General Relativity

Proff. Valeria Ferrari, Leonardo Gualtieri

AA 2011-2012
Contents

1 Introduction ................................................................. 1
  1.1 Non euclidean geometries ............................................ 1
  1.2 How does the metric tensor transform if we change the coordinate system .. 4
  1.3 Summary ............................................................... 6
  1.4 The Newtonian theory ................................................ 7
  1.5 The role of the Equivalence Principle in the formulation of the new theory of gravity ......................................................... 10
  1.6 The geodesic equations as a consequence of the Principle of Equivalence ... 11
  1.7 Summary ............................................................... 13
  1.8 Locally inertial frames ................................................ 13
  1.9 Appendix 1A ......................................................... 14
  1.10 Appendix 1B ......................................................... 15

2 Topological Spaces, Mapping, Manifolds .................................. 16
  2.1 Topological spaces .................................................... 16
  2.2 Mapping ............................................................... 18
  2.3 Composition of maps ................................................ 19
  2.4 Continuous mapping ................................................ 20
  2.5 Manifolds and differentiable manifolds ................................ 21

3 Vectors and One-forms ...................................................... 25
  3.1 The traditional definition of a vector .................................. 25
  3.2 A geometrical definition ............................................... 26
  3.3 The directional derivative along a curve form a vector space at P ............ 28
  3.4 Coordinate bases ...................................................... 31
  3.5 One-forms ............................................................. 34
  3.6 Vector fields and one-form fields ..................................... 39

4 Tensors ......................................................................... 41
  4.1 Geometrical definition of a Tensor .................................... 41
  4.2 Symmetries ............................................................ 46
  4.3 The metric Tensor ...................................................... 47
    4.3.1 The metric tensor allows to compute the distance between two points 50
    4.3.2 The metric tensor maps vectors into one-forms .................... 53
15.3 Diffeomorphisms on a manifold and Lie derivative .......................... 218
  15.3.1 The action of diffeomorphisms on functions and tensors ............. 219
  15.3.2 The action of diffeomorphisms on the metric tensor .................. 221
  15.3.3 Lie derivative ........................................................................ 221
  15.3.4 General Covariance and the role of diffeomorphisms .................. 223
15.4 Variational approach to General Relativity ................................. 223
  15.4.1 Action principle in special relativity ........................................ 223
  15.4.2 Action principle in general relativity ....................................... 224
  15.4.3 Variation of the Einstein-Hilbert action .................................... 228
  15.4.4 An example: the electromagnetic field ..................................... 229
  15.4.5 A comment on the definition of the stress-energy tensor ............ 230
Chapter 1

Introduction

General Relativity is the physical theory of gravity formulated by Einstein in 1916. It is based on the Equivalence Principle of Gravitation and Inertia, which establishes a fundamental connection between the gravitational field and the geometry of the spacetime, and on The Principle of General Covariance. General Relativity has changed quite dramatically our understanding of space and time, and the consequences of this theory, which we shall investigate in this course, disclose interesting and fascinating new phenomena, like for instance the existence of black holes and the generation of gravitational waves.

The language of General Relativity is that of tensor analysis, or, in a more modern formulation, the language of differential geometry. There is no way of understanding the theory of gravity without knowing what is a manifold, or a tensor. Therefore we shall dedicate a few lectures to the mathematical tools that are essential to describe the theory and its physical consequences. The first lecture, however, will be dedicated to answer the following questions:

1) why does the Newtonian theory become unappropriate to describe the gravitational field.

2) Why do we need a tensor to describe the gravitational field, and we why do we need to introduce the concept of manifold, metric, affine connections and other geometrical objects.

3) What is the role played by the equivalence principle in all that.

In the next lectures we shall rigorously define manifolds, vectors, tensors, and then, after introducing the principle of general covariance, we will formulate Einstein’s equations.

But first of all, since as we have already anticipated that there is a connection between the gravitational field and the geometry of the spacetime, let us introduce non-euclidean geometries, which are in some sense the precursors of general relativity.

1.1 Non euclidean geometries

In the prerelativistic years the arena of physical theories was the flat space of euclidean geometry which is based on the five Euclidean’s postulates. Among them the fifth has been the object of a millennia dispute: for over 2000 years geometers tried to show, without succeeding, that the fifth postulate is a consequence of the other four. The postulate states the following:
CHAPTER 1. INTRODUCTION

Consider two straight lines and a third straight line crossing the two. If the sum of the two internal angles (see figures) is smaller than $180^\circ$, the two lines will meet at some point on the side of the internal angles.

![Diagram of two lines with internal angles α and β, where α + β < 180°](image)

The solution to the problem is due to Gauss (1824, Germany), Bolyai (1832, Austria), and Lobachevski (1826, Russia), who independently discovered a geometry that satisfies all Euclidean’s postulates except the fifth. This geometry is what we may call, in modern terms, a two dimensional space of constant negative curvature. The analytic representation of this geometry was discovered by Felix Klein in 1870. He found that a point in this geometry is represented as a pair of real numbers $(x^1, x^2)$ with

\[(x^1)^2 + (x^2)^2 < 1, \quad (1.1)\]

and the distance between two points $x$ and $X$, $d(x, X)$, is defined as

\[d(x, X) = a \cosh^{-1} \left[ \frac{1 - x^1 X^1 - x^2 X^2}{\sqrt{1 - (x^1)^2 - (x^2)^2} \sqrt{1 - (X^1)^2 - (X^2)^2}} \right], \quad (1.2)\]

where $a$ is a lengthscale. This space is infinite, because

\[d(x, X) \to \infty\]

when

\[(X^1)^2 + (X^2)^2 \to 1.\]

The logical independence of Euclidean’s fifth postulate was thus established.

In 1827 Gauss published the *Disquisitiones generales circa superficies curvas*, where for the first time he distinguished the inner, or intrinsic properties of a surface from the outer, or extrinsic properties. The first are those properties that can be measured by somebody living on the surface. The second are those properties deriving from embedding the surface in a higher-dimensional space. Gauss realized that the fundamental inner property is the distance between two points, defined as the shortest path between them on the surface.

For example a cone or a cylinder have the same inner properties of a plane. The reason is that they can be obtained by a flat piece of paper suitably rolled, without distorting its metric relations, i.e. without stretching or tearing. This means that the distance between any two points on the surface is the same as it was in the original piece of paper, and parallel lines remain parallel. Thus the intrinsic geometry of a cylinder and of a cone is flat. This
is not true in the case of a sphere, since a sphere cannot be mapped onto a plane without distortions: the inner properties of a sphere are different from those of a plane. It should be stressed that the intrinsic geometry of a surface considers only the relations between points on the surface.

However, since a cylinder or a cone are “round” in one direction, we think they are curved surfaces. This is due to the fact that we consider them as 2-dimensional surfaces in a 3-dimensional space, and we intuitively compare the curvature of the lines which are on the surfaces with straight lines in the flat 3-dimensional space. Thus, the extrinsic curvature relies on the notion of higher dimensional space. In the following, we shall be concerned only with the intrinsic properties of surfaces.

The distance between two points can be defined in a variety of ways, and consequently we can construct different metric spaces. Following Gauss, we shall select those metric spaces for which, given any sufficiently small region of space, it is possible to choose a system of coordinates \((\xi^1, \xi^2)\) such that the distance between a point \(P = (\xi^1, \xi^2)\), and the point \(P' = (\xi^1 + d\xi^1, \xi^2 + d\xi^2)\) satisfies Pythagoras’ law

\[
ds^2 = (d\xi^1)^2 + (d\xi^2)^2.\tag{1.3}
\]

From now on, when we say the distance between two points, we mean the distance between two points that are infinitely close.

This property, i.e. the possibility of setting up a locally euclidean coordinate system, is a local property: it deals only with the inner metric relations for infinitesimal neighborhoods. Thus, unless the space is globally euclidean, the coordinates \((\xi^1, \xi^2)\) have only a local meaning. Let us now consider some other coordinate system \((x^1, x^2)\). How do we express the distance between two points? If we explicitly evaluate \(d\xi^1\) and \(d\xi^2\) in terms of the new coordinates we find

\[
\xi^1 = \xi^1(x^1, x^2) \quad \rightarrow \quad d\xi^1 = \frac{\partial \xi^1}{\partial x^1} dx^1 + \frac{\partial \xi^1}{\partial x^2} dx^2
\]

\[
\xi^2 = \xi^2(x^1, x^2) \quad \rightarrow \quad d\xi^2 = \frac{\partial \xi^2}{\partial x^1} dx^1 + \frac{\partial \xi^2}{\partial x^2} dx^2
\]

\[
ds^2 = \left[ \left( \frac{\partial \xi^1}{\partial x^1} \right)^2 + \left( \frac{\partial \xi^2}{\partial x^1} \right)^2 \right] (dx^1)^2 + \left[ \left( \frac{\partial \xi^1}{\partial x^2} \right)^2 + \left( \frac{\partial \xi^2}{\partial x^2} \right)^2 \right] (dx^2)^2
\]

\[+ 2 \left[ \left( \frac{\partial \xi^1}{\partial x^1} \right) \left( \frac{\partial \xi^1}{\partial x^2} \right) + \left( \frac{\partial \xi^2}{\partial x^1} \right) \left( \frac{\partial \xi^2}{\partial x^2} \right) \right] dx^1 dx^2\]

\[= g_{11}(dx^1)^2 + g_{22}(dx^2)^2 + 2g_{12}dx^1dx^2 = g_{\alpha\beta}dx^\alpha dx^\beta.
\]

In the last line of eq. (1.5) we have defined the following quantities:

\[
g_{11} = \left[ \left( \frac{\partial \xi^1}{\partial x^1} \right)^2 + \left( \frac{\partial \xi^1}{\partial x^2} \right)^2 \right]
\]

\[
g_{22} = \left[ \left( \frac{\partial \xi^2}{\partial x^1} \right)^2 + \left( \frac{\partial \xi^2}{\partial x^2} \right)^2 \right]
\]
CHAPTER 1. INTRODUCTION

\[ g_{12} = \left[ \left( \frac{\partial \xi^1}{\partial x^1} \right)^2 + \left( \frac{\partial \xi^2}{\partial x^1} \right)^2 \right] \]

namely, we have defined the metric tensor \( g_{\alpha\beta} \) i.e. the metric tensor is an object that allows us to compute the distance in any coordinate system. As it is clear from the preceding equations, \( g_{\alpha\beta} \) is a symmetric tensor, \( (g_{\alpha\beta} = g_{\beta\alpha}) \). In this way the notion of metric associated to a space, emerges in a natural way.

EINSTEIN’s CONVENTION

In writing the last line of eq. (1.5) we have adopted the convention that if there is a product of two quantities having the same index appearing once in the lower and once in the upper case (“dummy indices”), then summation is implied. For example, if the index \( \alpha \) takes the values 1 and 2

\[ v_\alpha V^\alpha = \sum_{i=1}^{2} v_i V_i = v_1 V^1 + v_2 V^2 \]  

(1.7)

We shall adopt this convention in the following.

EXAMPLE: HOW TO COMPUTE \( g_{\mu\nu} \)

Given the locally euclidean coordinate system \((\xi_1, \xi_2)\) let us introduce polar coordinates \((r, \theta) = (x_1, x_2)\). Then

\[ \xi_1 = r \cos \theta \quad \rightarrow \quad d\xi^1 = \cos \theta dr - r \sin \theta d\theta \]  

(1.8)

\[ \xi_2 = r \sin \theta \quad \rightarrow \quad d\xi^2 = \sin \theta dr + r \cos \theta d\theta \]  

(1.9)

\[ ds^2 = (d\xi^1)^2 + (d\xi^2)^2 = dr^2 + r^2 d\theta^2, \]  

(1.10)

and therefore

\[ g_{11} = 1, \quad g_{22} = r^2, \quad g_{12} = 0. \]  

(1.11)

1.2 How does the metric tensor transform if we change the coordinate system

We shall now see how the metric tensor transforms under an arbitrary coordinate transformation. Let us assume that we know \( g_{\alpha\beta} \) expressed in terms of the coordinate \((x^1, x^2)\), and we want to change the reference to a new system \((x'^1, x'^2)\). In section 1 we have shown that, for example, the component \( g_{11} \) is defined as (see eq. 1.7)

\[ g_{11} = \left[ \left( \frac{\partial \xi^1}{\partial x^1} \right)^2 + \left( \frac{\partial \xi^2}{\partial x^1} \right)^2 \right] \]  

(1.12)
where \( (\xi^1, \xi^2) \) are the coordinates of the locally euclidean reference frame, and \( (x^1, x^2) \) two arbitrary coordinates. If we now change from \( (x^1, x^2) \) to \( (x'^1, x'^2) \), where \( x^1 = x^1(x'^1, x'^2) \) and \( x^2 = x^2(x'^1, x'^2) \), the metric tensor in the new coordinate frame \( (x'^1, x'^2) \) will be

\[
g'_{11} = g_{1'1'} = \left[ (\frac{\partial \xi^1}{\partial x^1})^2 + (\frac{\partial \xi^2}{\partial x^2})^2 \right]^{1/2}
\]

This is the manner in which a tensor transforms under an arbitrary coordinate transformation

This point will be illustrated in more detail in following lectures.

In general we can write

\[
g'_{\alpha\beta} = g_{\mu\nu} \frac{\partial x^\mu}{\partial x'^\alpha} \frac{\partial x'^\nu}{\partial x^\beta}
\]

The metric determines the intrinsic properties of a metric space.

We now want to define a function of \( g_{\alpha\beta} \) and of its first and second derivatives, which depends on the inner properties of the surface, but does not depend on the particular coordinate system we choose. Gauss showed that in the case of two-dimensional surfaces this function can be determined, and it is called, after him, the gaussian curvature, defined as

\[
k(x^1, x^2) = \frac{1}{2g} \left[ 2\frac{\partial^2 g_{12}}{\partial x^1 \partial x^2} - \frac{\partial^2 g_{11}}{\partial x^2} - \frac{\partial^2 g_{22}}{\partial x^1} \right] 
\]

This concept can be generalized to a space of arbitrary dimensions.
CHAPTER 1. INTRODUCTION

\[ + \left( 2 \frac{\partial g_{12}}{\partial x^1} - \frac{\partial g_{11}}{\partial x^2} \right) \left( 2 \frac{\partial g_{12}}{\partial x^2} - \frac{\partial g_{22}}{\partial x^1} \right) \]

\[-\frac{g_{11}}{4g^2} \left( 2 \frac{\partial g_{12}}{\partial x^2} - \frac{\partial g_{11}}{\partial x^1} \right)^2 \]

where \( g \) is the determinant of the 2-metric \( g_{\alpha\beta} \)

\[ g = g_{11}g_{22} - g_{12}^2. \quad (1.18) \]

For example, given a spherical surface of radius \( a \), with metric \( ds^2 = a^2 d\theta^2 + a^2 \sin^2 \theta d\varphi^2 \) (polar coordinates) we find

\[ k = \frac{1}{a^2}; \quad (1.19) \]

no matter how we choose the coordinates to describe the spherical surface, we shall always find that the gaussian curvature has this value. For the Gauss-Bolyai-Lobachevski geometry where

\[ g_{11} = \frac{a^2 [1 - (x^2)^2]}{[1 - (x^1)^2 - (x^2)^2]^2}, \quad g_{22} = \frac{a^2 [1 - (x^1)^2]}{[1 - (x^1)^2 - (x^2)^2]^2}, \quad g_{12} = \frac{a^2 x^1 x^2}{[1 - (x^1)^2 - (x^2)^2]^2}, \quad (1.20) \]

we shall always find

\[ k = -\frac{1}{a^2}; \quad (1.21) \]

if the space is flat, the gaussian curvature is \( k = 0 \). If we choose a different coordinate system, \( g_{\alpha\beta}(x^1, x^2) \) will change but \( k \) will remain the same.

1.3 Summary

We have seen that it is possible to select a class of 2-dimensional spaces where it is possible to set up, in the neighborhoods of any point, a coordinate system \( (\xi^1, \xi^2) \) such that the distance between two close points is given by Pythagoras’ law. Then we have defined the metric tensor \( g_{\alpha\beta} \), which allows to compute the distance in an arbitrary coordinate system, and we have derived the law according to which \( g_{\alpha\beta} \) transforms when we change reference. Finally, we have seen that there exists a scalar quantity, the gaussian curvature, which expresses the inner properties of a surface: it is a function of \( g_{\alpha\beta} \) and of its first and second derivatives, and it is invariant under coordinate transformations.

These results can be extended to an arbitrary D-dimensional space. In particular, as we shall discuss in the following, we are interested in the case D=4, and we shall select those spaces, or better, those spacetimes, for which the distance is that prescribed by Special Relativity.

\[ ds^2 = -(d\xi^0)^2 + (d\xi^1)^2 + (d\xi^2)^2 + (d\xi^3)^2. \quad (1.22) \]

For the time being, let us only clarify the following point. In a D-dimensional space we need more than one function to describe the inner properties of a surface. Indeed, since \( g_{ij} \)
is symmetric, there are only \( D(D + 1)/2 \) independent components. In addition, we can choose \( D \) arbitrary coordinates, and impose \( D \) functional relations among them. Therefore the number of independent functions that describe the inner properties of the space will be

\[
C = \frac{D(D + 1)}{2} - D = \frac{D(D - 1)}{2}.
\]

If \( D=2 \), as we have seen, \( C=1 \). If \( D=4 \), \( C=6 \), therefore there will be 6 invariants to be defined for our 4-dimensional spacetime. The problem of finding these invariant quantities was studied by Riemann (1826-1866) and subsequently by Christoffel, LeviCivita, Ricci, Beltrami. We shall see in the following that Riemannian geometries play a crucial role in the description of the gravitational field.

1.4 The Newtonian theory

In this section we shall discuss why the Newtonian theory of gravity became unappropriate to correctly describe the gravitational field. The Newtonian theory of gravity was published in 1685 in the “Philosophiae Naturalis Principia Mathematica”, which contains an incredible variety of fundamental results and, among them, the cornerstones of classical physics:

1) *Newton’s law*

\[
\vec{F} = m \ddot{\vec{a}},
\]

2) *Newton’s law of gravitation*

\[
\vec{F}_G = m \vec{g},
\]

where

\[
\vec{g} = -\frac{G \sum_i M_i (\vec{r} - \vec{r}_i)}{r^3}
\]

depends on the position of the massive particle with respect to the other masses that generate the field, and it decreases as the inverse square of the distance \( g \sim \frac{1}{r^2} \). The two laws combined together clearly show that a body falls with an acceleration given by

\[
\ddot{a} = \left( \frac{m_G}{m_f} \right) \vec{g}.
\]

If \( \frac{m_G}{m_f} \) is a constant independent of the body, the acceleration is the same for every infalling body, and independent of their mass. Galileo (1564-1642) had already experimentally discovered that this is, indeed, true, and Newton itself tested the equivalence principle studying the motion of pendulum of different composition and equal length, finding no difference in their periods. The validity of the equivalence principle was the core of Newton’s arguments for the universality of his law of gravitation; indeed, after describing his experiments with different pendulum in the Principia he says:

*But, without all doubt, the nature of gravity towards the planets is the same as towards the earth.*

Since then a variety of experiments confirmed this crucial result. Among them Eotvos experiment in 1889 (accuracy of 1 part in \( 10^9 \)), Dicke experiment in 1964 (1 part in \( 10^{11} \)), Braginsky in 1972 (1 part in \( 10^{12} \)) and more recently the Lunar-Laser Ranging experiments
(1 part in $10^{13}$). All experiments up to our days confirm *The Principle of Equivalence of the gravitational and the inertial mass*. Now before describing why at a certain point the Newtonian theory fails to be a satisfactory description of gravity, let me briefly describe the reasons of its great success, that remained untouched for more than 200 years.

In the Principia, Newton formulates the universal law of gravitation, he develops the theory of lunar motion and tides and that of planetary motion around the Sun, which are the most elegant and accomplished descriptions of these phenomena.

After Newton, the law of gravitation was used to investigate in more detail the solar system; its application to the study of the perturbations of Uranus’ orbit around the Sun led, in 1846, Adams (England) and Le Verrier (France) to predict the existence of a new planet which was named Neptune. A few years later, the discovery of Neptune was a triumph of Newton’s theory of gravitation.

However, already in 1845 Le Verrier had observed anomalies in the motion of Mercury. He found that the perihelium precession of $35''/100$ years exceeded the value due to the perturbation introduced by the other planets predicted by Newton’s theory. In 1882 Newcomb confirmed this discrepancy, giving a higher value, of $43''/100$ year. In order to explain this effect, scientists developed models that predicted the existence of some interplanetary matter, and in 1896 Seelinger showed that an ellipsoidal distribution of matter surrounding the Sun could explain the observed precession.

We know today that these models were wrong, and that the reason for the exceedingly high precession of Mercury’s perihelium has a relativistic origin.

In any event, we can say that the Newtonian theory worked remarkably well to explain planetary motion, but already in 1845 the suspect that something did not work perfectly had some experimental evidence.

Let us turn now to a more philosophical aspect of the theory. The equations of Newtonian mechanics are invariant under Galileo’s transformations

$$
\begin{align*}
\vec{x}' &= R_0 \vec{x} + \vec{v} t + \vec{d}_0 \\
t' &= t + \tau
\end{align*}
$$

where $R_0$ is the orthogonal, constant matrix expressing how the second frame is rotated with respect to the first (its elements depend on the three Euler angles), $\vec{v}$ is the relative velocity of the two frames, and $\vec{d}_0$ the initial distance between the two origins. The ten parameters (3 Euler angles, 3 components for $\vec{v}$ and $\vec{d}$, + the time shift $\tau$ ) identify the Galileo group.

The invariance of the equations with respect to Galileo’s transformations implies the existence of inertial frames, where the laws of Mechanics hold. What then determines which frames are inertial frames? For Newton, the answer is that there exists an absolute space, and the result of the famous experiment of the rotating vessel is a proof of its existence:

1: inertial frames are those in uniform relative motion with respect to the absolute space.

---

1The vessel experiment: a vessel is filled with water and rotates with a given angular velocity about the symmetry axis. After some time the surface of the water assumes the typical shape of a paraboloid, being in equilibrium under the action of the gravity force, the centrifugal force and the fluid forces. Now suppose that the masses in the entire universe would rigidly rotate with respect to the vessel at the same angular
However this idea was rejected by Leibniz who claimed that there is no philosophical need for such a notion, and the debate on this issue continued during the next centuries. One of the major opponents was Mach, who argued that if the masses in the entire universe would rigidly rotate with respect to the vessel, the water surface would bend in exactly the same way as when the vessel was rotating with respect to them. This is because the inertia is a measure of the gravitational interaction between a body and the matter content of the rest of the Universe.

The problems I have described (the discrepancy in the advance of perihelium and the postulate absolute space) are however only small clouds: the Newtonian theory remains The theory of gravity until the end of the nineteenth century. The big storm approaches with the formulation of the theory of electrodynamics presented by Maxwell in 1864. Maxwell’s equations establish that the velocity of light is an universal constant. It was soon understood that these equations are not invariant under Galileo’s transformations; indeed, according to eqs. (1.28), if the velocity of light is \( c \) in a given coordinate frame, it cannot be \( c \) in a second frame moving with respect to the first with assigned velocity \( \vec{v} \). To justify this discrepancy, Maxwell formulated the hypothesis that light does not really propagate in vacuum: electromagnetic waves are carried by a medium, the \textit{luminiferous ether}; and the equations are invariant only with respect to a set of galilean inertial frames that are at rest with respect to the ether. However in 1887 Michelson and Morley showed that the velocity of light is the same, within 5km/s (today the accuracy is less than 1km/s), along the directions of the Earth’s orbital motion, and transverse to it. How this result can be justified? One possibility was to say the Earth is at rest with respect to the ether; but this hypothesis was totally unsatisfactory, since it would have been a coming back to an anthropocentric picture of the world. Another possibility was that the ether simply does not exist, and one has to accept the fact that the speed of light is the same in any direction, and whatever is the velocity of the source. This was of course the only reasonable explanation. But now the problem was to find the coordinate transformation with respect to which Maxwell’s equations are invariant. The problem was solved by Einstein in 1905; he showed that Galileo’s transformations have to be replaced by the Lorentz transformations

\[
x'^{\alpha} = L^\alpha_\gamma x^\gamma, \tag{1.29}
\]

where \( \gamma = (1 - \frac{v^2}{c^2})^{-\frac{1}{2}} \), and

\[
L^0_0 = \gamma, \quad L^0_j = L^i_j = \frac{\gamma}{c} v_j, \quad L^i_j = \delta^i_j + \frac{\gamma - 1}{v^2} v_i v_j, \quad i, j = 1, 3 \tag{1.30}
\]

and \( v^i \) are the components of the velocity of the boost.

As it was immediately realised, however, while Maxwell’s equations are invariant with respect to Lorentz transformations, Newton’s equations were not, and consequently one should face the problem of how to modify the equations of mechanics and gravity in such a way that they become invariant with respect to Lorentz transformations. It is at this point that Einstein made his fundamental observation.
1.5 The role of the Equivalence Principle in the formulation of the new theory of gravity

Let us consider the motion of a non relativistic particle moving in a constant gravitational field. Be $\vec{F}_k$ some other forces acting on the particle. According to Newtonian mechanics, the equation of motion are

$$ m_I \frac{d^2 \vec{x}}{dt^2} = m_G \vec{g} + \sum_k \vec{F}_k \tag{1.31} $$

Let us now jump on an elevator which is freely falling in the same gravitational field, i.e. let us make the following coordinate transformation

$$ \vec{x}' = \vec{x} - \frac{1}{2} \vec{g}t'^2, \quad t' = t. \tag{1.32} $$

In this new reference frame eq. (1.31) becomes

$$ m_I \left[ \frac{d^2 \vec{x}'}{dt'^2} + \vec{g}' \right] = m_G \vec{g} + \sum_k \vec{F}_k. \tag{1.33} $$

Since by the Equivalence Principle $m_I = m_G$, and since this is true for any particle, this equation becomes

$$ m_I \frac{d^2 \vec{x}'}{dt'^2} = \sum_k \vec{F}_k. \tag{1.34} $$

Let us compare eq. (1.31) and eq. (1.34). It is clear that that an observer $O'$ who is in the elevator, i.e. in free fall in the gravitational field, sees the same laws of physics as the initial observer $O$, but he does not feel the gravitational field. This result follows from the equivalence, experimentally tested, of the inertial and gravitational mass. If $m_I$ would be different from $m_G$, or better, if their ratio would not be constant and the same for all bodies, this would not be true, because we could not simplify the term in $\vec{g}$ in eq. (1.33)!

It is also apparent that if $\vec{g}$ would not be constant eq. (1.34) would contain additional terms containing the derivatives of $\vec{g}$. However, we can always consider an interval of time so short that $\vec{g}$ can be considered as constant and eq. (1.34) holds. Consider a particle at rest in this frame and no force $\vec{F}_k$ acting on it. Under this assumption, according to eq. (1.34) it will remain at rest forever. Therefore we can define this reference as a locally inertial frame. If the gravitational field is constant and uniform everywhere, the coordinate transformation (1.32) defines a locally inertial frame that covers the whole spacetime. If this is not the case, we can set up a locally inertial frame only in the neighborhood of any given point.

The points discussed above are crucial to the theory of gravity, and deserve a further explanation. Gravity is distinguished from all other forces because all bodies, given the same initial velocity, follow the same trajectory in a gravitational field, regardless of their internal constitution. This is not the case, for example, for electromagnetic forces, which act on charged but not on neutral bodies, and in any event the trajectories of charged particles depend on the ratio between charge and mass, which is not the same for all particles. Similarly, other forces, like the strong and weak interactions, affect different particles differently.
It is this distinctive feature of gravity that makes it possible to describe the effects of gravity in terms of curved geometry, as we shall see in the following.

Let us now state the Principle of Equivalence. There are two formulations:

**The strong Principle of Equivalence**

In an arbitrary gravitational field, at any given spacetime point, we can choose a locally inertial reference frame such that, in a sufficiently small region surrounding that point, all physical laws take the same form they would take in absence of gravity, namely the form prescribed by Special Relativity.

There is also a weaker version of this principle

**The weak Principle of Equivalence**

Same as before, but it refers to the laws of motion of freely falling bodies, instead of all physical laws.

The preceding formulations of the equivalence principle resembles very much to the axiom that Gauss chose as a basis for non-euclidean geometries, namely: at any given point in space, there exist a locally euclidean reference frame such that, in a sufficiently small region surrounding that point, the distance between two points is given by the law of Pythagoras.

The Equivalence Principle states that in a locally inertial frame all laws of physics must coincide, locally, with those of Special Relativity, and consequently in this frame the distance between two points must coincide with Minkowsky’s expression

\[ ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2 = -(d\xi^0)^2 + (d\xi^1)^2 + (d\xi^2)^2 + (d\xi^3)^2. \]  

We therefore expect that the equations of gravity will look very similar to those of Riemannian geometry. In particular, as Gauss defined the inner properties of curved surfaces in terms of the derivatives \( \frac{\partial \xi^\alpha}{\partial x^\mu} \) (which in turn defined the metric, see eqs. (1.5) and (1.7)), where \( \xi^\alpha \) are the “locally euclidean coordinates” and \( x^\mu \) are arbitrary coordinates, in a similar way we expect that the effects of a gravitational field will be described in terms of the derivatives \( \frac{\partial \xi^\alpha}{\partial x^\mu} \) where now \( \xi^\alpha \) are the “locally inertial coordinates”, and \( x^\mu \) are arbitrary coordinates. All this will follow from the equivalence principle. Up to now we have only established that, as a consequence of the Equivalence Principle there exist a connection between the gravitational field and the metric tensor. But which connection?

### 1.6 The geodesic equations as a consequence of the Principle of Equivalence

Let us start exploring what are the consequences of the Principle of Equivalence. We want to find the equations of motion of a particle that moves under the exclusive action of a gravitational field (i.e. it is in free fall), when this motion is observed in an arbitrary reference frame. We shall now work in a four-dimensional spacetime with coordinates \( (x^0 = ct, x^1, x^2, x^3) \).

First we start analysing the motion in a locally inertial frame, the one in free fall with the particle. According to the Principle of Equivalence, in this frame the distance between two neighboring points is

\[ ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2 = \eta_{\mu\nu}d\xi^\mu d\xi^\nu, \]  

(1.36)
CHAPTER 1. INTRODUCTION

where \( \eta_{\mu\nu} = \text{diag}(-1,1,1,1) \) is the metric tensor of the flat, Minkowsky spacetime. If \( \tau \) is the particle proper time, and if it is chosen as time coordinate, for what we said before the equations of motion are

\[
\frac{d^2 \xi^\alpha}{d\tau^2} = 0. \tag{1.37}
\]

We now change to a frame where the coordinates are labelled \( x^\alpha = x^\alpha(\xi^\alpha) \), i.e. we assign a transformation law which allows to express the new coordinates as functions of the old ones. In a following lecture we shall clarify and make rigorous all concepts that we are now using, such us metric tensor, coordinate transformations etc. In the new frame the distance is

\[
ds^2 = \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\nu} dx^\mu dx^\nu = g_{\mu\nu} dx^\mu dx^\nu, \tag{1.38}\]

where we have defined the metric tensor \( g_{\mu\nu} \) as

\[
g_{\mu\nu} = \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\nu} \eta_{\alpha\beta}. \tag{1.39}\]

This formula is the 4-dimensional generalization of the 2-dimensional gaussian formula (see eq. (1.5)). In the new frame the equation of motion of the particle (1.37) becomes:

\[
\frac{d^2 x^\alpha}{d\tau^2} + \Gamma^\alpha_{\mu\nu} \left[ \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \right] = 0. \tag{1.40}\]

(see the detailed calculations in appendix A). If we now define the following quantities

\[
\Gamma^\alpha_{\mu\nu} = \frac{\partial x^\alpha}{\partial \xi^\lambda} \frac{\partial^2 \xi^\lambda}{\partial x^\mu \partial x^\nu}, \tag{1.41}\]

eq (1.40) become

\[
\frac{d^2 x^\alpha}{d\tau^2} + \Gamma^\alpha_{\mu\nu} \left[ \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \right] = 0. \tag{1.42}\]

The quantities (1.41) are called the affine connections, or Christoffel’s symbols, the properties of which we shall investigate in a following lecture. Equation (1.42) is the geodesic equation, i.e. the equation of motion of a freely falling particle when observed in an arbitrary coordinate frame. Let us analyse this equation. We have seen that if we are in a locally inertial frame, where, by the Equivalence Principle, we are able to eliminate the gravitational force, the equations of motion would be that of a free particle (eq. 1.37). If we change to another frame we feel the gravitational field (and in addition all apparent forces like centrifugal, Coriolis, and dragging forces). In this new frame the geodesic equation becomes eq. (1.42) and the additional term

\[
\Gamma^\alpha_{\mu\nu} \left[ \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \right] \tag{1.43}\]

expresses the gravitational force per unit mass that acts on the particle. If we were in Newtonian mechanics, this term would be \( \vec{g} \) (plus the additional apparent accelerations,
but let us assume for the time being that we choose a frame where they vanish), and \( \vec{g} \) is the gradient of the gravitational potential. What does that mean? The affine connection \( \Gamma^\alpha_{\mu\nu} \) contains the second derivatives of \( (\xi^\alpha) \). Since the metric tensor (1.39) contains the first derivatives of \( (\xi^\alpha) \) (see eq. (1.39)), it is clear that \( \Gamma^\alpha_{\mu\nu} \) will contain first derivatives of \( g_{\mu\nu} \). This can be shown explicitly, and in a next lecture we will show that

\[
\Gamma^\sigma_{\lambda\mu} = \frac{1}{2} g^{\nu\sigma} \left\{ \frac{\partial g_{\mu\nu}}{\partial x^\lambda} + \frac{\partial g_{\nu\lambda}}{\partial x^\mu} - \frac{\partial g_{\lambda\mu}}{\partial x^\nu} \right\}. \tag{1.44}
\]

Thus, in analogy with the Newtonian law, we can say that the affine connections are the generalization of the Newtonian gravitational field, and that the metric tensor is the generalization of the Newtonian gravitational potential.

I would like to stress that this is a physical analogy, based on the study of the motion of freely falling particles compared with the Newtonian equations of motion.

### 1.7 Summary

We have seen that once we introduce the Principle of Equivalence, the notion of metric and affine connections emerge in a natural way to describe the effects of a gravitational field on the motion of falling bodies. It should be stressed that the metric tensor \( g_{\mu\nu} \) represents the gravitational potential, as it follows from the geodesic equations. But in addition it is a geometrical entity, since, through the notion of distance, it characterizes the spacetime geometry. This double role, physical and geometrical of the metric tensor, is a direct consequence of the Principle of Equivalence, as I hope it is now clear.

Now we can answer the question “why do we need a tensor to describe a gravitational field”: the answer is in the Equivalence Principle.

### 1.8 Locally inertial frames

We shall now show that if we know \( g_{\mu\nu} \) and \( \Gamma^\alpha_{\mu\nu} \) (i.e. \( g_{\mu\nu} \) and its first derivatives) at a point \( X \), we can determine a locally inertial frame \( \xi^\alpha(x) \) in the neighborhood of \( X \) in the following way. Multiply \( \Gamma^\beta_{\mu\nu} \) by \( \frac{\partial \xi^\beta}{\partial x^\mu} \)

\[
\frac{\partial \xi^\beta}{\partial x^\lambda} \Gamma^\lambda_{\mu\nu} = \frac{\partial \xi^\beta}{\partial x^\lambda} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial^2 \xi^\alpha}{\partial x^\nu \partial x^\nu} = \delta^\beta_\alpha \frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu}, \tag{1.45}
\]

i.e.

\[
\frac{\partial^2 \xi^\beta}{\partial x^\mu \partial x^\nu} = \frac{\partial \xi^\beta}{\partial x^\lambda} \Gamma^\lambda_{\mu\nu}. \tag{1.46}
\]
This equation can be solved by a series expansion near $X$

$$\xi^\beta (x) = \xi^\beta (X) + \left[ \frac{\partial \xi^\beta (x)}{\partial x^\lambda} \right]_{x = X} (x^\lambda - X^\lambda)$$

$$+ \frac{1}{2} \left[ \frac{\partial^2 \xi^\beta (x)}{\partial x^\lambda \partial x^\sigma} \right]_{x = X} (x^\lambda - X^\lambda) (x^\sigma - X^\sigma) + ...$$

$$= a^\beta + b^\beta_\lambda (x^\lambda - X^\lambda) + \frac{1}{2} b^\beta_\lambda \Gamma^\lambda_{\mu \nu} (x^\mu - X^\mu) (x^\nu - X^\nu) + ...$$

On the other hand we know by eq. (1.39) that

$$g^\mu_\nu (X) = \eta_{\alpha \beta} \frac{\partial \xi^\alpha (x)}{\partial x^\mu} \bigg|_{x = X} = \eta_{\alpha \beta} b^\alpha_\mu b^\beta_\nu,$$  

(1.48)

and from this equation we compute $b^\beta_\mu$. Thus, given $g^\mu_\nu$ and $\Gamma^\lambda_{\mu \nu}$ at a given point $X$ we can determine the local inertial frame to order $(x - X)^2$ by using eq. (1.47). This equation defines the coordinate system except for the ambiguity in the constants $a^\mu$. In addition we have still the freedom to make an inhomogeneous Lorentz transformation, and the new frame will still be locally inertial, as it is shown in appendix B.

### 1.9 Appendix 1A

Given the equation of motion of a free particle

$$\frac{d^2 \xi^\alpha}{dt^2} = 0,$$  

(A1)

let us make a coordinate transformation to an arbitrary system $x^\sigma$

$$\xi^\alpha = \xi^\alpha (x^\gamma), \quad \rightarrow \quad \frac{d \xi^\alpha}{dt} = \frac{\partial \xi^\alpha}{\partial x^\gamma} \frac{dx^\gamma}{dt}.$$  

(A2)

eq. (A1) becomes

$$\frac{d}{dt} \left( \frac{\partial \xi^\alpha}{\partial x^\gamma} \frac{dx^\gamma}{dt} \right) = \frac{d^2 x^\gamma}{dt^2} \frac{\partial \xi^\alpha}{\partial x^\gamma} + \frac{\partial^2 \xi^\alpha}{\partial x^\beta \partial x^\gamma} \frac{dx^\beta}{dt} \frac{dx^\gamma}{dt} = 0.$$  

(A3)

Multiply eq. (A3) by $\frac{dx^\sigma}{d\xi^\alpha}$ remembering that

$$\frac{\partial \xi^\alpha}{\partial x^\gamma} \frac{\partial x^\sigma}{\partial \xi^\alpha} = \frac{\partial x^\sigma}{\partial x^\gamma} = \delta^\sigma_\gamma,$$

where $\delta^\sigma_\gamma$ is the Kronecker symbol ($= 1$ if $\sigma = \gamma \quad 0$ otherwise), we find

$$\frac{d^2 x^\gamma}{dt^2} \delta^\sigma_\gamma + \frac{\partial x^\sigma}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x^\beta \partial x^\gamma} \frac{dx^\beta}{dt} \frac{dx^\gamma}{dt} = 0,$$  

(A4)

which finally becomes

$$\frac{d^2 x^\sigma}{dt^2} + \left[ \frac{\partial x^\sigma}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x^\beta \partial x^\gamma} \right] \frac{dx^\beta}{dt} \frac{dx^\gamma}{dt} = 0,$$  

(A5)

which is eq. (1.40).
1.10 Appendix 1B

Given a locally inertial frame $\xi^\alpha$

$$ds^2 = \eta_{\mu \nu} d\xi^\mu d\xi^\nu. \quad (B1)$$

let us consider the Lorentz transformation

$$\xi^i = L^i_{\ j} \xi^j', \quad (B2)$$

where

$$L_j^i = \delta_j^i + v_j \gamma \frac{1}{v^2} - \frac{1}{\gamma v_j}, \quad L_j^0 = \frac{\gamma v_j}{c}, \quad L_0^0 = \gamma, \quad \gamma = (1 - \frac{v^2}{c^2})^{-\frac{1}{2}}. \quad (B3)$$

The distance will now be

$$ds^2 = \eta_{\mu \nu} d\xi^\mu d\xi^\nu = \eta_{\mu \nu} \frac{\partial \xi^\mu}{\partial \xi^i} \frac{\partial \xi^\nu}{\partial \xi^j} d\xi^i d\xi^j', \quad (B4)$$

Since

$$\frac{\partial \xi^\mu}{\partial \xi^i} = L^\mu_{\ \beta} \delta^\beta_{\ \nu} = L^\mu_{\ \nu}, \quad (B5)$$

it follows that

$$ds^2 = \eta_{\mu \nu} L^\mu_{\ \nu} L^\nu_{\ j} d\xi^i d\xi^j'. \quad (B6)$$

Since $L_j^i$ is a Lorentz transformation,

$$\eta_{\mu \nu} L^\mu_{\ \nu} L^\nu_{\ j} = \eta_{4j},$$

consequently the new frame is still a locally inertial frame.

$$ds^2 = \eta_{\mu \nu} d\xi^\mu d\xi^\nu. \quad (B7)$$
Chapter 2

Topological Spaces, Mapping, Manifolds

In chapter 1 we have shown that the Principle of Equivalence allows to establish a relation between the metric tensor and the gravitational field. We used vectors and tensors, we made coordinate transformations, but we did not define the geometrical objects we were introducing, and we did not discuss whether we are entitled to use these notions. We shall now define in a more rigorous way what is the type of space we are working in, what is a coordinate transformation, a vector, a tensor. Then we shall introduce the metric tensor and the affine connections as geometrical objects and, after defining the covariant derivative, we shall finally be able to introduce the Riemann tensor. This work is preliminary to the derivation of Einstein’s equations.

2.1 Topological spaces

In general relativity we shall deal with topological spaces. The word topology has two distinct meanings: local topology (to which we are mainly interested), and global topology, which involves the study of the large scale features of a space.

Before introducing the general definition of a topological space, let us recall some properties of $\mathbb{R}^n$, which is a particular case of topological space; this will help us in the understanding of the general definition of topological spaces.

Given a point $y = (y^1, y^2, ..., y^n) \in \mathbb{R}^n$, a neighborhood of $y$ is the collection of points $x$ such that

$$|x - y| \equiv \sqrt{\sum_{i=1}^{n} (x^i - y^i)^2} < r,$$

(2.1)

where $r$ is a real number. (This is sometimes called an ‘open ball’).
A set of points $S \subseteq \mathbb{R}^n$ is open if every point $x \in S$ has a neighborhood entirely contained in $S$. This implies that an open set does not include the points on the boundary of the set. For instance, an open ball is an open set; a closed ball, defined by $|x - y| \leq r$, is not an open set, because the points of the boundary, i.e. $|x - y| = r$, do not admit a neighborhood contained in the set.

Intuitively we have an idea that this is a continuum space, namely that there are points of $\mathbb{R}^n$ arbitrarily close to any given point, that the line joining two points can be subdivided into arbitrarily many pieces which also join points of $\mathbb{R}^n$. A non continuous space is, for example, a lattice. A formal characterization of a continuum space is the Haussdorff criterion: any two points of a continuum space have neighborhoods which do not intersect.

The open sets of $\mathbb{R}^n$ satisfy the following properties:

1. if $O_1$ and $O_2$ are open sets, so it is their intersection.
2. the union of any collection (possibly infinite in number) of open sets is open.

Let us now consider a general set $T$. Furthermore, we consider a collection of subsets of $T$, say $O=\{O_i\}$, and call them open sets. We say that the couple $(T,O)$ formed by the set and the collection of subsets is a topological space if it satisfies the properties (1) and (2) above.

We remark that the space $T$ is not necessarily $\mathbb{R}^n$: it can be any kind of set; the only specification we give is the collection of subsets $O$, which are by definition the open sets, and that satisfy the properties (1), (2). In particular, in a topological space the notion of distance is a structure which has not been introduced: all definitions only require the notion of open sets.
2.2 Mapping

A map \( f \) from a space \( M \) to a space \( N \) is a rule which associates with an element \( x \) of \( M \), a unique element \( y = f(x) \) of \( N \).

\[ M \text{ and } N \text{ need not to be different. For example, the simplest maps are ordinary real-valued functions on } \mathbb{R} \]

\[ \text{EXAMPLE} \quad y = x^3, \quad x \in \mathbb{R}, \quad \text{and} \quad y \in \mathbb{R}. \quad (2.2) \]

In this case \( M \) and \( N \) coincide.

A map gives a unique \( f(x) \) for every \( x \), but not necessarily a unique \( x \) for every \( f(x) \).

\[ \text{EXAMPLE} \]

map \textbf{many to one} \quad \text{map \textbf{one to one}}

If \( f \) maps \( M \) to \( N \) then for any set \( S \) in \( M \) we have an \textit{image} in \( N \), i.e. the set \( T \) of all points mapped by \( f \) from \( S \) in \( N \).
Conversely the set $S$ is the inverse image of $T$

$$S = f^{-1}(T).$$  \hspace{1cm} (2.3)$$

Inverse mapping is possible only in the case of one-to-one mapping. The statement “$f$ maps $M$ to $N$” is indicated as

$$f : M \rightarrow N.$$ \hspace{1cm} (2.4)$$

$f$ maps a particular element $x \in M$ to $y \in N$ is indicated as

$$f : x \mid \rightarrow y$$ \hspace{1cm} (2.5)$$

the image of a point $x$ is $f(x)$.

### 2.3 Composition of maps

Given two maps $f : M \rightarrow N$ and $g : N \rightarrow P$, there exists a map $g \circ f$ that maps $M$ to $P$

$$g \circ f : M \rightarrow P.$$ \hspace{1cm} (2.6)$$

This means: take a point $x \in M$ and find the image $f(x) \in N$, then use $g$ to map this point to a point $g(f(x)) \in P$

**EXAMPLE**

\begin{align*}
  f &: x \mid \rightarrow y \quad y = x^3 \\
  g &: y \mid \rightarrow z \quad z = y^2 \\
  g \circ f &: x \mid \rightarrow z \quad z = x^6
\end{align*}

Map into: If a map is defined for all points of a manifold $M$, it is a mapping from $M$ into $N$.

Map onto: If, in addition, every point of $N$ has an inverse image (but not necessarily a unique one), it is a map from $M$ onto $N$.

**EXAMPLE**: be $N$ the unit open disc in $\mathbb{R}^2$, i.e. the set of all points in $\mathbb{R}^2$ such that the distance from the center is less than one, $d(0,x) < 1$. Be $M$ the surface of an emisphere $\theta < \frac{\pi}{2}$ belonging to the unit sphere.
There exists a one-to-one mapping $f$ from $M$ onto $N$.

### 2.4 Continuous mapping

A map $f : M \to N$ is continuous at $x \in M$ if any open set of $N$ containing $f(x)$ contains the image of an open set of $M$. $M$ and $N$ must be topological spaces, otherwise the notion of continuity has no meaning.

This definition is related to the familiar notion of continuous functions. Suppose that $f$ is a real-valued function of one real variable. That is $f$ is a map of $\mathbb{R}$ to $\mathbb{R}$

$$f : \mathbb{R} \to \mathbb{R}. \quad (2.8)$$

In the elementary calculus we say that $f$ is continuous at a point $x_0$ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - f(x_0)| < \epsilon, \quad \forall x \text{ such that } |x - x_0| < \delta. \quad (2.9)$$
Let us translate this definition in terms of open sets. From the figure it is apparent that any open set containing \( f(x_0) \), i.e. \( |f(x) - f(x_0)| < r \) with \( r \) arbitrary, contains an image of an open set of \( M \). This is true at least in the domain of definition of \( f \). This definition is more general than that of continuous functions, because it is based on the notion of open sets, and not on the notion of distance.

### 2.5 Manifolds and differentiable manifolds

The notion of manifold is crucial to define a coordinate system.

A **manifold** \( M \) is a topological space, which satisfies the Hausdorff criterion, and such that each point of \( M \) has an open neighborhood which has a continuous 1-1 map onto an open set of \( \mathbb{R}^n \). \( n \) is the dimension of the manifold.

In this definition we have used the concepts defined in the preceding pages: the space must be topological, continuous, and we want to associate an \( n \)-tuple of real numbers, i.e. a **set of coordinates** to each point. For example, when we consider the diagram

we are just using the notion of manifold: we take a point \( P \), and map it to the point \((x^1, y^1) \in \mathbb{R}^2\). And this operation can be done for any open neighborhood of \( P \). It should be stressed that the definition of manifold involves open sets and not the whole of \( M \) and \( \mathbb{R}^n \), because we do not want to restrict the global topology of \( M \). Moreover, at this stage we only require the map to be 1-1. We have not yet introduced any geometrical notion as
length, angles etc. At this level we only require that the local topology of $M$ is the same as that of $\mathbb{R}^n$. A manifold is a space with this topology.

