infinity*. The only way to get to i^0 is to go along a line that is more that 45°, to the vertical, i.e a space-like line therefore i^0 is called a space-like infinity.

- $\mathscr{I}^+$ is called a "scri plus" and the only way to get there is to take a null like trajectory in the future. So its called a future null infinity.

- $\mathscr{I}^-$ is called a "scri minus" and is the past null infinity.

Notice that the R.H.S of both Figure 37 and 35 is the same, therefore the causal structure of both space-times is the same (even though the metric is not), however the Schwarzschild metric does asymptotically approach the Minkowski space-time). The L.H.S we see a different picture. The $r = 0$ has been stretched out and instead of $r = 0$, there is an additional null surface, which is an event horizon.

EXAMPLE 23. In Schwarzschild de-Sitter space the Penrose diagram looks like:

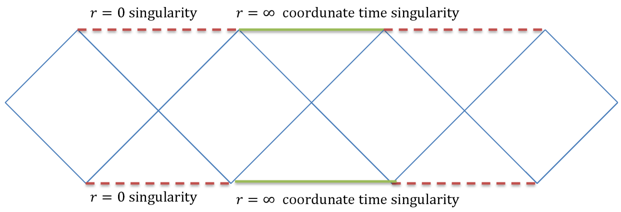


FIGURE 38. Penrose diagram for Schwarzschild de-Sitter space

EXAMPLE 24. In Schwarzschild adS space the Penrose diagram looks like:

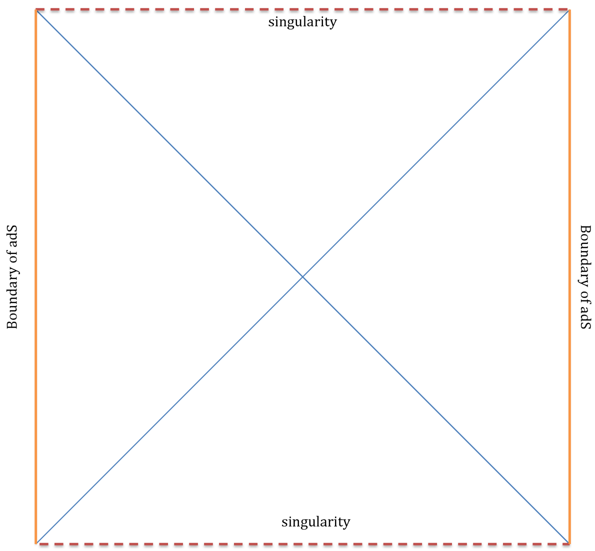


FIGURE 39. Penrose diagram for Schwarzschild adS space

2. Euclidean "black holes"?

It is a mathematically interesting question to ask what happens when we take the Schwarschild metric and make it Euclidean:

$$ds^2 = \left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\Omega_2^2 \tag{7.26}$$

The metric also has $R = 0$ (Ricci scalar). What happens as $r \to 2GM$; in the Minkowski like metric, $r \to 2GM$ is not a singularity, it is an event horizon. So one is lead to ask, weather there is a way to make this metric non-singular aswell at $r = 2GM$. Ignoring the solid angle term (as it simply adds a numerical factor); we want to see what happens to the term:

$$ds_2^2 \equiv \left(\frac{r - 2GM}{r} dt^2 + \frac{r}{r - 2GM} dr^2\right) \tag{7.27}$$

Define:

$$\rho^2 \equiv \lambda(r - 2GM) \tag{7.28}$$

Then:

$$2\rho d\rho = \lambda dr \tag{7.29}$$

and the new metric is:

$$ds_2^2 = \frac{\rho^2}{\rho^2 + 2GM\lambda} dt^2 + \left(\frac{\rho^2 + 2GM\lambda}{\rho^2}\right) \frac{4\rho^2}{\lambda^2} d\rho^2 \tag{7.30}$$

As $r \to 2GM$, $\rho \to 0$ (i.e 2GM $\lambda >> \rho$) :

$$ds^2 \approx \frac{p^2}{2m\lambda} dt^2 + \frac{8GM}{\lambda} d\rho^2 \tag{7.31}$$

Choose $\lambda = 8GM$, then define new coordinates:

$$\theta \equiv \frac{t}{4GM} \tag{7.32}$$

therefore:

$$ds_2^2 = \rho^2 d\theta^2 + d\rho^2 \tag{7.33}$$

i.e $\rho \to 0 \in \mathbb{R}^2$

This is near the origin of the plane in plane polar coordinates, if θ has periodicity 2π. Therefore the Euclidean metric can be made regular at $r = 2GM$, provided the θ coordinate is periodic in 2π. However, from the definition of θ in Eq 7.32, this implies that:

$$\text{period of } t = 8\pi GM \tag{7.34}$$

In thermal field theory periodic Euclidean time corresponds to a finite temperature:

$$T \approx \frac{1}{\beta} \frac{\hbar c^3}{8\pi G m k_B} \tag{7.35}$$

where the constants have been put in to get the correct units. β represents the periodicity of Euclidean time. So it seems like the black hole has some finite temperature. Also note that larger mass would decrease the temperature and hence eventually the black hole would decrease its temperature by evaporating. This is of course what Hawking has already shown; that Black holes will radiate via Hawking radiation.

CHAPTER 8

Gravitational field theory

The most successful theory in physics is the Standard model of particle physics which is a Quantum field theory. One of the major problems in physics (if not the major problem) is that of finding a quantum filed theory for gravity. However to do that, the first step is to describe general relativity as a classical field theory.

The natural place to start is to look for an action that describes gravity. This action is known as the Einstein-Hilbert action (sometimes just known as the Einstein action). The idea is to find the corresponding Lagrangian, which, when varied w.r.t to the fields will provide the Einstein field equations of motion. This means one needs to use the scalar gravitational field.

Since the action is an integral over the Lagrangian and the Lagrangian has fields in it which are defined over a manifold, one needs to understand the concept of integrating over a manifold aswell.

1. Integrating over a manifold

In field theory, one usually has a Lagrangian density, $\mathcal{L}$. In flat space, this is integrated over to give the full Lagrangian, $\mathbb{L}$:

$$\int d^3x\mathcal{L} = \mathbb{L} \tag{8.1}$$

One cannot simply take this expression and lift it to an arbitrary manifold. This is because changing coordinates in a manifold gives a factor of a Jacobian matrix:

$$d^4Y \to det\left(\frac{\partial y}{\partial x}\right)d^4x \tag{8.2}$$

This is not a tensor, its more like a tensor density (as seen for the ϵ symbol), therefore multiplying it by $\sqrt{g}$, gives a quantity that transforms as a tensor. Thinking about this in the context of the Lagrangian density, we are integrating over a volume element and the volume should have some information about the underlying the metric, because a metric is what gives information about distance on a manifold. In fact talking about a volume is meaningless without a metric to define distances on a manifold. Thus it is natural to take:

$$\sqrt{|g|}d^x \tag{8.3}$$

under a coordinate transformation, $\sqrt{g}$, get a factor of $\frac{\partial x}{\partial y}$, then under a coordinate transformation one gets:

$$\sqrt{|g|}d^4x \to \sqrt{|g(y)|}d^4y \tag{8.4}$$

Thus we have to write down a co-variant volume element, which means we need to how the right information about the metric, which is why a $\sqrt{g}$ is inserted. Now we need to vary the action w.r.t the physical observables. The metric represents the distances, so the obvious choice is to vary the action w.r.t to the metric, so we need to know how the determinant of the metric varies. To do this, we need to use the identity:

$$det(M) = exp(tr(log(M))) \tag{8.5}$$

where M is an arbitrary matrix. Varying this:

$$\delta(det(M)) = \delta(tr(log(m)))det(M) = tr(M^{-1}\delta M)det(M) \tag{8.6}$$

Using this identity for g:

$$\begin{aligned}
\delta\sqrt{-g} &= \frac{1}{2}\frac{1}{\sqrt{-g}}\delta(-g) \\
&= \frac{1}{2}\frac{1}{\sqrt{-g}}(-g)g^{ab}\delta g_{ab} \\
&= \frac{1}{2}\sqrt{-g}g^{ab}\delta g_{ab} \\
&= -\frac{1}{2}\sqrt{-g}g_{ab}\delta g^{ab}
\end{aligned} \tag{8.7}$$

Another result from this identity is:

$$\begin{aligned}
\partial_a\sqrt{-g} &= \frac{1}{2}\sqrt{-g}g^{cd}g_{cd,a} \\
&= -\sqrt{-g}\Gamma^d_{ad}
\end{aligned} \tag{8.8}$$

2. Einstein action

Consider a mass less scalar field:

$$\mathcal{L}_\phi = \frac{1}{2}(\partial\phi)^2 \tag{8.9}$$

Therefore the action is:

$$\begin{aligned}
S_\phi &= \int d^4x\sqrt{-g}\frac{1}{2}(\partial\phi)^2 \\
&= \frac{1}{2}\int d^4x\sqrt{-g}\,\underbrace{\partial_a\phi\partial_b\phi g^{ab}}_{(\partial\phi)^2}
\end{aligned} \tag{8.10}$$

In general, the action is varied by:

$$\delta S_\phi = \int d^4x\sqrt{-g}\left[\frac{\delta S}{\delta\phi}\delta\phi + \frac{\delta S}{\delta g^{ab}}\delta g^{ab}\right] \tag{8.11}$$

Computing the terms:

$$\begin{aligned}
\delta S_\phi &= \frac{1}{2}\delta\int d^x\sqrt{-g}(\partial\phi)^2 \\
&= \frac{1}{2}\delta\int d^4x\sqrt{-g}\partial_a\phi\partial_b\phi g^{ab} \\
&= \frac{1}{2}\int d^4x\delta(\sqrt{-g}\partial_a\phi\partial_b\phi g^{ab}) \\
&= \frac{1}{2}\int d^4x(\delta\sqrt{-g})\partial_a\phi\partial_b\phi g^{ab} + (\sqrt{-g})\delta(\partial_a\phi\partial_b\phi)g^{ab} + (\sqrt{-g})\partial_a\phi\partial_b\phi(\delta g^{ab}) \\
&= \int d^4x\sqrt{-g}\left[\partial_a\phi\partial_b\delta\phi g^{ab} + \frac{1}{2}\partial_a\phi\partial_b\phi\delta g^{ab} - \frac{1}{4}g_{ab}(\partial\phi)^2\delta g^{ab}\right]
\end{aligned} \tag{8.12}$$

To get Eq 8.12 into the form of Eq 8.11, we have to integrate by parts. The first term is:

$$\begin{aligned}\frac{\delta S}{\delta\phi} &= -\frac{1}{\sqrt{-g}}\left[\partial_b\sqrt{-g}g^{ab}\partial_a\phi\right] \\ &= -\partial_b\nabla^b\phi - \frac{1}{\sqrt{-g}}(\partial_b\sqrt{-g})\nabla^b\phi \end{aligned} \tag{8.13}$$

But $\partial_b\sqrt{-g} = \sqrt{-g}\,\Gamma^d_{bd}$, therefore:

$$\begin{aligned}\frac{\delta S}{\delta\phi} &= -\partial_b\nabla^b\phi - \Gamma^d_{bd}\nabla^b\phi \\ &= -\Box\phi \end{aligned} \tag{8.14}$$

where:

$$\Box \equiv \partial_b\nabla^b + \Gamma^d_{bd}\nabla^b \tag{8.15}$$

This is the wave equation in curved space-time. The only difference is that the wave operator is the curved space wave operator i.e it has a curvature term, Γ, in it. This is the scalar equation of motion and it is as expected. The variation of the scalar field action w.r.t the metric:

$$\frac{\delta S_\phi}{\delta g^{ab}} = \frac{1}{2}(\partial_a\phi\partial_b\phi - \frac{1}{2}(\partial\phi)^2 g_{ab}) \tag{8.16}$$

This is like the energy momentum tensor of free space:

$$\frac{\delta S_\phi}{\partial g^{ab}} = \frac{1}{2}T_{\phi ab} \tag{8.17}$$

If one can construct a gravitational action such that when it is varied w.r.t the gravitational field, i.e the metric, g, then we get the Einstein tensor. If one varies the matter Lagrangian, w.r.t, g, we expect to get the energy momentum tensor.

Now we want to construct a gravitational Lagrangian. The obvious thing to use in the Lagrangian that requires a scalar is the Ricci scalar. A Lagrangian will, in general have a kinetic energy term and a potential energy term. The kinetic energy term will involve first order derivatives of the quantity being varied. However we have already shown that to first order all derivatives of curvature can be set to zero, thus it is good that R already has first order derivatives of g, thus differentiating again means only second order derivatives of the metric are present, which cannot be set to zero by coordinate transformation.

The first ingredient required is the variation of the Ricci scalar:

$$\delta R = \delta(R_{ab}g^{ab}) \tag{8.18}$$

This is just the trace of the Ricci tensor:

$$\delta R = \underbrace{\delta R_{ab}g^{ab}}_{T_1} + \underbrace{R_{ab}\delta g^{ab}}_{T_2} \tag{8.19}$$

T_2 is the first part of the Einstein tensor. In the integral, one also has a $\sqrt{-g}$ (which came from demanding that the coordinate transformations did not change the integral) which gives $-\frac{1}{2}Rg_{ab}$. Therefore the T_2 with the integral gives the Einstein tensor. Any contribution from T_1 will change the structure of the Einstein equations.

Lemma 1. Variation of R_{ab} given by Palatini's Lemma:

$$\delta R_{ab} = \nabla_c\delta T^c_{ab} - \nabla_b\delta T^c_{ca} \tag{8.20}$$

So the variation in R_{ab} is given by the covariant derivative of the discrepancy in the connection, which comes when one goes from $g_{ab} \to g_{ab} + \delta g_{ab}$. The difference of two connections is a tensor, therefore Eq 8.20 is actually a tensor.

There is an easy way to motivate this lemma using *normal coordinates*. Normal coordinates are ones which have no curvature locally, i.e $\Gamma = 0$. In other words, coordinates transformations are made in the neighborhood of a point b such that the connection vanishes at point b:

$$x \to x' \in \text{neighborhood of } b\{|\Gamma^a_{bc} \equiv 0 \; \forall \; a, b, c \text{ at } b \tag{8.21}$$

This is because under a coordinate transformation a connection can be chosen such that all second and first derivatives of g can be set to zero at b, meaning in the neighborhood of b, Γ is nearly zero and thus $d\Gamma$ is non-zero:

$$\begin{aligned} \delta\Gamma^c_{ab,c} &= \frac{1}{2} g^{cd} \left[\delta g_{da,bc} + \delta g_{db.ac} - \delta g_{ab,dc}\right] \\ &= -\frac{1}{2} \left(\nabla_c \nabla_b \delta g^{ce}\right) g_{ea} - \frac{1}{2} (\nabla_c \nabla_a \delta g^{ce}) g_{bc} + \frac{1}{2} g_{ac} g_{bd} \Box \delta g^{cd} \end{aligned} \tag{8.22}$$

In normal coordinates, all Γ's are zero locally and the Ricci tensor is given by:

$$R_{ab} = R^c_{acb} = \Gamma^c_{ab;c} - \Gamma^c_{ac;b} \tag{8.23}$$

Thus the variation of the Ricci tensor is:

$$\delta R_{ab} = \delta\Gamma^c_{ab,c} - \delta\Gamma^c_{ac;b} \tag{8.24}$$

Hence:

$$\begin{aligned} g^{ab}\delta R_{ab} &= g^{ab}(\delta\Gamma^c_{ab,c} - \delta\Gamma^c_{ac,b}) \\ &= g^{ab}\delta\Gamma^c_{ab,c} - g^{ab}\delta\Gamma^c_{ac,b} \\ &= g^{ab}\left(-\frac{1}{2}(\nabla_c\nabla_b\delta g^{ce})g_{ea} - \frac{1}{2}(\nabla_c\nabla_a\delta g^{ce})g_{be} + \frac{1}{2}g_{ac}g_{bd}\Box\delta g^{cd}\right) \\ &- g^{ab}\left(-\frac{1}{2}(\nabla_b\nabla_c\delta g^{bc})g_{ea} - \frac{1}{2}(\nabla_e\nabla_a\delta g^{be})g_{ce} + \frac{1}{2}g_{ae}g_{cd}\Box\delta g^{bd}\right) \\ &= -\nabla_a\nabla_b\delta g^{ab} + g_{ab}\Box\delta g^{ab} \end{aligned} \tag{8.25}$$

This looks like a total derivative. Now one can look at the variation of the action:

$$\begin{aligned} \delta S_g &= C\int d^4x\sqrt{-g}\left[R_{ab}\delta g^{ab} - \frac{1}{2}Rg_{ab}\delta g^{ab}\right] - \nabla_a\nabla_b\delta g^{ab} + g_{ab}\Box g^{ab} \\ &= C\int d^4x\sqrt{-g}G_{ab}\delta g^{ab} + \underbrace{C\int_{dM} d^3x\sqrt{g_3}\int\left[-\nabla_b\delta g^{ab} + \nabla^a g_{cd}\delta g^{cd}\right]\hat{n}_a}_{T_1} \end{aligned} \tag{8.26}$$

where C is an arbitrary constant, $\hat{n}_a$ is the normal vector to the surface and g_3 is a metric of a 3D sub-manifold on the 4D manifold. Term T_1 is the boundary term that comes from integrating by parts and has been written as a 3D surface integral as supposed to a 4D volume integral by an extension of Gauss's divergence theorem to 4D. By equating the expression of the Einstein action obtained for the stress energy tensor in Eq 8.17 to the first term in Eq 8.26:

$$C \equiv -\frac{1}{16\pi G} \tag{8.27}$$

where G has been inserted to get the correct units from the Einstein field equations. This gives the Einstein action:

$$S_E = -\frac{1}{16\pi G}\int d^4x\sqrt{-g}R \tag{8.28}$$

For now the boundary term is ignored. Using this action, we know how to incorporate general relativity into a field theory.

3. Beyond the Einstein action

3.1. Scalar-tensor theories. Consider an extension Lagrangian (density):

$$\mathcal{L}_{extension} = \phi R \qquad \text{where } \phi = \frac{1}{16\pi G} \tag{8.29}$$

Here Newton's constant is not taken to be constant, instead it form a time varying scalar field. This is known as *Brans-Dickie gravity.* It was first introduced to satisfy Mach's principle in general relativity, which is how to know weather one is rotating in an empty universe.

Having a scalar in front of the action has a more general range of applicability in terms of scalar-tensor theories of gravity. The equation of motion is:

$$g^{ab}\phi\delta R_{ab} = -\phi\nabla_a\nabla_b\delta g^{ab} + \phi g_{ab}\Box\delta g^{ab} \tag{8.30}$$

Now when we integrate by parts, we get new contributions to the equations of motion as one is picking up derivatives of ϕ:

$$g^{ab}\phi\delta R_{ab} = -\delta g^{ab}\nabla_a\nabla_b\phi + \delta g^{ab}\Box\phi g_{ab} \tag{8.31}$$

Einstein tensor becomes:

$$\phi G_{ab} - \nabla_a\nabla_b\phi + g_{ab}\Box\phi \tag{8.32}$$

In Brans-Dickie theory, the full Lagrangian is:

$$\mathcal{L} = -\phi R + \omega(\partial\phi)^2 + 16\pi\mathcal{L}_{matter} \tag{8.33}$$

ω is a dimensionless coupling constant, known as the *Dickie* constant. This is the full Lagrangian with a kinetic term plus the usual matter Lagrangian. The equations of motion are:

$$\phi G_{ab} = 8\pi T_{ab} + \nabla_a\nabla_b\phi - g_{ab}\Box\phi + \frac{\omega}{\phi}\left(\partial_a\phi\partial_b\phi - \frac{1}{2}g_{ab}(\partial\phi)^2\right) \tag{8.34}$$

This is obtained by varying the action w.r.t metric and is an analog of Einstein's field equations. To elaborate further I will quote what Brans and Dickie wrote in there paper in 1961[**6**]:

> " *The left side of Eq. (11)* ***[which refers to Eq 8.34]*** *is completely familiar and needs no comment. Note that the first term on the right is the usual source term of general relativity, but with the variable gravitational coupling parameter ϕ^{-1}. Note also that the second term is the energy-momentum tensor of the scalar field, also coupled with the gravitational coupling ϕ^{-1}. The third term is foreign and results from the presence of second derivatives of the metric tensor in R in Eq. (6)****[which refers to the variation on the action of Brans-Dickie theory]****. These second derivatives are eliminated by integration by parts to give a divergence and the extra terms.*"

If we vary w.r.t ϕ:

$$R + 2\omega\frac{\Box\phi}{\phi} = \omega\frac{(\Box\phi)^2}{\phi^2} \tag{8.35}$$

combining both of these equations of motion gives:

$$\frac{\Box\phi}{\phi} = \frac{8\pi T}{3+2\omega} \tag{8.36}$$

where $T = T^a + a$ is the trace of the stress-energy tensor. Scalar fields are used many times in the action of general relativity. Note that matter couples minimally to g:

$$(\partial\phi)^2 = g^{ab}\partial_a\phi\partial_b\phi \tag{8.37}$$

But the gravitational action is not Einstein-Hilbert thus the gravitational equations are more complex. To simplify these one often changes coordinates to make the gravitational equations like Einstein, but the matter is non-minimally couple i.e one rescales the metric to cancel out the scalar function ϕ. This is done via a conformal transformation of coordinates:

$$\hat{g}_{ab} = \Omega^2(x)g_{ab} \tag{8.38}$$

where $\Omega(x)$ is a local rescale of the metric. This doesn't change angles, but changes the lengths. After this transformation, the new connection terms are:

$$\hat{\Gamma}^a_{bc} = \Gamma^a_{bc} + \Omega^{-1}(\Omega_{,c}\delta^a_b + \Omega_{,b}\delta^a_c - g^{ac}\Omega_c g_{bc}) \tag{8.39}$$

and:

$$\hat{R}_{bd} = R_{bd} + (2-D)\Omega^{-1}\nabla_b\nabla_d\Omega - g_{bd}\Omega^{-1}\Box\Omega + 2(D-2)\Omega^{-2}\nabla_b\Omega\nabla_d\Omega - (D-3)g_{bd}\Omega^{-2}(\nabla\Omega)^2 \tag{8.40}$$

where D is the dimension of the manifold. This is called the *Einstein frame.* So we have seen here how one of the scalar-tensor theories of gravity works. This is a very interesting area of research as it seems that in the Einstein gravity one has to include new terms such as dark energy, dark matter to obtain correct results from the Einstein field equations. However, it is possible that Einstein gravity does not actually work on the largest of scales and the simplest extension to Einstein's theory are these scalar-tensor theories.

4. Non-perturbative field theory

Typically, when one studies a filed theory, one looks at perturbative field theory. Generally the aim is to quantise the field theory, by looking at excitations around the vacuum. One can approach Einstein's gravity in the same way, i.e take a vacuum solution to the field and look at small perturbations around the field in the vacuum state.

However this approach cannot be applied to things like black holes, as even though the mass of a black hole is totally arbitrary, it has an event horizon, that makes it different to a small perturbation. Hawking's initial calculation for the evaporation of a black hole was done by putting field theory on the classical background of a black hole[**7**].

4.1. Domain wall. First we will consider what the non-perturbed solution is in the context of field theory and then add gravity. Consider a scalar field:

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \underbrace{\frac{\lambda}{2}(\phi^2 - \eta^2)}_{V(\phi)} \tag{8.41}$$

Here the potential has two distinct vacuum states (minimum's). This is an example of spontaneous symmetry breaking in field theory, in the sense that this potential has a symmetry around the vertical axis. But if the theory needs to be quantised one has to pick a vacuum, $\pm\eta$, and then look at the excitations around that vacuum. The equation of motion for ϕ is:

$$\Box\phi + 2\lambda\phi(\phi^2 - \eta^2) = 0 \tag{8.42}$$

And suppose we want to look at a solution where one side of the universe has the vacuum at $-\eta$ and the other side of the universe has a positive η. In this potential, we have not included gravity, therefore one can say:

$$\phi \to \pm\eta \qquad \text{as } z \to \pm\infty \tag{8.43}$$

as we expect to get the usual wave equation $\Box\phi = 0$ (z is just a Cartesian coordinate).

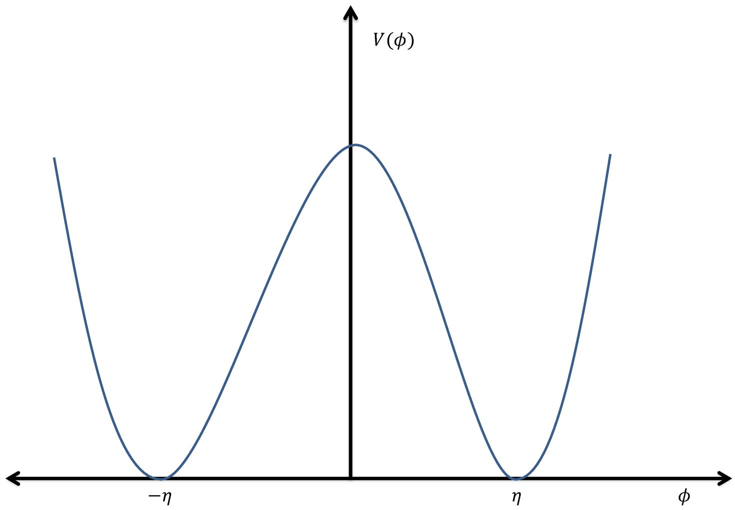


FIGURE 40. Potential of scalar field, showing two vacuum states

The equation of motion is simply:

$$-\phi'' + 2\lambda\phi(\phi^2 - \eta^2) = 0 \tag{8.44}$$

Integrating once by multiply by ϕ':

$$-\frac{1}{2}\phi'^2 + \frac{\lambda}{2}(\phi^2 - \eta^2) = C(\text{constant}) \tag{8.45}$$

Which has a solution:

$$\phi = \eta \tanh(\sqrt{\lambda}\eta z) \qquad C = 0 \tag{8.46}$$

This solution interpolates between these two vacuum states. By a perturbative solution, one usually means something which is intuitively close to the vacuum solution, so we might imagine that we can get to the perturbative solution from the vacuum through a set of field configurations, which may or may not solve the equations of motion themselves, but there is a continuous route and each of these field configurations has a finite energy. It is possible to go from this solution to the true vacuum, by moving the regions of ϕ with opposite sign of η, to regions with the same sign of η. However, since the regions of space have a finite energy, this means moving an infinite region of space from one configuration to another would cost an infinite amount of energy. Therefore there is a barrier between either of the two vacuum solutions as seen in Figure 41. tanh is well approximated by ± 1, unless its argument is $\mathcal{O}(1)$. Which means we can associate a width with this "kink" of order $\frac{1}{\sqrt{\lambda}\eta}$, which is representative of the mass of the scalar excitations $\frac{1}{\sqrt{\lambda}\eta} \approx m_\phi^{-1}$. Now if we look at the energy momentum tensor:

$$\begin{aligned} T_{\mu\nu} &= \phi_{,\mu}\phi_{,\nu} - g_{\mu\nu}\mathcal{L}_\phi \\ &= \lambda\eta^4 sech^4(\sqrt{\lambda}\eta z)\delta^z_\mu\delta^z_\nu - g_{\mu\nu}\left(-\frac{\phi'^2}{2} - V\right) \end{aligned} \tag{8.47}$$

The δ symbols in the first term shows that it is only non-zero if we are looking at a z index. This can be simplified to:

$$T_{\mu\nu} = \lambda\eta^4 sech^4(\sqrt{\lambda}\eta z)\left[\delta^z_\mu\delta^z_\nu + g_{\mu\nu}\right] \tag{8.48}$$

If we look at $\mu = \nu = z$, then $T^z_z = 0$ as $(\delta^z_\mu\delta^z_z + g_{zz}) = 0$. But when $\mu\nu \neq z$, the δ's give zero and we get:

$$T^0_0 = T^x_x = T^y_y = \lambda\eta^4 sech^4(\sqrt{\lambda}\eta z) \tag{8.49}$$

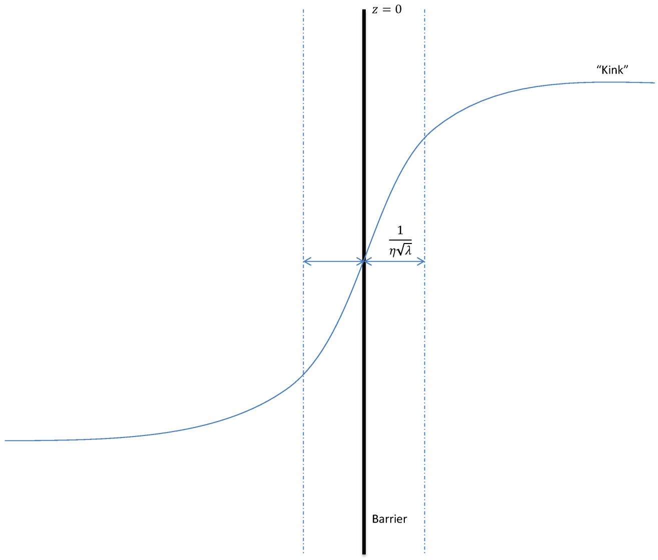


Figure 41. Barrier between two vacuum solutions

This is sort off a cosmological constant on the sub-manifold of x, y, t dimensions of the 4D Minkowski metric. The curve, $\eta^4 sech^4(\sqrt{\lambda}\eta z)$, can be approximated by a Dirac delta function:

$$\approx \eta^2(z)(\lambda\eta^2) \tag{8.50}$$

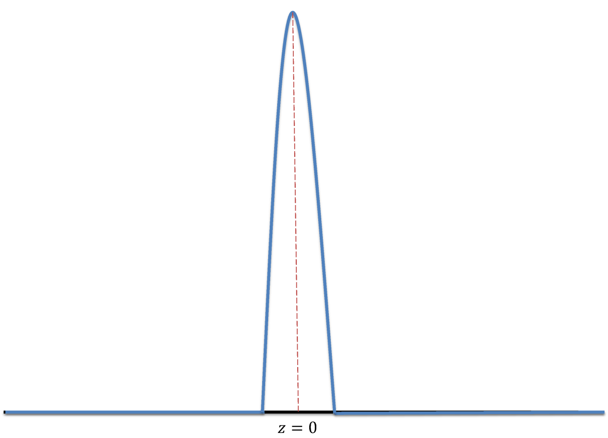


Figure 42. Energy density of this pseudo cosmological constant

We find the energy by integrating through the kink:

$$\begin{aligned}\int T_0^0 dz &= \int \lambda\eta^4 sech^4(\sqrt{\lambda}\eta z)dz \\ &= \frac{4}{3}\sqrt{\lambda}\eta^3 \qquad (8.51)\end{aligned}$$

This gives a finite value of this energy, but its on energy density (i.e energy per unit area in this case) and $z = 0$ has an infinite area, therefore the overall energy is infinite. Typically in gravity, we look at things as isolated systems on average. Even though this solution to the energy-momentum tensor has infinite energy, let's ask how does it gravitates. Just because something has an infinite area does not mean that it will be singular, or not have a good gravitational solution.

When thinking about solutions to Einstein's equations, one always has to think about a choice of gauge/coordinates. The kink has t, x, y Lorentz symmetry, which suggests constant curvature t, x, y space-time. This can now be used to guide the search for an appropriate gauge. We have done the same thing for the FRW metric, when defining it's coordinates by assuming isotropy and homogeneity.

In this case, we consider a metric:

$$ds^2 = A^2(z)\gamma_{\mu\nu}dx^\mu dx^\nu - dz^2 \qquad (8.52)$$

Which is very similar to FRW metric, except for the variation of the metric comes from z, not t. This is sometimes known as a *warped compactification*. In the sense that $\gamma_{\mu\nu}$ has a dimensions (D-1) and z adds an extra dimension, where D is the dimensionality of the full space-time. Using the Cartan formalism:

$$\begin{aligned}\vec{w}^{\hat{z}} &= \vec{d}z \\ \vec{w}^{\hat{a}} &= A(z)\vec{w}_0^{\hat{a}} = A(z)e_\mu^{\hat{a}}\vec{d}x^\mu \qquad (8.53)\end{aligned}$$

The one forms of the first part of the metric, i.e ignoring the dz' terms are simply the $A^2(z)\gamma_{\mu\nu}$. The $\gamma_{\mu\nu}$ is left completely arbitrary at the moment:

$$\vec{d}\vec{w}^{\hat{a}} = \frac{A'}{A}\vec{w} : \wedge\vec{w}^a - A\vec{\theta}_0^a b \wedge \vec{w}_0^b \qquad (8.54)$$

$$\vec{\theta}_b^a = \vec{\theta}_{0b}^a \qquad (8.55)$$

$$\vec{\theta}_z^a = \frac{A'}{A}\vec{w}^a \qquad (8.56)$$

By keeping the indices in the (D-1) dimensional subspace one can read off, the background connection one forms. By looking at a mixed components we pick up a factor of $\frac{A'}{A}$. The curvature two form can be found from Cartan's second equation:

$$R_b^a = \vec{d}\vec{\theta}_{cb}^a + \underbrace{\vec{\theta}_{0c}^a \wedge \vec{\theta}_{0b}^c}_{R_{0b}^a} + \vec{\theta}_c^a \wedge \vec{\theta}_b^z \qquad (8.57)$$

The second part can be written as:

$$\vec{\theta}_{0c}^a \wedge \vec{\theta}_{0b}^c = \vec{\theta}_z^a \wedge \vec{\theta}_z^c \eta_{cb} \qquad (8.58)$$

Therefore the curvature two form is:

$$\begin{aligned}R_b^a &= R_{0b}^a + \vec{\theta}_z^a \wedge \vec{\theta}_z^c \eta_{cb} \\ &= \frac{1}{2}R_{0bcd}^a \vec{w}_0^c \wedge \vec{w}_0^d + \left(\frac{A'}{A}\right)^2 \eta_{bc}\vec{w}^a \wedge \vec{w}^c (8.59)\end{aligned}$$

The components are:

$$\begin{aligned} R^a_z &= \frac{A''}{A}\vec{w}^z \wedge \vec{w}^a - \frac{A'}{A}\vec{\theta}^a \wedge \vec{w}^b + \vec{\theta}^a_{0b} \wedge \vec{\theta}^b_z \\ &= \frac{A''}{A}\vec{w}^z \wedge \vec{w}^a - \frac{A'}{A}\left(\frac{A}{A'}\right)\vec{\theta}^a_{0b} \wedge \vec{\theta}^b_z + \vec{\theta}^a_{0b} \wedge \vec{\theta}^b_z \\ &= \frac{A''}{A}\vec{w}^z \wedge \vec{w}^a \end{aligned} \tag{8.60}$$

Therefore:

$$R^a_{bcd} = \frac{1}{A^2}R^a_{0\,bcd} + \left(\frac{A'}{A}\right)^2 [\delta^a_c \eta_{bd} - \delta^a_d \eta_{bc}] \tag{8.61}$$

$$R^{\mu\nu}_{\lambda\tau} = \frac{1}{A}R_0^{\mu\nu}{}_{\lambda\tau} + \left(\frac{A'}{A}\right)^2 [\delta^\mu_\lambda \delta^\nu_\tau - \delta^\mu_\tau \delta^\nu_\lambda] \tag{8.62}$$

Looking at the z terms:

$$\begin{aligned} R^{\mu z}_{\nu z} &= \frac{A''}{A}\delta^\mu_\nu \\ R^z_z &= \frac{A''}{A}(D-1) \\ R^\mu_\nu &= \frac{A''}{A}\delta^\mu_\nu + \frac{R^\mu_{0\,\nu}}{A^2} + (D-2)\left(\frac{A'}{A}\right)^2 \delta^\mu_\nu \end{aligned} \tag{8.63}$$

if $D = 4$:

$$\begin{aligned} R^z_z &= \frac{4A''}{A} \\ R^\mu_\nu &= \left(\frac{A''}{A} + 2\left(\frac{A'}{A}\right)^2 + \frac{\kappa}{l^2 A^2}\right)\delta^\mu_\nu \end{aligned} \tag{8.64}$$

where l is some arbitrary length scale. We want to solve the Einstein equations:

$$G^y_y = G^x_x = G^0_0 = \frac{\kappa}{l^2 A^2} - \frac{A'^2}{A^2} - \frac{2A''}{A} = 8\pi G(V + \frac{1}{2}\phi'^2) \tag{8.65}$$

$$G^z_z = \frac{3\kappa}{l^2 A^2} - \frac{3A'^2}{A^2} = 8\pi G(V - \frac{1}{2}\phi'^2) \tag{8.66}$$

$$\Box\phi = -\frac{1}{A^3}\partial_z A^3 \partial_z \phi = -2\lambda\phi(\phi^2 - \eta^2) \tag{8.67}$$

In flat space Eq 8.65 is a $sech^4$ function and Eq 8.66 = 0 (from Eq 8.47). Suppose:

$$8\pi G\eta^2 << 1 \tag{8.68}$$

Then:

$$8\pi GV = \underbrace{\mathcal{O}(8\pi G\eta^3)}_{\epsilon(<<1)} \times \underbrace{\lambda\eta^2}_{\text{Area}^{-1}} \tag{8.69}$$

This is stating that $\eta << M_p$, where M_p, i.e η is much less than the Planck scale, which is a reasonable assumption. The R.H.S of Eq 8.65 and 8.66 has dimensions of Area^{-1}, which has the same units as Eq 8.69, thus $\lambda\eta^2$ sets a length (energy) scale and $\mathcal{O}(8\pi G\eta^2)$ gives the gravitational interaction and we are going to look at a situation in which the gravitational interaction is small.