**DEFINITION OF COORDINATE SYSTEMS**

A coordinate system, or a chart, is a pair consisting of an open set of $M$ and its map to an open set of $\mathbb{R}^n$. The open set does not necessarily include all $M$, thus there will be other open sets with the associated maps, and each point of $M$ must lie in at least one of such open sets.

AND NOW WE WANT TO MAKE A COORDINATE TRANSFORMATION.

Let us consider, for example, the following situation: $U$ and $V$ are two overlapping open sets of $M$ with two distinct maps onto $\mathbb{R}^n$.

$$f(U) \quad \quad g(V)$$

The overlapping region is open (since it is the intersection of two open sets), and is given two different coordinate systems by the two maps, thus there must exist some equation relating the two. We want to find it.

$$f(U) \quad \quad g(V)$$

Pick a point in the image of the overlapping region belonging to $f(U)$, say the point $(x^1, \ldots, x^n)$. The map $f$ has an inverse $f^{-1}$ which brings to the point $P$. Now from $P$, by using the map $g$, we go to the image of $P$ belonging to $g(V)$, i.e. to the point $(y^1, \ldots, y^n)$. 

$$\begin{align*}
    f^{-1} & : f(U) \rightarrow P \\
g & : P \rightarrow g(V)
\end{align*}$$
in $\mathbb{R}^n$
\[ g \circ f^{-1} : \mathbb{R}^n \to \mathbb{R}^n. \]  

The result of this operation is a functional relation between the two sets of coordinates:
\[
\begin{align*}
  y^1 &= y^1(x^1, \ldots, x^n) \\
  \\
  y^n &= y^n(x^1, \ldots, x^n),
\end{align*}
\]  

If the partial derivatives of order $\leq k$ of all the functions $\{y^i\}$ with respect to all $\{x^i\}$ exist and are continuous, then the charts $(U, f)$ and $(V, g)$ are said to be $C^k$ related. If it is possible to construct a system of charts such that each point of $M$ belongs at least to one of the open sets, and every chart is $C^k$ related to every other one it overlaps with, then the manifold is said to be a $C^k$ manifold. If $k=1$, it is called a differentiable manifold.

The notion of differentiable manifold is crucial, because it allows to add “structure” to the manifold, i.e. one can define vectors, tensors, differential forms, Lie derivatives etc.

In order to complete our definition of a coordinate transformation we still need another element. Eqs. (2.11) can be written as
\[ y^i = f^i(x^1, \ldots x^n), \quad i = 1, \ldots, n, \]  

where $f^i$ are $C^k$ differentiable. Be $J$ the jacobian of the transformation
\[ J = \frac{\partial(f^1, \ldots f^n)}{\partial(x^1, \ldots x^n)} = \begin{vmatrix} \frac{\partial f^1}{\partial x^1} & \frac{\partial f^1}{\partial x^2} & \cdots & \frac{\partial f^1}{\partial x^n} \\ \frac{\partial f^2}{\partial x^1} & \frac{\partial f^2}{\partial x^2} & \cdots & \frac{\partial f^2}{\partial x^n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f^n}{\partial x^1} & \frac{\partial f^n}{\partial x^2} & \cdots & \frac{\partial f^n}{\partial x^n} \end{vmatrix} \]  

If $J$ is non zero at some point $P$, then the inverse function theorem ensures that the map $f$ is 1-1 and onto in some neighborhood of $P$. If $J$ is zero at some point $P$ the transformation is singular.

AN EXAMPLE OF MANIFOLD.

Consider the 2-sphere (also called $S^2$). It is defined as the set of all points in $\mathbb{R}^3$ such that $(x^1)^2 + (x^2)^2 + (x^3)^2 = \text{const}$. Suppose that we want to map the whole sphere to $\mathbb{R}^2$ by using a single chart. For example let us use spherical coordinates $\theta \equiv x^1$, and $\varphi \equiv x^2$. The sphere appears to be mapped onto the rectangle $0 \leq x^1 \leq \pi, \quad 0 \leq x^2 \leq 2\pi$.
(note that this manifold has no boundary). But now consider the north pole $\theta = 0$: this point is mapped to the entire line

$$x^1 = 0, \quad 0 \leq x^2 \leq 2\pi.$$  \hfill (2.14)

Thus there is no map at all.

In addition all points of the emicircle $\varphi = 0$ are mapped in two places

$$x^2 = 0, \quad \text{and} \quad x^2 = 2\pi.$$  \hfill (2.15)

Again there is no map at all. In order to avoid these problems, we must restrict the map to open regions

$$0 < x^1 < \pi, \quad 0 < x^2 < 2\pi.$$  \hfill (2.16)

The two poles and the semicircle $\varphi = 0$ are left out. Then we may consider a second map, again in spherical coordinates but “rotated” in such a way that the line $\varphi = 0$ would coincide with the equator of the old system. Then every point of the sphere would be covered by one of the two charts, and in principle one should be able to find the coordinate transformation for the overlapping region. It is interesting to note that

1) this mapping does not preserve angles and lengths.

2) there exist manifolds that cannot be covered by a single chart, i.e. by a single coordinate system.
Chapter 3

Vectors and One-forms

3.1 The traditional definition of a vector

Let us consider an \(N\)-dimensional manifold, and a generic coordinate transformation

\[
x'^\alpha = x'^\alpha(x^\mu), \quad \alpha', \mu = 1, \ldots, N.
\]  

---

A comment on notation

Here and in the following, we shall use indices with and without primes to refer to different coordinate frames.

Strictly speaking, eq. (3.1) should be written as

\[
x'^\alpha = x'^\alpha(x^\mu), \quad \alpha', \mu = 1, \ldots, N,
\]  

because the coordinate with (say) \(\alpha' = 1\) belongs to the new frame, and is then different from the coordinate with \(\mu = 1\), belonging to the old frame. However, for brevity of notation, we will omit the primes in the coordinates, keeping only the primes in the indices.

---

A contravariant vector

\[ \vec{V} \rightarrow_0 \{V^\mu\}, \quad \mu = 1, 2, \ldots, N, \]  

where the symbol \(\rightarrow_0\) indicates that \(\vec{V}\) has components \(\{V^\mu\}\) with respect to a given frame \(O\), is a collection of \(N\) numbers which transform under the coordinate transformation (3.1) as follows:

\[
V'^\nu = \sum_{\alpha=1}^{N} \frac{\partial x'^\nu}{\partial x^\alpha} V^\alpha = \frac{\partial x'^\nu}{\partial x^\alpha} V^\alpha.
\]  

Notice that in writing the last term we have used Einstein’s convention. \(V'^\nu\) are the components of the vector in the new frame. If we now define the \(N \times N\) matrix

\[
(\Lambda^\nu_\beta) = \begin{pmatrix}
\frac{\partial x^\nu}{\partial x^\mu} & \frac{\partial x^\nu}{\partial x^x} & \cdots \\
\vdots & \ddots & \vdots \\
\frac{\partial x^N}{\partial x^\mu} & \frac{\partial x^N}{\partial x^x} & \cdots 
\end{pmatrix},
\]  

\[ (3.5) \]
the transformation law can be written in the general form
\[ V'^\alpha = \Lambda'^\alpha_\beta V^\beta. \]  
(3.6)
In addition, covariant vectors are defined as objects that transform according to the following rule
\[ A'_\mu = \frac{\partial x^\beta}{\partial x'^\mu} A_\beta = \Lambda^\beta_\mu A_\beta, \]  
(3.7)
where \( \Lambda^\beta_\mu \) is the inverse matrix of \( \Lambda_\beta^\mu \). However, a vector is a geometrical object. In fact it is an oriented segment that joins two points of a given space. We can associate to this object the components with respect to an assigned reference frame; when we change frame the vector components change, but the vector itself does not change. We shall now give a more adequate definition.

3.2 A geometrical definition

In order to define a vector as a geometrical object we need to introduce the notions of paths and curves.

**PATH**
A path is a connected series of points in the plane (or in any arbitrary \( N \)-dimensional manifold).

**CURVE**
A curve is a path with a real number associated with each point of the path, i.e. it is a mapping of an interval of \( \mathbb{R}^1 \) into a path in the plane (or in the \( N \)-dimensional manifold). The number is called the parameter. For example
\[ \text{curve} \ : \{ x^1 = f(s), x^2 = g(s), a \leq s \leq b \}, \]
(3.8)
means that each point of the path has coordinates that can be expressed as functions of \( s \). The path is called the image of the curve in the plane (or in the manifold). What happens if we change the parameter? If \( s' = s'(s) \) we shall get a new curve
\[ \{ x^1 = f'(s'), x^2 = g'(s'), a' \leq s' \leq b' \}, \]
(3.9)
where $f', g'$ are new functions of $s'$. This is a new curve, but with the same image. Thus there are an infinite number of curves corresponding to the same path.

FOR EXAMPLE: The position of a bullet shot by a gun in the 2-dimensional plane $(x,z)$ is a PATH; when we associate the parameter $t$ (time) at each point of the trajectory, we define a CURVE; if we change the parameter, say for instance the curvilinear abscissa, we define a new curve.

VECTORS
A vector is a geometrical object defined as the tangent vector to a given curve at a point $P$.
The set of numbers \( \{dx^i ds\} = \left(\frac{dx^1}{ds}, \frac{dx^2}{ds}\right) \) are the components of a vector tangent to the curve. (In fact if \( \{dx^i\} \) are infinitesimal displacements along the curve, dividing them by $ds$ only changes the scale but not the direction of the displacement). Every curve has a unique tangent vector

\[
\vec{V} \rightarrow \left\{\frac{dx^i}{ds}\right\}.
\]

One must be careful and not to confuse the curve with the path. In fact a path has, at any given point, an infinite number of tangent vectors, all parallel, but with different length. The length depends on the parameter $s$ that we choose to label the points of the path, and consequently it is different for different curves having the same image. A curve has a unique tangent vector, since the path and the parameter are given.

It should be reminded that a vector is tangent to an infinite number of different curves, for two different reasons. The first is that there are curves that are tangent to one another in $P$, and therefore have the same tangent vector:
CHAPTER 3. VECTORS AND ONE-FORMS

The second is that a path can be reparametrized in such a way that its tangent vector remains the same.

We shall now derive how does a vector transform if we change the coordinate system, and put for example \( x^1' = x^1(x^1, x^2), x^2' = x^2(x^1, x^2) \). The parameter \( s \) is unaffected, thus

\[
\begin{pmatrix}
\frac{dx^1'}{ds} = \frac{\partial x^1'}{\partial x^1} \frac{dx^1}{ds} + \frac{\partial x^1'}{\partial x^2} \frac{dx^2}{ds} \\
\frac{dx^2'}{ds} = \frac{\partial x^2'}{\partial x^1} \frac{dx^1}{ds} + \frac{\partial x^2'}{\partial x^2} \frac{dx^2}{ds}
\end{pmatrix} = \begin{pmatrix}
\frac{\partial x^1'}{\partial x^1} & \frac{\partial x^1'}{\partial x^2} \\
\frac{\partial x^2'}{\partial x^1} & \frac{\partial x^2'}{\partial x^2}
\end{pmatrix} \cdot \begin{pmatrix}
\frac{dx^1}{ds} \\
\frac{dx^2}{ds}
\end{pmatrix}
\]

As expected, this is the same transformation as (3.6) that was used to define a contravariant vector

\( V^\mu' = \Lambda^\mu_\beta V^\beta \). (3.11)

3.3 The directional derivative along a curve form a vector space at \( P \)

In order to understand the meaning of the statement contained in the heading of this section, let us consider a curve, parametrized with an assigned parameter \( \lambda \), and a differentiable function \( \Phi(x^1, \ldots x^N) \), in a general \( N \)-dimensional manifold. The directional derivative of \( \Phi \) along the curve will be

\[
\frac{d\Phi}{d\lambda} = \frac{\partial \Phi}{\partial x^1} \frac{dx^1}{d\lambda} + \ldots + \frac{\partial \Phi}{\partial x^N} \frac{dx^N}{d\lambda} = \frac{\partial \Phi}{\partial x^i} \frac{dx^i}{d\lambda}, \quad i = 1, \ldots, N.
\]

(3.12)

Since the function \( \Phi \) is totally arbitrary, we can rewrite this expression as

\[
\frac{dx^i}{d\lambda} = \frac{d}{d\lambda} \frac{\partial}{\partial x^i},
\]

(3.13)

where \( \frac{dx^i}{d\lambda} \) is now the operator of directional derivative, while \( \{ \frac{dx^i}{d\lambda} \} \) are the components of the tangent vector.

Let us consider two curves \( x^i = x^i(\lambda) \) and \( x^i = x^i(\mu) \) passing through the same point \( P \), and write the two directional derivatives along the two curves

\[
\frac{d}{d\lambda} = \frac{dx^i}{d\lambda} \frac{\partial}{\partial x^i}, \quad \frac{d}{d\mu} = \frac{dx^i}{d\mu} \frac{\partial}{\partial x^i}.
\]

(3.14)

\( \{ \frac{dx^i}{d\mu} \} \) are the components of the vector tangent to the second curve. Let us also consider a real number \( a \).

- We define the sum of the two directional derivatives as the directional derivative

\[
\frac{d}{d\lambda} + \frac{d}{d\mu} \equiv \left( \frac{dx^i}{d\lambda} + \frac{dx^i}{d\mu} \right) \frac{\partial}{\partial x^i}.
\]

(3.15)
The numbers \( \left( \frac{dx^i}{d\lambda} + \frac{dx^i}{d\mu} \right) \) are the components of a new vector, which is certainly tangent to some curve through P. Thus there must exist a curve with a parameter, say, \( s \), such that at P
\[
\frac{d}{ds} = \left( \frac{dx^i}{d\lambda} + \frac{dx^i}{d\mu} \right) \frac{\partial}{\partial x^i} = \frac{d}{d\lambda} + \frac{d}{d\mu}.
\] (3.16)

• We define the product of the directional derivative \( \partial/\partial\lambda \) with the real number \( a \) as the directional derivative
\[
a \frac{d}{d\lambda} = \left( a \frac{dx^i}{d\lambda} \right) \frac{\partial}{\partial x^i}.
\] (3.17)

The numbers \( \left( a \frac{dx^i}{d\lambda} \right) \) are the components of a new vector, which is certainly tangent to some curve through P. Thus there must exist a curve with a parameter, say, \( s' \), such that at P
\[
\frac{d}{ds'} = \left( a \frac{dx^i}{d\lambda} \right) \frac{\partial}{\partial x^i} = a \frac{d}{d\lambda}.
\] (3.18)

In this way we have defined two operations on the space of the directional derivatives along the curves passing through a point \( P \): the sum of two directional derivatives, and the multiplication of a directional derivative with a real number.

We remind the mathematical definition of a vector space.

A vector space is a set \( V \) on which two operations are defined:

1. Vector addition
\[
(\vec{v}, \vec{w}) \rightarrow \vec{v} + \vec{w}
\] (3.19)

2. Multiplication by a real number:
\[
(a, \vec{v}) \rightarrow a\vec{v}
\] (3.20)

(where \( \vec{v}, \vec{w} \in V, a \in \mathbb{R} \)), which satisfy the following properties:

• Associativity and commutativity of vector addition
\[
\vec{v} + (\vec{w} + \vec{u}) = (\vec{v} + \vec{w}) + \vec{u}
\]
\[
\vec{v} + \vec{w} = \vec{w} + \vec{v}.
\] \( \forall \vec{v}, \vec{w}, \vec{u} \in V \).

• Existence of a zero vector, i.e. of an element \( \vec{0} \in V \) such that
\[
\vec{v} + \vec{0} = \vec{v} \quad \forall \vec{v} \in V.
\]

• Existence of the opposite element: for every \( \vec{w} \in V \) there exists an element \( \vec{v} \in V \) such that
\[
\vec{v} + \vec{w} = \vec{0}.
\]

\(^1\)To be precise, what we are defining here is a real vector space, but we will omit this specification, because in this book only real vector spaces will be considered.
• Associativity and distributivity of multiplication by real numbers:

\[
\begin{align*}
  a(b\vec{v}) &= (ab)\vec{v} \\
  a(\vec{v} + \vec{w}) &= a\vec{v} + a\vec{w} \\
  (a + b)\vec{v} &= a\vec{v} + b\vec{v} \quad \forall \vec{v} \in V, \forall a, b \in \mathbb{R}.
\end{align*}
\] (3.22)

• Finally, the real number 1 must act as an identity on vectors:

\[
1 \vec{v} = \vec{v} \quad \forall \vec{v}.
\] (3.23)

Coming back to directional derivatives (taken at a given point \(P\) of the manifold), it is easy to verify that the operations of addition and multiplication by a real number defined in (3.15),(3.17) respectively, satisfy the above properties. For instance:

• Commutativity of the addition:

\[
\frac{d}{d\lambda} + \frac{d}{d\mu} = \left(\frac{dx^i}{d\lambda} + \frac{dx^i}{d\mu}\right) \frac{\partial}{\partial x^i} = \left(\frac{dx^i}{d\mu} + \frac{dx^i}{d\lambda}\right) \frac{\partial}{\partial x^i} = \frac{d}{d\mu} + \frac{d}{d\lambda}.
\] (3.24)

• Associativity of multiplication by real numbers:

\[
\begin{align*}
  a \left( b \frac{d}{d\lambda} \right) &= a \left( b \frac{dx^i}{d\lambda} \right) \frac{\partial}{\partial x^i} \\
  &= \left( a b \frac{dx^i}{d\lambda} \right) \frac{\partial}{\partial x^i} = \left( ab \frac{dx^i}{d\lambda} \right) \frac{\partial}{\partial x^i} \\
  &= ab \frac{d}{d\lambda}.
\end{align*}
\] (3.25)

• Distributivity of multiplication by real numbers:

\[
\begin{align*}
  a \left( \frac{d}{d\lambda} + \frac{d}{d\mu} \right) &= a \left( \frac{dx^i}{d\lambda} + \frac{dx^i}{d\mu} \right) \frac{\partial}{\partial x^i} \\
  &= \left( a \frac{dx^i}{d\lambda} + a \frac{dx^i}{d\mu} \right) \frac{\partial}{\partial x^i} = \left( a \frac{dx^i}{d\lambda} + \frac{dx^i}{d\mu} \right) \frac{\partial}{\partial x^i} \\
  &= \left( a \frac{dx^i}{d\lambda} \right) \frac{\partial}{\partial x^i} + \left( a \frac{dx^i}{d\mu} \right) \frac{\partial}{\partial x^i} = a \frac{d}{d\lambda} + a \frac{d}{d\mu}.
\end{align*}
\] (3.26)

• The zero element is the vector tangent to the curve \(x^\mu \equiv \text{const.}\), which is simply the point \(P\).

• The opposite of the vector \(\vec{v}\) tangent to a given curve is obtained by changing sign to the parametrization

\[
\lambda \rightarrow -\lambda.
\] (3.27)
The proof of the remaining properties is analogous. **Therefore, the set of directional derivatives is a vector space.**

In any coordinate system there are special curves, the coordinates lines (think for example to the grid of cartesian coordinates). The directional derivatives along these lines are

\[ \frac{d}{dx^i} = \frac{dx^k}{dx^i} \frac{\partial}{\partial x^k} = \delta^k_i \frac{\partial}{\partial x^k} = \frac{\partial}{\partial x^i}, \]

Eq. (3.13) shows that the generic directional derivative \( \frac{d}{d\lambda} \) can always be expressed as a linear combination of \( \frac{\partial}{\partial x^i} \). It follows that \( \frac{d}{dx^i} \equiv \frac{\partial}{\partial x^i} \) are a basis for this vector space, and \( \{ \frac{dx^i}{d\lambda} \} \) are the components of \( \frac{d}{d\lambda} \) on this basis. But \( \{ \frac{dx^i}{d\lambda} \} \) are also the components of a tangent vector at \( P \). Therefore the space of all tangent vectors and the space of all derivatives along curves at \( P \) are in 1-1 correspondence. For this reason we can say that \( \frac{d}{d\lambda} \) is the vector tangent to the curve \( x^i(\lambda) \).

TO SUMMARIZE: the vectors tangent to the coordinate lines in a point \( P \), i.e. the directional derivatives in \( P \) along these lines in a coordinate system \( (x^1, \ldots x^N) \), have the following components

\[ \frac{\partial}{\partial x^1} = (1, 0, \ldots 0), \quad \frac{\partial}{\partial x^2} = (0, 1, \ldots 0), \quad \ldots, \quad \frac{\partial}{\partial x^N} = (0, 0, \ldots 1). \]

If we use the \( \{ \frac{\partial}{\partial x^i} \} \) as a basis for vectors, the vector \( \frac{d}{d\lambda} \), tangent to the curve \( x^i(\lambda) \), with respect to this basis has components \( \{ \frac{dx^i}{d\lambda} \} \).

**Vectors do not lie in \( M \), but in the tangent space to \( M \), called \( T_P \).** For example in the two-dimensional case analysed above the tangent plane was the plane itself, but if the manifold is a sphere, since we cannot define a vector as an “arrow” on the sphere, we need to define the tangent space, i.e. the plane tangent to the sphere at each point. For more general manifolds it is not easy to visualize \( T_P \). In any event \( T_P \) has the same dimensions as the manifold \( M \).

### 3.4 Coordinate bases

Any collection of \( n \) linearly independent vectors of \( T_P \) is a basis for \( T_P \). However, a natural basis is provided by the vectors that are tangent to the coordinate lines, i.e. \( \{ e_{(i)} \} \equiv \{ \frac{\partial}{\partial x^{(i)}} \} \); this is the coordinate basis.

IMPORTANT:

To hereafter, we shall enclose within () the indices that indicate which vector of the basis we are choosing, not to be confused with the index which indicates the vector components. For instance \( e_1^{(2)} \) indicates the component 1 of the basis vector \( e_{(2)} \).
Any vector \( \vec{A} \) at a point \( P \), can be expressed as a linear combination of the basis vectors

\[
\vec{A} = A^i \vec{e}_i,
\]

(Remember Einstein’s convention: \( \sum_i A^i \vec{e}_i \equiv A^i \vec{e}_i \)) where the numbers \( \{ A^i \} \) are the components of \( \vec{A} \) with respect to the chosen basis.

If we make a coordinate transformation to a new set of coordinates \( (x^1', x^2', ..., x^n') \), there will be a new coordinate basis: \( \{ \vec{e}_i' \} \equiv \left\{ \frac{\partial}{\partial x^i} \right\} \).

We now want to find the relation between the new and the old basis, i.e., we want to express each new vector \( \vec{e}_i' \) as a linear combination of the old ones \( \{ \vec{e}_i \} \). In the new basis, the vector \( \vec{A} \) will be written as

\[
\vec{A} = A^j' \vec{e}_j',
\]

where \( \{ A^j' \} \) are the components of \( \vec{A} \) with respect to the basis \( \{ \vec{e}_j' \} \). But the vector \( \vec{A} \) is the same in any basis, therefore

\[
A^i' \vec{e}_i = A^j' \vec{e}_j'.
\]

From eq. (3.11) we know how to express \( A^j' \) as functions of the components in the old basis, and substituting these expressions into eq. (3.30) we find

\[
A^i' \vec{e}_i = \Lambda^j_k A^k \vec{e}_j'.
\]

By relabelling the dummy indices this equation can be written as

\[
\left[ \Lambda^j_k \vec{e}_j' - \vec{e}_k \right] A^k = 0,
\]

i.e.,

\[
\vec{e}_k = \Lambda^j_k \vec{e}_j'.
\]
Multiplying both members by $\Lambda^{k'}_j$ and remembering that
\[
\Lambda^{k'}_j \Lambda^i_k = \frac{\partial x^k}{\partial x'^j} \frac{\partial x'^j}{\partial x^i} = \delta^i_j
\]  
we find the transformation we were looking for
\[
\tilde{e}(j') = \Lambda^{k'}_j \tilde{e}(k). \tag{3.35}
\]
Summarizing:
\[
\begin{align*}
\tilde{e}(k) &= \Lambda^{i'}_k \tilde{e}(i') , \\
\tilde{e}(i') &= \Lambda^k_{i'} \tilde{e}(k).
\end{align*}
\]  
We are now in a position to compute the new basis vectors in terms of the old ones.

**EXAMPLE**

Consider the 4-dimensional flat spacetime of Special Relativity, but let us restrict to the (x-y) plane, where we choose the coordinates $(c t, x, y) \equiv (x^0, x^1, x^2)$. The coordinate basis is the set of vectors
\[
\begin{align*}
\vec{∂}_{x^0} &= \tilde{e}(0) = (1, 0, 0) , \\
\vec{∂}_{x^1} &= \tilde{e}(1) = (0, 1, 0) , \\
\vec{∂}_{x^2} &= \tilde{e}(2) = (0, 0, 1),
\end{align*}
\]  
or, in a compact form
\[
\tilde{e}^\beta_\alpha = \delta^\beta_\alpha. \tag{3.38}
\]
(The superscript $\beta$ now indicates the $\beta$-component of the $\alpha$-th vector). In this basis any vector $\vec{A}$ can be written as
\[
\vec{A} = A^0 \tilde{e}(0) + A^1 \tilde{e}(1) + A^2 \tilde{e}(2) = A^\alpha \tilde{e}_\alpha, \quad \alpha = 0, 1, 2 \tag{3.39}
\]
where $\{A^\alpha\} = (A^0, A^1, A^2)$ are the components of $\vec{A}$ with respect to this basis. Let us consider the following coordinate transformation
\[
\begin{align*}
(x^0, x, y) &\rightarrow (x^0, r, \theta) \\
x^0 &= x^0' \\
x^1 &= r \cos \theta \\
x^2 &= r \sin \theta,
\end{align*}
\]  
i.e. $x^1' = r, x^2' = \theta$. The new coordinate basis is
\[
\begin{align*}
\vec{∂}_{x^0'} &= \tilde{e}(0'), \\
\vec{∂}_r &= \tilde{e}(1'), \\
\vec{∂}_\theta &= \tilde{e}(2'),
\end{align*}
\]  
From eq. (3.35) we find
\[
\tilde{e}(\alpha') = \Lambda^\alpha_\alpha \tilde{e}(\alpha), \quad \Lambda^\alpha_\alpha = \frac{\partial x^\alpha}{\partial x'^\alpha}. \tag{3.42}
\]
In the example we are considering only $\Lambda^0_{0'} \neq 0$ and it is equal to 1. It follows that

$$\vec{e}_{(0')} \equiv \frac{\partial}{\partial x^{0'}} = \vec{e}_{(0')}.$$ (3.43)

In addition

$$\vec{e}_{(1')} = \Lambda^\alpha_1 \vec{e}_{(\alpha)},$$ (3.44)

and since

$$\Lambda^0_1 = \frac{\partial x^0}{\partial r} = 0, \quad \Lambda^1_1 = \frac{\partial x^1}{\partial r} = \cos \theta, \quad \Lambda^2_1 = \frac{\partial x^2}{\partial r} = \sin \theta,$$ (3.45)

$$\vec{e}_{(1')} \equiv \frac{\partial}{\partial r} = \cos \theta \vec{e}_{(1)} + \sin \theta \vec{e}_{(2)}.$$ (3.46)

Similarly

$$\vec{e}_{(2')} = \Lambda^\alpha_2 \vec{e}_{(\alpha)},$$ (3.47)

and since

$$\Lambda^0_2 = 0, \quad \Lambda^1_2 = \frac{\partial x^1}{\partial \theta} = -r \sin \theta, \quad \Lambda^2_2 = \frac{\partial x^2}{\partial \theta} = r \cos \theta,$$ (3.48)

hence

$$\vec{e}_{(2')} \equiv \frac{\partial}{\partial \theta} = -r \sin \theta \vec{e}_{(1)} + r \cos \theta \vec{e}_{(2)}.$$ (3.49)

Summarizing,

$$\begin{cases}
\vec{e}_{(0')} = \vec{e}_{(0)} \\
\vec{e}_{(1')} = \vec{e}_{(r)} = \cos \theta \vec{e}_{(1)} + \sin \theta \vec{e}_{(2)} \\
\vec{e}_{(2')} = \vec{e}_{(\theta)} = -r \sin \theta \vec{e}_{(1)} + r \cos \theta \vec{e}_{(2)}.
\end{cases}$$ (3.50)

It should be noted that we do not need to choose necessarily a coordinate basis. We may choose a set of independent basis vectors that are not tangent to the coordinate lines. In this case the matrix which allows to transform from one basis to another has to be assigned and will not be $\Lambda^\alpha_\beta'$ as in eq. (3.35).

### 3.5 One-forms

A one-form is a linear, real valued function of vectors. This means the following: a one-form (or 1-form) $\tilde{q}$ at the point $P$ takes the vector $\vec{V}$ at $P$ and associates a number to it, which we call $\tilde{q}(\vec{V})$. To hereafter a “~” will indicate 1-forms, as an arrow “→” indicates vectors.

By definition, a one-form is linear. This means that, for every couple of vectors $\vec{V}, \vec{W}$, for every couple of real numbers $a, b$, for every one-form $\tilde{q}$,

$$\tilde{q}(a\vec{V} + b\vec{W}) = a\tilde{q}(\vec{V}) + b\tilde{q}(\vec{W}).$$ (3.51)

We define two operations acting on the space of one-forms:
CHAPTER 3. VECTORS AND ONE-FORMS

• Multiplication by real numbers: given a one-form \( \tilde{q} \) and a real number \( a \), we define the new one-form \( a\tilde{q} \) such that, for every vector \( \vec{V} \),

\[
(a\tilde{q})(\vec{V}) = a[\tilde{q}(\vec{V})].
\]

(3.52)

• Addition: given two one-forms \( \tilde{q}, \tilde{\sigma} \), we define the new one-form \( \tilde{q} + \tilde{\sigma} \) such that, for every vector \( \vec{V} \),

\[
[\tilde{q} + \tilde{\sigma}](\vec{V}) = \tilde{q}(\vec{V}) + \tilde{\sigma}(\vec{V}).
\]

(3.53)

One-forms satisfy the axioms (3.21-3.23). Let us show this for some of the axioms.

• Commutativity of addition. Given two one-forms \( \tilde{q}, \tilde{\sigma} \), we have that, for every vector field \( \vec{V} \),

\[
(\tilde{q} + \tilde{\sigma})(\vec{V}) = \tilde{q}(\vec{V}) + \tilde{\sigma}(\vec{V}) = (\tilde{\sigma} + \tilde{q})(\vec{V}).
\]

(3.54)

• Distributivity of multiplication with real numbers. Given two one-forms \( \tilde{q}, \tilde{\sigma} \) and a real number \( a \), we have that, for every vector field \( \vec{V} \),

\[
[a(\tilde{q} + \tilde{\sigma})](\vec{V}) = a[\tilde{q}(\vec{V}) + \tilde{\sigma}(\vec{V})] = a\tilde{q}(\vec{V}) + a\tilde{\sigma}(\vec{V}) = a\tilde{q}(\vec{V}) + (a\tilde{\sigma})(\vec{V}) = [(a\tilde{q}) + (a\tilde{\sigma})](\vec{V}).
\]

(3.55)

then, being this true for every \( \vec{V} \),

\[
a(\tilde{q} + \tilde{\sigma}) = (a\tilde{q}) + (a\tilde{\sigma}).
\]

(3.56)

• Existence of the zero element. The zero one-form \( \tilde{0} \) is the one-form such that, for every \( \vec{V} \),

\[
\tilde{0}(\vec{V}) = 0.
\]

(3.57)

The other axioms can be proved in a similar way.

Therefore, one-forms form a vector space, which is called the dual vector space to \( T_p \), and it is indicated as \( T_p^* \); this is also called the cotangent space in \( P \).

\( T_p^* \) is the space of the maps (the 1-forms) that associate to any given vector a number, i.e. that map \( T_p \) on \( \mathbb{R} \). The reason why \( T_p^* \) is called dual to \( T_p \) is that vectors also can be regarded as linear, real valued functions of one-forms: a vector \( \vec{V} \) takes a 1-form \( \tilde{q} \) and associates a number to it, which we call \( \vec{V}(\tilde{q}) \), and

\[
\tilde{q}(\vec{V}) \equiv \vec{V}(\tilde{q}),
\]

(3.58)
in the sense that the two “operations” give as a result the same number. This point will be further clarified in the following. Once we choose a basis for vectors, say \( \{\vec{e}_i\} \), we can introduce a dual basis for one-forms defined as follows: the dual basis \( \{\tilde{\omega}^{(i)} \}, i = 1, \ldots, N \), takes any vector \( \vec{V} \) in \( T_p \) and produces its components

\[
\tilde{\omega}^{(i)}(\vec{V}) = V^i.
\]

(3.59)
It should be remembered that an index in parenthesis does not refer to a component, but selects the \( i \)-th one-form (or vector) of the basis. Thus the \( i \)-th basis one-form applied to \( \vec{V} \) gives as a result a number, which is the component \( V^i \) of the vector \( \vec{V} \). As expected, this operation is linear in the argument

\[
\tilde{\omega}^{(i)}(\vec{V} + \vec{W}) = V^i + W^i, \tag{3.60}
\]

since \( \vec{V} + \vec{W} \) is a vector whose \( i \)-th component is \( V^i + W^i \). In particular, if the argument of a one-form is \( \vec{e}^{(j)} \), i.e. one of the basis vectors of the tangent space at the point \( P \), since only the \( j \)-th component of \( \vec{e}^{(j)} \) is different from zero and equal to 1, we have

\[
\tilde{\omega}^{(i)}(\vec{e}^{(j)}) = \delta^i_j. \tag{3.61}
\]

We now want to answer the questions:

1. Who tells us that \( \{\tilde{\omega}^{(i)}\} \) form a basis for one-forms?
2. Can we define the components of a 1-form as we define the components of a vector?

---

1. Consider any one-form \( \tilde{q} \) acting on an arbitrary vector \( \vec{V} \). By expressing \( \vec{V} \) as a linear combination of the basis vectors \( \vec{e}^{(j)} \), and using the linearity of one-forms we can write

\[
\tilde{q}(\vec{V}) = \tilde{q}(V^j \vec{e}^{(j)}) = V^j \tilde{q}(\vec{e}^{(j)}) = \tilde{\omega}^{(j)}(\vec{V}) \tilde{q}(\vec{e}^{(j)}), \tag{3.62}
\]

where the last equality follows from eq.(3.59). This equation holds for any vector \( \vec{V} \) therefore we can write

\[
\tilde{q} = \tilde{\omega}^{(j)} \tilde{q}(\vec{e}^{(j)}); \tag{3.63}
\]

since \( \tilde{q}(\vec{e}^{(j)}) \) are real numbers, this equation shows that any one-form \( \tilde{q} \) can be written as a linear combination of the \( \{\tilde{\omega}^{(j)}\} \); consequently \( \{\tilde{\omega}^{(j)}\} \) form a basis for one-forms.

2. We now define the components of \( \tilde{q} \) on the basis \( \{\tilde{\omega}^{(i)}\} \) as

\[
q_j = \tilde{q}(\vec{e}^{(j)}), \tag{3.64}
\]

and consequently we can write

\[
\tilde{q} = q_j \tilde{\omega}^{(j)}. \tag{3.65}
\]

Consider an open region \( U \) of the manifold \( M \), and choose a coordinate system \( \{x^i\} \). We have seen that this defines a natural coordinate basis for vectors \( \vec{e}^{(i)} \equiv \{\partial / \partial x^{(i)}\} \). Furthermore, it also defines a natural coordinate basis for one-forms (dual to the natural basis for vectors), often indicated as \( \{dx^{(i)}\} \), whose components are

\[
\tilde{\omega}^{(i)} \equiv \tilde{dx}^{(i)} = \tilde{d}x^{(i)} \left( \frac{\partial}{\partial x^{(j)}} \right) = \delta^i_j. \]
And now the most important thing. From eq. (3.65) it follows that for any vector $\vec{V}$

$$\tilde{q}(\vec{V}) = q_j \tilde{\omega}^{(j)}(\vec{V}).$$  \hspace{1cm} (3.66)

Since $\tilde{\omega}^{(j)}(\vec{V}) = V^j$, we find

$$\tilde{q}(\vec{V}) = q_j V^j.$$ \hspace{1cm} (3.67)

This operation is called **contraction** and tells us how to compute the number which results from the application of $\tilde{q}$ on $\vec{V}$ (or vice versa), once we know the components of $\tilde{q}$ and $\vec{V}$.

From eq. (3.67) we can now better understand why vectors and one-forms are dual of each other. In fact, if $q_j$ and $V^j$ are respectively the components of the one-form $\tilde{q}$ and of the vector $\vec{V}$

$$\tilde{q}(\vec{V}) = q_j V^j = q_1 V^1 + \ldots + q_N V^N;$$ \hspace{1cm} (3.68)

The right-hand side of this equation can be considered as a linear combination of the components of $\vec{V}$ with coefficients $q_j$, or alternatively, as a linear combination of the components of $\tilde{q}$ with coefficients $V^j$, and this follows from the linearity of the previous expression. Therefore, we can define vectors as those linear functions that, when applied to one-forms, produce a number.

Let us now make a coordinate transformation $x^k' = x^k(x^i)$ and let us consider the following questions.

1. How do the components of one-forms transform?

2. Will the new coordinate basis for one-forms be a linear combination of the old ones, and if so, and which combination?

1. By definition

$$q_j = \tilde{q}(\vec{e}(j)).$$ \hspace{1cm} (3.69)

If we change coordinates, we will have a new set of basis vectors $\{\vec{e}'(j')\}$, and we have seen that they are related to the old ones by

$$\vec{e}'(j') = \Lambda^k_r \vec{e}(k),$$ \hspace{1cm} (3.70)

where $\Lambda^k_r = \frac{\partial x^k}{\partial x'^r}$. The new components of $\tilde{q}$ will be

$$q_{j'} = \tilde{q}(\vec{e}'(j')) = \tilde{q}[\Lambda^k_{j'} \vec{e}(k)] = \Lambda^k_{j'} \tilde{q} (\vec{e}(k)) = \Lambda^k_{j'} q_k,$$ \hspace{1cm} (3.71)

hence

$$q_{j'} = \Lambda^k_{j'} q_k.$$ \hspace{1cm} (3.72)

If we compare this result with eq. (3.7) we immediately recognize that this is the way covariant vectors transform, thus **covariant vectors are one-forms**.
2. We now want to check whether the new basis one-forms can be expressed as a linear combination of the old ones. We shall proceed along the same lines of section 3.4. From eq. (3.65) we see that
\[
\tilde{q}_j \tilde{\omega}^{(j)} = q_{kj} \tilde{\omega}^{(k)}, \tag{3.73}
\]
(sum removed according to Einstein’s convention), where \( \{ \tilde{\omega}^{(k')} \} \) are the new basis one-forms. But
\[
q_{kj} \tilde{\omega}^{(j)} = \Lambda_{kj} q_i, \tag{3.74}
\]
therefore
\[
q_j \tilde{\omega}^{(j)} = \Lambda_{kj} q_i \tilde{\omega}^{(k)}. \tag{3.75}
\]
This equation can be rewritten as
\[
[\Lambda_{kj} \tilde{\omega}^{(k)} - \tilde{\omega}^{(i)}] q_i = 0, \tag{3.76}
\]
hence
\[
\tilde{\omega}^{(i)} = \Lambda_{kj} \tilde{\omega}^{(k)}. \tag{3.77}
\]
The matrix \( \Lambda_{ij}^{'} \) is inverse of \( \Lambda_{ik} \). Thus
\[
\Lambda_{kj}^{'} \Lambda_{ji}^{'} = \delta_{ij}^{'} \quad \text{or} \quad \Lambda_{kj}^{'} \Lambda_{ik'} = \delta_{ij}'. \tag{3.78}
\]
Multiplying both sides of eq. (3.77) by \( \Lambda_{ji}^{'} \) we find
\[
\Lambda_{ji}^{'} \tilde{\omega}^{(i)} = \Lambda_{ji}^{'} \Lambda_{kj} \tilde{\omega}^{(k)} = \delta_{ij}' \tilde{\omega}^{(k)}, \tag{3.79}
\]
hence
\[
\tilde{\omega}^{(j)} = \Lambda_{ji}^{'} \tilde{\omega}^{(i)}. \tag{3.80}
\]
Summarizing, the transformation laws for the basis one-forms are
\[
\left\{ \begin{array}{l}
\tilde{\omega}^{(i)} = \Lambda_{kj} \tilde{\omega}^{(k)} \\
\tilde{\omega}^{(k)} = \Lambda_{kj}^{'} \tilde{\omega}^{(j)}
\end{array} \right. \tag{3.81}
\]

**EXAMPLE**

Let us consider the same coordinate transformation analyzed in section 3.4. We start with Minkowskian coordinates \((x^0, x^1, x^2)\). The coordinate basis for vectors is \( \{ \frac{\partial}{\partial x^\alpha} \} \) and the dual basis for one-forms is \( \{ \hat{dx}^\alpha \} \)
\[
\tilde{dx}^{(0)} \rightarrow (1, 0, 0) \tag{3.82}
\]
\[
\tilde{dx}^{(1)} \rightarrow (0, 1, 0) \tag{3.83}
\]
\[
\tilde{dx}^{(2)} \rightarrow (0, 0, 1) \tag{3.84}
\]
If we now change to polar coordinates \((x^0' = x^0, x^1' = r, x^2' = \theta)\), according to eq. (3.80) we find
\[
\tilde{\omega}^{(0')} = \Lambda_{0'}^\alpha \hat{dx}^{(\alpha)}. \tag{3.85}
\]
Since \( \Lambda^0' \alpha = \frac{\partial \omega'}{\partial x^\alpha} \), only \( \Lambda^0'0 = 1 \neq 0 \), thus
\[
\tilde{\omega}^{(0')} = \tilde{dx}^{(0)}.
\] (3.86)

Similarly
\[
\tilde{\omega}^{(1')} = \Lambda^1'_{\alpha} \tilde{dx}^{(\alpha)} = \frac{\partial x^{1'}}{\partial x^\alpha} \tilde{dx}^{(\alpha)} = \frac{\partial x^{1'}}{\partial x^1} \tilde{dx}^{(1)} + \frac{\partial x^{1'}}{\partial x^2} \tilde{dx}^{(2)}.
\] (3.87)

Since
\[
\frac{\partial x^{1'}}{\partial x^1} = \frac{x^1}{x^{1'}} = \cos \theta, \quad \text{and} \quad \frac{\partial x^{1'}}{\partial x^2} = \frac{x^2}{x^{1'}} = \sin \theta
\] (3.88)

it follows that
\[
\tilde{\omega}^{(1')} = \cos \theta \tilde{dx}^1 + \sin \theta \tilde{dx}^2.
\] (3.89)

Moreover
\[
\tilde{\omega}^{(2')} = \frac{\partial x^{2'}}{\partial x^1} \tilde{dx}^{(1)} + \frac{\partial x^{2'}}{\partial x^2} \tilde{dx}^{(2)},
\] (3.90)

hence
\[
\tilde{\omega}^{(2')} = -\frac{1}{r} \sin \theta \tilde{dx}^{(1)} + \frac{1}{r} \cos \theta \tilde{dx}^{(2)}.
\] (3.91)

Summarizing,
\[
\begin{cases}
\tilde{\omega}^{(0')} = \tilde{\omega}^{(0)} \\
\tilde{\omega}^{(1')} = \cos \theta \tilde{\omega}^{(1)} + \sin \theta \tilde{\omega}^{(2)} \\
\tilde{\omega}^{(2')} = -\frac{1}{r} \sin \theta \tilde{\omega}^{(1)} + \frac{1}{r} \cos \theta \tilde{\omega}^{(2)}.
\end{cases}
\] (3.92)

AN EXAMPLE OF ONE-FORM.
Consider a scalar field \( \Phi(x^1, \ldots, x^N) \). The gradient of a scalar field is
\[
\tilde{\Phi} \to \left( \frac{\partial \Phi}{\partial x^1}, \ldots, \frac{\partial \Phi}{\partial x^N} \right).
\] (3.93)

It is easy to see, for example, that the components transform according to eq. (3.72), in fact
\[
\tilde{\Phi}_j = \frac{\partial \Phi}{\partial x^j}, \quad \text{and} \quad \tilde{\Phi}_{j'} = \frac{\partial \Phi}{\partial x^{j'}} = \frac{\partial \Phi}{\partial x^k} \cdot \frac{\partial x^k}{\partial x^{j'}};
\] (3.94)

since \( \Lambda^k_{j'} = \frac{\partial x^k}{\partial x^{j'}} \), it follows that
\[
\tilde{\Phi}_{j'} = \Lambda^k_{j'} \tilde{\Phi}_k,
\] (3.95)

same as eq. (3.72). Thus the gradient of a scalar field is a one-form.

### 3.6 Vector fields and one-form fields

The vectors and one-forms are defined on a point \( P \) of the manifold, and belong to the vector spaces \( T_P \) and \( T^*_P \), respectively, which also refer to a specific point \( P \) of the manifold; to make this explicit, we could also denote a vector in \( P \) as \( \tilde{V}_P \), a one-form in \( P \) as \( \tilde{W}_P \). We shall now define vector fields and one-form fields.
Given an open set $U$ of a differentiable manifold $M$, we define the vector spaces
\[ T_U \equiv \bigcup_{P \in U} T_P \]
\[ T^*_U \equiv \bigcup_{P \in U} T^*_P, \]
i.e., the union of the tangent spaces on the points $P \in U$, and the union of the cotangent spaces on the points $P \in U$.

A vector field $\vec{V}$ is a mapping
\[ \vec{V} : U \rightarrow T_U \]
\[ P \rightarrow \vec{V}_P \]
which associates, to every point $P \in U$, a vector $\vec{V}_P$ defined on the tangent space in $P$, $T_P$. A one-form field $\tilde{W}$ is a mapping
\[ \tilde{W} : U \rightarrow T^*_U \]
\[ P \rightarrow \tilde{W}_P \]
which associates, to every point $P \in U$, a one-form $\tilde{W}_P$ defined on the cotangent space in $P$, $T^*_P$. If a coordinate system (a chart) $\{x^\mu\}$ is defined on $U$, we can indicate the vector field and the one-form field as $\vec{V}(x)$, $\tilde{W}(x)$.

In the following, we will mainly consider vector fields and one-form fields; however, for brevity of notation, we will often refer to them simply as vectors and one-forms.
Chapter 4

Tensors

4.1 Geometrical definition of a Tensor

The definition of a tensor is a generalization of the definition of one-forms. Consider a point \( P \) of an \( n \)-dimensional manifold \( M \). A tensor of type \( \left( \frac{N}{N'} \right) \) at \( P \) is defined to be a \textit{linear, real valued function, which takes as arguments \( N \) one-forms and \( N' \) vectors and associates a number to them.}

For example if \( F \) is a \( \left( \frac{2}{2} \right) \) tensor this means that

\[
F(\tilde{\omega}, \tilde{\sigma}, \vec{V}, \vec{W})
\]

is a number and the linearity implies that

\[
F(a\tilde{\omega} + b\tilde{\eta}, \tilde{\sigma}, \vec{V}, \vec{W}) = aF(\tilde{\omega}, \tilde{\sigma}, \vec{V}, \vec{W}) + bF(\tilde{\eta}, \tilde{\sigma}, \vec{V}, \vec{W})
\]

and

\[
F(\tilde{\omega}, \tilde{\eta}, a\vec{V}_1 + b\vec{V}_2, \vec{W}) = aF(\tilde{\omega}, \tilde{\eta}, \vec{V}_1, \vec{W}) + bF(\tilde{\omega}, \tilde{\eta}, \vec{V}_2, \vec{W})
\]

and similarly for the other arguments.

This definition of tensors is rather abstract, but we shall see how to make it concrete with specific examples.

The order in which the arguments are placed is important, as it is true for any function of real variables. For example if

\[
f(x, y) = 4x^3 + 5y, \quad \text{then} \quad f(1, 5) \neq f(5, 1). \tag{4.1}
\]

In the same way

\[
F(\tilde{\omega}, \tilde{\eta}, \vec{V}, \vec{W}) \neq F(\tilde{\eta}, \tilde{\omega}, \vec{V}, \vec{W}). \tag{4.2}
\]

**EXAMPLES**

A \( \left( \frac{0}{1} \right) \) tensor is a function that takes a vector as argument, and produces a number.

This is precisely what one-forms do (on the other hand this is the definition of one-forms).
Thus, a \( \begin{pmatrix} 0 & 1 \end{pmatrix} \) tensor is a one-form.

\[
\tilde{q}(\vec{V}) = \sum_{\alpha} q_{\alpha} V^\alpha \equiv q_{\alpha} V^\alpha. \tag{4.3}
\]

A \( \begin{pmatrix} 1 & 0 \end{pmatrix} \) tensor is a function that takes a one-form as an argument, and produces a number. Thus a \( \begin{pmatrix} 1 & 0 \end{pmatrix} \) tensor is a vector

\[
\vec{V}(\tilde{q}) = q_{\alpha} V^\alpha. \tag{4.4}
\]

Let us now consider a \( \begin{pmatrix} 0 & 2 \end{pmatrix} \) tensor. It is a function that takes 2 vectors and associates a number to them.