Therefore are can solve the equations of motion, perturbation. Note that the does not mean that we are close to flat space necessarily. It is saying that one can write the A function as:

$$A(z) = 1 + \mathcal{O}(\epsilon) \tag{8.70}$$

and:

$$\phi = \eta \tanh(\sqrt{\lambda}\eta z) + \mathcal{O}(\epsilon) \tag{8.71}$$

Away from the wall, $T^{\mu}_{\nu} = 0$, therefore $G^{\mu}_{\nu} = R^{\mu}_{\nu} = 0$ aswell. Therefore from Eq 8.65:

$$A'' = 0 \tag{8.72}$$

$$(A')^2 = \frac{\kappa}{l^2} \tag{8.73}$$

Therefore κ must be positive and the $\gamma_{\mu\nu}$ is a de Sitter type solution i.e an analog of a sphere:

$$A = 1 \pm \frac{z}{l} \tag{8.74}$$

This is the general shape of the solution away from the wall. Now we have to relation this solution to the perturbative terms. So now we use the perturbed solution and put that into the R.H.S of the Einstein equations:

$$A'' = -\epsilon\frac{\lambda\eta^2}{2} sech^4(\sqrt{\lambda}\eta z) \tag{8.75}$$

Integrate:

$$A' = -\epsilon\sqrt{\epsilon}\eta(\tanh(\sqrt{\lambda}\eta z) - \frac{1}{3}\tanh^3(\sqrt{\lambda}\eta z)) \tag{8.76}$$

Set:

$$A'(0) \equiv 0 \tag{8.77}$$

by looking for a symmetric solution. This implies:

$$l = \frac{3}{2\epsilon\sqrt{\lambda}\eta} \tag{8.78}$$

so if ϵ is small, which it is, then the length l, scale is large. Integrating Eq 8.76 again, we get the solution:

$$A(z) = 1 - \epsilon\left[\frac{2}{3}\log(\cosh(\sqrt{\lambda}\eta z)) - \frac{1}{6}sech^2(\sqrt{\lambda}\eta z) + \frac{2}{3}\log z\right] \tag{8.79}$$

For large $\sqrt{\lambda}\eta z$; the solution is approximately:

$$A(z) \approx 1 - \frac{2}{3}\epsilon(\pm\sqrt{\lambda}\eta z) \tag{8.80}$$

Which has the correct form of Eq 8.70. This looks correct, but one must check what happens when $z = \pm l$, as it appears to be an event horizon or singularity. One would not expect it to be a singularity as the metric was flat space. What has actually happened, is that we have written out flat space in a strange coordinate system:

$$ds^2 = \left(1 - \frac{z}{l}\right)^2 (dt^2 - \cosh^2 t d\Omega_2^2) - dz^1 \tag{8.81}$$

A constant curvature space time (de-Sitter space) written in global coordinates has a compact spatial section. This is hidden away in the γ as we did not specify what γ was. Therefore the wall actually gives two space-times as seen in Figure 43.

To summarise; the idea behind discussing domain walls, was to solve the Einstein equations with some matter i.e non-zero $T_{\mu\nu}$. Non-perturbative field theory was introduced to show that one could have a stable solution, that was distinct from the vacuum solution as it took an infinite energy to go from the solution to the vacuum solution.

The domain wall itself has one direction removed from it, therefore it is a 3D sub-manifold of the 4D space-time. As far as the gravity was concerned, for the pure wall space-time, most of the space was vacuum except for that narrow region in the wall (as shown in Figure 42), which was approximated by a Delta function.

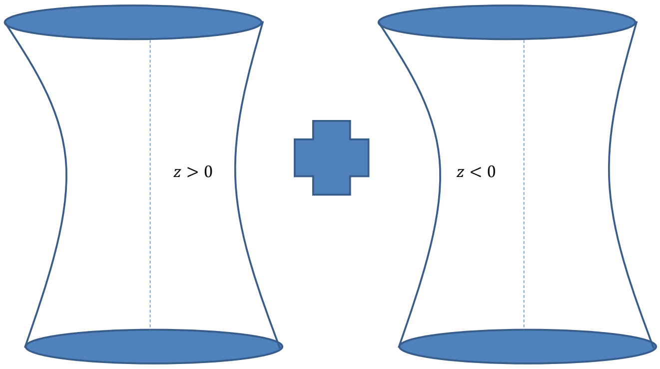


FIGURE 43. Non-perturbative solution to Einsteins equations gives rise to two distinct de-Sitter space-times

5. Sub-manifolds

The reason behind discussing sub-manifolds comes directly from the example of a domain wall approximated by a (D-1) sub-manifold. A natural extension would be to look at sub-manifolds in (D-2), (D-3) etc.

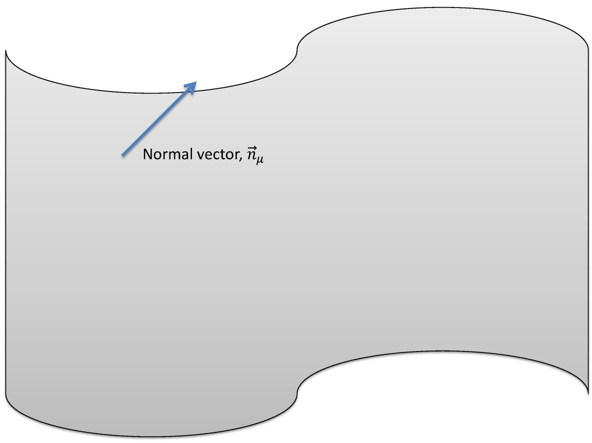


FIGURE 44. Sub-manifold Σ with normal vector, $\vec{n}_\mu$

5.1. Gauss-Codazzi formalism. Imagine a sub-manifold Σ, of a manifold M, of dimension (p+1):

$$\Sigma \subset M \tag{8.82}$$

The +1 can be thought of as a time dimensions and p being the spatial dimension. There is no real mathematical motivation for this way of labeling dimensions. It's simply because when sub-manifolds are used in general relativity, they will be in the context of physical objects, hence a time component is needed. Σ is going to be a differentiable manifold in its own right, of dimension (p+1) and it is also a subset of an underlying manifold M.

The co-dimension of Σ, n, is defined as:

$$D - (p+1) = n \tag{8.83}$$

which is the complement of $dim(\Sigma)$, i.e the number of independent dimensions which do not lie in Σ but are present in M. Hence there exist n linearly independent normal vectors to the sub-manifold Σ. Imagine taking internal coordinate charts on Σ, which has a set of basis vectors, σ_a, then normal vectors will follow:

$$\sigma_a . \vec{n}_\mu = 0 \tag{8.84}$$

where σ_a are the basis vectors on the vector space that is used to parametrize Σ and the dot represents the usual inner product. This is a statement that there is no component of $\vec{n}_\mu$ along any of the basis vectors, hence are perpendicular. Since the basis vectors can be defined as directional derivatives, this statement can be re-written as:

$$\vec{n}_\mu \frac{\partial x^\mu}{\partial \sigma^A} = 0 \qquad x^\mu(\sigma^A) \tag{8.85}$$

the $x^\mu(\sigma^A)$ defines the coordinate functions of Σ within of M.

5.2. 1st fundamental form. Since Σ is a sub-manifold of M, it must inherit some of the structure of the manifold M. We have a very good intuitive picture of this, for example, consider a 2D surface like a paper sheet embedded on a 3D surface like a sphere. The paper sheet, which is initially flat, will take on a spherical shape on the surface of the sphere, as long as the paper sheet is *not rigid*.

DEFINITION 24. The 1st fundamental form of Σ is given by:

$$h_{ab} \equiv g_{ab} + \sum_{i=1}^{b} (-)^i n_{ia} n_{ib} \tag{8.86}$$

where:

$$i = \begin{cases} -1 \text{ for time-like } n \\ +1 \text{ for space-like } n \end{cases} \tag{8.87}$$

h is defined as the first fundamental form of Σ and is like taking the metric on the manifold M and projecting out all normal directions. Therefore h is a projection operator, acting within the tangent space of M, which takes a general vector and projects it down to its component parallel to the sub-manifold Σ. The minus sign being in the definition has no mathematical motivation. Again, we are using the fact that we know that space-time manifolds have one time dimension and the time and space dimensions differ in sign.

h_{ab} is equation to the metric that the sub-space inherits from M. The *induced* metric is defined by:

$$\gamma_{AB} = g_{\mu\nu} \frac{x^\mu}{\partial \sigma^A} \frac{\partial x^\nu}{\partial \sigma^B} \tag{8.88}$$

γ and h are physically the same thing and contain the same information. Mathematically these are different objects as h_{ab} has ab indicies and these refer to the tangent space of the manifold, M, restricted to a sub-space of this manifold. γ_{AB} on the other hand refers to Σ as a manifold in its own right, and thus the metric is intrinsic to the sub-manifold.

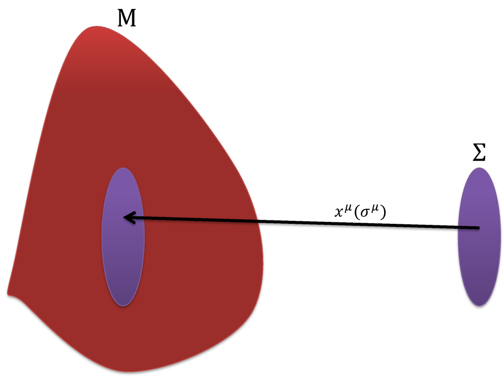


FIGURE 45. Sub-manifold projected onto a manifold

$\frac{\partial x^\mu}{\partial \sigma^A}$ is a map from the cotangent space of Σ to the cotangent space of M.

5.3. 2nd fundamental form.

DEFINITION 25. The 2nd fundamental form is also known as the extrinsic curvature and is defined as:

$$K_{iab} \equiv \nabla_c n_{id} h^c_{(a} h^d_{b)} \tag{8.89}$$

where ∇_c represents the covariant derivative. Therefore what we are doing is taking the covariant derivatives of the normals and projecting them parallel to the sub-manifold using the first fundamental form.

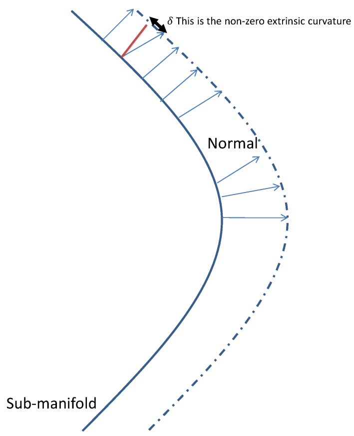


FIGURE 46. Tangent vectors to the sub-manifold varying from point to point, showing extrinsic curvature.

We can recognize extrinsic curvature when surface is not flat. For example a sheet of paper can be folded in on itself so that it appears curved, its the space in which it sits, that is curved. As one can simply unravel the sheet of paper and it would appear flat. This type of curvature is extrinsic curvature.

EXAMPLE 25. Consider a sub-manifold of $\mathbb{R}^3$, defined by a cylinder of radius a:

$$x^2 + y^2 = a^2 \qquad \subset \mathbb{R}^3 \tag{8.90}$$

A unit normal, $\vec{n}_a$, would be just $(\cos\theta, \sin\theta, 0)$. In terms of h_{ab} (first fundamental form):

$$\begin{aligned} h_{ab} &= \delta_{ab} - n_a n_b \\ &= \begin{pmatrix} \sin^2\theta & -\sin\theta\cos\theta & 0 \\ \sin\theta\cos\theta & \cos^2\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(Cylindrical coordinates)} \\ &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & a^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(Spherical polar coordinates)} \end{aligned} \tag{8.91}$$

In spherical polar coordinates it becomes obvious that one has projected from $\mathbb{R}^3$ to some 2D sub-manifold since the h_{ab} contains a block matrix:

$$\begin{pmatrix} a^2 & 0 \\ 0 & 1 \end{pmatrix} \tag{8.92}$$

with all other elements zero, show that even though the manifold is 3D, the object represented by the 1st fundamental form is a sub-manifold of a 3D space. Therefore a natural set of coordinates for this surface is:

$$\sigma^A = \{\theta, z\} \tag{8.93}$$

The intrinsic metric is:

$$\gamma_{AB} = \begin{pmatrix} a^2 & 0 \\ 0 & 1 \end{pmatrix} \tag{8.94}$$

The extrinsic curvature is:

$$K_{ab} = \nabla_a n_b = -\Gamma^r_{ab} \qquad \text{(in polar coordinates)} \tag{8.95}$$

The normal in polar coordinates is given by $\frac{\partial}{\partial r}$:

$$K_{ab} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{8.96}$$

Thus the extrinsic curvature is non-zero.

Another way of writing K_{iAB} is:

$$\begin{aligned} K_{iAB} &= \frac{\partial x^\mu}{\partial \sigma^A}\frac{\partial x^\nu}{\partial \sigma^B}\nabla_\mu n_{i\nu} \\ &= -n_{i\nu}\frac{\partial x^\mu}{\partial \sigma^A}\nabla_\mu\left(\frac{\partial x^\nu}{\partial \sigma^B}\right) \\ &\equiv -n_{i\nu} D_A D^A x^\nu \end{aligned} \tag{8.97}$$

The i index is there to show that in principle there are many normals. What can happen, if the co-dimension is greater than 1, the normals can vary of "twist", through the sub-manifold as shown in Figure 46.

Thus we have to be able to link the normals as they vary from point to point in the sub-manifold.

DEFINITION 26. This is done by *normal fundamental forms*:

$$\beta_{u,j} = n_{,\nu}\nabla_\mu n_j^\nu \tag{8.98}$$

This takes the covariant derivative of a normal and instead of projecting it parallel to the manifold, it is projected perpendicular to the manifold. Therefore for co-dimension greater than 1, the tangent space splits and can have a non-trivial structure itself. We will assume that n_i^μ extend geodesically out of Σ into M and $\nabla_j n_j = 0$, i.e the normal covariant derivative of the normals is zero.

5.4. Gauss equation for curvature. The Gauss equation relates the curvature of M and Σ:

$$^{(p+1)}R^a_{bcd} = \underbrace{^{(D)}R^{a'}_{b'c'd'}h^a_{a'}h^{b'}_b h^{c'}_c h^{d'}_d}_{T_1} - \underbrace{\sum_{i=1}^{n}(-)^i[K^a_c K_{ibd} - K^a_{id}K_{ibc}]}_{T_2} \tag{8.99}$$

where R^a_{bcd} is the Riemann curvature of the sub-manifold and $R^{a'}_{b'c'd'}$ is the Riemann curvature of the manifold. T_1 gives the *intrinsic* curvature; we project down the manifold intrinsic curvature. The T_2 term shows that one also needs some extra parts that are related to the extrinsic curvature, to get the overall curvature.

EXAMPLE 26. Let's try to reproduce this for a co-dimension 1, sub-manifold. Use Riemann identity:

$$2^{(p+1)}\nabla^{(p+1)}_{[c}\nabla_{d]}V^a =^{(p+1)} R^a_{bcd}V^b \tag{8.100}$$

where:

$$V^b h^a_b = V^a \tag{8.101}$$

Here we are imagining a vector, which is parallel to Σ (it lies in the subspace of the tangent space parallel to Σ) and then using the Riemann identity acting on this vector.

$$\begin{aligned} ^{(p+1)}R^a_{bcd}V^b &= h^a_e h^f_c h^p_d \nabla^{(p+1)}_f \nabla_p V^e - T(c \leftrightarrow d) \\ &= h^a_e h^f_c h^p_d \nabla_f (h^q_p h^e_r \nabla_q V^r) - T(c \leftrightarrow d) \end{aligned} \tag{8.102}$$

the $T(c \leftrightarrow d)$ term denotes the first term, with c and d indices swapped, this comes from the anti-symmetry of the Riemann identity. Here we are taking the manifold covariant derivative and projecting down with every free index, in the first line this has been done for the first index.

Now we act on all the terms in the bracket with the covariant derivative, ∇_f:

$$\begin{aligned} ^{(p+1)}R^a_{bcd}V^b &= \underbrace{h^a_e h^f_c h^q_d h^e_r \nabla_f \nabla_q V^r 0}_{T_1} - T_1(c \leftrightarrow d) \\ &+ \underbrace{h^a_e h^f_c h^p_d (\nabla_f h^q_p) h^e_r \nabla_q V^r}_{T_2} - T_2(c \leftrightarrow d) \\ &+ \underbrace{h^a_e h^f_c h^f_d h^q_p (\nabla_f h^e_r)\nabla_q V^r}_{T_3} - T_3(c \leftrightarrow d) \end{aligned} \tag{8.103}$$

Substitute in for the expression of the first fundamental form from Eq 8.86. The first part is the metric and for no torsion, the covariant derivative of the metric is zero. The remaining terms give:

$$h_p^q = \delta_p^q = \delta_p^q (-)^i n_p n^q \tag{8.104}$$

this is Eq 8.86 without the metric term (as it's covariant derivative is zero) and so on for the rest of the h's aswell. However by contracting a normal with a first fundamental form we get zero by definition:

$$n_p h_d^p = 0 \tag{8.105}$$

Therefore acting with the covariant derivative on the h_q^p terms will give a product of two terms and then contacting with the 1st fundamental forms that remain, will give zero, thus Eq 8.103 simplifies[1]:

$$\begin{aligned}
{}^{(p+1)}R^a_{bcd}V^b \quad &= \quad \underbrace{h_e^a h_c^f h_d^q h_r^e \nabla_f \nabla_q V^r}_{T_1} - T_1(c \leftrightarrow d) \\
(-)^i \quad & \underbrace{h_e^a h_c^f h_d^p h_r^e n^q (\nabla_f n_p) \nabla_q V^r}_{T_2} - T_2(c \leftrightarrow d) \\
(-)^i \quad & \underbrace{h_e^a h_c^f h_d^q n_r \nabla_f n^e \nabla_q V^r}_{T_3} - T_3(c \leftrightarrow d)
\end{aligned} \tag{8.106}$$

In the first line we are anti-symmetrizing of c and d which means we are also anti-symmetrizing on f and q and thus in the derivatives in front of V^r. Therefore we can use the Riemann identity. In the second derivative of V, which is zero, as V is defined to be parallel to the manifold. In the last line, there is a covariant derivative of n, with index f and e and there are terms $h_e^a h_c^f$ outside, which means we will get an extrinsic curvature. The $n_r \nabla_q V^r$ term follows:

$$n_r \nabla_q V^r = -V^r \nabla_q n_r \tag{8.107}$$

because V and n are also orthogonal:

$$\begin{aligned}
{}^{(p+1)}R^a_{bcd}V^b &= h_e^a h_c^f h_d^q h_r^{e(D)} R^r_{qpfq} V^p - (-)^i h_e^a h_e^f h_d^q n_r (\nabla_f n^e) \nabla_q V^r - T(c \leftrightarrow d) \\
&= h_e^a h_c^f h_d^q h_r^{e(D)} R^r_{qpfq} V^p - (-)^i K_c^a V^r K_{dr} + (-)^i K_d^a V^r K_{cr}
\end{aligned} \tag{8.108}$$

Re-labeling:

$$ {}^{(p+1)}R^a_{bcd}V^b = \underbrace{h^a_{a'} h_c^{c'} h_d^{d'} R^{a'}_{bc'd'} V^b}_{T_\alpha} (-)^i \underbrace{[K_d^a K_{bc} - K_c^a K_{bd}] V^b}_{T_\beta} \tag{8.109}$$

Which reproduces the Gauss equation in 1D. So the Riemann curvature of the sub-manifold, is composed of the Riemann curvature of the manifold and the extrinsic curvature coming from the way in which the sub-manifold is embedded in the manifold.

Going back to the cylinder, it has a extrinsic curvature, but no intrinsic curvature (as $\mathbb{R}^3$ has no curvature as it is just the Euclidean space), therefore T_1 is zero. However since the co-dimensionality is 1 and the K's are anti-symmetrised, they will cancel leaving:

$$R^a_{bcd} \equiv 0 \tag{8.110}$$

for a cylinder.

[1]If simplifies is a word that can be used for this expression!

6. Applying Gauss-Codazzi formulation

6.1. Israel equations. Israel equations are a set of equations of motion for a co-dimension 1 sub-manifold, when viewed as being a physical object. Recall the domain wall depended on one direction, which we choose to be z. To begin with consider the for dimensional Riemann tensor, with some normal vector n:

$$R_{abcd}n^b n^d \equiv n^d(\nabla_c\nabla_d - \nabla_d\nabla_c)n_a \tag{8.111}$$

Here we assume n is geodesic, therefore $\nabla_n n = 0$:

$$R_{abcd}n^b n^d = K_c^d K_{da} - \nabla_n K_{ac} \tag{8.112}$$

The normal covariant derivative of K, can be replaced by a Lie derivative:

$$\begin{aligned} R_{abcd}n^b n^d &= -(\mathcal{L}_n K_{ac} - \nabla_a n^d K_{dc} - \nabla_c n^d K_{da}) - K_c^d K_{da} \\ &= -\mathcal{L}_n K_{ac} + K_a^d K_{dc} \end{aligned} \tag{8.113}$$

So we have re-written the Riemann tensor with two normal components in terms of extrinsic curvature. Eventually we want to discuss the Einstein equations, which can be re-written as:

$$R_{ab} = 8\pi G(T_{ab} - \frac{1}{2}Tg_{ab}) \tag{8.114}$$

foe the domain wall:

$$T_{ab} = \sigma(z)h_{ab} \tag{8.115}$$

i.e there was no normal component of the stress-energy tensor and all the parallel components to be the same (proportional to the metric). This is what motivates the re-writing of the Einstein field equations, in the form Eq 8.114. Recall the Gauss equation in terms of a manifold in 4D with a sub-manifold in 3D:

$${}^{(3)}R_{abcd} ={}^{(4)} R_{a'b'c'd'}h_a^{a'}h_b^{b'}h_c^{c'}h^{d'}d - (K_{ac}K_{bd} - K_{ad}K_{bc}) \tag{8.116}$$

The normal has been taken to be space-like, therefore has only one component and only one K. The first term on the R.H.S is a contraction between the 4D Riemann tensor, with the induced metric h, but h is just $g + nn$, therefore the Ricci tensor is:

$${}^{(3)}R_{bd} ={}^{(4)} R_{b'd'}h_b^{b'}h_d^{d'} + \underbrace{{}^{(4)}R_{abcd}n^a n^c}_{T_1} - KK_{bd} + K_{bc}K_d^c \tag{8.117}$$

This is the reason behind writing Eq 8.113 in that form is that now it can be substituted into T_1 in the equation above:

$${}^{(3)}R_{bd} = \underbrace{{}^{(4)}R_{b'd'}h_b^{b'}h_d^{d'}}_{T_2} + \mathcal{L}_n K_{bd} + 2K_{bc}K_d^c - KK_{bd} \tag{8.118}$$

T_2 can be substituted for from the Einstein equations in the form of Eq 8.114:

$${}^{(3)}R_{bd} = 8\pi G(T_{a'b'}h_a^{a'}h_b^{b'} - \frac{1}{2}Th_{ab}) - \mathcal{L}_n K_{bd} + 2K_{bc}K_d^c - KK_{bd} \tag{8.119}$$

So far, the Gauss equation has been to analyse the parallel components of the Einstein equations. If T_{ab} is sharply localised (at z), then one can approximate it by a delta function:

$$T_{ab} \approx \delta(z)S_{ab} \tag{8.120}$$

This was done for the wall, where $\delta(z) \approx^4 z$. Thus defines S_{ab}:

$$S_{ab} \equiv \int_{-\infty}^{\infty} dz T_{ab} \tag{8.121}$$

The Ricci tensor, ${}^{(3)}R_{bd}$, is an intrinsic curvature, of the sub-manifold, therefore it does not have a delta function(i.e not sharply localised). ${}^{(3)}R_{bd}$, is regular as the width of the object goes to zero. K must also be regular, as K is a property of Σ. However, the projection of the Einstein equations parallel to the wall, give terms involving the stress-energy tensor, which is strongly localised as defined by Eq 8.120. The last term remaining, $\mathcal{L}_n K_{bd}$, must also be sharply peaked in order to cancel with the stress-energy tensor.

If we integrate Eq 8.119 from one side of the wall to the other, the K terms will give a finite value times the width of the wall. Whereas the T and $\mathcal{L}K$ terms will give something $\mathcal{O}(1)$, therefore ${}^{(3)}R_{bd}$, will be of $\mathcal{O}$(width of wall):

$$8\pi G(S_{ab} - \frac{1}{2}Sh_{ab}) = \int \mathcal{L}_n K_{ab} = K^+_{ab} - K^-_{ab} \tag{8.122}$$

This looks like some form of Einstein equation. In states that the extrinsic curvature has to vary, by looking at the difference across the wall, and that difference has to be equal to the combination of the stress-energy tensor. Taking the trace of the equation above:

$$-4\pi GS = K^+ - K^- \tag{8.123}$$

Which can be re-written as:

$$\Delta K_{ab} - \Delta K h_{ab} = 8\pi G S_{ab} \tag{8.124}$$

these are the Israel equations. These are the analog of the Einstein equations (for co-dimension 1), integrated out over the strongly localised source of energy/matter. These equations describe a highly localised physical object that is a hyper-surface. Normal points into the bulk on (+) region and out of the bulk on the (-) region. The normal is continuous over the bulk field.

Let's check these equations by re-deriving the domain wall solution, which looked like they had a horizon. Let's look at a hyperboloid in Minkowski space:

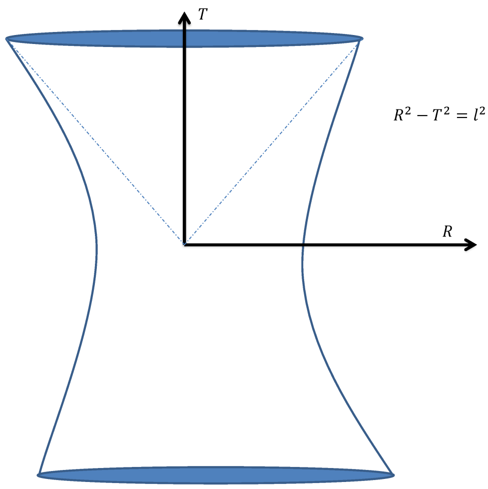


Figure 47. Hyperboloid in Minkowski space.

The hyperboloid in Minkowski space can be parametrised by:

$$X^\mu = (l \sinh\left(\frac{\tau}{l}\right), l\cosh\left(\frac{\tau}{l}\right), \theta, \phi) \tag{8.125}$$

Thus Σ can be written as:

$$\Sigma : x^\mu(\sigma^A), \sigma^A = \{\tau, \theta, \phi\} \tag{8.126}$$

This defines the hyperbolic surface. Now we want to check what happens when we apply Israels equations to the surface (and check that the result is the same). The normal to the surface is:

$$n_\mu = \pm(-\sinh\left(\frac{\tau}{l}\right), \cosh\left(\frac{\tau}{l}\right), 0, 0) \tag{8.127}$$

This is a normal to the surface. The two signs are present to represent the inward and outward normals respectively. Note that:

$$\gamma_{AB} d\sigma^A d\sigma^B = d\tau^2 - l^2\cosh^2\left(\frac{\tau}{l}\right) d\Omega_2^2 \tag{8.128}$$

This is just stating that if we look at the hyperboloid on the induced metric, we get de-Sitter space. Computing one of the K's:

$$\begin{aligned} K_{\theta\theta} &= -\Gamma^R_{\theta\theta} n_R = R n_R \\ &= \pm l \cosh^2\left(\frac{\tau}{l}\right) = \mp\frac{g_{\theta\theta}}{l} \end{aligned} \tag{8.129}$$

as $R = l\cosh\left(\frac{\tau}{l}\right), n_R = \pm\cosh\left(\frac{\tau}{l}\right)$. This just shows that $K \propto g$ (in this case constant of proportionality is $\mp l$). If we take the inside of the hyperboloid and assume that the wall is symmetric on both sides (with the different signs of each side):

$$K^+_{ab} - K^-_{ab} = 2K^+_{ab} = -\frac{2}{l} h_{ab} \tag{8.130}$$

Therefore from Eq 8.124:

$$\Delta K_{\theta\theta} = -\frac{2}{l} g_{\theta\theta} = -4\pi G\sigma g_{\theta\theta} \tag{8.131}$$

Therefore:

$$l = \frac{1}{2\pi G\sigma} \tag{8.132}$$

Which is the energy value per area, same as was obtained for the domain wall in Eq 8.78. To get the original expression, make a coordinates transformation:

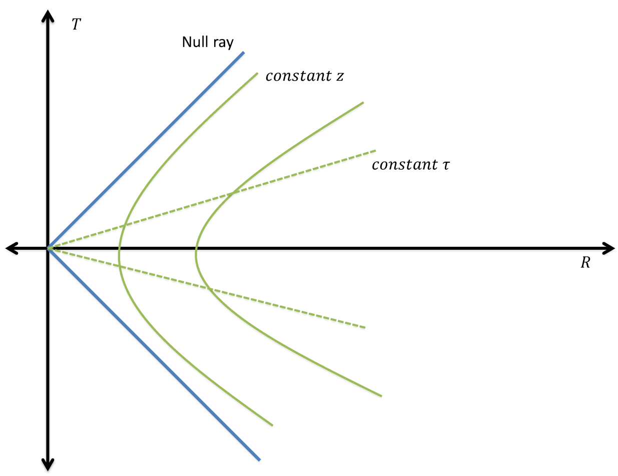


FIGURE 48. Metric shown in Eq 8.134

$$\begin{aligned} R &= l\left(1 - \frac{z}{l}\right)\cosh\left(\frac{\tau}{l}\right) \\ T &= l\left(1 - \frac{z}{l}\right)\sinh\left(\frac{\tau}{l}\right) \end{aligned} \tag{8.133}$$

this gives the $z > 0$ metric for the domain wall:

$$ds^2 = \left(1 - \frac{z}{l}\right)^2 (d\tau^2 - l^2 \cosh^2\left(\frac{\tau}{l}\right) d\Omega_2^2 - dz^2) \tag{8.134}$$

R and T are the flat space, radial coordinates. This shows that the Israel equations give a physical meaning to a sub-manifold.

7. Boundary term in Einstein action: Gibbons-Hawking term

Going back to the Einstein action and the calculation of it's variation in Eq 8.26. Recall that during the integration by parts we had a boundary term, T_1 in Eq 8.26, which we simply ignored. Now let's go back and check weather this boundary term can be incorporated into the variation of the action. In general, T_1 in Eq 8.26 cannot be made zero by the variation of the action because it depends on how δg extends into the bulk (no matter/energy region). The idea now is to add a boundary term, that will cancel these normal derivatives.