Let us first define the tensor components: generalizing the definition (3.64) for the components of a one-form, they are the numbers that are obtained when the \( \begin{pmatrix} 0 & 2 \end{pmatrix} \) tensor is applied to the basis vectors:

\[
F_{\alpha\beta} = F(\vec{e}(\alpha), \vec{e}(\beta)); \tag{4.5}
\]

since there are n basis vectors, \( F_{\alpha\beta} \) will be a \( n \times n \) matrix.

If we now take as arguments of \( \vec{F} \) two arbitrary vectors \( \vec{A} \) and \( \vec{B} \) we find

\[
F(\vec{A}, \vec{B}) = F(A^\alpha \vec{e}(\alpha), B^\beta \vec{e}(\beta)) = A^\alpha B^\beta F(\vec{e}(\alpha), \vec{e}(\beta)) = F_{\alpha\beta} A^\alpha B^\beta. \tag{4.6}
\]

It should be stressed that in going from the first to the second line of eq. (4.6) we have used the property that tensors are linear functions of the arguments.

It is now clear what is the number that \( \vec{F} \) associates to the two vectors: the number is \( F_{\alpha\beta} A^\alpha B^\beta \).

We shall now construct a basis for \( \begin{pmatrix} 0 & 2 \end{pmatrix} \) tensors as we did for one-forms.

We want to write

\[
F = F_{\alpha\beta} \omega^{(\alpha)(\beta)} \tag{4.7}
\]

where \( \omega^{(\alpha)(\beta)} \) are the basis \( \begin{pmatrix} 0 & 2 \end{pmatrix} \) tensors.

If the arguments of \( \vec{F} \) are two arbitrary vectors \( \vec{A} \) and \( \vec{B} \), eq. (4.7) gives

\[
F(\vec{A}, \vec{B}) = F_{\alpha\beta} \omega^{(\alpha)(\beta)}(\vec{A}, \vec{B}). \tag{4.8}
\]

On the other hand, since \( A^\alpha = \tilde{\omega}^{(\alpha)}(\vec{A}) \) and \( B^\beta = \tilde{\omega}^{(\beta)}(\vec{B}) \), eq. (4.6) gives

\[
F(\vec{A}, \vec{B}) = F_{\alpha\beta} \tilde{\omega}^{(\alpha)}(\vec{A}) \tilde{\omega}^{(\beta)}(\vec{B}), \tag{4.9}
\]

and, by equating eqs. (4.8) and (4.9) we find

\[
\omega^{(\alpha)(\beta)}(\vec{A}, \vec{B}) = \tilde{\omega}^{(\alpha)}(\vec{A}) \tilde{\omega}^{(\beta)}(\vec{B}).
\]
CHAPTER 4. TENSORS

The previous equation holds for any two vectors \( \vec{A} \) and \( \vec{B} \), consequently we write

\[
\omega^{(\alpha)(\beta)} = \tilde{\omega}^{(\alpha)} \otimes \tilde{\omega}^{(\beta)},
\]

(4.10)

where the symbol \( \otimes \) indicates the “outer product” of the two basis one-forms, and means precisely that if \( \omega^{(\alpha)(\beta)} \) is applied to the vectors \( \vec{A} \) and \( \vec{B} \), the result is a number, which coincides with the number produced by the application of \( \tilde{\omega}^{(\alpha)} \) to \( \vec{A} \), times that produced by the application of \( \tilde{\omega}^{(\beta)} \) to \( \vec{B} \) (the order is important!).

Thus the basis for \( \left( \begin{array}{c} 0 \\ 2 \end{array} \right) \) tensors can be constructed by taking the outer product of the basis one-forms. Finally, we can write

\[
F = F_{\alpha\beta} \tilde{\omega}^{(\alpha)} \otimes \tilde{\omega}^{(\beta)}.
\]

(4.11)

It is now clear that we can construct any sort of tensors using the procedure that we have developed in the previous pages. Thus for example a \( \left( \begin{array}{c} 2 \\ 0 \end{array} \right) \) tensor \( T \) is a function that associates to two one-forms \( \tilde{\alpha} \) and \( \tilde{\sigma} \) a number, \( T(\tilde{\alpha}, \tilde{\sigma}) \).

The components of this tensor are found by applying \( T \) to the basis one-forms

\[
T^{\mu\nu} = T(\tilde{\omega}^{(\mu)}, \tilde{\omega}^{(\nu)}),
\]

(4.12)

and the number produced when \( T \) is applied to any two one-forms \( \tilde{\alpha}, \tilde{\sigma} \) will be

\[
T(\tilde{\alpha}, \tilde{\sigma}) = T(\alpha_{\mu} \tilde{\omega}^{(\mu)}, \sigma_{\nu} \tilde{\omega}^{(\nu)}) = \alpha_{\mu} \sigma_{\nu} T(\tilde{\omega}^{(\mu)}, \tilde{\omega}^{(\nu)}) = \alpha_{\mu} \sigma_{\nu} T^{\mu\nu},
\]

(4.13)

where again use has been made of the linearity of tensors with respect to their arguments. By following the same procedure used to find the basis for a \( \left( \begin{array}{c} 2 \\ 0 \end{array} \right) \) tensor, it is easy to show that the basis appropriate for a \( \left( \begin{array}{c} 2 \\ 0 \end{array} \right) \) tensor will be

\[
e_{(\alpha)(\beta)} = \tilde{e}_{\alpha} \otimes \tilde{e}_{\beta},
\]

(4.14)

and consequently

\[
T = T^{\alpha\beta} e_{\alpha} \otimes e_{\beta}.
\]

(4.15)

Exercise: prove that the \( \left( \begin{array}{c} 1 \\ 1 \end{array} \right) \) tensor \( \vec{V} \otimes \tilde{\sigma} \) has components \( V^{\mu} \sigma_{\nu} \) and find the basis for \( \left( \begin{array}{c} 1 \\ 1 \end{array} \right) \) tensors.

Now we ask the following question: how do the components of a tensor transform if we make a coordinate transformation? We start with a \( \left( \begin{array}{c} 0 \\ 2 \end{array} \right) \) tensor

\[
F = F_{\alpha\beta} \tilde{\omega}^{(\alpha)} \otimes \tilde{\omega}^{(\beta)}
\]

(4.16)
If we change coordinates, we shall have a new set of basis one forms \( \tilde{\omega}^{(\mu')} \) which are related to the old ones by the equations

\[
\tilde{\omega}^{(\alpha)} = \Lambda_{\alpha}^{\mu'} \tilde{\omega}^{(\mu')}, \quad \tilde{\omega}^{(\mu')} = \Lambda^{\mu'}_{\alpha} \tilde{\omega}^{(\alpha)}
\] (4.17)

In the new basis the tensor \( \begin{pmatrix} 0 \\ 2 \end{pmatrix} \) will be

\[
F = F_{\alpha'\beta'} \tilde{\omega}^{(\alpha')} \otimes \tilde{\omega}^{(\beta')}. \quad \tag{4.18}
\]

By equating (4.16) and (4.18)

\[
F_{\alpha'\beta'} \tilde{\omega}^{(\alpha')} \otimes \tilde{\omega}^{(\beta')} = F_{\alpha\beta} \tilde{\omega}^{(\alpha)} \otimes \tilde{\omega}^{(\beta)}.
\]

Replacing \( \tilde{\omega}^{(\alpha)} \) and \( \tilde{\omega}^{(\beta)} \) by using the first of eqs. 4.17

\[
F_{\alpha'\beta'} \tilde{\omega}^{(\alpha')} \otimes \tilde{\omega}^{(\beta')} = F_{\alpha\beta} \Lambda_{\alpha}^{\mu'} \Lambda^{\beta}_{\nu} \tilde{\omega}^{(\mu')} \otimes \tilde{\omega}^{(\nu')},
\]

or by relabelling the dummy indices

\[
F_{\mu'\nu'} \tilde{\omega}^{(\mu')} \otimes \tilde{\omega}^{(\nu')} = F_{\alpha\beta} \Lambda_{\mu'}^{\alpha} \Lambda_{\nu}^{\beta} \tilde{\omega}^{(\mu')} \otimes \tilde{\omega}^{(\nu')},
\]

and finally

\[
F_{\mu'\nu'} = F_{\alpha\beta} \Lambda_{\mu'}^{\alpha} \Lambda_{\nu}^{\beta}, \quad \tag{4.19}
\]

or, by writing explicitly the elements of the matrix \( \Lambda_{\mu'}^{\alpha} \)

\[
F_{\mu'\nu'} = F_{\alpha\beta} \frac{\partial x^{\alpha}}{\partial x_{\mu'}} \frac{\partial x^{\beta}}{\partial x_{\nu'}}, \quad \tag{4.20}
\]

where \( \{ x^{\mu'} \} \) are the new coordinates.

In a similar way, by using eqs. 3.33 and 3.35 we would find that

\[
T^{\mu'\nu'} = T^{\alpha\beta} \Lambda_{\mu'}^{\alpha} \Lambda_{\nu'}^{\beta}, \quad \tag{4.21}
\]

and

\[
T^{\mu'\nu'} = T^{\alpha\beta} \Lambda_{\mu'}^{\alpha} \Lambda_{\nu'}^{\beta}. \quad \tag{4.22}
\]

**IMPORTANT**

The following point should be stressed: the notion of tensor we have introduced is independent of which coordinates, i.e. which basis, we use. In fact the number that an \( \begin{pmatrix} N \\ N' \end{pmatrix} \) tensor associates to \( N \) one-forms and \( N' \) vectors does not depend on the particular basis we choose.

This is the reason why, for example, we can equate eqs. (4.16) and (4.18).

The operations that we are allowed to make with tensors are the following.
Multiplication by a real number

Given a tensor $T$ of type $\left( \begin{array}{c} N \\ N' \end{array} \right)$ and a real number $a$, we define the tensor, of the same type,

$$W = aT.$$ 

Let the components of $T$, in a given frame, be $\{T^{\alpha...}_{\beta...}\}$. The components of $W$ are

$$W^{\alpha...}_{\beta...} = aT^{\alpha...}_{\beta...}.$$ 

Addition of tensors

Given two tensors $T, G$ of the same type $\left( \begin{array}{c} N \\ N' \end{array} \right)$, we define the tensor, of the same type,

$$W = T + G.$$ 

Let the components of $T, G$, in a given frame, be $\{T^{\alpha...}_{\beta...}\}, \{G^{\alpha...}_{\beta...}\}$. The components of $W$ in that frame are

$$W^{\alpha...}_{\beta...} = T^{\alpha...}_{\beta...} + G^{\alpha...}_{\beta...}.$$ 

Outer product

Given two tensors $T, G$ of types $\left( \begin{array}{c} N_1 \\ N'_1 \end{array} \right), \left( \begin{array}{c} N_2 \\ N'_2 \end{array} \right)$, respectively. We define the tensor, of type $\left( \begin{array}{c} N_1 + N_2 \\ N'_1 + N'_2 \end{array} \right)$,

$$W = T \otimes G.$$ 

Let the components of $T, G$, in a given frame, be $\{T^{\alpha...}_{\beta...}\}, \{G^{\gamma...}_{\delta...}\}$. The components of $W$ in that frame are

$$W^{\alpha...\gamma...}_{\beta...\delta...} = T^{\alpha...}_{\beta...}G^{\gamma...}_{\delta...}.$$ 

For instance, if both $T, G$ are of type $\left( \begin{array}{c} 0 \\ 2 \end{array} \right)$,

$$W^{\alpha\beta\gamma\delta}_{\alpha\beta\gamma\delta} = T^{\alpha\beta}_{\alpha\beta}G^{\gamma\delta}_{\gamma\delta}.$$ 

Contraction

Given a tensor $T$ of type $\left( \begin{array}{c} N \\ N' \end{array} \right)$, we choose one of the contravariant (i.e. upper) indexes, and one of the covariant (i.e. lower) indexes of the tensor; then, we define a tensor $W$ of type $\left( \begin{array}{c} N - 1 \\ N' - 1 \end{array} \right)$. Let the components of $T$, in a given frame, be $\{T^{\alpha_1\alpha_2...}_{\beta_1\beta_2...}\}$; the components of $W$ in that frame are

$$W^{...\alpha_1\alpha_{i+1}...}_{...\beta_{j-1}\beta_j...} = T^{...\alpha_1\alpha_{i+1}...}_{...\beta_{j-1}\beta_j...}.$$
For instance, if $T$ is of type $\begin{pmatrix} 3 & \cdot & \cdot \\ \cdot & 3 & \cdot \\ \cdot & \cdot & 3 \end{pmatrix}$ and we choose to contract the first contravariant index with the first covariant index,

$$W^\beta_\gamma = T^\alpha_\beta \gamma_\alpha \delta_\sigma = T^0_0 \gamma_0 \delta + T^1_1 \gamma_1 \sigma + T^2_2 \gamma_2 \sigma + \ldots$$

These are called tensor operations and an equation involving tensor components and tensor operations is a tensor equation.

Finally, we remark that since a tensor $T$ has been defined as an application from vectors and one-forms, it is defined on the product of a certain number of copies of the tangent and the cotangent spaces on a point $P$, $T_P$, $T^*_P$. Then, we can define tensor fields, i.e., a tensor for each point $P$ of an open subset of the manifold; in a given coordinate system $\{x^\mu\}$, we can write a tensor field as $T(x)$. For brevity of notation, in the following we will often refer to a tensor field simply as a tensor.

## 4.2 Symmetries

A $\begin{pmatrix} 0 & B \\ A & 2 \end{pmatrix}$ tensor $F$ is **Symmetric** if

$$F(\vec{A}, \vec{B}) = F(\vec{B}, \vec{A}) \quad \forall \vec{A}, \vec{B}. \quad (4.23)$$

As a consequence of eq. (4.6) we see that if the tensor is symmetric

$$F_{\alpha\beta} A^\alpha B^\beta = F_{\alpha\beta} B^\alpha A^\beta, \quad (4.24)$$

and, by relabelling the indices on the RHS

$$F_{\alpha\beta} A^\alpha B^\beta = F_{\beta\alpha} B^\beta A^\alpha, \quad (4.25)$$

i.e.

$$F_{\alpha\beta} = F_{\beta\alpha} \quad (4.26)$$

i.e. if a $\begin{pmatrix} 0 & B \\ A & 2 \end{pmatrix}$ tensor is symmetric the matrix representing its components is symmetric.

Given any $\begin{pmatrix} 0 & B \\ A & 2 \end{pmatrix}$ tensor $F$ we can always construct from it a symmetric tensor $F^{(s)}$

$$F^{(s)}(\vec{A}, \vec{B}) = \frac{1}{2} [F(\vec{A}, \vec{B}) + F(\vec{B}, \vec{A})]. \quad (4.27)$$

In fact $\forall \vec{A}, \vec{B}$

$$\frac{1}{2} [F(\vec{A}, \vec{B}) + F(\vec{B}, \vec{A})] = \frac{1}{2} [F(\vec{B}, \vec{A}) + F(\vec{A}, \vec{B})].$$
Moreover

\[ F^{(s)}(\vec{A}, \vec{B}) = F_{\alpha\beta}^s A^\alpha B^\beta = \frac{1}{2} [F_{\alpha\beta} A^\alpha B^\beta + F_{\alpha\beta} B^\alpha A^\beta] = \frac{1}{2} [F_{\alpha\beta} A^\alpha B^\beta + F_{\beta\alpha} B^\beta A^\alpha] \]

and consequently the components of the symmetric tensor are

\[ F_{\alpha\beta}^{(s)} = \frac{1}{2} [F_{\alpha\beta} + F_{\beta\alpha}] . \] (4.28)

The components of a symmetric tensor are often indicated as

\[ F_{(\alpha\beta)} = \frac{1}{2} [F_{\alpha\beta} + F_{\beta\alpha}] . \] (4.29)

A \( \left( \begin{array}{c} 0 \\ 2 \end{array} \right) \) tensor \( F \) is antisymmetric if

\[ F(\vec{A}, \vec{B}) = -F(\vec{B}, \vec{A}) \quad \forall \vec{A}, \vec{B}, \quad \text{i.e.} \quad F_{\alpha\beta} = -F_{\beta\alpha} . \] (4.30)

Again from any \( \left( \begin{array}{c} 0 \\ 2 \end{array} \right) \) tensor we can construct an antisymmetric tensor \( F^{(a)} \) defined as

\[ F^{(a)}(\vec{A}, \vec{B}) = \frac{1}{2} [F(\vec{A}, \vec{B}) - F(\vec{B}, \vec{A})] . \]

Proceeding as before, we find that its components are

\[ F_{\alpha\beta}^{(a)} = \frac{1}{2} [F_{\alpha\beta} - F_{\beta\alpha}] , \]

also indicated as

\[ F_{[\alpha\beta]} = \frac{1}{2} [F_{\alpha\beta} - F_{\beta\alpha}] . \] (4.31)

It is clear that any tensor \( \left( \begin{array}{c} 0 \\ 2 \end{array} \right) \) can be written as the sum of its symmetric and antisymmetric part

\[ h[\vec{A}, \vec{B}] = \frac{1}{2} [h(\vec{A}, \vec{B}) + h(\vec{B}, \vec{A})] + \frac{1}{2} [h(\vec{A}, \vec{B}) - h(\vec{B}, \vec{A})] \]

4.3 The metric Tensor

In chapter 1 we have seen that the metric tensor has a central role in the relativistic theory of gravity. In this section we shall discuss its geometrical meaning.

Definition: the metric tensor \( g \) is a \( \left( \begin{array}{c} 0 \\ 2 \end{array} \right) \) tensor that, having two arbitrary vectors \( \vec{A} \) and \( \vec{B} \) as arguments, associates to them a real number that is the **inner product** (or scalar product) \( \vec{A} \cdot \vec{B} \)

\[ g(\vec{A}, \vec{B}) = \vec{A} \cdot \vec{B} . \] (4.32)
The scalar product is usually defined to be a linear function of two vectors that satisfies the following properties

\[
\vec{U} \cdot \vec{V} = \vec{V} \cdot \vec{U} \\
(a \vec{U}) \cdot \vec{V} = a(\vec{U} \cdot \vec{V}) \\
(\vec{U} + \vec{V}) \cdot \vec{W} = \vec{U} \cdot \vec{W} + \vec{V} \cdot \vec{W}
\] (4.33)

From the first eq. (4.33) it follows that \( g \) is a symmetric tensor. In fact

\[
\vec{U} \cdot \vec{V} = g(\vec{U}, \vec{V}) = \vec{V} \cdot \vec{U} = g(\vec{V}, \vec{U}).
\] (4.34)

The second and third eqs. (4.33) imply that \( g \) is a linear functions of the arguments, a condition which is automatically satisfied since \( g \) is a tensor.

As usual the components of the metric tensor are obtained by replacing \( \vec{A} \) and \( \vec{B} \) with the basis vectors

\[
g_{\alpha \beta} = g(\vec{e}_\alpha, \vec{e}_\beta) = \vec{e}_\alpha \cdot \vec{e}_\beta.
\] (4.35)

Thus the metric tensor allows to compute the scalar product of two vectors in any space and whatever coordinates we use:

\[
\vec{A} \cdot \vec{B} = g(\vec{A}, \vec{B}) = g(A^\alpha \vec{e}_\alpha, B^\beta \vec{e}_\beta) = A^\alpha B^\beta g(\vec{e}_\alpha, \vec{e}_\beta) = A^\alpha B^\beta g_{\alpha \beta}.
\] (4.36)

EXAMPLES

1) The metric of four dimensional Minkowski spacetime, in Minkowskian coordinates \( x^\alpha = (ct, x, y, z) \) is

\[
g_{\alpha \beta} = \begin{pmatrix}
-1 & 0 & 0 & 0 \\
0 & +1 & 0 & 0 \\
0 & 0 & +1 & 0 \\
0 & 0 & 0 & +1
\end{pmatrix} \equiv \eta_{\alpha \beta}
\]

i.e.

\[
ds^2 = g_{\alpha \beta} dx^\alpha dx^\beta = -c^2 dt^2 + dx^2 + dy^2 + dz^2.
\] (4.37)

This implies that the basis vectors in the coordinate basis

\[
\vec{e}_{(0)} = \vec{e}_{(ct)} \rightarrow (1, 0, 0, 0) \\
\vec{e}_{(1)} = \vec{e}_{(x)} \rightarrow (0, 1, 0, 0) \\
\vec{e}_{(2)} = \vec{e}_{(y)} \rightarrow (0, 0, 1, 0) \\
\vec{e}_{(3)} = \vec{e}_{(z)} \rightarrow (0, 0, 0, 1)
\]

are, in this case, mutually orthogonal:

\[
\vec{e}_{(\alpha)} \cdot \vec{e}_{(\beta)} = g_{\alpha \beta} = 0 \quad \text{if} \quad \alpha \neq \beta.
\]
In addition, since
\[ g_{11} = g_{22} = g_{33} = 1, \quad \text{and} \quad g_{00} = -1, \]
the basis vectors are unit vectors, \( \vec{e}_{(0)} \) is a timelike vector, and \( \vec{e}_{(i)} \) \((i = 1, 2, 3)\) are spacelike vectors:
\[ \vec{e}_{(k)} \cdot \vec{e}_{(k)} = 1 \quad \text{if} \quad k = 1, \ldots, 3, \]
\[ \vec{e}_{(0)} \cdot \vec{e}_{(0)} = -1. \]

From now on we shall indicate as \( \eta_{\alpha\beta} \) the components of the metric tensor of the Minkowski spacetime when expressed in cartesian coordinates.

2) Let us now consider the metric of Minkowski spacetime in three dimensions, i.e. we suppress the coordinate \( z \):
\[ g_{\alpha\beta} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & +1 \end{pmatrix} \equiv \eta_{\alpha\beta} \tag{4.38} \]
with \( \alpha, \beta = 0, \ldots, 2 \). The vectors of the coordinate basis have components
\[ \vec{e}_{(0)} \rightarrow (1, 0, 0) \]
\[ \vec{e}_{(1)} \rightarrow (0, 1, 0) \]
\[ \vec{e}_{(2)} \rightarrow (0, 0, 1) \].

We now change to polar coordinates
\[ x^0 = x^0', \quad x^1 = r \cos \theta, \quad x^2 = r \sin \theta. \tag{4.39} \]
The vectors of the coordinate basis in the new coordinate system have been computed in Sec. 3.4, and are
\[ \vec{e}_{(0')} = \vec{e}_{(0)} \tag{4.40} \]
\[ \vec{e}_{(1')} = \vec{e}_{(r)} = \cos \theta \vec{e}_{(1)} + \sin \theta \vec{e}_{(2)} \tag{4.41} \]
\[ \vec{e}_{(2')} = \vec{e}_{(\theta)} = -r \sin \theta \vec{e}_{(1)} + r \cos \theta \vec{e}_{(2)}. \]

We can determine the metric tensor in the new frame by computing the scalar product of the vectors of this frame:
\[ g_{0'0'} = \vec{e}_{(0')} \cdot \vec{e}_{(0')} = \vec{e}_{(0)} \cdot \vec{e}_{(0)} = -1 \]
\[ g_{0'1'} = 0 \quad i' = 1, 2 \]
\[ g_{1'1'} = \vec{e}_{(1')} \cdot \vec{e}_{(1')} = (\cos \theta \vec{e}_{(1)} + \sin \theta \vec{e}_{(2)}) \cdot (\cos \theta \vec{e}_{(1)} + \sin \theta \vec{e}_{(2)}) = \cos^2 \theta + \sin^2 \theta = 1 \]
\[ g_{2'2'} = \vec{e}_{(2')} \cdot \vec{e}_{(2')} = r^2 \sin^2 \theta + r^2 \cos^2 \theta = r^2 \]
\[ g_{1'2'} = -r \cos \theta \sin \theta + r \cos \theta \sin \theta = 0 \]
i.e.
\[ g_{\alpha'\beta'} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & r^2 \end{pmatrix} \tag{4.42} \]
i.e.

\[ ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta = g_{\alpha'\beta'} dx^{\alpha'} dx^{\beta'} = -dt^2 + dr^2 + r^2 d\theta^2 . \]  

(4.43)

We note that although the metric tensor is the same, its components in the two coordinate frames, (4.38) and (4.42), are different, since \( g^{22'} = \hat{e}(2') \cdot \hat{e}(2') = r^2 \neq 1 \). Thus, \( \hat{e}(2') \) is not an unit vector. In general the basis vectors are not required to have unitary norm, even in a coordinate frame.

Usually, to determine the components of the metric tensor in a new frame, one does not use the procedure above, based on the computation of the scalar products. One rather employs the transformation law

\[ g_{\mu'\nu'} = \Lambda^\alpha_\mu' \Lambda^\beta_\nu' g_{\alpha\beta} \]

which, in this case, has the form

\[ g_{\mu'\nu'} = \Lambda^\alpha_\mu' \Lambda^\beta_\nu' \eta_{\alpha\beta} . \]

Since \( \eta_{\alpha\beta} \) is diagonal, we only need to consider the components with \( \alpha = \beta \).

\[ g_{00'} = \Lambda^0_0' \Lambda^0_0 \eta_{00} = \left( \frac{\partial x^0}{\partial x^{0'}} \right)^2 \eta_{00} = 1 \cdot (-1) = -1 \]

\[ g_{01'} = \Lambda^0_0' \Lambda^1_1 \eta_{01} = \frac{\partial x^0}{\partial x^{0'}} \frac{\partial x^0}{\partial x^{1'}} \eta_{00} + \frac{\partial x^1}{\partial x^{0'}} \frac{\partial x^1}{\partial x^{1'}} \eta_{11} + \frac{\partial x^2}{\partial x^{0'}} \frac{\partial x^2}{\partial x^{1'}} \eta_{22} = 0 \quad i' = 1, 2 \]

because \( \frac{\partial x^0}{\partial x^{i'}} = \frac{\partial x^1}{\partial x^{i'}} = \frac{\partial x^2}{\partial x^{i'}} = 0. \)

\[ g_{11'} = (\Lambda^0_1 \Lambda^1_1 \eta_{00} + (\Lambda^1_1)^2 \eta_{11} + (\Lambda^2_1)^2 \eta_{22} = \]

\[ = \left( \frac{\partial x^0}{\partial x^{0'}} \right)^2 (-1) + \left( \frac{\partial x^1}{\partial x^{0'}} \right)^2 1 + \left( \frac{\partial x^2}{\partial x^{0'}} \right)^2 1 = \left( \frac{\partial x}{\partial r} \right)^2 + \left( \frac{\partial y}{\partial r} \right)^2 \]

\[ g_{11'} = \cos^2 \theta + \sin^2 \theta = 1 \]

Proceeding in this way we find the metric in the frame \((x^0', x^1', x^2') = (ct, r, \theta)\), i.e. (4.42).

4.3.1 The metric tensor allows to compute the distance between two points

Let us consider, for example, a three-dimensional space.

\[ (x^0, x^1, x^2) \equiv (ct, x, y) \]

The distance between two points infinitesimally close, \( P(x^0, x^1, x^2) \) and \( P'(x^0 + dx^0, x^1 + dx^1, x^2 + dx^2) \), is

\[ \bar{ds} = dx^0 \hat{e}(0) + dx^1 \hat{e}(1) + dx^2 \hat{e}(2) = dx^\alpha \hat{e}(\alpha) \]

(4.44)
By definition of the metric tensor $d\vec{s}$ is the norm of the vector $\vec{ds}$, i.e. the square of the distance between $P$ and $P'$:

$$
\begin{align*}
    ds^2 &= \vec{ds} \cdot \vec{ds} = (dx^0 \vec{e}_0 + dx^1 \vec{e}_1 + dx^2 \vec{e}_2) \cdot (dx^0 \vec{e}_0 + dx^1 \vec{e}_1 + dx^2 \vec{e}_2) \\
    &= (dx^0)^2 (\vec{e}_0 \cdot \vec{e}_0) + dx^0 dx^0 (\vec{e}_0 \cdot \vec{e}_0) + dx^0 dx^2 (\vec{e}_2 \cdot \vec{e}_0) + \\
    &\quad + dx^0 dx^1 (\vec{e}_0 \cdot \vec{e}_1) + (dx^1)^2 (\vec{e}_1 \cdot \vec{e}_1) + dx^1 dx^2 (\vec{e}_2 \cdot \vec{e}_1) + \\
    &\quad + dx^0 dx^2 (\vec{e}_0 \cdot \vec{e}_2) + dx^2 dx^1 (\vec{e}_2 \cdot \vec{e}_1) + (dx^2)^2 (\vec{e}_2 \cdot \vec{e}_2)
\end{align*}
$$

By definition of the metric tensor

$$(\vec{e}_i \cdot \vec{e}_j) = g(\vec{e}_i, \vec{e}_j) = g_{ij},$$

therefore

$$ds^2 = g(\vec{ds}, \vec{ds}) = (dx^0)^2 g_{00} + 2 dx^0 dx^1 g_{01} + 2 dx^0 dx^2 g_{02} + 2 dx^1 dx^2 g_{12} + (dx^1)^2 g_{11} + (dx^2)^2 g_{22},$$

(4.45)

where we have used the fact that $g_{\alpha\beta} = g_{\beta\alpha}$.

This calculation is simplified if we use the following notation

$$ds^2 = g(\vec{ds}, \vec{ds}) = \sum_{\alpha=0}^{2} dx^\alpha g(\vec{e}_\alpha, \vec{e}_\alpha) \equiv g(dx^\alpha \vec{e}_\alpha) = g_{\alpha\beta} dx^\alpha dx^\beta,$$

(4.46)

with $\alpha, \beta = 0, \ldots, 2$.

This way of writing is completely equivalent to eq. (4.45). For example, if the space is Minkowski spacetime $g_{\alpha\beta} = \eta_{\alpha\beta} = diag(-1,1,1)$, and eq. (4.46) gives

$$ds^2 = -(dx^0)^2 + (dx^1)^2 + (dx^2)^2,$$

(4.47)

as expected.

If we now change to a coordinate system $(x^0', x^1', x^2')$, the distance $P\vec{P}'$ will be $ds^2 = ds^2$, i.e.

$$g(\vec{ds}', \vec{ds}') = \vec{ds}' \cdot \vec{ds}' = ds'^2 = ds^2 = g(dx^\alpha' \vec{e}_{(\alpha')}, dx^\beta' \vec{e}_{(\beta')}) = dx^\alpha' dx^\beta' g(\vec{e}_{(\alpha')}, \vec{e}_{(\beta')}),$$

where $\{\vec{e}_{(\alpha')}\}$ are the new basis vectors. Therefore

$$ds^2 = g_{\alpha'\beta'} dx^\alpha' dx^\beta'$$

(4.48)

where now $g_{\alpha'\beta'}$ are the components of the metric tensor in the new basis. For example, if we change from cartesian to polar coordinates $(x^0', x^1', x^2') \equiv (ct, r, \theta)$,

$$ds^2 = (dx^0')^2 g_{00'} + (dx^1')^2 g_{11'} + (dx^2')^2 g_{22'} = -(dx^0)^2 + dr^2 + r^2 d\theta^2.$$  

(4.49)

Thus if we know the components of the metric tensor in any reference frame, we can compute the distance between two points infinitesimally close, $ds^2$. 

The “infinitesimal” interpretation of $ds^2$ we have discussed above is useful to understand the role of the metric in measuring distances. A more mathematically rigorous discussion involves the measure of finite distances, through the definition of the length of a path: let us consider a curve, i.e. a path $\mathcal{C}$ and a map

$$[a, b] \subset \mathbb{R} \to \mathcal{C}$$

$$\lambda \to P(\lambda)$$

which, in a given coordinate system $\{x^\mu\}$, corresponds to the real functions

$$\lambda \to \{x^\mu(\lambda)\}.$$  

We can define the length of the path $\mathcal{C}$ as

$$\Delta s = \int_a^b d\lambda ds = \int_a^b d\lambda \sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}.$$  

This definition corresponds, in infinitesimal form, to $ds = \sqrt{g_{\mu\nu} dx^\mu dx^\nu}$.

In other words, if we have a curve, characterized, in a given coordinate system, by the functions $\{x^\mu(\lambda)\}$, and then by the tangent vector

$$U^\mu = \frac{dx^\mu}{d\lambda},$$

the measure element on the curve $ds/d\lambda$ (which, integrated in $d\lambda$, gives the length of the path) is

$$\frac{ds}{d\lambda} = \sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} = \sqrt{g_{\mu\nu} U^\mu U^\nu}.$$  

This can be expressed in a coordinate-independent way:

$$\frac{ds}{d\lambda} = \sqrt{g(U, U)}.$$  

Note that if we change coordinate system, $\{x^\mu\} \to \{x'^\alpha\}$, the quantity (4.54) does not change. Furthermore, if we change the parametrization of the curve, $\lambda \to \lambda' = \lambda'(\lambda)$, the new measure element is

$$\lambda \to \lambda' = \lambda'(\lambda),$$

the new measure element is

$$\frac{ds}{d\lambda'} = \sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda'} \frac{dx^\nu}{d\lambda'}} = \sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \frac{d\lambda}{d\lambda'} \frac{d\lambda'}{d\lambda}} = \frac{ds}{d\lambda} \frac{d\lambda'}{d\lambda}.$$  

and

$$\Delta s = \int_a^b d\lambda ds = \int_a^b \left( \frac{d\lambda'}{d\lambda} \frac{ds}{d\lambda} \right) d\lambda = \int_a^{b'} d\lambda' ds = \int_a^{b'} d\lambda' \frac{ds}{d\lambda'}.$$  

Therefore, $\Delta s$ does not depend on the parametrization, and is a characteristic of the path, given the metric, not of the curve.
4.3.2 The metric tensor maps vectors into one-forms

As we have seen, the metric tensor is a linear function of two vectors: this means that it takes two vectors and associates a number to them. The number is their scalar product.

But now suppose that we write \( g(\boldsymbol{\cdot}, \tilde{\nu}) \), namely we leave the first slot empty. What is this? We know that if we fill the first slot with a generic vector \( \tilde{\nu} \) we will get a number, thus \( g(\cdot, \tilde{\nu}) \) must be a linear function of a generic vector that we can put in the empty slot, and that associates a number to this vector.

But this is the definition of one-forms! Thus \( g(\cdot, \tilde{\nu}) \) is a one-form.

In addition, it is a particular one-form because it depends on \( \tilde{\nu} \): if we change \( \tilde{\nu} \), the one-form will be different. Let us indicate this one-form as

\[
\tilde{V} = \tilde{V}. \tag{4.57}
\]

By definition the components of \( \tilde{V} \) are

\[
V_\alpha = \tilde{V}(\tilde{\nu}_\alpha) = g(\tilde{\nu}_\alpha, \tilde{V}) = g(\tilde{\nu}_\alpha, V^\beta \tilde{\nu}_\beta) = V^\beta g(\tilde{\nu}_\alpha, \tilde{\nu}_\beta) = V^\beta g_{\alpha\beta},
\]

hence

\[
V_\alpha = g_{\alpha\beta} V^\beta. \tag{4.58}
\]

Thus the tensor \( g \) associates to any vector \( \tilde{\nu} \) a one-form \( \tilde{V} \), dual of \( \tilde{V} \), whose components can be computed if we know \( g_{\alpha\beta} \) and \( V^\alpha \).

In addition, if we multiply eq. (4.58) by \( g^{\alpha\gamma} \), where \( g^{\alpha\gamma} \) is the matrix inverse to \( g_{\alpha\gamma} \)

\[
g_{\alpha\gamma} g^{\gamma\beta} = \delta_\alpha^\beta, \tag{4.59}
\]

we find

\[
g^{\alpha\gamma} V_\alpha = g^{\alpha\gamma} g_{\alpha\beta} V^\beta = \delta_\beta^\gamma V^\beta = V^\gamma,
\]

i.e.

\[
V^\gamma = g^{\alpha\gamma} V_\alpha. \tag{4.60}
\]

Consequently the metric tensor also maps one-forms into vectors. In a similar way the metric tensor can map a \( \begin{pmatrix} 2 \\ 0 \end{pmatrix} \) tensor in a \( \begin{pmatrix} 1 \\ 1 \end{pmatrix} \) tensor

\[
A^{\alpha\beta} = g_{\beta\gamma} A^{\alpha\gamma},
\]

or in a \( \begin{pmatrix} 0 \\ 2 \end{pmatrix} \) tensor

\[
A_{\alpha\beta} = g_{\alpha\mu} g_{\beta\nu} A^{\mu\nu},
\]

or vice versa

\[
A^{\alpha\beta} = g^{\alpha\mu} g^{\beta\nu} A_{\mu\nu}.
\]

These maps are called index raising and lowering.

Summarizing, the metric tensor

1) allows to compute the inner product of two vectors \( g(\tilde{A}, \tilde{B}) = \tilde{A} \cdot \tilde{B} \), and consequently
the norm of a vector \( g(\vec{A}, \vec{A}) = \vec{A} \cdot \vec{A} = A^2 \).

2) As a consequence it allows to compute the distance between two points
\[ ds^2 = g(d\vec{s}, d\vec{s}) = g_{\alpha\beta} dx^\alpha dx^\beta. \]

3) It maps one-forms into vectors and vice versa.

4) It allows to raise and lower indices.
Chapter 5

Affine Connections and Parallel Transport

In chapter 1 we showed that there are two quantities that describe the effects of a gravitational field on moving bodies by virtue of the Equivalence Principle: the metric tensor and the affine connections. In chapter 4 we discussed the geometrical properties of the metric tensor. In this chapter we shall define the affine connections as the quantities that allow to compute the derivative of a vector in an arbitrary space, and we shall show that they coincide with the \( \Gamma \)’s introduced in chapter 1.

5.1 The covariant derivative of vectors

Let us consider a vector (field) \( \vec{V} = V^\mu \vec{e}_\mu \). The derivative of \( \vec{V} \) is

\[
\frac{\partial \vec{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta} \vec{e}_\alpha + V^\alpha \frac{\partial \vec{e}_\alpha}{\partial x^\beta}.
\]  

(5.1)

The first term on the right-hand side is a linear combination of the basis vectors, therefore it is a vector and we know how to compute it. The second term involves the derivative of the basis vectors, for which we need to compute the quantities \( \vec{e}_\alpha(p') - \vec{e}_\alpha(p) \), i.e. to subtract vectors which are applied in different points of the manifold \( M \). Note that the vectors \( \vec{e}_\alpha(p) \) and \( \vec{e}_\alpha(p') \) belong to the tangent space to \( M \), respectively, in \( p \) and \( p' \), and that \( T_p \neq T_{p'} \). Thus, to define the derivative of a vector field on a manifold, we need to specify a rule to compare vectors belonging to different tangent spaces; such a rule is called a connection.

Let us start considering Minkowski’s spacetime, where it is possible to define a global coordinate system \((ct, x, y, z)\) which covers the entire spacetime; at any given point \( p \) of the manifold there exists the coordinate basis \( \vec{e}_{M(\alpha)}(p) \) which belongs to the tangent space \( T_p \). In this case a simple rule to compare vectors on different tangent spaces is to impose that each basis vector in a point \( p \) of the manifold is equal to the corresponding basis vector in any other point \( p' \), i.e.

\[
\vec{e}_{M(\alpha)}(p) = \vec{e}_{M(\alpha)}(p').
\]  

(5.2)
This rule is the affine connection in Minkowski’s spacetime. Note that, with this choice the basis vectors of the Minkowskian frame are, by definition, constant:

$$\frac{\partial \vec{e}_{M(\alpha)}}{\partial x^\beta} = \vec{0}. \tag{5.3}$$

Let us now consider a general spacetime. The equivalence principle tells us that at any point of the manifold we can choose a locally inertial frame, in which the laws of physics are (locally) those of Special Relativity. Thus, the natural choice for the affine connection in a general spacetime is the following: we impose that in a locally inertial frame the basis vectors are constant. We shall now show that, using this rule, we will be able to compute the derivative of a vector 5.1 at a given point \( p \) of the manifold.

Let us make a coordinate transformation to the local inertial frame in \( p \), introducing the new basis vectors \( \vec{e}_{M'(\alpha')} \), related to the old basis vectors \( \vec{e}_{M(\alpha)} \) by the equation

$$\vec{e}_{M(\alpha)} = \Lambda^\alpha_{\alpha'} \vec{e}_{M'(\alpha')} \tag{5.4}$$

From (5.3) we know that the vectors \( \vec{e}_{M'(\alpha')} \) are constant. Consequently

$$\frac{\partial \vec{e}_{(\alpha)}}{\partial x^\beta} = \left( \frac{\partial}{\partial x^\beta} \Lambda^\alpha_{\alpha'} \right) \vec{e}_{M'(\alpha')}. \tag{5.5}$$

The R.H.S. of (5.5) is a linear combination of the basis vectors \( \{ \vec{e}_{M'(\alpha')} \} \), therefore it is a vector.

Since \( \frac{\partial \vec{e}_{(\alpha)}}{\partial x^\beta} \) is a vector, we must be able to express it as a linear combination of the basis vectors \( \{ \vec{e}_{(\mu)} \} \) we are working with, i.e.:

$$\frac{\partial \vec{e}_{(\alpha)}}{\partial x^\beta} = \Gamma^\mu_{\alpha\beta} \vec{e}_{(\mu)}, \tag{5.6}$$

where the constants \( \Gamma^\mu_{\alpha\beta} \) have three indices because \( \alpha \) indicates which basis vector \( \vec{e}_{(\alpha)} \) we are differentiating, and \( \beta \) indicates the coordinate with respect to which the differentiation is performed. The \( \Gamma^\mu_{\beta\alpha} \) are called affine connection or Christoffel symbols. Note that in the case of Minkowski space, the basis vectors in the Minkowskian frame are constant, thus \( \Gamma^\mu_{\alpha\beta} = 0 \).

Thus, coming back to eq. (5.1), the derivative of \( \vec{V} \) becomes

$$\frac{\partial \vec{V}}{\partial x^\beta} = \frac{\partial V^\alpha}{\partial x^\beta} \vec{e}_{(\alpha)} + V^\alpha \Gamma^\mu_{\beta\alpha} \vec{e}_{(\mu)},$$

or relabelling the dummy indices

$$\frac{\partial \vec{V}}{\partial x^\beta} = \left[ \frac{\partial V^\alpha}{\partial x^\beta} + V^\sigma \Gamma^\alpha_{\beta\sigma} \right] \vec{e}_{(\alpha)}. \tag{5.7}$$

For any fixed \( \beta \), \( \frac{\partial \vec{V}}{\partial x^\beta} \) is a vector field because it is a linear combination of the basis vectors \( \{ \vec{e}_{(\alpha)} \} \) with coefficients \( \left[ \frac{\partial V^\alpha}{\partial x^\beta} + V^\sigma \Gamma^\alpha_{\beta\sigma} \right] \).
If we introduce the following notation
\[ V^\alpha;_\beta = \frac{\partial V^\alpha}{\partial x^\beta}, \quad \text{and} \quad V^\alpha;_\beta = \frac{\partial V^\alpha}{\partial x^\beta} + V^\mu \Gamma^\alpha_{\beta \mu}, \] (5.8)
eq (5.7) becomes
\[ \frac{\partial \vec{V}}{\partial x^\beta} = V^\alpha;_\beta \vec{e}_{(\alpha)}. \] (5.9)

5.1.1 \( V^\alpha;_\beta \) are the components of a tensor

Let us define the following quantity:
\[ \nabla \vec{V} = \left[ V^\alpha;_\beta \vec{e}_{(\alpha)} \right] \otimes \tilde{\omega}^{(\beta)}. \] (5.10)

As shown in section 5.1, for any fixed value of \( \beta \) the quantity \( \left[ V^\alpha;_\beta \vec{e}_{(\alpha)} \right] \) is a vector; thus, \( \nabla \vec{V} \) defined in eq. (5.10) is the outer product between these vectors and the basis one-forms, i.e. it is a \( (1,1) \) tensor. This tensor field is called Covariant derivative of a vector, and its components are
\[ (\nabla \vec{V})^\alpha_{\beta} \equiv \nabla_\beta V^\alpha \equiv V^\alpha;_\beta. \] (5.11)

NOTE THAT
In a locally inertial frame the basis vectors are constant, and consequently, according to eq. (5.6) the affine connections vanish and from eq. (5.8) it follows that
\[ V^\alpha;_\beta = V^\alpha;_\beta \implies \frac{\partial \vec{V}}{\partial x^\beta} = V^\alpha;_\beta \vec{e}_{(\alpha)}. \] (5.12)
Thus, in a locally inertial frame covariant and ordinary derivative coincide.

5.2 The covariant derivative of one-forms and tensors

In order to find the covariant derivative of a one-form consider a scalar \( \Phi \) field. At any point of space it is a number, therefore it does not depend on the coordinate basis: this implies that ordinary and covariant derivative coincide
\[ \nabla_\mu \Phi = \frac{\partial \Phi}{\partial x^\mu} = (\tilde{d}\Phi)_\mu. \] (5.13)

Now remember the definition of a one-form: it is a linear function that takes a vector and associates to it a number according to the rule
\[ \tilde{q}(\vec{V}) = q^\alpha V^\alpha, \] (5.14)
where \( q^\alpha \) and \( V^\alpha \) are the components of the one-form and of the vector. We can therefore assume that the scalar function \( \Phi \) is
\[ \Phi = q^\alpha V^\alpha, \] (5.15)
and consequently the covariant derivative will be

$$\nabla_\mu \Phi \equiv \frac{\partial \Phi}{\partial x^\mu} = \frac{\partial q_\alpha}{\partial x^\mu} V^\alpha + q_\alpha \frac{\partial V^\alpha}{\partial x^\mu}.$$  

Substituting $\frac{\partial V^\alpha}{\partial x^\mu}$ from eq. (5.8) we find

$$\nabla_\mu \Phi = \frac{\partial q_\alpha}{\partial x^\mu} V^\alpha + q_\alpha [V^\alpha ;_\mu - V^\beta \Gamma^\alpha_{\mu \beta}].$$

We relabel the indices to put $V^\alpha$ in evidence

$$\nabla_\mu \Phi = \frac{\partial q_\alpha}{\partial x^\mu} V^\alpha + q_\sigma V^\alpha ;_\mu = \frac{\partial q_\alpha}{\partial x^\mu} - q_\sigma \Gamma^\alpha_{\sigma \mu} ]V^\alpha + q_\sigma V^\sigma ;_\mu. \quad (5.16)$$

$\nabla_\mu \Phi$ are the components of a tensor ( $\nabla_\mu \Phi \equiv \tilde{d} \Phi_\mu$ ) as well as $V^\alpha$, $q_\beta$ and $V^\alpha ;_\mu$  

1. Therefore in order this equation to be true, also the terms in parenthesis must be the components of a tensor. We call covariant derivative of the one-form $\tilde{q}$ a \( \begin{pmatrix} 0 \\ 2 \end{pmatrix} \) tensor $\nabla \tilde{q}$ whose components are

$$\left( \nabla \tilde{q} \right)_{\alpha \beta} \equiv \nabla_\beta q_\alpha = q_{\alpha, \beta} = q_{\alpha, \beta} - q_\sigma \Gamma^\sigma_{\beta \alpha}.$$  

(5.17)

We now observe that eq. (5.16) can be written as

$$\nabla_\mu \Phi = \nabla_\mu (q_\alpha V^\alpha) = q_{\alpha, \mu} V^\alpha + q_\alpha V^\alpha ;_\mu. \quad (5.18)$$

From this equation it follows that the covariant derivative satisfies the usual property of a derivative of a product.

The same procedure that led to define the covariant derivative of one-forms can be used to define the covariant derivative of \( \begin{pmatrix} N \\ N' \end{pmatrix} \) tensors.

(do it as an exercise)

$$\left( \nabla T^\mu_{\nu, \beta} \right) \beta = T^\mu_{\nu, \beta} - T^\alpha_{\nu \beta} \Gamma^\alpha_{\beta \mu} - T^\mu_{\alpha \beta} \Gamma^\alpha_{\nu} \quad (5.19)$$

$$\left( \nabla A^\mu_{\nu, \beta} \right) \beta = A^\mu_{\nu, \beta} + A^\alpha_{\nu} \Gamma^\mu_{\alpha \beta} + A^\mu_{\alpha} \Gamma^\nu_{\alpha \beta} \quad (5.20)$$

$$\left( \nabla B^\mu_{\nu} \right) \beta = B^\mu_{\nu, \beta} + B^\alpha_{\nu} \Gamma^\mu_{\beta \alpha} - B^\mu_{\alpha} \Gamma^\nu_{\beta \alpha} \quad (5.21)$$

what is the rule?