The obvious term to include in the boundary term is a scalar and thus a good choice would be the trace of the extrinsic curvature, because it is the other curvature term (other than the Ricci scalar, which forms the Einstein action). The trace of the extrinsic curvature is the covariant derivative of the normal:

$$K = \nabla_a n^a \tag{8.135}$$

Since we are thinking about a variational principle, δh (h is the induced metric) is held fixed on the boundary. Therefore the variation of the normal at the boundary will give the variation of the metric:

$$\delta g_{ab} = n_a \delta n_b + n_b \delta n_a \tag{8.136}$$

We can choose n to be space-like, without any loss of generality:

$$\delta K = \nabla_a \delta n^a + \delta \Gamma^a_{ab} n^b \tag{8.137}$$

δn^a is given in terms of the normal components of δg:

$$\begin{aligned}
\delta K &= -\frac{1}{2}\nabla_a(n^a n_b n_c \delta g^{bc}) - \frac{1}{2}\nabla_n \delta g \\
&= -\frac{1}{2}\nabla_n \delta g - \frac{1}{2}(\nabla_a n^a)(n_b n_c \delta g^{bc}) - \frac{1}{2}(\nabla_a n_b)(n^a n_c \delta g^{bc}) - \frac{1}{2}(\nabla_a n_c)(n^a n_b \delta g^{bc}) - \frac{1}{2}(\nabla_a \delta g^{bc})(n^a n_b n_c) \\
&= -\frac{1}{2}\nabla_n \delta g - \frac{1}{2}K n_b n_c \delta g^{bc} - \frac{1}{2}K n_b n_c \delta g^{bc} - \frac{1}{2}K n_c n_b \delta g^{bc} - \frac{1}{2}n_b n_c n^a \nabla_a \delta g^{bc} \\
&= -\frac{1}{2}\nabla_n \delta g - \frac{1}{2}K n_b n_c \delta g^{bc} - \frac{1}{2}n_b n_c n^e \nabla_a \delta g^{bc}
\end{aligned} \tag{8.138}$$

Substitute for T_1 from induced metric:

$$\begin{aligned}
\delta K &= -\frac{1}{2}\nabla_n \delta g - \frac{1}{2}K n_b n_c \delta g^{bc} - \frac{1}{2}K n_b n_c \delta g^{bc} - \frac{1}{2}n_b (h^d_c - \delta^d_c)\nabla_d \delta g^{bc} \\
&= -\frac{1}{2}\nabla_n \delta g + \frac{1}{2}n_b \delta^d_c \nabla_d \delta g^{bc} - \frac{1}{2}K n_b n_c \delta g^{bc} - \frac{1}{2}h^d_c \nabla_d (n_b \delta g^{bc}) + \frac{1}{2}\delta g^{bc} K_{bc}
\end{aligned} \tag{8.139}$$

Now lets look at the variation of this boundary term (ignoring the Ricci scalar for now, as we already know how that works):

$$\delta \int K\sqrt{q} d^3x = \int d^3x \sqrt{g} \underbrace{\left[\frac{1}{2}\delta g^{bc}(K_{bc} - K_{hbc})\right.}_{T_1} - \frac{1}{2}\underbrace{\nabla_c(n_b \delta g^{bc})}_{T_2} + \frac{1}{2}\underbrace{(n_b \nabla_c \delta g^{bc} - \nabla_n \delta h)}_{T_3} \tag{8.140}$$

T_2 is zero as the metric is on the boundary and therefore its derivative is zero. T_1 has a δg^{bc} in it, which is good as we expect the variation of the curvature on the boundary to be proportional to the metric on the boundary. T_3 is precisely the term needed to cancel the derivative coming from the variation of the bulk action. Therefore:

$$\delta\left[S_E - \underbrace{\frac{1}{8\pi G}\int_{dM}[K\sqrt{q}]}_{T_\alpha}\right] = 0 \tag{8.141}$$

T_α is known as the Gibbons-Hawking boundary term, S_{GH}. Then the variation principle is satisfied and we no longer need to fudge the calculations by setting the boundary terms to zero.

7.1. Revisiting the Israel equations. If we split up the space-time into the two regions shown in Figure 49, M^+ and M^-, then the action is:

$$S = S_E^+ + S_E^- + S_{GH}^+ + S_{GH}^- + S_{wall} \tag{8.142}$$

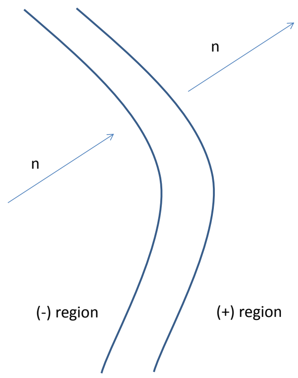


FIGURE 49. Space-time region, represented by manifold M, with a sub-manifold Σ, representing the domain wall (a small localised region of matter)

where:

$$\begin{aligned} S_E^+ &= \text{Einstein action in (+) region of space-time} \\ S_E^- &= \text{Einstein action in (-) region of space-time} \\ S_{GH}^+ &= \text{Gibbons-Hawking action term in (+) region of space-time} \\ S_{GH}^- &= \text{Gibbons-Hawking action term in (-) region of space-time} \\ S_{wall} &= \text{Matter part of the action in the region of the wall} \end{aligned} \tag{8.143}$$

Varying the action:

$$\delta S = 0 \Rightarrow G_{ab} = 0 \tag{8.144}$$

This is the solution for the bulk (as there is no matter/energy). The action in the domain wall:

$$\frac{\delta S_{wall}}{\delta g^{ab}} = \frac{1}{16\pi G}\left[\left(K_{ab}^+ - K^+ h_{ab}\right) + \left(K_{ab}^- - K^- h_{ab}\right)\right] \tag{8.145}$$

In this region of the domain wall (described by a sharp localised peak in the matter and it is described on a sub-manifold Σ). From the Israel equation, 8.124:

$$\frac{1}{2}S_{ab}|_{wall} = \frac{1}{16\pi G}\left[\Delta K_{ab} - \Delta K h_{ab}\right]_{Israel} \tag{8.146}$$

CHAPTER 9

Black holes

This chapter is strongly related to the first few chapters of [**8**].

1. Black hole thermodynamics

Within Einstein gravity, black holes are very simple when there in the vacuum equations. An example that has been previously studied is the usual Kerr metric:

$$ds^2 = dt^2 - \frac{\Sigma}{\Delta}dr^2 - \Sigma d^2 - \frac{2GMr}{\Sigma}(dt^2 - a\sin^2 d\phi)^2 - (r^2 + a^2)\sin^2\theta d\phi^2 \tag{9.1}$$

also written as:

$$ds^2 = \frac{\Delta}{\Sigma}(t - a\sin^2\theta d\phi)^2 - \frac{\Sigma}{\Delta}dr^2 - \Sigma d\theta^2 - \frac{\sin^2\theta}{\Sigma}((r^2 + a^2)d\phi - adt)^2 \tag{9.2}$$

where:

$$\begin{aligned} \Delta &= r^2 + a^2 - 2GMr \\ \Sigma &= r^2 + a^2\cos^2\theta \\ a &= \frac{J}{M} \\ J &= \text{Angular momentum} \\ M &= \text{Mass of black hole} \\ \frac{\partial}{\partial t}, \frac{\partial}{\partial \phi} &= \text{Killing vectors} \end{aligned} \tag{9.3}$$

This represents a rotating black hole. Notice that there are very charges, i.e M, a are the only ones. There would also be Q if the black hole had an electric charge. In which case:

$$\Delta = r^2 + a^2 + GQ^2 - 2GMr \tag{9.4}$$

The area of the black hole:

$$\text{Area} = 4\pi(r_+^2 + a^2) \tag{9.5}$$

and:

$$r_+ = GM + \sqrt{G^2M^2 - a^2 - GQ^2} \tag{9.6}$$

Suppose a material object falls into the black hole. One would expect that on its journey into the black hole, the solutions might carry some dynamics. However once the particles has fallen into the black hole to settle down into a new stable configuration with a new mass, angular momentum and charge, if the particle falling in has an electric charge. First let's look at what happens to the area of the event horizon:

$$\begin{aligned}
\delta A &= 8\pi r_+ \delta r_+ + a\delta a \\
&= 8\pi r_+ \left[G\delta M + \frac{G^2 M \delta M - a\delta a - GQ\delta Q}{r_+ - GM} \right] + 8\pi a \delta a \\
&= 8\pi \left[\frac{G r_+^2 \delta M}{r_+ - GM} - \frac{GMa\delta a}{r_+ - GM} - \frac{GQr_+\delta Q}{r_+ - GM} \right] \qquad (9.7)
\end{aligned}$$

But:

$$a = \frac{J}{M} \qquad (9.8)$$

therefore:

$$\delta a = \frac{\delta J}{M} - \frac{J\delta M}{M^2} \qquad (9.9)$$

Substitute Eq 9.9, 9.8 into Eq 9.7:

$$\delta A = 8\pi \left(\frac{G r_+^2}{r_+ - GM} \delta M - \frac{Ga\delta J}{r_+ - GM} + \frac{Ga^2 \delta M}{r_+ - GM} - \frac{G r_+ Q \delta Q}{r_+ - GM} \right) \qquad (9.10)$$

Re-arrange for charge in mass of black hole:

$$\delta M = \underbrace{\frac{r_+ - GM}{2\pi(r_+^2 + a^2)}}_{T_1} \frac{\delta A}{4G} + \underbrace{\Omega \delta J}_{T_2} + \underbrace{\phi \delta Q}_{T_3} \qquad (9.11)$$

$$\begin{aligned}
\Omega &= \frac{a}{r_+^2 + a^2} \equiv \text{Angular velocity} \\
\phi &= \frac{Q r_+}{r_+^2 + a^2} \equiv \text{Electrostatic potential} \qquad (9.12)
\end{aligned}$$

Now consider the thermodynamic equation:

$$dU = TdS + u_i dQ_i \qquad (9.13)$$

By analogy between Eq 9.13 and Eq 9.11, it appears that there is a direct correspondence between thermodynamics and black holes. To see it explicitly, let's look at the Schwarzschild solution, i.e set $r_+ = 2GM$ and $a = 0$, thus T_1 is:

$$\frac{GM}{2\pi(4G^2M^2)} = \frac{1}{8\pi GM} \qquad (9.14)$$

Which is what we previously argued was the Schwarzschild temperature which suggests that T_1 is actually a generalised expression for the temperature and thus $\frac{\delta A}{4G}$ should be equal to the entropy:

$$\delta S = \frac{\delta A}{4G} \qquad (9.15)$$

At the moment this is not a formal derivation of any king, it is just a curiosity that there appears to be a correspondence between area and entropy.

1.1. Euclidean path integral approach. To formalise this correspondence between black holes and thermodynamics, we embed the classical solutions to Einstein's equations, into a quantum background. This is not attempting to quantise these solutions in any way, it is simply a different way to look at the solutions. We will try to construct a partition function for the gravitational theory. In general one would take a trace of states:

$$Z \approx tr(e^{-\beta H}) \tag{9.16}$$

where $\beta = \frac{1}{T}$ in units that $k_B \equiv 1$, thus T has units of energy. H is the Hamiltonian:

$$H \approx \int d^3x \mathcal{H} \approx \frac{1}{\beta} \int d^3x dt H \approx \frac{I_E}{\beta} \tag{9.17}$$

where I_E is the *Euclidean action*:

$$I_E = \underbrace{\int d^4x \frac{\sqrt{g}}{16\pi G}}_{\equiv I_{EH}} + \underbrace{\int d^3x \frac{\sqrt{h}K}{8\pi G}}_{\equiv I_{GH}} \tag{9.18}$$

We expect Z to be dominated by saddle points:

$$Z \approx \sum_{\text{classical solutions, } \beta} e^{-I_E} \tag{9.19}$$

We postulate that the contribution is dominated by classical solutions with *periodicity* β. There is no Ricci scalar, $R = 0$, for Schwarzschild and Kerr, therefore there is no Einstein action and $I_{EH} = 0$. Therefore the total action must come from the Gibbon-Hawking term, I_{GH}:

$$I_E = \int_{dM} d^3x \frac{\sqrt{h}K}{8\pi G} \tag{9.20}$$

dM is the surface at the boundary. Remember h is the induced metric. Let's look at the Schwarzschild solution as it is simple yet shows all the important points of a black hole:

$$ds^2 = \left(1 - \frac{2GM}{r}\right) d\tau^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2 d\Omega_2^2 \tag{9.21}$$

We have already seen that in polar coordinates, light rays can reach the singularity at $r = 0$, if the τ coordinates have periodicity $8\pi GM$.

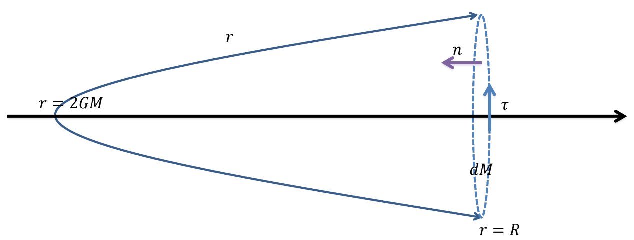


FIGURE 50. Metric in Eq 9.21.

Therefore the Euclidean Schwarzschild metric looks like this, as at large r, the coefficient of $d\tau$ goes to 1. As $r \to 2GM$, the space looks, locally as r^2 in the vicinity of the origin. The size of the τ circles is $\left(1 - \frac{2GM}{r}\right)$, and eventually reaches 0 as $r \to 2GM$. Topologically this is $\mathbb{R}^2$ and then one has a two sphere on top, S^2, $S^2 \times \mathbb{R}^2$. But this is all just the Einstein part of the solution which has the action zero as $R = 0$. The question is, how do we find a boundary to this space-tie to get the Gibbons-Hawking term. The idea is to evaluate this solution to some finite radius, $R >> GM$ and then see what happens to the solution in this limit. On the surface , the metric is 3D:

$$dS_3^2 = \left(1 - \frac{2GM}{R}\right) d\tau^2 + R^2 d\Omega_2^2 \tag{9.22}$$

this has a 2 sphere of radius R and then a single S_1. τ has a periodicity $8\pi GM$, as usual. The inward pointing normal n, has the form:

$$n = -\sqrt{1 - \frac{2GM}{R}} \frac{\partial}{\partial r} \tag{9.23}$$

The extrinsic curvature is:

$$\begin{aligned} K &= \nabla_a n^a = \frac{1}{r^2}\partial_r(r^2 n^r) \\ &= -\frac{2}{R}\sqrt{1 - \frac{2GM}{R}} - \frac{GM}{R^2}\frac{1}{\sqrt{1 - \frac{2GM}{R}}} \end{aligned} \tag{9.24}$$

Therefore the boundary term is:

$$\int_{dM} K\sqrt{-h} d\tau d\phi d\theta \tag{9.25}$$

but:

$$K \neq K(\tau, \phi, \theta) \tag{9.26}$$

therefore integrating over them simply gives numerical factor of $4\pi\beta$:

$$\begin{aligned} \int_{dM} K\sqrt{-h} d\tau d\theta d\phi &= -4\pi\beta R^2 \left[\frac{2}{R}\underbrace{\left(1 - \frac{2GM}{R}\right)}_{T_\alpha} + \frac{GM}{R^2} \right] \\ &= -4\pi\beta \left[2R - 3GM\right] \end{aligned} \tag{9.27}$$

Now, the leading order term is proportional to R, therefore we cannot take the boundary to be at $R \to \infty$ as the action will also becomes ∞. However, by comparing the equation above, especially, T_α, with the Schwarzschild metric solution. Then the extrinsic curvature for *flat* space-time, K_0, must be:

$$K_0 = -\frac{2}{R} \tag{9.28}$$

Therefore:

$$\begin{aligned} \int K_0 \sqrt{h} d^x &= -4\pi\beta R^2 \sqrt{1 - \frac{2GM}{R}} \left(\frac{2}{R}\right) \\ &\approx -4\pi\beta[2R - 2GM + \mathcal{O}\left(\frac{1}{r}\right)] \end{aligned} \tag{9.29}$$

Therefore the flat space answer is also ∞ as $R \to \infty$. This means that this divergent term is not physics (as it is also present in the flat space). Therefore we can simply subtract this vacuum solution by the Schwarzschild solution, to get the true value for the action:

$$I_{SCHW} - I_{Vacuum} = \frac{1}{8\pi G}\left[-4\pi\beta[-GM]\right] = \frac{\beta M}{2} \tag{9.30}$$

This is sometimes called the *re-normalised* action. The entropy is given by the usual statistical formula:

$$S = \beta^2 \frac{\partial}{\partial \beta} \left[-\beta^{-1} \ln Z \right] \tag{9.31}$$

where $Z = -e^{\frac{\beta M}{2}}$ in this case. Therefore:

$$\begin{aligned} S &= \frac{\beta^2}{2} \frac{\partial M}{\partial \beta} \\ &= \frac{\beta^2}{16\pi G} = 4G\pi M^2 \\ &= \frac{4\pi}{4G}(2GM)^2 \end{aligned} \tag{9.32}$$

By comparison to Eq 9.15:

$$\text{Area} = 4\pi(2GM)^2 \tag{9.33}$$

Therefore we have shown that for the Schwarzschild black hole, the entropy is indeed $\frac{A}{4G}$. From the microscopic viewpoint:

$$S \approx \log\left(\#\text{of micro states}\right) \tag{9.34}$$

but classically, black holes have very few charges. Also notice that $T \propto \frac{1}{M}$, therefore black holes have a negative heat capacity, just like stars do. The energy released is given by the rate of change of mass:

$$\begin{aligned} \dot{M} &\approx \sigma T^4 A \\ &\approx \frac{\sigma (GM)^2}{(8\pi GM)^4} \\ &\approx \frac{1}{M^2} \end{aligned} \tag{9.35}$$

Therefore the lifetime of a black hole can be found by integrating over this expression, therefore the lifetime is $\propto M^{-3}$.

2. Kaluza-Klein theory

As far as we observe, we live in a 3D space, with 1 time dimension, however the idea of extra dimensions keeps on coming up in physics. The first concrete model of extra dimensions was thought up by Kaluza and Klein after Einstein came up with general relativity.

In 1920/21, Kaluza had an idea of adding new dimensions to space-time in order to unify Maxwell's equations and the theory of electro-magnetism with general relativity and gravity. In fact, Einstein resisted this idea for a number of years, until accepting it as a valid concept to discuss in theoretical physics at the time. By adding an extra dimension, Kaluza showed that one could almost obtain the Maxwell-Einstein theory that he wanted, accept that there was an extra scalar field, which then Klein suggested, needed to be stabilised in order for the theory to work (he never gave a mechanism for stabilising it).

Now, a century later, the whole concept of extra-dimensions, where these extra-dimensions has been stabilised as Klein suggested. This is of course in the theory of super strings. The idea in Kaluza-Klein is to look at gravity in 5D, except for how the 5D geometry would look like in 4D perspective, i.e the geometry depends only on 4 of the 5 dimensions:

$$ds^2 = \underbrace{g_{\mu\nu}dx^\mu dx^\nu}_{T_1} - \underbrace{e^{2\sigma}[d\psi + A_\mu dx^\mu]^2}_{T_2} \tag{9.36}$$

where:

$$\nu, \mu \in \{0, 1, 2, 3\} \tag{9.37}$$

and:

$$\sigma = \sigma(x^\mu) \qquad A_\mu = A_\mu(x^\mu) \tag{9.38}$$

T_1 represents the 4D space that we live in. ψ is the 5th dimensions, but none of the terms have any ψ dependence in them. Therefore $\frac{\partial}{\partial\psi}$ is a killing vector. To obtain Einstein's equations, we need the Ricci tensor and scalar. Using the Cartan formalism; Identify basis:

$$\begin{aligned} \vec{w}^{\hat{\psi}} &= e^\sigma[d\psi + A] \\ \vec{w}^{\hat{a}} &= e^{\hat{a}}_\mu \vec{d}x^\mu \end{aligned} \tag{9.39}$$

$$\begin{aligned} \vec{d}\vec{w}^{\hat{\psi}} &= \sigma_{,\hat{a}}\vec{w}^{\hat{a}} \wedge \vec{w}^{\hat{\psi}} + e^\sigma F(\equiv \vec{d}a) \\ \vec{d}\vec{w}^{\hat{a}} &= e^{\hat{a}}_{\mu\nu}\vec{d}x^\nu \wedge \vec{d}x^\mu \end{aligned} \tag{9.40}$$

Reading of the connection 1 forms:

$$\begin{aligned} \vec{\theta}^{\hat{\psi}}_{\hat{a}} &= \sigma_{,\hat{a}}\vec{w}^{\hat{\psi}} + \frac{1}{2}e^\sigma F_{\hat{a}\hat{b}}\vec{w}^{\hat{b}} \\ \vec{\theta}^{\hat{a}}_{\hat{b}} &= \vec{\theta}^{\hat{a}}_{0\,\hat{b}} + \frac{1}{2}e^\sigma F^{\hat{a}}_{\hat{b}}\vec{w}^{\hat{\psi}} \end{aligned} \tag{9.41}$$

Recall the Cartan equation:

$$R^{\hat{a}}_{\hat{b}} = \underbrace{\vec{d}\vec{\theta}^{\hat{a}}_{\hat{b}}}_{T_1} + \underbrace{\vec{\theta}^{\hat{a}}_{\hat{c}} \wedge \vec{\theta}^{\hat{c}}_{\hat{b}}}_{T_2} + \underbrace{\vec{\theta}^{\hat{a}}_{\hat{\psi}} \wedge \vec{\theta}^{\hat{\psi}}{}_{\hat{b}}}_{T_3} \tag{9.42}$$

Computing the terms individually:

$$\begin{aligned} T_1 &= \vec{d}\vec{\theta}^{\hat{a}}_{\hat{b}} + \frac{1}{2}e^\sigma[\sigma_{,\hat{c}}\vec{w}^{\hat{c}} \wedge \vec{w}^{\hat{\psi}}F^{\hat{a}}_{\hat{b}} + F^{\hat{a}}_{\hat{b}} + F^{\hat{a}}_{\hat{b}\hat{c}}\vec{w}^{\hat{c}} \wedge \vec{w}^{\hat{\psi}} + F^{\hat{a}}_{\hat{b}}(\sigma_{,\hat{c}}\vec{w}^{\hat{c}} \wedge \vec{w}^{\hat{\psi}} + e^\sigma F)] \\ T_2 &= \vec{\theta}^{a}_{0c} \wedge \vec{\theta}^{c}_{0b} + \frac{1}{2}e^\sigma[\vec{\theta}^{\hat{a}}_{0\hat{c}} \wedge F^{\hat{c}}_{\hat{b}}\vec{w}^{\hat{\psi}} + F^{\hat{a}}_{\hat{c}}\vec{w}^{\hat{\psi}} \wedge \vec{\theta}^{\hat{c}}_{0\hat{b}}] \\ T_3 &= \sigma^{,\hat{a}}\vec{w}^{\hat{\psi}} \wedge e^\sigma F_{\hat{b}\hat{c}}\vec{w}^{\hat{c}} + \frac{1}{2}F^{\hat{a}}_{\hat{c}}\vec{w}^{\hat{c}} \wedge \sigma_{,\hat{b}}\vec{w}^{\hat{\psi}} + \frac{1}{4}e^{2\sigma}F^{\hat{a}}_{\hat{c}}\vec{w}^{\hat{c}} \wedge F_{\hat{b}\hat{a}}\vec{w}^{\hat{d}} \end{aligned} \tag{9.43}$$

This long expression gives the curvature 2 form. Now we need to identify the parts that are needed for the Ricci scalar. The Ricci scalar only depends on the $\{0, 1, 2, 3\}$ indices therefore we can ignore any terms with ψ in them:

$$R^{\hat{a}}_{\hat{b}\hat{c}\hat{d}} = R^{\hat{a}}_{0\,\hat{b}\hat{c}\hat{d}} + \frac{1}{4}e^{2\sigma}[F^{\hat{a}}_{\hat{c}}F)\hat{b}\hat{a} - F^{\hat{a}}_{\hat{d}}F_{\hat{b}\hat{c}} + 2F^{\hat{a}}_{\hat{b}}F_{\hat{c}\hat{d}}] \tag{9.44}$$

The other curvature two form that contributes is:

$$\begin{aligned}
R^{\hat{\psi}}_{\hat{a}} &= \vec{d\theta}^{\hat{\psi}}_{\hat{a}} + \vec{\theta}^{\hat{\psi}}_{\hat{b}} \wedge \vec{\theta}^{\hat{b}}_{\hat{a}} \\
&= \sigma_{,\hat{a}\hat{b}}\vec{w}^{\hat{b}} \wedge \vec{w}^{\hat{\psi}} + \sigma_{,\hat{a}}[\sigma_{,b}\vec{w}^{\hat{b}} \wedge \vec{w}^{\hat{\psi}} + e^{\sigma}F] \\
&+ \frac{1}{2}e^{\sigma}[\sigma_{,\hat{b}}\vec{w}^{\hat{b}} \wedge F_{\hat{a}\hat{c}}\vec{w}^{\hat{c}} + F_{\hat{a}\hat{b},\hat{c}}\vec{w}^{\hat{c}} \wedge \vec{w}^{\hat{b}} - F_{\hat{a}\hat{b}}\vec{\theta}^{\hat{b}}_{\hat{c}}\vec{w}^{\hat{c}}] \\
&+ \sigma_{,\hat{b}}\vec{w}^{\hat{\psi}} \wedge [\vec{\theta}^{\hat{b}}_{\hat{a}} + \frac{1}{2}e^{\sigma}F^{\hat{b}}_{\hat{a}}\vec{w}^{\hat{\psi}}] \\
&+ \frac{1}{2}e^{\sigma}F_{\hat{b}\hat{c}}\vec{w}^{\hat{c}} \wedge [\vec{\theta}^{\hat{b}}_{0\hat{a}} + \frac{1}{2}e^{\sigma}F^{\hat{b}}_{\hat{a}}\vec{w}^{\hat{\psi}}] \qquad (9.45)
\end{aligned}$$

Again we only take terms that will contribute to the $R^{\hat{\psi}}_{\hat{a}\hat{\psi}\hat{b}}$, thus we need terms which has one 4D index and one 5D index (which is ψ), i.e we ignore all terms with purely space-time indices:

$$R^{\hat{\psi}}_{\hat{a}\hat{\psi}\hat{b}} = -\sigma_{,\hat{a}\hat{b}} - \sigma_{,\hat{a}}\sigma_{,\hat{b}} + \sigma_{,\hat{c}}\Gamma^{\hat{c}}_{\hat{b}\hat{a}} - \frac{1}{4}e^{2\sigma}F_{\hat{a}\hat{c}}F^{\hat{c}}_{\hat{b}} \qquad (9.46)$$

Therefore:

$$\begin{aligned}
R^{\hat{\psi}}_{\hat{\psi}} &= -\Box\sigma - (\nabla\sigma)^2 - \frac{1}{4}e^{2\sigma}F^2 \\
R_{ab} &= R_{0ab} - \nabla_a\nabla_b\sigma - \nabla_{a\sigma}\nabla_{b\sigma} + \frac{1}{2}e^{2\sigma}F_{ac}F^c_b \\
R_s &= R_0 + \frac{1}{4}e^{2\sigma}F^2 - 2\Box\sigma - 2(\nabla\sigma)^2 \qquad (9.47)
\end{aligned}$$

So the 5D Ricci scalar is the 4D Ricci scalar plus an F^2 terms and then some scalar terms. If we look at Einstein's action and look at the last component, $det(g_5)$:

$$det(g_5) = e^{2\sigma}det(g_\psi) \qquad (9.48)$$

Therefore the 5 dimensional Einstein action, S_5 becomes:

$$\begin{aligned}
S_5 &= -\frac{1}{16\pi G_5}\int d^5x\sqrt{g_5}R_5 \\
&= -\frac{1}{16\pi G_5}\int d^4x\sqrt{g_4}d\psi e^{\sigma}\left(R_0 + \frac{1}{4}e^{2\sigma}F^2 - 2e^{-\sigma}\Box e^{\sigma}\right) \\
&= -\frac{L}{16\pi G_5}\int d^5x\sqrt{g_4}e^{\sigma}\left[R_4 + \frac{1}{4}e^{2\sigma}R^2\right] \qquad (9.49)
\end{aligned}$$

where in the 2nd line we have integrated over ψ, which is just a constant, L (Periodicity of ψ). So we seem to have reduced the Einstein action in 5D, to an action in 4D which looks almost exactly like the Einstein-Maxwell action. In fact, if σ is fixed, this Einstein-Maxwell action. Finally, we can conformally transform this action into the Einstein frame, by the following coordinate transformations:

$$g_{\mu\nu} \equiv e^{-\sigma}\tilde{g}_{\mu\nu} \qquad \phi \equiv \frac{\sigma}{3} \qquad (9.50)$$

Which gives the action:

$$S = \frac{L}{8\pi G_5}\int d^4x\sqrt{\tilde{g}}\left(-\tilde{R} + \frac{1}{2}(\partial\phi)^2 - \frac{e^{\sqrt{3}\phi}}{4}F^2\right) \qquad (9.51)$$

and the metric is:

$$ds^2 = e^{-\sqrt{3}\phi}\tilde{g}_{\mu\nu}dx^\mu dx^\nu - e^{2\sqrt{3}\phi}\left[d\psi + A_\mu dx^\mu\right]^2 \qquad (9.52)$$

The main problem with this theory (in fact with all higher dimensional theories) is that, when we apply this to the whole universe, the expansion of the universe would also expand the smaller dimension, which would make it visible after a "sufficiently" large amount of time. The fact that we do not observe these dimensions today, means they must have been even smaller in the past, leading to another fine tuning problem.

3. Black holes in Kaluza Klein theory

Kaluza-Klein (KK) black holes are vacuum solutions of standard Einstein gravity. Since the extra dimension, ψ, in KK is expected to be small, $\frac{\partial}{\partial\psi}$ is a Killing vector. One of the solutions is the Schwarzschild like solution in 5D:

$$ds^2 = \left(1 - \frac{r_+}{r}\right) dt^2 - \left(1 - \frac{r_+}{r}\right)^{-1} dr^2 - r^2 d\Omega_2^2 - d\psi^2 \tag{9.53}$$

This is called a *black string*, as it the Schwarzschild solution in 4D, extended along the ψ direction.

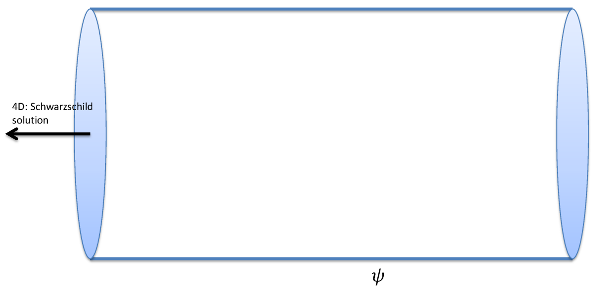


FIGURE 51. Black string with a 4D boundary represented by the Schwarzschild solution and an extra smooth direction ψ added on to the metric

This the simplest example of a KK black hole. To make it more general, firstly, one can introduce cross terms in the metric. This can be done via Lorentz boosting the black string, however, at first sight one would think that a Lorentz transformation should leave the metric unchanged. This is because once ψ is fixed, it is periodic as stated before and therefore fixes a rest frame.

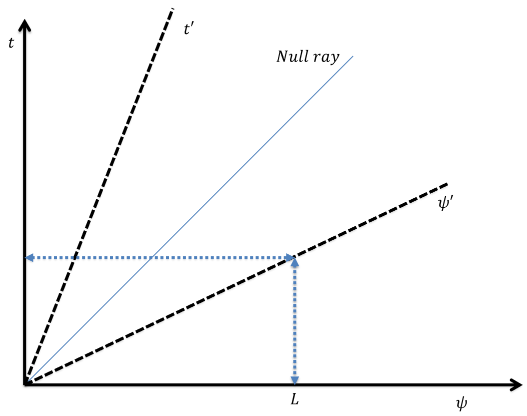


FIGURE 52. Lorentz boost of the black string in the ψ dimension, with L being the periodicity of ψ.

L is the periodicity of ψ:

$$\psi = \psi + L \qquad \text{(at constant } t, x, y, z) \tag{9.54}$$

Hence doing a Lorentz transformation means one has to start identifying coordinates at different points, ψ' and t':

$$\begin{aligned} t' &= \gamma(t - v\psi) \\ \psi' &= \gamma(\psi - vt) \end{aligned} \tag{9.55}$$

Therefore:

$$(t', \psi') \approx (t' - v\gamma L, \psi' + \gamma L) \tag{9.56}$$

Which means there is a difference between between identifying ψ' and then boosting or boosting first and then identifying ψ'. By making a boost, the ψ and t directions are mixed therefore the motion in ψ gives a $g_{t\psi}$ term which one can related to A_t, a charge. Let's see this explicitly:

$$\begin{aligned} ds^2 &= \left(1 - \frac{r_+}{r}\right) \frac{(dt + vd\psi)^2}{1 - v^2} - \frac{(d\psi + vdt)^2}{1 - v^2} - \left(1 - \frac{r_+}{r}\right)^{-1} dr^2 - r^2 d\Omega_2^2 \\ &= -\left(1 - \frac{r + v^2}{(1 + v^2)r}\right) \left(d\psi + \frac{r + vdt}{(1 - v^2)r + r_s V^2}\right)^2 + \left(1 - \frac{r_+}{(1 - v)r + r_+ v^2}\right) dt^2 - \left(1 - \frac{r_+}{r}\right)^{-1} dr^2 - r^2 d\Omega_2^2 \end{aligned} \tag{9.57}$$

Shifting the origin the radial coordinate:

$$\hat{h} = r + \frac{r + v^2}{1 - v^2} \tag{9.58}$$

Identify charge:

$$q = \frac{r_+ v}{1 - v^2} \tag{9.59}$$

which gives:

$$A = \frac{q}{\hat{r}} dt \tag{9.60}$$

which is a familiar looking potential and the scalar field:

$$e^{2\sqrt{3}\phi} = \frac{\hat{r}}{\hat{r} - vq} = \frac{\hat{r}}{r} \tag{9.61}$$

The metric then, in the Einstein frame is:

$$ds_4^2 = \left(1 - \frac{r_s}{(1 - v^2)\hat{r}}\right) \left(1 - \frac{Vq}{\hat{r}}\right)^{-\frac{1}{2}} dt^2 - \left(1 - \frac{\hat{r}_s}{\hat{r}}\right)^{-1} \left(1 - \frac{Vq}{\hat{r}}\right)^{\frac{1}{2}} d\hat{r}^2 - \left(1 - \frac{Vq}{\hat{r}}\right)^{\frac{1}{2}} r^2 d\Omega_2^2 \tag{9.62}$$

If we take the extremal limit:

$$\hat{r}_+ \to 0, \qquad v \to 1, \qquad \hat{r}_s = \text{fixed} \tag{9.63}$$

Which gives:

$$q = \hat{r}_+ \left(1 - \frac{vq}{r^2}\right) \to \left(1 - \frac{\hat{r}_+}{r}\right) \tag{9.64}$$

which has a null singularity at $\hat{r}_s$.