### 5.3 The covariant derivative of the metric tensor

The covariant derivative of $g_{\mu \nu}$ is zero

$$g_{\mu \nu ;\alpha} = 0.$$

---

1. here we use the word ‘tensor’ to indicate any type of tensors, included vectors and one-forms.
The reason is the following. We know from the principle of equivalence that at each point of spacetime we can choose a coordinate system such that $g_{\mu\nu}$ reduces to $\eta_{\mu\nu}$. The coordinate basis associated to these coordinates has constant basis vectors, therefore the affine connections also vanish (see eq. 5.6). In this frame

$$g_{\alpha\beta;\mu} = \eta_{\alpha\beta;\mu} = \frac{\partial \eta_{\alpha\beta}}{\partial x^\mu} - \Gamma_{\alpha\mu}^\nu \eta_{\nu\beta} - \Gamma_{\beta\mu}^\nu \eta_{\alpha\nu} = 0$$

$g_{\alpha\beta;\mu}$ is a $\begin{pmatrix} 0 & 3 \end{pmatrix}$ tensor, and if all components of a tensor are zero in a coordinate system, they are zero in any coordinate system therefore

$$g_{\alpha\beta;\mu} = 0 \quad (5.22)$$

always.

### 5.4 Symmetries of the affine connections

Consider an arbitrary scalar field $\Phi$. Its first covariant derivative is a one-form and coincides with the ordinary derivative. Its second covariant derivative $\nabla \nabla \Phi$ is a $\begin{pmatrix} 0 & 2 \end{pmatrix}$ tensor of components $\Phi_{,\beta;\alpha}$. In minkowskian coordinates, i.e. in a locally inertial frame, covariant derivative reduces to ordinary derivative:

$$\Phi_{,\beta;\alpha} = \Phi_{,\beta,\alpha} = \frac{\partial}{\partial x^\alpha} \frac{\partial}{\partial x^\beta} \Phi, \quad (5.23)$$

and since partial derivatives commute

$$\Phi_{,\beta;\alpha} = \Phi_{,\alpha;\beta} \Rightarrow \Phi_{,\beta;\alpha} = \Phi_{,\alpha;\beta}. \quad (5.24)$$

Thus, the tensor $\nabla \nabla \Phi$ is symmetric. But if a tensor is symmetric in one basis, it is symmetric in any basis, therefore

$$\Phi_{,\beta;\alpha} - \Phi_{,\mu} \Gamma_{\beta\alpha}^\mu = \Phi_{,\alpha;\beta} - \Phi_{,\mu} \Gamma_{\alpha\beta}^\mu$$

in any coordinate system. It follows that for any $\Phi$

$$\Phi_{,\mu} \Gamma_{\beta\alpha}^\mu = \Phi_{,\mu} \Gamma_{\alpha\beta}^\mu$$

and consequently

$$\Gamma_{\beta\alpha}^\mu = \Gamma_{\alpha\beta}^\mu \quad (5.25)$$

in any coordinate system.
5.5 The relation between the affine connections and the metric tensor

From eq. (5.22) it follows that

$$g_{\alpha\beta;\mu} = g_{\alpha\beta,\mu} - \Gamma^\nu_{\alpha\mu} g_{\nu\beta} - \Gamma^\nu_{\beta\mu} g_{\alpha\nu} = 0,$$

therefore

$$g_{\alpha\beta,\mu} = \Gamma^\nu_{\alpha\mu} g_{\nu\beta} + \Gamma^\nu_{\beta\mu} g_{\alpha\nu}. \quad (5.26)$$

Let us now consider the following equations

$$g_{\alpha\mu,\beta} = \Gamma^\nu_{\alpha\beta} g_{\nu\mu} + \Gamma^\nu_{\mu\beta} g_{\alpha\nu},$$

$$-g_{\beta\mu,\alpha} = -\Gamma^\nu_{\beta\alpha} g_{\nu\mu} - \Gamma^\nu_{\mu\alpha} g_{\beta\nu},$$

It follows that

$$g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha} = (\Gamma^\nu_{\alpha\mu} - \Gamma^\nu_{\mu\alpha}) g_{\nu\beta} +$$

$$+ (\Gamma^\nu_{\beta\mu} + \Gamma^\nu_{\mu\beta}) g_{\alpha\nu} + (\Gamma^\nu_{\alpha\beta} - \Gamma^\nu_{\beta\alpha}) g_{\nu\mu},$$

where we have used $g_{\alpha\beta} = g_{\beta\alpha}$.

Since $\Gamma^\alpha_{\beta\gamma}$ are symmetric in $\beta$ and $\gamma$, it follows that

$$g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha} = 2\Gamma^\nu_{\beta\mu} g_{\alpha\nu}. \quad (5.27)$$

If we multiply by $g^{\alpha\gamma}$ and remember that since $g^{\alpha\gamma}$ is the inverse of $g_{\alpha\gamma}$

$$g^{\alpha\gamma} g_{\alpha\nu} = \delta^\gamma_\nu,$$

we finally find

$$\Gamma^\gamma_{\beta\mu} = \frac{1}{2} g^{\alpha\gamma} (g_{\alpha\beta,\mu} + g_{\alpha\mu,\beta} - g_{\beta\mu,\alpha}) \quad (5.28)$$

This expression is extremely useful, since it allows to compute the affine connection in terms of the components of the metric.

Are the $\Gamma^\alpha_{\beta\gamma}$ components of a tensor?

They are not, and it is easy to see why. In a locally inertial frame the $\Gamma^\alpha_{\beta\gamma}$ vanish, but in any other frame they don’t. If it would be a tensor they should vanish in any frame.

In the first chapter we defined the Christoffel symbols as

$$\Gamma^\alpha_{\mu\nu} = \frac{\partial x^\alpha}{\partial \xi^\mu} \frac{\partial^2 \xi^\lambda}{\partial x^\nu \partial x^\mu}. \quad (5.29)$$

This definition was a consequence of the equivalence principle. We did the following: We considered a free particle in a locally inertial frame $\{\xi^\alpha\}$:

$$\frac{d^2 \xi^\alpha}{dt^2} = 0. \quad (5.29)$$
Then we transformed this equation to an arbitrary coordinate system \( \{ x^\alpha \} \) and we showed that it becomes

\[
\frac{d^2 x^\alpha}{d\tau^2} + \Gamma^\alpha_{\mu \nu} \left[ \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \right] = 0,
\]

with \( \Gamma^\alpha_{\mu \nu} \) defined in eq. (5.28).

In this chapter we have defined the \( \Gamma \)'s as those functions that satisfy the equation

\[
\frac{\partial \bar{e}_{(\mu)}}{\partial x^\nu} = \Gamma^\alpha_{\mu \nu} \bar{e}_{M(\alpha)}.
\]

In a local inertial frame \( \{ \xi^\alpha \} \) be \( \bar{e}_{M(\alpha)} \) the constant basis vectors. If we make a coordinate transformation to a new coordinate system \( \{ x^\alpha \} \), the new basis \( \{ \bar{e}_{(\mu')} \} \) will be

\[
\bar{e}_{(\mu')} = \Lambda^\alpha_{\mu} \bar{e}_{M(\alpha)} = \frac{\partial \xi^\alpha}{\partial x^\mu'} \bar{e}_{M(\alpha)}.
\]

In this frame, eq. (5.31) which defines the affine connections can be re-written as

\[
\frac{\partial}{\partial x^\nu} \left[ \Lambda^\beta_{\mu \nu} \bar{e}_{M(\beta)} \right] = \Gamma^\mu'_{\mu \nu} \Lambda^\gamma_{\alpha \nu} \bar{e}_{M(\gamma)}
\]

or, being the \( \bar{e}_{M(\beta)} \) constant

\[
\frac{\partial \Lambda^\beta_{\mu \nu}}{\partial x^\nu} \bar{e}_{M(\beta)} = \Gamma^\mu'_{\mu \nu} \Lambda^\gamma_{\alpha \nu} \bar{e}_{M(\gamma)}.
\]

This equation can be re-written as

\[
\left( \frac{\partial \Lambda^\beta_{\mu \nu}}{\partial x^\nu} - \Gamma^\alpha_{\mu \nu} \Lambda^\beta_{\alpha} \right) \bar{e}_{M(\beta)} = 0.
\]

We now multiply eq. (5.35) by \( \Lambda^\sigma_{\beta} \) and find

\[
\Lambda^\sigma_{\beta} \frac{\partial \Lambda^\beta_{\mu \nu}}{\partial x^\nu} - \Gamma^\alpha_{\mu \nu} \Lambda^\sigma_{\beta} \Lambda^\beta_{\alpha} = 0.
\]

Since \( \Lambda^\sigma_{\beta} \Lambda^\beta_{\alpha} = \delta^\sigma_{\alpha} \), it follows that

\[
\Gamma^\sigma_{\mu \nu} = \Lambda^\sigma_{\beta} \frac{\partial \Lambda^\beta_{\mu \nu}}{\partial x^\nu} = \frac{\partial x^\sigma}{\partial \xi^\beta} \frac{\partial^2 \xi^\beta}{\partial x^\nu \partial x^\mu'}
\]

which coincides with eq. (5.28). Thus, as expected, the two definitions are equivalent. How do the \( \Gamma^\beta_{\gamma} \) transform?

The easiest way to see it is from the definition (5.28). In an arbitrary coordinate system \( \{ x^\mu' \} \) they are

\[
\Gamma^{\nu'}_{\mu' \nu} = \frac{\partial x^\nu}{\partial \xi^\alpha} \frac{\partial^2 \xi^\alpha}{\partial x^\mu' \partial x^\nu'} = \\
= \frac{\partial x^\nu}{\partial x^\rho} \frac{\partial}{\partial x^\rho} \left( \frac{\partial \xi^\alpha}{\partial x^\sigma} \frac{\partial x^\sigma}{\partial x^\nu'} \right) = \\
= \frac{\partial x^\nu}{\partial x^\rho} \frac{\partial}{\partial \xi^\alpha} \left[ \frac{\partial x^\sigma}{\partial x^\rho} \frac{\partial^2 \xi^\alpha}{\partial x^\nu' \partial x^\mu'} + \frac{\partial \xi^\alpha}{\partial x^\sigma} \frac{\partial^2 x^\sigma}{\partial x^\nu' \partial x^\mu'} \right] = \\
= \frac{\partial x^\nu}{\partial x^\rho} \frac{\partial}{\partial x^\rho} \frac{\partial x^\sigma}{\partial x^\mu'} \frac{\partial x^\tau}{\partial x^\mu'} \Gamma^\rho_{\sigma} + \frac{\partial x^\nu}{\partial x^\rho} \frac{\partial^2 x^\sigma}{\partial x^\rho \partial x^\mu'} \Gamma^\rho_{\sigma}.
\]
CHAPTER 5. AFFINE CONNECTIONS AND PARALLEL TRANSPORT

The first term is what we should get if \( \Gamma_{\beta \gamma}^{\alpha} \) were a tensor. But we know it is not, and in fact there is an additional term.

5.6 Non coordinate basis

In Sec. 3.4 we have seen that if we pass from Minkowskian coordinates \( \{x^\alpha\} \equiv (ct, x, y) \) to polar coordinates \( \{x'^\alpha\} \equiv (ct, r, \theta) \) the coordinate basis

\[
\{\vec{e}(\alpha)\} \rightarrow \begin{cases} 
\vec{e}(0) & \rightarrow (1, 0, 0) \\
\vec{e}(1) & \rightarrow (0, 1, 0) \\
\vec{e}(2) & \rightarrow (0, 0, 1) 
\end{cases}
\]  

transforms to \( \{\vec{e}'(\alpha')\} \)

\[
\begin{cases} 
\vec{e}'(0') = \vec{e}(0) \\
\vec{e}'(1') = \vec{e}(r) = \cos \theta \vec{e}(1) + \sin \theta \vec{e}(2) \\
\vec{e}'(2') = \vec{e}(\theta) = -r \sin \theta \vec{e}(1) + r \cos \theta \vec{e}(2) 
\end{cases}
\]

according to the law

\[ \vec{e}'(\alpha') = \Lambda^\mu_{\alpha'} \vec{e}(\mu). \]

The new basis is a coordinate basis and the matrix \( \Lambda^\mu_{\alpha'} = \frac{\partial x^\mu}{\partial x'^{\alpha'}} \) is the matrix associated to the coordinate transformation. However we may choose a different basis for vectors. For example the vectors \( \{\vec{e}'(\alpha')\} \) in the previous example are not normalized. In fact

\[ \vec{e}'(\alpha') \cdot \vec{e}'(\beta') = g_{\alpha'\beta'} = \begin{pmatrix} -1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & r^2 \end{pmatrix} \neq \eta_{\alpha'\beta'}. \]

We may decide that we want a basis composed by unit vectors, and choose

\[
\begin{cases} 
\vec{e}_r = \vec{e}_r \\
\vec{e}_t = \vec{e}_t \\
\vec{e}_\theta = \frac{1}{r} \vec{e}_\theta 
\end{cases}
\]

In this case we would find

\[ \vec{e}_{(\hat{\alpha})} \cdot \vec{e}_{(\hat{\beta})} = \eta_{\hat{\alpha}\hat{\beta}}. \]

But now the question is: do there exist coordinates \( \{x^{\hat{\alpha}}\} \) such that

\[ e_{(\hat{\alpha})} = \Lambda^\mu_{\hat{\alpha}} \vec{e}(\mu) = \frac{\partial x^\mu}{\partial x^{\hat{\alpha}}} \vec{e}(\mu). \]

so that the basis \( \{\vec{e}_{(\hat{\alpha})}\} \) is a coordinate basis? Alternatively, we can formulate the same question for the basis one-forms: if \( \{\tilde{\omega}^{(\alpha')}\} \) is the coordinate basis for one-forms and \( \{\omega^{(\hat{\alpha})}\} \) is the normalized basis, is \( \{\tilde{\omega}^{(\hat{\alpha})}\} \) a new coordinate basis associated to some coordinates \( \{x^{\hat{\alpha}}\} \) ? i.e.

\[ \tilde{\omega}^{(\hat{\alpha})} = \Lambda^{\hat{\beta}}_{\alpha} \tilde{\omega}^{(\beta)} = \frac{\partial x^{\hat{\alpha}}}{\partial x^{\beta}} \tilde{\omega}^{(\beta)}? \]
For instance, in the previous example,
\[ \tilde{\omega}^1 = \tilde{\omega}^\rho = \tilde{\omega}^r = \cos \theta \, \tilde{d}x + \sin \theta \, \tilde{d}y \]
\[ \tilde{\omega}^2 = \tilde{\omega}^\theta = r \, \tilde{\omega}^\theta = - \sin \theta \, \tilde{d}x + \cos \theta \, \tilde{d}y \]  \hspace{1cm} (5.41)

The point is that if this is true, \( \Lambda^\alpha_{\beta} \) must coincide with the partial derivative \( \frac{\partial \hat{x}^\alpha}{\partial x^\beta} \), and consequently the following condition must be satisfied for any \( \Lambda^\alpha_{\gamma} \):
\[ \frac{\partial}{\partial x^\gamma} \Lambda^\alpha_{\beta} = \partial^2 x^\alpha_{\partial x^\gamma \partial x^\beta} = \partial^2 x^\alpha_{\partial x^\beta \partial x^\gamma} = \frac{\partial}{\partial x^\gamma} \Lambda^\alpha_{\gamma}. \]  \hspace{1cm} (5.42)

This is an “integrability condition” that all the components of \( \Lambda^\alpha_{\gamma} \) must satisfy in order the coordinates \( \{ x^\alpha \} \) do exist.

For example, let us check whether the basis (5.41) is a coordinate basis. From the expression of \( \tilde{\omega}^\theta \) we find that
\[ \Lambda^2_{\hat{1}} = \frac{\partial x^2}{\partial x} = - \sin \theta \quad \Lambda^2_{\hat{2}} = \frac{\partial x^2}{\partial y} = \cos \theta, \]

eq. (5.42) gives
\[ \frac{\partial}{\partial y} \Lambda^2_{\hat{1}} = \frac{\partial}{\partial x} \Lambda^2_{\hat{2}} \Rightarrow \frac{\partial}{\partial y} (- \sin \theta) = \frac{\partial}{\partial x} (\cos \theta), \]

But
\[ x = r \cos \theta \quad y = r \sin \theta \quad r = \sqrt{x^2 + y^2}, \]

so that it should be
\[ \frac{\partial}{\partial y} \left[ - \frac{y}{\sqrt{x^2 + y^2}} \right] = \frac{\partial}{\partial x} \left[ \frac{y}{\sqrt{x^2 + y^2}} \right], \]

which is certainly not true.

We conclude that the basis \( \{ \tilde{\omega}^{(\alpha)} \} \) is not a coordinate basis, since we cannot associate to it a coordinate transformation.

What are the consequences of choosing a noncoordinate basis?

As we have seen at the end of section 3.5, the gradient of a scalar field \( \Phi \) is a one-form:
\[ \tilde{d} \Phi \rightarrow \left\{ \frac{\partial \Phi}{\partial x^\alpha} \right\} \equiv \{ \Phi_{,\alpha} \}. \]  \hspace{1cm} (5.43)

For example let us start in a 2-dimensional plane with coordinates \( (x, y) = (x^1, x^2) \). Then change to polar coordinates \( (r, \theta) = (x^1, x^2) \). The gradient will transform as one-forms do:
\[ \tilde{d} \Phi_{,\alpha} = \Lambda^\beta_{,\alpha} \tilde{d} \Phi_{,\beta} \]

where \( \tilde{d} \Phi_{,x} = \Phi_{,x} = \frac{\partial \Phi}{\partial x} \) and \( \tilde{d} \Phi_{,y} = \Phi_{,y} = \frac{\partial \Phi}{\partial y} \).

The components of the gradient in the new coordinate basis are
\[ \tilde{d} \Phi_r = \Lambda^x_{,r} \tilde{d} \Phi_x + \Lambda^y_{,r} \tilde{d} \Phi_y = \frac{\partial x}{\partial r} \tilde{d} \Phi_x + \frac{\partial y}{\partial r} \tilde{d} \Phi_y \]
\[ \tilde{d} \Phi_\theta = \Lambda^x_{,\theta} \tilde{d} \Phi_x + \Lambda^y_{,\theta} \tilde{d} \Phi_y = \frac{\partial x}{\partial \theta} \tilde{d} \Phi_x + \frac{\partial y}{\partial \theta} \tilde{d} \Phi_y. \]  \hspace{1cm} (5.44)
Being

\begin{align*}
x &= r \cos \theta, \\
y &= r \sin \theta,
\end{align*}

\begin{align*}
\hat{d}\Phi_r &= \cos \theta \hat{d}\Phi_x + \sin \theta \hat{d}\Phi_y = \frac{\partial \Phi}{\partial r} = \Phi_r, \\
\hat{d}\Phi_\theta &= -r \sin \theta \hat{d}\Phi_x + r \cos \theta \hat{d}\Phi_y = \frac{\partial \Phi}{\partial \theta} = \Phi_\theta.
\end{align*}

(5.45)

Thus the components of the gradient in the new coordinate basis \((\vec{e}_r, \vec{e}_\theta)\) will still be

\[\tilde{\partial} \Phi_j \rightarrow \frac{\partial \Phi}{\partial x^j}.\]

But this is certainly non true if we use the non coordinate basis \(\{\vec{e}(\hat{\alpha})\}\): there are no coordinates associated to this basis, thus we cannot define \(\tilde{d}\Phi_j = \frac{\partial \Phi}{\partial x^j}\)!

Let us see what happens to the affine connections if we use a non-coordinate basis. We have defined \(\Gamma^\alpha_{\beta\gamma}\) as

\[\nabla_\alpha \vec{e}(\beta) = \frac{\partial \vec{e}(\beta)}{\partial x^\alpha} = \Gamma^\nu_{\beta\alpha} \vec{e}(\nu).\]

(5.46)

This is a definition valid in any basis, therefore in terms of a noncoordinate basis \(\{\vec{e}(\hat{\alpha})\}\) eq. (5.46) becomes

\[\nabla_\alpha \vec{e}(\hat{\beta}) = \Gamma^{\hat{\nu}}_{\hat{\beta} \hat{\alpha}} \vec{e}(\nu).\]

(5.47)

But now, since the \(\{x^\hat{\alpha}\}\) do not exist, is not longer true that

\[\Phi_{,\hat{\beta} \hat{\alpha}} = \Phi_{,\hat{\alpha} \hat{\beta}}.\]

If we go back to eq.(5.24) we see that we used this condition to show the simmetry of the affine connection in the two lower indices. Thus if the basis is a non coordinate basis

\[\Gamma^{\hat{\alpha}}_{\hat{\beta} \hat{\gamma}} \neq \Gamma^{\hat{\alpha}}_{\hat{\gamma} \hat{\beta}},\]

and moreover eq (5.27) which gives the connections in terms of \(g_{\alpha\beta}\) is no longer true as well.

In the following of this course we shall use mainly coordinate basis, and we shall explicitly specify when we will use a non coordinate basis.

---

**EXERCISE**

In this chapter we have introduced the connections as those quantities that allow to find the covariant derivative of a vector in an arbitrary frame. Given the metric components, the simplest way to compute the connection is to use eq. (5.27). As an exercise, let us compute the connection \(\Gamma^\mu_{\alpha\beta}\) in a different way, using directly the definition

\[\frac{\partial \vec{e}(\alpha)}{\partial x^\beta} = \Gamma^\mu_{\alpha\beta} \vec{e}(\mu).\]

(5.48)
CHAPTER 5. AFFINE CONNECTIONS AND PARALLEL TRANSPORT

Let us consider for example a 2-dimensional flat space in polar coordinates, i.e. \((x^1', x^2') \equiv (r, \theta)\) and remember that the basis vectors are related to the coordinate basis associated to cartesian coordinates by the equations (3.50)

\[
\vec{e}_{(1')} = \cos \theta \vec{e}_{(1)} + \sin \theta \vec{e}_{(2)} \\
\vec{e}_{(2')} = -r \sin \theta \vec{e}_{(1)} + r \cos \theta \vec{e}_{(2)}.
\]

Let us indicate, for simplicity \((\vec{e}_{(1)}, \vec{e}_{(2)})\) with \((\vec{e}_x, \vec{e}_y)\), and \((\vec{e}_{(1')}, \vec{e}_{(2')})\) with \((\vec{e}_r, \vec{e}_\theta)\). From these expressions we find

\[
\frac{\partial \vec{e}_r}{\partial r} = \frac{\partial}{\partial r} (\cos \theta \vec{e}_x + \sin \theta \vec{e}_y) = 0,
\]

and consequently

\[
\Gamma^\mu_{rr} \vec{e}_\mu = \Gamma^r_{rr} \vec{e}_r + \Gamma^\theta_{rr} \vec{e}_\theta = 0 \implies \Gamma^r_{rr} = \Gamma^\theta_{rr} = 0.
\]

Moreover

\[
\frac{\partial \vec{e}_\theta}{\partial \theta} = \frac{\partial}{\partial \theta} (\cos \theta \vec{e}_x + \sin \theta \vec{e}_y) =
\]

\[
= -\sin \theta \vec{e}_x + \cos \theta \vec{e}_y - \frac{1}{r} \vec{e}_\theta,
\]

therefore

\[
\frac{1}{r} \vec{e}_\theta = \Gamma^\mu_{\theta\mu} \vec{e}_\mu = \Gamma^r_{\theta\mu} \vec{e}_r + \Gamma^\theta_{\theta\mu} \vec{e}_\theta \implies \Gamma^r_{\theta\theta} = 0, \quad \Gamma^\theta_{r\theta} = \frac{1}{r}.
\]

Proceeding along these lines one can show that

\[
\Gamma^r_{\theta r} = 0, \quad \Gamma^\theta_{r \theta} = \frac{1}{r}, \quad \Gamma^\theta_{\theta \theta} = -r, \quad \Gamma^\theta_{\theta \theta} = 0.
\]

It should be noted that although we have used the cartesian basis to express \(\vec{e}_r\) and \(\vec{e}_\theta\) and compute their derivatives, at the end the \(\Gamma\)'s depend only on the coordinates \(r\) and \(\theta\). Note also that the same result can be obtained by using eq. (5.27) and the metric

\[
g_{\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}.
\]

5.7 Summary of the proceeding Sections

In chapter 1 we have seen that the equation of motion of a particle which moves under the exclusive action of a gravitational field is

\[
\frac{d^2 x^\alpha}{d\tau^2} + \Gamma^\alpha_{\mu\nu} \left[ \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \right] = 0.
\]

(5.49)

In the frame associated to the coordinates \(\{x^\mu\}\) the line element is

\[
ds^2 = g_{\mu\nu} dx^\mu dx^\nu.
\]

(5.50)
Then we have seen that the Equivalence Principle allows to find a locally inertial frame \( \{ \xi^\alpha \} \) where eq. (5.49) becomes

\[
\frac{d^2 \xi^\alpha}{d\tau^2} = 0,
\]

(5.51)

and the line element reduces to

\[
\frac{ds^2}{\eta_{\mu\nu} dx^\mu dx^\nu}.
\]

(5.52)

However we do not know if this transformation holds everywhere, i.e. if the spacetime is really flat, or if it holds only locally, which would mean that there is a non constant and non uniform gravitational field. It follows that the study of the motion of a single particle and the knowledge of the \( \Gamma^\alpha_{\mu\nu} \)'s do not allow to decide whether there is a non constant and non uniform gravitational field.

Then we have introduced vectors and tensors on a manifold, we have defined the metric tensor as a geometric object and we have shown that its role is not only that of defining the distance between points, but also that of mapping vectors into one-forms, and of computing the scalar product between vectors. We have shown that if we introduce at each point of the manifold a basis for vectors \( \{ \vec{e}^{(\alpha)} \} \) (and a dual basis for one forms \( \{ \tilde{\omega}^{(\beta)} \} \) ) any vector (or one-form) can be assigned “components” with respect to the basis

\[
\vec{A} = A^\alpha \vec{e}^{(\alpha)}.
\]

(5.53)

Then we have introduced an operator of covariant derivative, which generates a tensor according to the following rule

\[
\nabla_\beta V^\alpha = V^{\alpha,\beta} + \Gamma^\alpha_{\mu\beta} V^\mu.
\]

(5.54)

(and similar rules for tensors). The covariant derivative coincides with ordinary derivative in two particular cases:

1) the spacetime is flat and we are using a basis where the vectors \( \vec{e}^{(\alpha)} \) are constant. Consequently from the definition (5.6) it follows that \( \Gamma^\alpha_{\mu\beta} = 0 \).

2) the spacetime is curved, but we are in a locally inertial frame. Indeed, in this frame eq. (5.49) reduces to eq. (5.51), which means again that \( \Gamma^\alpha_{\mu\beta} = 0 \).

The fact that we can always find a frame where \( g_{\mu\nu} \) reduces to \( \eta_{\mu\nu} \) and the \( \Gamma^\alpha_{\mu\beta} = 0 \) (and consequently the first derivatives of \( g_{\mu\nu} \) vanish) implies that in order to know if we are in the presence of a gravitational field, (i.e. if the spacetime is curved), we need to know the second derivatives of the metric tensor \( g_{\mu\nu,\alpha,\beta} \). This result should not be surprising: in chapter 1 we introduced the 2-dimensional Gaussian geometry and we said that one can always choose a frame where the metric looks flat, but there exists a quantity, the Gaussian curvature, which tells us that the space is curved. The gaussian curvature depends on the first derivatives (non linearly) and on the second derivatives (linearly) of the metric; thus, we shall now look for a generalization of the Gaussian curvature. We already mentioned that in four dimensions we need more than one invariant to describe the intrinsic properties of a curved surface: we need six functions, and it is clear that a vector would not be enough. Thus, we need a tensor, but which tensor? The only thing we know is that it should contain the second derivatives of \( g_{\mu\nu} \). In order to introduce the curvature tensor we first need to introduce the notion of parallel transport of a vector along a curve.
5.8 Parallel Transport

In chapter 1 we discussed and compared the intrinsic geometry of cones, cylinders and spheres, and we noticed that while it is flat for cones and cylinders, it is curved for spheres. That means, for example, that two lines which start parallel do not remain parallel when prolonged:

Consider two segments in A and B, perpendicular to the equator, i.e. parallel.

The same lines when prolonged: they do not remain parallel.

It is also interesting to see what happens when we parallely transport a vector along a path. Parallel Transport means that for each infinitesimal displacement, the displaced vector must be parallel to the original one, and must have the same length. Let us consider first the case when the path belongs to a flat space.

a) FLAT SPACE
When we return to A the displaced vector coincides with the original vector in A.

b) ON A SPHERE
(remember that the vector must always be tangent to the sphere)

When the vector goes back to A it is rotated of 90 degrees This is a consequence of the curvature of the sphere.

On a curved manifold it is impossible to define a globally parallel vector field. The parallel transport of a vector depends on the path along which it is transported.

Let us now compute how does a vector change when it is paralllely transported. Consider a curve of parameter $\lambda$ and a vector field $\vec{V}$ defined at every point of the curve. Be $\vec{U} \to \left\{ \frac{dx^\mu}{d\lambda} \right\}$ the vector tangent to the curve
At every point of the curve we can choose a locally inertial frame \( \{ \xi^\alpha \} \). In this frame, if we move \( \vec{V} \) along the curve of an infinitesimal \( d\lambda \), parallel to itself and keeping its length unchanged, its components do not change

\[
\frac{dV^\alpha}{d\lambda} = 0.
\]

But

\[
\frac{dV^\alpha}{d\lambda} = \frac{\partial V^\alpha}{\partial \xi^\beta} \frac{d\xi^\beta}{d\lambda} = U^\beta V'^{\alpha,\beta} = 0.
\]

Since we are in a locally inertial frame, ordinary and covariant derivative coincide and therefore we can write

\[
U^\beta V'^{\alpha,\beta} = 0.
\]

If this equation is true in a locally inertial frame, since it is a tensor equation it must be true in any other frame. Therefore eq. (5.57) is the frame-invariant definition of the parallel transport of \( \vec{V} \) along the curve identified by the tangent vector \( \vec{U} \).

Eq. (5.57) is written in terms of the components of \( \vec{V} \) and \( \vec{U} \); if we want to write it in a frame-independent form we shall write

\[
\nabla_{\vec{U}} \vec{V} = 0,
\]

which means that the covariant derivative along the direction of the vector \( \vec{U} \) is zero. Written explicitly for a generic reference frame with coordinates \( \{ x^\alpha \} \) eq. (5.58) gives

\[
\left( \nabla_{\vec{U}} \vec{V} \right)^\alpha = U^\beta V'^{\alpha,\beta}
\]

\[
= \frac{dV^\alpha}{d\lambda} \left[ \frac{\partial V^\alpha}{\partial x^\beta} + \Gamma^\alpha_{\beta\nu} V^\nu \right] = \frac{dV^\alpha}{d\lambda} + \Gamma^\alpha_{\beta\nu} V^\nu U^\beta = 0.
\]

Thus, contrary to what happens in flat space the components of a vector parallely transported along a curve in curved space do change, and the change is given by

\[
\frac{dV^\alpha}{d\lambda} = -\Gamma^\alpha_{\beta\nu} V^\nu U^\beta.
\]

5.9 The geodesic equation

In Chapter 1 we introduced the geodesics, as the curves which describe the motion of free particles; “free” here means that no other force than gravity is acting on them. We showed that they are the solution of the geodesic equation (1.37)

\[
\frac{d^2 x^\alpha}{d\tau^2} + \Gamma^\alpha_{\beta\nu} \left[ \frac{dx^\mu}{d\tau} \frac{dx^\beta}{d\tau} \right] = 0.
\]

A different derivation of this equation, simpler than that given in Chapter 1, makes use of the notion of covariant derivative. Let us consider a “free” particle, with worldline \( x^\mu(\tau) \) and four-velocity (i.e. tangent vector to the worldline) \( U^\mu = dx^\mu/d\tau \). By the equivalence
principle, at any point of the worldline we can define a locally inertial frame \( \{ x^\alpha' \} \), in which the laws of special relativity hold; then, in this frame the particle four-acceleration is zero, i.e.

\[
\frac{dU^\mu'}{d\tau} = \frac{dx^\alpha'}{d\tau} \frac{\partial U^\mu'}{\partial x^\alpha} = U^\alpha' U^\mu',_{\alpha'} = 0 .
\] (5.61)

In a locally inertial frame ordinary and covariant derivative coincide, thus

\[
U^\alpha' U^\mu',_{\alpha'} = 0 .
\] (5.62)

This is a tensorial equation, and the covariance principle establishes that it holds in any coordinate frame; therefore, in a generic frame we can write

\[
U^\alpha U^\mu,_{\alpha} = 0 .
\] (5.63)

Equations (5.63) and (5.60) coincide; indeed

\[
U^\alpha U^\mu,_{\alpha} = U^\alpha U^\mu,_{\alpha} + U^\alpha \Gamma^\mu_{\alpha\beta} U^\beta ,
\] (5.64)

and by substituting \( U^\mu = dx^\mu / d\tau \) this equation becomes

\[
\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} = 0 ,
\] (5.65)

which is eq. (5.60).

The parameter along a geodesic need not to be the proper time. Be \( s \) the new parameter chosen to parametrize the geodesic. Since

\[
\frac{d}{d\tau} = \frac{d}{ds} \frac{ds}{d\tau} ,
\] (5.66)

equation (5.65) becomes

\[
\frac{d^2 x^\alpha}{ds^2} + \Gamma^\alpha_{\mu\nu} \left[ \frac{dx^\mu}{ds} \frac{dx^\nu}{ds} \right] = - \left[ \frac{d^2 s}{d\tau^2} / \left( \frac{ds}{d\tau} \right)^2 \right] \frac{dx^\alpha}{ds} ;
\] (5.67)

From this equation we see that the new curve is a geodesic, i.e. has the form of equation (5.65), only if the new parameter is related to the proper time \( \tau \) by a linear transformations

\[
s = a\lambda + b , \quad a, b = \text{const} ;
\] (5.68)
in which case the right hand side of equation (5.67) vanishes. \( \tau \) and \( s \) are called affine parameters.

Equation (5.63) was derived assuming that the geodesic was the worldline of a massive particle, i.e a timelike curve. However, this equation has a more general validity, since a geodesic can be either timelike, spacelike or null. If a geodesic is timelike, i.e. \( \vec{U} \cdot \vec{U} < 0 \), it can represent the wordline of a massive particle; in this case, by performing the linear
transformation (5.68) it is possible to change the affine parameter in such a way that \( \vec{U} \cdot \vec{U} = -1 \), so that the new parameter is the particle proper time.

If, instead, a geodesic is a null curve, i.e. \( \vec{U} \cdot \vec{U} = 0 \), it can represent the wordline of a massless particle; in this case the affine parameter is a generic parameter, since proper time is not defined for massless particles.

If the geodesic is spacelike, i.e. \( U \cdot \vec{U} > 0 \), it does not represent the worldline of a particle of any kind.

According to the equation of parallel transport (5.57), the geodesic equation written in the form (5.63) is the equation of the parallel transport of the tangent vector \( \vec{U} \) along the geodesic. This means that if we take the tangent vector at a point \( p \), and parallelly transport it to a point \( p' \) along the geodesic line, the transported vector is tangent to the curve at \( p' \). Thus, a curve \( C \) with tangent vector \( \vec{U} \) is a geodesic if

\[
\nabla_\xi \vec{U} = 0.
\]

(5.69)

For this reason we say that: geodesics are those curves which parallel-transport their own tangent vectors.
Chapter 6

The Curvature Tensor

We are now in a position to introduce the curvature tensor. We will do it in two different ways.

6.1 a) A Formal Approach

Let us start writing the transformation rule for affine connections

\[ \Gamma^\lambda_{\mu\nu} = \frac{\partial x^\lambda}{\partial x^{\alpha'}} \frac{\partial x^{\alpha'}}{\partial x^\rho'} \Gamma^\rho_{\sigma\tau} + \frac{\partial x^\lambda}{\partial x^{\alpha'}} \frac{\partial^2 x^{\alpha'}}{\partial x^\rho' \partial x^\nu}. \]  

(6.1)

As we already noticed (Chapter V sec. 5) if the last term on the right-hand side would be zero \( \Gamma^\lambda_{\mu\nu} \) would transform as a tensor. Let us isolate the ‘bad term’, by multiplying eq. (6.1) by \( \frac{\partial x^\tau'}{\partial x^\lambda} \):

\[ \frac{\partial^2 x^{\tau'}}{\partial x^\rho' \partial x^\nu} = \frac{\partial x^{\tau'}}{\partial x^\lambda} \Gamma^\lambda_{\mu\nu} - \frac{\partial x^\rho'}{\partial x^\mu} \frac{\partial x^{\sigma'}}{\partial x^\nu} \Gamma^\tau_{\rho\sigma}. \]  

(6.2)

We now differentiate this equation with respect to \( x^\kappa \):

\[ \frac{\partial^3 x^{\tau'}}{\partial x^\kappa \partial x^\mu \partial x^\nu} = \frac{\partial^2 x^{\tau'}}{\partial x^\kappa \partial x^\lambda} \Gamma^\lambda_{\mu\nu} + \frac{\partial x^{\tau'}}{\partial x^\lambda} \left( \frac{\partial}{\partial x^\kappa} \Gamma^\lambda_{\mu\nu} \right) 
- \frac{\partial^2 x^\rho'}{\partial x^\mu \partial x^\nu} \frac{\partial x^{\sigma'}}{\partial x^\nu} \Gamma^\tau_{\rho\sigma} - \frac{\partial x^\rho'}{\partial x^\mu} \frac{\partial^2 x^{\sigma'}}{\partial x^\nu \partial x^\kappa} \Gamma^\tau_{\rho\sigma} - \frac{\partial x^\rho'}{\partial x^\mu} \frac{\partial x^{\sigma'}}{\partial x^\nu} \left( \frac{\partial}{\partial x^\kappa} \Gamma^\tau_{\rho\sigma} \right). \]

(6.3)

We now use eq. (6.2):

\[ \frac{\partial^3 x^{\tau'}}{\partial x^\kappa \partial x^\mu \partial x^\nu} = \frac{\partial^2 x^{\tau'}}{\partial x^\kappa \partial x^\lambda} \Gamma^\lambda_{\mu\nu} + \frac{\partial x^{\tau'}}{\partial x^\lambda} \left( \frac{\partial}{\partial x^\kappa} \Gamma^\lambda_{\mu\nu} \right) 
+ \Gamma^\lambda_{\mu\nu} \left[ \frac{\partial x^{\tau'}}{\partial x^\kappa} \Gamma^\kappa_{\alpha\lambda} - \frac{\partial x^\beta'}{\partial x^\alpha} \Gamma^\tau_{\beta\gamma} \frac{\partial x^\gamma}{\partial x^\kappa} \right] + \frac{\partial x^{\tau'}}{\partial x^\kappa} \left[ \frac{\partial}{\partial x^\kappa} \Gamma^\lambda_{\mu\nu} \right] 
- \frac{\partial x^{\sigma'}}{\partial x^\nu} \Gamma^\tau_{\rho\sigma} \left[ \frac{\partial x^{\rho'}}{\partial x^\alpha} \Gamma^\kappa_{\alpha\mu} - \frac{\partial x^\beta'}{\partial x^\alpha} \Gamma^\tau_{\beta\gamma} \frac{\partial x^\gamma}{\partial x^\mu} \right]. \]

(6.4)
CHAPTER 6. THE CURVATURE TENSOR

\[ -\frac{\partial x'^\rho}{\partial x^\mu} \Gamma^\nu_{\rho \sigma} \left[ \frac{\partial x'^\alpha}{\partial x^\kappa} \Gamma^\gamma_{\alpha \nu} - \frac{\partial x'^\beta}{\partial x^\kappa} \frac{\partial x'^\gamma}{\partial x^\nu} \Gamma^\alpha_{\beta \gamma} \right] \]

\[ + \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} \left( \frac{\partial}{\partial x^\kappa} \Gamma^\nu_{\rho \sigma} \right). \]

Let us rewrite the last term as

\[ \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} \left( \frac{\partial}{\partial x^\kappa} \Gamma^\nu_{\rho \sigma} \right). \tag{6.5} \]

(The reason is that the indices of \( \Gamma \) have a prime, thus the derivatives must be computed with respect to the \( \{x'^{\sigma}\} \)). We now rewrite eq. (6.5) in the following way

\[
\frac{\partial^3 x'^\tau}{\partial x^\kappa \partial x^\mu \partial x^\nu} = \frac{\partial x'^\tau}{\partial x^\lambda} \left( \frac{\partial}{\partial x^\kappa} \Gamma^\mu_{\lambda \nu} \right) + \left( \frac{\partial x'^\tau}{\partial x^\gamma} \Gamma^\mu_{\kappa \lambda} \Gamma^\lambda_{\mu \nu} \right)
\]

\[ - \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} \left( \frac{\partial}{\partial x^\kappa} \Gamma^\nu_{\rho \sigma} \right) - \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\sigma}{\partial x^\nu} \frac{\partial x'^\gamma}{\partial x^\nu} \Gamma^\tau_{\rho \sigma} \Gamma^\nu_{\beta \gamma} \]

\[ - \left[ \frac{\partial x'^\sigma}{\partial x^\nu} \Gamma^\tau_{\rho \sigma} \frac{\partial x'^\alpha}{\partial x^\kappa} \Gamma^\lambda_{\nu \mu} \Gamma^\nu_{\beta \gamma} \right]. \]

We now relabel the indices in the following way

\[ \frac{\partial x'^\tau}{\partial x^\alpha} \Gamma^\alpha_{\kappa \lambda} \Gamma^\lambda_{\mu \nu} - \frac{\partial x'^\tau}{\partial x^\gamma} \Gamma^\mu_{\kappa \lambda} \Gamma^\lambda_{\mu \nu} \tag{6.6} \]

\[ \frac{\partial x'^\tau}{\partial x^\nu} \frac{\partial x'^\alpha}{\partial x^\kappa} \frac{\partial x'^\beta}{\partial x^\mu} \frac{\partial x'^\gamma}{\partial x^\nu} \Gamma^\tau_{\rho \sigma} \Gamma^\rho_{\beta \gamma} \rightarrow \frac{\partial x'^\tau}{\partial x^\nu} \frac{\partial x'^\sigma}{\partial x^\kappa} \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\gamma}{\partial x^\nu} \Gamma^\tau_{\rho \sigma} \Gamma^\rho_{\beta \gamma} \Gamma^\lambda_{\mu \nu} \]

\[ \frac{\partial x'^\tau}{\partial x^\nu} \frac{\partial x'^\alpha}{\partial x^\kappa} \frac{\partial x'^\beta}{\partial x^\mu} \frac{\partial x'^\gamma}{\partial x^\nu} \Gamma^\tau_{\rho \sigma} \Gamma^\rho_{\beta \gamma} \rightarrow \frac{\partial x'^\tau}{\partial x^\nu} \frac{\partial x'^\sigma}{\partial x^\kappa} \frac{\partial x'^\alpha}{\partial x^\mu} \frac{\partial x'^\gamma}{\partial x^\nu} \Gamma^\tau_{\rho \sigma} \Gamma^\rho_{\beta \gamma} \Gamma^\lambda_{\mu \nu} \]

With these changes the terms can be collected in the following way

\[ \frac{\partial^3 x'^\tau}{\partial x^\kappa \partial x^\mu \partial x^\nu} = \frac{\partial x'^\tau}{\partial x^\lambda} \left( \frac{\partial}{\partial x^\kappa} \Gamma^\lambda_{\mu \nu} \right) + \Gamma^\lambda_{\kappa \eta} \Gamma^\mu_{\eta \nu} \tag{6.8} \]

\[ - \frac{\partial x'^\rho}{\partial x^\mu} \frac{\partial x'^\tau}{\partial x^\nu} \frac{\partial x'^\sigma}{\partial x^\kappa} \left( \frac{\partial}{\partial x^\nu} \Gamma^\tau_{\rho \sigma} \right) - \Gamma^\tau_{\lambda \sigma \tau} \Gamma^\mu_{\eta \rho} - \Gamma^\tau_{\rho \lambda} \Gamma^\lambda_{\mu \nu} \]

\[ - \frac{\partial x'^\rho}{\partial x^\mu} \Gamma^\nu_{\rho \sigma} \left[ \Gamma^\lambda_{\kappa \mu} \frac{\partial x'^\sigma}{\partial x^\nu} + \Gamma^\lambda_{\kappa \nu} \frac{\partial x'^\rho}{\partial x^\mu} + \Gamma^\lambda_{\mu \nu} \frac{\partial x'^\rho}{\partial x^\kappa} \right]. \]
We now subtract from this expression the same expression with \( \kappa \) and \( \nu \) interchanged

\[
\frac{\partial^3 x^\tau}{\partial x^\kappa \partial x^\mu \partial x^\nu} - \frac{\partial^3 x^\tau}{\partial x^\nu \partial x^\kappa \partial x^\mu} = 0 = (6.9)
\]

\[
\frac{\partial x^\tau}{\partial x^\lambda} \left[ \frac{\partial}{\partial x^\kappa} \Gamma^\lambda_{\mu \nu} + \Gamma^\lambda_{\kappa \eta} \Gamma^\eta_{\mu \nu} \right]
\]

\[
- \frac{\partial x^\tau}{\partial x^\mu} \frac{\partial x^\sigma}{\partial x^\nu} \frac{\partial x^\rho}{\partial x^\kappa} \left[ \frac{\partial}{\partial x^\sigma} \Gamma^\tau_{\rho \sigma} - \frac{\partial}{\partial x^\rho} \Gamma^\tau_{\sigma \rho} + \Gamma^\tau_{\lambda \rho} \Gamma^\lambda_{\sigma \rho} - \Gamma^\tau_{\lambda \sigma} \Gamma^\lambda_{\rho \rho} \right]
\]

collecting all terms we find

\[
\frac{\partial x^\tau}{\partial x^\lambda} \left[ \frac{\partial}{\partial x^\kappa} \Gamma^\lambda_{\mu \nu} - \frac{\partial}{\partial x^\mu} \Gamma^\lambda_{\kappa \nu} \Gamma^\eta_{\mu \nu} - \Gamma^\lambda_{\nu \eta} \Gamma^\eta_{\mu \kappa} \right] - \frac{\partial x^\tau}{\partial x^\mu} \frac{\partial x^\sigma}{\partial x^\nu} \frac{\partial x^\rho}{\partial x^\kappa} \left[ \frac{\partial}{\partial x^\sigma} \Gamma^\tau_{\rho \sigma} - \frac{\partial}{\partial x^\rho} \Gamma^\tau_{\sigma \rho} + \Gamma^\tau_{\lambda \rho} \Gamma^\lambda_{\sigma \rho} - \Gamma^\tau_{\lambda \sigma} \Gamma^\lambda_{\rho \rho} \right] = 0.
\]

If we now define the following

\[
R^\lambda_{\mu \nu \kappa} = - \left[ \frac{\partial}{\partial x^\kappa} \Gamma^\lambda_{\mu \nu} - \frac{\partial}{\partial x^\mu} \Gamma^\lambda_{\kappa \nu} \Gamma^\eta_{\mu \nu} - \Gamma^\lambda_{\nu \eta} \Gamma^\eta_{\mu \kappa} \right], \quad (6.10)
\]

we can write eq. (6.10) as the transformation law for the tensor

\[
R^\sigma_{\alpha \beta \gamma \tau} = \frac{\partial x^\sigma}{\partial x^\lambda} \frac{\partial x^\tau}{\partial x^\alpha} \frac{\partial x^\nu}{\partial x^\beta} \frac{\partial x^\kappa}{\partial x^\gamma} R^\lambda_{\mu \nu \kappa}. \quad (6.12)
\]

The tensor (6.11) is the \textbf{Curvature Tensor}, also called the \textbf{Riemann Tensor}, and it can be shown that it is the only tensor that can be constructed by using the metric, its first and second derivatives, and which is linear in the second derivatives.

This way of defining the Riemann tensor is the “old-fashioned way”: it is based on the transformation properties of the affine connections. The idea underlying this derivation is that the information about the curvature of the space must be contained in the second derivative of the metric, and therefore in the first derivative of the \( \Gamma^\alpha_{\mu \nu} \). But since we want to find a tensor out of them, we must eliminate in eq. (6.1) the part which does not transform as a tensor, and we do this in eq. (6.9).

\footnote{The - sign does not agree with the definition given in Weinberg, but it does agree with the definition given in many other textbooks. As we shall see in the next section it is irrelevant. What is important is to write the Einstein equations with the right signs!}
6.2 b) The curvature tensor and the curvature of the spacetime

We shall now rederive the curvature tensor in a different way that explicitly shows why it expresses the curvature of a spacetime. This derivation, due to Levi Civita, will use the notion of parallel transport of a vector along a closed loop.

Consider a closed loop whose four sides are the coordinate lines \( x^1 = a \), \( x^1 = a + \delta a \), \( x^2 = b \), \( x^2 = b + \delta b \).