3.1. Magnetic Kaluza-Klein black hole. One of the curious things in 4D, is that aswell as having an electrically charged black hole, one can also get a magnetically charged black hole. Which means that:

$$F \approx Q \sin\theta \vec{d\theta} \wedge \vec{d\phi} \tag{9.65}$$

Which basically means one has a radial magnetic field. It has the same stress-energy tensor as an electrically charged black hole:

$$A = \frac{Q}{r}\vec{dt} \tag{9.66}$$

Which means the same geometry, since the magnetic field is the dual of F; the geometry for F and its dual is the same. Therefore there is a duality between electric and magnetic charges. For KK theory we want the same F, which means we need to construct a gauge potential:

$$A = Q(\pm 1 - \cos\theta)\vec{\theta} \tag{9.67}$$

This may not be obvious as naively one would just expect it to be something that goes as $-\cos\theta$. But this term is part of a metric and in a metric, any singularities are actually coordinate singularities. Normally one does not worry about singularities in polar coordinates on the north and south pole, as they are relatively straight forward to remove via coordinate transformations.

In this case, there is a term $g_{\psi\phi}$ and therefore A must be regular at north and south poles, which in the reason behind adding the ± 1 part. These different signs actually give rise to two distinct gauge fields, one that works on the north pole and are that works on the south pole:

$$ds^2 = \underbrace{\left(\frac{r-r_+}{r-r_-}\right)}_{T_\alpha} dt^2 - \frac{dr^2}{\left(1-\frac{r_\pm}{r}\right)} - r(r-r_-)d\Omega_2^2 - \left(1-\frac{r_-}{r}\right)\left[d\psi + \underbrace{\sqrt{r_+ r_-}(1-\cos)}_{T_1} d\phi\right]^2 \tag{9.68}$$

This is what the 5D metric looks like and this is regular at the north pole, and the south pole cannot be included in this. Instead, for the south pole:

$$A_s = -Q(1+\cos\theta)d\phi \tag{9.69}$$

This is a coordinate transformation:

$$\psi_S = \psi_N + 2Q\phi \tag{9.70}$$

Therefore shifting the ψ angle between the North and South patches by $2Q\phi$, then the $1 \to -1$ in T_1 is taken care off. The periodicity of ψ implies:

$$4\pi Q = nL \qquad (n \in Z) \tag{9.71}$$

This is giving a very striking result. It is saying that this solution only works if the magnetic charge is quantised. Or in other words, if there exists quantised magnetic charges, L, can only take quantised values and L is related to electric charge of fields, therefore the existence of a single magnetic monopole means electric charge is quantised.

3.2. Extremal limit, $r_+ = r_-$. In this limit T_α in Eq 9.68 becomes zero, therefore there is no potential in front of the time components of the metric. If we set:

$$Q = r_+ = \frac{L}{4\pi} \qquad (i.e n = 1) \tag{9.72}$$

and redefine the extra dimension:

$$\chi = \frac{4\pi\psi}{L} \tag{9.73}$$

therefore the periodicity of χ is 4π:

$$ds^2 = dt^2 - \frac{dr^2}{1-\frac{r_+}{r}} - \left(1 - \frac{r_+}{r}\right)[r^2 d\Omega^2 + 2 + r_+^2(d\chi(1+\cos\theta)d\phi)^2 \tag{9.74}$$

It appears that there is a singularity as $r \to r_+$. Define a new radial coordinate:

$$\rho^2 = 4r_+(r - r_+) \tag{9.75}$$

Now as $r \to r_+$:

$$ds^2 = dt^2 - d\rho^2 - \frac{\rho^2}{4}\left[d\Omega_2^2 + (d\chi + (1-\cos\theta)d\rho)^2\right] \tag{9.76}$$

where $d\Omega_3^2$ is the distance element in a unit 3 sphere in Euler angle.

$\rho^2 \to 0$ is the origin of $\mathbb{R}^4$. The KK monopole is topologically $\mathbb{R}^5$, smooth. As we go to $r \to 0$, the θ, ϕ directions grow, whereas the χ remains fixed as it is independent of r. Therefore it really looks like a KK solution at large r, as the usual $3+1$ dimensions increases in size and the 1 remaining dimension does not grow and hence is not observable (of course, here we have to assume that the dimension is small in the beginning, relative to the size of the other dimension today). The patching of coordinates between the north and south pole is known as *Hopf fibration* of S^4 over S^2 giving S^3 (discovered by Gross-Perry-Sordin).

This concludes the pure KK solution. We have seen that adding in extra dimensions not not gives rise to black holes, it also gives Black strings, black branes etc. An example of a Higher dimensional black hole is:

$$ds^2 = \left(1 - \frac{r_+}{r}^{D-3}\right)dt^2 - \frac{dr^2}{\left(1-\left(\frac{r_+}{r}\right)^{D-3}\right)} - r^2 d\Omega_{D-2}^2 \tag{9.77}$$

D is the dimension of the overall manifold. This is just the Schwarzschild solution in higher dimensions. All we see is that the potential just changes its form. These black holes have a mass:

$$16\pi G_D M = (D-2)A_{D-2}r_+^{D-3} \tag{9.78}$$

A is the area of corresponding to a surface of radius ρ^{D-2}. Similarly, there exist objects called *Black branes*, described by the metric:

$$\begin{aligned} ds^2 &= \underbrace{ds_{p+1}^2}_{T_1} - \underbrace{dx^i dx^j \delta_{ij}}_{T_2} \\ T_1 &= \text{black hole in (p+1) dimension} \\ T_2 &= \text{more dimensions added to the black hole metric} \end{aligned} \tag{9.79}$$

Here we are simply taking the Black hole solution in (p+1) dimensions and then adding extra dimensions. It is interesting to see what charges these branes can carry. To find charged brane solutions, one has to have form fields $B_{\mu\nu}, C_{\mu\nu\lambda}$ etc.

4. Perturbation theory

In 4D we think black holes are stable as they are observed in our universe. To check this stability via general relativity, the solutions obtained analytically can be perturbed and then the effects are analysed. If a solution is stable it will not change too much under the perturbation. Consider a 5D black hole, with radius, r_+ and compare it to to a black string.

Note that the entropy of the Black string has a factor of $\frac{1}{L}$ which means that as the length increases with fixed mass, eventually the entropy of the black strings will drop below the entropy of

	5D Black hole	Black string
Area	$2\pi^2 r_+^3$	$4\pi r_s^2 L$
Mass	$\frac{3\pi r_+^2}{8G_5}$	$\frac{r_s L}{2G_5}$
Entropy	$\frac{8}{3}\sqrt{\frac{2\pi G_5}{3}}M^{\frac{3}{2}}$	$\frac{4\pi M^2 G_5}{L}$

TABLE 7. Comparison of black hole and black string

black holes. Thus if black holes thermo-dynamics is believed then as the compactification length, L, increases the entropy of the black string will decrease, which means the black strings would prefer to organise its mass as a black hole. This can be seen by expressing the black string entropy as a black hole entropy:

$$S_{BS} = S_{BH}\sqrt{\frac{27\pi r_s}{16L}} \tag{9.80}$$

The important thing to notice here is the ration of $\frac{r_s}{L}$. Therefore there exists a critical length for a string, above which the black hole is a preferred state. This indicates that a black string will have a *long wavelength instability.* To check whether the instability is real, one needs to understand how perturbation theory works in gravity. Not only is perturbation theory useful to check stability, it is also useful because it is hard to find exact solutions to Einstein's field equations therefore once a solution has been found, we can perturb it and see what the equations give. The idea behind perturbation theory, is to take a metric g_{ab} and perturb it:

$$\begin{aligned} g_{ab} &= \underbrace{g_{oab}}_{T_1} + \underbrace{h_{ab}}_{T_2} \\ T_1 &= \text{Background known solution} \\ T_2 &= \text{Perturbation} \end{aligned} \tag{9.81}$$

We only keep terms linear in h:

$$\delta T^a_{bc} = \frac{1}{2}(g_0^{ad} - h^{ad})(g_{0db,c} + h_{db,c}) \tag{9.82}$$

We can choose normal coordinates, such that the connection is zero as all first order derivatives of g are zero in normal coordinates (remember, it is the second order derivative that cannot be set to zero, by a coordinate transformation) and therefore the partial derivative is the same as the covariant derivative. In this case:

$$\begin{aligned} \frac{1}{2}(g_0^{ad} - h^{ad})(g_{0db,c} + h_{db,c}) &= \frac{1}{2}(h^{ad}h_{0db,c}) \\ &= \frac{1}{2}(\nabla_c h^a_b + \nabla_b h^a_c - \nabla^a h_{bc}) \end{aligned} \tag{9.83}$$

But since this is a tensor equation, it must be the same in all coordinate systems:

$$\delta\Gamma^a_{bc} = \frac{1}{2}(\nabla_c h^a_b + \nabla_b h^a_c - \nabla^a h_{bc}) \tag{9.84}$$

Therefore the perturbation of the Ricci tensor:

$$\delta R_{ab} = \nabla_c \delta\Gamma^c_{ab} - \nabla_b \delta\Gamma^c_{ac} \tag{9.85}$$

From the Palatini lemma we known how to compute the variation of this Ricci tensor:

$$\delta R_{ab} = \frac{1}{2}\nabla_c\nabla_a h^c_b + \frac{1}{2}\nabla_c\nabla_b h^c_a - \frac{1}{2}\nabla_c\nabla^c h_{ab} - \underbrace{\frac{1}{2}\nabla_b\nabla_a h}_{T_\alpha} \quad (9.86)$$

The first three terms come from T_1 in Eq 9.85 by substituting Eq 9.84 into it. The final terms is the covariant derivative of the trace of the perturbation in Eq 9.84. we can re-order the derivatives using the Riemann identity:

$$\begin{aligned}\delta R_{ab} &= \frac{1}{2}\nabla_a\nabla_c h^c_b + \frac{1}{2}R^c_{dca}h^d_b + R_{bdca}h^{cd} \\ &+ \frac{1}{2}\nabla_b\nabla_c h^c_a + \frac{1}{2}R^c_{dcb}h^d_a - \frac{1}{2}\Box h_{ab} - \frac{1}{2}\nabla_b\nabla_a h \\ &= -\frac{1}{2}(\Box h_{ab} + 2R_{acbd}h^{cd} - 2R^d_{(a}h_{b)d} - \nabla_{(a}\nabla^d\bar{h}_{b)d})\end{aligned} \quad (9.87)$$

where:

$$\begin{aligned}\bar{h}_{bd} &\equiv h_{bd} - \frac{1}{2}hg_{0bd} \\ &\equiv -\frac{1}{2}\nabla_L h_{ab}\end{aligned} \quad (9.88)$$

$$\nabla_L \equiv \Box h_{ab} + 2R_{acbd}h^{cd} - 2R^d_{(a}h_{b)d} - \nabla_{(a}\nabla^d\bar{h}_{b)d} \quad (9.89)$$

This is called the *Lichneronicz* operator and its the curved space wave operator for a tensor h_{AB}. If the Einstein vacuum equations are true, then this term must be zero. There is a problem with this formalism however, and that is to do with the meaning if h_{ab} being "small". As it cannot be said that every component of h_{ab} is small, for example in the Schwarzschild solution, the metric has no off diagonal elements, therefore any finite component of h_{ab} will be much greater than 0. And hence the meaning of h_{ab} being small has to be evaluated from the context it is being applied to. Also of the background metric has singularities any finite value of h_{ab} components will appear "small" compared to ∞. For the flat space, Minkowski metric it is obvious what to do, since all the components of η_{ab} are 1 in magnitude, the condition on the perturbation is simply:

$$|h_{ab}| << 1 \quad (9.90)$$

4.1. Gauge choices. As stressed before, coordinates are important in general relativity as we are finding exact solutions and coordinates are chosen to make the solutions as simple as possible. A gauge transformation can be written as:

$$x^a \to x^a + \xi^a \quad (9.91)$$

so we move a small distance ξ^a along a vector field:

$$g_{ab} \to g_{ab} + \mathcal{L}_\xi g_{ab} \quad (9.92)$$

$\mathcal{L}_\xi g_{ab}$ is the Lie derivative of g along ξ, which is the symmetrised covariant derivative of ξ:

$$g_{ab} \to g_{ab} + 2\nabla_{(a}\xi_{b)} \quad (9.93)$$

Therefore if ξ was a killing vector, g, would not change. Define:

$$h_{\xi ab} \equiv \nabla_a\xi_b + \nabla_b\xi_a \quad (9.94)$$

So a coordinate transformation induces a perturbation in the metric, which was what makes doing perturbation theory in gravity quite challenging. This can be used to fix a gauge to make problems easier. From Eq 9.94:

$$\bar{h}_{\xi ab} = \Box\xi_b + \underbrace{\nabla^a\nabla_b\xi_a}_{T_1} - \underbrace{\nabla_b(\nabla\cdot\xi)}_{T_2} \tag{9.95}$$

By swapping the order of differentiation in T_1, we can cancel out T_2, using the Riemann identity:

$$\nabla_a \bar{h_{\xi ab}} = \Box\xi_b + R^a_b\xi_a \tag{9.96}$$

This is again a wave operator on a vector field, but corrected for the fact that the space is curved. This gives a well posed set of differential equations:

$$\Box\xi^a + R^a_b\xi^b = V^a(\equiv \nabla^a\bar{h}_{\xi ab}) \tag{9.97}$$

where V^a has been introduce as a source, if the diverges of $\bar{h}_{\xi ab}$ is non-zero. This means, we can solve Eq 9.97 to obtain:

$$\nabla^a\bar{h}_{\xi ab} \equiv \nabla_a\bar{h}^{ab} = 0 \tag{9.98}$$

A *de Donder* gauge is defined by:

$$\delta R_{ab} = -\frac{1}{2}\left(\Box h_{ab} + 2R_{acbd}h^{cd} - 2R^c_{(a}h_{b)c}\right) \tag{9.99}$$

We still have some gauge freedom:

$$x^a \to x^a + \chi^a \tag{9.100}$$

such that:

$$\Box x^a + R^a_b\chi^b = 0 \tag{9.101}$$

Let's count the degrees of freedom:

$$\begin{aligned} h_{ab} &= \frac{D(D+1)}{2} && \text{Components} \\ \nabla_a\bar{h}^{ab} &= D && \text{Constraints} \\ \chi^a &= D && \text{Constraints} \end{aligned} \tag{9.102}$$

Therefore the total physical degrees of freedom are:

$$\frac{D(D+1)}{2} - 2D = \frac{D(D-3)}{2} \tag{9.103}$$

In 4D, there are 2 physical propagating degrees of freedom. Going back to the instability of the black string. For perturbations, decompose with respect to symmetries of background metric. The black string has a symmetry of rotations on a sphere (SO(3)), the killing vectors are $\frac{\partial}{\partial t}, \frac{\partial}{\psi}$. This means the perturbation are written in terms of e^{iwt} or e^{wt} (for instability), and $e^{\frac{2\pi i\mu\psi}{L}}$. In principle the perturbations could have angular momentum. In general we can have:

- A scalar perturbation, h_{ss}.
- A vector perturbation, $h_{s\mu}$.
- A tensor perturbation, $h_{\mu\nu}$.

The equations of motion for the scalar perturbation h_{ss}:

$$\begin{aligned} \nabla_L h_{ss} &= (\Box_4 + \mu^2)h_{ss} \\ &= -Vh''_{ss} - \frac{2(r-GM)}{r^2}h'_{ss} + \left(\mu^2 + \frac{\Omega^2}{V}\right)h_{ss} = 0 \end{aligned} \tag{9.104}$$

where:

$$V \equiv 1 - \frac{2GM}{r} \tag{9.105}$$

μ^2 is like the Fourier mode for the extra dimension. As:

$$\begin{aligned} r &\rightarrow \infty \rightarrow h_{ss} \propto e^{\pm\sqrt{\Omega^2+\mu^2 r}} \\ r &\rightarrow 2GM \rightarrow h_{ss} \propto (r-2GM)^{2GM\Omega} \end{aligned} \tag{9.106}$$

In this case the scalar mode goes to zero at both $r \rightarrow 2GM$ and $r \rightarrow \infty$ as $h_{ss} \rightarrow 0$. Therefore there has to be a turning point from one solution to the other. From Eq 9.104, at a turning point, h'' is positive if h is positive, therefore one has a minimum. If h'' is negative, h is negative and we have a maximum. However if h goes from 0 to 0, one has to have a maximum, which is a contradiction to the equation of motion in Eq 9.104. Therefore there cannot be an instability in the scalar mode.

The vector mode gives a similar expression. Looking at the tensor mode:

$$h_{\mu\nu} = e^{\Omega t} e^{i\mu z} \begin{pmatrix} h_{tt} & h_{tr} & \multicolumn{2}{c}{0} \\ h_{tr} & h_{rr} & & \\ \multicolumn{2}{c}{0} & K & 0 \\ & & 0 & K\sin^2\theta \end{pmatrix} \tag{9.107}$$

This does not depend on any angles. Again one has to use the gauge conditions:

$$\nabla_a \bar{h}^{ab} = 0 \tag{9.108}$$

This can be used to re-express all of these components as a single function:

$$h(r) e^{\Omega t} e^{i\mu z} \tag{9.109}$$

Then one has to check the issue about regularity and weather h remains small. It turns out that there is a particular function of $\Omega = \Omega(\mu)$, where μ is the wave-number and satisfies:

$$\mu < \mathcal{O}\left(\frac{1}{GM}\right) \tag{9.110}$$

In other words, when the length is bigger then r_s. The horizon of the black string wobbles:

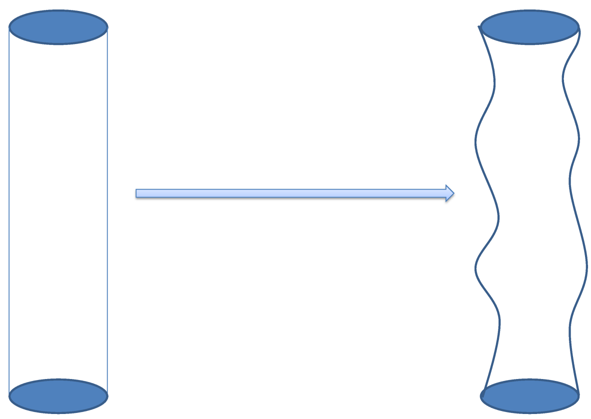


FIGURE 53. Wobble of black string

This method of perturbation theory cannot be used to see what happens beyond the linear regime. It is simply hinting that it appears large for large $L(< /(GM))$.

5. Interpreting metrics

Consider four observers; A, I, A^+, D:

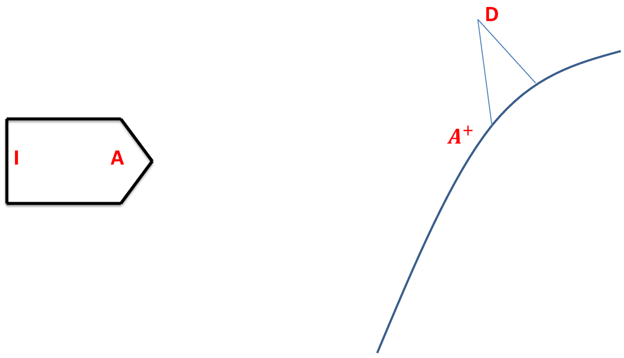


FIGURE 54. Observes A, I are in an accelerating frame in a rocket, and observers A^+, D are on a spherically mass, like the earth, and D is on a higher surface then A^+.

The gravitational field of the earth describes a Schwarzschild metric. The world line of an accelerating observer can be parametrised by:

$$x^\mu = \left(\frac{1}{a}\sinh a\tau, \frac{1}{a}\cosh a\tau, 0, 0\right) \tag{9.111}$$

in flat space (i.e the coordinates are Cartesian). From Eq 9.111 it follows:

$$\begin{aligned} \dot{x}^\mu &= (\cosh a\tau, \sinh a\tau, 0, 0) \\ \ddot{x}^\mu &= a^2 x^\mu \end{aligned} \tag{9.112}$$

The absolute value of the acceleration is a (as there is an $\frac{1}{a}$ in x^μ). Therefore this is indeed a uniformly accelerating observer.

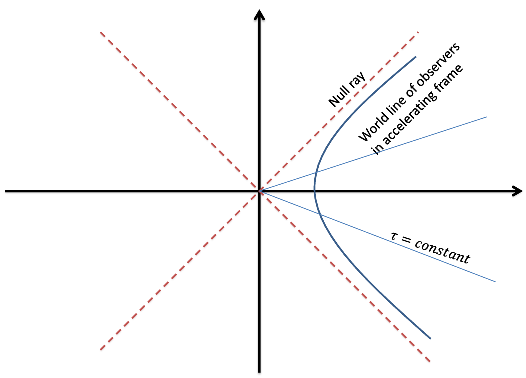


FIGURE 55. Space-time diagram for accelerating frame.

After a given amount of time the observer on the world line above will not be able to communicate with inertial observers as the world line will asymptotically approach the null ray. Now we can change to coordinates that are centered on the observers:

$$\begin{aligned} t &= \rho \sinh(a\tau) \\ x &= \rho \cosh(a\tau) \\ dt^2 - dx^2 &= a^2\rho^2 d\tau^2 - d\rho^2 \end{aligned} \tag{9.113}$$

when $\rho \to 0, g_{tt} \to 0$ and we get something like a horizon. The observer is at $\rho = \frac{1}{a}$, in this coordinate system and this agrees with the definition of x^μ. ρ represents the distance of an observer accelerating at a constant rate. Suppose we center the coordinate system on observer A, we assume I is at a constant distance from A, which would represent a constant ρ.

If I is at a constant distance, then its acceleration is $\frac{1}{a^{-1}+d}$ or $\frac{a}{1+ad}$, where d is the distance between A and I.

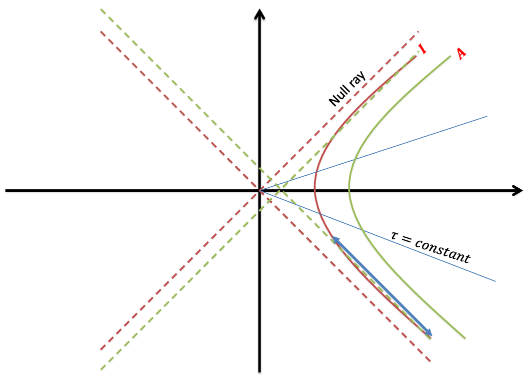


FIGURE 56. Space-time diagram for two accelerating frames. The region marked by blue line represents where I and A are in causal constant

Since I is accelerating at the same rate A, I will also have his our horizon, and eventually I and A will go out of causal constant. Now lets look at $A^+\&D$, the coordinates for A^+:

$$\begin{aligned} x^\mu_A &= (t, r_e, 0) \\ \dot{x}^\mu_A &= (\dot{t}, 0) \\ \dot{t} &= V_{r_e}^{-\frac{1}{2}} = \left(1 - \frac{2GM}{r_e}\right)^{-\frac{1}{2}} \end{aligned} \tag{9.114}$$

The second derivative can be written as:

$$\begin{aligned} \ddot{x}^\mu_A &= \dot{A}^\nu_A \nabla_\mu \dot{x}^\mu \\ &= \frac{\partial}{\partial\tau}\dot{x}^\mu_A + \Gamma^\mu_{\nu\lambda}\dot{x}^\nu_a\dot{x}^\lambda_\mu \end{aligned} \tag{9.115}$$

since $\dot{x}$ is a constant, T_1 is zero and:

$$\Gamma^\mu_{\nu\lambda}\dot{x}^\nu_a\dot{x}^\lambda_\mu = \Gamma^r_{tt}V_e^{-1}\delta^\mu_r = \frac{GM}{r_e^2} \tag{9.116}$$

For observer D, that is a distance d away from the surface of the planet:

$$\ddot{x}^\mu_D \approx \frac{GM}{(r_E+d)^2} = \frac{GM}{r_E^2}\left(1-\frac{2d}{r_E}\right) \approx 10^{-6} d \text{ metres} \quad (9.117)$$

This shows that the observers separated by a distance feel different accelerations.

5.1. Acceleration and gravity. Here we will look at frames that are accelerating but no longer in flat space, but in curved space-time. Let's start with the metric:

$$\begin{aligned} ds^2 &= \frac{1}{A^2(xy)^2}\left[f(y)\partial\tau^2 - \frac{\partial y^2}{f(y)} - \frac{\partial x^2}{g(x)} - g(x)d\phi^2\right] \\ f(y) &= -1+y^2-2GMAy^3 \\ g(x) &= 1-x^2-2GMAx^3 \end{aligned} \quad (9.118)$$

This metric does not look familiar. First thing to notice; there are two killing vectors, $\frac{\partial}{\partial t}, \frac{\partial}{\partial\psi}$. Let's take ψ to be a periodic coordinate aswell. Next thing to check is the ranges over which these coordinates are defined. For the metric to be Minkowski, like in the signature, f and g must be positive.

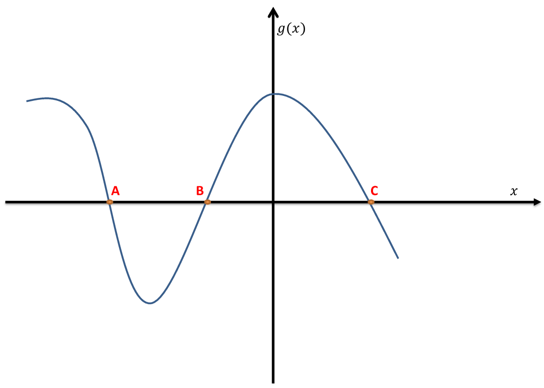


FIGURE 57. Form of $g(x)$

$g(x)$ is a cubic, therefore can have 3 roots. If $2GM$ is small, then the roots will be close to ± 1. If $GMA = 0, x = \cos\theta$, which gives $d\Omega_2^2$ for x, ψ part of the metric, therefore we can take x to be between B and C:

$$x \in [B, C] \quad (9.119)$$

Similarly:

$$f(y) \geq 0 \rightarrow y \in [-B, -A] \quad (9.120)$$

if $2GMA << 1$, then root $-A$ is roughly:

$$-A \approx \frac{1}{2GMA} \quad (9.121)$$

So the y coordinate looks like it has a limit, roughly 1 or $\frac{1}{2GMA}$. This looks like an event horizon and if $r = 2GM$ is an event horizon, then it suggests that we try a coordinate transformation:

$$r \equiv \frac{1}{Ay} \Rightarrow dr = -\frac{1}{Ay^2}dy \quad (9.122)$$

Then the root looks like $r \approx 2GM$, of course these are all orders of magnitude calculation for the roots of the cubic. With this transformation, $f(g)$ becomes:

$$\begin{aligned} f(y) &= -1 + \frac{1}{A^2r^2} - \frac{2GM}{Ar^2} \\ &= \frac{1}{A^2r^2}\left(1 - A^2r^2 - \frac{2GM}{r}\right) \end{aligned} \tag{9.123}$$

and the metric is:

$$ds^2 = \frac{1}{(1+Arx)^2}\left[\left(1 - \frac{2GM}{r} - A^2r^2\right)\frac{\partial t^2}{A^2} - \frac{dr^2}{\left(r - \frac{2GM}{r} - A^2r^2\right)} - \text{(same terms as before)}\right] \tag{9.124}$$

Comparing to the Schwarzschild, de Sitter black holes, this metric looks familiar, with A being a potential. But the solution we write down was a solution to the Einstein vacuum equations, therefore there was no cosmological constant in there, Eq **??**, has two horizons:

$$\begin{aligned} r_1 &\approx 2GM \rightarrow \text{Black hole} \\ r_2 &\approx \frac{1}{A} \rightarrow \text{Acceleration horizon} \end{aligned} \tag{9.125}$$

Going back to the x and ψ coordinates. When $x = \cos\theta$, this looks roughly like a sphere not exactly. Again we have two coordinate singularities. One is close to $x = 1$, and the other is close to $x = -1$. If this were a sphere and we didn't have this GMA terms of the $g(x)$ function, we would say $x = \cos\theta$ and not worry about the north pole and south pole singularities, as they are easy to remove. Here we have a north pole and a south pole, but because the metric function has been changed, one cannot assume that the singularities at the north and south pole are the same as they would be with any other sphere. So one has to examine:

$$\begin{aligned} x &\rightarrow B \quad \text{(south pole)} \\ x &\rightarrow C \quad \text{(north pole)} \end{aligned} \tag{9.126}$$

In each case, we have to check that the metric is regular. Near $x = B$:

$$g(x) \approx g'(B)(x - B) \tag{9.127}$$

where g' is the gradient of g. Defining:

$$\begin{aligned} \theta_B &\equiv 2\sqrt{\frac{x-B}{g'_B}} \\ d\theta_B^2 &\equiv \frac{dx^2}{g'B(x - x_B)} \approx \frac{\partial x^2}{g(x)} \end{aligned} \tag{9.128}$$

So we have re-defined a coordinate that looks like a radius. This looks like the usual south pole, but only if ψ has correct periodicity. Near the south pole, θ_B:

$$g(x)d\psi^2 \approx g'_B(x - B)d\psi^2 = g'^2_B\frac{\theta_B^2}{4}d\psi^2 \tag{9.129}$$

To identify a ϕ_B angle:

$$\phi_B = \frac{|g'_B|}{2}\psi \tag{9.130}$$

ϕ_B would have periodicity 2π. If we look at g':

$$g' = -2x - 6GMAx^2 \tag{9.131}$$

Looking at the g' at both roots:

$$|g'_B| \neq |g'_C| \Rightarrow 2 - 6GMA \neq 2 + 6GMA \tag{9.132}$$

In other words, if the periodicity is chosen in the ψ coordinate such that it is regular at one pole, then it is not regular at the other pole. This tells us that the periodicity is not exactly 2π, instead this is a conical singularity. A canonical singularity can be imagined by imagining a sheet of paper, cutting out a wedge and fold it up to form a cone. The key feature of a conical singularity is that the constant r is not $2\pi r$, this can be seen them the metric as it has a $(1-\delta)$ factor multiplying the r^2 term. There exists a physical solution that gives a canonical singularity (at large scales):

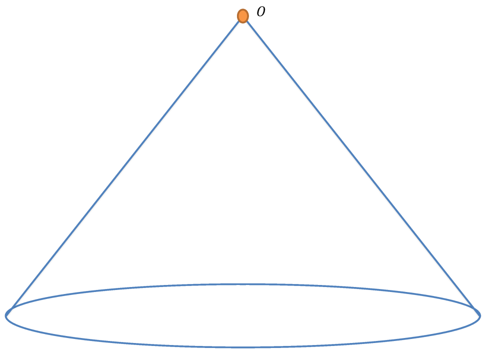


FIGURE 58. Conical singularity on black string

The singularity at point O is smoothed out over large distances, This is object is known as a *cosmic string*. This has many singularities with the domain wall. A cosmic string forms when the vacuum has non-trivial $U(1)$ or circle. This arises in the abelian Higgs model. This solution has feature of a black hole, and acceleration and these conical singularities on one side (i.e south or north pole).

So this metric actually represents a black hole that is accelerating of to infinity, and it is pulled by a cosmic string. Just like the domain wall, cosmic strings also have a stress energy tensor that is proportional to the induced metric on this string world sheet. Therefore the string has a very strong tension, so the interpretation is that it is this tension of the cosmic string that is accelerating this black hole. As shown in Eq 9.132, the difference in the metric at both points is of order GMA, therefore the greater the mass of the black hole, the greater the canonical deficit, and the greater the acceleration, A, the greater the deficit. Which is just saying that if the tension in the string is bigger, then the acceleration is bigger. If the black hole is heavier, then it will take more effort to pull the black hole of the same acceleration, which again gives the bigger difference between the metrics, so this picture is actually quite intuitive when thought about psychically. This metric is called a *C-metric*.

6. Gravitational instantons

Several quantum affects in general relativity have been discussed so far, however the underlying frame work or field theory has always been classical. Here we discuss the process of tunneling of the space-itself. This was first discussed by *Sidney Coleman*[**10**] and then *Frank de Luccia*[**9**] and it is the same method that is followed here.

An instanton is an event centered on an instant in time, but this time is *Euclidean time*. Recall the tunneling in QM; as long as $V_0 < \infty$, the incoming wave with energy E will be able to go through the potential, even if $E < V_0$, with a finite probability:

$$
\begin{aligned}
|T|^2 &= \frac{1}{1+\frac{\sinh^2 \Omega d}{4E(V_0-E)}} \approx e^{-2\Omega d} \\
\Omega^2 &= \frac{2M(V_0-E)}{\hbar^2} \\
\Omega_d &= \frac{1}{\hbar}\int_0^d \sqrt{2m(V-E_0)}dx
\end{aligned}
\tag{9.133}
$$

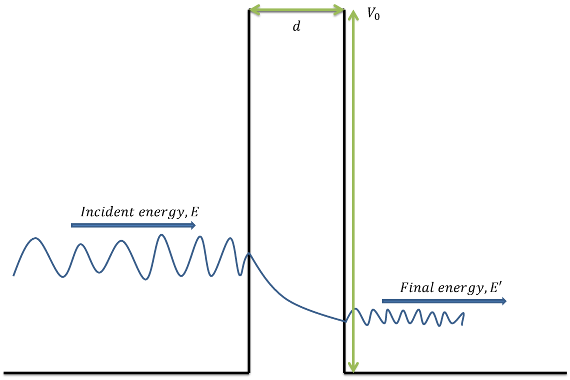


FIGURE 59. Quantum mechanical tunneling

Ω_d is like the volume of the barrier. This can be compared to a classical particle moving in a well:

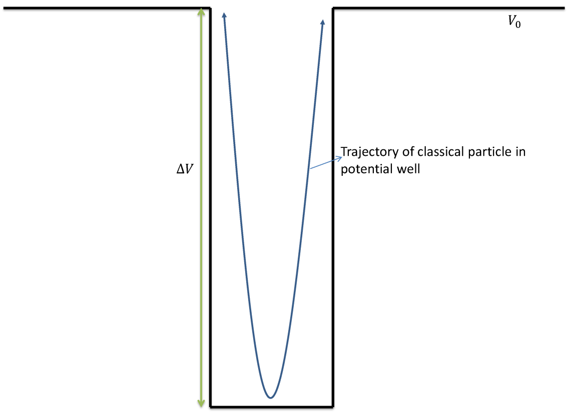


FIGURE 60. Classical motion in potential well

The motion in this classical potential well would be described by:

$$
\begin{aligned}
\frac{\dot{x}^2}{2} &= \Delta V \\
\int \sqrt{2\Delta V}dx &= \int 2\Delta V d\tau \\
&= \int \left(\Delta V + \frac{1}{2}\dot{x}^2\right) d\tau
\end{aligned}
\tag{9.134}
$$

Therefore by comparing Eq 9.133 and Eq 9.134, the same equations seems to describe the motion of a classical particle in a well and a quantum particle in a well and a quantum particle in a barrier. Generally to compute a tunneling amplitude, we take classical path of particle moving in an inverted potential. The idea is that in computing the equations of motion of a tunneling process, one is analytically continuing to Euclidean time and then calculating a classical problem. The tunneling amplitude is given by:

$$e^{-S_B}\hbar \tag{9.135}$$

where S_B is the Euclidean action of this "bounce solution". A particle, calculation tunneling amplitude is quite simple, but in field theory, the easiest way of computing is to look at the Euclidean equation of motion, calculate the classical bounce problem and its corresponding Euclidean action and then we can obtain the tunneling amplitude from Eq 9.135.