Take a generic vector \( \vec{V} \) and parallelly transport \( \vec{V} \) along AB, i.e. consider \( \nabla_{\vec{e}_1} \vec{V} = 0 \). From eq. (5.58) it follows that

\[
\epsilon^\mu_{(1)} V^\alpha;\mu = 0. \tag{6.13}
\]

Since \( \vec{e}_1 \) has only \( \epsilon^1_{(1)} \neq 0 \) then

\[
\frac{\partial V^\alpha}{\partial x^1} + \Gamma^\alpha_{\beta 1} V^\beta = 0. \tag{6.14}
\]

This equation can be integrated along the line AB:

\[
\delta V^\alpha_{AB} = - \int_{A(x^2=b)}^{B} \Gamma^\alpha_{\beta 1} V^\beta dx^1. \tag{6.15}
\]

In a similar way, if we go from B to C along the line \( x^1 = a + \delta a \)

\[
\frac{\partial V^\alpha}{\partial x^2} = - \Gamma^\alpha_{\beta 2} V^\beta \rightarrow \delta V^\alpha_{BC} = - \int_{B(x^1=a+\delta a)}^{C} \Gamma^\alpha_{\beta 2} V^\beta dx^2. \tag{6.16}
\]

From C to D

\[
\frac{\partial V^\alpha}{\partial x^1} = - \Gamma^\alpha_{\beta 1} V^\beta \rightarrow \delta V^\alpha_{CD} = - \int_{C(x^2=b+\delta b)}^{D} \Gamma^\alpha_{\beta 1} V^\beta dx^1, \tag{6.17}
\]

and from D back to A

\[
\frac{\partial V^\alpha}{\partial x^2} = - \Gamma^\alpha_{\beta 2} V^\beta \rightarrow \delta V^\alpha_{DA} = - \int_{D(x^1=a)}^{A} \Gamma^\alpha_{\beta 2} V^\beta dx^2. \tag{6.18}
\]
The change in $\vec{V}$ due to this parallel transport will be a vector $\delta \vec{V}$ whose components can be found by adding eqs. (6.15)-(6.18):

$$\delta V^\alpha = - \int_{D(x^1=a)}^A \Gamma^\alpha_{\beta 2} V^\beta \, dx^2$$
$$- \int_{B(x^1=a+\delta a)}^C \Gamma^\alpha_{\beta 2} V^\beta \, dx^2 - \int_{C(x^2=b+\delta b)}^D \Gamma^\alpha_{\beta 1} V^\beta \, dx^1$$
$$- \int_{A(x^2=b)}^B \Gamma^\alpha_{\beta 1} V^\beta \, dx^1.$$

If the spacetime is flat, $V^\alpha$ do not change when the vector is parallelly transported, i.e. $\delta V^\alpha = 0$. But in curved spacetime $\delta V^\alpha$ will in general be different from zero.

If we consider an infinitesimal loop, i.e. $\delta a$ and $\delta b$ tend to zero, we can take an expansion of eq. (6.19) to first order in $\delta a$ and $\delta b$:

$$\delta V^\alpha \simeq - \int_{A(x^2=b)}^B \Gamma^\alpha_{\beta 1} V^\beta \, dx^1 - \left[ \int_{B(x^1=a)}^C \Gamma^\alpha_{\beta 2} V^\beta \, dx^2 + \frac{\partial}{\partial x^1} \left( \int_{B}^C \Gamma^\alpha_{\beta 2} V^\beta \, dx^2 \right) \delta a \right]$$
$$- \left[ \int_{D(x^2=b)}^C \Gamma^\alpha_{\beta 1} V^\beta \, dx^1 + \frac{\partial}{\partial x^2} \left( \int_{C}^D \Gamma^\alpha_{\beta 1} V^\beta \, dx^1 \right) \delta b \right]$$
$$- \int_{D(x^1=a)}^A \Gamma^\alpha_{\beta 2} V^\beta \, dx^2,$$

Since

$$A = (a, b), \quad C = (a + \delta a, b + \delta b), \quad B = (a + \delta a, b), \quad \text{and} \quad D = (a, b + \delta b),$$

the previous equation becomes

$$\delta V^\alpha \simeq - \int_{a}^{a+\delta a} \Gamma^\alpha_{\beta 1} V^\beta \, dx^1$$
$$- \int_{b}^{b+\delta b} \Gamma^\alpha_{\beta 2} V^\beta \, dx^2 - \left[ \int_{b}^{b+\delta b} \frac{\partial}{\partial x^1} \left( \Gamma^\alpha_{\beta 2} V^\beta \right) \, dx^2 \right] \delta a$$
$$+ \int_{a}^{a+\delta a} \Gamma^\alpha_{\beta 1} V^\beta \, dx^1 + \left[ \int_{a}^{a+\delta a} \frac{\partial}{\partial x^2} \left( \Gamma^\alpha_{\beta 1} V^\beta \right) \, dx^1 \right] \delta b$$
$$+ \int_{b}^{b+\delta b} \Gamma^\alpha_{\beta 2} V^\beta \, dx^2,$$

i.e.

$$\delta V^\alpha \simeq - \delta a \int_{b}^{b+\delta b} \frac{\partial}{\partial x^1} \left( \Gamma^\alpha_{\beta 2} V^\beta \right) \, dx^2$$
$$+ \delta b \int_{a}^{a+\delta a} \frac{\partial}{\partial x^2} \left( \Gamma^\alpha_{\beta 1} V^\beta \right) \, dx^1 \simeq \delta a \delta b \left[ - \frac{\partial}{\partial x^1} \left( \Gamma^\alpha_{\beta 2} V^\beta \right) + \frac{\partial}{\partial x^2} \left( \Gamma^\alpha_{\beta 1} V^\beta \right) \right].$$
Eq. (6.23) can be further developed by using eq. (6.14)

$$\frac{\partial V^\kappa}{\partial x^1} = -\Gamma^\kappa_{\beta_1} V^\beta, \quad \frac{\partial V^\kappa}{\partial x^2} = -\Gamma^\kappa_{\beta_2} V^\beta; \quad (6.24)$$

it becomes

$$\delta V^\alpha = \delta a \delta b \left[ \frac{\partial \Gamma^\alpha_{\beta_1}}{\partial x^2} V^\beta + \Gamma^\alpha_{\kappa_1} \frac{\partial V^\kappa}{\partial x^2} - \frac{\partial \Gamma^\alpha_{\beta_2}}{\partial x^1} V^\beta - \Gamma^\alpha_{\kappa_2} \frac{\partial V^\kappa}{\partial x^1} \right] (6.25)$$

Note that:

- $\delta a$ and $\delta b$ are the non-vanishing components of the displacement vectors $\delta \vec{x}^{(1)}$ and $\delta \vec{x}^{(2)}$ along the direction of the basis vectors $\vec{e}^{(1)}$ and $\vec{e}^{(2)}$, i.e.

$$\delta \vec{x}^{(1)} = \delta a \vec{e}^{(1)}, \quad \delta \vec{x}^{(2)} = \delta a \vec{e}^{(2)} \quad (6.26)$$

whose components in the basis $\{\vec{e}^{(\alpha)}\}$ are

$$\delta x^{\mu}_{(1)} = (0, \delta a, 0, 0) = \delta a \delta^\mu_1, \quad (6.27)$$

$$\delta x^{\mu}_{(2)} = (0, 0, \delta b, 0) = \delta b \delta^\mu_2.$$

Thus, we can write eq. (6.25) as follows

$$\delta V^\alpha = \delta \vec{x}^{(1)} \cdot \delta \vec{x}^{\mu \mu \mu} \left[ \frac{\partial \Gamma^\alpha_{\beta_\nu \mu}}{\partial x^\mu} - \frac{\partial \Gamma^\alpha_{\beta_\mu \nu}}{\partial x^\mu} - \Gamma^\alpha_{\kappa_\nu} \Gamma^\kappa_{\beta_\mu} + \Gamma^\alpha_{\kappa_\mu} \Gamma^\kappa_{\beta_\nu} \right] V^\beta. \quad (6.28)$$

- The term in square brackets is the curvature tensor which we have already defined in eq. (6.11): $R^\alpha_{\beta \mu \nu} = \Gamma^\alpha_{\beta_\nu \mu} - \Gamma^\alpha_{\beta_\mu \nu} - \Gamma^\alpha_{\kappa_\nu} \Gamma^\kappa_{\beta_\mu} + \Gamma^\alpha_{\kappa_\mu} \Gamma^\kappa_{\beta_\nu}. \quad (6.29)$

Note that it is antisymmetric in $\nu$ and $\mu$; indeed, it must be because, if we interchange $\delta \vec{x}^{(1)}$ and $\delta \vec{x}^{(2)}$, $\delta V^\alpha$ changes sign, because we would go around the loop in the opposite direction. This shows that the sign of (6.29) can be chosen arbitrarily, and this is the reason why the definitions of the Riemann tensor given in textbooks may differ for a sign.

We have already shown that the object given in eq. (6.29) is a tensor, by looking at the way it transforms under a coordinate transformation (eq. 6.12). However, we want to see if it also agrees with the definition of tensors given in chapter 4. Let us contract eq. (6.28) with $V^\alpha$.

$$\delta V^\alpha V^\alpha = \delta x^{\nu \nu \nu \mu \mu} \left[ \frac{\partial \Gamma^\alpha_{\beta_\nu}}{\partial x^\mu} - \frac{\partial \Gamma^\alpha_{\beta_\mu}}{\partial x^\nu} - \Gamma^\alpha_{\kappa_\nu} \Gamma^\kappa_{\beta_\mu} + \Gamma^\alpha_{\kappa_\mu} \Gamma^\kappa_{\beta_\nu} \right] V^\beta V^\alpha. \quad (6.30)$$

The result of this contraction is, of course, a number. On the right-hand side there are the components of 3 vectors i.e.: $\delta \vec{x}^{(1)}$, $\delta \vec{x}^{(2)}$ and $V^\beta$. Moreover there are the components of the one-form $V^\alpha$. The four geometrical objects (three vectors and one one-form) are contracted
with the quantity within brackets, and the result is a number. In addition, we note that (6.30) is linear in \(V^\beta, \delta x^{(1)}_\alpha \delta x^{(2)}_\mu\). For instance, if we consider a displacement \(\delta x^{(1a)}_\nu + \delta x^{(1b)}_\nu\) along \(\vec{e}_{(1)}\) it is immediate to see that
\[
\delta V^\alpha V_\alpha = \delta x^{(1)}_\nu \delta x^{(2)}_\mu [\ldots] V^\beta V_\alpha + \delta x^{(1b)}_\nu \delta x^{(2)}_\mu [\ldots] V^\beta V_\alpha ,
\]
and similarly for the other quantities. If we consider a generic \(13\) tensor, \(T^\alpha_{\beta\gamma\delta}\), since by definition it is a linear function of one one-form and three vectors, when supplied with these arguments (for example the one-form \(\tilde{V}\), and the three vectors \(\vec{V}, \delta x^{(1)}\) and \(\delta x^{(2)}\) it will produce the following number
\[
T(\tilde{V}, \vec{V}, \delta x^{(1)}, \delta x^{(2)}) = T^\alpha_{\beta\rho\delta} V_\alpha V^\beta \delta x^{\rho}_{(1)} \delta x^{\delta}_{(2)}.
\]
Eq. (6.32) has the same structure of eq. (6.30). Therefore we are entitled to define the components of the Riemann tensor as in eq. (6.29).

It should now be clear why the Riemann tensor deserves its name of **Curvature Tensor**: it tells us how does a vector change when it is parallely transported along a loop, due to the curvature of the spacetime. If the spacetime is flat
\[
\delta V^\alpha = 0 \quad \text{along any closed loop} \quad \rightarrow \quad R^\alpha_{\beta\gamma\delta} = 0,
\]
in any reference frame. Indeed, if a tensor vanishes in a given frame, then it vanishes in any other frame.

The components of the Riemann tensor assume a very nice form when computed in a locally inertial frame:
\[
R^\alpha_{\beta\mu\nu} = \frac{1}{2} g^{\alpha\sigma} \left[ g_{\sigma\nu,\beta\mu} - g_{\sigma\mu,\beta\nu} + g_{\beta\mu,\sigma\nu} - g_{\beta\nu,\sigma\mu} \right],
\]
or lowering the index \(\alpha\)
\[
R_{\alpha\beta\mu\nu} = g_{\alpha\lambda} R^\lambda_{\beta\mu\nu} = \frac{1}{2} \left[ g_{\alpha\nu,\beta\mu} - g_{\alpha\mu,\beta\nu} + g_{\beta\mu,\alpha\nu} - g_{\beta\nu,\alpha\mu} \right].
\]
It should be stressed that

1) The Riemann tensor is linear in the second derivatives of \(g_{\mu\nu}\), and non linear in the first derivatives.

2) In a locally inertial frame the \(\Gamma^\alpha_{\nu\sigma}\) vanish and therefore the non-linear part of the Riemann tensor vanishes as well.

### 6.3 Symmetries

From eq. (6.35) it is easy to verify that
\[
R_{\alpha\beta\mu\nu} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu} = R_{\mu\nu\alpha\beta},
\]
\[
R_{\alpha\beta\mu\nu} + R_{\alpha\nu\beta\mu} + R_{\mu\alpha\nu\beta} = 0.
\]
Since \(R_{\alpha\beta\mu\nu}\) is a tensor, these symmetry properties are valid in any reference frame. The symmetries of the Riemann tensor reduce the number of independent components to 20.
6.4 The Riemann tensor gives the commutator of covariant derivatives

Let us consider the second covariant derivatives of a vector field $\vec{V}$

$$\nabla_\alpha \nabla_\beta V^\mu = \nabla_\alpha (V_\beta^\mu) = (V_\beta^\mu)_{,\alpha} + \Gamma^\mu_{\sigma \alpha} V_\beta^{,\sigma} - \Gamma^\sigma_{\beta \alpha} V_\mu^{,\sigma}.$$  \hspace{1cm} (6.38)

In a locally inertial frame $\Gamma^\mu_{\sigma \alpha} = 0$, and eq. (6.38) becomes

$$\nabla_\alpha \nabla_\beta V^\mu = (V_\beta^\mu)_{,\alpha} = V^\mu_{,\beta,\alpha} + \Gamma^\mu_{\nu \beta,\alpha} V^\nu.$$  \hspace{1cm} (6.39)

By interchanging $\alpha$ and $\beta$

$$\nabla_\beta \nabla_\alpha V^\mu = (V_\alpha^\mu)_{,\beta} = V^\mu_{,\alpha,\beta} + \Gamma^\mu_{\nu \alpha,\beta} V^\nu.$$  \hspace{1cm} (6.40)

The commutator of the covariant derivatives then is

$$[\nabla_\alpha, \nabla_\beta] V^\mu = \nabla_\alpha \nabla_\beta V^\mu - \nabla_\beta \nabla_\alpha V^\mu = (\Gamma^\mu_{\nu \beta,\alpha} - \Gamma^\mu_{\nu \alpha,\beta}) V^\nu.$$  \hspace{1cm} (6.41)

Since in a locally inertial frame

$$R^\mu_{\nu \alpha \beta} = \Gamma^\mu_{\nu \beta,\alpha} - \Gamma^\mu_{\nu \alpha,\beta}$$  \hspace{1cm} (6.42)

(equivalent to eq. 6.34), eq. (6.41) becomes

$$[\nabla_\alpha, \nabla_\beta] V^\mu = R^\mu_{\nu \alpha \beta} V^\nu.$$  \hspace{1cm} (6.43)

This is a tensor equation and since it is valid in a given reference frame, it will be valid in any frame. Eq. (6.43) implies that in curved spacetime covariant derivatives do not commute and therefore the order in which they appear is important.

6.5 The Bianchi identities

Let us differentiate eq. (6.35) with respect to $x^\lambda$ (and remember that it is valid in a locally inertial frame)

$$R_{\alpha \beta \mu \nu,\lambda} = \frac{1}{2} [g_{\alpha \nu,\beta \mu,\lambda} - g_{\alpha \mu,\beta \nu,\lambda} + g_{\beta \mu,\alpha \nu,\lambda} - g_{\beta \nu,\alpha \mu,\lambda}].$$  \hspace{1cm} (6.44)

By using the fact that $g_{\alpha \beta}$ is symmetric and eq. (6.44) one can show that

$$R_{\alpha \beta \mu \nu,\lambda} + R_{\alpha \beta \lambda \mu,\nu} + R_{\alpha \beta \nu \lambda,\mu} = 0.$$  \hspace{1cm} (6.45)

Since it is valid in a locally inertial frame and it is a tensor equation, it will be valid in any frame:

$$R_{\alpha \beta \mu \nu,\lambda} + R_{\alpha \beta \lambda \mu,\nu} + R_{\alpha \beta \nu \lambda,\mu} = 0,$$  \hspace{1cm} (6.46)

where we have replaced the ordinary derivative with the covariant derivative. These are the Bianchi identities that, as we shall see, play an important role in the development of the theory.
Chapter 7

The stress-energy tensor

Now we know that there exists a tensor which allows to understand if the spacetime is curved or flat, i.e. if we are in the presence of a non-constant, non-uniform gravitational field. But in order to derive Einstein’s equations, we still need to answer the following question: how do we describe matter and fields in general relativity? This question is relevant because we want to find out what to put on the right-hand-side of the equations as a source of the gravitational field.

We shall first define the stress-energy tensor in a flat spacetime, and then generalize this notion to a generic spacetime.

In Special Relativity, we define the energy-momentum four-vector of a particle of mass \( m \) and velocity \( \mathbf{v} = \frac{d\xi}{dt} \) in the following way

\[
p^\alpha = m c u^\alpha, \quad \alpha = 0, 3,
\]

where \( u^\alpha = \frac{d\xi^\alpha}{d\tau} \) is the four-velocity \( (u^\alpha u_\alpha = -1) \); \( \tau \), which has the dimensions of a length, is related to the particle proper time by the equation: \( \text{proper time} = \frac{1}{c^2} \tau \). In what follows, we shall indicate in bold face tri-vectors, for instance \( \mathbf{v} \), whereas four-vectors will be indicated with an arrow, i.e. \( \vec{A} \). Also remember that \( \{\xi^\alpha\} \) are the coordinates of a flat spacetime, or of a locally inertial frame.

Note that \( \xi^0 = ct \) and, defining

\[
\gamma = \frac{d\xi^0}{d\tau},
\]

we have:

\[
\begin{align*}
    u^0 &= \gamma \\
    u^i &= \frac{d\xi^i}{d\tau} = \frac{d\xi^i}{dt} \frac{dt}{d\tau} = \frac{v^i}{c} \\
    u^\alpha u^\beta \eta_{\alpha\beta} &= -\gamma^2 \left(1 - \frac{v^2}{c^2}\right) = -1 \quad \Rightarrow \quad \gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2}.
\end{align*}
\]

We have then

\[
p^\mu = m(c\gamma, \gamma \mathbf{v}).
\]
The time-component of the energy-momentum vector does represent the energy of the particle
\[ p^0 = \frac{E}{c}, \quad \text{and} \quad E = mc^2 \gamma. \] (7.5)
The space-components are the components of the three-dimensional momentum
\[ \mathbf{p} = m \gamma \mathbf{v}. \] (7.6)
What does it change if we are dealing with a continuous or discrete distribution of matter and energy? In that case we should be able to measure some other quantities, as the mass and the energy which are contained in a unitary volume, or the flux of energy and momentum that flows across the different faces of this volume. These informations are contained in the stress-energy tensor we are now going to define.

Let us consider the simple case of a system of \( n \) non-interacting particles located at some points \( \xi_n(t) \), each with an energy-momentum vector \( p_n^a \).

We define the **density of energy** as
\[ T^{00} \equiv \sum_n c p_n^0(t) \delta^3(\xi - \xi_n(t)) = \sum_n E_n \delta^3(\xi - \xi_n(t)), \] (7.7)
the **density of momentum** \( \frac{1}{c} T^{0i} \), where \( T^{0i} \) is defined as
\[ T^{0i} \equiv \sum_n c p_n^i(t) \delta^3(\xi - \xi_n(t)), \quad i = 1, 3 \] (7.8)
and the **current** of momentum as
\[ T^{ki} \equiv \sum_n p_n^k(t) \frac{d\xi_n^i(t)}{dt} \delta^3(\xi - \xi_n(t)), \quad k = 1, 3 \quad i = 1, 3. \] (7.9)
\( \delta^3(\xi - \xi_n) \) is the Dirac delta-function defined by the statement that for any smooth function \( f(\xi) \)
\[ \int d^3\xi \, f(\xi) \delta^3(\xi - \xi_n) = f(\xi_n), \] (7.10)
and if \( \xi_n = (x_0, y_0, z_0) \)
\[ \delta^3(\xi - \xi_n) = \delta(x - x_0) \delta(y - y_0) \delta(z - z_0), \] (7.11)
or, in polar coordinates
\[ \delta^3(\xi - \xi_n) = \frac{1}{r^2 \sin^2 \theta} \delta(r - r_0) \delta(\theta - \theta_0) \delta(\varphi - \varphi_0). \] (7.12)

Thus, according to the definition (7.10) the three-dimensional \( \delta \)-function has the dimensions of the inverse of a cubic length \( l^{-3} \). For this reason, for example, \( T^{00} \) is, dimensionally, an
energy ([\text{cp}^0]) divided by a volume ([\delta^3]) and therefore it is the energy density of the system. The definitions (7.7), (7.8) and (7.9) can be unified into a single formula

\[ T^{\alpha\beta} = \sum_n p_n^\alpha \frac{dE_n^\beta}{dt} \delta^3(\xi - \xi_n(t)), \quad \alpha, \beta = 0, 3, \] (7.13)

or, since

\[ p_n^\alpha = \frac{E_n}{c^2} \frac{dE_n^\beta}{dt}, \] (7.14)

eq. (7.13) can also be written as

\[ T^{\alpha\beta} = c^2 \sum_n p_n^\alpha p_n^\beta \delta^3(\xi - \xi_n(t)), \] (7.15)

which clearly shows that \( T^{\alpha\beta} \) is symmetric

\[ T^{\alpha\beta} = T^{\beta\alpha}. \] (7.16)

Finally, an alternative way of writing eq. (7.13) is

\[ T^{\alpha\beta} = c \sum_n \int p_n^\alpha \frac{dE_n^\beta}{d\tau_n} \delta^4(\xi - \xi_n(\tau_n)) d\tau_n, \] (7.17)

where

\[ \delta^4(\xi - \xi_n) = \delta(\xi^0 - \xi_n^0)\delta(\xi^1 - \xi_n^1)\delta(\xi^2 - \xi_n^2)\delta(\xi^3 - \xi_n^3); \] (7.18)

indeed, using the property (7.10) of the \( \delta \)-function it is easy to see that

\[ T^{\alpha\beta} = c \sum_n \int p_n^\alpha \frac{dE_n^\beta}{d\tau_n} \delta^4(\xi - \xi_n(\tau_n)) d\tau_n = \sum_n p_n^\alpha \frac{dE_n^\beta}{dt} \delta^3(\xi - \xi_n(\xi^0)) = \sum_n p_n^\alpha \frac{dE_n^\beta}{dt} \delta^3(\xi - \xi_n(\xi^0)) \] (7.19)

which coincides with (7.15)

---

1Properties of the \( \delta \)-function

\[ \delta(x) = \delta(-x), \quad \delta(nx) = \frac{1}{|n|} \delta(x) \]

\[ \delta[g(x)] = \sum_j \frac{1}{|g'(x_j)|} \delta(x - x_j) \quad x\delta(x) = 0 \]

\[ \int dx f(x) \delta'(x - x_0) = -f'(x_0). \]
Summarizing, the meaning of the different components is the following

\( T^{00} \) = energy-density. In the non-relativistic case \( v << c, \ p^0_n \sim m_n c \) and \( T^{00} \sim \sum_n m_n c^2 \delta(\xi - \xi_n(t)) \) reduces to the density of matter \( \rho c^2 \) where

\[
\rho = \sum_n m_n \delta^3(\xi - \xi_n(t))
\]  
(7.20)

(remember the dimensions of the \( \delta \)-function).

\( \frac{1}{c} T^{0i} \) = density of momentum. Since the dimensions of the momentum \( \vec{p} \) are those of an energy divided by a velocity, \( [p^0] = [E/c] \), it follows that \( cT^{0i} \) has the dimensions of \( [E_t S] \), i.e. it is the energy which flows across the unit surface orthogonal to the axis \( \xi^i \) per unit time \( i=1,3 \) (see eq. (7.8)).

Similar dimensional considerations allow us the interpret \( T^{ik} \) as the flux of the i-th component of the three-momentum \( \vec{p} \) across the unit surface orthogonal to the axis \( \xi^k \) \( (i,k=1,3) \) (see eq. (7.9)).

Now we must check several things:

1) is \( T^{\alpha\beta} \) a tensor?

2) does it satisfy any conservation law? (remember that the energy-momentum four vector does satisfy a conservation law).

3) if it does, how to write this law in a curved spacetime, i.e. in the presence of a gravitational field?

**1) is \( T^{\alpha\beta} \) a tensor?**

Let us consider a generic coordinate transformation

\[
\{\xi^\alpha\} \rightarrow \{x^\alpha'\} \rightarrow \xi^\alpha = \Lambda^\alpha_{\gamma'} x^{\gamma'},
\]  
(7.21)

The four-momentum and the four-velocity transform as

\[
p^\alpha = \Lambda^\alpha_{\gamma'} p^{\gamma'},
\]
\[
u^\alpha = \frac{d\xi^\alpha}{d\tau} = \Lambda^\alpha_{\gamma'} u^{\gamma'}.
\]  
(7.22)

In order to see how \( T^{\alpha\beta} \) transforms we need a brief digression to show how to transform \( \delta^4(x) \).

In a four dimensional spacetime the volume element which is invariant under a generic coordinate transformation is \( \sqrt{-g} \ d^4 x \), i.e.

\[
\sqrt{-g} \ d^4 x = \sqrt{-g'} \ d^4 x'.
\]  
(7.23)

Indeed,

\[
d^4 x = |J| \ d^4 x',
\]  
(7.24)

where \( J = \det \left( \frac{\partial x^\alpha}{\partial x'^\mu} \right) \) is the Jacobian associated to the coordinate transformation. Since

\[
g_{\alpha'\beta'} = \frac{\partial x^\mu}{\partial x'^\alpha} \frac{\partial x^\nu}{\partial x'^\beta} g_{\mu\nu},
\]  
(7.25)
CHAPTER 7. THE STRESS-ENERGY TENSOR

84

taking the determinant of both member we get

\[ g' = J^2 g \quad \text{and therefore} \quad \sqrt{-g} = \frac{1}{|J|} \sqrt{-g'}. \] (7.26)

Thus, if \( \{ \xi^\alpha \} \) is a Minkowskian frame, and \( \{ x'^\alpha \} \) is a generic frame,

\[ d^4 \xi = \sqrt{-g} \ d^4 x. \] (7.27)

Let us now consider a delta-function in Minkowski’s spacetime; its definition is

\[ \int d^4 \xi \ \delta^4(\vec{\xi} - \vec{\xi}_n) = 1, \] (7.28)

and, in a generic frame,

\[ \int d^4 x \ \delta^4(\vec{x} - \vec{x}_n) = 1, \] (7.29)

i.e.

\[ \int \sqrt{-g} \ d^4 x \ \delta^4(\vec{x} - \vec{x}_n) \sqrt{-g} = \int \delta^4(\vec{x} - \vec{x}_n) \ d^4 \xi = 1. \] (7.30)

By comparing eq. (7.28) and eq. (7.30) we find the sought transformation formula for the delta-function

\[ \delta^4(\vec{\xi} - \vec{\xi}_n) = \frac{\delta^4(\vec{x} - \vec{x}_n)}{\sqrt{-g}}. \] (7.31)

Using eqs. (7.17), (7.22) and (7.31) it is now easy to find the transformation rule for \( T^{\alpha\beta} \):

\[ T^{\alpha\beta} = c \sum_n \int \Lambda^\alpha_\gamma \Lambda^\beta_\delta \ p_n^\gamma \frac{dx^\delta_n}{d\tau_n} \delta^4(\vec{x} - \vec{x}_n) \ d\tau_n. \] (7.32)

Therefore if we define

\[ T^{\alpha\beta} = c \sum_n \int \frac{1}{\sqrt{-g}} \ p_n^\alpha \frac{dx^\beta_n}{d\tau_n} \delta^4(\vec{x} - \vec{x}_n) \ d\tau_n, \] (7.33)

under a generic coordinate transformation it will transform like

\[ T^{\alpha\beta} = \Lambda^\alpha_\gamma \Lambda^\beta_\delta \ T^{\gamma\delta}. \] (7.34)

and therefore it is a tensor. In flat spacetime, and in a locally inertial frame \( \sqrt{-g} = 1 \) and we recover the definition (7.17). In conclusion, eq. (7.33) is the stress-energy tensor appropriate to describe a cloud of non interacting particles both in flat and in curved spacetime. Of course we may have different kind of matter and/or energy: a fluid, an electromagnetic field, etc. In that case it is possible to show that the corresponding stress-energy tensor can be derived by writing the action of the considered field, and by varying this action with respect to \( g_{\mu\nu} \). However, the physical meaning of the different components of \( T^{\alpha\beta} \) will be the same.
We shall now use the tensor we have derived to answer the second important question we raised. The answer will be valid for the stress-energy tensor of any sort of matter-energy.

2) Does \( T^{\alpha\beta} \) satisfy a conservation law?

Let us assume that we are in flat spacetime, and let us differentiate eq. (7.9):

\[
\frac{\partial T^{\alpha i}}{\partial \xi^i} = \sum_n p_n^\alpha(t) \frac{d\xi_n^i(t)}{dt} \frac{\partial}{\partial \xi^i} \delta^3(\mathbf{\xi} - \mathbf{\xi}_n(t)),
\]

(7.35)

where \( \alpha = 0,3 \) and \( i = 1,3 \). Since

\[
\frac{\partial}{\partial \xi^i} \delta^3(\mathbf{\xi} - \mathbf{\xi}_n(t)) = -\frac{\partial}{\partial \xi^i_n} \delta^3(\mathbf{\xi} - \mathbf{\xi}_n(t)),
\]

(7.36)
eq (7.35) becomes

\[
\frac{\partial T^{\alpha i}}{\partial \xi^i} = -\sum_n p_n^\alpha(t) \frac{d\xi_n^i(t)}{dt} \frac{\partial}{\partial \xi^i} \delta^3(\mathbf{\xi} - \mathbf{\xi}_n(t))
\]

(7.37)

By making use of eqs. (7.7) and (7.8), eq. (7.37) gives

\[
\frac{\partial T^{\alpha i}}{\partial \xi^i} = -\frac{1}{c} \frac{\partial}{\partial t} T^{\alpha 0} + \sum_n \frac{dp_n^\alpha(t)}{dt} \delta^3(\mathbf{\xi} - \mathbf{\xi}_n(t)).
\]

(7.38)

Since

\[
\frac{dp_n^\alpha(t)}{dt} = \frac{dp_n^\alpha(\tau)}{d\tau} \frac{d\tau}{dt} = \frac{d\tau}{dt} f_n^\alpha,
\]

(7.39)

where \( f_n^\alpha \) is the relativistic force, the last term in eq. (7.38) can be considered as a density of force \( G^\alpha \) defined as

\[
G^\alpha(\mathbf{\xi}, t) = \sum_n \frac{dp_n^\alpha(t)}{dt} \delta^3(\mathbf{\xi} - \mathbf{\xi}_n(t)) = \sum_n \delta^3(\mathbf{\xi} - \mathbf{\xi}_n(t)) \frac{d\tau}{dt} f_n^\alpha.
\]

(7.40)

It is a density because the \( \delta \)-function is \([l^{-3}]\). If the particles are free, \( f_n^\alpha = 0 \) and eq. (7.38) becomes

\[
\frac{\partial}{\partial \xi^i} T^{\alpha i} + \frac{1}{c} \frac{\partial}{\partial t} T^{\alpha 0} = \frac{\partial}{\partial \xi^i} T^{\alpha i} = 0,
\]

(7.41)

or

\[
T^{\alpha \beta} = 0,
\]

(7.42)

which is the conservation law we were looking for.

Why is \( T^{\alpha \beta} = 0 \) a conservation law? To answer this question, let us start with a familiar equation in classical electrodynamics. Consider, as an example, a collection of charged particles of density \( \rho \) enclosed in a volume \( V \).

\[
\frac{\partial}{\partial t} \int_V \rho dV
\]

(7.43)
will be the variation of charge inside the volume \( V \). Be \( S \) the surface enclosing the volume, and \( \mathbf{n} \) the normal vector, which is assumed to be positive if pointing outward.

\[
\rho \mathbf{v} \cdot \mathbf{n} dS \tag{7.44}
\]

will be the charge which flows across \( dS \) per unit time. It is positive if the charge goes out, negative if it flows in. Thus

\[
\int_S \rho \mathbf{v} \cdot \mathbf{n} dS \tag{7.45}
\]
is the total charge per unit time, which flows across the surface \( S \) enclosing the volume \( V \).

The continuity equation then says that

\[
\frac{\partial}{\partial t} \int_V \rho dV = - \int_S \rho \mathbf{v} \cdot \mathbf{n} dS. \tag{7.46}
\]
The minus sign is because the right-hand side is positive if the charge contained in \( V \) increases. If we now introduce the three-dimensional current

\[
\mathbf{J} = \rho \mathbf{v}, \tag{7.47}
\]
eq (7.46) becomes

\[
\frac{\partial}{\partial t} \int_V \rho dV = - \int_S \mathbf{J} \cdot \mathbf{n} dS. \tag{7.48}
\]

We now apply the Gauss theorem:

\[
\int_S \mathbf{J} \cdot \mathbf{n} dS = \int_V \text{div}\, \mathbf{J} dV, \tag{7.49}
\]
and eq. (7.48) becomes

\[
\frac{\partial}{\partial t} \int_V \rho dV = - \int_V \text{div}\, \mathbf{J} dV. \tag{7.50}
\]

Since the volume \( V \) is arbitrary, we can write

\[
\text{div}\, \mathbf{J} = - \frac{\partial \rho}{\partial t}, \tag{7.51}
\]
or

\[
\frac{\partial J_x}{\partial x} + \frac{\partial J_y}{\partial y} + \frac{\partial J_z}{\partial z} = - \frac{\partial \rho}{\partial t}, \tag{7.52}
\]
which is the continuity equation in a differential form. Let us now transform eq. (7.51) in a four-dimensional form. We define a four-current

\[
J^\alpha = \rho \frac{d\xi^\alpha}{dt} = (\rho c, \mathbf{J}), \tag{7.53}
\]

Then eq. (7.51) becomes

\[
\frac{\partial}{\partial \xi^\alpha} J^\alpha = 0, \quad \alpha = 0, 3. \tag{7.54}
\]
We are now going to show that any current \( J^\alpha(x) \) which satisfies the conservation law (7.54) is associated to a total charge \( Q \) defined as

\[
Q = \int_V J^0 dV, \tag{7.55}
\]

which is conserved. The integral in eq. (7.55) is evaluated at some fixed time, thus we say that the integration is performed on a hypersurface \( \xi^0 = \text{const} \) over the whole three-dimensional space. The total charge \( Q \) is a conserved quantity for the following reason. By virtue of eq. (7.54)

\[
\frac{1}{c} \frac{dQ}{dt} = \int_{\text{all space}} \frac{1}{c} \frac{\partial}{\partial t} J^0 dV = - \int_{\text{all space}} \text{div} J dV = - \int_{\text{surface}} J^k dS_k. \tag{7.56}
\]

The last equality follows from the application of the Gauss theorem, and the subscript ‘surface’ means that we are considering the flux of \( J \) across the surface which encloses the whole space. \( dS_k \) are the element of surface orthogonal to \( \xi^k \). If \( J \) goes to zero at infinity, the last term in eq. (7.56) vanishes, and therefore the total charge \( Q \) is a conserved quantity.

And now let us go back to equation (7.42). Let us assume for example that \( \alpha = 0 \):

\[
\partial T^{01}/\partial \xi^1 + \partial T^{02}/\partial \xi^2 + \partial T^{03}/\partial \xi^3 = - \partial T^{00}/\partial \xi^0. \tag{7.57}
\]

If we integrate over a volume \( V \) as we did before, we get

\[
- \frac{\partial}{\partial \xi^0} \int_V T^{00} dV = \int_V \text{div}(T^{0k}) dV = \int_S T^{0k} dS_k. \tag{7.58}
\]

Remembering that \( T^{00} \) is the energy-density and \( T^{0k} \) is the energy which flows across the unit surface orthogonal to \( \xi^k \) it is clear that eq. (7.58) expresses a law of conservation of energy, and a similar procedure can be used to find the conservation of momentum by putting \( \alpha = 1, 2, 3 \). In analogy with eq. (7.55) we can define a vector

\[
P^\alpha = \int_V T^{\alpha 0} dV, \quad \alpha = 0, 3, \tag{7.59}
\]

which can be identified as the conserved energy-momentum vector of the system. For example

\[
P^0 = \int_V T^{00} dV, \tag{7.60}
\]

does represent the total energy of the system. It is conserved because

\[
\frac{1}{c} \frac{dP^0}{dt} = \frac{1}{c} \int_{\text{all space}} \frac{\partial}{\partial t} T^{00} dV = - \int_{\text{all space}} \frac{\partial}{\partial \xi^i} T^{0i} dV = - \int_{\text{surface}} T^{0i} dS_i = 0. \tag{7.61}
\]

It should be reminded that this derivation has been carried out in the framework of Special Relativity.

3) How do we write this conservation law in curved spacetime?

In order to answer this question we need to state The Principle of General Covariance which will be the foundation of the theory of General Relativity:
7.1 The Principle of General Covariance

A physical law is true if:

1) it is true in the absence of gravity, i.e. it reduces to the laws of special relativity when $g_{\mu\nu} \rightarrow \eta_{\mu\nu}$ and $\Gamma^{\alpha}_{\mu\nu}$ vanish. It is clear that this first proposition includes the Equivalence Principle.

2) In order to preserve their form under an arbitrary coordinate transformation, all equations must be generally covariant. This means that all equations must be expressed in a tensor form.

The physical content of the Principle of General Covariance is that if a tensor equation is true in absence of gravity, then it is true in the presence of an arbitrary gravitational field. It should also be stressed that the Principle of General Covariance can be applied only on scales that are small compared with the typical distances associated to the gravitational field, (for example to the curvature) , because only on these scales one can construct locally inertial frames.

And now we can give an answer to the question 3). First we note that eq. (7.42) is valid in special relativity, i.e. in the absence of gravity, therefore, according to the Principle of Equivalence, it will hold in a locally inertial frame of a curved spacetime. In this frame, the covariant and ordinary derivative coincide, therefore we can write eq. (7.42) in the alternative form

$$T^{\alpha\beta ; \gamma} = 0. \quad (7.62)$$

Then we observe that in the light of the Principle of General Covariance, since the conservation law (7.42) is a tensor equation, it will hold in any arbitrary frame. Thus in order to transform a generic tensor equation valid in Special Relativity to a generally covariant form it will suffice to replace the comma with a semi-colon. The general conservation law satisfied by the stress-energy tensor therefore is eq. (7.62).

Is this a conservation law?

To answer this question we need to compute the covariant divergence of a tensor. From the expression of the affine connections in terms of the metric we find that

$$\Gamma^\mu_{\lambda\mu} = \frac{1}{2} g^{\mu\rho} \left( \frac{\partial g_{\rho\lambda}}{\partial x^\mu} + \frac{\partial g_{\mu\lambda}}{\partial x^\rho} - \frac{\partial g_{\rho\mu}}{\partial x^\lambda} \right). \quad (7.63)$$

The first and the third term give

$$g^{\mu\rho} \frac{\partial g_{\rho\lambda}}{\partial x^\mu} - g^{\mu\rho} \frac{\partial g_{\rho\mu}}{\partial x^\lambda} = g^{\mu\rho} \frac{\partial g_{\rho\lambda}}{\partial x^\mu} - g^{\rho\mu} \frac{\partial g_{\mu\lambda}}{\partial x^\rho} = 0, \quad (7.64)$$

due to the symmetry of $g_{\alpha\beta}$, therefore

$$\Gamma^\mu_{\lambda\mu} = \frac{1}{2} g^{\mu\rho} \frac{\partial g_{\rho\mu}}{\partial x^\lambda}. \quad (7.65)$$

For any arbitrary matrix $M$

$$\text{Tr} \left[ M^{-1}(x) \frac{\partial}{\partial x^\lambda} M(x) \right] = \frac{\partial}{\partial x^\lambda} \ln |\text{Det} M(x)|. \quad (7.66)$$
CHAPTER 7. THE STRESS-ENERGY TENSOR

But this is what we have on the right-hand side of eq. (7.65), therefore, if we call \( \text{Det}(g) = g \), eq. (7.65) becomes (since \( g < 0 \))

\[
\Gamma^\mu_{\lambda\mu} = \frac{1}{2} \frac{\partial}{\partial x^\lambda} \ln[-g] = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\lambda} \sqrt{-g}.
\]  

(7.67)

Thus for example, if \( V^\mu \) is a vector

\[
V^\lambda;\lambda = V^\lambda,\lambda + \Gamma^\lambda_{\alpha\lambda} V^\alpha = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\lambda} (\sqrt{-g} V^\lambda),
\]  

(7.68)

and for \( T^{\mu\nu} \)

\[
T^{\mu\nu};\mu = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\nu} (\sqrt{-g} T^{\mu\nu}) + \Gamma^\nu_{\mu\lambda} T^{\mu\lambda}.
\]  

(7.69)

In particular, if \( F^{\mu\nu} \) is antisymmetric, the last term in eq. (7.69) is zero and

\[
F^{\mu\nu};\mu = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\lambda} (\sqrt{-g} F^{\mu\nu}).
\]  

(7.70)

Now we go back to eq. (7.62). By using eq. (7.69) it becomes

\[
\frac{\partial}{\partial x^\lambda} (\sqrt{-g} T^{\mu\nu}) = -\sqrt{-g} \Gamma^\nu_{\lambda\mu} T^{\mu\lambda},
\]  

(7.71)

and this is not a conservation law. Thus we cannot define a conserved four-momentum as we did in Special Relativity. We may be tempted to define

\[
P^\alpha = \int_V \sqrt{-g} T^{\alpha\theta} dV, \quad \alpha = 0, 3,
\]  

(7.72)

but this would not be a vector. The physical reason for this failure is that now we are in General Relativity, and we must take into account not only the energy and momentum associated to matter, but also the energy which is carried by the gravitational field itself, and the momentum which may be carried by gravitational waves. However we shall see that if the spacetime admits some symmetry (for example if it is spherically or plane-symmetric, or it is invariant under time-translations etc.) conserved quantities can be defined.
Chapter 8

The Einstein equations

We now have all the elements needed to derive the equations of the gravitational field. We expect they will be more complicated than the linear equations of the electromagnetic field. For example electromagnetic waves are produced as a consequence of the motion of charged particles, but the energy and the momentum they carry are not a source for the electromagnetic field, and they do not appear on the right-hand side of the equations. In gravity the situation is different. The equation

\[ E = mc^2, \quad (8.1) \]

establishes that mass and energy can transform one into another: they are different manifestation of the same physical quantity. It follows that if the mass is the source of the gravitational field, so must be the energy, and consequently both mass and energy should appear on the right-hand side of the field equations. This implies that the equations we are looking for will be non linear. For example a system of arbitrarily moving masses will radiate gravitational waves, which carry energy, which is in turn source of the gravitational field and must appear on the right-hand-side of the equations. However, since newtonian gravity works remarkably well when we are dealing with non relativistic particles, or in general when the gravitational field is weak, in formulating the new theory we shall require that in the weak field limit the new equations reduce to the Poisson equation

\[ \nabla^2 \Phi = 4\pi G \rho, \quad (8.2) \]

where \( \rho \) is the matter density, \( \Phi \) is the newtonian potential and \( \nabla^2 \) is the Laplace operator in cartesian coordinates

\[ \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (8.3) \]

Let us start by asking how the equations should look in the weak field limit.
8.1 The geodesic equations in the weak field limit

Consider a non-relativistic particle which moves in a weak and stationary gravitational field. Let \( \tau / c \) be the proper time. Since \( v \ll c \), it follows that
\[
\frac{dx^i}{dt} \ll c \quad \rightarrow \quad \frac{dx^i}{d\tau} \ll \frac{cdt}{d\tau} = \frac{dx^0}{d\tau}.
\] (8.4)

In an arbitrary coordinate system the geodesic equations are
\[
\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} = 0 \quad \rightarrow \quad \frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{00} \left( \frac{cdt}{d\tau} \right)^2 = 0.
\] (8.5)

From the expressions of the affine connections in terms of \( g_{\mu\nu} \), we easily find that
\[
\Gamma^\mu_{00} = \frac{1}{2} g^{\mu\sigma} (2g_{0\sigma,0} - g_{00,\sigma}).
\] (8.6)

In addition, if the field is stationary \( g_{0\sigma,0} = 0 \), and
\[
\Gamma^\mu_{00} = -\frac{1}{2} g^{\mu\sigma} \frac{\partial g_{00}}{\partial x^\sigma}.
\] (8.7)

Since we have assumed that the gravitational field is weak, we can choose a coordinate system such that
\[
g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad |h_{\mu\nu}| \ll 1,
\] (8.8)
where \( h_{\mu\nu} \) is a small perturbation of the flat metric. In other words, we are assuming that the field is so weak that the metric is nearly flat. Since we are interested only in first order terms, we shall raise and lower indices with the flat metric \( \eta^{\mu\nu} \). For example
\[
h^\lambda_{\nu} = g^{\lambda\rho} h_{\rho\nu} \sim \eta^{\lambda\rho} h_{\rho\nu} + O(h^2_{\mu\nu}).
\]

If we substitute eq. (8.8) into eq. (8.7), and retain only the terms up to first order in \( h_{\mu\nu} \), we find
\[
\Gamma^\mu_{00} \sim -\frac{1}{2} \eta^{\mu\sigma} \frac{\partial h_{00}}{\partial x^\sigma},
\] (8.9)
and the geodesic equation becomes
\[
\frac{d^2x^\mu}{d\tau^2} = \frac{1}{2} \eta^{\mu\alpha} \frac{\partial h_{00}}{\partial x^\alpha} \left( \frac{cdt}{d\tau} \right)^2,
\] (8.10)
or, splitting the time- and the space-components
\[
\frac{d^2x}{d\tau^2} = \frac{1}{2} \nabla h_{00} \left( \frac{cdt}{d\tau} \right)^2, \quad \text{and} \quad \frac{d^2ct}{d\tau^2} = -\frac{1}{2} \frac{\partial h_{00}}{\partial ct} \left( \frac{cdt}{d\tau} \right)^2 = 0.
\] (8.11)

where
\[
\nabla \rightarrow \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)
\] (8.12)
is the gradient in cartesian coordinates. The second equation vanishes because we have assumed that the field is stationary ($\frac{\partial h_{00}}{\partial t} = 0$). We can rescale the time coordinate in such a way that $\frac{d\tau}{dt} = 1$ and the first of eqs. (8.11) becomes

$$\frac{d^2 x}{d\tau^2} = \frac{1}{2} \nabla^2 h_{00}. \quad (8.13)$$

We should remember that the corresponding newtonian equation is

$$\frac{d^2 x}{dt^2} = -\nabla^2 \Phi, \quad (8.14)$$

where $\Phi$ is the gravitational potential given by the Poisson equation (8.2). By comparing eqs. (8.14) and (8.13), and since $\tau = ct$ we see that it must be

$$h_{00} = -2\frac{\Phi}{c^2} + \text{const.} \quad (8.15)$$

For example if the field is stationary and spherically symmetric, the newtonian potential is

$$\Phi = -\frac{GM}{r}, \quad (8.16)$$

and if we require that $h_{00}$ vanishes at infinity, the constant must be zero and eq. (8.15) gives

$$h_{00} = -2\frac{\Phi}{c^2}, \quad \text{and} \quad g_{00} = -(1 + 2\frac{\Phi}{c^2}). \quad (8.17)$$

Thus we have shown that in the weak field limit the geodesic equations reduce to the newtonian law of gravitation. This suggests the form that the field equations should have. In fact if the field is weak, matter will behave non-relativistically, i.e. $T^{00} = T_{00} \sim \rho c^2$ and therefore the generalization of Laplace’s equation (8.2) could be

$$\nabla^2 g_{00} = -\frac{8\pi G}{c^4} T_{00}. \quad (8.18)$$

But this equation is not even Lorentz-invariant! It doesn’t work. However it suggests that if in place of a stationary field, we would have an arbitrary distribution of energy and matter, we should construct a tensor starting from $g_{\mu \nu}$ and its derivatives such that the field equations are

$$G_{\mu \nu} = \frac{8\pi G}{c^4} T_{\mu \nu}, \quad (8.19)$$

where $G_{\mu \nu}$ is an operator acting on $g_{\mu \nu}$ which we shall now define. It should be stressed that, by the Principle of General Covariance, if equation (8.19) holds in a given reference frame, it will hold in any other frame.
8.2 Einstein’s field equations

Let us first see which derivatives and of which order do we expect in $G_{\mu\nu}$. A comparison with the Laplace equation shows that $G_{\mu\nu}$ must have the dimensions of a second derivative. In fact, suppose that it contains terms of this type

$$\frac{\partial^3 g_{\mu\nu}}{\partial x^3_{\mu}}, \quad \frac{\partial^2 g_{\mu\nu}}{\partial x^2_{\mu}} \frac{\partial g_{\mu\nu}}{\partial x_{\nu}}, \quad \frac{\partial g_{\mu\nu}}{\partial x_{\nu}}.$$

(8.20)

then, in order to be dimensionally homogeneous each term should be multiplied by a constant having the dimensions of a suitable power of a length

$$\frac{\partial^3 g_{\mu\nu}}{\partial x^3_{\mu}} \cdot l, \quad \frac{\partial^2 g_{\mu\nu}}{\partial x^2_{\mu}} \frac{\partial g_{\mu\nu}}{\partial x_{\nu}} \cdot l, \quad \frac{\partial g_{\mu\nu}}{\partial x_{\nu}} \cdot l.$$

(8.21)

In this case, a gravitational field acting on small or on very large scale would be described by equations where some of the terms would be negligible with respect to some others. This is unacceptable, because we want a set of equations that are valid at any scale, and consequently the only terms we can accept in $G_{\mu\nu}$ are those containing the second derivatives of $g_{\mu\nu}$ in a linear form and products of first derivatives. Let us summarize the assumptions that we need to make on $G_{\mu\nu}$:

1) it must be a tensor
2) it must be linear in the second derivatives, and it must contain products of first derivatives of $g_{\mu\nu}$.
3) Since $T_{\mu\nu}$ is symmetric, $G_{\mu\nu}$ also must be symmetric.
4) Since $T_{\mu\nu}$ satisfies the “conservation law” $T_{\mu\nu;\mu} = 0$, $G_{\mu\nu}$ must satisfy the same conservation law.
5) In the weak field limit it must reduce to (compare with eq. (8.18))

$$G_{00} \sim -\nabla^2 g_{00}.$$

(8.22)

(8.23)

In this last assumption the Principle of Equivalence and the weak field limit explicitly appear.