6.1. False vacuum decay. The Lagrangian of a field, ϕ, with a false vacuum is:

$$L_\phi = \frac{1}{2}(\partial\phi)^2 - V(\phi) \tag{9.136}$$

where ϕ is described by the potential.

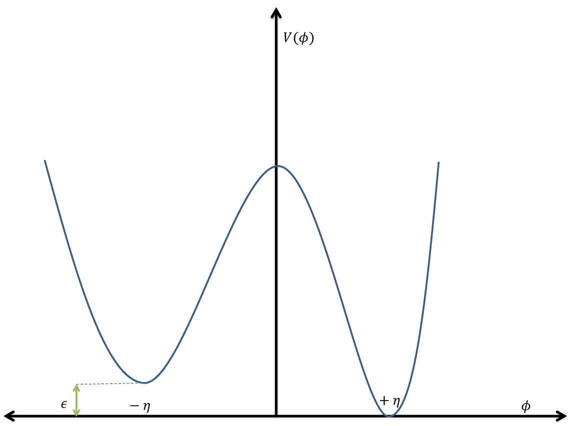


Figure 61. False vacuum potential

The potential is:

$$V(\phi) = \frac{\lambda}{2}(\phi^2 - \eta^2)^2 - \frac{\epsilon(\phi - \eta)}{2\eta} \tag{9.137}$$

The $-\eta$ has a "false vacuum" state as the potential is not actually a minimum:

$$\begin{aligned} \phi &= -\eta \equiv \text{ False vacuum }, V \approx \epsilon \\ \phi &= +\eta \equiv \text{ True vacuum }, V \approx 0 \end{aligned} \tag{9.138}$$

If we are in the false vacuum, then in quantum mechanics the wavefunctions would be affected by the true vacuum at $\phi = +\eta$ and therefore it is possible to tunnel through to the true vacuum. In this situation the False vacuum state is metastable (i.e it is stable to small perturbations), however a tunneling process might take it into the true vacuum.

Now we look for a Euclidean time ($t \to i\tau$) solution:

$$\frac{\partial^2\phi}{\partial\tau^2} + \nabla^2\phi = \frac{\partial V}{\partial\phi} = 2\lambda\phi(\phi^2 - \eta^2) + \mathcal{O}(\epsilon) \tag{9.139}$$

Asymptotically, the idea of a bounce means we have to imagine that the background state is going to be a false vacuum (FV), but there should be a true vacuum (TV), solution in the false vacuum.

In other words, if we go along from left to right in the schematic solution in Figure 62, we go from the false vacuum, to true vacuum, to false vacuum. In the solution, spherical symmetry is expected, which means we can go from a 4D partial differential equation to a 1D ordinary differential equation:

$$\begin{aligned} \phi'' + \frac{3}{\rho}\phi' &= 2\lambda\phi(\phi^2 - \eta^2) \\ \rho^2 &\equiv \tau^2 + \vec{x}^2 \end{aligned} \tag{9.140}$$

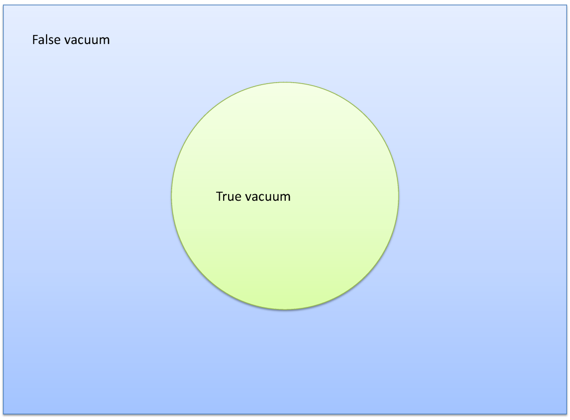


FIGURE 62. Solutions to equations of motion in terms of Euclidean time

The potential has a small energy difference, ϵ, between vacuum solutions, relative to the size of the barrier. Therefore one might expect that ϕ is roughly constant inside the true vacuum and false vacuum and goes through the potential barrier between the true and false vacuum quickly in some region. This is like a thin wall problem. Suppose ϕ varies rapidly in a thin region of ρ, centered on $\rho >> \frac{1}{\sqrt{\lambda}\eta}$. This gives the same solution as the domain wall:

$$\phi \approx \eta \tanh(\sqrt{\lambda}\eta(\rho - \rho_0)) \tag{9.141}$$

As long as we are not in the limit $\sqrt{\lambda}\eta$ is large, ϕ is essentially constant and the only deviation is around $\rho \approx \rho_0$. The action for this solution is:

$$S_B = \underbrace{2\pi^2\rho_0^3\sigma}_{\text{Wall action}} - \underbrace{\frac{\pi^2\epsilon\rho_0^4}{3}}_{\text{Action inside true vacuum}} \tag{9.142}$$

σ is the energy per unit area. If ρ is a solution to the equations of motion, then ρ_0 will be a stationary point of the action, i.e:

$$\frac{\partial S_B}{\partial \rho_0} \equiv 0 \tag{9.143}$$

which gives:

$$2\pi^2\rho_0^2(3\sigma - \epsilon\rho_0) = 0 \Rightarrow \rho_0 \frac{3\sigma}{\epsilon} \tag{9.144}$$

Substitute this into Eq 9.142:

$$S_B = \frac{27\sigma^4\pi^2}{2\epsilon^3} \tag{9.145}$$

This is the bounce action. Since we assumed that ϵ is a lot smaller than the overall potential scale, which is represented by σ, the value of ρ, will always be greater than 1:

$$\rho_0 >> 1 \tag{9.146}$$

hence that assumption that everything happens near ρ_0 is a good one (as the tanh function will then start to move away from a constant value in Eq 9.141). Therefore in Figure 62, we have constructed a bubble of the true vacuum in Euclidean space, $\mathbb{R}^4$, which sits at some fixed ρ_0. So we have a true vacuum for $\rho < \rho_0$ a wall at ρ_0, and a fixed vacuum at $\rho > \rho_0$. We can now continue analytically. Recall:

$$\rho^2 = \tau^2 + \vec{x}^2 \equiv \rho_0^2 (\text{At wall}) \tag{9.147}$$

which is, in terms of Lorentzian (i.e non-Euclidean) time:

$$\rho^2 = -t^2 + \vec{x}^2 \equiv \rho_0^2 (\text{At wall}) \tag{9.148}$$

Which is a hyperboloid in Lorentzian flat space:

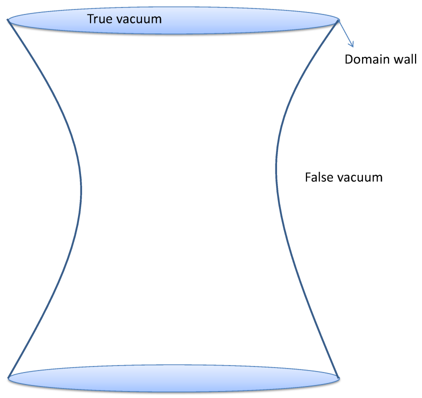


Figure 63. Hyperboloid in Lorentzian flat space

The tunneling process can be thought of as a Euclidean bubble forming in the past and then expanding into a Lorentzian space at a given time. To match the solutions of the Euclidean space and the Lorentzian space, the surface needs to be mapped to a surface which has no time derivatives (as the real and imaginary parts can only equal each other if the solution is trivial). In the false vacuum state, the energy momentum tensor is:

$$\begin{aligned} T_{\mu\nu} &= \phi_{,\mu}\phi_{,\nu} - L g_{\mu\nu} \\ &= V g_{\mu\nu} \\ &= \epsilon g_{\mu\nu} \end{aligned} \tag{9.149}$$

Therefore the false vacuum is not flat (as there is a cosmological constant), instead it is a de-Sitter space-time, and the length scale, L, is:

$$L^2 = \frac{3}{\Lambda} = \frac{3}{8\pi G\epsilon} \tag{9.150}$$

where the $8\pi G$ has been put in to obtain the correct units. If ϵ is small, the length scale of the de-Sitter space time is big, thus the metric is nearly flat, however the geometry is not exactly flat and that needs to be accounted for, should one require a consistent solution. Gravitationally, we need to replace the hyperboloid with dS. Now we can use the Gauss-Cadazzi formalism, for inside the bubble and outside.

- **Inside**, the metric is:

$$ds^2 = \partial\rho^2 + \rho^2 \partial\Omega_3^2 \tag{9.151}$$

Wall at ρ_0, and an outward normal is $\frac{\partial}{\partial\rho}$. Therefore from the Israel equations, the extrinsic curvature is:

$$K^-_{\alpha\beta} = -\Gamma^\rho_{\alpha\beta} \qquad (\alpha, \beta \in \{1,2,3\} \text{ on } S^2) \tag{9.152}$$

- **Outside**, we have Euclidean de-Sitter space:

$$ds^2 = (\partial\chi^2 + \sin^2\chi\partial\Omega_3^2)L^2 \tag{9.153}$$

This is a 4 sphere of radius, L. Where:

$$L = \frac{3}{8\pi G\epsilon} \tag{9.154}$$

Therefore in this metric, the wall is at χ_0 and here we need an inward normal (as we are looking into the true vacuum, Minkowski space), so the inward normal is:

$$\frac{1}{L}\frac{\partial}{\partial\chi} \tag{9.155}$$

and:

$$K^+_{\alpha\beta} = -\frac{\Gamma^\chi_{\alpha 0}}{L} = \frac{\cot\chi_0}{L}g_{\alpha\beta} \tag{9.156}$$

Combining the two solutions for inside and outside solutions, the intrinsic/induced metrics must be the same on each side, therefore equating Eq 9.151 and Eq 9.153:

$$\rho_0 = L\sin\chi_0 \tag{9.157}$$

By equating the metrics at the wall. From the Israel equation, Eq 8.124:

$$\underbrace{\frac{\cot\chi_0}{L}}_{K^+_{\alpha\beta}} - \underbrace{\frac{\csc\chi_0}{L}}_{K^-\alpha\beta} = -4\pi G\sigma \tag{9.158}$$

Which can be re-written as:

$$4\pi G\sigma L = \frac{(1-\cos\chi_0)}{\sin\chi_0} \tag{9.159}$$

So we get a relation between the tension of the wall, σ, the curvature of the de-Sitter space and the place at which the two different metrics are being identified. The next step is to calculate the action of the Euclidean solution:

$$S_E = -\int_{M^+}(R+2\Lambda) - \int_{M_-}(R+2\Lambda) + \int_{\partial M^+}\frac{K^+}{8\pi G} - \int_{\partial M_-}\frac{K^-}{8\pi G} + \int_{\text{Wall}}\sigma \tag{9.160}$$

This is the full action, but we also have to subtract off what was already present:

$$S_b = \int_{M_-} -\frac{(R+2\Lambda)}{16\pi G} + \int_{\text{Wall}}\left(\frac{\Delta K}{8\pi G} + \sigma\right) \tag{9.161}$$

For the de-Sitter space:

$$R = -4\pi\Lambda \tag{9.162}$$

where:

$$\Lambda = 8\pi G\epsilon \tag{9.163}$$

. From the Israel equation, $\Delta K = -12\pi H\sigma$:

$$\begin{aligned} S_B &= 2\epsilon\pi L^4 \int_0^{\chi_0} \sin^3\chi d\chi \\ &= -\frac{\sigma}{2}2\pi^2\rho_0^3 \end{aligned} \tag{9.164}$$

Replace ϵ in terms of L:

$$\begin{aligned} S_B &= \frac{3\pi}{4G}L^2\left[\frac{1}{3}\cos^3\chi_0 - \cos\chi_0 + \frac{2}{3}\right] - \pi^2 L^3 \sin^3\chi_0 \frac{(1-\cos\chi_0)}{(\sin\chi_0)4\pi GL} \\ &= \frac{\pi L^2}{4G}(1-\cos\chi_0)^2 \end{aligned} \tag{9.165}$$

Which can be re-written as:

$$S_B = \frac{\pi L^2(4\pi G\sigma L)^4}{G(1+(4\pi G\sigma L)^2)^2} \tag{9.166}$$

This is known as the Coleman-de Luccia result and its the action of a gravitational instanton which goes from de-Sitter space to Minkowski space. It was this procedure of metastability of false vacuum solution that first led Alan Guth to propose the inflationary model.

Part 3

Cosmology

CHAPTER 10

Metrics of the cosmos

Cosmology is arguably the broadest subject one can have, as it is about the universe as a whole. It brings together all other disciplines of physics to describe phenomena, from stars to black holes to the cosmic microwave background and the Big bang itself.

1. The important metrics

There are three metrics that are very useful in cosmology:

- Maximally symmetric metrics:

$$\begin{aligned} E_n &= n \text{ dimensional Euclidean space} \\ S_n &= n \text{ dimensional sphere} \\ H_n &= n \text{ dimensional hyperbolic space} \\ dS_n &= n \text{ dimensional de-Sitter space} \\ M_n &= n \text{ dimensional Minkowski space} \\ AdS_n &= n \text{ dimensional Anti-de-Sitter space} \end{aligned} \tag{10.1}$$

  All of these spaces are actually related to each other.
- The Friedmann-Robertson-Walker, FRW, metric. As far as we can observe, the universe at large scales appears to be described by this metric. In this metric, every 3D slice of this space, is one of the maximally symmetric spaces, which can be scaled in an arbitrary way (usually as time evolves).
- Black hole metrics, there are two very important ones; the Schwarzschild metric for a non-rotating black hole and the Kerr metric, which describes a rotating black hole.

1.1. ***n*** **dimensional sphere metric.** To begin with, let's construct the metric for an n dimensional sphere. When we imagine a 2D sphere, it is embedded in a flat 3D space. More generally when we want to construct the n dimensional sphere, it is embedded in a $n+1$ dimensional space. Consider:

$$x^A \quad (A = dim(n+1)) \tag{10.2}$$

In general (for this subsection), capital indices will run over $n+1$ dimensions, i.e the dimension of the embedding space. Lower case indices run over n dimensions, i.e the dimension of the manifold. The metric of the embedding space is just a flat Euclidean, $n+1$ dimensional space:

$$\begin{aligned} ds^2 &= \delta_{AB} dx^A dx^B \\ &= \delta_{ab} dx^a dx^b + (dx^{n+1})^2 \end{aligned} \tag{10.3}$$

Now we consider, a surface defined by an equation:

$$\begin{aligned} \rho^2 &= \delta_{AB} x^A x^B \\ &= \delta_{ab} x^a x^b + (x^{n+1})^2 \end{aligned} \tag{10.4}$$

Now, we can solve for (x^{n+1}) in this expansion:

$$x^{n+1} = \pm(\rho^2 - \delta_{ab}x^a x^b)^{\frac{1}{2}} \tag{10.5}$$

Differentiating:

$$dx^{n+1} = \pm\frac{1}{2}(\rho^2 - \delta_{ab}x^a x^b)^{\frac{-1}{2}}(-2\delta_{ab}dx^a dx^b) \tag{10.6}$$

Putting this into Eq 10.3:

$$ds^2 = \left(\delta_{ab} + \frac{x_a x_b}{\rho^2 - x^2}\right) dx^a dx^b \tag{10.7}$$

where:

$$\begin{aligned} x^2 &\equiv \delta_{ab}x^a x^b \\ x_a &\equiv \delta_{ab}x^b \\ x_b &\equiv \delta_{ab}x^a \end{aligned} \tag{10.8}$$

This is the metric for the n sphere.

1.2. Constructing a general metric. Now we want to construct an n dimensional manifold with a general signature (p, q), i.e p is a spatial direction and q is a time direction. The curvature, K, can be:

$$K = +1, 0, -1 \tag{10.9}$$

and the radius of curvature is ρ. Again, let's take a flat $n+1$ dimensional embedding space:

$$ds^2 = \eta_{ab}^{(p,q)} dx^a dx^b + K(dx^{n+1})^2 \tag{10.10}$$

where $\eta_{ab}^{(p,q)}$ is the flat Minkowski metric with signature (p, q). We also take the same embedding equation:

$$K\rho^2 = \eta_{a,b}^{(p,q)} x^a x^b + K(x^{n+1})^2 \tag{10.11}$$

Now the same procedure used for the sphere is used; eliminate x^{n+1} from the embedding equation, differentiate it, put in into the metric of the embedding space, to get an n dimensional line element:

$$ds^2 = \left(\eta_{ab}^{(p,q)} + \frac{K x_a x_b}{\rho^2 - K x^2}\right) dx^a dx^b \tag{10.12}$$

where:

$$\begin{aligned} x^2 &\equiv \eta_{ab}^{(p,q)} x^a x^b \\ x_a &\equiv \eta_{ab}^{(p,q)} x^b \\ x_b &\equiv \eta_{ba}^{(p,q)} x^a \end{aligned} \tag{10.13}$$

This is the metric for maximally symmetric space. A summary of the different spaces is given below:

	$K=+1$	$K=0$	$K=-1$
Signature $(n, 0)$	S_n	E_n	H_n
Signature $(n-1, 1)$	dS_n	M_n	AdS_n

TABLE 8. Comparison of different maximally symmetric metrics

Maximally symmetric metrics have a signature (p, q) with $q \geq 2$ (i.e two or more time-like directions) are not observed in nature. These types of metrics would have some strange properties. For example, they would have closed time-like curves, therefore it would be (theoretically) possible to travel back in time and this leads to the usual paradoxes, such as the grandfather paradox.

As in general relativity, an important object is the inverse of the metric tensor, as that provides a mechanism for converting contravariant objects to covariant objects and vice-versa. Suppose $\eta^{ab}_{(p,q)}$ is the inverse of $\eta^{(p,q)}_{ab}$, then the inverse of $g^{(p,q)K}_{ab}$ is

$$g^{ab}_{(p,q)K} = \eta^{ab}_{(p,q)} - \frac{Kx^a x^b}{\rho^2} \tag{10.14}$$

This can be explicitly shown as follows.

CLAIM 20.

$$g^{(p,q)K}_{ab} g^{bc}_{(p,q)K} = \delta^c_a \tag{10.15}$$

PROOF 20.

$$\begin{aligned}
g^{(p,q)K}_{ab} g^{bc}_{(p,q)K} &= \left(\eta^{(p,q)}_{ab} + K\frac{x_a x_b}{\rho^2 - Kx^2}\right)\left(\eta^{bc}_{(p,q)} - K\frac{x^b x^c}{\rho^2}\right) \\
&= \eta^{(p,q)}_{ab}\eta^{bc}_{(p,q)} - \eta^{p,q}_{ab} K\frac{x^b x^c}{\rho^2} + \eta^{bc}_{(p,q)} K\frac{x_a x_b}{\rho^2 - Kx^2} - \frac{K^2 x_a x_b x^b x^c}{\rho^2(\rho^2 - Kx^2)} \\
&= \delta^c_a - \frac{Kx_a x^c}{\rho^2} + K\frac{x_a x^c}{\rho^2 - Kx^2} - \frac{K^2 x_a x^c x^2}{\rho^2(\rho^2 - Kx^2)} \\
&= \delta^c_a
\end{aligned} \tag{10.16}$$

Next we calculate the Christoffel symbols.

CLAIM 21. The Christoffel symbols are

$$\Gamma^a_{bc} = K\frac{x^a}{\rho^2} g^{(p,q)K}_{bc}. \tag{10.17}$$

PROOF 21. The generic form of the Christoffel symbols is given in Eq 3.111. Computing the terms individually

$$\begin{aligned}
\frac{\partial g^{(p,q)K}_{bd}}{\partial x^c} &= \frac{\partial}{\partial x^c}\left(\eta^{(p,q)}_{bd} + K\frac{x_b x_d}{\rho^2 - Kx^2}\right) \\
&= \frac{\partial \eta^{(p,q)}_{bd}}{\partial x^c} + \frac{\partial}{\partial x^c}\left(\frac{Kx_b x_d}{\rho^2 - Kx^2}\right) \\
&= \frac{\partial \eta^{(p,q)}_{bd}}{\partial x^c} + \frac{Kx_d}{\rho^2 - Kx^2}\frac{\partial x_b}{\partial x^c} + \frac{Kx_b}{\rho^2 - Kx^2}\frac{\partial x_d}{\partial x^c} + Kx_b x_d \frac{\partial}{\partial x^b}\left(\frac{1}{\rho^2 - Kx^2}\right)
\end{aligned} \tag{10.18}$$

and similarly

$$\frac{\partial g^{(p,q)K}_{cd}}{\partial x^b} = \frac{\partial \eta^{(p,q)}_{cd}}{\partial x^b} + \frac{Kx_d}{\rho^2 - Kx^2}\frac{\partial x_c}{\partial x^b} + \frac{Kx_c}{\rho^2 - Kx^2}\frac{\partial x_d}{\partial x^b} + Kx_c x_d \frac{\partial}{\partial x^b}\left(\frac{1}{\rho^2 - Kx^2}\right), \tag{10.19}$$

$$\frac{\partial g^{(p,q)K}_{bc}}{\partial x^d} = \frac{\partial \eta^{(p,q)}_{bc}}{\partial x^d} + \frac{Kx_c}{\rho^2 - Kx^2}\frac{\partial x_b}{\partial x^d} + \frac{Kx_a}{\rho^2 - Kx^2}\frac{\partial x_c}{\partial x^d} + Kx_b x_c \frac{\partial}{\partial x^d}\left(\frac{1}{\rho^2 - Kx^2}\right). \tag{10.20}$$

Now use the identities; $x^a_{,b} = \delta^a_b$, $x_{a,b} = \eta^{(p,q)}_{ab}$, $x^2_{,a} = 2x_a$, $x^a, x_a = x^2$ and substitute the individual components into Eq 3.111 to get

$$\Gamma^a_{bc} = K\frac{x^a}{\rho^2} g^{(p,q)K}_{bc}. \tag{10.21}$$

Once we have the Christoffel symbols, its relatively straightforward to show the following relations,

$$R^a_{bcd} = \frac{K}{\rho^2}(\delta^a_c g^{(p,q)K}_{bd} - \delta^a_d g^{(p,q)K}_{bc}) \tag{10.22}$$

$$R_{abcd} = \frac{K}{\rho^2}(g^{(p,q)K}_{ac} g^{(p,q)K}_{bd} - g^{(p,q)K}_{ad} g^{(p,q)K}_{bc}) \tag{10.23}$$

$$R_{ab} = \frac{(n-1)K}{\rho^2} g^{(p,q)K}_{ab} \tag{10.24}$$

$$R = \frac{n(n-1)K}{\rho^2} \tag{10.25}$$

Note that these four equations are tensor equations, therefore they are completely general and are valid in any coordinate system. The metrics in these equations are soutions to Einstein's equations in the vacuum, $G_{\mu\nu} + \Lambda g_{\mu\nu} = 0$, where

$$\Lambda = \frac{K(n-1)(n-2)}{2\rho^2} \tag{10.26}$$

2. FRW space-times

From the assumptions of isotropy and homogeneity, FRW space-times are symmetric in all spatial directions, which means that all of the changes in this space can be encompassed in one term, $a(t)$, which is known as the scale factor. The $(n+1)$ dimensional FRW metric can therefore be written as

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -dt^2 + a^{(t)} g^{(n,0)K}_{ij} dx^i dx^j \tag{10.27}$$

By a coordinate transformation of the form $x^i = \rho \tilde{x}^i$ and $\rho a(t) = \tilde{a}(t)$, the ρ in $g^{(n,0)K}_{ij}$ can be set to 1. From Eq 10.14, we can find the inverse metric components

$$g_{00} = -1, \quad g^{0i} = 0, \quad g^{ij} = \frac{g^{ij}_{(n,0)K}}{a^2(t)}. \tag{10.28}$$

And the Christoffel symbols follow

$$\begin{aligned} &\Gamma^0_{00} = \Gamma^i_{00} = \Gamma^0_{0i} = 0 \\ &\Gamma^i_{0j} = H\delta^i_j, \quad \Gamma^0_{ij} = H g_{ij}, \quad \Gamma^k_{ij} = K x^k g^{(n,0)K}_{ij}. \end{aligned} \tag{10.29}$$

The components of the Ricci tensor are

$$R_{00} = -n\frac{\ddot{a}}{a}, \quad R_{0i} = 0, \quad R_{ij} = \left(\frac{\ddot{a}}{a} + (n-1)\left(H^2 + \frac{K}{a^2}\right)\right) g_{ij}. \tag{10.30}$$

Finally the components of the Einstein tensor, $G^\mu_\nu = g^{\mu\rho} R_{\rho\nu} - \frac{1}{2}\delta^\mu_\nu R$, are

$$G_{00} = \frac{n(n-1)}{2}\left(H^2 + \frac{K}{a^2}\right) \tag{10.31}$$

$$G_{0i} = 0 \tag{10.32}$$

$$G_{ij} = -(n-1)\left(\frac{\ddot{a}}{a} + \frac{(n-2)}{2}\left(H^2 + \frac{K}{a^2}\right)\right) g_{ij} \tag{10.33}$$

The stress-energy tensor in FRW space-time is that of an ideal fluid

$$T^{\mu}_{\nu} = \begin{pmatrix} -\rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0P & \end{pmatrix}, \tag{10.34}$$

which can be written as

$$T_{00} = \rho, \qquad T_{0i} = 0, \qquad T_{ij} = pg_{ij}. \tag{10.35}$$

Putting the ingredients back into the Einstein equations, Eq 4.43, gives two of the most important equations in cosmology. First, is the Friedmann equation

$$H^2 = \frac{16G_N\rho + 2\Lambda}{n(n-1)} - \frac{K}{a^2}, \tag{10.36}$$

and the continuity equation,

$$\dot{\rho} = -nH(\rho + P) \tag{10.37}$$

The continuity equation is obtained by the 00 component of the Einstein equation, and its time derivative, to simplify the ii component. There is another (quicker) way to derive this relation from the conservation of the stress-energy tensor, $T^{\mu\nu}_{;\nu} = 0$.

2.1. Standard coordinates for FRW space-time. Since the FRW universe is homogenous and isotropic, the natural coordinates for the FRW metric are the spherical polar coordinates under which the line element becomes

$$ds^2 = -dt^2 + a^2(t)\left(\frac{dr^2}{1-Kr} + r^2 d\Omega^2\right). \tag{10.38}$$

It is conventional to define a new radial coordinate, χ, as $r \equiv S_k(\chi)$, where

$$S_K(\chi) = \begin{cases} \sin\chi & (K = +1) \\ \chi & (K = 0) \\ \sinh\chi. & (K = -1) \end{cases} \tag{10.39}$$

Now, the line element can be written as

$$ds^2 = -dt^2 + a^2(t)(d\chi^2 + S_K^2(\chi)d\Omega^2). \tag{10.40}$$

It is also conventional to define a conformal time, η as $dt = ad\eta$. Sometimes the definition of conformal time is given in integral form as

$$\eta(t) - \eta_0 = \int_{t_0}^{t} \frac{dt}{a(t)} \Leftrightarrow t(\eta) - t_0 = \int_{\eta_0}^{\eta} a(\eta)d\eta. \tag{10.41}$$

In terms of this conformal time, the line element becomes

$$\begin{aligned} ds^2 &= a^2(\eta)\left(-d\eta^2 + \frac{dr^2}{1-Kr^2} + r^2 d\Omega^2\right) \\ &= a^2(\eta)\left(-d\eta^2 + d\chi^2 + S_K^2(\chi)d\Omega^2\right). \end{aligned} \tag{10.42}$$

A word on notation; $\dot{a} \equiv \frac{da}{dt}$ and $a' \equiv \frac{da}{d\eta}$.

2.2. Motion in FRW space-time. Now let's see how free particles behave as they move along the geodesics in FRW space-time. Let the geodesic, $x^\mu(\lambda)$, followed by the particles, to be defined by the affine parameter, λ. The tangent vectors to the geodesic is the 4-momentum, p^μ,

$$p^\mu \equiv \frac{dx^\mu}{d}. \tag{10.43}$$

Therefore for a particle with mass m, λ is related to proper time τ by $d\tau = md\lambda$. The geodesic equation, Eq 2.25, in terms of λ and the coefficients of affine connection (i.e the Christoffel symbols) is

$$\frac{dp^\mu}{d\lambda} + \Gamma^\mu_{\alpha\beta} p^\alpha p^\beta = 0. \tag{10.44}$$

The $\mu = 0$ component of the geodesic equation is

$$\frac{dp^0}{d\lambda} + HP^2 = 0, \tag{10.45}$$

where, $P \equiv (g_{ij} P^i P^j)^{\frac{1}{2}}$, is the magnitude of the spatial momentum. Taking the differential of the equation, $m^2 = -g_{\mu\nu} P^\mu P^\nu = (p^0)^2 - P^2$ yields $p^0 dp^0 = PdP$. This can be used to re-write Eq 10.45 as

$$P\left(\frac{dP}{dt}\right) + HP^2 = 0. \tag{10.46}$$

Which can be simply integrated to get

$$P \propto \frac{1}{a}. \tag{10.47}$$

This result applies to both massive and mass-less particles. First let's consider massive particles. From Eq 10.47, if the particles are initially at rest, relative to the spatial coordinates, then they will remain at rest. These are called *co-moving observes.* It is quite common to actually define a coordinate system in which particles remain stationary with respect to the coordinate system, such coordinate systems are called co-moving coordinates. Even if a massive particle initially has a large, non-vanishing 3-velocity relative to the co-moving coordinates (called its peculiar velocity). Eq 10.47, shows that the cosmological expansion causes the particle to gradually come to rest relative to them; an affect known as *Hubble drag.* Although the "coordinate distance" or "co-moving distance" between any two co-moving observes remains constant with time; the "physical distance" between grows $\propto a(t)$; this is what is meant by the expansion of the universe.

Coming to mass-less particles, even though a mass-less particle does not come to rest (it always move at the speed of light), it feels a kind of Hubble drag through the relation; $P = p^0 = \hbar\omega$; that is the energy and frequency decrease (from the de Broglie relation) and its wavelength stretches as the universe expands. This phenomena is known as cosmological redshift. This effect had been measured by Hubble before the full understanding of general relativity came about to describe it.

Now let's analyse the horizon structure of these space-times. To get the basic idea, we can keep things simple and consider a spatially flat FRW metric expanding as $a(t) = a_0 \left(\frac{t}{t_0}\right)^p$ $(0 < t < \infty)$. If we switch to conformal time η, this becomes

$$a(t) = a_0 \left(\frac{\eta}{\eta_0}\right)^{\frac{p}{1-p}} \tag{10.48}$$

where $0 < \eta < \infty$ if $p < 1$ (i.e if $\ddot{a} < 0$, deceleration) and $-\infty < \eta < 0$ if $p > 1$ (i.e if $\ddot{a} > 0$, acceleration). First suppose, that at time, t_* and conformal time $\eta_* = \eta(t_*)$ an observer emits a photon from $\chi = 0$. Let's follow this photon's trajectory forward in time. Firstly, a photon will have a null trajectory, as $ds^2 = 0$, and hence $d\eta = d\chi$, so at a later time $t > t_*(\eta > \eta_*)$ it will have reached a co-moving radius $\chi(t) = \eta(t) - \eta_*$. If we trace the photon all the way to $t = \infty$, it will reach a co-moving radius $\chi_\infty = \chi(t = \infty)$. Note that χ_∞ is infinite for $p < 1$, and finite for $p > 1$; we say that when $p > 1$, the observer is surrounded by a *future horizon* or event horizon- a surface beyond which the observer cannot send signals (This is analogous to the role that a black hole horizon plays for all observers inside the black hole).

Next suppose that at time t_* and conformal time $\eta_* = \eta(t_*)$, an observer receives a photon at $\chi = 0$. Let's follow the trajectory back in time. Again, it is traveling on an null trajectory, therefore $ds^2 = 0$ and hence $d\eta = -d\chi$, so at an earlier time $t < t_*(\eta < \eta_*)$ it was at a co-moving radius $\chi(t) = \eta_* - \eta(t)$. If we trace the photon all the way back to $t = 0$, it came from a co-moving radius $\chi_0 = \chi(t = 0)$. Note that χ_0 is finite for $p < 1$, and infinite for $p > 1$: we say that when $p < 1$, the observer is surrounded by a *past horizon* or a particle horizon, i.e the observer cannot receive signals from beyond the surface (this is analogous to the role that the black hole plays for all observers outside the black hole).

2.3. Motion of FRW space-times. The expansion history of a FRW space-time is determined by the amount and type of matter it consists of. To see this, first look at the Friedmann equation

$$H^2 = \frac{8\pi G}{3}\rho - \frac{K}{a^2} \tag{10.49}$$

If $\rho(t)$ is {equal to, greater then, less then} the instantaneous critical density, ρ_c

$$\rho_c \equiv \frac{3H^2(t)}{8\pi G}, \tag{10.50}$$

at some time, then it will be {equal to, greater then, less then} the critical density at all times; such a space-time has $\{K = 0, K = +1, K = -1\}$ and is called a {flat, closed, open} universe.

To get a feel for how these types of universes behave, imagine a universe that contains only a single matter component with a constant ratio, $\omega = \frac{p}{\rho}$. This ω is often known as the equation of state parameter. The continuity equation, Eq 5.41, integrates to give

$$\rho = \rho_0 \left(\frac{a}{a_0}\right)^{-3(1+\omega)}. \tag{10.51}$$

Putting this into the Friedmann equation Eq 10.49:

$$H^2 = \frac{8\pi G\rho_0}{3}\left(\frac{a}{a_0}\right)^{-3(1+\omega)} - \frac{K}{a^2}. \tag{10.52}$$

Let's first consider what the different values of K mean. First suppose the universe is "flat" (K=0), then $H^2 \geq 0$ implies $\rho \geq 0$; we can integrate the Friedmann equation to find the evolution of the scale factor;

$$\begin{aligned} a(t) &= a_0\left(\frac{t}{t_0}\right)^{\frac{2}{3(1+\omega)}} \qquad (K = 0, \omega \neq -1) \\ a(t) &= a_0 \exp\left(H_0(t - t_0)\right) \qquad (K = 0, \omega = -1), \end{aligned} \tag{10.53}$$

equivalent expression in terms of conformal times are;

$$\begin{aligned} a(\eta) &= a_0\left(\frac{\eta}{\eta_0}\right)^{\frac{2}{1+3\omega}} \qquad \left(K = 0, \omega \neq -\frac{1}{3}\right) \\ a(\eta) &= a_0 \exp\left(a_0 H_0(\eta - \eta_0)\right) \qquad \left(K = 0, \omega = -\frac{1}{3}\right). \end{aligned} \tag{10.54}$$

Such a universe either expands monotonically from $a = 0$ to $a = \infty$, or contracts monotonically from $a = \infty$ to $a = 0$. The age of the universe, t, is related to the instantaneous expansion rate $H(t)$ by $t = \frac{2}{3(1+\omega)H(t)}$.

Next suppose the universe is "closed", $(K = +1)$, then $H^2 \geq 0$ implies $\rho > 0$ and, even more strongly $\frac{8\pi G\rho}{3}\rho_0 \left(\frac{a}{a_0}\right)^{-3(1+\omega)} \geq \frac{1}{a^2}$. For $\omega > -\frac{1}{3}$, this means the universe expands from a big bang, (a=0), up to maximum size $a = a_*$, and then re-contracts to a big crunch (a=0); and for $\omega > -\frac{1}{3}$, the universe contracts from $a = \infty$ to a minimum size $a = a_*$ and then re-expands to $a = \infty$; in each case the extremal size a_* is given by

$$a_* = a_0 \left(\frac{8\pi G}{3}\rho_0 a_0^2\right)^{\frac{1}{1+3\omega}}. \tag{10.55}$$

Finally, suppose the universe is "open" (K=-1), then $H^2 \geq 0$ does not imply $\rho > 0$; instead it implies the weaker condition:

$$\frac{8\pi G}{3}\rho_0 \left(\frac{a}{a_0}\right)^{-3(1+\omega)} \geq -\frac{1}{a^2}. \tag{10.56}$$

If $\rho > 0$, this condition is never saturated, the universe expands monotonically from $a = 0$ to $a = \infty$, or contracts monotonically from $a = \infty$ to $a = 0$, much like the $K = 0$ case. The only difference is that the evolution is divided into an "early phase" and a "later phase" which have different characters, because the RHS of the Friedmann equation is dominated by the $\frac{8\pi G\rho}{3}$ terms in one phase, and the $-\frac{K}{a^2}$ term in the other. The final possibility is $\rho < 0$; then for $\omega > -\frac{1}{3}$, the universe contracts from $a = \infty$ to a minimum size $a = a_*$, and then re-expands to $a = \infty$ and for $\omega < -\frac{1}{3}$ the universe expands from a big bang ($a = 0$) to a maximum size $a = a_*^m$ and then re-contracts to a big crunch ($a = 0$); the extremal size a_* is given by Eq 10.55, with $\rho_0 \to -\rho_0$.

In fact, the general solution for ω and arbitrary $K = 0, +1, -1$, may be found as follows; move the $-\frac{K}{a^2}$ to the left side of the Friedmann equation; divide both sides by $\frac{8\pi G\rho}{3}$, so that the RHS equals 1; now note that if we define $A\tilde{a} = \left(\frac{a}{a_0}\right)^{\frac{(1+3\omega)}{3}}$, where $A = \sqrt{8\pi G\rho a_0^2/3}$ is a constant, the equation becomes

$$1 = \left(\frac{1+3\omega}{2}\right)^2 (\tilde{a}')^2 + K\tilde{a}^2. \tag{10.57}$$

The solution to this equation is

$$\tilde{a}(\eta) = S_k\left(\frac{(1+3\omega)\eta}{2}\right), \tag{10.58}$$

and hence[1]

$$a(\eta) = a_0 A^\alpha S_K^\alpha\left(\frac{\eta}{\alpha}\right), \tag{10.59}$$

where S_K was defined before and $\alpha = \frac{2}{1+3\omega}$. It used to be thought our universe contained two types of matter; a gas a of relativistic particles with $\omega = \frac{1}{3}$ ("radiation") and a gas of non-relativistic particles with $\omega = 0$ ("matter")

$$\rho = \rho_0^r \left(\frac{a}{a_0}\right)^{-4} + \rho_0^m \left(\frac{a}{a_0}\right)^{-3}. \tag{10.60}$$

In such a model, the universe should have been initially radiation dominated with $\rho = \rho_0^r \left(\frac{a}{a_0}\right)^{-4}$, and then "matter dominates", with $\rho = \rho_0^m \left(\frac{a}{a_0}\right)^{-3}$. Then the model we have also predicts that universe should become "curvature dominated" at late times, i.e the $-\frac{K}{a^2}$ should eventually overwhelm the $\frac{8\pi G\rho}{3}$ term on the RHS of the Friedmann equation. The universe we

[1] This is a result I have been trying to find ever since coming across universe models in the FRW metric!

observe does not appear to have entered the curvature dominated phase yet, even though it is 13.7 billion years old. This is known as the *flatness* problem.

CHAPTER 11

Matter in the universe

1. Thermodynamics and statistical mechanics of the universe

The fundamental concept in statistical mechanics is phase space,

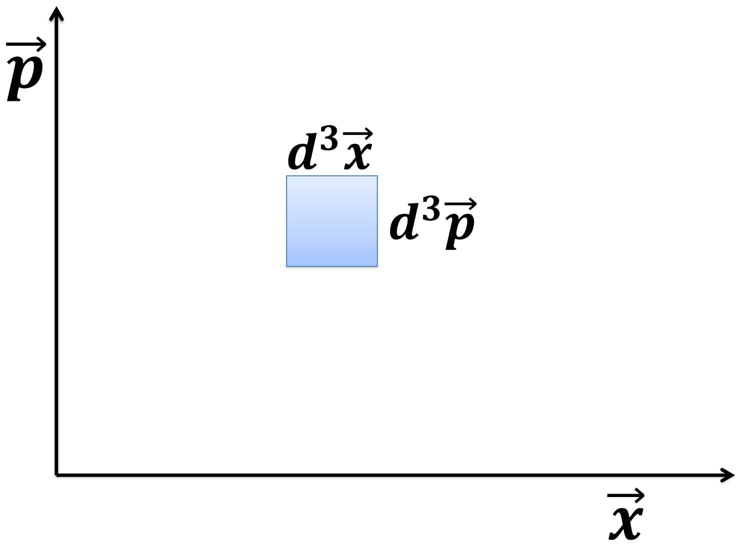


FIGURE 64. Schematic of phase space, with section in blue showing an infinitesimal volume V.

The most basic object is the distribution function in phase space, $f_i(t, \vec{x}, \vec{p})$. This function gives the expectation values of the number of particles of species, i, in the volume, V, above, by $f_i(t, \vec{x}, p)d^3\vec{x}d^3\vec{p}$. To obtain the number density in physical space of that species of particles, $n_i(t, \vec{x})$, one has to integrate over the momentum,

$$n_i(t, \vec{x}) = \int d^3\vec{p} f_i(t, \vec{x}, \vec{p}). \tag{11.1}$$

If we want to calculate the energy density in physical space, $\rho_i(\vec{x}, t)$, then one requires the energy of a point in phase space, which is also a distribution, $E(\vec{p}, m_i)$. Now the energy density in physical space is simply

$$\rho_i(t, \vec{x}) = \int d^3\vec{p} f_i(\vec{x}, \vec{p}, t) E(\vec{p}, m_i). \tag{11.2}$$

Finally the pressure, $P_i(\vec{x}, t)$ is given by

$$P_i(t, \vec{x}) = \int d^3\vec{p} f_i(t, \vec{x}, \vec{p}) \frac{\vec{p}^2}{2E} \tag{11.3}$$

These are the general expression that works for any distribution function. There are two special distribution functions that are often used in physics. If the particles being considered can be approximated by an ideal gas of bosons or fermions, then the distribution function for these particles are the Bose-Einstein and the Fermi-Dirac distribution respectively. If the particles of mass m_i are in thermal equilibrium at temperature T, and chemical potential μ_i, then the expected number of particles in any tiny volume $h^3 = (2\pi\hbar)^3$ in phase space is given by

$$f_i = \frac{1}{\exp\left(\frac{E-\mu_i}{k_B T}\right) \mp 1}, \tag{11.4}$$

where the minus sign is for bosons and the positive sign is for the fermions, and $E = \sqrt{\vec{p}^2 - m^2}$. The next step is to change the integration variables in Eq 11.1, 11.2 and 11.3 from $\vec{p}$ to E (we can do this because $E = \sqrt{\vec{p}^2 - m^2}$) and substitute in for the distribution function in Eq 11.4 to get

$$n_i = \frac{1}{2\pi^2}\int_{m_i}^{\infty} \frac{(E^2 - m_i^2)^{\frac{1}{2}} E dE}{\exp\left(\frac{E-\mu_i}{T_i}\right) \mp 1} \tag{11.5}$$

$$\rho_i = \frac{1}{2\pi^2}\int_{m_i}^{\infty} \frac{(E^2 - m_i^2)^{\frac{1}{2}} E^2 dE}{\exp\left(\frac{E-\mu_i}{T_i}\right) \mp 1} \tag{11.6}$$

$$P_i = \frac{1}{6\pi^2}\int_{m_i}^{\infty} \frac{(E^2 - m_i^2)^{\frac{3}{2}} dE}{\exp\left(\frac{E-\mu_i}{T_i}\right) \mp 1}. \tag{11.7}$$

As is done with most complex equations, to get an idea of what they mean, we take the equations to appropriate limiting cases to get an appropriate equation and analyse this. In this case the limiting cases that will be used are when the particles are highly relativistic or when the particles are non-relativistic. In the relativistic limits ($T >> m_i, \mu_i$, n_i) simplifies to

$$n_i = \frac{\zeta(3)}{\pi^2} T_i^3 \tag{11.8}$$

and for ρ_i,

$$\rho_i = \frac{\pi^2}{30} T_i^4. \tag{11.9}$$

This is for bosons, for fermions one has

$$n_i = \left(\frac{3}{4}\right) \frac{\zeta(3)}{\pi^2} T_i^3 \tag{11.10}$$

$$\rho_i = \left(\frac{7}{8}\right) \frac{\pi^2}{30} T_i^4. \tag{11.11}$$

The pressure is evaluated using the relation,

$$P_i = \omega_i \rho_i \tag{11.12}$$

and $\omega_i = \frac{1}{3}$ for both bosons and fermions (in the relativistic case). Going to the non-relativistic limit, $T << m_i$, with $\mu_i < m_i$ we have,

$$n_i \approx \left(\frac{m_i T}{2\pi}\right)^{\frac{3}{2}} \exp\left(\frac{\mu_i - m_i}{T}\right) \tag{11.13}$$

$$\rho_i = m_i n_i \tag{11.14}$$

$$P_i \approx T n_i. \tag{11.15}$$

Note that the equation for n_i is saying that in thermal equilibrium, as soon as $T < m_i$, the number density of the particles starts to be exponentially suppressed. The equation for ρ_i is intuitive as in the non-relativistic limit one expects the total energy of a system to be dominated by the mass energy (i.e the kinetic energy is negligible). Note that from $\omega = \frac{P_i}{\rho_i}$, $\omega \approx 0$ as $m_i >> T_i$.

Another point that needs pointing out is the fact that we have immediately assumed an ideal gas distribution function to compute n_i, ρ_i and P_i. These distribution functions describe a lot of the universe that we observe, however, they are limited by the fact that they only work in thermal equilibrium. If one tries to describe the very early universe, this distribution function no longer works, therefore one has to generalise the distribution function. The equation of motion followed by the distribution equation is called the *Boltzmann equation*. This equation needs to be solved for the most general solution of $f_i(t, \vec{p}, \vec{x})$.

2. Entropy

The universe has been expanding for a long time, so it is obviously not in thermal equilibrium. This raises the question; why are we considering an expression that only holds in thermodynamical equilibrium. The answer is that we can find the universe at a point in time where it is in thermal equilibrium and then use the entropy to find out its state at a later time, when it is not in thermal equilibrium. Under adiabatic conditions the universe would actually remain thermal equilibrium throughout its expansion. Adiabatic conditions are when the rate of expansion of the universe, $H = \frac{\dot{a}}{a}$, is much slower than the rate of collisions between particles.

In a fluid, one can think of entropy density, s, such that the total entropy, S, is given by $S = sV$, where V is the spatial volume. One can get S using the first law of thermodynamics,

$$dE = TdS - PdV \tag{11.16}$$

Substituting for $E = \rho V$ and the total entropy,

$$d(\rho V) = Td(sV) - PdV \tag{11.17}$$

This can be written as

$$(\rho + P - Ts)dV = \left(T\frac{dS}{dT} - \frac{d\rho}{dT} \right) VdT. \tag{11.18}$$

The coefficients of dV and dT must be 0, therefore

$$s = \frac{\rho + P}{T}. \tag{11.19}$$

For relativistic particles $P = \frac{1}{3}\rho$;

$$s = \frac{4\rho}{3T}. \tag{11.20}$$

Substituting for ρ from Eq 11.11 gives

$$s_f = \frac{2\pi^2}{45}T^3, \tag{11.21}$$

this is for fermions, for bosons it is,

$$s_b = \frac{7}{8}s_f. \tag{11.22}$$

Now if one is in adiabatic conditions, there is no heat transfer therefore $dS = 0$, i.e S is a constant. Since $V \propto a^3$, $s \propto a^{-3}$, so from Eq 11.21, $T \propto a^{-1}$, i.e temperature drops linearly with the expansion of the universe.

3. Big Bang Nucleosynthesis (BBN)

BBN was discovered by Gamow. When we look out into the universe we see an abundance of the elements present in the periodic table out in the universe. At large, the universe contains about 75% of hydrogen, 25% helium and the rest of the elements are simply negligible.

In the middle of the 20^{th} century, it was worked out that stars actually generate their energy from nuclear reactions, converting, predominantly, $4H \rightarrow He$ (the p-p chain to be more general). Fusion reactions continue until we get to Iron, Fe^{56}, as that is the most stable nuclei (has the maximum binding energy per nucleus). All of the elements heavier than Fe^{56}, are produced in non-fusion processes like supernova explosions. Gamow and other people realised that if the universe started off 100% hydrogen then it was not possible to generate 25% helium simply by stellar fusion reactions.

Gamow showed with a rough back of an envelope calculation that this observed ratio of hydrogen to helium could be explained by the production of helium at the big bang. Suppose we start the universe when it was at a temperature of 10 GeV. At this energy, the particles that are present are the usual protons, neutrons, electrons, photons, neutrino's. These are all still relativistic. Since $n_i \propto T^3$, all of these species has a number density that is roughly the same. When the universe cools down by a factor of 10, i.e $T \approx 1GeV$, the protons and neutrons start to become non-relativistic. The number density of a non-relativistic species is given in Eq 11.13, and so we see that in the non-relativistic case the number density of these non-relativistic particles drops exponentially (i.e the particles decay to lighter particles that are thermodynamically more favorable). Here in-lies the problem; by he fact that there still is matter in the universe, all of the matter did not annihilate with anti-matter. It turns out that for every 10^9 $\bar{p}$, there were $10^9 + 1, p$ (i.e one p of a billion was left over an average).

So now, the cosmological zoo of particles has $e, e^+, \gamma, \nu's$ all with roughly the same number density with a very small number of $p's$ left. It is an observed fact about the universe, that, η, which is defined as the ratio of baryons to photons is 10^{-9}. This slight access of matter over anti-matter is one of the biggest problems in physics and demands an explanation (other than the anthropic one!). This process is known as baryogensis and will be discussed in a later section.

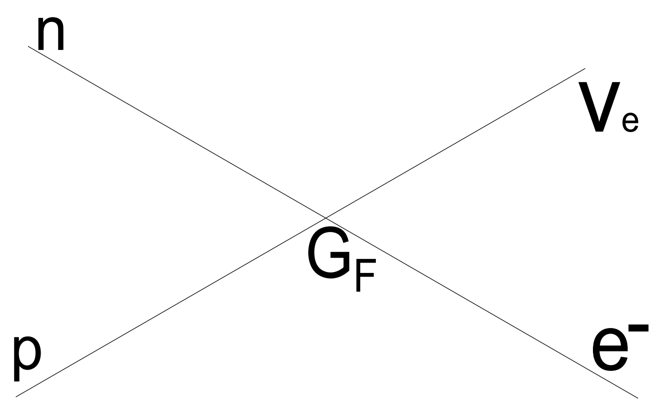


FIGURE 65. Feynman diagram of Eq 11.23 in Fermi theory.

Now let's continue to decrease the energy of the universe. The neutrons and protons have now annihilated and a very small number of them are left over. However, the energy scale is still much longer than the energy scale of the binding energy, MeV, and thus the neutrons and protons have not formed nuclei. The remaining protons and neutrons are in thermal equilibrium with the remaining particles (electrons, neutrino's etc) by the reaction

$$p^{+} + e^{-} \leftrightarrow n + \nu_e. \tag{11.23}$$

The idea now is to find the point at which the rate of the reaction in Eq 11.23 is slower than the rate of expansion, H, at that point, as at that point the equilibrium of this reaction is broken. At the time in which this reaction was first being thought about, Fermi's theory of the weak interactions was generally accepted, therefore the reaction in Eq 11.23 has a diagram shown in figure 65. In the standard model however, the diagram is given below.

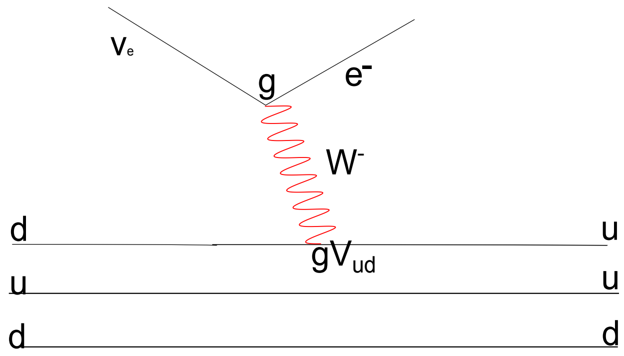


FIGURE 66. Feynman diagram in of Eq 11.23 in the standard model.

Ignoring CKM elements the rate is $\propto g^2 \frac{1}{\rho^2 - M_w^2}$, however since $M_w >> p$, we can approximate this by $\frac{g^2}{M_w^2}$, which is equivalent to G_F according to Fermi theory,

$$G_F \approx \frac{g^2}{M_w^2} \approx 10^{-5} GeV^2. \tag{11.24}$$

The rate, Γ, is approximately $G_F^2 T^5$ (by dimensional analysis). Comparing this to the expansion rate, H. H is given by the Friedmann equation,

$$H^2 = \frac{8\pi G\rho}{3}. \tag{11.25}$$

The energy density, ρ, of the universe can be re-written in terms of the temperature, $\rho \propto T^4$, from Eq 11.9

$$H^2 \approx \frac{8\pi G}{3} g_* T^4 \tag{11.26}$$

where, g_* is the effective number of degrees of freedom. Therefore H is of the order of $\frac{T^2}{M_p}$ (as $G = \frac{1}{M_p}$). Equating, H to Γ solve for T, we get $T = T_* \approx 1MeV$ for the temperature at which these reactions can no longer take place. After this temperature, thermal equilibrium is no longer maintained. The number density of protons, n_p, at this temperature was

$$n_p = \left(\frac{m_P T_*}{2\pi}\right)^{\frac{3}{2}} \exp\left(-\frac{M_p}{T_*}\right). \tag{11.27}$$

Similarly, the number density of the neutrons, n_N, was

$$n_N = \left(\frac{M_N T_*}{2\pi}\right)^{\frac{3}{2}} \exp\left(-\frac{M_N}{T_*}\right) \tag{11.28}$$

As a first approximation, after thermodynamics ceases to be effective, the protons and neutrons do not do anything. The proton is stable, it's lifetime is much longer than the present age of the universe. The lifetime of a neutron is roughly 15 minutes, however all of the processes being discussed happening in the very early universe, therefore neutrons can also be considered stable. The ratio of the number of neutrons to the number of protons at T_* is

$$\frac{n_N}{n_p} \approx \exp\left(\frac{M_N - M_p}{T^*}\right). \tag{11.29}$$

T_* is sometimes called the freeze out temperature, $M_N - M_p \approx 1.3MeV$, and $T \approx 0.8MeV$ (comes from the full calculation). This gives a value for this ratio as about $\frac{1}{6}$. As soon as T drops below the binding energy of helium, it is favorable for neutron/proton to be in a helium nuclei and thus helium nuclei are formed. However in the time it takes for T to drop below the binding energy some of the neutrons will decay, therefore the rate actually becomes smaller, $\approx \frac{1}{7}$. The mass density in the helium nuclei will be twice the mass density in neutrons, and the ratio to the mass density is,

$$\frac{2M_{N,p}n_n}{M_{N,p}(M_p + M_N)} = \frac{2}{\left(\frac{n_p}{n_n} + 1\right)} = \frac{1}{4}. \tag{11.30}$$

Which shows that 25% of the mass of the universe is helium, which agrees with observation. One possible discrepancy between BBN predictions and observations has emerged; the predicted abundances of lithium appears to be about a factor of 3 higher then the abundances observed in outer layers of the lower metalicity stars.

There is analogous calculation to check when the CMB was formed, i.e when did the photons stop interacting with matter. Electrons and protons can bind together to form hydrogen atoms (or helium atoms). When the T is greater than the binding energy of these atoms (not nuclei!), the electrons do not orbit the nucleus due to their large kinetic energy. A similar calculation to the one just done (for protons and neutrons to form helium nuclei), we can find a T_{CMB} which is the temperature at which electrons are bound up with nuclei to form atoms. Prior to this transition, photons could scatter of a lot of particles and hence had a very small mean free path (relative to t he size of the universe). As soon as atoms formed, the photons stopped colliding with matter and there mean free path become almost infinite. Calculating, T_{CMB} will give an idea of the wavelength of the photons at the time of T_{CMB} temperature and these photons are predicted to still be around today at a much higher wavelength (microwaves). The temperature of these photons is measured today as $T \approx 3K$. T_{CMB} was $\approx 3.3 \times 10^3 K$ and since $T \propto a^{-1}$, the universe was smaller by a factor of about 1000 at that time. The moment at which the electrons combined with the nuclei is known as *recombination.*

4. Dark matter

4.1. Why dark matter? Dark matter was first discovered by Fritz Zwicky in the 1930's. He observed that individual galaxies in a cluster of galaxies were moving much faster than could be accounted for by the gravitational energy present in that cluster. Which lead Zwicky to believe that there was more mass then was visible in the cluster, to account for this motion. Zwicky's work was largely ignored until the 1970's when Vera Rubin analysed the galaxy rotation curves (the velocity at which stars in a galaxy are orbiting the center). This observation has been backed up by many different types of experiments that all appear to point to the same abundance of dark matter, which is why dark matter is strongly expected to exist in the universe.

In BBN, we saw that when $T \approx 1MeV$, the protons and neutrons combined to form the light elements in the periodic table. In particular we showed that the matter in the universe is distributed into about 75% hydrogen and 25% helium (by mass). This prediction is quite robust, more specifically, it does not depend on η (the baryon to photon ratio). On the other hand the percentage of Deuterium does depend of η. As η is varied at the time of BBN, the predicted amount of deuterium varies a lot, and so by observing the amount of deuterium in our universe one can put a limit upon the value of η, which is $\approx 10^{-9}$. This value of η agrees with value of predicted by the number of protons and neutrons observed in galaxies and the universe in general. If one assumes that the matter needed to account for the light curves of the galaxies comes from protons and neutrons (i.e baryonic matter), then the predicted value of η is completely different, therefore the matter required to account for this gravitational affect must not be baryonic matter.

4.2. Baryonic acoustic oscillations. Another measurement independent of BBN of this quantity η, is in the CMB. At the time when the CMB photons were released, (recombination) the universe contained a plasma and also (we suppose) dark matter, that did not interact with anything. Suppose that the density of matter of the universe was varying slightly in position.

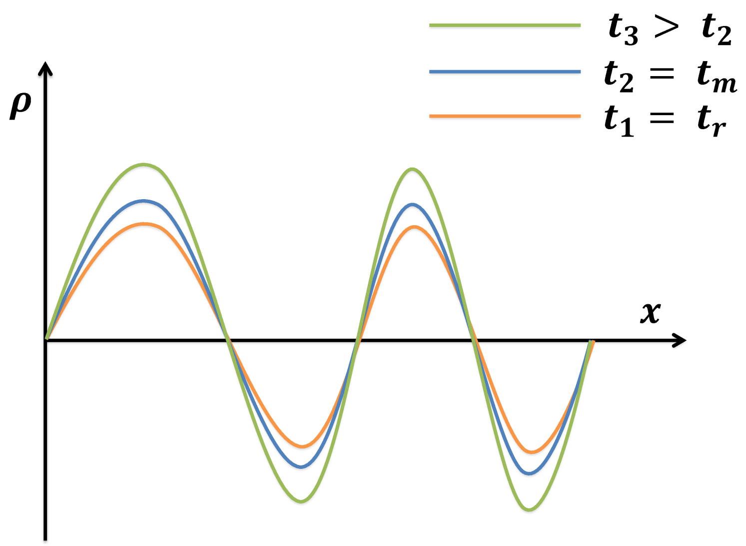


FIGURE 67. Schematic showing the baryonic acoustic oscillations. t_m represents the time at which matter is dominant, t_r is the time at which radiation is dominant.

When matter started to dominate over radiation, the energy density at the peaks increased (as matter started to attract more matter from less dense regions). However the plasma also contains photons, which do not interact with gravity and are bound by the their collisions with the electrons and protons. Therefore there is a constant tug of war between the gravitational forces and the pressure coming from the radiation. As photons move around in the plasma , they oscillate and these oscillations can become "quantised" into phonon's that produce sound waves. So the spectrum of the CMB shows the peaks corresponding to wavelengths that happen to be at the peak of the amplitude of their oscillations when recombination happened. The first peak is the peak corresponding to the longest wavelength that from the beginning of the matter dominated era to the time of recombination only had time to complete half an oscillations (i.e get to its peak), the second peak corresponds to wavelength that had time to complete one full oscillation and so on. Dark matter played a crucial role in this spectrum of the CMB, by providing deep potential wells that set the size of these oscillations.

Everything described so far relies on the gravitational interaction of dark matter. Therefore the obvious thing to ask is, what if it is gravity that is not behaving as we expect. If gravity is described by general relativity (as we think it is), then one is forced to postulate the existence of this dark matter[1]. If dark matter was matter from the baryonic sector, it would have had to

[1]There is a new theory proposed by Mukhanov, called Mimetic dark matter, http://arxiv.org/pdf/1308.5410v1, which modifies Einstein's gravity to account for the effect of dark matter.

collapse into a black hole (due to the restriction imposed by η). These black holes would have had to have formed prior to BBN. These types of suggestions generically fall under the name of massively compact halo objects, MACHOs, which contain any matter that is highly dense but just not visible, for example brown dwarf stars.

The other possibility is that they are some new elementary particles that interact via gravity. These types of particles are called weakly interaction massive particles, WIMPs. MACHOs are heavily constrained by observational evidence, however are not completely ruled out. If MACHOs were really around, one should be able to observe them using gravitational micro-lensing. Micro-lensing works as follows; if one observes the intensity of light coming from a star, it would be expected to be roughly constant. Now if a MACHO, say a black hole, goes in front of it, then the light from the edges of the star will be bent towards the center and if the observer is in the same line, the intensity of light will appear to *increase*. This increase in light can be predicted for different MACHOs and observations have basically ruled out MACHOs in the range, $10^{-8}M_{sun} < M_{MACHO} < 2M_{sun}$.

If there are massive new elementary particles in the universe, we can break them down into two categories.

(1) Thermal relics: Particles that were in thermal equilibrium in the early universe and once they reach their freeze out temperature, they would have stopped interacting and traveled through the universe ever since.
(2) Non-thermal relics: Particles that were never in thermal equilibrium and had their abundances set by some other process. The classic example for this particle is the axion. There is a problem in the standard model of particle physics called the strong CP problem, which can be solved by the introduction of this particle which also happens to be a good candidate for dark matter, that is a non-thermal relic.

Thermal relics can be subdivided further into *hot-thermal relics* or *cold-thermal relics*. If the particle was relativistic after its freeze out temperature, it is said to be a hot thermal relic and the opposite definition for a cold thermal relic.

As an example, lets take a close look at hot thermal relics. Consider a hot thermal relic particle, χ. After freeze out, one can assume the number density n_χ decreases with increasing volume. Similarly the entropy density, s, decreases with volume if we assume adiabatic conditions (i.e total entropy is the same). Therefore, the ratio $\frac{n_\chi}{s}$ is constant. Let's denote the number density today by $n_{\chi,0}$ and at freeze out it is $n_{\chi,f}$. Similarly the entropy density today is s_0 and at freeze out it is s_f. Now we can form the equation

$$\frac{n_{\chi,0}}{s_0} = \frac{n_{\chi f}}{s_f}. \tag{11.31}$$

All the quantities except $n_{\chi,0}$ are knoww, therefore we can solve for $n_{\chi,0}$

$$n_{\chi,0} = \left(\frac{s_0}{s_f}\right) n_{\chi,f}, \tag{11.32}$$

where,

$$\begin{aligned} s_0 &= g_{s,0}\left(\frac{2\pi^2}{45}\right)T_0^3 \\ s_f &= g_{s,f}\left(\frac{2\pi^2}{45}\right)T_f^3 \\ n_{\chi,f} &= g_\chi\left(\frac{\zeta(3)}{\pi^2}\right)T_f^3. \end{aligned} \tag{11.33}$$

Substituting these into Eq 11.32,

$$n_{\chi,0} = \frac{g_{s,0}}{g_{s,f}} \left(\frac{T_0}{T_f}\right)^3 g_\chi \left(\frac{\zeta(3)}{\pi^2}\right) T_f^3$$
$$= \frac{g_{s,0}}{g_{s,f}} T_0^3 g_\chi \frac{\zeta(3)}{\pi^2}. \tag{11.34}$$

It is safe to assume the χ's are non-relativistic today, the present energy density due to the χ's can therefore be approximated by $\rho_{\chi,0} = m_\chi n_{\chi,0}$,

$$\rho_{\chi,0} = m_\chi \frac{g_{s,0}}{g_{s,f}} T_0^3 g_\chi \frac{\zeta(3)}{\pi^2}. \tag{11.35}$$

Comparing this to the critical density, $\rho_{c,0} = \frac{3H_0^3 m_{pl}^2}{8\pi}$ gives,

$$\frac{\rho_{\chi,0}}{\rho_{c,0}} = \frac{8\zeta(3)}{3\pi} \frac{g_\chi g_{s,0}}{g_{s,f}} \frac{m_\chi T_0^3}{H_0^2 m_{pl}^2} \tag{11.36}$$

Some of the numerical values are $g_{s,0} = 3.91, H_0 = (2.13 \times 10^{-42}) h_{100} GeV, m_{pl} \approx 1.22 \times 10^{19} GeV$ and $T_0 = 2.35 \times 10^{-13} GeV$. This gives

$$\frac{\rho_{\chi,0}}{\rho_{c,0}} = \frac{g_\chi}{g_{s,f}} \frac{8 \times 10^{-2}}{h_{100}^2} \frac{m_\chi}{eV}. \tag{11.37}$$

This result is for χ particles that are bosons. If one has fermions, then there is an extra multiplicative factor of $\frac{7}{8}$ on the RHS of Eq 11.37. From observations we know that $\frac{\rho_{\chi,0}}{\rho_{c,0}} \leq \frac{1}{4}$, which provides a constraint on m_χ via Eq 11.37. Note that the neutrino (which decouples around $T \approx 1MeV$) is an example of a hot thermal relic, therefore we get a limit on the mass of neutrino's via this calculation aswell.

After the χ particle decouples they are still relativistic, so they stream across the universe like mass-less particles until the Hubble drag finally makes them non-relativistic, and brings them to rest relative to co-moving observers. Before coming to rest, they travel a co-moving distance which is roughly equal to the "co-moving Hubble radius", $\frac{1}{aH}$, at the moment they come to rest. This would wipe out the formation of all cosmic structure formation on length scales less than or equal to this co-moving scale, in conflict with observations. Thus hot thermal relics are observationally ruled out as dark matter. This procedure of equating the ratio of number density to entropy density at different times, is quite common and can be applied to cold thermal relics aswell.

5. Vacuum energy

The action for general relativity is

$$S = \int d^4x \sqrt{-g} \left(\frac{R - 2\Lambda}{16\pi G} + L_{matter}\right). \tag{11.38}$$

Where L_{matter} is any matter Lagrangian. This action is varied w.r.t $g_{\mu\nu}$ to obtain Einstein's equations

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{11.39}$$

$G_{\mu\nu}$ comes from the variation of R in the action, $\Lambda g_{\mu\nu}$ comes from the variation of the Λ term in the Lagrangian and $8\pi G T_{\mu\nu}$ comes from varying the L_{matter} w.r.t $g_{\mu\nu}$. One can also move the Λ term on to the RHS, where it is considered a component of the energy density (as supposed to the curvature),

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} - \Lambda g_{\mu\nu}. \tag{11.40}$$

The Λ acts as a source of energy now, so we are at liberty to include the Λ term inside $T_{\mu\nu}$ to define a new stress energy tensor, $T'_{\mu\nu}$, to get

$$G_{\mu\nu} = 8\pi G T'_{\mu\nu}. \tag{11.41}$$

If L_{matter} is that of a scalar field, then it takes the form

$$L_{matter} = -\frac{1}{2}(\partial_\mu \phi)^2 - V(\phi). \tag{11.42}$$

The potential, $V(\phi)$, can be chosen arbitrary according to the field theory being analysed. Suppose $V(\phi)$ takes the form in figure 68,

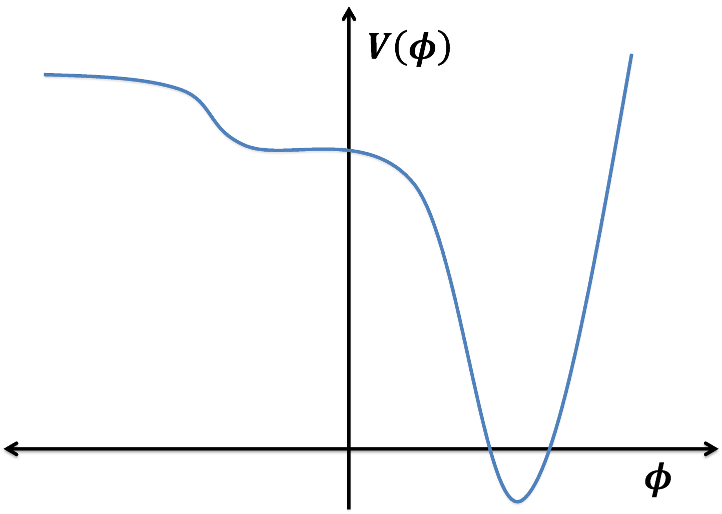


FIGURE 68. The blue line shows the shape of the potential, $V(\phi)$ in Eq 11.42.

the vacuum state is when the scalar field is constant in space and time (therefore $\partial_\mu \phi = 0$) and it is sitting in the minimum potential value aswell. The minimum value of this potential is called vacuum energy and it basically plays the role of shifting the potential minimum up or down by a given value. There is observational evidence for the existence of a positive cosmological constant. The Friedman equation with the cosmological constant is

$$H^2 = \frac{8\pi G\rho}{3} - \frac{K}{a^2} + \frac{\Lambda}{3}. \tag{11.43}$$

5.1. Old and new cosmological problem. It appears that we happen to live at the time at which the Λ is starting to dominate over the matter, causing the universe to accelerate. If we lived in an era in which Λ did not dominate, then we could have never observed the acceleration of the universe's expansion and would have remained oblivious to the fact that the Λ even existed! This may same too much of a coincidence to some[2] and is known as the new cosmological constant problem.