In the preceding section we have shown that there exists a tensor which is linear in the second derivatives of $g_{\mu\nu}$ and non linear in the first derivatives. It is the Riemann tensor, given in eq. (6.34), and it contains the information on the gravitational field. However we cannot use it directly in the field equations we are looking for, since it has four indices (it is a $^{1\times3}$ tensor) while we need a $^{2\times0}$ (or $^{0\times2}$) tensor. In addition, the covariant divergence of the stress-energy tensor vanishes, and so must be also for the tensor we shall put on the left-hand side of eq. (8.19).

By contracting the Riemann tensor with the metric we can construct a $^{0\times2}$ tensor, i.e. the Ricci tensor:

$$R_{\mu\nu} = g^{\kappa\alpha} R_{\kappa\mu\nu\alpha} = R^{\alpha}_{\mu\alpha\nu},$$

(8.24)
which is a symmetric tensor because of the symmetry property of the Riemann tensor

\[ R_{\kappa \mu \alpha \nu} = R_{\alpha \kappa \mu \nu}, \] (8.25)

and a scalar, called the \textbf{scalar curvature}

\[ R = R_{\alpha}^{\alpha}. \] (8.26)

The contraction in eq. (8.26) has the following meaning

\[ R_{\alpha}^{\alpha} = R_{\alpha 0}^{\alpha 0} + R_{\alpha 1}^{\alpha 1} + R_{\alpha 2}^{\alpha 2} + R_{\alpha 3}^{\alpha 3}. \] (8.27)

It can be shown, by using the symmetries of the Riemann tensor, that \( R_{\mu \nu} \) and \( R \) are the only second rank tensor and scalar that can be constructed by contraction of \( R_{\kappa \mu \alpha \nu} \) with the metric. Both in \( R_{\mu \nu} \) and \( R \), the second derivatives of \( g_{\mu \nu} \) appear linearly. Therefore the tensor we are looking for should have the following form

\[ G_{\mu \nu} = C_1 R_{\mu \nu} + C_2 g_{\mu \nu} R, \] (8.28)

where \( C_1 \) and \( C_2 \) are constants to be determined. The tensor \( G_{\mu \nu} \) satisfies the points 1, 2, and 3. Condition 4 requires that

\[ G^{\mu \nu; \mu} = C_1 R^{\mu \nu; \mu} + C_2 g^{\mu \nu} R;\mu = 0. \] (8.29)

(remember that the covariant derivative of \( g_{\mu \nu} \) vanishes). Now a very remarkable thing happens: eq. (8.29) is satisfied because of the Bianchi identities

\[ R_{\lambda \mu \nu; \eta} + R_{\lambda \mu \eta; \nu} + R_{\lambda \nu \eta; \mu} = 0. \] (8.30)

In fact by contracting these equations we find

\[ g^{\lambda \nu} (R_{\lambda \mu \nu; \eta} + R_{\lambda \mu \eta; \nu} + R_{\lambda \nu \eta; \mu}) = g^{\lambda \nu} (R_{\lambda \mu \nu; \eta} - R_{\lambda \mu \eta; \nu}) + g^{\lambda \nu} R_{\lambda \mu \nu; \eta} \]

\[ = (R_{\mu \nu; \eta} - R_{\mu \eta; \nu} + R^{\nu} \mu \eta; \eta) = 0. \] (8.31)

Contracting again

\[ g^{\mu \kappa} (R_{\mu \kappa; \eta} - R_{\mu \eta; \kappa} + R^{\eta} \mu \kappa; \nu) = R_{\eta \nu} - R^{\kappa} \eta \kappa - R^{\nu} \eta; \nu = 0. \] (8.32)

The last expression can be rewritten in the following form

\[ \left( R^{\mu \nu} - \frac{1}{2} g^{\mu \nu} R \right); \nu = 0. \] (8.33)

Therefore, the Bianchi identities say that if

\[ \frac{C_2}{C_1} = \frac{1}{2}, \] (8.34)

eq. (8.33) will be satisfied. We still need \( C_1 \).
In the weak field limit \(^1\)
\[ |T_{ij}| << |T_{00}|, \quad i, j = 1, 3, \] (8.37)
and therefore
\[ |G_{ij}| << |G_{00}|, \quad i, j = 1, 3. \] (8.38)
From eqs. (8.28) and (8.34) it follows
\[ |C_1\left(R_{ij} - \frac{1}{2}g_{ij}R\right)| << |G_{00}|, \] (8.39)
hence
\[ R_{ij} \simeq \frac{1}{2}g_{ij}R. \] (8.40)
Since \(g_{ij} \simeq \eta_{ij}\)
\[ R_{kk} \simeq \frac{1}{2}R, \quad k = 1, 3 \] (8.41)
consequently
\[ R = g^{\mu\nu}R_{\mu\nu} \simeq \eta^{\mu\nu}R_{\mu\nu} = -R_{00} + \sum_k R_{kk} = -R_{00} + \frac{3}{2}R, \] (8.42)
and
\[ R \simeq 2R_{00}. \] (8.43)
Since
\[ G_{\mu\nu} = C_1\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right), \] (8.44)
we find
\[ G_{00} \simeq C_12R_{00}. \] (8.45)
If we now compute \(R_{00}\) in the weak field limit (assuming the field is stationary), we find that the non linear part is second order. Retaining only the first order terms and imposing stationarity we get
\[ R_{00} \simeq -\frac{1}{2}\eta^{ij}\frac{\partial^2 g_{00}}{\partial x^i \partial x^j} = -\frac{1}{2}\nabla^2 g_{00}, \quad i, k = 1, 3 \] (8.46)
namely
\[ G_{00} \simeq -C_1\nabla^2 g_{00}. \] (8.47)

\(^1\)The fact that in the weak field limit \(|T_{ik}| << T_{00}\) can be easily understood if we consider, as an example, a system on non-interacting particles. If \(\rho\) is the mass density
\[ \rho = \sum_n m_n\delta^3(r - r_n), \] (8.35)
where \(r_n\) denotes the positions of the particles, the stress-energy tensor (7.15) can be also written as
\[ T^{\mu\nu} = \rho c^2 \frac{dx^{\mu}}{d\tau} \frac{dx^{\nu}}{d\tau}. \] (8.36)
It is clear that, if \(\frac{dx^i}{d\tau} << \frac{dx^i}{d\tau}\) \(i = 1, 3\) the dominant term will be \(T^{00}\).
A comparison of this equation with eq. (8.23) shows that if we require that the relativistic field equations reduce to the newtonian equations in the weak field limit it must be

\[ C_1 = 1. \]  

(8.48)

In conclusion, the Einstein’s field equations are

\[ G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \]  

(8.49)

where

\[ G_{\mu\nu} = \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right), \]  

(8.50)

and it is called The Einstein tensor. An alternative form is

\[ R_{\mu\nu} = \frac{8\pi G}{c^4} \left( T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T \right). \]  

(8.51)

In vacuum \( T_{\mu\nu} = 0 \) and the Einstein equations reduce to

\[ R_{\mu\nu} = 0. \]  

(8.52)

Therefore, in vacuum the Ricci tensor vanishes, but the Riemann tensor does not, unless the gravitational field vanishes or is constant and uniform. We may still add to eqs. (8.49) the following term

\[ \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \lambda g_{\mu\nu} \right) = \frac{8\pi G}{c^4} T_{\mu\nu}. \]  

(8.53)

where \( \lambda \) is a constant. This term satisfies the conditions 1,2,3 and 4, but not the condition 5. This means that it must be very small in such a way that in the weak field limit the equations reduce to the newtonian equations.

### 8.3 Gauge invariance of the Einstein equations

Since there are 10 independent components of \( G_{\mu\nu} \), Einstein’s equations provide 10 equations for the 10 independent components of \( g_{\mu\nu} \). However these equations are not independent, because, as we have seen, the Bianchi identities imply the “conservation law” \( G^{\mu\nu}_{\nu} = 0 \), which provides 4 relations that the Einstein tensor must satisfy. Thus the number of independent equations reduces to six.

Do we have six equations and 10 unknown functions? Why do we have these four degrees of freedom? The reason is the following. Be \( g_{\mu\nu} \) a solution of the equations. If we make a coordinate transformation \( x'^\mu = x^\mu(x^\alpha) \) the ‘transformed’ tensor \( g'_{\mu\nu} = g_{\mu\nu} \) is again
a solution, as established by the Principle of General Covariance. This also means that $g_{\mu\nu}$ and $g'_{\mu\nu}$ do represent the same physical solution (the same geometry) seen in different reference frames.

The coordinate transformation involves 4 arbitrary functions $x'{}^\alpha(x^\beta)$, therefore the four degrees of freedom derive from the freedom of choosing the coordinate system, and disappear when we choose it. For example, we may choose a frame where four of the ten $g_{\mu\nu}$ are zero.

Thus Einstein’s equations do not determine the solution $g_{\mu\nu}$ in a unique way, but only up to an arbitrary coordinate transformation. A similar situation arises in the case of Maxwell’s equations in Special Relativity. In that case the equations for the vector potential $A^\mu$ are

$$\Box A^\alpha - \frac{\partial^2 A^\beta}{\partial x^\alpha \partial x^\beta} = - \frac{4\pi}{c} J_\alpha. \quad (8.54)$$

(where $\Box = -\frac{\alpha^2}{c^2} \gamma^{\alpha\beta} + \nabla^2 = \eta^{\alpha\beta} \frac{\partial}{\partial x^\alpha} \frac{\partial}{\partial x^\beta}$). These are four equations for the four components of the vector potential. However they do not determine $A^\mu$ uniquely, because of the conservation law

$$J^\mu{}_{,\mu} = 0, \quad \text{i.e.} \quad \frac{\partial}{\partial x^\mu} \left( \Box A^\mu - \eta^{\mu\alpha} \frac{\partial^2 A^\beta}{\partial x^\alpha \partial x^\beta} \right) = 0. \quad (8.55)$$

Equation (8.55) plays the same role as the Bianchi identities do in our context. It provides one condition which must be satisfied by the components of $A^\mu$, therefore the number of independent Maxwell equations is three. The extra degree of freedom corresponds to a gauge invariance, which means the following.

If $A_\alpha$ is a solution,

$$A'_\alpha = A_\alpha + \frac{\partial \Phi}{\partial x^\alpha}, \quad (8.56)$$

will also be a solution. In fact, by direct substitution we find

$$\Box A'_\alpha - \frac{\partial}{\partial x^\alpha} \Box \Phi - \eta^{\beta\delta} \frac{\partial^2 A^\beta}{\partial x^\alpha \partial x^\beta} \frac{\partial \Phi}{\partial x^\delta} = - \frac{4\pi}{c} J_\alpha, \quad (8.57)$$

and since the second and the last term on the left hand-side cancel, it becomes

$$\Box A'_\alpha = \frac{\partial^2 A^\beta}{\partial x^\alpha \partial x^\beta} = - \frac{4\pi}{c} J_\alpha, \quad (8.58)$$

q.e.d.

Since $\Phi$ is arbitrary, we can chose it in such a way that

$$\frac{\partial}{\partial x^\beta} A^{\beta\gamma} = 0 \quad (8.59)$$

Eq. (8.54) is the four-dimensional version of the wave equation for the vector potential

$$\Box \mathbf{A} = \text{grad}(\text{div} \mathbf{A}) = - \frac{4\pi}{c} \mathbf{J}.$$
and eq. (8.58) becomes
\[ \Box A'_\alpha = -\frac{4\pi}{c} J_\alpha, \]
(8.60)

This is the Lorentz gauge.

Summarying: in the electromagnetic case the extra degree of freedom on \( A_\mu \) is due to the fact that the vector potential is defined up to a function \( \Phi \) defined in eq. (8.56). In our case the four extra degrees of freedom are due to the fact that \( g_{\mu\nu} \) is defined up to a coordinate transformation. This gauge freedom is particularly useful when one is looking for exact solutions of Einstein’s equations.
8.4 Example: The armonic gauge.

The armonic gauge is defined by the condition

\[ \Gamma^\lambda = g^{\mu\nu} \Gamma^\lambda_{\mu\nu} = 0. \tag{8.61} \]

As we shall see in a next lecture, this gauge is of particular interest when we study the propagation of gravitational waves, because it simplifies the equations in a way similar to that of Maxwell’s equations when written in the Lorentz gauge. It is always possible to choose this gauge indeed, given a generic coordinate transformation, the affine connections \( \Gamma^\alpha_{\beta\gamma} \) transform as (see eq. (6.1))

\[ \Gamma^\lambda_{\mu\nu}' = \partial_x^\lambda \partial_x^\rho \partial_x^\sigma \Gamma^\rho_{\tau\sigma} \partial_x^\tau \partial_x^\nu \partial_x^\mu \partial_x^\sigma \partial_x^2 x^\lambda. \tag{8.62} \]

When contracted with \( g^{\mu\nu} \) this equation gives

\[ \Gamma^\lambda = \partial_x^\lambda \Gamma^\rho + g^{\rho\sigma} \partial^2 x^\lambda \partial_x^\rho \partial_x^\sigma, \tag{8.63} \]

where we have made use of the relation

\[ g^{\tau\sigma} = g^{\mu\nu} \partial_x^\tau \partial_x^\nu. \tag{8.64} \]

Therefore, if \( \Gamma^\lambda \) is non zero, we can always find a frame where \( \Gamma^\rho = 0 \) and reduce to the armonic gauge. The condition \( \Gamma^\lambda = 0 \) can be rewritten in a more elegant form remembering the expression of the affine connections in terms of the metric tensor

\[ \Gamma^\lambda = \frac{1}{2} g^{\mu\nu} \left\{ \frac{\partial g_{\kappa\mu}}{\partial x^\nu} + \frac{\partial g_{\kappa\nu}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\kappa} \right\} = 0. \tag{8.65} \]

Since

\[ g^{\lambda\kappa} \frac{\partial g_{\kappa\mu}}{\partial x^\nu} = -g_{\kappa\mu} \frac{\partial g^{\lambda\kappa}}{\partial x^\nu}, \tag{8.66} \]

\[ \frac{1}{2} g^{\mu\nu} \frac{\partial g_{\mu\nu}}{\partial x^\kappa} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\kappa} \sqrt{-g}, \]

it follows that

\[ \Gamma^\lambda = \frac{1}{2} g^{\mu\nu} \left\{ -g_{\kappa\mu} \left[ \frac{\partial g^{\lambda\kappa}}{\partial x^\nu} \right] - g_{\kappa\nu} \left[ \frac{\partial g^{\lambda\kappa}}{\partial x^\mu} \right] - \frac{g^{\lambda\kappa}}{\sqrt{-g}} \frac{\partial}{\partial x^\kappa} \sqrt{-g} \right\} = 0. \tag{8.67} \]
The term in brackets is symmetric in $\mu$ and $\nu$, therefore

$$\Gamma^\lambda = -\frac{1}{2} \left\{ 2g^{\mu\sigma} g_{\kappa\mu} \frac{\partial g^{\lambda\kappa}}{\partial x^\sigma} \right\} - g^{\lambda\kappa} \frac{\partial}{\sqrt{-g}} \frac{\partial}{\partial x^\kappa} \sqrt{-g} = 0,$$

(8.68)

and, since $g^{\mu\sigma} g_{\kappa\mu} = \delta^\sigma_\kappa$

$$\Gamma^\lambda = -\frac{\partial g^{\lambda\kappa}}{\partial x^\kappa} - g^{\lambda\kappa} \frac{\partial}{\sqrt{-g}} \frac{\partial}{\partial x^\kappa} \sqrt{-g} = 0,$$

(8.69)

from which we find

$$-\frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\kappa} \left( \sqrt{-g} g^{\lambda\kappa} \right) = 0.$$  

(8.70)

This means that

$$\Gamma^\lambda = 0 \quad \text{implies} \quad \frac{\partial}{\partial x^\kappa} \left( \sqrt{-g} g^{\lambda\kappa} \right) = 0.$$  

(8.71)

The reason why this gauge is called ‘armonic’ is the following. A function $\Phi$ is armonic if

$$\Box \Phi = 0,$$

(8.72)

where the operator $\Box$ is the covariant d’Alambertian operator defined as

$$\Box \Phi = g^{\lambda\kappa} \nabla_\lambda \nabla_\kappa \Phi,$$

(8.73)

and $\nabla_\lambda$ is the covariant derivative. Since

$$g^{\lambda\kappa} \nabla_\lambda \nabla_\kappa \Phi = g^{\lambda\kappa} \left( \frac{\partial \Phi}{\partial x^\kappa} - \Gamma^\alpha_\kappa \Phi_{,\alpha} \right) =$$

$$g^{\lambda\kappa} \left[ \frac{\partial^2 \Phi}{\partial x^\kappa \partial x^\lambda} - \Gamma^\alpha_\kappa \frac{\partial \Phi}{\partial x^\alpha} \right] = g^{\lambda\kappa} \frac{\partial^2 \Phi}{\partial x^\kappa \partial x^\lambda} - \Gamma^\alpha_\kappa \frac{\partial \Phi}{\partial x^\alpha},$$

(8.74)

If $\Gamma^\lambda = 0$ the armonic gauge condition becomes

$$\Box \Phi = g^{\lambda\kappa} \frac{\partial^2 \Phi}{\partial x^\kappa \partial x^\lambda} = 0.$$  

(8.75)

If $\Gamma^\lambda = 0$ then the coordinates itself are armonic functions, in fact putting $\Phi = x^\mu$ in eq. (8.75) one finds

$$\Box x^\mu = g^{\lambda\kappa} \frac{\partial^2 x^\mu}{\partial x^\kappa \partial x^\lambda} = g^{\lambda\kappa} \frac{\partial}{\partial x^\kappa} \delta^\mu_\lambda = 0,$$

(8.76)

q.e.d. If the spacetime is flat, armonic coordinates coincide with minkowskian coordinates.
Chapter 9

Symmetries

H. Weyl: “Symmetry, as wide or as narrow as you may define its meaning, is one idea by which man through the ages has tried to comprehend and create order, beauty, and perfection.”

The solution of a physical problem can be considerably simplified if it allows some symmetries. Let us consider for example the equations of Newtonian gravity. It is easy to find a solution which is spherically symmetric, but it may be difficult to find the analytic solution for an arbitrary mass distribution.

In euclidean space a symmetry is related to an invariance with respect to some operation. For example plane symmetry implies invariance of the physical variables with respect to translations on a plane, spherically symmetric solutions are invariant with respect to translation on a sphere, and the equations of Newtonian gravity are symmetric with respect to time translations

\[ t' \to t + \tau. \]

Thus, a symmetry corresponds to invariance under translations along certain lines or over certain surfaces. This definition can be applied and extended to Riemannian geometry. A solution of Einstein’s equations has a symmetry if there exists an n-dimensional manifold, with \( 1 \leq n \leq 4 \), such that the solution is invariant under translations which bring a point of this manifold into another point of the same manifold. For example, for spherically symmetric solutions the manifold is the 2-sphere, and \( n=2 \). This is a simple example, but there exist more complicated four-dimensional symmetries. These definitions can be made more precise by introducing the notion of Killing vectors.

9.1 The Killing vectors

Consider a vector field \( \tilde{\xi}(x^{\mu}) \) defined at every point \( x^{\alpha} \) of a spacetime region. \( \tilde{\xi} \) identifies a symmetry if an infinitesimal translation along \( \tilde{\xi} \) leaves the line-element unchanged, i.e.

\[ \delta(ds^2) = \delta(g_{\alpha\beta}dx^\alpha dx^\beta) = 0. \]  

(9.1)

This implies that

\[ \delta g_{\alpha\beta}dx^\alpha dx^\beta + g_{\alpha\beta} \left[ \delta(dx^\alpha)dx^\beta + dx^\alpha \delta(dx^\beta) \right] = 0. \]  

(9.2)
\( \xi \) is the tangent vector to some curve \( x^\alpha(\lambda) \), i.e. \( \xi^\alpha = \frac{dx^\alpha}{d\lambda} \), therefore an infinitesimal translation in the direction of \( \xi \) is an infinitesimal translation along the curve from a point \( P \) to the point \( P' \) whose coordinates are, respectively,

\[
P = (x^\alpha) \quad \text{and} \quad P' = (x^\alpha + \delta x^\alpha).
\]

Let us consider, for example, the 2-dimensional space indicated in the following figure

Since

\[
\delta x^1 = \frac{dx^1}{d\lambda} \delta \lambda = \xi^1 \delta \lambda \quad \text{and} \quad \delta x^2 = \frac{dx^2}{d\lambda} \delta \lambda = \xi^2 \delta \lambda
\]
the coordinates of the point \( P' \) can be written as

\[
x'^\alpha = x^\alpha + \xi^\alpha \delta \lambda.
\]

When we move from \( P \) to \( P' \) the metric components change as follows

\[
g_{\alpha\beta}(P') \simeq g_{\alpha\beta}(P) + \frac{\partial g_{\alpha\beta}}{\partial \lambda} \delta \lambda + \ldots
\]

\[
= g_{\alpha\beta}(P) + \frac{\partial g_{\alpha\beta}}{\partial x^\mu} \frac{dx^\mu}{d\lambda} \delta \lambda + \ldots
\]

\[
= g_{\alpha\beta}(P) + g_{\alpha\beta,\mu} \xi^\mu \delta \lambda,
\]

hence

\[
\delta g_{\alpha\beta} = g_{\alpha\beta,\mu} \xi^\mu \delta \lambda.
\]

Moreover, since the operators \( \delta \) and \( d \) commute, we find

\[
\delta(dx^\alpha) = d(\delta x^\alpha) = d(\xi^\alpha \delta \lambda) = d\xi^\alpha \delta \lambda
\]

\[
= \frac{\partial \xi^\alpha}{\partial x^\mu} dx^\mu \delta \lambda = \xi^\alpha_{\mu} dx^\mu \delta \lambda.
\]

Thus, using eqs. (9.7) and (9.6), eq. (9.2) becomes

\[
g_{\alpha\beta,\mu} \xi^\mu \delta \lambda dx^\alpha dx^\beta + g_{\alpha\beta} \left[ \xi^\alpha_{\mu} dx^\mu \delta \lambda dx^\beta + \xi^\beta_{\gamma} dx^\gamma \delta \lambda dx^\alpha \right] = 0,
\]

and, after relabelling the indices,

\[
\left[ g_{\alpha\beta,\mu} \xi^\mu + g_{\delta\beta} \xi^\delta_{\alpha} + g_{\alpha\delta} \xi^\delta_{\beta} \right] dx^\alpha dx^\beta \delta \lambda = 0.
\]
In conclusion, a solution of Einstein’s equations is invariant under translations along $\vec{\xi}$, if and only if

$$g_{\alpha\beta,\mu}\xi^\mu + g_{\delta\beta}\xi^\delta_{,\alpha} + g_{\alpha\delta}\xi^\delta_{,\beta} = 0.$$ \hspace{1cm} (9.10)

In order to find the Killing vectors of a given a metric $g_{\alpha\beta}$ we need to solve eq. (9.10), which is a system of differential equations for the components of $\vec{\xi}$. If eq. (9.10) does not admit a solution, the spacetime has no symmetries. It may look like eq. (9.10) is not covariant, since it contains partial derivatives, but it is easy to show that it is equivalent to the following covariant equation (see appendix A)

$$\xi_{\alpha;\beta} + \xi_{\beta;\alpha} = 0.$$ \hspace{1cm} (9.11)

This is the Killing equation.

### 9.1.1 Lie-derivative

The variation of a tensor under an infinitesimal translation along the direction of a vector field $\vec{\xi}$ is the Lie-derivative ( $\vec{\xi}$ must not necessarily be a Killing vector), and it is indicated as $L_{\vec{\xi}}$. For a $\begin{pmatrix} 0 & 2 \end{pmatrix}$ tensor

$$L_{\vec{\xi}}T_{\alpha\beta} = T_{\alpha\beta,\mu}\xi^\mu + T_{\delta\beta}\xi^\delta_{,\alpha} + T_{\alpha\delta}\xi^\delta_{,\beta}.$$ \hspace{1cm} (9.12)

For the metric tensor

$$L_{\vec{\xi}}g_{\alpha\beta} = g_{\alpha\beta,\mu}\xi^\mu + g_{\delta\beta}\xi^\delta_{,\alpha} + g_{\alpha\delta}\xi^\delta_{,\beta} = \xi_{\alpha;\beta} + \xi_{\beta;\alpha} ;$$ \hspace{1cm} (9.13)

if $\vec{\xi}$ is a Killing vector the Lie-derivative of $g_{\alpha\beta}$ vanishes.

### 9.1.2 Killing vectors and the choice of coordinate systems

The existence of Killing vectors remarkably simplifies the problem of choosing a coordinate system appropriate to solve Einstein’s equations. For instance, if we are looking for a solution which admits a timelike Killing vector $\vec{\xi}$, it is convenient to choose, at each point of the manifold, the timelike basis vector $\vec{e}_{(0)}$ aligned with $\vec{\xi}$; with this choice, the time coordinate lines coincide with the worldlines to which $\vec{\xi}$ is tangent, i.e. with the congruence of worldlines of $\vec{\xi}$, and the components of $\vec{\xi}$ are

$$\xi^\alpha = (\xi^0, 0, 0, 0).$$ \hspace{1cm} (9.14)

If we parametrize the coordinate curves associated to $\vec{\xi}$ in such a way that $\xi^0$ is constant or equal unity, then

$$\xi^\alpha = (1, 0, 0, 0),$$ \hspace{1cm} (9.15)

and from eq. (9.10) it follows that

$$\frac{\partial g_{\alpha\beta}}{\partial x^0} = 0.$$ \hspace{1cm} (9.16)
This means that if the metric admits a timelike Killing vector, with an appropriate choice of the coordinate system it can be made independent of time.

A similar procedure can be used if the metric admits a spacelike Killing vector. In this case, by choosing one of the spacelike basis vectors, say the vector $\vec{e}_{(1)}$, parallel to $\vec{\xi}$, and by a suitable reparametrization of the corresponding congruence of coordinate lines, one can write

$$\xi^\alpha = (0, 1, 0, 0) ,$$

and with this choice the metric is independent of $x^1$, i.e. $\partial g_{\alpha\beta}/\partial x^1 = 0$.

If the Killing vector is null, starting from the coordinate basis vectors $\vec{e}_{(0)}, \vec{e}_{(1)}, \vec{e}_{(2)}, \vec{e}_{(3)}$, it is convenient to construct a set of new basis vectors

$$\vec{e}_{(\alpha')} = \Lambda^\beta_\alpha \vec{e}_{(\beta)} ,$$

such that the vector $\vec{e}_{(0')}$ is a null vector. Then, the vector $\vec{e}_{(0')}$ can be chosen to be parallel to $\vec{\xi}$ at each point of the manifold, and by a suitable reparametrization of the corresponding coordinate lines

$$\xi^\alpha = (1, 0, 0, 0) ,$$

and the metric is independent of $x^{0'}$, i.e. $\partial g_{\alpha\beta}/\partial x^{0'} = 0$.

The map

$$f_t : \mathcal{M} \to \mathcal{M}$$

under which the metric is unchanged is called an isometry, and the Killing vector field is the generator of the isometry.

The congruence of worldlines of the vector $\vec{\xi}$ can be found by integrating the equations

$$\frac{dx^\mu}{d\lambda} = \xi^\mu(x^\alpha) .$$

### 9.2 Examples

1) Killing vectors of flat spacetime

The Killing vectors of Minkowski’s spacetime can be obtained very easily using cartesian coordinates. Since all Christoffel symbols vanish, the Killing equation becomes

$$\xi_{\alpha,\beta} + \xi_{\beta,\alpha} = 0 .$$

By combining the following equations

$$\xi_{\alpha,\beta\gamma} + \xi_{\beta,\alpha\gamma} = 0 , \quad \xi_{\beta,\gamma\alpha} + \xi_{\gamma,\beta\alpha} = 0 , \quad \xi_{\gamma,\alpha\beta} + \xi_{\alpha,\gamma\beta} = 0 ,$$

and by using eq. (9.21) we find

$$\xi_{\alpha,\beta\gamma} = 0 ,$$

whose general solution is

$$\xi^\alpha = c_\alpha + \epsilon_{\alpha\gamma} x^\gamma ,$$
where \( c_\alpha, \epsilon_{\alpha\beta} \) are constants. By substituting this expression into eq. (9.21) we find

\[
\epsilon_{\alpha\gamma} x_\beta^{\gamma} + \epsilon_{\beta\gamma} x_\alpha^{\gamma} = \epsilon_{\alpha\gamma} \delta_\beta^{\gamma} + \epsilon_{\beta\gamma} \delta_\alpha^{\gamma} = \epsilon_{\alpha\beta} + \epsilon_{\beta\alpha} = 0
\]

Therefore eq. (9.24) is the solution of eq. (9.21) only if

\[
\epsilon_{\alpha\beta} = -\epsilon_{\beta\alpha} .
\]

(9.25)

The general Killing vector field of the form (9.24) can be written as the linear combination of ten Killing vector fields \( \xi_{\alpha}^{(A)} = \{ \xi_{\alpha}^{(1)}, \xi_{\alpha}^{(2)}, \ldots, \xi_{\alpha}^{(10)} \} \) corresponding to ten independent choices of the constants \( c_\alpha, \epsilon_{\alpha\beta} \):

\[
\xi_{\alpha}^{(A)} = c_{\alpha}^{(A)} + \epsilon_{\alpha\gamma}^{(A)} x^{\gamma} \quad A = 1, \ldots, 10 .
\]

(9.26)

For instance, we can choose

<table>
<thead>
<tr>
<th>( \xi_{\alpha}^{(1)} )</th>
<th>( \xi_{\alpha}^{(2)} )</th>
<th>( \xi_{\alpha}^{(3)} )</th>
<th>( \xi_{\alpha}^{(4)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 0, 0, 0)</td>
<td>(0, 1, 0, 0)</td>
<td>(0, 0, 1, 0)</td>
<td>(0, 0, 0, 1)</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\epsilon_{\alpha\beta}^{(1)} &= 0 \\
\epsilon_{\alpha\beta}^{(2)} &= 0 \\
\epsilon_{\alpha\beta}^{(3)} &= 0 \\
\epsilon_{\alpha\beta}^{(4)} &= 0
\end{align*}
\]

<table>
<thead>
<tr>
<th>( \xi_{\alpha}^{(5)} )</th>
<th>( \xi_{\alpha}^{(6)} )</th>
<th>( \xi_{\alpha}^{(7)} )</th>
<th>( \xi_{\alpha}^{(8)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 0, 1, 0)</td>
<td>(0, 0, 0, 0)</td>
<td>(0, 0, 0, 0)</td>
<td>(0, 0, 0, 1)</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\epsilon_{\alpha\beta}^{(5)} &= \begin{pmatrix}
0 & 1 & 0 & 0 \\
-1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix} \\
\epsilon_{\alpha\beta}^{(6)} &= \begin{pmatrix}
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix} \\
\epsilon_{\alpha\beta}^{(7)} &= \begin{pmatrix}
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
-1 & 0 & 0 & 0
\end{pmatrix} \\
\epsilon_{\alpha\beta}^{(8)} &= \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix} \\
\epsilon_{\alpha\beta}^{(9)} &= \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0
\end{pmatrix} \\
\epsilon_{\alpha\beta}^{(10)} &= \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & -1 & 0
\end{pmatrix}
\]
Therefore, flat spacetime admits ten linearly independent Killing vectors. The symmetries generated by the Killing vectors with \( A = 1, \ldots, 4 \) are spacetime translations; the symmetries generated by the Killing vectors with \( A = 5, 6, 7 \) are Lorentz’s boosts; the symmetries generated by the Killing vectors with \( A = 8, 9, 10 \) are space rotations.

2) Killing vectors of a spherical surface

Let us consider a sphere of unit radius

\[
 ds^2 = d\theta^2 + \sin^2 \theta d\varphi^2 = (dx^1)^2 + \sin^2 x^1(dx^2)^2.
\]

Eq. (9.10)

\[
 g_{\alpha\beta,\mu} \xi^\mu + g_{\delta\beta} \xi^\delta + g_{\alpha\delta} \xi^\delta = 0
\]
gives

1) \( \alpha = \beta = 1 \)

\[
 2g_{\delta1} \xi^\delta = 0 \rightarrow \xi^1 = 0 \quad (9.29)
\]

2) \( \alpha = 1, \beta = 2 \)

\[
 g_{\delta2} \xi^\delta + g_{1\delta} \xi^\delta = 0 \rightarrow \xi^1 + \sin^2 \theta \xi^2 = 0
\]

3) \( \alpha = \beta = 2 \)

\[
 g_{22,\mu} \xi^\mu + 2g_{\delta2} \xi^\delta = 0 \rightarrow \cos \theta \xi^1 + \sin \theta \xi^2 = 0.
\]

The general solution is

\[
 \xi^1 = A \sin(\varphi + a), \quad \xi^2 = A \cos(\varphi + a)cot\theta + b. \quad (9.30)
\]

Therefore a spherical surface admits three linearly independent Killing vectors, associated to the choice of the integration constants \((A, a, b)\).

### 9.3 Conserved quantities in geodesic motion

Killing vectors are important because they are associated to conserved quantities, which may be hidden by an unsuitable coordinate choice.

Let us consider a massive particle moving along a geodesic of a spacetime which admits a Killing vector \( \vec{\xi} \). The geodesic equations written in terms of the particle four-velocity \( \vec{U} = \frac{dx}{d\tau} \) read

\[
 \frac{dU^\alpha}{d\tau} + \Gamma^\alpha_{\beta\nu} U^\beta U^\nu = 0. \quad (9.31)
\]

By contracting eq. (9.31) with \( \vec{\xi} \) we find

\[
 \xi_\alpha \left[ \frac{dU^\alpha}{d\tau} + \Gamma^\alpha_{\beta\nu} U^\beta U^\nu \right] = \frac{d(\xi_\alpha U^\alpha)}{d\tau} - U^\alpha \frac{d\xi_\alpha}{d\tau} + \Gamma^\alpha_{\beta\nu} U^\beta U^\nu \xi_\alpha.
\]

Since

\[
 U^\alpha \frac{d\xi_\alpha}{d\tau} = U^\beta \frac{d\xi_\beta}{d\tau} = U^\beta \frac{\partial \xi_\beta}{\partial x^\nu} \frac{dx^\nu}{d\tau} = U^\beta U^\nu \frac{\partial \xi_\beta}{\partial x^\nu},
\]

eq. (9.32) becomes

\[
 \frac{d(\xi_\alpha U^\alpha)}{d\tau} - U^\beta U^\nu \left[ \frac{\partial \xi_\beta}{\partial x^\nu} - \Gamma^\alpha_{\beta\nu} \xi_\alpha \right] = 0,
\]

(9.34)
i.e.

\[
\frac{d(\xi^\alpha U^\alpha)}{d\tau} - U^\beta U^\nu \xi_{\beta\nu} = 0 .
\] (9.35)

Since \(\xi_{\beta\nu}\) is antisymmetric in \(\beta\) and \(\nu\), while \(U^\beta U^\nu\) is symmetric, the term \(U^\beta U^\nu \xi_{\beta\nu}\) vanishes, and eq. (9.35) finally becomes

\[
\frac{d(\xi^\alpha U^\alpha)}{d\tau} = 0 \quad \rightarrow \quad \xi^\alpha U^\alpha = \text{const} ,
\] (9.36)

i.e. the quantity \((\xi^\alpha U^\alpha)\) is a constant of the particle motion. Thus, for every Killing vector there exists an associated conserved quantity.

Eq. (9.36) can be written as follows:

\[
g_{\alpha\beta} \xi^\beta U^\alpha = \text{const} .
\] (9.37)

Let us now assume that \(\xi^\beta\) is a timelike Killing vector. In section 9.1.2 we have shown that the coordinate system can be chosen in such a way that \(\xi^\beta = \{1, 0, 0, 0\}\), in which case eq. (9.37) becomes

\[
g_{0\alpha} \xi^0 U^\alpha = \text{const} \quad \rightarrow \quad g_{00} U^0 = \text{const} .
\] (9.38)

If the metric is asymptotically flat, as it is for instance when the gravitational field is generated by a distribution of matter confined in a finite region of space, at infinity \(g_{\alpha\beta}\) reduces to the Minkowski metric \(\eta_{\alpha\beta}\), and eq. (9.38) becomes

\[
\eta_{00} U^0 = \text{const} \quad \rightarrow \quad U^0 = \text{const} .
\] (9.39)

Since in flat spacetime the energy-momentum vector of a massive particle is \(p^\alpha = mcU^\alpha = \{E/c, mv^i\gamma\}\), the previous equation becomes

\[
\frac{E}{c} = \text{const} ,
\] (9.40)

i.e. at infinity the conservation law associated to a timelike Killing vector reduces to the energy conservation for the particle motion. For this reason we say that, when the metric admits a timelike Killing vector, eq. (9.36) expresses the energy conservation for the particle motion along the geodesic.

If the Killing vector is spacelike, by choosing the coordinate system such that, say, \(\xi^\mu = \{0, 1, 0, 0\}\), eq. (9.36) reduces to

\[
g_{01} \xi^1 U^\alpha = \text{const} \quad \rightarrow \quad g_{01} U^1 = \text{const} .
\]

At infinity this equation becomes

\[
\eta_{11} U^1 = \text{const} \quad \rightarrow \quad \frac{p^1}{mc} = \text{const} ,
\]

showing that the component of the energy-momentum vector along the \(x^1\) direction is constant; thus, when the metric admit a spacelike Killing vector eq. (9.36) expresses momentum conservation along the geodesic motion.
If the particle is massless, the geodesic equation cannot be parametrized with the proper
time. In this case the particle worldline has to be parametrized using an affine parameter
λ such that the geodesic equation takes the form (9.31), and the particle four-velocity is
\( U^\alpha = \frac{dx^\alpha}{d\lambda} \). The derivation of the constants of motion associated to a spacetime symmetry,
i.e. to a Killing vector, is similar as for massive particles, reminding that by a suitable choice
of the parameter along the geodesic \( p^\alpha = \{E, p^i\} \).

It should be mentioned that in Riemannian spaces there may exist conservation laws
which cannot be traced back to the presence of a symmetry, and therefore to the existence
of a Killing vector field.

### 9.4 Killing vectors and conservation laws

In Chapter 7 we have shown that the stress-energy tensor satisfies the “conservation law”

\[
T^\mu{}_{\nu;\nu} = 0, \quad (9.41)
\]

and we have shown that in general this is not a genuine conservation law. If the spacetime
admits a Killing vector, then

\[
(\xi^\mu T^\mu{}_{\nu})_{;\nu} = \xi^\mu{}_{;\nu} T^\mu{}_{\nu} + \xi^\mu T^\mu{}_{;\nu} = 0. \quad (9.42)
\]

Indeed, the second term vanishes because of eq. (9.41) and the first vanishes because \( \xi^\mu{}_{;\nu} \)
is antisymmetric in \( \mu \) and \( \nu \), whereas \( T^\mu{}_{\nu} \) is symmetric.

Since there is a contraction on the index \( \mu \), the quantity \( (\xi^\mu T^\mu{}_{\nu}) \) is a vector, and according
to eq. (7.68)

\[
V^\nu{}_{;\nu} = \frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\nu} \left( \sqrt{-g} V^\nu \right), \quad (9.43)
\]

therefore eq. (9.42) is equivalent to

\[
\frac{1}{\sqrt{-g}} \frac{\partial}{\partial x^\nu} \left[ \sqrt{-g} (\xi^\mu T^\mu{}_{\nu}) \right] = 0, \quad (9.44)
\]

which expresses the conservation of the following quantity and accordingly, a conserved
quantity can be defined as

\[
T = \int_{(x^0 = \text{const})} \sqrt{-g} \left( \xi^\mu T^{\mu0} \right) \, dx^1 dx^2 dx^3, \quad (9.45)
\]
as shown in Chapter 7.

In classical mechanics energy is conserved when the hamiltonian is independent of time;
thus, conservation of energy is associated to a symmetry with respect to time translations.
In section 9.1.2 we have shown that if a metric admits a timelike Killing vector, with a
suitable choice of coordinates it can me made time independent (where now “time” indicates
more generally the \( x^0 \)-coordinate). Thus, in this case it is natural to interpret the quantity
defined in eq. (9.45) as a conserved energy.

In a similar way, when the metric adimits a spacelike Killing vector, the associated
conserved quantities are indicated as “momentum” or “angular momentum”, although this
is more a matter of definition.

It should be stressed that the energy of a gravitational system can be defined in a non
ambiguous way only if there exists a timelike Killing vector field.
9.5 Hypersurface orthogonal vector fields

Given a vector field $\vec{V}$ it identifies a congruence of worldlines, i.e. the set of curves to which the vector is tangent at any point of the considered region. If there exists a family of surfaces $f(x^\mu) = const$ such that, at each point, the worldlines of the congruence are perpendicular to that surface, $\vec{V}$ is said to be hypersurface orthogonal. This is equivalent to require that $\vec{V}$ is orthogonal to all vectors $\vec{t}$ tangent to the hypersurface, i.e.

$$t^\alpha V^\beta g_{\alpha\beta} = 0$$ \hspace{1cm} (9.46)

We shall now show that, as consequence, $\vec{V}$ is parallel to the gradient of $f$. As described in Chapter 3, section 5, the gradient of a function $f(x^\mu)$ is a one-form

$$df \rightarrow \left( \frac{\partial f}{\partial x^0}, \frac{\partial f}{\partial x^1}, ..., \frac{\partial f}{\partial x^n} \right) = \{f,\alpha\}.$$ \hspace{1cm} (9.47)

When we say that $\vec{V}$ is parallel to $\tilde{df}$ we mean that the one-form dual to $\vec{V}$, i.e. $\tilde{V} \rightarrow \{g_{\alpha\beta}V^\beta \equiv V^\alpha\}$ satisfies the equation

$$V^\alpha = \lambda f,\alpha,$$ \hspace{1cm} (9.48)

where $\lambda$ is a function of the coordinates $\{x^\mu\}$. This equation is equivalent to eq. (9.46). Indeed, given any curve $x^\alpha(s)$ lying on the hypersurface, and being $t^\alpha = dx^\alpha / ds$ its tangent vector, since $f(x^\mu) = const$ the directional derivative of $f(x^\mu)$ along the curve vanishes, i.e.

$$\frac{df}{ds} = \frac{\partial f}{\partial x^\alpha} \frac{dx^\alpha}{ds} = f,\alpha t^\alpha = \lambda^{-1}V^\alpha t^\alpha = 0,$$ \hspace{1cm} (9.49)

i.e. eq. (9.46).

If (9.48) is satisfied, it follows that

$$V_{\alpha;\beta} - V_{\beta;\alpha} = (\lambda f,\alpha)_{;\beta} - (\lambda f,\beta)_{;\alpha}$$

$$= \lambda (f,\alpha;\beta - f,\beta;\alpha) + f,\alpha \lambda,\beta - f,\beta \lambda,\alpha =$$

$$= \lambda (f,\alpha;\beta - f,\beta;\alpha - \Gamma^\mu_{\beta\alpha}f,\mu + \Gamma^\mu_{\alpha\beta}f,\mu) + f,\alpha \lambda,\beta - f,\beta \lambda,\alpha$$

$$= V^\alpha \frac{\lambda,\beta}{\lambda} - V^\beta \frac{\lambda,\alpha}{\lambda},$$

i.e.

$$V_{\alpha;\beta} - V_{\beta;\alpha} = V^\alpha \frac{\lambda,\beta}{\lambda} - V^\beta \frac{\lambda,\alpha}{\lambda}.$$ \hspace{1cm} (9.51)

If we now define the following quantity, which is said rotation

$$\omega^\delta = \frac{1}{2} \epsilon^{\delta\alpha\beta\mu} V_{[\alpha;\beta]} V_\mu,$$ \hspace{1cm} (9.52)

using the definition of the antisymmetric unit pseudotensor $\epsilon^{\delta\alpha\beta\mu}$ given in Appendix B, it follows that

$$\omega^\delta = 0.$$ \hspace{1cm} (9.53)

Then, if the vector field $\vec{V}$ is hypersurface orthognal, (9.53) is satisfied. Actually, (9.53) is a necessary and sufficient condition for $\vec{V}$ to be hypersurface orthogonal; this result is the Frobenius theorem.
### 9.5.1 Hypersurface-orthogonal vector fields and the choice of coordinate systems

The existence of a hypersurface-orthogonal vector field allows to choose a coordinate frame such that the metric has a much simpler form. Let us consider, for the sake of simplicity, a three-dimensional spacetime \((x^0, x^1, x^2)\).

Be \(S_1\) and \(S_2\) two surfaces of the family \(f(x^\mu) = \text{cost}\), to which the vector field \(\vec{V}\) is orthogonal. As an example, we shall assume that \(\vec{V}\) is timelike, but a similar procedure can be used if \(\vec{V}\) is spacelike. If \(\vec{V}\) is timelike, it is convenient to choose the basis vector \(\vec{e}_{(0)}\) parallel to \(\vec{V}\), and the remaining basis vectors as the tangent vectors to some curves lying on the surface, so that

\[
\begin{align*}
g_{00} &= g(\vec{e}_{(0)}, \vec{e}_{(0)}) = \vec{e}_{(0)} \cdot \vec{e}_{(0)} \neq 0 \quad (9.54) \\
g_{0i} &= g(\vec{e}_{(0)}, \vec{e}_{(i)}) = 0, \quad i = 1, 2.
\end{align*}
\]

Thus, with this choice, the metric becomes

\[
ds^2 = g_{00}(dx^0)^2 + g_{ik}(dx^i)(dx^k), \quad i, k = 1, 2. \quad (9.55)
\]

The generalization of this example to the four-dimensional spacetime, in which case the surface \(S\) is a hypersurface, is straightforward.

In general, given a timelike vector field \(\vec{V}\), we can always choose a coordinate frame such that \(\vec{e}_{(0)}\) is parallel to \(\vec{V}\), so that in this frame

\[
V^\alpha(x^\mu) = (V^0(x^\mu), 0, 0, 0). \quad (9.56)
\]

Such coordinate system is said **comoving**. If, in addition, \(\vec{V}\) is hypersurface-orthogonal, then \(g_{0i} = 0\) and, as a consequence, the one-form associated to \(\vec{V}\) also has the form

\[
V_\alpha(x^\mu) = (V_0(x^\mu), 0, 0, 0), \quad (9.57)
\]

since \(V_i = g_{\mu}V^\mu = g_{00}V^0 + g_{ik}V^k = 0\).
9.6 Appendix A

We want to show that eq. (9.10) is equivalent to eq. (9.11).

\[
\begin{align*}
\xi_{\alpha;\beta} &= (g_{\alpha\mu} \xi^\mu)_{;\beta} \\
&= g_{\alpha\mu} \xi^\mu_{;\beta} = g_{\alpha\mu} \left( \xi^\mu_{;\beta} + \Gamma^\mu_{\delta\beta} \xi^\delta \right),
\end{align*}
\]

hence

\[
\begin{align*}
\xi_{\alpha;\beta} + \xi_{\beta;\alpha} &= g_{\alpha\mu} \left( \xi^\mu_{;\beta} + \Gamma^\mu_{\delta\beta} \xi^\delta \right) \\
&+ g_{\beta\mu} \left( \xi^\mu_{;\alpha} + \Gamma^\mu_{\alpha\delta} \xi^\delta \right) \\
&= g_{\alpha\mu} \xi^\mu_{;\beta} + g_{\beta\mu} \xi^\mu_{;\alpha} + (g_{\alpha\mu} \Gamma^\mu_{\delta\beta} + g_{\beta\mu} \Gamma^\mu_{\alpha\delta}) \xi^\delta.
\end{align*}
\]

The term in parenthesis can be written as

\[
\begin{align*}
&\frac{1}{2} \left[ g_{\alpha\mu} g_{\mu\sigma} \left( g_{\delta\sigma,\beta} + g_{\sigma\beta,\delta} - g_{\delta\beta,\sigma} \right) + g_{\beta\mu} g_{\mu\sigma} \left( g_{\alpha\sigma,\delta} + g_{\sigma\delta,\alpha} - g_{\alpha\delta,\sigma} \right) \right] \\
= &\frac{1}{2} \left[ \delta_\sigma^\alpha \left( g_{\delta\sigma,\beta} + g_{\sigma\beta,\delta} - g_{\delta\beta,\sigma} \right) + \delta_\beta^\sigma \left( g_{\alpha\sigma,\delta} + g_{\sigma\delta,\alpha} - g_{\alpha\delta,\sigma} \right) \right] \\
= &\frac{1}{2} \left[ g_{\delta\sigma,\beta} + g_{\alpha\beta,\delta} - g_{\delta\beta,\alpha} + g_{\alpha\beta,\delta} + g_{\beta\delta,\alpha} - g_{\beta\delta,\beta} \right] \\
= &g_{\alpha\beta,\delta},
\end{align*}
\]

and eq. (9.59) becomes

\[
\xi_{\alpha;\beta} + \xi_{\beta;\alpha} = g_{\alpha\mu} \xi^\mu_{;\beta} + g_{\beta\mu} \xi^\mu_{;\alpha} + g_{\alpha\beta,\delta} \xi^\delta
\]

which coincides with eq. (9.10).

9.7 Appendix B: The Levi-Civita completely antisymmetric pseudotensor

We define the Levi-Civita symbol (also said Levi-Civita tensor density), \( e_{\alpha\beta\gamma\delta} \), as an object whose components change sign under interchange of any pair of indices, and whose non-zero components are \( \pm 1 \). Since it is completely antisymmetric, all the components with two equal indices are zero, and the only non-vanishing components are those for which all four indices are different. We set

\[
e_{0123} = 1.
\]

Under general coordinate transformations, \( e_{\alpha\beta\gamma\delta} \) does not transform as a tensor; indeed, under the transformation \( x^\alpha \to x'^\alpha \),

\[
\partial x^\alpha \partial x^\beta \partial x^\gamma \partial x^\delta \frac{\partial x^\alpha}{\partial x'^\alpha} \frac{\partial x^\beta}{\partial x'^\beta} \frac{\partial x^\gamma}{\partial x'^\gamma} \frac{\partial x^\delta}{\partial x'^\delta} e_{\alpha\beta\gamma\delta} = J e_{\alpha'\beta'\gamma'\delta'}
\]

where \( J \) is defined (see Chapter 7) as

\[
J \equiv \det \left( \frac{\partial x^\alpha}{\partial x'^\alpha} \right)
\]
and we have used the definition of determinant.

We now define the Levi-Civita pseudo-tensor as

$$\epsilon_{\alpha\beta\gamma\delta} \equiv \sqrt{-g} \epsilon_{\alpha\beta\gamma\delta}.$$  \hspace{1cm} (9.65)

Since, from (7.26), for a coordinate transformation $x^\alpha \to x'^\alpha$

$$|J| = \frac{\sqrt{-g'}}{\sqrt{-g}},$$  \hspace{1cm} (9.66)

then

$$\epsilon_{\alpha\beta\gamma\delta} \rightarrow \epsilon_{\alpha'\beta'\gamma'\delta'} = \text{sign}(J) \frac{\partial x'^\alpha}{\partial x^{\alpha'}} \frac{\partial x'^\beta}{\partial x^{\beta'}} \frac{\partial x'^\gamma}{\partial x^{\gamma'}} \frac{\partial x'^\delta}{\partial x^{\delta'}} \epsilon_{\alpha\beta\gamma\delta}.$$  \hspace{1cm} (9.67)

Thus, $\epsilon_{\alpha\beta\gamma\delta}$ is not a tensor but a pseudo-tensor, because it transforms as a tensor times the sign of the Jacobian of the transformation. It transforms as a tensor only under a subset of the general coordinate transformations, i.e. that with $\text{sign}(J) = +1$.

Warning: do not confuse the Levi-Civita symbol, $\epsilon_{\alpha\beta\gamma\delta}$, with the Levi-Civita pseudo-tensor, $\epsilon_{\alpha\beta\gamma\delta}$. 

Chapter 10

The Schwarzschild solution

The Schwarzschild solution was first derived by Karl Schwarzschild in 1916, although a complete understanding of the Schwarzschild spacetime was achieved much recently. The paper was communicated to the Berlin Academy by Einstein on 13 January 1916, just about two months after he had published the seminal papers on the theory of General Relativity. In those years Schwarzschild was very ill. He had contracted a fatal disease in 1915 while serving the German army at the eastern front. He died on 11 May 1916, and during his illness he wrote two papers in General Relativity, one describing the solution for the gravitational field exterior to a spherically symmetric non rotating body, which we are going to derive, and the second describing the interior solution for a star of constant density which we shall discuss later.

We now want to find an exact solution of Einstein’s equations in vacuum, which is spherically symmetric and static. This will be the relativistic generalization of the newtonian solution for a pointlike mass

\[ V = -\frac{GM}{r}, \]  

(10.1)

and it will describe the gravitational field in the exterior of a non rotating body. Let us first discuss the symmetries of the problem.

10.1 The symmetries of the problem

a) Symmetry with respect to time.

Time-symmetric spacetimes can be stationary or static. A spacetime is said to be stationary if it admits a timelike Killing vector \( \xi \). It follows from the Killing equations that the metric of a stationary spacetime does not depend on time

\[ \frac{\partial g_{\alpha\beta}}{\partial x^0} = 0. \]  

(10.2)

A spacetime is static if it admits a hypersurface-orthogonal, timelike Killing vector. In this case, as shown in Chapter 12, we can choose the coordinates in such a way that
\[ \vec{\xi} \to (1, 0, 0, 0), \] and the line-element takes the simple form

\[ ds^2 = g_{00}(x^i)(dx^0)^2 + g_{kn}(x^i)dx^kd\bar{x}^n, \quad i, k, n = 1, 3, \]

where \[ g_{00} = g(\vec{\xi}, \vec{\xi}) = \vec{\xi} \cdot \vec{\xi}. \]

From this equation we see that the metric is not only independent on time, but also invariant under time reversal \( t \to -t \). (If terms like \( dx^0 dx^i \) were present this would not be true).

b) Spatial symmetry.

We now take care of the spatial part of the metric. The basic idea is that we want to “fill” the space with concentric spherical surfaces. We start with the 2-sphere of radius \( a \) in flat space

\[ ds^2_{(2)} = g_{22}(dx^2)^2 + g_{33}(dx^3)^2 = a^2(d\theta^2 + \sin^2 \theta d\varphi^2). \]

The surface of this sphere is

\[ A = \int \sqrt{g} d\theta d\varphi = \int_0^\pi a^2 \sin \theta d\theta \int_0^{2\pi} d\varphi = 4\pi a^2, \]

and the length of the circumference

\[ \theta = \frac{\pi}{2}, \quad dl = ad\varphi, \quad C = 2\pi a. \]

These results continue to hold if \( a \) is an arbitrary function of the remaining coordinates \( x^0, x^1 \)

\[ ds^2_{(2)} = a^2(x^0, x^1)(d\theta^2 + \sin^2 \theta d\varphi^2). \]

But since we have already established that the metric does not depend on time, we put \( a = a(x^1) \). We are now free to make a coordinate transformation and put

\[ r = a(x^1). \]

Thus we define the radial coordinate as being half the ratio between the surface and the circumference of the 2-sphere. However, it should be noted that in principle the coordinate \( r \) has nothing to do with the distance between the center of the sphere and the surface, as we shall later show.

Then we go to the next sphere at \( r + dr \). We may label the points of the second sphere with different \( (\theta', \varphi') \) as indicated in the figure.
If the poles of the two spheres are not aligned, the vector \( \vec{\eta} \) which maps the point \( P = (\theta_0, \varphi_0) \) on the internal sphere \( (\theta_0, \varphi_0 \text{ constants}) \), to the point \( P' = (\theta'_0, \varphi'_0) \) on the external sphere \( (\text{with } \theta'_0 = \theta_0 \text{ and } \varphi'_0 = \varphi_0) \), is directed as indicated in the figure. Conversely, if the poles are aligned \( \vec{\eta} \) is orthogonal to the two spheres, and therefore it is orthogonal to \( \frac{\partial}{\partial \theta} = \vec{e}_\theta \) and \( \frac{\partial}{\partial \varphi} = \vec{e}_\varphi \), which are the basis vectors on the sphere. Thus in this case \( \vec{\eta} \) is \textbf{hypersurface-orthogonal}. Since we want angular coordinates \( (\theta, \varphi) \) defined in a unique way on the whole set of spheres filling the space, we require that \( \vec{\eta} \) is indeed orthogonal to the spheres. In this case \( \vec{\eta} \) is the vector tangent to the coordinate line \( (\theta = \text{const}, \varphi = \text{const}) \), therefore \( \vec{\eta} = \frac{\partial}{\partial r} = \vec{e}_r \). The orthogonality condition then gives
\[
\vec{e}_r \cdot \vec{e}_\theta = g_{rr} = 0, \quad \vec{e}_r \cdot \vec{e}_\varphi = g_{r\varphi} = 0.
\] (10.9)

Under these assumptions, the metric of the three-space becomes
\[
ds^2 = g_{rr} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2),
\] (10.10)
and that of the four-dimensional spacetime finally is
\[
ds^2 = g_{00} (dx^0)^2 + g_{rr} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2).
\] (10.11)

At this point the two metric components \( g_{00} \) and \( g_{rr} \) should, in principle, depend on \( (r, \theta, \phi) \). However, this is not the case. Indeed, if we consider a set of new polar coordinates \( (\theta', \phi') \) to label the points on the two spheres that fill the space, neither the vector \( \vec{e}_0 \), nor the vector \( \vec{e}_r \) will change and therefore they cannot depend on the angular coordinates we choose. As a consequence \( g_{00} \) and \( g_{rr} \) do not depend on \( (\theta, \phi) \) either, and we can write
\[
g_{00} = g_{00}(r), \quad \text{and} \quad g_{rr} = g_{rr}(r).
\]

It is convenient to rewrite the metric in the following form
\[
ds^2 = -e^{2\nu} (dx^0)^2 + e^{2\lambda} dr^2 + r^2 (d\theta^2 + \sin^2 \theta d\varphi^2),
\] (10.12)
where \( \nu = \nu(r) \) and \( \lambda = \lambda(r) \). Let us now compute the distance between two points \( P1 = (x^0_s, r_1, \theta_s, \varphi_s) \) and \( P2 = (x^0_s, r_2, \theta_s, \varphi_s) \)

\[
l = \int_{r_1}^{r_2} e^\lambda dr \tag{10.13}
\]
(we can compute this finite length because we are in a time-independent spacetime). This distance does not coincide with \( (r_2 - r_1) \).

We now write the components of the Einstein tensor in terms of the metric (10.13). They are

\[
a) \quad G_{00} = \frac{1}{r^2} e^{2\nu} \frac{d}{dr} \left[ r(1 - e^{-2\lambda}) \right] \tag{10.14}
b) \quad G_{rr} = -\frac{1}{r^2} e^{2\lambda} \left[ (1 - e^{-2\lambda}) \right] + \frac{2}{r} \nu_r 
c) \quad G_{\theta\theta} = r^2 e^{-2\lambda} \left[ \nu_{rr} + \nu_r^2 + \frac{\nu_r}{r} - \nu_r \lambda_r - \frac{\lambda_r}{r} \right] 
d) \quad G_{\varphi\varphi} = \sin^2 \theta G_{\theta\theta}
\]

The remaining components identically vanish. Since we are looking for a vacuum solution, the equations to solve are

\[
G_{\mu\nu} = 0, \quad \tag{10.15}
\]
and eq. (10.14a) gives

\[
r(1 - e^{-2\lambda}) = K, \quad \tag{10.16}
\]
where \( K \) is an integration constant. Hence

\[
e^{2\lambda} = \frac{1}{1 - \frac{K}{r}}, \quad \tag{10.17}
\]

From eq. (10.14b) we find

\[
\nu_r = \frac{1}{2} \frac{K}{r(r - K)}, \quad \tag{10.18}
\]
and therefore

\[
\nu = \frac{1}{2} \log \left( 1 - \frac{K}{r} \right) + \nu_0, \quad \rightarrow \quad e^{2\nu} = \left( 1 - \frac{K}{r} \right) e^{2\nu_0}, \quad \tag{10.19}
\]
where \( \nu_0 \) is a constant. We can rescale the time coordinate

\[
t \rightarrow e^{\nu_0} t,
\]
in such a way that \( e^{2\nu} \) becomes

\[
e^{2\nu} = 1 - \frac{K}{r}, \quad \tag{10.20}
\]
The final form of the solution is

\[
ds^2 = -\left( 1 - \frac{K}{r} \right) e^{2\nu} dt^2 + \frac{1}{1 - \frac{K}{r}} dr^2 + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right). \quad \tag{10.21}
\]
This is the **Schwarzschild solution**. When \( r \to \infty \) the metric reduces to that of a flat spacetime, therefore we say that the metric is **asymptotically flat**.

Now we want to understand what is the meaning of the integration constant \( K \). In Chapter 8 section 1, we showed that in the weak-field limit, the geodesic equations reduce to the newtonian equations of motion, and consequently

\[
g_{00} \sim - \left(1 + \frac{2\Phi}{c^2}\right) = - \left(1 - \frac{2GM}{c^2r}\right), \quad \text{where}
\]

\( \Phi = -\frac{GM}{r} \) is the newtonian potential generated by a spherical distribution of matter. From eq. (10.20) we see that when \( r \to \infty \) \( g_{00} \) tends to unity as

\[
-g_{00} = e^{2\nu} = 1 - \frac{K}{r}.
\]

By comparing eq. (10.22) and (10.23) we find

\[
K = \frac{2GM}{c^2}.
\]

Therefore the constant \( K \) is the physical mass multiplied by \( \frac{2G}{c^2} \). It is easy to check that the solution (10.21) satisfies eq. (10.14c).

### 10.2 The Birkhoff theorem

The solution (10.21) has been found by imposing that the spacetime is static and spherically symmetric, therefore it represents the gravitational field external to a non-rotating, spherically symmetric body whose structure is time-independent. However, it is more general than that. In fact Birkhoff’s theorem establishes that it is the only spherically symmetric and asymptotically flat solution of the vacuum Einstein field equations. Let us assume that the functions \((\nu, \lambda)\) in the metric (10.12) depend both on the radial coordinate \( r \) and on time.

To prove Birkhoff’s theorem we only need the components \( R_{0r} \) and \( R_{\theta\theta} \) of the Ricci tensor:

\[
a) \quad R_{0r} = \frac{2}{r} \frac{\partial \lambda}{\partial x^0} = 0,
\]

\[
b) \quad R_{\theta\theta} = 1 - e^{-2\lambda} \left[1 + r \frac{\partial(\nu - \lambda)}{\partial r}\right] = 0.
\]

From eq. (10.25a) it follows that \( \lambda \) must depend only upon the radial coordinate \( r \). Then from eq. (10.25b) it follows that also \( \frac{\partial \nu}{\partial r} \) must be independent on \( x^0 \) and consequently

\[
\nu = \nu(r) + f(x^0).
\]

This means that the coefficient of \( (dx^0)^2 \) in the line element is \( e^{2\nu(r)}e^{2f(x^0)} \). But the term \( e^{2f(x^0)} \) can be ‘reabsorbed’ by a coordinate transformation

\[
dt' = e^{f(x^0)}dt,
\]
so that the new metric coefficients are
\[ \nu = \nu(r), \quad \lambda = \lambda(r), \quad (10.28) \]
and the metric is time independent. This means that even if we impose that the central object evolves in time, as it would be for example in the case of a star radially pulsating, or in a spherical collapse, we would find, in the exterior, the same Schwarzschild metric, and since the spacetime remains static even in these cases, gravitational waves could not be emitted. The conclusion is that spherically symmetric systems can never emit gravitational waves. A similar situation occurs in electrodynamics: a spherically symmetric distribution of charges and currents does not radiate.

10.3 Geometrized unities

From eq. (10.23) and (10.24) is easy to see that \( K = \frac{2GM}{c^2} \) must have the dimension of a length. In fact the ratio \( \frac{G}{c^2} \) is
\[ \frac{G}{c^2} = 0.7425 \times 10^{-28} \text{cm} \cdot \text{g}^{-1}. \quad (10.29) \]
It is often convenient to put
\[ G = c = 1, \quad (10.30) \]
which means that we measure the mass, as the lengths, in cm. We shall often adopt this convention, and we will indicate the geometrical mass (i.e. the mass in cm) as \( m \).

In these unities the Schwarzschild solution becomes
\[ ds^2 = - \left( 1 - \frac{2m}{r} \right) dt^2 + \frac{1}{1-\frac{2m}{r}} dr^2 + r^2 \left( d\theta^2 + \sin^2 \theta d\phi^2 \right). \quad (10.31) \]

10.4 The singularities of Schwarzschild solution

Let us examine the metric (10.31) in some more detail. We immediately see that there is a problem when \( r \to 2m \): \( g_{00} \to 0 \), and \( g_{rr} \to \infty \). Moreover, when \( r \to 0 \), \( g_{00} \to \infty \), and \( g_{rr} \to 0 \). In both cases we say that there is a singularity, but of a different nature. In order to check whether a singularity is a genuine curvature singularity, we should compute the scalars which we can construct from the Riemann tensor and see if they diverge. To check whether the Riemann tensor is well-behaved is not enough, in fact for the Schwarzschild metric the components of \( R^{\alpha\beta\gamma\delta} \) are
\[ R^t_{trr} = -2 \frac{m}{r^3} \left( 1 - \frac{2m}{r} \right)^{-1} \quad (10.32) \]
\[ R^t_{\theta\theta} = \frac{1}{\sin^2 \theta} R^t_{\varphi\varphi} = \frac{m}{r^5} \]
\[ R^\theta_{\varphi\phi} = 2 \frac{m}{r^5} \sin^2 \theta \]
\[ R^r_{\theta\theta} = \frac{1}{\sin^2 \theta} R^r_{\varphi\varphi} = -\frac{m}{r^5} \]
and they diverge both at \( r = 0 \), and at \( r = 2m \). However, if we compute the scalar invariants, like \( R^{abcd}R_{abcd} \), we find that they diverge only at \( r = 0 \). We conclude that \( r = 0 \) is a true curvature singularity, while \( r = 2m \) is only a coordinate singularity, due to an unappropriate choice of the coordinates.

We shall now analyse the properties of the surface \( r = 2m \).

### 10.5 Spacelike, Timelike and Null Surfaces

In a curved background hypersurfaces are classified in the following way. Consider a generic hypersurface \( \Sigma \)

\[
\Sigma(x^\alpha) = 0,
\]

(10.33)

Be \( \vec{n} \) the normal vector dual to the gradient one-form

\[
n_\alpha = \Sigma_{,\alpha}
\]

(10.34)

If \( t^\alpha \) is a tangent vector to the surface, then \( t^\alpha n_\alpha = 0 \). Indeed, \( t^\alpha = \frac{dx^\alpha}{d\lambda} \) with \( x^\alpha(\lambda) \) curve on \( \Sigma \); therefore,

\[
t^\alpha n_\alpha = \frac{dx^\alpha}{d\lambda} \frac{\partial \Sigma}{\partial x^\alpha} = \frac{d\Sigma}{d\lambda} = 0.
\]

(10.35)

At any point of the hypersurface we can introduce a locally inertial frame, and rotate it in such a way that the components of \( \vec{n} \) are

\[
n^\alpha = (n^0, n^1, 0, 0) \quad \text{and} \quad n_\alpha n^\alpha = (n^1)^2 - (n^0)^2.
\]

(10.36)

Consider a vector \( t^\alpha \) tangent to \( \Sigma \) at the same point. \( t^\alpha \) must be orthogonal to \( \vec{n} \)

\[
n_\alpha t^\alpha = -n^0 t^0 + n^1 t^1 = 0 \quad \rightarrow \quad \frac{t^0}{t^1} = \frac{n^1}{n^0}.
\]

(10.37)

From eq. (10.37) it follows that

\[
t^\alpha = \Lambda(n^1, n^0, a, b) \quad \text{with a, b e } \Lambda \quad \text{costant and arbitrary.}
\]

(10.38)

Consequently the norm of \( \vec{t} \) is

\[
t_\alpha t^\alpha = \Lambda^2[-(n^1)^2 + (n^0)^2 + (a^2 + b^2)] = \Lambda^2[-n_\alpha n^\alpha + (a^2 + b^2)].
\]

(10.39)

There are three possibilities:

1) \( n_\alpha n^\alpha < 0 \), \( \rightarrow \) \( n^\alpha \) is a timelike vector \( \rightarrow \) \( \Sigma \) is spacelike

2) \( n_\alpha n^\alpha > 0 \), \( \rightarrow \) \( n^\alpha \) is a spacelike vector \( \rightarrow \) \( \Sigma \) is timelike

3) \( n_\alpha n^\alpha = 0 \), \( \rightarrow \) \( n^\alpha \) is a null vector \( \rightarrow \) \( \Sigma \) is null

We shall now see how the normal and the tangent vectors are directed in order to understand the disposition of the light-cones with respect to the hypersurface.

1) If \( n_\alpha n^\alpha < 0 \), then \( t_\alpha t^\alpha > 0 \) and \( \vec{t} \) is spacelike. Consequently no tangent vector to \( \Sigma \) in \( P \) lies inside, or on the light-cone through \( P \). Since a massive particle which starts
at P must move inside the light-cone (or on the light-cone if it is massless), this means that a spacelike hypersurface can be crossed only in one direction.

2) If $n_\alpha n^\alpha > 0$, then $t_\alpha t^\alpha$ can be positive, negative or null depending on the value of $a^2 + b^2$. Therefore there will be some tangent vectors which lie inside the light-cone. Consequently a **timelike hypersurface can be crossed inward and outward.**

3) If $n_\alpha n^\alpha = 0$, then $t_\alpha t^\alpha$ is positive ($t^\alpha$ is spacelike), or null if $a = b = 0$. In this case there will be only one tangent vector (and all its multiples) at P which lies on $\Sigma$ and on the light-cone.
For example, in Minkowski spacetime \( t = \text{const} \) is a spacelike surface, and any physical object can pass it in only one direction without violating causality. \( x = \text{const} \) is a timelike surface, and physical objects can pass it in either direction, \( x - ct = 0 \) is a null surface. Let us now try to understand what kind of surface \( r = 2m \) is.

Consider a generic hypersurface \( r = \text{const} \) in the Schwarzschild geometry

\[
\Sigma = r - cost = 0. \tag{10.40}
\]

The norm of the normal vector is

\[
n_\alpha n^\alpha = g^{\alpha\beta} n_\alpha n_\beta = g^{\alpha\beta} \Sigma_\alpha \Sigma_\beta =
\]

\[
g^{00} \Sigma^2 + g^{rr} \Sigma^2 + g^{\theta\theta} \Sigma^2 + g^{\varphi\varphi} \Sigma^2 = g^{rr} \Sigma^2 = \left(1 - \frac{2m}{r}\right). \tag{10.41}
\]

From eq. (10.41) it follows that

\[
\begin{align*}
  r &> 2m \quad \rightarrow \quad n_\alpha n^\alpha > 0, \quad \Sigma \text{ is timelike} \\
  r &= 2m \quad \rightarrow \quad n_\alpha n^\alpha = 0, \quad \Sigma \text{ is null} \\
  r &< 2m \quad \rightarrow \quad n_\alpha n^\alpha < 0, \quad \Sigma \text{ is spacelike}
\end{align*}
\]

Consider for example \( S_1 \) and \( S_2 \) as shown in the following figure.
Any signal which starts at some point of $S_1$ can be sent both toward the origin and outward, since $S_1$ is timelike. Conversely, a signal which starts at a point of $S_2$ in the interior of $r = 2m$ must necessarily go inward, and be captured by the singularity at $r = 0$, since $S_2$ is spacelike. The surface $r = 2m$ is a null surface, which is basically the transition from a spacelike to a timelike hypersurface. On the surface $r = 2m$ the timelike Killing vector $\xi(t)$ becomes null and it is spacelike for $r < 2m$.

The Schwarzschild solution is said to represent the gravitational field of a black hole, and the hypersurface $r = 2m$ is called the event horizon. The reason for these names is that if we are outside $r = 2m$ we can send a signal both inward and outward, but as soon as we cross the horizon any signal will inevitably bend toward the singularity: there is no way to know what happens inside the horizon.

As we mentioned before, $r = 0$ is a genuine curvature singularity. Thus General Relativity predicts the existence of singularities hidden by a horizon.
10.6 How to remove a coordinate singularity

In general, it is not a simple problem to understand whether a singularity is a genuine curvature singularity or it is only a coordinate singularity. The first thing to do is to compute the Riemann tensor and the scalars which can be computed from it, like $R_{abcd}R^{abcd}$ and check whether they diverge somewhere. If this is not the case, the singularity is due to a bad choice of the coordinate system, and a suitable choice of a new set of coordinates should remove it. If this can be done, we say that we are extending our original spacetime $(M, g_{\alpha\beta})$ to a larger spacetime $(\tilde{M}, \tilde{g}_{\alpha\beta})$ which includes the original one. Before analyzing the Schwarzschild case, let us consider two examples.

Consider the two-dimensional spacetime
\[ ds^2 = -\frac{1}{t^4}dt^2 + dx^2, \quad 0 < t < \infty, \quad -\infty < x < \infty. \] (10.42)

$(c = G = 1.)$ The metric is singular at $t = 0$. The coordinate transformation
\[ t' = \frac{1}{t}, \quad dt' = -\frac{1}{t^2}dt, \] (10.43)
gives
\[ ds^2 = -(dt')^2 + dx^2, \] (10.44)
Thus the metric (10.42) represents a flat spacetime. The metric (10.44) is defined for any $t'$, i.e. $-\infty < t' < \infty$, therefore it describes regions of the spacetime which where “unaccessible” to the coordinates (10.42). In fact in that case $0 < t < \infty$, which corresponds only to the section $0 < t' < \infty$, of our new spacetime. This is the reason why we say that the new coordinates provide an extension of the spacetime. The coordinate singularity $t = 0$ is mapped onto the line $t' \to \infty$. The new spacetime is said to be geodesically complete because any geodesic which starts at any given point of the spacetime, can be extended for arbitrarily large values of the affine parameter. Conversely, the original spacetime (10.42) was geodesically incomplete for the following reason. We have established that the spacetime is flat, and it extends from $-\infty$ to $+\infty$ in both coordinates $(t', x')$. In eq. (10.42) we were trying to cover our infinite spacetime with coordinates which vary in a semi-infinite range $(0 < t < \infty)$. This is the reason why the singularity $t = 0$ appears. With those coordinates we were able to cover only the region $(0 < t' < \infty)$ of the complete spacetime, but not the region $(-\infty < t' < 0)$. Consequently, geodesics which start in the region $t' < 0$, cross the axis and continue in the region $t' > 0$, cannot be completely represented in the spacetime described by $(t, x)$: they will terminate for a finite value of the proper time.

Another example is the Rindler spacetime, which has interesting similarities with the Schwarzschild geometry. The line-element is
\[ ds^2 = -x^2dt^2 + dx^2, \quad -\infty < t < \infty, \quad 0 < x < \infty. \] (10.45)
The metric is singular at $x = 0$. The determinant $g$ vanishes at $x = 0$, therefore $g^{\mu\nu}$ is also singular. Let us consider geodesics in this spacetime. Since the metric is independent on time, it admits a timelike Killing vector $\xi(t) \to (1, 0)$. According to eq. (9.37)
\[ \xi(t)\alpha U^\alpha = g_{\alpha\beta}\xi(t)\beta U^\beta = const = -E, \quad \rightarrow \quad U^0 = \frac{E}{x^2}, \] (10.46)
where $U^\beta = \frac{dx^\beta}{d\lambda}$ and $\lambda$ is an affine parameter (not necessarily the proper time). Therefore,

$$\frac{dt}{d\lambda} = \frac{E}{x^2}. \quad (10.47)$$

Since the norm of the vector tangent to the worldline of a massive particle is $-1$, then

$$U^\mu U^\nu g_{\mu\nu} = -x^2 \left( \frac{dt}{d\lambda} \right)^2 + \left( \frac{dx}{d\lambda} \right)^2 = -1, \quad (10.48)$$

thus

$$\left( \frac{dx}{d\lambda} \right)^2 = x^2 \left( \frac{dt}{d\lambda} \right)^2 - 1 = \frac{E^2}{x^2} - 1. \quad (10.49)$$

Hence

$$\frac{dx}{d\lambda} = \pm \sqrt{\frac{E^2}{x^2} - 1}, \quad \rightarrow \quad \lambda = \int^x \frac{xdx}{\sqrt{E^2 - x^2}} = -\sqrt{E^2 - x^2} + \text{const.} \quad (10.50)$$

Thus a particle starting at some point $x$ reaches $x = 0$ in a finite interval of the affine parameter: Rindler spacetime is **geodesically incomplete**. However, since the Riemann tensor and the curvature scalars do not diverge at $x = 0$, there must exist a coordinate transformation which brings the metric into a non-singular form. Unfortunately a systematic approach to the problem of finding the “right” coordinates to extend the metric does not exist. We shall describe a procedure which is based on the behaviour of null geodesics. In two dimension the situation is easier, since null geodesics belong, at least locally, to two classes: ingoing and outgoing. Two geodesics belonging to the same class cannot cross, because the two tangent vectors should coincide at that point, and consequently the two geodesic should coincide everywhere (remember that geodesics parallel-transport their own tangent vector).

If

$$\vec{K} \rightarrow \{ \frac{dx^\mu}{d\lambda} \}, \quad (10.51)$$

is the vector tangent to the null geodesic whose affine parameter is $\lambda$, we must have that

$$g_{\mu\nu}K^\mu K^\nu = 0. \quad (10.52)$$

In the case we are considering it becomes

$$0 = g_{\mu\nu}K^\mu K^\nu = -x^2 \left( \frac{dt}{d\lambda} \right)^2 + \left( \frac{dx}{d\lambda} \right)^2, \quad (10.53)$$

from which we find

$$\left( \frac{dt}{dx} \right)^2 = \frac{1}{x^2}. \quad (10.54)$$

Therefore along the null geodesic

$$t = \pm \log x + \text{const}, \quad (10.55)$$
where the + identifies the outgoing geodesics and the - the ingoing geodesics. Accordingly, we define the null ingoing and outgoing coordinates as

\[ u = t - \log x \quad \text{and} \quad v = t + \log x \]  

and the metric in the new coordinates becomes

\[ ds^2 = -e^{v-u} du dv. \]  

The coordinates \( u \) and \( v \) vary in the range \((-\infty, +\infty)\), and they cover the original region \( x > 0 \), (they do not extend the spacetime yet!), thus we haven’t solved the problem of eliminating the singularity. An extension of the spacetime can be accomplished if we reparametrize the null geodesics with new coordinates

\[ U = U(u) \] 
\[ V = V(v). \]  

The form of the metric is so simple that we may define \( U \) and \( V \) immediately. But to have a feeling on what one should do in general we proceed in a more systematic way. From eqs. (10.46) and (10.51) it follows that for a massless particle

\[ \xi_{(i)} \gamma^{\alpha} = g_{\alpha \beta} \xi^{\alpha}_i \gamma^\beta = \text{const} = -E, \quad \rightarrow \quad d\lambda = \frac{x^2}{E} dt. \]  

Since \( dt = \frac{1}{2} d(u + v) \), if we put \( u = \text{const} \) and move along a null direction parallel to the \( v \)-axis, i.e. along an outgoing null geodesic, eq. (10.59) becomes

\[ d\lambda = \frac{x^2}{2E} dv, \quad \text{or, since} \quad 2 \log x = v - u \rightarrow x = e^{\frac{v-u}{2}}, \] 
\[ \lambda = \frac{1}{2E} \int e^{(v-u)} dv = C + \frac{e^{-u}}{2E} e^v, \]  

where \( C \) is a constant. If we shift \( \lambda \rightarrow \frac{\lambda - C}{2E} \), then the affine parameter along outgoing null geodesics becomes

\[ \lambda_{out} = e^v. \]  

Proceeding in a similar way we find that the affine parameter along ingoing null geodesics is

\[ \lambda_{in} = -e^{-u}. \]  

If we now choose

\[ U = -e^{-u} \] 
\[ V = e^v, \]  

the metric becomes

\[ ds^2 = -dU dV, \quad \text{or if we put} \quad T = \frac{(U + V)}{2}, \quad X = \frac{(V - U)}{2} \] 
\[ ds^2 = -dT^2 + dX^2, \]  

where
which is again a flat spacetime.

Summarying: 1) we find the equations for the ingoing and outgoing null geodesics, 2) we choose the affine parameters along these geodesics as coordinates, then we introduce $X$ and $T$. $U$ and $V$ range between $\infty$ and $+\infty$. The original spacetime $(x, t)$ coincides with the quadrant $[U < 0, V > 0]$, but since everything is regular at $[U = 0, V = 0]$, the metric is extended to the regions $U > 0$, and $V < 0$, which were not included before. The relation between the old and the new coordinates is

$$x = (X^2 - T^2)^{\frac{1}{2}}$$
$$t = \tanh^{-1}\left(\frac{T}{X}\right) = \frac{1}{2} \log \left(\frac{X + T}{X - T}\right)$$

A picture of the spacetime is given in the following figure:

The singularity $x = 0$ corresponds to the lines $X = \pm T$, where the metric in the new coordinates is perfectly well behaved. From the second of eqs. (10.65)

$$X = -T \quad \text{corresponds to} \quad t \to -\infty$$
$$X = T \quad \text{corresponds to} \quad t \to +\infty.$$  

The curves $x = \text{const}$ are now mapped onto the hyperbolae $X^2 - T^2 = \text{const}$, while the curves $t = \text{const}$ are mapped onto $T = \text{const}X$. The original Rindler space corresponds to the dashed region in the figure. Therefore we have finally extended the spacetime across the barrier $x = 0$.

If we now go back to Rindler’s metric and consider the following coordinate transformation

$$y = \frac{1}{4}x^2,$$
the metric becomes
\[ ds^2 = -4y dt^2 + \frac{1}{y} dy^2, \tag{10.68} \]
and rescaling the time coordinate \( t \to 2t \)
\[ ds^2 = -y dt^2 + \frac{1}{y} dy^2. \tag{10.69} \]
This is similar to the form of the two-dimensional \((t, r)\) part of the Schwarzschild metric.

### 10.7 The Kruskal extension

First we compute the null geodesics of the two-dimensional Schwarzschild metric
\[ ds^2 = \left(1 - \frac{2m}{r}\right) dt^2 + \frac{1}{1 - \frac{2m}{r}} dr^2 \tag{10.70} \]
by imposing
\[ 0 = g_{\mu\nu} K^\mu K^\nu = - \left(1 - \frac{2m}{r}\right) \left(\frac{dt}{d\lambda}\right)^2 + \left(1 - \frac{2m}{r}\right)^{-1} \left(\frac{dr}{d\lambda}\right)^2. \tag{10.71} \]
Hence
\[ \left(\frac{dr}{dt}\right)^2 = \left(1 - \frac{2m}{r}\right)^2 \quad \rightarrow \quad \frac{dt}{dr} = \pm \frac{r}{r - 2m}, \tag{10.72} \]
whose solution is
\[ t = \pm r_* + \text{const} \tag{10.73} \]
where
\[ r_* = r + 2m \log \left(\frac{r}{2m} - 1\right), \quad \text{and} \quad \frac{dr}{dr_*} = \left(1 - \frac{2m}{r}\right). \tag{10.74} \]
The coordinate \( r_* \) is called the “tortoise” coordinate, since if \( r \to +\infty \) then \( r_* \sim r \), but if \( r \to 2m \) then \( r \to -\infty \), thus as \( r \to 2m \) \( r_* \) pushes the horizon to \( -\infty \). We now define the null ingoing and outgoing coordinates
\[ u = t - r_* \quad -\infty < u < +\infty, \tag{10.75} \]
\[ v = t + r_* \quad \rightarrow \quad r_* = \frac{v - u}{2} \quad -\infty < v < +\infty \]
and the two-dimensional metric becomes
\[ ds^2 = \left(1 - \frac{2m}{r}\right) \left[ dt^2 - \frac{dr^2}{\left(1 - \frac{2m}{r}\right)^2} \right] = \left(1 - \frac{2m}{r}\right) dudv. \tag{10.76} \]
hence

\[ ds^2 = - \left( 1 - \frac{2m}{r} \right) du dv = \frac{2m}{r} e^{-\frac{r}{2m}} dudv. \] (10.80)

Since \( r = 2m \) corresponds to \( u \to \infty \) and \( v \to -\infty \), the metric (10.80) is regular everywhere. A comparison with the Rindler case shows that a convenient choice for \( U \) and \( V \) is

\[ U = -e^{-\frac{r}{2m}}, \quad \to \quad -\infty < U < 0 \] (10.81)
\[ V = e^{\frac{r}{2m}}, \quad \to \quad 0 < V < +\infty \]

The metric becomes

\[ ds^2 = -\frac{32m^3 e^{-\frac{r}{2m}}}{r} dUdV. \] (10.82)

The surface \( r = 2m \) now corresponds to \( U = 0 \) or \( V = 0 \) where the metric (10.82) is non-singular. Therefore it can be extended across these two hypersurfaces to cover the whole two-dimensional spacetime. By introducing the coordinates \( T \) and \( X \)

\[ T = \frac{V + U}{2}, \quad X = \frac{V - U}{2} \] (10.83)

the four-dimensional metric finally becomes

\[ ds^2 = \frac{32m^3 e^{-\frac{r}{2m}}}{r} \left[ -dT^2 + dX^2 \right] + r^2 \left( d\theta^2 + \sin^2 \theta d\varphi^2 \right). \] (10.84)

This extension was independently found by Kruskal and Szekeres in 1960. The relation between the old and the new coordinates is \(^2\)

\(^1\)From the definition of \( r_* \) we find

\[ \frac{r_* - r}{2m} = \log \left( \frac{r - 2m}{2m} \right) \quad \to \quad e^{\frac{r}{2m}} e^{-\frac{r}{2m}} = \frac{r - 2m}{2m}. \] (10.77)
\[ (1 - \frac{2m}{r}) = \frac{r - 2m}{2m} = \frac{2m}{r} e^{-\frac{r}{2m}} e^{\frac{r}{2m}}. \] (10.78)

Since \( r_* = (v - u)/2 \), it follows

\[ (1 - \frac{2m}{r}) = \frac{2m}{r} e^{-\frac{r}{2m}} e^{\frac{r}{2m}} \] (10.79)

\(^2\)The derivation of eqs. (10.85) and (10.86):

\[ (X^2 - T^2) = \left( \frac{V - U}{2} \right)^2 - \left( \frac{V + U}{2} \right)^2 = -UV = + e^{-\frac{r}{2m}} = \frac{r}{2m} \left( 1 - \frac{2m}{r} \right) e^{\frac{r}{2m}}, \]

and

\[ \log \left( \frac{X + T}{X - T} \right) = \log \left( \frac{V}{U} \right) = \log e^{\frac{r}{2m}} = \frac{v + u}{4m} = \frac{t}{2m}. \]
\[ (X^2 - T^2) = \left( \frac{r}{2m} - 1 \right) e^{\frac{r}{2m}} \]  
(10.85)

\[ \frac{t}{2m} = \log \left( \frac{X + T}{X - T} \right) = 2\tanh^{-1} \left( \frac{T}{X} \right) \]  
(10.86)

The extended two-dimensional spacetime is shown in the following figure.

If \( r = \text{const} > 2m \), from eq. (10.85) it follows that \( X^2 - T^2 > 0 \) and constant, and consequently \( X = \pm \sqrt{T^2 + k} \), where \( k = \left[ \left( 1 - \frac{2m}{r} \right) e^{\frac{r}{2m}} \right]_{r=\text{const}} \). These curves are indicated as continuous lines in the quadrants I and IV of the preceding figure.

If \( r = \text{const} < 2m \), \( X^2 - T^2 < 0 \) and constant, and \( X = \pm \sqrt{T^2 - |k|} \). These curves are the dashed lines in the quadrants II and III. The curvature singularity \( r = 0 \)
corresponds to the curves \( X^2 - T^2 = -1 \) and \( X = \pm \sqrt{T^2 - 1} \) also represented in regions II and III. Radial null geodesics correspond to \( U = const \) (ingoing) or to \( V = const \) (outgoing). This diagram has the remarkable property that null geodesics are \( 45^\circ \)-straight lines. The curves \( t = const \) are straight lines passing through the origin.

The original spacetime (\( r > 0 \)), i.e. the Schwarzschild spacetime in the exterior of the horizon, corresponds to the quadrant \([-\infty < U < 0, 0 < V < \infty]\), labeled as region I. What is the meaning of the other regions? Consider a physical observer which starts at some point \( r \) in the exterior of the horizon, i.e. in region I, as indicated in the next figure.

He can move only in the interior of the light-cones, which, at every point are \( 45^\circ \)-straight lines. As one can see from the figure, as long as the observer is outside the horizon, he can still invert its direction of motion and escape free at infinity. But as soon as he crosses the surface \( U = 0 \) and enters in region II, this is no longer possible, and he gets captured by the singularity \( r = 0 \), (compare with the discussion on the nature of the hypersurfaces \( r = const \) in section 10.5. The singularity \( r = 0 \) is a spacelike singularity). Thus region II represents the spacetime in the interior of the horizon. Regions III and IV have the same characteristics as regions I and II, but they are time-reversed with respect to them: a particle in region III must necessarily have been emitted by the singularity sitting in that region. Then it will cross the surface \( r = 2m \) (\( U = 0 \) or \( V = 0 \)) and will escape free at infinity either in region I, or in its mirror image region IV. It should be noted that region I and IV are causally unrelated, since a signal emitted by an observer in region I will never reach region IV and vice versa. It is interesting to ask whether regions IV and III do exist.
or not. Suppose that a black hole has formed, and we really have a singularity concealed by
a horizon. We live in the exterior of the horizon (we can move inward and outward). We
can send signals to region II, but no signal emitted by us will reach regions III and IV for
the reasons explained above. On the other hand, no signal coming from region IV can reach
us. A signal emitted in region III (the white hole region) might, in principle, reach region
I. However it is reasonable to assume that the black hole has formed at some time as the
result of some physical process (the collapse of a massive star, as we shall soon see), and
since any signal emitted in region III would take an infinite time $t$ to reach region I, region
III cannot communicate with us. If we want to take a pragmatical point of view, we can
conclude that since we cannot communicate with regions III and IV (and viceversa), they
do not exist for us. To speculate on the existence of ‘other universes’, although intriguing,
is outside the scope of this course.

The Kruskal extension is very useful to investigate the causal structure of the spacetime
in the vicinity of the horizon. However it is unappropriate to describe the spacetime at
infinity, due to the exponential behaviour of $g_{TT}$ and $g_{XX}$. 
Chapter 11

Experimental Tests of General Relativity

11.1 Gravitational redshift of spectral lines

Time intervals are measured using clocks, which are instruments whose functioning is based on the repetition of a periodic phenomenon, such as atomic oscillations or the oscillations of a quartz crystal. We choose as time unit the interval of proper time between two successive repetitions of the periodic phenomenon. The definition of proper time in general relativity is

\[ dT = \frac{1}{c} \sqrt{-ds^2} = \frac{1}{c} \sqrt{-g_{\mu\nu}(x^\mu)dx^\mu dx^\nu}. \]  

(11.1)

In this expression \( g_{\mu\nu} \) must be evaluated at the (spacetime) position of the body to which it refers; in the case under consideration it has to be evaluated at the clock position. Thus, if the clock is at rest with respect to the reference frame, \( dx^i = 0, i = 1, 3 \) and the proper time interval between two ticks is

\[ dT = \frac{1}{c} \sqrt{-g_{00}(x^0)dx^0} = \sqrt{-g_{00}(x^0)dt}, \]  

(11.2)

were \( dt \) is the interval of coordinate time between two ticks. Note that we are assuming that \( dT \) is very small, so that we can use the infinitesimal expression of the proper time without integrating over the clock worldline.

By dividing the proper time interval by \( dt \) we find

\[ \frac{dT}{dt} = \frac{1}{c} \sqrt{-g_{\mu\nu}(x^\mu) \frac{dx^\mu}{dt} \frac{dx^\nu}{dt}}. \]  

(11.3)

\( \frac{dT}{dt} \) is called time dilation factor; it gives the ratio between the interval of proper time between two events and the corresponding interval of coordinate time, and depends both on the metric and on the clock velocity. If the clock is at rest with respect to the reference frame it becomes

\[ \frac{dT}{dt} = \sqrt{-g_{00}(x^0)}. \]  

(11.4)
We shall now show that, in a gravitational field, the frequency of a signal detected at a point different from the emission point is different from the emission frequency. Let us assume that the gravitational field is stationary, which implies that there exists a timelike Killing vector and that, by a suitable coordinates choice, the metric can be made independent of time. In this case, the coordinate $x^0$ will be referred to as the universal time. This choice is not univocal, because we can always shift the origin of time, and rescale $x^0$ by an arbitrary constant. Be $S$ a light source and $O$ an observer, located at two different points.

The source $S$ emits a wave crest which reaches $O$ after an interval of coordinate time $\Delta x^0$. Since for a light signal $ds^2 = 0$, we can compute $\Delta x^0$ by solving this equation with respect to $dx^0$, and by integrating over the light path as follows:

$$ds^2 = g_{00}(dx^0)^2 + 2g_{0i}dx^0dx^i + g_{ik}dx^idx^k = 0, \quad i, k = 1, \ldots, 3$$

$$\Delta x^0 = \int_{\text{light path}} dx^0 = \int_{\text{light path}} \frac{-g_{0i}dx^i \pm \sqrt{(g_{0i}dx^i)^2 - g_{00}g_{ik}dx^idx^k}}{g_{00}}.$$ 

The physical solution is that corresponding to the $-$ sign.  

1 Why do we have two solutions for $\Delta x^0$ corresponding to the $\pm$ sign? Firstly note that since $g_{00}$ is negative and $g_{ik}$ are positive $\sqrt{(g_{0i}dx^i)^2 - g_{00}g_{ik}dx^idx^k} > g_{0i}dx^i$; consequently the solution with the $+$ sign is negative and that with the $-$ is positive, i.e. $(\Delta x^0)_+ < 0$ and $(\Delta x^0)_- > 0$. Clearly the physical solution is $(\Delta x^0)_- > 0$, whereas $(\Delta x^0)_+ < 0$ would correspond to a signal that being emitted by $O$ would reach $S$ at $x^0 = 0$. The situation is not symmetric; indeed if we imagine that $S$ is near a massive body and $O$ is far away, in one case the signal would travel from $S$ to $O$ “against” the gravitational force, in the other case it would travel inward, from $O$ to $S$, favoured by the gravitational attraction.
Similarly, the period measured by the observer is the interval of its own proper time, which elapses between the detection of two wave crests, i.e.
\[ \Delta T_{\text{obs}} = \sqrt{-g_{00}(x^\mu_{\text{obs}})} \Delta t_{\text{obs}}, \]
and the observed frequency is
\[ \nu_{\text{obs}} = \frac{1}{\Delta T_{\text{obs}}} = \frac{1}{\sqrt{-g_{00}(x^\mu_{\text{obs}})} \Delta t_{\text{obs}}}. \]

Using the fact that \( \Delta t_{\text{em}} = \Delta t_{\text{obs}} \) we finally find
\[ \frac{\nu_{\text{obs}}}{\nu_{\text{em}}} = \frac{\lambda_{\text{em}}}{\lambda_{\text{obs}}} = \sqrt{\frac{g_{00}(x^\mu_{\text{em}})}{g_{00}(x^\mu_{\text{obs}})}}. \]  
\[ (11.5) \]

Thus, in general the frequency of a signal emitted in a gravitational field at a given point, is different from that detected at a different point, since the metric in the two points is different.