The old cosmological constant problem is that the zero point energy of particles in space appears to be much higher than the vacuum energy observed in the universe.

[2]Including me.

6. Baryogenesis

Today we observe the ratio of baryons to photons, η, with a value of 10^{-9}, i.e for any 10^9 anti-baryons, there were roughly $10^9 + 1$ baryons, so when baryons and anti-baryons eventually annihilated, a small number of baryons (and no anti-baryons) were left over (note we also assume that the number of photons and baryons was of the same order). The question is what causes this imbalance, i.e why were the initial conditions such that there were more baryons than anti-baryons.

To do this day, we have never observed a process which violates the conservation of baryon number. However it is now believed that the baryon number conservation is actually violated in the standard model of particle physics itself and in any grand unified theory (GUT) into which it may be embedded. Hence the baryons could have been decayed by standard model processes. Furthermore, an epoch of inflation in the early universe would have exponentially suppressed any pre-existing baryon number density, requiring that they really are generated after the inflationary epoch.

6.1. Sakharov's conditions. In the 1960's, Andrei Sakharov produced the modern picture taken up by Baryogenesis; the universe started out with an equal abundance of baryons and anti-baryons and the slight excess of baryons was generated was generated during its subsequent expansion. In other words, he suggested that the asymmetry could be explained by the laws of physics, rather than the initial conditions. He came up with three conditions that would be required for baryogenesis;

Condition 1 Violation of baryon number conserving processes.
Condition 2 Violation of C and CP symmetry.
Condition 3 Violation of thermal equilibrium (in thermal equilibrium, any reaction which creates baryon number will be precisely balanced by the inverse reaction which destroys baryon number).

Let's look at **Condition 2** more closely. Consider the distribution functions $f_j(\vec{x}, \vec{p}, t)$ for the j^{th} species in a FRW universe: homogeneity and isotropy imply that f_j is independent of $\vec{x}$ and of the direction of $\vec{p}$, but not its magnitude,

$$f_j(\vec{x}, \vec{p}, t) = f_j(p, t). \tag{11.44}$$

Sakharov's proposal was that the particle and anti-particle distributions were equal at some initial time, t_i

$$f_j(p, t_i) = \bar{f}_j(p, t_i) \tag{11.45}$$

and unequal and some final t_f,

$$f_j(p, t_f) \neq \bar{f}_j(p, t_f). \tag{11.46}$$

A parity transformation, $\mathbb{P}$, is defined as,

$$\vec{x}, \vec{p} \overset{\mathbb{P}}{\Longrightarrow} -\vec{x}, -\vec{p}. \tag{11.47}$$

A charge conjugation transformation, C, swaps particles and anti-particles;

$$f_i \overset{C}{\Longrightarrow} \bar{f}_i. \tag{11.48}$$

Therefore P leaves $f_j(p, t)$ unchanged, whereas C and CP both swap $f_j(p, t) \Leftrightarrow f_j(p, t)$. Thus the initial state (at t_i) is invariant under P, C and CP, but the final state (at t_f) is only invariant under P. Thus, the evolution from t_i and t_f must violate C and CP (but not necessarily P).

6.2. Example of baryogenesis. Consider a particle X with two decay channels;

$$X \to f_1, \tag{11.49}$$

where f_1 is a final state with baryon number B_1, lepton number L_1, and branching ratio r and the other particle decay is

$$X \to f_c \tag{11.50}$$

where f_2 has baryon number B_2, lepton number L_2 and branching ratio $1 - r$. The antiparticle of X, $\bar{X}$ also has two decay channels; $\bar{X} \to \bar{f}_1$, (with baryon number $-B_1$, lepton number $-L_1$ and a branching ratio $\bar{r}$) and $\bar{X} \to \bar{f}_2$ (with baryon number $-B_2$, lepton number $-L_2$ and branching ratio $1 - \bar{r}$).

CPT symmetry requires that X and $\bar{X}$ have the same mass m_X and total decay rate Γ_x; but if C and CP are violated, they can have different branching ratios (i.e $r \neq \bar{r}$). Now we imaging that X's freeze out, and later decay; the ratio of baryon density to entropy density, n_B/s, after the decay will be related to the ratio, $\frac{n_X}{s}$ (the ratio of the number density of X particles to the ratio of the energy density) before the decay by,

$$\frac{n_B}{s} = \frac{n_X}{s}(sB_1 + (1-r)B_2 - \bar{r}B_2 - \bar{r}B_1 - (1-\bar{r})B_2) = \frac{n_X}{s}(r - \bar{r})(B_1 - B_2). \tag{11.51}$$

Note how the Sakharov conditions show up in this calculation: in order for $\frac{n_B}{s}$ to be non-zero, we need $r \neq \bar{r}$ (requiring C and CP violation) and $\beta_1 \neq \beta_2$ (requiring baryon number violation).

We also need violation of thermal equilibrium: before the X particles decay, they freeze out and T drops below m_X: otherwise the inverse decay processes (which create X's and $\bar{X}$'s from the thermal bath, and which were neglected in the calculation above) cancel the baryon number we have just calculated. We want $\frac{n_B}{s}$ to match the observed values (10^{-9}), and the above formula tells us how to check weather it holds in a given model of particle physics: given the Lagrangian for the model, we look for any particles with several decay channels with different baryon (or lepton) numbers; we estimate the freeze out abundances of thus particle ($\frac{n_X}{s}$) using techniques similar to those employed in the dark matter calculation above. We calculate the branching ratios $(r, \bar{r}, ..)$ for the relevant decays, and finally check that the particle decays after it has been frozen out and the temperature has dropped below the mass; if so, and the above formulation leads to a predicted value, which is similar to the observed one, then the theory is good. The scenario described above was inspired by models of GUT's in the 1970's.

CHAPTER 12

Inflationary cosmology

There are two important length scales in the universe, at any given time. One is the Hubble radius, given by $\frac{c}{H}$, the other is the scale factor $a(t)$ of the universe. The ratio of these two scales, $\frac{c}{Ha}$, which is known as the *co-moving Hubble radius*. When the universe is decelerating, the co-moving Hubble radius is increasing. When the universe is accelerating, the co-moving Hubble radius is getting smaller with time. If the universe was radiation dominated all the way to the big bang, then the Hubble co-moving radius would have started at zero size and gotten bigger until the end of the radiation dominated era.

Inflation postulates that at the beginning of the universe, before the radiation dominated era, the universe underwent an exponential expansion, in which the Hubble co-moving radius decreased to a very small size, and then inflation stopped and the universe had been radiation dominated and started decelerating and the Hubble co-moving radius has been increasing every since. If the co-moving Hubble sphere before inflation, is longer than the co-moving Hubble sphere today, then there are three problems (that were bothering cosmologists at the time), that will be solved. We describe each of them below.

1. The three problem; flatness, horizon and monopole problems

1.1. Flatness problem. There are two key equations that describe the FRW universe. One is the Friedmann equation

$$H^2 = \frac{8\pi G\rho}{3} - \frac{K}{a^2} \tag{12.1}$$

and the other is the continuity equation

$$\dot{\rho} = -3H(P+\rho). \tag{12.2}$$

Eq 12.1 can be written as

$$1 = \frac{8\pi G\rho}{3H^2} - \frac{K}{a^2H^2}. \tag{12.3}$$

We then define the critical density, $\rho_c \equiv \frac{3H^2}{8\pi G}$ and Eq 12.3 becomes

$$1 = \frac{\rho}{\rho_c} - \frac{K}{(aH)^2}. \tag{12.4}$$

Differentiate Eq 12.1 w.r.t time to get

$$2H\dot{H} = \frac{8\pi G\dot{\rho}}{3} + \frac{2KH}{a^2} \tag{12.5}$$

substitute Eq 12.2 into Eq 12.5, for $\dot{\rho}$,

$$2H\dot{H} = \frac{8\pi G}{3}(-3H)(\rho+P) + \frac{2KH}{a^2} \tag{12.6}$$

cancel a factor of H from both sides,

$$2\dot{H} = -\frac{8\pi G}{3}(\rho+P) + \frac{2K}{a^2} \tag{12.7}$$

but $\dot{H}$ is simply,

$$\dot{H} = \frac{dH}{dt} = \frac{d}{dt}\left(\frac{\dot{a}}{a}\right) = \frac{\ddot{a}}{a} - \frac{\dot{a}}{a^2}. \tag{12.8}$$

Substitute Eq 12.8 into Eq 12.7

$$2\left(\frac{\ddot{a}}{a} - \frac{\dot{a}}{a^2}\right) = -\frac{8\pi G}{3}(\rho + P) + \frac{2K}{a^2} \tag{12.9}$$

this must be equal to Eq 12.6

$$\frac{2\ddot{a}}{a} - 2\left(\frac{8\pi G\rho}{3} - \frac{K}{a^2}\right) = -8\pi G(\rho + P) + \frac{2K}{a^2} \tag{12.10}$$

which can be re-written to given the *acceleration equation*

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + 3P\right). \tag{12.11}$$

For ordinary matter ρ is positive and P is negligible, therefore $\ddot{a}$ is negative, as expected (attractive nature of gravity). To get acceleration, one needs a sufficiently negative pressure, which is what the Λ provides. In fact, from Eq 12.11 we can see that for acceleration $\omega(\equiv \frac{P}{\rho})$ must be less than $-\frac{1}{3}$,

$$\ddot{a} > 0 \quad \text{, iff, } \omega < -\frac{1}{3}. \tag{12.12}$$

Now, let's look at Eq 12.4. We observe today that $\rho \approx \rho_c$ (to within 1%), therefore $\frac{K}{(aH)^2}$ must make a very small contribution as the LHS is 1. However, in a deceleration phase, if we kook backward in time, $(aH)^2$ would becomes larger, therefore this terms made an even smaller contribution as we look back in time. This is known as the *flatness problem*. For the universe to appear as flat as it is today, it must have been even flatter at the start of the radiation dominated era.

Inflation's explanation is that at some early time, the Hubble sphere was not a small fraction of the whole Hubble sphere, maybe it was the same order of magnitude as the total sphere. Now there was an exponentially increase in the size, therefore the sphere would have shrunk to an exponentially small size and therefore at the beginning of the radiation era, the sphere would have been very small and slowly accelerated into the larger sphere we see today, which would still appear flat. If $\frac{1}{aH}$ decreased by a factor e^N during inflation (where N is called the "number of e-folds of inflation") then the condition is $N \geq \ln\left(\frac{(aH)_r}{(aH)_0}\right)$, where the subscripts r and 0 denotes the start of the radiation era or the present day, respectively.

1.2. Horizon problem. Around an observer in a FRW universe, there exists a future horizon. This is unlike a static black hole metric where there is a horizon around the black hole. The future horizon is the maximum limit to which a signal can be sent to. Similarly a past horizon is one in which any signal sent beyond this horizon will not reach an observer at the present day. In an accelerating universe, observers will have future horizons, in a decelerating universe, observers will have past horizons. This is easily seen by the FRW metric,

$$ds^2 = a^2(\eta)[-d\eta^2 + d\chi^2 + \chi^2 \sin^2 \chi d\theta^2], \tag{12.13}$$

where η is the conformal time. Photons follow null geodesics and taking a radial path of the photon (without loss of generality due to spherical symmetry),

$$\eta = \int \frac{dt}{a} = \int \frac{1}{\dot{a}}\frac{da}{a} \tag{12.14}$$

but $\dot{a} = a\frac{\dot{a}}{a} \equiv Ha$,

$$\eta = \int_0^{a_0} \frac{1}{aH}\frac{da}{a}. \tag{12.15}$$

If the co-moving Hubble radius was constant, then η would be proportional to $\ln a$, which means we would have a singularity at the origin. However, if the co-moving Hubble radius $\frac{1}{aH}$ goes to zero at the origin, then the singularity will be removed as the $\ln a$ will diverge very slowly compared to $\frac{1}{aH}$ going to zero, and inflation does exactly this. It forces $\frac{1}{aH}$ to go to zero and get rid of the singularity.

Another way of expressing the Horizon problem is words, is to say that the CMB appears to be extremely to be extremely uniform in every direction (to one part in 10^5). Yet according to the big bang model, the only causally connected regions of space at the time of re-combination should have been around one degree in angular size as observed today. Therefore it is not possible that the entire universe at that time could have thermally equilibriated to such a uniform temperature. Which then raises the question of why the universe appeared to be so uniform at that time. Inflation provides the solution to this problem by saying taking some small patch of the universe that was initially in thermal equilibrium and driving to exponentially in size to make it very large, such that the original causally connected patch would no longer appear to be in causal contact.

1.3. Monopole problem. In the 1970's, the first GUT's had first been discovered. John Preskill noticed that if the universe is described by a GUT, then as the universe cooled through the GUT temperature (energy scale of GUT's) scale, one gets spontaneous symmetry breaking. The GUT symmetry break down to the symmetry of the standard model. The bosonic fields are disordered at the horizon scales at that time and as a result, these stable field configurations that get formed are called magnetic monopoles, each of which is then stable and has a mass roughly at the GUT scale, $\approx 10^{16}GeV$. There would be about one of these per Hubble volume at the time. The Hubble volume today is much larger and therefore the overall mass of these monopoles today would be extremely large and would certainly be observable, therefore these monopoles are ruled out by observations.

Inflation solves this problem in a similar way to the solution of the Horizon problem. If the GUT symmetry breaking happened before inflation, then each small Hubble co-moving sphere will have about one monopole and then it is exactly one of these spheres that is blown up exponentially into the universe we see today and therefore there would be only one magnetic monopole floating around in the universe today (no wonder we haven't found it!).

2. Quantum field theory in curved space-time

Let's start with the simplest action; the action for a free real scalar field, ϕ, and take the space to be flat initially, i.e $g_{\mu\nu} = \eta_{\mu\nu}$,

$$S = \int d^4xL = \int d^4x \left(-\frac{1}{2}\eta^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\phi) - \frac{m^2\phi^2}{2} \right). \tag{12.16}$$

2.1. Classical field theory. We start with classical field theory. By differentiating L w.r.t $\dot{\phi}$, we obtain the momentum canonically conjugate to ϕ,

$$\pi \equiv \frac{\partial L}{\partial \dot{\phi}} = \dot{\phi}. \tag{12.17}$$

By varying the action w.r.t ϕ, we obtain the classical equations of motion, which is the flat space Klein-Gordon equation,

$$-\Box\phi + m^2\phi = 0. \tag{12.18}$$

where $\Box \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu$, is the wave operator in flat space. It we define the inner product between two field configurations to be,

$$\langle\phi_1|\phi_2\rangle \equiv -i\int d^3\vec{x}(\phi_1(\partial_t\phi_2^*) - (\partial_t\phi_1)\phi_2^*). \tag{12.19}$$

Note that ∂_t is used as ϕ only depends on time due to the homogeneity of the field and thus all the spatial derivatives are zero. The solutions to Eq 12.18 are

$$\phi_{\vec{k}} = \frac{1}{\sqrt{2\omega}}e^{i(\vec{k}\vec{x}-\omega t)}, \qquad (\omega = \sqrt{\vec{k}^2 + m^2}). \tag{12.20}$$

Indeed, these positive-frequency plane-waves $\phi_{\vec{k}}$ form a complete orthonormal basis for the solutions of Eq 12.18, so that we can expand any such solution in the form

$$\phi(\vec{x},t) = \int \frac{d^3\vec{k}}{(2\pi)^2}(a_{\vec{k}}\phi_{\vec{k}} + a^*_{\vec{k}}\phi^*_{\vec{k}}). \tag{12.21}$$

2.2. Quantum field theory. We quantise by promoting π and ϕ to be space-time operator that satisfy canonical commutation relations given below

$$\begin{aligned} &[\phi(\vec{x},t), \phi(\vec{x}',t)] = [\pi(\vec{x},t), \pi(\vec{x}',t)] = 0 \\ &[\phi(\vec{x},t), \pi(\vec{x}',t)] = i\delta^3(\vec{x}-\vec{x}'). \end{aligned} \tag{12.22}$$

It is more convenient to discuss these fields in terms of particles. We do that by going to Fourier space. So we promote the Fourier coefficients $a_{\vec{k}}$ and $a^*_{\vec{k}}$ in Eq 12.21 to operators $a_{\vec{k}}$ and $a^\dagger_{\vec{k}}$ satisfy the algebra of creation and annihilation operators,

$$\begin{aligned} &[a_{\vec{k}}, a_{\vec{k}'}] = [a^\dagger_{\vec{k}}, a^\dagger_{\vec{k}'}] = 0 \\ &[a_{\vec{k}}, a^\dagger_{\vec{k}'}] = \delta^3(\vec{k}-\vec{k}'). \end{aligned} \tag{12.23}$$

The vacuum state $|0\rangle$ is the state defined by the fact that it is annihilated by all of the annihilation operators,

$$a_{\vec{k}}|0\rangle = 0, \qquad \forall\vec{k} \tag{12.24}$$

and all the subsequent higher energy states are created by action with creation operators, $a^\dagger_{\vec{k}}$. Acting with a creation operator once on the ground state creates a quantised excitation that is interpreted to be a particle of that field. In fact one can define an operator, $N_{\vec{k}}$, that measures the number of excitations in a field (at the wavelength, $\vec{k}$)

$$N_{\vec{k}} \equiv a^\dagger_{\vec{k}}a_{\vec{k}}. \tag{12.25}$$

2.3. Quantum field theory in curved space-time. In curved space-time the metric is no longer Minkowski, therefore the action becomes

$$S = \int d^4x\sqrt{-g}\left(-\frac{1}{2}g^{\mu\nu}(\partial_\mu\phi)(\partial_\nu\phi) - \frac{m^2\phi^2}{2}\right). \tag{12.26}$$

The calculation follows the same procedure as we did in the flat metric. We obtain the conjugate momenta to ϕ, π_ϕ

$$\pi_\phi = \frac{\partial L}{\partial\dot{\phi}} = -\sqrt{-g}m^2\phi. \tag{12.27}$$

The action is varied w.r.t ϕ, which gives the same equations of motion as before except the metric is no longer Minkowski therefore the wave operator is different;

$$-\Box\phi + m^2\phi = 0, \qquad \Box \equiv g^{\mu\nu}\nabla_\mu\nabla_\nu. \tag{12.28}$$

The inner product is now defined as

$$\langle\phi_1|\phi_2\rangle \equiv -i\int d^3\vec{x}\sqrt{g_3}u^\mu(\phi_1(\partial_\mu\phi_2^*) - (\partial_\mu\phi_1)\phi_2^*) \tag{12.29}$$

where the integral is over a spatial (constant time) slice parametrised by the spatial coordinates, $\vec{x}$, and g_3 is the determinant of the spatial 3 metric of the slice, and u^μ is a time-like unit vector orthogonal to the slice. As before, we expand the field in terms of a complete set of positive frequency solutions ϕ_n and negative frequency solutions ϕ_n^* of the wave equation

$$\phi(\vec{x},t) = \sum_n(a_n\phi_n + a_n^*\phi_n^*). \tag{12.30}$$

The solutions ϕ_n, ϕ_n^* to the equations of motion are no longer simple plane waves, since they are in curved space-time. The key point is that in curved space, different observers (e.g at different points in space-time, or at the same points but in different states of motion) will in general disagree about how to split up the solutions into positive and negative frequency. Suppose one observer, O expands ϕ as Eq 12.30 and another observer, $\bar{O}$, expands ϕ as

$$\phi(\vec{x},t) = \sum_n(\bar{a}_n\bar{\phi}_n + \bar{a}_n^\dagger\bar{\phi}_n^\dagger) \tag{12.31}$$

where $\bar{\phi}$ can be expanded in terms of the ϕ's as they form a complete orthonormal set,

$$\bar{\phi}_n = \alpha_{nm}\phi_m + \beta_{nm}\phi_m^\dagger. \tag{12.32}$$

The coefficients α and β are called Bogliubov coefficients. Once we have the expressions for ϕ_n and $\bar{\phi}_n$, we can extract the Bogliubov coefficients by calculating the appropriate inner products,

$$\alpha_{mn} = \langle\bar{\phi}_m|\phi_n\rangle \qquad \beta_{mn} = -\langle\bar{\phi}_m|\phi_n^\dagger\rangle. \tag{12.33}$$

Substituting in for $\bar{\phi}_n$ from Eq 12.32 into Eq 12.31 and comparing to Eq 12.30, the observer O's creation and annihilation operators are related to $\bar{O}$'s by the so-called "Bogliubov transformation":

$$a_n = \bar{a}_m\alpha_{mn} + \bar{a}_m^\dagger\beta_{mn}^\dagger. \tag{12.34}$$

Since the $\bar{a}_n$'s and $\bar{a}_n^\dagger$;s must obey the creation/annihilation commutation relations that are identical to the ones obeyed by a_n and $a_n^\dagger$, we find that, if we regard the Bogliubov coefficients α_{mn} and β_{mn} are matrices, they must obey the following constraints

$$\beta_\alpha^\dagger - (\beta_\alpha^\dagger)^\dagger = 0 \qquad 1 = \alpha^\dagger\alpha - (\beta^\dagger\beta)^\dagger. \tag{12.35}$$

O and $\bar{O}$ will each define their own vacuum state to be the state that is annihilated by all of their own annihilation operators

$$a_n|0\rangle = 0 \qquad \bar{a}_n|\bar{0}\rangle = 0 \tag{12.36}$$

and they will each define their own number operators in terms of their creation and annihilation operators,

$$N_n \equiv a_n^\dagger a_n, \quad \bar{N}_n \equiv \bar{a}_n^\dagger\bar{a}_n. \tag{12.37}$$

So according to $\bar{O}$, the state $|\bar{O}\rangle$ is vacuum, containing no particles. But according to O, it is full of particles. Indeed, according to O, the expectation value of the number of particles of type n in the state $|\bar{0}\rangle$ is given by

$$\begin{aligned}\langle\bar{0}|N_n|\bar{0}\rangle &= \langle\bar{0}|(\alpha_n^*\bar{a}_n^\dagger + \beta_n\bar{a}_n)(\alpha_n\bar{a}_n + \beta_n^*\bar{a}_n^\dagger)|\bar{0}\rangle \\ &= |\beta_n|^2\langle\bar{0}|\bar{a}_n\bar{a}_n^\dagger|\bar{0}\rangle \\ &= |B_n|^2\langle\bar{0}|\underbrace{1 - \bar{a}_n^\dagger\bar{a}_n}_{T1}|\bar{0}\rangle \\ &= |B_n|^2\end{aligned} \tag{12.38}$$

where in $T1$ we have used the commutation relation, $[\bar{a}_n \bar{a}_n^\dagger] = 1$.

2.4. Unruh effect. We look at two different coordinates of flat space. Consider a 2D flat space with time coordinates, t and space coordinates, x. The metric of this space is

$$ds^2 = -dt^2 + dx^2. \tag{12.39}$$

We want to, eventually, go to coordinates that correspond to a uniformly accelerating observer. A uniformly accelerating observer, moves along a hyperbola, that asymptotes to the light cone null rays (look at figure 48 for picture). First, let's switch to coordinates in the frame of the light cone, $u_\pm \equiv x \pm t$; which is equivalent to $du_\pm = dx \pm dt$. Substituting this into Eq 12.39 gives

$$ds^2 = du_+ du_-. \tag{12.40}$$

Now we introduce $u_\pm = x \pm t = \frac{1}{a} e^{a\bar{u}_\pm}$, sometimes called the "warped" light cone coordinates; the metric now becomes

$$ds^2 = e^{a(\bar{u}_+ + \bar{u}_-)} d\bar{u}_+ d\bar{u}_-. \tag{12.41}$$

Finally we switch to "warped Cartesian coordinates", $\bar{u}_\pm \equiv \bar{x} \pm \bar{t}$,

$$ds^2 = e^{2\bar{x}}(-d\bar{t}^2 + d\bar{x}^2). \tag{12.42}$$

The metric in the barred coordinate looks just like the original metric, except it has an extra conformal factor in front. These barred coordinates are called Rindler coordinates. Notice that while $U_\pm$ runs from $-\infty$ to $+\infty$, the $\bar{u}_\pm$ only run from 0 to ∞ as the exponential is always positive. Now let's consider a mass-less scalar field on the Rindler space. The equation of motion for this field in Minkowski or Rindler coordinates is

$$\Box\phi = 0 \qquad \Box = g^{\mu\nu}\nabla_\mu\nabla_\nu. \tag{12.43}$$

In Minkowski, $g^{\mu\nu} \equiv \eta^{\mu\nu}$, therefore Eq 12.43 is simply

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right)\phi = 0. \tag{12.44}$$

In Rindler space however, Eq 12.43 becomes

$$e^{-2a\bar{x}}\left(\frac{\partial^2}{\partial \bar{t}^2} - \frac{\partial^2}{\partial \bar{x}^2}\right)\phi = 0. \tag{12.45}$$

However one can just divide through by $e^{-2a\bar{x}}$ to get an identical equation as Eq 12.44. The positive frequency solutions for Eq 12.44 are

$$e^{i(kx-\omega t)} \propto \phi_k \qquad (\omega = |k|). \tag{12.46}$$

The positive frequency solutions for Eq 12.45 are

$$\bar{\phi}_k \propto e^{i(k\bar{x}-\omega\bar{t})}. \tag{12.47}$$

Firstly, we want the wavefunctions to be normalised, so the inner product in Eq 12.39 must be satisfied. The normalisation factor is $\frac{1}{\sqrt{2\omega}}$. Imposing the conditions, $\langle\phi_n|\phi_m\rangle = \delta_{mn}, \langle\phi_n^*|\phi_m^*\rangle = -\delta_{mn}, \langle\phi_n|\phi_m^*\rangle = 0$, one can get the β coefficients via,

$$\langle\phi_n|\bar{\phi}_m^*\rangle = \beta_{mn}. \tag{12.48}$$

To compute the inner product in Eq 12.48, one has to convert the wavefunctions so that both of them are in the same coordinate system, either Minkowski space or Rindler space. Computing the inner product and obtaining $|\beta|^2$, which gives the measure of the number of particles, shows that the number of particles is that given by the Bose-Einstein distribution for an ideal has of bosons at finite temperature. Starting from the Minkowski line element

$$ds^2 = -dt^2 + dx2 \tag{12.49}$$

we can make the coordinate transformation,

$$t = \rho \sinh \sigma \qquad x = \rho \cosh \sigma \tag{12.50}$$

to obtain

$$ds^2 = d\rho^2 - \rho^2 d\sigma^2. \tag{12.51}$$

Along the imaginary time direction, $\phi = i\sigma$, the line element becomes,

$$ds^2 = d\rho^2 + \rho^2 d\phi^2 \tag{12.52}$$

which is precisely the flat Euclidean line element expressed in polar coordinates. This space is, of course, periodic in the ϕ direction, with period 2π. Now, a Rindler observer riding along a curve of constant $\rho = \rho_0$ has constant proper acceleration $a = \frac{1}{\rho}$ and if we look at the line element, the corresponding proper time coordinate is not σ, but $\tau = \rho_0\sigma$. Therefore, $\sin\phi$ has a period 2π, this observer's proper time is periodic in the imaginary time direction, with period $-2\pi\rho i$. Identifying this with $-i\beta$, we find that this observer sees a temperature,

$$T = \frac{1}{2\pi\rho_0} = \frac{a}{2\pi}. \tag{12.53}$$

The Unruh effect is an example of a more general phenomena; *thermal states are periodic in imaginary time.* We can see this by simple quantum mechanics and statistical mechanics. In quantum mechanics, if we have a state $|\psi_1\rangle$ at time t_1 and $|\psi_2\rangle$ at time t_2, and we want to know what is the amplitude, $\mathbb{T}$, of the state $|\psi_1\rangle$ evolving into $|\psi_2\rangle$, in the time $t_2 - t_1$, we use the time evolution operator, e^{-iHt},

$$\mathbb{T} \equiv \langle\psi_2|e^{-iH(t_2-t_1)}|\psi_1\rangle \tag{12.54}$$

In statistical mechanics, a state in thermal equilibrium at temperature T is described by the density matrix, $\rho \propto \exp(-\beta H)$, where $\beta = \frac{1}{k_B T}$ and the key object is the partition function, Z, given by

$$Z = Tr(e^{-\beta H}) = \sum_n \langle\psi_n|e^{-\beta H}|\psi_n\rangle. \tag{12.55}$$

The expectation value of a physical observable A, which has a corresponding Hermitian operator, $\hat{A}$, $\langle A\rangle$, is given by $Tr(\rho, A)$. Let's look at a single term in Eq 12.55,

$$\langle\psi_m|e^{-\beta H}|\psi_m\rangle. \tag{12.56}$$

By comparing Eq 12.56 to 12.54, we see that if $-i\beta = (t_2 - t_1)$, then these two expressions are the same. So what this is saying that calculating the transition amplitude of going from state $|\psi_2\rangle$ to $|\psi_1\rangle$ in time (t_0, t_1), is equivalent to taking the state $|\psi_m\rangle$ and evolving it in the imaginary time direction, by a factor $-\beta$, and finding the amplitude that it comes back to itself.

2.5. De-Sitter space temperature. Let's apply the formulation used for the Unruh affect for de-Sitter space. In static coordinates the de-Sitter line element is

$$ds^2 = -\left(1 - \frac{r^2}{\rho_0^2}\right)dt^2 + \left(1 - \frac{r^2}{\rho_0^2}\right)^{-1} dr^2 + r^2 d\Omega_2^2 \tag{12.57}$$

where $\rho_0 = \frac{1}{H_0}$ is the de-Sitter radius. Let's look at radii closer to the horizon, i.e define coordinates $R = \rho_0 - r$; and look at the region $0 \le R << \rho_0$. In this case, we get

$$\left(1 - \frac{r^2}{\rho_0^2}\right) = \left(1 - \frac{(\rho_0 - R)^2}{\rho_0^2}\right) \approx \frac{2R}{\rho_0}. \tag{12.58}$$

Substituting Eq 12.58 into 12.57,

$$ds^2 = -\left(2\frac{R}{\rho_0}\right)dt^2 + \underbrace{\left(\frac{1}{2R/\rho_0}\right)dR^2}_{T2} \tag{12.59}$$

where the angular part has been ignored for now. We define $T_2 = d\chi^2$, therefore

$$d\chi = \frac{dk}{\sqrt{2R/\rho}}. \tag{12.60}$$

Integrating, $\chi = \sqrt{2\rho_0}R^{\frac{1}{2}}$. Substituting Eq 12.60 into Eq 12.58,

$$ds^2 = -\frac{\chi^2}{\rho_0^2}dt^2 + d\chi^2. \tag{12.61}$$

Now if we switch to imaginary time coordinates, $\phi = i(t/\rho_0)$, we find;

$$ds^2 = d\chi^2 + \chi^2 d\phi^2 \tag{12.62}$$

which is once again the flat Euclidean metric $ds^2 = dx^2 + dy^2$, expressed in polar coordinates. Again by the fact that ϕ is periodic in 2π, the t must have period $-2\pi\rho_0 i = -i\beta$, or in other words the de-Sitter space has temperature

$$T = \frac{1}{2\pi\rho_0} = \frac{H}{2\pi}. \tag{12.63}$$

The constant temperature of de-Sitter space excites a scale invariant spectrum of fluctuations in a mass-less scalar field. The fluctuations are probably the most important thing in cosmology as without them, we wouldn't be here, infact nothing interesting would happen in the universe if it were perfectly homogenous and isotropic.

3. Density fluctuations

If the universe has some density fluctuations in it, then the first thing to do is to Fourier decompose this function that describes the fluctuations in the universe,

$$\rho(\vec{x}) = \int \frac{d^3k}{(2\pi)^3}\rho_{\vec{k}}e^{-i\vec{k}\vec{x}} \tag{12.64}$$

this is done because the universe is, to first approximation, translationally invariant in space, then at linear order, as long as these perturbations are small (i.e quadratic terms can be ignored), the individual Fourier components remain independent of each other, and each Fourier mode evolves independently (this is an observed fact about the universe). So we have a single sine wave, which in co-moving coordinates has a fixed length. Therefore the only thing to think about is how does it's amplitude evolve in time. Since the physical wavelength of the Fourier mode, $\lambda_a = a(t)\lambda$ (λ is the co-moving wavelength), increases with time, it is important to see weather it becomes larger than the Hubble radius, r_H, given by $\frac{c}{H(t)}$, as there will be a dramatic transition when these two length's cross. Usually the situation is that if λ_p is bigger then r_H, the amplitude of the Fourier mode is fixed. When r_H becomes larger than λ_P, the Fourier mode starts to oscillate. If this Fourier mode belongs to a gravitation wave, the oscillations dampen down the amplitude and make the signal very weak. If the Fourier mode belongs to a density perturbation, then the amplitude increases and turns into galaxies we see today. Ever since the universe has been in the radiation dominated era, the universe has been decelerating, (neglecting the relatively recent Λ dominated era). During the time that it has been increasing faster then the physical wavelengths proportional to a, λ_a. Therefore if the Fourier modes are around for long enough, as soon as r_H is greater than λ_p, they will oscillate as described above.

The idea of inflation is that we assume that there was an epoch before to the radiation dominated epoch, during which the universe was exponentially accelerating, thus the Hubble radius would have been growing slower than the physical length scale of the Fourier modes, λ_p. If the universe was accelerated by some form of matter, described by the equation of state parameter, $\omega = -1$ (i.e vacuum energy), then during inflation, r_H, is constant and $a(t)$ grows exponentially. So the physical length scale starts off as very small compared to the Hubble scale.

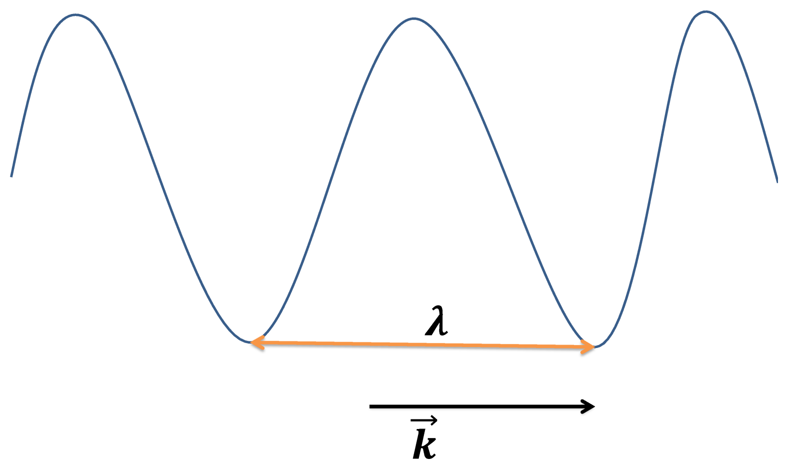


FIGURE 69. This represents one Fourier mode oscillation with wave-vector $\vec{k}$. In co-moving coordinates, this wave does not stretch its physical length stretches proportional to a, but its co-moving wavelength remains constant.

This means that the Fourier mode can be considered to be confined locally in the overall manifold that describes the geometry of the universe. Therefore initially the Fourier mode will think it is in flat Minkowski space. The accelerated expansion of the universe makes this physical wavelength grow exponentially, relative to the Hubble length scale (as r_H is roughly constant) and will soon become longer than the Hubble radius. In cosmology, when λ_p is smaller than, r_H it is said to be "inside the horizon", when λ_p is larger than r_H it is said to be "outside the horizon". So in the inflationary scenario, the Fourier modes start of inside the horizon, inflation causes rapid expansion and therefore takes the modes outside the horizon. Once inflation is finished, the modes re-enter the horizon at some later time, as the universe decelerates in the radiation and matter dominated era. Physically, this Fourier mode is sitting in the vacuum state, with some zero point energy, say $\frac{\hbar\omega}{2}$, and oscillates with frequency $\omega = \frac{k_c}{a(t)}$ (as $\omega = \sqrt{\frac{k^2}{a} + m^2}$ and we have a mass-less field), where $k_c = \frac{2\pi}{\lambda_c}$.

In the beginning when $a(t)$ is small, ω is very large and therefore the energy of the oscillations is high. Recall that for de-Sitter space the temperature is $\frac{H}{2\pi}$ and since r_H is constant and is less than λ_c, during inflation, the temperature of the space is not large enough to excite this Fourier mode. However at the point when $\lambda_c \approx r_H$, ω becomes small, therefore there comes a point at which the temperature increases enough to excite the Fourier mode and cause it to oscillate. This is what is meant when one says that perturbations are created.

When we look at the plot of the CMB in figure 70, we see the values of C_l (which is the mean square value of the spherical harmonics over all l at a given m) plotted against l. The data agrees beautifully with the predicted spectrum. The spectrum we see is the spectrum of fluctuations roughly 3×10^5 years after the big bang. When we are calculating the theoretical curve that fits it, the idea is that the initial spectrum of perturbations, had no none of the bumps seen in the spectrum. In other words, if one plots the mean square size of the perturbations in the early universe, at the beginning of the radiation dominated epoch as a power law in k, i.e k^n, then the value of n the so-called spectral index is roughly zero. So if a Fourier mode is chosen randomly in the early universe and we ask what it's amplitude is; first of all the universe is isotropic, therefore the amplitude does not depend on the direction of $\vec{k}$ corresponding to that Fourier mode, it only depends on the magnitude of k. The observational fact that makes the theoretical curve fit the observed bumps in the CMB spectra is saying that the power spectrum of k^n has a spectral index that is roughly zero (-0.04 $\pm$ 0.01). Therefore to first approximation every Fourier mode is roughly the same in amplitude (to first approximation) in the early universe. Therefore we say that the perturbations in the early universe were "scale invariant".

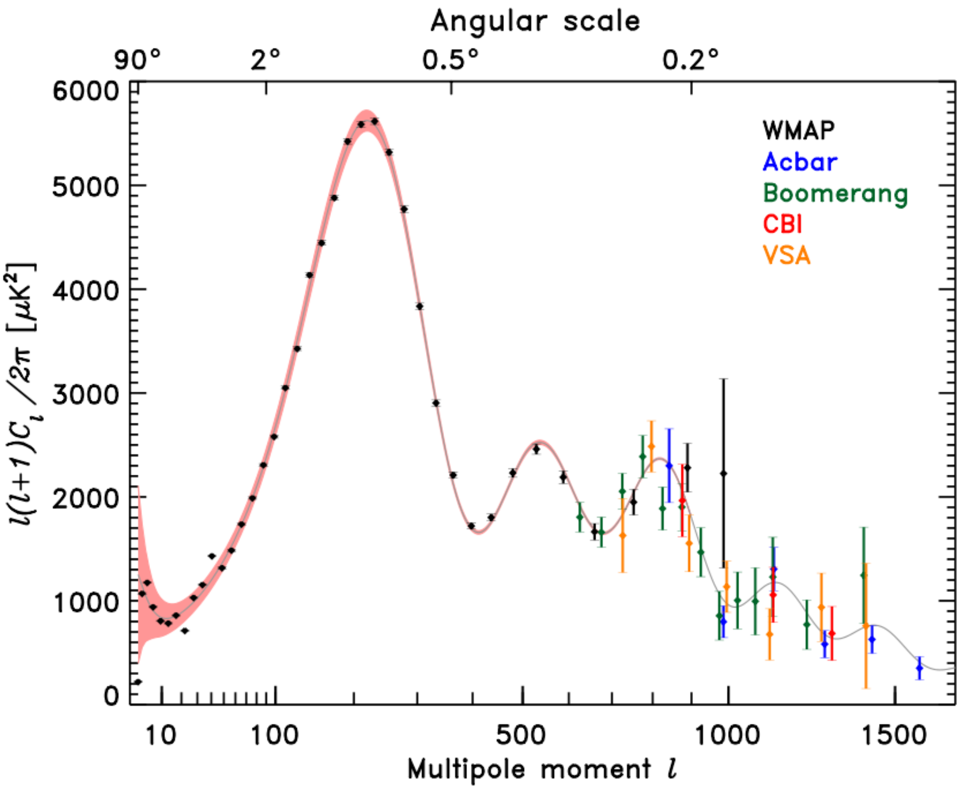


FIGURE 70. CMB spectrum taken from Wikipedia. The peaks come from the baryonic acoustic oscillations that carry an imprint of the primordial density fluctuations in the amplitude of these peaks.

Now, if we go to the time where inflation ended and the radiation dominated epoch began. The universe is decelerating, therefore r_H is increasing relative to λ_c and the Fourier modes eventually come inside the horizon and then the CMB photons were released. At the beginning of the radiation dominated era, all the Fourier modes had the same amplitude. As $\lambda_c \propto H_c$, the modes become excited again. If the Fourier mode corresponds to a density perturbation, then the perturbation will grow as the regions with more matter will have more gravitational pull and will attract more matter (whereas if the density has a component of pressure, i.e photons, then we get a competition between pressure and gravitational force which leads to baryon acoustic oscillations as described in section 4.2). The role of inflation here is to provide a scale invariant primordial spectrum of perturbations.

4. Inflationary perturbations

4.1. Single-field slow-roll inflation. There were problems with the homogeneity of the universe in the 1980's (flatness, horizon, monopole) originally lead to the idea of inflation. It was soon realised that the simplest way to obtain inflation was with a slowly varying scalar field, generally called the inflaton field. To start of with, we take a closer look at this single-field slow roll inflation.

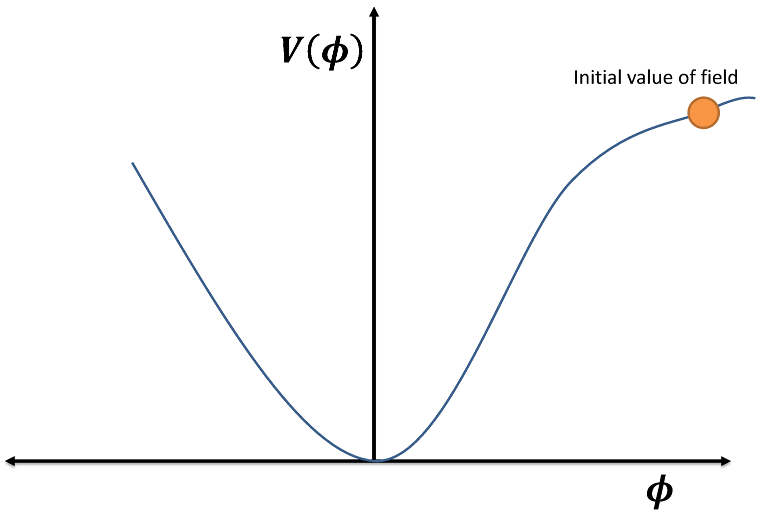


FIGURE 71. The shape of the inflaton field, ϕ, potential, $V(\phi)$.

The field starts out at a potential that is not minimum (for unexplained reasons) and slowly rolls down the potential, following it's equations of motion. The action is that of general relativity with the matter Lagrangian of a general scalar field

$$S = \int d^4x\sqrt{-g}\left(\frac{R}{16\pi G} - \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right). \tag{12.65}$$

We define the slow-roll parameters,

$$\epsilon = \frac{1}{2}\frac{m_{pl}^2}{8\pi}\left(\frac{V'(\phi)}{V(\phi)}\right)^2 \tag{12.66}$$

$$\eta = \frac{m_{pl}^2}{8\pi}\frac{V''(\phi)}{V(\phi)} \tag{12.67}$$

We assume that the potential $V(\phi)$ is such that the field can "roll" monotonically down it's potential (with ϵ and η are both $<< 1$), to a local minimum at $\phi = \phi_{min}$ with $V(\phi_{\rm min}) = 0$. This is the framework for a "single-field slow roll inflation". Varying this action w.r.t ϕ, gives the equation of motion

$$-\Box\phi + V(\phi) = 0. \tag{12.68}$$

By varying w.r.t g we get the usual Einstein field equations,

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{12.69}$$

To begin with we look at the homogenous solutions, i.e the metric $g_{\mu\nu}$ is the FRW metric and ϕ is only a function of time, $\phi_0(t)$. In this case, Eq 12.68 becomes

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0 \tag{12.70}$$

and Eq 12.69 gives

$$H^2 = \frac{8\pi G\rho}{3} \tag{12.71}$$

where ρ is the energy density of the inflaton field,

$$\rho = \frac{\dot{\phi}^2}{2} + V(\phi). \tag{12.72}$$

Eq 12.70 is like the equation of motion for a particle experiencing two different forces; a potential force (the V term) and a friction force (the $3H\phi$ term). The field starts in the slow-roll regime: it is strongly over-damped and quickly relaxes to it's "terminal velocity", where it's acceleration is negligible in Eq 12.70 and the drag approximately balances the potential force

$$\dot{\phi} \approx -\frac{V'(\phi)}{3H}. \tag{12.73}$$

This means that the condition $\epsilon << 1$ becomes $\dot{\phi} << V(\phi)$, which implies $w \approx -1$ and here $\ddot{a} > 0$. Eventually when the field gets close to the minimum so that $V(\phi) \approx \frac{1}{2}m^2\phi^2 + \mathcal{O}(\phi^4)$ and $H < m$, the field begins under-damped oscillations $\phi(t) \propto a^{-\frac{3}{2}} \cos{(mt)}$ and the energy density decays as $\rho \propto a^{-3}$ which implies $w = 0$ and hence $\ddot{a} < 0$. So the field starts in it's slow-roll regime, gradually rolls down it's potential as the universe accelerates; finally close enough to the minimum, the slow-roll conditions cease to hold and the ϕ begins under-damped oscillations about it's minimum as the universe stops accelerating and begins to decelerate.

By quantum mechanics we know that the vacuum state of this field cannot be empty, but must have zero point fluctuations. As the fields rolls down the potential, these fluctuations get amplified by the background stretching of the universe into a scale invariant spectrum of perturbations, which matches the primordial spectrum of perturbations.

4.2. Perturbations. So far we have Eq 12.68 and Eq 12.69 which are the general equations of motion and then Eq 12.70 and Eq 12.71 are obtained in the homogenous approximation. Now we don't assume that the field is homogenous. Instead we take the homogenous case as a first approximation and perturb it. The single field, $\phi(t, \vec{x})$ is perturbed as follows,

$$\phi(t, \vec{x}) = \phi_0(t) + \delta\phi(t, \vec{x}) \tag{12.74}$$

where $\phi_0(t)$ is the zeroth order homogeneous limit field. Similarly, the metric is perturbed as,

$$g_{\mu\nu}(t, \vec{x}) = g^{(0)}_{\mu\nu}(t) + \delta g_{\mu\nu}(t, \vec{x}) \tag{12.75}$$

where $g^{(0)}_{\mu\nu}(t)$ is the FRW metric. Putting Eq 12.74 into 12.68 gives,

$$-\Box(\phi_0 + \delta\phi) + V'(\phi_0 + \delta\phi) = 0. \tag{12.76}$$

Taylor expanding V' in small quantity $\delta\phi$ gives

$$-\Box(\phi_0 + \delta\phi) + V'(\phi_0) + V''(\phi_0)\delta\phi = 0. \tag{12.77}$$

However from Eq 12.68, this simplifies to

$$\Box\delta\phi + V''(\phi_0)\delta\phi = 0, \tag{12.78}$$

similarly, perturbing Eq 12.69 gives

$$G_{\mu\nu} = G^{(0)}_{\mu\nu} + G^{(1)}_{\mu\nu} \tag{12.79}$$

$$8\pi G T_{\mu\nu} = 8\pi G(T^{(0)}_{\mu\nu} + T^{(1)}_{\mu\nu}) \tag{12.80}$$

Equating Eq 12.79 to 12.80 (again the zeroth order terms cancel from Eq 12.69)

$$G^{(1)}_{\mu\nu} = 8\pi G T^{(1)}_{\mu\nu}. \tag{12.81}$$

Now Eq 12.78 and 12.81 would need to be solved to obtain the mode functions. When we perturb a single in FRW space one would naively think that there are 10 + 1 perturbations (i.e 10 components from the symmetric metric and 1 from the scalar field). The metric can be decomposed using a so-called ADM decomposition as

$$ds^2 = Ndt^2 + \gamma_{ij}(dx^i + N^i dt)(dx^j + N^j dt). \tag{12.82}$$

The metric has 10 independent degrees of freedom, here we have 6 that go into γ_{ij} (3D metric), which represents coordinates at constant t. The remaining 4 components go into the so called "lapse", N, "shift" N^i. These 4 N and N^i, degrees of freedom can be removed by constraint equations in general relativity (which come from the Bianchi identities $R^{\alpha}_{\beta[\gamma\delta;\epsilon]} = 0$). Now γ_{ij} has 6 degrees of freedom, however we can still impose 4 gauge conditions to get the number of degrees of freedom down to 2. The gauge chosen to work in is,

$$\delta\phi = 0, \qquad \gamma_{ij} = a^2(t)e^{2\zeta(t,\vec{x})}e^{2h_{ij}(t,\vec{x})} \tag{12.83}$$

where $\delta^{ij}h_{ij} = \delta^{ij}\partial_i h_{jk} = 0$. Physically, the 3 physical modes are ζ (which, in this $\delta\phi = 0$ gauge, physically corresponds to the perturbation of the Ricci 3 curvature of the spatial slices) and the two independent components of a traceless transverse 3×3 symmetric matrix h_{ij} (which physically corresponds to the two polarizations of a gravitational wave). ζ is referred to a "primordial scalar perturbations" or "primordial curvature perturbations" it is observed more or less directly by observing the fluctuations in the temperature of the CMB between different points on the sky; it is the primordial perturbation that is ultimately responsible for all of the density perturbations and structure that we see in the universe today. The 2 independent components of h_{ij} are referred to as "primordial tensor perturbation"; they are primordial gravitational waves predicted by inflation, but are extremely difficult to observe.

4.3. Spectra. The starting point is to take the perturbations, put them into the original action and expand them up to quadratic in order in small perturbations, ζ, h_{ij}. The zeroth order part just gives back the homogenous equations, the first order terms are just zero, since the action is defined to be a minimum in a theory and so the second order (quadratic) part of the action describes the behavior of the perturbations[1]. The action is split up into the scalar part ζ and the tensor part h_{ij},

$$S_\zeta = -\frac{1}{2}\int d\eta d^3\vec{x} z^2 \eta^{\mu\nu}(\partial_\mu\zeta)(\partial_\nu\zeta) \tag{12.84}$$

$$S_h = -\frac{1}{2}\int d\eta d^3\vec{x}((m_{pl}a)^2\eta^{\mu\nu}h^s_{,\mu}h^s_{,\nu}) \tag{12.85}$$

where $\vec{x}$ is a co-moving coordinate, η is the conformal time, $\eta^{\mu\nu}$ is the Minkowski metric. The s superscript in Eq 12.85 needs to be summed over ± 1 to get both polarizations of the gravitational waves. z is not a perturbation. It is a background function of time, $z(\eta)$, is defined by

$$z^2(\eta) = 2\epsilon a^2(\eta) \tag{12.86}$$

where ϵ is the slow-roll parameter. In other words, if ϵ is constant (which it nearly is for inflation), $z \propto a$. z acts as the effective scale factor felt by the mass-less scalar field ζ. In fact, if z^2 is replaced by a^2, Eq 12.84 would be exactly the action for a mass-less scalar field on the Minkowski metric.

The two scalar fields ($s = \pm$) are also defined as,

$$h_s(\eta, \vec{x}) = \int \frac{d^3\vec{k}}{(2\pi)^3} h^s_{\vec{k}}(\eta)e^{i\vec{k}\vec{x}} \tag{12.87}$$

which are dimensionless. Let's also split up the ζ and h_{ij} into Fourier components;

[1]The full derivation is given in section 2 of http://arxiv.org/abs/astro-ph/0210603

$$\zeta(\eta, \vec{x}) = \int \frac{d^3\vec{k}}{(2\pi)^{\frac{3}{2}}} \zeta_{\vec{k}}(\eta) e^{i\vec{k}\vec{x}} \tag{12.88}$$

$$h_{ij}(\eta, \vec{x}) = \int \frac{d^3\vec{k}}{(2\pi)^{\frac{3}{2}}} \epsilon^s_{ij}(\vec{k}) h^s_{\vec{k}}(\eta) e^{i\vec{k}\vec{x}} \tag{12.89}$$

where we sum over perturbations $s = \pm$ in Eq 12.89 and the 3×3 polarization tensors $\epsilon^\dagger_{ij}$ and ϵ_{ij} are real ($\epsilon^*_{ij} = \epsilon_{ij}$), symmetric ($\epsilon_{ij} = \epsilon_{ji}$), traceless ($\epsilon_{ii} = 0$), transverse ($\epsilon_{ij}k^j = 0$), even parity ($\epsilon_{ij}(\vec{k}) = \epsilon_{ij}(-\vec{k})$) and "ortho-normal" ($\epsilon^s_{ij}(\vec{k})\epsilon^{s'}_{ij}(\vec{k}) = 4\delta^{ss'}$). Of course, the final expression for the tensor spectrum will not depend on the normalisation of the polarization tensor ϵ^s_{ij}, but this particular choice is convenient, because it canonically normalizes the scalar fields.

The conjugate momenta for these actions is;

$$\begin{aligned} \pi_\zeta &= \frac{\partial L_\zeta}{\partial \zeta'} = z^2 \zeta' \\ \pi_h &= \frac{\partial L_h}{\partial h'_s} = (m_{pl} a)^2 h'_s \end{aligned} \tag{12.90}$$

where the primes denote the derivative w.r.t η and the equation of motion coming from these two actions are,

$$\begin{aligned} \zeta'' + 2\left(\frac{z'}{z}\right)\zeta' - \nabla^2\zeta = 0 \\ h''_s + 2\left(\frac{a'}{a}\right)h'_s - \nabla^2 h_s = 0 \end{aligned} \tag{12.91}$$

Now we canonically quantise;

$$\begin{aligned} [\hat{\zeta}(\eta, \vec{x}), \hat{\pi}_\zeta(\eta, \vec{x}')] = i\delta^{(3)}(\vec{x} - \vec{x}') \\ [\hat{\zeta}(\eta, \vec{x}), \hat{\zeta}(\eta, \vec{x}')] = [\hat{\pi}_\zeta(\eta, \vec{x}), \hat{\pi}_\zeta(\eta, \vec{x}')] = 0, \end{aligned} \tag{12.92}$$

and similarly for h_s

$$\begin{aligned} [\hat{h}_s(\eta, \vec{x}), \hat{\pi}_{s'}(\eta, \vec{x}')] = i\delta^{(3)}(\vec{x} - \vec{x}')\delta_{ss'} \\ [\hat{h}_s(\eta, \vec{x}), \hat{h}_{s'}(\eta, \vec{x}')] = [\hat{\pi}_s(\eta, \vec{x}), \hat{\pi}_{s'}(\eta, \vec{x}')] = 0. \end{aligned} \tag{12.93}$$

Since $\hat{\zeta}$ is a real and Hermitian, its Fourier components satisfy $\hat{\zeta}_{\vec{k}} = \hat{\zeta}^\dagger_{-\vec{k}}$, Similarly for $\hat{h}_s, \hat{h}^s_{\vec{k}} = \hat{h}^s_{-\vec{k}}$. Now we expand $\hat{\zeta}$ in terms of creation and annihilation operators;

$$\hat{\zeta} = \int \frac{d^3\vec{k}}{(2\pi)^3} \left(a_{\vec{k}} \zeta_{\vec{k}} + a^\dagger_{\vec{k}} \zeta^\dagger_{\vec{k}} \right). \tag{12.94}$$

In FRW, we want to write each of the positive frequency $\zeta_{\vec{k}}$ and negative frequency, $\zeta_{-\vec{k}}$ solutions as a part that depends pn space and a part that depends on time

$$\zeta_{\vec{k}} = \zeta_k(\eta) e^{i\vec{k}\vec{x}}. \tag{12.95}$$

Note $\zeta(\eta)$ only depends on the magnitude of $\vec{k}$ due to the isotropy of FRW. Putting Eq 12.95 into 12.94 and using $\zeta_{\vec{k}} = \zeta^\dagger_{\vec{k}}$ we get

$$\hat{\zeta}_{\vec{k}} = \int \frac{d^3k}{(2\pi)^3}\left(a_{\vec{k}}\zeta_k(\eta)e^{i\vec{k}\vec{x}} + a^\dagger_{-\vec{k}}\zeta^\dagger_k(\eta)e^{i\vec{k}\vec{x}}\right)$$
$$= \int \frac{d^3\vec{k}}{(2\pi)^3}\left(a_{\vec{k}}\zeta_k(\eta) + a_{-\vec{k}}\zeta^\dagger_k(\eta)\right)e^{i\vec{k}\vec{x}}. \quad (12.96)$$

Similarly for $\hat{h}^s_{\vec{k}}$

$$\hat{h}^s_{\vec{k}} = \int \frac{d^3k}{(2\pi)^3}\left(h_k(\eta)a^s_{\vec{k}} + h^*_k(\eta)a^{s\dagger}_{\vec{k}}\right). \quad (12.97)$$

Now the ζ creation and annihilation operators ($a^\dagger_{\vec{k}}$ and $a_{\vec{k}}$) and the h_s creation and annihilation operators ($a^{s\dagger}_{\vec{k}}$ and $a^s_{\vec{k}}$) satisfy the usual commutation relations,

$$[a_{\vec{k}}, a^\dagger_{\vec{k}}] = \delta^3(\vec{k} - \vec{k}')$$
$$[a_{\vec{k}}, a_{\vec{k}'}] = [a^\dagger_{\vec{k}}, a^\dagger_{\vec{k}'}] = 0$$
$$[a^s_{\vec{k}'}, a^{s'\dagger}_{\vec{k}'}] = \delta^3(\vec{k} - \vec{k}')\delta^{ss'}$$
$$[a^s_{\vec{k}}, a^{s'}_{\vec{k}'}] = [a^{s\dagger}_{\vec{k}}, a^{s'\dagger}_{\vec{k}'}] = 0, \quad (12.98)$$

while the classical mode functions, $\{\zeta_k(\tau), \zeta^*_k(\tau)\}$ and $\{h_k(\tau), h^*_k(\tau)\}$ are linearly independent solutions of the (Fourier transformed) equations of motion,

$$\zeta''_k + 2\left(\frac{z'}{z}\right)\zeta'_k + k^2\zeta_k = 0$$
$$h''_k + 2\left(\frac{a'}{a}\right)h'_k + k^2 h_k = 0. \quad (12.99)$$

The aim is to calculate the power spectrum of ζ and h. That means, if take the expectation value of the ζ and h at any point in space,

$$\langle 0|\hat{\zeta}^2(\vec{x}, t)/\hat{h}^{s2}(\vec{x}, t)|0\rangle, \quad (12.100)$$

where $a_{\vec{k}}|0\rangle \equiv 0$. This gives the integral below

$$\int d^3\vec{k}\frac{1}{(2\pi)^3}|\zeta_k|^2 = \int \frac{1}{(2\pi)^3}4\pi k^2 dk|\zeta_k|^2$$
$$\equiv \int d\ln k\frac{k^3}{2\pi^2}|\zeta_k|^2 \quad (12.101)$$

Where we have used the fact that the physics is independent of $\vec{k}$, therefore the integral can be evaluated in terms of $k = |\vec{k}|$. A similar calculation for the tensor perturbations can also be done. The power spectrum is then defined as,

$$P_\zeta \equiv \frac{d\langle 0|\hat{\zeta}^2(\eta, \vec{x})|0\rangle}{d\ln k} = \frac{k^3}{2\pi^3}|\zeta_k|^2$$
$$P_h \equiv \frac{d\langle 0|\hat{h}^2_{ij}(\eta, \vec{x})|0\rangle}{d\ln k} = \frac{8k^3}{2\pi^2}|h_k|^2, \quad (12.102)$$

where $\zeta_k(\eta)$ and $h_k(\eta)$ are obtained by solving the equations of motion Eq 12.99 w.r.t the following initial conditions,

$$\begin{aligned}\zeta_k &\rightarrow \frac{1}{2\sqrt{2k}}e^{-ik\eta}\\ h_k &\rightarrow \frac{1}{m_{pl}a\sqrt{2k}}e^{-ik\eta}.\end{aligned} \tag{12.103}$$

The $2\frac{z'}{z}\zeta_k'$ and $2\left(\frac{a'}{a}\right)h_k'$ act as friction terms. So if there was no friction term, the equations of motion are the same as those of a harmonic oscillator. The equation of motions cannot be solved exactly, but they can be solved in 2 important regimes. We previously saw that during inflation, there are two important length scales. One was the co-moving wavelength, λ_c,

$$\lambda_c = \frac{2\pi}{k} \tag{12.104}$$

of some perturbation. In the very early universe it was much smaller than the Hubble radius, $r_H(\equiv \frac{c}{H})$. When the universe accelerates during inflation, the Hubble radius is constant (or grows very slowly as $H = \frac{\dot{a}}{a}$ and $\dot{a}$ increases very rapidly as $\ddot{a}$ is very large). But the wavelength ($\propto a(t)$) grows exponentially. The equations of motion can solved exactly in the case where $\lambda_c >> r_H$ or $\lambda_c << r_H$. Let's start in the limit $\lambda_c << r_H$. The frictional terms are very weak compared to the harmonic terms. Therefore the equations of motion simplify to the equations of motion of an under-damped oscillator.

A good approximation of a solution is given by defining $\zeta(k) = \frac{v_k}{z}$, and $h(k) = \frac{w_k}{a}$, the equations of motion simplify to

$$\begin{aligned} v_k'' + \left(k^2 - \frac{z''}{z}\right)v_k = 0\\ w_k'' + \left(k^2 - \frac{a''}{a}\right)w_k = 0,\end{aligned} \tag{12.105}$$

in the regime when $k^2 >> \frac{z''}{z}$, therefore two equations simplify to,

$$\begin{aligned} v_k'' + k^2 v_k = 0\\ w_k'' + k^2 w_k = 0,\end{aligned} \tag{12.106}$$

with solutions

$$\begin{aligned} v &= A(k)_\pm e^{\pm ik\eta}\\ w &= B(k)_\pm e^{\pm ik\eta}.\end{aligned} \tag{12.107}$$

So in terms of the usual field ζ and h,

$$\begin{aligned} \zeta_k(\eta) &\approx \frac{A(k)_\pm}{z}e^{\pm ik\eta}\\ h_k^s(\eta) &\approx \frac{B(k)_\pm}{z}e^{\pm ik\eta}.\end{aligned} \tag{12.108}$$

These are not normalised. To normalise we need the Wronskians

$$W(\zeta_k, \zeta_k^\dagger) \equiv \zeta_k(\eta)\zeta_k^{*\prime}(\eta) - \zeta_k^*(\eta)\zeta_k'(\eta) = \frac{i}{z^2} \tag{12.109}$$

$$W(h_k, h_k^*) \equiv h_k(\eta)h_k^{*\prime}(\eta) - h_k^*(\eta)h_k'(\eta) = \frac{i}{(m_{pl}a)^2}. \tag{12.110}$$

The normalisation turns out to give,

$$\begin{aligned}\zeta_k(\eta) &\approx \frac{1}{z\sqrt{2k}}(A_1(k)e^{-ik\eta}+A_2(k)e^{ik\eta}),\\ h_k(\eta) &\approx \frac{1}{(m_p a)\sqrt{2k}}(B_1(k)e^{-ik\eta}+B_2(k)e^{ik\eta}).\end{aligned} \tag{12.111}$$

Now let's look at the case outside the horizon. The equations of motion are even simpler in this case as the k^2 term can be ignored and therefore the friction term halts the motion of the field very quickly. Once friction "freezes" this mode (i.e halts the motion), it remains the same for the remaining of the universe's lifetime and is what is observed today. So we want to calculate the value of the mode functions at the point of horizon crossing. The horizon crossing happens when

$$r_H = \frac{\lambda_p}{2\pi} = \frac{1}{H} \qquad (c \equiv 1), \tag{12.112}$$

where the 2π is inserted to get the correct value. λ_p is the physical wavelength which is $a(t)\lambda_c$, where λ_c is the co-moving wavelength, therefore $r_H = \frac{\lambda_c a(t)}{2\pi} = \frac{1}{H} \Rightarrow \frac{\lambda_c}{2\pi} = \frac{1}{a(t)H} \Rightarrow k = aH$, where k is the co-moving wave-vector. Recall that $z = \sqrt{2\epsilon}a$, so at the moment of horizon crossing we define $z = z_*$, $\epsilon = \epsilon_*$ and $a = a_*$, therefore at the moment of horizon crossing, we get

$$z_* = \sqrt{2\epsilon_*}\left(\frac{k}{H}\right). \tag{12.113}$$

In this case ζ_k becomes

$$\zeta_k = \frac{1}{\sqrt{2k}}\frac{1}{\sqrt{2\epsilon_*}}\left(\frac{H}{k}\right) \tag{12.114}$$

where the phase term has been ignored, as only the modulus square comes into the power spectrum (note that we are only using a positive or negative mode here, therefore there are no interference terms). Therefore we get

$$P_\zeta(k) = \left(\frac{k}{2\pi^2}\right)^2 \approx \frac{1}{2\epsilon}\left(\frac{H}{2\pi m_p}\right) \tag{12.115}$$

and analogously for tensor modes,

$$P_n(k) = 8\left(\frac{k}{2\pi(m_p a)}\right)^2 = 8\left(\frac{H}{2\pi m_{pl}}\right)^2. \tag{12.116}$$

During inflation, k is growing exponentially, while H is growing very slowly, which ultimately means the spectrum is independent of k, also called *scale-invariant* (i.e we have $\approx k^0$). A lot of different perturbations with different values of k cross the horizon at a given time as they are all expanding exponentially and H is basicly constant. So each wave-number that we look at, we are evaluating the value of the slow roll parameter ϵ and the value of H at the moment a mode with wave-vector k crosses the horizon, but during inflation $\approx e^{60}$ are crossing the horizon, while ϵ and H have barely had time to change at all, i.e as independent of k.

Bibliography

[1] Pedro G.Ferreira , *The Perfect Theory*, Little, Brown Publishing Company, London-U.K., 1st edition, 2014, ISBN 978-1-4087-0430-1

[2] A.Zee , *Einstein gravity in a nutshell*, Published by Princeton University Press, 41 William Street, Princeton, New Jersey 08540, 2014, ISBN 978-0-691-14558-7

[3] S.Weinberg , *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, Publication Date: July 1972 ISBN-13: 978-0471925675

[4] A. K. Raychaudhuri, Phys. Rev. 98, 1123 (1955).

[5] N. J. Poplawski, "Radial motion into the Einstein-Rosen bridge," Phys. Lett. B **687** (2010) 110 [arXiv:0902.1994 [gr-qc]].

[6] C. Brans and R. H. Dicke, "Mach's principle and a relativistic theory of gravitation," Phys. Rev. **124** (1961) 925

[7] S. W. Hawking, "Particle Creation by Black Holes," Commun. Math. Phys. **43** (1975) 199 [Erratum-ibid. **46** (1976) 206].

[8] Editor: Gary T. Horowitz , *Black Holes in Higher Dimensions* Cambridge University Press (19 April 2012), ISBN 9781107013452

[9] S. R. Coleman and F. De Luccia, "Gravitational Effects on and of Vacuum Decay," Phys. Rev. D **21** (1980) 3305.

[10] S. R. Coleman, "The Fate of the False Vacuum. 1. Semiclassical Theory," Phys. Rev. D **15** (1977) 2929 [Erratum-ibid. D **16** (1977) 1248].