### 11.1.1 Some useful numbers

At this point it is useful to recall some numbers which allow us to establish when a gravitational field can be considered weak. Let us consider the Sun first. Its mass and radius are
\[ M_\odot = 1.989 \cdot 10^{33} \text{ g}, \quad R_\odot = 6.9599 \cdot 10^5 \text{ km}; \]  
\[ (11.6) \]
moreover
\[ \frac{GM_\odot}{c^2} = \frac{1.989 \cdot 10^{33} \times 6.673 \cdot 10^{-8}}{(2.998 \cdot 10^{10})^2} \sim 1.4768 \text{ km}, \quad \text{and} \quad \frac{GM_\odot}{R_\odot c^2} \sim 0.21 \cdot 10^{-5}. \]  
\[ (11.7) \]
The quantity \( \frac{GM}{Rc^2} \) is said surface gravity, and it is a measure of how strong are the effects of general relativity. The surface gravity of the Sun is much smaller than unity, therefore we can say that its gravitational field is weak.

The Earth has mass \( M_\oplus = 5.98 \cdot 10^{27} \text{ g} \) and equatorial radius \( R_\oplus = 6.378 \cdot 10^3 \text{ km} \). Since
\[ M_\odot/M_\oplus \simeq 3 \cdot 10^5, \quad \text{and} \quad R_\odot/R_\oplus \simeq 10^2, \]
\[ (11.8) \]
i.e. the surface gravity of the Sun is about 3000 times larger than that of the Earth.

Conversly, if we consider a neutron star with typical mass and radius
\[ M_{\text{NS}} \sim 1.4 \, M_\odot, \quad R_{\text{NS}} \sim 10 \text{ km}, \]  
\[ (11.9) \]
the surface gravity is
\[ \frac{GM_{\text{NS}}}{R_{\text{NS c}^2}} \sim 0.21, \]  
\[ (11.10) \]
which is close to unity and much larger than that of the Sun. Thus, the effects of general relativity will be much more important for a neutron star than for the Sun.
11.1.2 Redshift of spectral lines in the weak field limit

Let us now consider eq. (11.5) in the weak field limit. In section 8.1 we have seen that if we assume that the gravitational field is stationary and weak, the geodesic equations show that the 00-component of the metric tensor is related to the Newtonian potential $\Phi$, solution of the equation $\nabla^2 \Phi = 4\pi G\rho$, by the equation

$$g_{00} \simeq - \left(1 + \frac{2\Phi}{c^2}\right).$$

Consequently, if the gravitational field is weak and stationary eq. (11.5) becomes

$$\frac{\nu_{\text{obs}} - \nu_{\text{em}}}{\nu_{\text{em}}} = \frac{\lambda_{\text{em}} - \lambda_{\text{obs}}}{\lambda_{\text{obs}}} = \sqrt{1 + \frac{2\Phi_{\text{em}}}{c^2}} \sqrt{1 + \frac{2\Phi_{\text{obs}}}{c^2}} - 1 \simeq \sqrt{1 + \frac{2}{c^2} (\Phi_{\text{em}} - \Phi_{\text{obs}})} - 1 \simeq \sqrt{1 + \frac{2}{c^2} (\Phi_{\text{em}} - \Phi_{\text{obs}})} - 1 \simeq \frac{1}{c^2} (\Phi_{\text{em}} - \Phi_{\text{obs}})$$

and finally

$$\frac{\Delta \nu}{\nu} \simeq \frac{1}{c^2} (\Phi_{\text{em}} - \Phi_{\text{obs}}).$$ \hspace{1cm} (11.11)

Let us suppose that the source of light is on the Sun, whose gravitational field is weak, and that the observer is on the Earth. We shall neglect the gravitational field of the Earth since it is much smaller than that of the Sun. In this case, $\Phi = -\frac{GM}{r}$, where $r$ is the distance from the Sun center, $r_{\text{em}} = R_{\odot}$ and $r_{\text{obs}} = r_{\text{Sun-Earth}}$. Thus eq. (11.11) becomes

$$\frac{\Delta \nu}{\nu} \simeq \frac{GM_{\odot}}{c^2} \left( -\frac{1}{R_{\odot}} + \frac{1}{r_{\text{Sun-Earth}}} \right);$$

Since the average distance between the Sun and the Earth is $r_{\text{Sun-Earth}} = 149.6 \cdot 10^6$ km, which is about 210 times the Sun radius, we can assume $r_{\text{Sun-Earth}} \gg R_{\odot}$, so that

$$\frac{\Delta \nu}{\nu} \simeq -\frac{GM_{\odot}}{R_{\odot} c^2} \simeq -0.21 \cdot 10^{-5}. \hspace{1cm} (11.12)$$

Note the following:

- $\Delta \nu < 0$, i.e. the observed spectral lines are shifted toward lower frequencies, i.e. the light reddened.
- The redshift of spectral lines produced by the Sun is of the order of its surface gravity.
11.1.3 Redshift of spectral lines in a strong gravitational field

Let us now consider the case when the source emitting light and the observer are located in the gravitational field of a neutron star or of a black hole. The metric appropriate to describe the exterior of a neutron star, from $r = R_{NS}$ up to radial infinity, and a black hole is the Schwarzschild metric

$$ds^2 = -(1 - \frac{2m}{r})dt^2 + \frac{1}{1 - \frac{2m}{r}}dr^2 + r^2(d\theta^2 + \sin^2\theta d\varphi^2),$$

where $m = GM/c^2$ is the mass, either of the star or of the black hole, in geometric units. If we assume that the observer is located very far from the source emitting light, i.e. $r_{obs} \gg r_{em}$, eq. (11.5) gives

$$\frac{\nu_{obs}}{\nu_{em}} = \sqrt{\frac{-g_{00}(x^\mu_{em})}{-g_{00}(x^\mu_{obs})}} = \sqrt{1 - \frac{2m}{r_{em}}} \sim \sqrt{1 - \frac{2m}{r_{obs}}} \sim \sqrt{1 - \frac{2m}{r_{em}}}.$$ (11.13)

If the light source is located on a neutron star surface, i.e. $r_{em} = R_{NS}$, this equation gives

$$\frac{\nu_{obs}}{\nu_{em}} \sim \sqrt{1 - \frac{2GM_{NS}}{R_{NS}c^2}} \sim \sqrt{1 - 2 \times 0.21} \sim 0.76 \rightarrow \frac{\Delta\nu}{\nu} = \frac{\nu_{obs} - \nu_{em}}{\nu_{em}} \sim -0.24,$$

where we have used eq. (11.10). This means that an observer located at infinity with respect to the neutron star will see the emitted light reddened ($\Delta\nu < 0$) by quite a large amount, much larger than that produced by the Sun which we computed in eq. (11.12).

Let us now suppose that the source of the gravitational field is a black hole, and that the source emitting light is on a spacecraft orbiting around it. From eq. (11.13) we see that as the light source approaches the horizon $r = 2m$,

$$\nu_{obs} \sim \sqrt{1 - \frac{2m}{r_{em}}} \nu_{em} \rightarrow 0,$$

i.e. the observed signal will fade away since the observed frequency tends to zero. Thus, the signal emitted by a source falling into a black hole has a distinctive feature, i.e. its frequency will progressively decrease tending to zero near the horizon.

NOTE THAT: to derive the gravitational redshift, we have used only the fact that the effects of the gravitational field are described by the metric tensor, i.e. we have used basically only the Equivalence Principle.
11.2 The geodesic equations in the Schwarzschild background

The geodesic equations can be derived not only from the Equivalence Principle as shown in previous chapters, but also from a variational principle, as we shall now show.

11.2.1 A variational principle for geodesic motion

Let us define the Lagrangian of a free particle as

\[ L \left( x^\alpha, \frac{dx^\alpha}{d\lambda} \right) = \frac{1}{2} g_{\mu\nu}(x^\alpha) \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = \frac{1}{2} g_{\mu\nu}(x^\alpha) \dot{x}^\mu \dot{x}^\nu, \quad (11.14) \]

in the space of the curves \( \{ x^\mu(\lambda), \lambda \in [\lambda_0, \lambda_1] \} \), and the action

\[ S = \int L (x^\alpha, \dot{x}^\alpha) d\lambda = \frac{1}{2} \int g_{\mu\nu}(x^\alpha) \dot{x}^\mu \dot{x}^\nu d\lambda, \]

where we have set

\[ \dot{x}^\mu = \frac{dx^\mu}{d\lambda}. \quad (11.15) \]

\( \lambda \) can be the proper time if we consider massive particles, or an affine parameter which parametrizes the geodesic, if we consider massless particles. The Euler-Lagrange equations are obtained, as usual, by varying the action with respect to the coordinates, and by setting the variation equal to zero. By varying a curve \( x^\mu(\lambda) \)

\[ x^\mu(\lambda) \rightarrow x^\mu(\lambda) + \delta x^\mu(\lambda) \]

with \( \delta x^\mu(\lambda_0) = \delta x^\mu(\lambda_1) = 0 \), the action variation is

\[ \delta S = \int \left( \frac{\partial L}{\partial x^\sigma} \delta x^\sigma + \frac{\partial L}{\partial \dot{x}^\sigma} \delta (\dot{x}^\sigma) \right) d\lambda. \quad (11.16) \]

Since

\[ \delta (\dot{x}^\sigma) = \delta \left( \frac{dx^\sigma}{d\lambda} \right) = \frac{d\delta x^\sigma}{d\lambda}, \]

the last term in eq. (11.16) can be written as

\[ \frac{\partial L}{\partial \dot{x}^\sigma} \delta (\dot{x}^\sigma) = \frac{\partial L}{\partial \dot{x}^\sigma} \frac{d\delta x^\sigma}{d\lambda} = \frac{d}{d\lambda} \left( \frac{\partial L}{\partial x^\sigma} \delta x^\sigma \right) - \frac{d}{d\lambda} \left( \frac{\partial L}{\partial x^\sigma} \right) \delta x^\sigma. \]

When integrated between \( \lambda_0 \) and \( \lambda_1 \) the first term on the RHS vanishes because \( \delta x^\mu(\lambda_0) = \delta x^\mu(\lambda_1) = 0 \), therefore eq. (11.16) becomes

\[ \delta S = \int \left[ \frac{\partial L}{\partial x^\sigma} \delta x^\sigma - \frac{d}{d\lambda} \left( \frac{\partial L}{\partial \dot{x}^\sigma} \right) \delta x^\sigma \right] d\lambda, \quad (11.17) \]
which vanishes for all \( \delta x^\sigma \) if and only if
\[
\frac{\partial L}{\partial x^\sigma} - \frac{d}{d\lambda} \frac{\partial L}{\partial (\dot{x}^\sigma)} = 0. \tag{11.18}
\]

These are the Euler–Lagrange equations. We shall now show that these equations, when written for the action (11.14), are the geodesic equations
\[
\ddot{x}^\gamma + \Gamma^\gamma_{\mu\nu} \dot{x}^\mu \dot{x}^\nu = 0. \tag{11.19}
\]

By substituting the Lagrangian (11.14) in the Euler–Lagrange equations (11.18), and remembering that \( g_{\mu\nu} = g_{\mu\nu}(x^\alpha) \) and \( \dot{x}^\mu = \dot{x}^\mu(\lambda) \), we find
\[
\frac{1}{2} g_{\mu\nu,\alpha} \dot{x}^\mu \dot{x}^\nu - \frac{d}{d\lambda} \left[ \frac{1}{2} g_{\mu\nu} (\delta^\mu_{\alpha} \dot{x}^\nu + \dot{x}^\mu \delta^\nu_{\alpha}) \right] = 0 \tag{11.20}
\]
\[
= g_{\mu\nu,\alpha} \dot{x}^\mu \dot{x}^\nu - \frac{d}{d\lambda} \left[ g_{\alpha\nu} \dot{x}^\nu + g_{\alpha\mu} \dot{x}^\mu \right]
\]
\[
= g_{\mu\nu,\alpha} \dot{x}^\mu \dot{x}^\nu - \frac{d}{d\lambda} \left[ 2 g_{\alpha\nu} \dot{x}^\nu \right]
\]
\[
= g_{\mu\nu,\alpha} \dot{x}^\mu \dot{x}^\nu - 2 g_{\alpha\nu,\beta} \dot{x}^\beta \dot{x}^\nu - 2 g_{\alpha\nu} \ddot{x}^\nu = 0
\]

By contracting this equation with \( g^{\alpha\gamma} \) we find
\[
\delta^\gamma_{\nu} \dot{x}^\nu + \frac{1}{2} g^{\alpha\gamma} \left[ - g_{\mu\nu,\alpha} + g_{\alpha\mu,\nu} + g_{\alpha\nu,\mu} \right] \dot{x}^\mu \dot{x}^\nu = 0
\]
i.e.
\[
\ddot{x}^\gamma + \frac{1}{2} g^{\alpha\gamma} [g_{\alpha\nu,\gamma} g_{\alpha\nu,\mu} - g_{\mu\nu,\alpha}] \dot{x}^\mu \dot{x}^\nu = 0 \tag{11.21}
\]
which coincides with eq. (11.19).

### 11.2.2 Geodesics in the Schwarzschild metric

For the Schwarzschild metric, the Lagrangian of a free particle is
\[
\mathcal{L} = \frac{1}{2} \left[ - \left(1 - \frac{2m}{r}\right) \dot{t}^2 + \frac{\dot{r}^2}{(1 - \frac{2m}{r})} + r^2 \dot{\theta}^2 + r^2 \sin^2 \phi \dot{\phi}^2 \right],
\]

(we put \( G = c = 1 \)), and a dot indicates differentiation with respect to \( \lambda \). The equations of motion for \( \dot{t}, \dot{\phi} \) and \( \dot{\theta} \) are:

1) Equation for \( \dot{t} \):
\[
\frac{\partial \mathcal{L}}{\partial \dot{t}} - \frac{d}{d\lambda} \frac{\partial \mathcal{L}}{\partial (\dot{t})} = 0 \quad \rightarrow \quad \frac{d}{d\lambda} \left[ \left(1 - \frac{2m}{r}\right) 2\dot{t} \right] = 0
\]
i.e. 
\[ \dot{t} = \frac{\text{const}}{(1 - \frac{2m}{r})}. \] (11.22)

It should be reminded that, since the Schwarzschild metric admits a timelike Killing vector 
\[ \frac{\partial}{\partial t} \rightarrow \xi_{(t)}^\alpha = (1, 0, 0, 0), \]
there exists an associated conserved quantity for the geodesic motion
\[ g_{\alpha\beta} \xi_{(t)}^\alpha u^\beta = \text{const} \rightarrow g_{00} u^0 = \text{const} \rightarrow \left(1 - \frac{2m}{r}\right) \dot{t} = \text{const} \] (11.23)

where \( u^0 = \dot{t} = \frac{dx^0}{d\lambda} \). Note that this equation coincides with eq. (11.22). As discussed in section 9.3, at radial infinity, where the Schwarzschild metric tends to Minkowski’s metric in spherical coordinates, \( g_{00} \) becomes \( \eta_{00} \) and the equation \( g_{00} u^0 = \text{const} \) reduces to \( u^0 = \text{const} \).

In flat spacetime (putting \( G = c = 1 \)) the energy-momentum vector of a massive particle is \( p^\alpha = m u^\alpha = \{E, m v^\gamma\} \); therefore \( u^0 = \text{const} \) means \( E/m = \text{const} \). Therefore we are entitled to interpret the constant in eqs. (11.22) and (11.23) as the energy per unit mass of the particle at infinity. In this case the parameter \( \lambda \) is the particle proper time. If the particle is massless \( \lambda \) must be an affine parameter which parametrizes the null geodesic, ad it can be chosen in such a way that the constant is the particle energy at infinity. In the following we shall put \( \text{const} = E \) and write eq. (11.22) as
\[ \dot{t} = \frac{E}{(1 - \frac{2m}{r})}. \] (11.24)

2) Equation for \( \dot{\phi} \):

since the Lagrangian does not depend on \( \phi \) it is easy to show that
\[ \frac{d}{d\lambda} \frac{\partial L}{\partial (\dot{\phi})} = 0 \rightarrow \dot{\phi} = \frac{\text{const}}{r^2 \sin^2 \theta}. \] (11.25)

Due to its spherical symmetry, the Schwarzschild metric admits the spacelike Killing vector
\[ \frac{\partial}{\partial \phi} \rightarrow \xi_{(\phi)}^\alpha = (0, 0, 0, 1), \]
which is associate to the conserved quantity
\[ g_{\alpha\beta} \xi_{(\phi)}^\alpha u^\beta = \text{const} \rightarrow r^2 \sin^2 \theta \dot{\phi} = \text{const}; \] (11.26)

again eqs. (11.25) and (11.26) coincide. To understand the meaning of the constant, let us consider the simple case of a particle in circular orbit on the equatorial plane; in this case the conservation equation becomes
\[ r^2 \dot{\phi} = \text{const}; \]
from Newtonian mechanics we know that the particle angular momentum \( \vec{\ell} = \vec{r} \wedge \vec{m} \vec{v} \) is conserved so that, being \( v = r \dot{\phi} \), it follows that \( |\vec{\ell}| = mr^2 \dot{\phi} = \text{const} \). Thus we can interpret the constant as the particle angular momentum per unit mass (or as the particle angular momentum if it is a massless particle) at infinity. In the following we shall put \( \text{const} = L \) and write eq. (11.25) as
\[ \dot{\phi} = \frac{L}{r^2 \sin^2 \theta}. \] (11.27)
3) Equation for $\dot{\theta}$:

$$\frac{\partial L}{\partial \theta} - \frac{d}{d\lambda} \frac{\partial L}{\partial (\dot{\theta})} = 0 \quad \rightarrow \quad \frac{d}{d\lambda}(r^2 \dot{\theta}) = r^2 \sin \theta \cos \phi \dot{\phi}^2.$$ 

Therefore the equation for $\theta$ is

$$\ddot{\theta} = -\frac{2}{r} \dot{r} \dot{\theta} + \sin \theta \cos \phi \dot{\phi}^2. \quad (11.28)$$

We will prove that this equations implies that, as in Newtonian theory, orbits are planar.

Due to the spherical symmetry, the metric is invariant under rotations of the polar axes. Using this freedom, we choose them such that, for a given value of the affine parameter, say $\lambda = 0$, the particle is on the equatorial plane $\theta = \frac{\pi}{2}$ and its three-velocity ($\dot{r}, \dot{\theta}, \dot{\phi}$) lays on the same plane, i.e. $\theta(\lambda = 0) = \frac{\pi}{2}$ and $\dot{\theta}(\lambda = 0) = 0$. Thus, we have to solve the following Cauchy problem

$$\ddot{\theta} = -\frac{2}{r} \dot{r} \dot{\theta} + \sin \theta \cos \phi \dot{\phi}^2 \quad (11.29)$$

$$\dot{\theta}(\lambda = 0) = 0 \quad \dot{\theta}(\lambda = 0) = 0$$

which admits a unique solution. Since

$$\theta(\lambda) \equiv \frac{\pi}{2} \quad (11.30)$$

satisfies the differential equation and the initial conditions, it must be the solution. Thus, the orbit is plane and to hereafter we shall assume $\theta = \frac{\pi}{2}$ and $\dot{\theta} = 0$.

4) Equation for $\dot{r}$:

it is convenient to derive this equation from the condition $u_\alpha u^\alpha = -1$, or $u_\alpha u^\alpha = 0$, respectively valid for massive and massless particles.

A) massive particles:

$$g_{\alpha\beta} u^\alpha u^\beta = -\left(1 - \frac{2m}{r}\right) \dot{t}^2 + \frac{\dot{r}^2}{\left(1 - \frac{2m}{r}\right)} + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2 = -1 \quad (11.31)$$

which becomes, by substituting the equations for $\dot{t}$ and $\dot{\Phi}$

$$r^2 + \left(1 - \frac{2m}{r}\right) \left(1 + \frac{L^2}{r^2}\right) = E^2 \quad (11.32)$$

B) massless particles:

$$g_{\alpha\beta} u^\alpha u^\beta = -\left(1 - \frac{2m}{r}\right) \dot{t}^2 + \frac{\dot{r}^2}{\left(1 - \frac{2m}{r}\right)} + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2 = 0 \quad (11.33)$$
which becomes

\[ \dot{r}^2 + \frac{L^2}{r^2} \left( 1 - \frac{2m}{r} \right) = E^2. \]  (11.34)

Finally, the geodesic equations are:

A) For massive particles:

\[ \theta = \frac{\pi}{2}, \quad \dot{\theta} = \frac{E}{\left(1 - \frac{2m}{r}\right)} \]

\[ \dot{\phi} = \frac{L}{r^2}, \quad \dot{r}^2 = E^2 - \left(1 - \frac{2m}{r}\right) \left(1 + \frac{L^2}{r^2}\right) \]  (11.35)

B) For massless particles:

\[ \theta = \frac{\pi}{2}, \quad \dot{\theta} = \frac{E}{\left(1 - \frac{2m}{r}\right)} \]

\[ \dot{\phi} = \frac{L}{r^2}, \quad \dot{r}^2 = E^2 - \frac{L^2}{r^2} \left(1 - \frac{2m}{r}\right) \]  (11.36)

11.3 The orbits of a massless particle

Let us write the radial equation (11.34) in the following form

\[ \dot{r}^2 = E^2 - V(r) \]  (11.37)

where

\[ V(r) = \frac{L^2}{r^2} \left(1 - \frac{2m}{r}\right). \]  (11.38)
Note that:
- For massless particles the angular momentum $L$ acts as a scale factor for the potential
- $V(r)$ tends to $-\infty$ as $r \to 0$, and approaches zero at $r \to \infty$
- $V(r)$ has only one maximum at $r_{\text{max}} = 3m$, where it takes the value
  \[ V_{\text{max}} = \frac{L^2}{27m^2} \tag{11.39} \]

It is useful to consider also the radial acceleration, obtained by differentiating eq. (11.37) with respect to $\lambda$

\[ 2\ddot{r}\dot{r} = -\frac{dV(r)}{dr}\dot{r} \quad \rightarrow \quad \ddot{r} = -\frac{1}{2} \frac{dV(r)}{dr}. \tag{11.40} \]

Let us assume that the particle, say a photon, starts its path from $+\infty$ with $\dot{r} < 0$. The energy of the particle can be:

1) $E^2 > V_{\text{max}}$

according to eq. (11.37) $\dot{r}^2 > 0$ always, and the particle falls into the central body with increasing radial velocity, possibly making several revolutions around the central body before falling in.

2) $E^2 = V_{\text{max}}$

as the particle approaches $r_{\text{max}}, |\dot{r}|$ decreases and tends to zero as $r = r_{\text{max}}$. Since at $r = r_{\text{max}}$ the radial acceleration is zero (see eq. 11.40), if a particle, at a given time, has $r = r_{\text{max}}$ and $\dot{r} = 0$ (i.e. $E = V_{\text{max}}$), it remains at the same $r$ at later times, i.e. its orbit is circular. This is, however, an unstable orbit; indeed if the position is perturbed, the particle will
either fall into the central body; this happens when the radial coordinate of the particle is displaced to \( r < r_{\text{max}} \), since there the radial acceleration is negative

- or escape toward infinity; this happens when the radial coordinate is displaced to \( r > r_{\text{max}} \), since there the radial acceleration is positive.

Thus, for massless particles, there exists only one circular, unstable orbit, and for this orbit

\[
E^2 = \frac{L^2}{2\gamma m^2}.
\]  

(11.41)

3) \( E^2 < V_{\text{max}} \)

be \( r_0 \) the abscissa of the point where \( E^2 = V(r) \) (see figure); for \( r > r_0 \), \( i^2 \) is always positive and becomes zero at \( r = r_0 \). This is a turning point: the particle cannot penetrate the potential barrier and reach values of \( r < r_0 \) because \( \dot{r} \) would become imaginary; since at \( r = r_0 \) the radial acceleration is positive, the particle is forced to invert its radial velocity, and it escapes toward infinity on an open trajectory.

Thus, according to General relativity a light ray is deflected by the gravitational field of a massive body, provided its energy satisfies the following condition

\[
E^2 < \frac{L^2}{2\gamma m^2}.
\]  

(11.42)

11.3.1 The deflection of light

We shall now compute the deflection angle that a massive body induces on the trajectory of a massless particle, say a photon. Referring to the figure 11.1, we shall use the following notation:

- \( r \) is the radial coordinate of the particle in a frame centered in the center of attraction; \( r \) forms an angle \( \phi \) with the \( y \)-axis.
- \( b \) is the impact parameter, i.e. the distance between the direction of the incoming particle (dashed vertical line) and the center of attraction.
- \( \delta \) is the deflection angle which we are going to evaluate: it is the angle between the incoming direction and the outgoing direction (dashed, green line).

Note that, since the Schwarzschild metric is invariant under time reflection, the particle can go through the red trajectory on the figure either in the direction indicated by the red arrow, or in the opposite one. Thus, the trajectory must be symmetric. The periastron is indicated in the figure as \( r_0 \).

We choose the orientation of the frame axes such that the initial value of \( \phi \) when the particle starts its motion at radial infinity be

\[
\phi^{\text{in}} = 0.
\]  

(11.43)

The outgoing particle will escape to \( r \to \infty \) at

\[
\phi^{\text{out}} = \pi + \delta.
\]  

(11.44)

Our only assumption will be that, for all values of \( r \) reached by the particle,

\[
\frac{m}{r} \ll 1.
\]  

(11.45)
This condition is satisfied, for instance, in the case of a photon deflected by the Sun; indeed, if $R_s$ is the radius of the Sun, then $r \geq R_s$, and

$$\frac{m}{r} \leq \frac{m}{R_s} \sim 10^{-6}. \tag{11.46}$$

From the figure we see that

$$b = \lim_{\phi \to 0} r \sin \phi. \tag{11.47}$$

We shall now express the impact parameter $b$ in terms of the energy and the angular momentum of the particle.

When the particle arrives from infinity, $r$ is large, $\phi \simeq 0$ and

$$b \simeq r \phi \quad \rightarrow \quad \frac{d\phi}{dr} \simeq -\frac{b}{r^2}. \tag{11.48}$$

$\frac{d\phi}{dr}$ can also be derived combining together the third and the fourth eqs. (11.36)

$$\frac{d\phi}{dr} = \pm \frac{L}{r^2 \sqrt{E^2 - V(r)}}. \tag{11.49}$$
taking the limit for $r \to \infty$ it gives

$$\frac{d\phi}{dr} \sim \pm \frac{L}{r^2 E},$$

(11.50)

thus, combining together eqs. (11.48) and (11.50), we find that $b$ can be written as

$$b = \frac{L}{E}.$$  

(11.51)

In order the particle being deflected its energy must satisfy eq. (11.42), and this imposes a constraint on $b$, i.e.

$$b \geq \sqrt{27}m \equiv b_{\text{crit}};$$  

(11.52)

if $b$ is smaller than this critical value, the particle is captured by the central body. Note that if the central body is not a black hole but a star, its radius $R$ is in general larger than $\sqrt{27}m$, so the critical value of the impact parameter is $R$: if the particle reaches the stellar surface, it is not deflected.

To find the deflection angle, let us consider the third and the fourth eqs. (11.36); we introduce a new variable

$$u \equiv \frac{1}{r};$$  

(11.53)

by construction, it must be

$$u(\phi = 0) = 0.$$  

(11.54)

Furthermore, $u$ must also vanish when $\phi = \pi + \delta$, because this value of $\phi$ corresponds to the particle escaping to infinity.

In terms of the variable $u$, the third equation (11.36) for $\dot{\phi}$ becomes

$$\dot{\phi} = Lu^2.$$  

By indicating with a prime differentiation with respect to $\phi$ we find that

$$\dot{r} = r'\dot{\phi} = -\frac{1}{u^2}u'\dot{\phi} = -Lu'.$$

By substituting this expression in the fourth eq. (11.36), it becomes

$$L^2(u')^2 + u^2L^2 - 2mL^2u^3 = E^2,$$

and differentiating with respect to $\phi$,

$$2L^2u''u' + 2uu'L^2 - 6ML^2u'u^2 = 0.$$  

Dividing by $2L^2u'$, we finally find the equation $u$ must satisfy

$$u'' + u - 3mu^2 = 0,$$  

(11.55)

to which we associate the boundary condition

$$u(\phi = 0) = 0,$$

$$u'(\phi = 0) = \frac{1}{b}.$$  

(11.56)
The second condition is obtained by the relation
\[ u(\phi \simeq 0) = \frac{1}{b} \sin \phi \]
which derives from eq. (11.47).

If the mass of the central body vanishes, equations (11.55) becomes
\[ u'' + u = 0 \] (11.57)
the solution of which
\[ u(\phi) = \frac{1}{b} \sin \phi \quad \rightarrow \quad b = r \sin \phi \] (11.58)
describes the trajectory of a particle which is not deflected.

If there is a central body with a finite mass \( m \), the solution of (11.55) is different from (11.58), and the light ray is deflected. We note that equations (11.55) and (11.57) differ by a term, \( 3mu^2 \), which is much smaller than, say, the term \( u \) by a factor
\[ \frac{3mu^2}{u} = \frac{3m}{r} \ll 1. \] (11.59)

Consequently, it is appropriate to solve eq. (11.55) using a perturbative approach; we shall proceed as follows. We put
\[ u = u^{(0)} + u^{(1)} \] (11.60)
where \( u^{(0)} \) is the solution of equation (11.57),
\[ u^{(0)} = \frac{1}{b} \sin \phi \] (11.61)
and we assume that
\[ u^{(1)} \ll u^{(0)}. \] (11.62)

By substituting (11.60) in eq. (11.55) we find
\[ (u^{(0)})'' + u^{(0)} - 3m(u^{(0)})^2 + (u^{(1)})'' + u^{(1)} - 3m(u^{(1)})^2 - 6mu^{(0)}u^{(1)} = 0. \] (11.63)

Since \( u^{(0)} \) satisfies (11.57), eq. (11.63) becomes
\[ (u^{(1)})'' + u^{(1)} - 3m(u^{(0)})^2 - 3m(u^{(1)})^2 - 6mu^{(0)}u^{(1)} = 0. \] (11.64)
The terms \( 3m(u^{(1)})^2 \) and \( 6mu^{(0)}u^{(1)} \) are of higher order with respect to \( 3m(u^{(0)})^2 \), therefore the leading terms in equation (11.55) are
\[ (u^{(1)})'' + u^{(1)} - 3m(u^{(0)})^2 = 0. \] (11.65)

Consequently,
\[ (u^{(1)})'' + u^{(1)} = \frac{3m}{b^2} \sin^2 \phi = \frac{3m}{2b^2} (1 - \cos 2\phi). \] (11.66)
The solution of (11.66) which satisfies the boundary conditions (11.57) is
\[ u^{(1)} = \frac{3m}{2b^2} \left( 1 + \frac{1}{3} \cos 2\phi - \frac{4}{3} \cos \phi \right), \] (11.67)
as can be checked by direct substitution. It should be noticed that the boundary conditions (11.57) must be satisfied by the complete solution \( u = u^{(0)} + u^{(1)} \). Therefore,
\[ u = \frac{1}{b} \sin \phi + \frac{3m}{2b^2} \left( 1 + \frac{1}{3} \cos 2\phi - \frac{4}{3} \cos \phi \right). \] (11.68)

We now want to find the deflection angle, i.e., the small angle \( \delta \) such that \( u(\pi + \delta) = 0 \). By substituting \( \phi = \pi + \delta \) in (11.68) we finally find
\[ u(\pi + \delta) \simeq -\delta \frac{b}{b} + \frac{3m}{2b^2} \cdot \frac{8}{3}, \] (11.69)
which vanishes for
\[ \delta = \frac{4m}{b}. \] (11.70)

For a light ray which passes close to the surface of the Sun
\[ \delta \sim 1.75 \text{ seconds of arc} \] (11.71)
The first measurement of the deflection of light was done by Eddington, Dayson and Davidson during the solar eclipse in 1919. What was measured was the apparent position of a star behind the Sun (see figure) during the eclipse, when some light coming from the star was able to reach the Earth because the luminosity of the Sun was obscured by the eclipse. Comparing this apparent position with the position of the star as measured when the Earth is on the opposite side of its orbit around the Sun, one finds \( \delta \phi \). The deflection was measured with an accuracy of about 10% at that time. Today, the bending of radio waves by quasars has been measured with an accuracy of 1%. 

![Diagram of light deflection](image)
11.4 The orbits of a massive particle

Let us first discuss the orbits that a massive particle is allowed to move on. The equations of motion are

\[ \theta = \frac{\pi}{2}, \quad \dot{\theta} = \frac{E}{1 - \frac{2m}{r}} \] \hspace{1cm} (11.72)

\[ \dot{\phi} = \frac{L}{r^2}, \quad \dot{r}^2 = E^2 - \left(1 - \frac{2m}{r}\right) \left(1 + \frac{L^2}{r^2}\right). \] \hspace{1cm} (11.73)

Let us study the radial equation

\[ \dot{r}^2 = E^2 - V(r), \] \hspace{1cm} (11.73)

where

\[ V(r) = \left(1 - \frac{2m}{r}\right) \left(1 + \frac{L^2}{r^2}\right). \] \hspace{1cm} (11.74)

First of all we note that, contrary to the massless case, the potential does not scale with the angular momentum and that \( V(r) \rightarrow 1 \) when \( r \rightarrow \infty \). To plot the potential, let us first see if it admits a minimum or a maximum by solving

\[ \frac{\partial V}{\partial r} = 2mr^2 - L^2 r + 3mL^2 = 0; \]

this equation has two roots

\[ r_{\pm} = \frac{L^2 \pm \sqrt{L^4 - 12m^2L^2}}{2m}. \] \hspace{1cm} (11.75)

If \( L^2 < 12m^2 \) the roots are complex and there are no extrema; the potential will have the shape shown in the figure.
from which is clear that a particle arriving from infinity with $\dot{r} \leq 0$ and having $L^2 < 12m^2$ will be captured by the black hole.

If $L^2 > 12m^2$ the potential has the following form
$V(r)$ has a maximum in $r = r_-$ followed by a minimum in $r = r_+$; thus, a particle with energy $E^2 = V(r_-) \equiv V_{\text{max}}$ will move on an unstable circular orbit at $r = r_-$, whereas if $E^2 = V(r_+) \equiv V_{\text{min}}$ it will move on a stable circular orbit at $r = r_+$. (See the discussion for $E^2 = V_{\text{max}}$ in section 11.3.) Depending on the value of $L$ the maximum of the potential can be greater or smaller than 1, i.e.

\begin{align*}
\text{a)} \quad & L^2 > 16m^2 \quad V_{\text{max}} > 1, \\
\text{b)} \quad & 12m^2 < L^2 < 16m^2 \quad V_{\text{max}} < 1.
\end{align*}

Case b) is shown in the following figure.
Therefore:
in case a) a particle with $V_{\text{min}} < E^2 < 1$ will move on an ellipse (actually, an approximate ellipse as we will see below), if $1 < E^2 < V_{\text{max}}$ and $\dot{r} \leq 0$ it will approach the black hole, reach a turning point $r_0$ where $E^2 = V(r_p)$ and $\dot{r} = 0$ then, since it cannot penetrate the barrier, it will invert its radial velocity and escape free at infinity. (See the discussion for $E^2 < V_{\text{max}}$ in section 11.3).
Conversely, if $E^2 > V_{\text{max}}$ and $\dot{r} \leq 0$ it will fall in the black hole.

In case b) a particle with $V_{\text{min}} < E^2 < V_{\text{max}}$ will move on an elliptic orbit, whereas if $E^2 > 1$ and $\dot{r} \leq 0$, since $\dot{r}^2 = E^2 - V(r)$, it will approach the black hole horizon with increasing velocity and finally fall in.

From the expression of $r_\pm$ given in eq. (11.75) we see that if $L^2 = 12m^2$ the two roots coincide and $r_- = r_+ = 6m$; furthermore, $r_+$ is an increasing function of $L$ and, as $L \to \infty$, $r_+ \to \infty$. This means that there cannot exist stable circular orbits with radius smaller than $6m$. In addition, it is easy
to show that $r_-$ is a decreasing function of $L$ and, as $L \to \infty$,

$$r_- = \frac{L^2}{2m} \left( 1 - \sqrt{1 - \frac{12m^2}{L^2}} \right) \to \frac{L^2}{2m} \left( 1 - (1 - \frac{6m^2}{L^2} + ...) \right) = 3m + O\left(\frac{m}{L}\right);$$

therefore, unstable circular orbits exist only between

$$3m < r_- < 6m.$$

11.4.1 The radial fall of a massive particle

Let us consider a massive particle falling radially into a Schwarzschild black hole. In this case $d\phi/d\tau = 0$, therefore $L = 0$; moreover, since the particle is moving inwards, $\dot{r} < 0$. Equations (11.72) become

$$\frac{dt}{d\tau} = \frac{E}{1 - \frac{2m}{r}}, \quad \frac{dr}{d\tau} = -\sqrt{E^2 - 1 + \frac{2m}{r}}, \quad \frac{d\theta}{d\tau} = 0, \quad \frac{d\phi}{d\tau} = 0. \quad (11.76)$$

If we consider a particle which is at rest at infinity, i.e. such that

$$\lim_{r \to \infty} \frac{dr}{d\tau} = 0 \quad (11.78)$$

from (11.76) it follows that

$$E = 1 \quad (11.79)$$

and the equations for $t$ and $r$ reduce to

$$\frac{dt}{d\tau} = \frac{1}{1 - \frac{2m}{r}}, \quad \frac{dr}{d\tau} = -\sqrt{\frac{2m}{r}}. \quad (11.80)$$

We shall now integrate these equations.

- Putting $r_0 \equiv r(\tau = 0)$, eq. (11.81) gives

$$\tau(r) = -\int_{r_0}^{r} dr' \sqrt{\frac{r'}{2m}} = -\frac{1}{\sqrt{2m}} \int_{r_0}^{r} dr'(r')^{1/2} = \frac{2}{3} \frac{1}{\sqrt{2m}} \left( r_0^{3/2} - r^{3/2} \right); \quad (11.82)$$

the radial trajectory $r(\tau)$ is the inverse function of $\tau(r)$; it is not known analytically, but it can be found numerically.
To find \( t(r) \), we combine equations (11.80) and (11.81):

\[
\frac{dt}{dr} = -\frac{1}{1 - \frac{2m}{r}} \sqrt{\frac{r}{2m}}. \tag{11.83}
\]

If we set \( t = 0 \) when \( \tau = 0 \) we find

\[
t(r) = \int_0^t dt' = -\int_{r_0}^r \frac{1}{1 - \frac{2m}{r'}} \sqrt{\frac{r'}{2m}}; \tag{11.84}
\]

by solving the integral in (11.84) we get (we omit the explicit computation and give only the result):

\[
t(r) = \frac{2}{3} \sqrt{2m} \left[ r_0^{3/2} - r^{3/2} + 6mr_0^{1/2} - 6mr^{1/2} \right] + 2m \ln \left[ \frac{\sqrt{r_0} - \sqrt{2m} \sqrt{r} + \sqrt{2m}}{\sqrt{r_0} + \sqrt{2m} \sqrt{r} - \sqrt{2m}} \right]; \tag{11.85}
\]

\( r(t) \) is the inverse function of \( t(r) \) and, as \( r(\tau) \), is not known analytically.

In figure 11.2 we plot \( t(r) \) and \( \tau(r) \).

![Graph of t(r) and \( \tau(r) \)](image)

Figure 11.2:

Assuming for simplicity \( r_0 \gg 2m \), the behaviour of \( t(r) \) for \( r \to 2m \) and \( r \gg 2m \) is:

- for \( r \approx 2m \)
  \[
t \simeq -2m \ln(\sqrt{r} - \sqrt{2m}) + \text{const.} \to \infty \tag{11.86}
\]

- for \( r \gg 2m \)
  \[
t \simeq 2 \sqrt{2m} \left( r_0^{3/2} - r^{3/2} \right) \equiv \tau. \tag{11.87}
\]

From eq. (11.86) we see that for \( r \to 2m \), \( t(r) \) diverges\(^2\) while eq. (11.82) shows that \( \tau(r) \) is \textit{regular} at \( r = 2m \). The inverse functions \( r(\tau) \) and \( r(t) \) are plotted in figure 11.3. From figure

\(^2\)We also note that even if the coordinate frame \( \{t,r,\theta,\phi\} \) is defined in \( \{0 < r < 2m\} \cup \{r > 2m\} \), namely, outside and inside the horizon, these coordinates are really meaningful (i.e., they are useful to describe physical processes) only for \( r > 2m \), i.e. outside the horizon.
11.3 we also see that $r(\tau)$, which is the radial trajectory as a function of the proper time, i.e. as seen by an observer moving with the particle, for $r = 2m$ has a regular behaviour: this observer does not feel anything strange in crossing the horizon, and after crossing it he reaches the singularity in a finite amount of proper time.

The function $r(t)$, instead, approaches $r = 2m$ only asymptotically. In order to understand what is the meaning of this behaviour, let us consider a spaceship which, while falling radially into the black hole, sends an SOS in the form of a sequence of equally spaced electromagnetic pulses; these signals are received by an observer at radial infinity (the spaceship and the observer have the same $\phi = \text{const}$), located at $r = r^{\text{obs}}$. The SOS travels along null geodesics $t = t(\lambda)$, $r = r(\lambda)$, with $\theta, \phi$ constants and $L = 0$; $\lambda$ is the affine parameter along the geodesic. Therefore, from eqs. (11.36) we find

$$\theta = \frac{\pi}{2}, \quad \phi = \text{const}, \quad \frac{dt}{d\lambda} = \frac{E}{1 - \frac{2m}{r}}, \quad \frac{dr}{d\lambda} = \pm E$$

(11.88)

hence

$$\frac{dt}{dr} = \pm \frac{r}{r - 2m},$$

(11.89)

and the solution is

$$t = \pm r_* + \text{const},$$

(11.90)

where $r_*$ is the tortoise coordinate already introduced in eq. (10.74)

$$r_* \equiv r + 2m \log \left(\frac{r}{2m} - 1\right),$$

(11.91)

so that

$$\frac{dr}{dr_*} = 1 - \frac{2m}{r}.$$  

(11.92)

As in (10.75) we define the outgoing coordinate

$$u \equiv t - r_*$$

(11.93)

so that a given outgoing null geodesic is characterized by a constant value of $u$.  

Let us consider two electromagnetic pulses sent from the spaceship as it approaches the horizon, the first at $\tau = \tau_1$, the second at $\tau = \tau_2$ (see figure 11.4.1). The two pulses correspond to $u = u_1$ and $u = u_2$, respectively. The observer at infinity detects the pulses at two values of its own proper time, which coincides with the coordinate time, i.e. at $t = t^{\text{obs}}_1$ and $t = t^{\text{obs}}_2$. Thus, while the person on the spaceship measures a proper time interval between the pulses

$$\Delta \tau = \tau_2 - \tau_1, \quad (11.94)$$

the observer at infinity measures a corresponding coordinate time interval

$$\Delta t^{\text{obs}} = t^{\text{obs}}_2 - t^{\text{obs}}_1 = (u_2 + r^{\text{obs}}_*) - (u_1 + r^{\text{obs}}_*) = u_2 - u_1 = \Delta u. \quad (11.95)$$

Since $u$ is constant along the two null geodesics, $u_1$ and $u_2$ can also be evaluated in terms of points along the spaceship geodesic:

$$u_1 = t(\tau_1) - r_*(\tau_1)$$
$$u_2 = t(\tau_2) - r_*(\tau_2) \quad (11.96)$$

thus $\Delta u = \Delta t(\tau) - \Delta r_*(\tau)$. Therefore, assuming that the pulses are emitted at very short time intervals, we can write (11.93), we find

$$\frac{\Delta t^{\text{obs}}}{\Delta \tau} = \frac{\Delta u}{\Delta \tau} \simeq \frac{dt}{d\tau} - \frac{dr_*}{d\tau} = \frac{dt}{d\tau} - \frac{dr_*}{dr} \frac{dr}{d\tau} = \frac{1}{1 - \frac{2m}{r}} \left(1 + \sqrt{\frac{2m}{r}}\right). \quad (11.97)$$

![Figure 11.4: A spaceship radially falling into the black hole sends electromagnetic signals to a distant observer](image-url)
This equation shows that as $r \to 2m$, $\frac{\Delta t_{\text{obs}}}{\Delta r} \to \infty$, which means that the time interval between pulses as detected by the observer at infinity increases, and finally diverges, as the spaceship approaches the horizon.

It is interesting to note that the right hand side of eq. (11.97) has two terms: the first is the square of the gravitational redshift, the second is a Doppler contribution due to the fact that, while sending the pulses, the ship is moving away from the observer.

### 11.4.2 The motion of a planet around the Sun

Let us now use the geodesic equations (11.72) to study the motion of a planet around the Sun. We can consider the limit

$$\frac{m}{r} \ll 1,$$

indeed, if we consider Mercury, which is the closest planet to the Sun, since the Mercury-Sun distance is $r \sim 5.8 \cdot 10^7$ km, we find

$$\frac{GM_\odot}{rc^2} = \frac{1.4768}{5.8 \cdot 10^7} \sim 2.5 \cdot 10^{-8}.$$ 

In what follows, we shall indicate with a prime differentiation with respect to $\phi$, and use the variable $u \equiv \frac{1}{r}$, as we did in Section 11.3.1. In terms of $u$, the equation for $\dot{\phi}$ becomes

$$\dot{\phi} = Lu^2,$$

and

$$\dot{r} = r' \dot{\phi} = -\frac{1}{u^2} u' \dot{\phi} = -Lu'.$$

By substituting in eq. (11.72), it becomes

$$L^2 (u')^2 + 1 - 2mu + u^2 L^2 - 2mL^2 u^3 = E,$$

and differentiating with respect to $\phi$,

$$2L^2 uu' - 2mu' + 2uu' L^2 - 6mL^2 u'u^2 = 0.$$ 

Dividing by $2L^2 u'$, we find the equation for $u$

$$u'' + u - \frac{m}{L^2} - 3mu^2 = 0.$$ 

#### The Newtonian equation

The Newtonian equation which corresponds to the third eq. (11.72) is derived from the energy conservation law

$$\frac{1}{2} m_p \left[(\dot{r})^2 + r^2 (\dot{\phi})^2\right] - \frac{m_p m}{r} = \text{const} \quad \Rightarrow \quad (\dot{r})^2 - \frac{2m}{r} + \frac{L^2}{r^2} = \text{const}.$$
where $m_p$ is the particle mass and we have set $G = 1$. By expressing (11.4.2) in terms of $u$ and differentiating with respect to $\phi$, we find
\[2L^2u''u' - 2mu' + 2uu'L^2 = 0,\]
which becomes
\[u'' + u - \frac{m}{L^2} = 0.\quad (11.100)\]
Equation (11.100) differs from equation (11.99) only by the term $3mu^2$, which is smaller than, say, $u$ by a factor
\[3mu = \frac{3m}{r} \ll 1.\]
Equation (11.100) can be written as
\[(u - \frac{m}{L^2})'' + (u - \frac{m}{L^2}) = 0,\]
the solution of which is
\[u - \frac{m}{L^2} = A \cos(\phi - \phi_0) \quad \Rightarrow \quad u = \frac{m}{L^2} \left[1 + \frac{L^2A}{m} \cos(\phi - \phi_0)\right],\]
where $\phi_0$ and $A$ are integration constants. In terms of $r$ the solution is
\[r = \frac{L^2}{m} \frac{1}{1 + \frac{L^2A}{m} \cos(\phi - \phi_0)}.\]
If we set
\[e = \frac{L^2A}{m},\quad (11.101)\]
the previous equation becomes
\[r = \frac{L^2}{m} \frac{1}{1 + e \cos(\phi - \phi_0)},\quad (11.102)\]
which describes an ellipse with eccentricity $e$ in polar coordinates $(r, \phi)$. If we set for example $\phi_0 = 0$, we see that the periastron, i.e. the minimum distance the planet reaches in its motion around the central body (perihelion if the central body is the Sun) occurs when $\phi = 0$, i.e.
\[r_{\text{periastron}} = \frac{L^2}{m} \frac{1}{1 + e}.\quad (11.103)\]
The apastron (the maximum distance from the central body, aphelion in the case of the Sun) is
\[r_{\text{apastron}} = \frac{L^2}{m} \frac{1}{1 - e}.\quad (11.104)\]
It is worth noting that, since
\[\frac{m}{L^2} = \frac{1}{r_{\text{periastron}}(1 + e)} \quad \Rightarrow \quad \frac{m^2}{L^2} = \frac{m}{r_{\text{periastron}}(1 + e)},\]
and since $m/r \ll 1$, it follows that
\[\frac{m^2}{L^2} \ll 1.\quad (11.105)\]
The relativistic equations

In order to solve equation (11.99)

\[ u'' + u - \frac{m}{L^2} - 3mu^2 = 0 \]  \hspace{1cm} (11.106)

we adopt the same perturbative approach used in Section 11.3.1 to study the deflection of light by a massive body. We search for a solution in the form

\[ u = u^{(0)} + u^{(1)} \]

where \( u^{(0)} \) is the solution of the Newtonian equation, i.e.

\[ u^{(0)} = \frac{m}{L^2} (1 + e \cos \phi) , \]

and

\[ u^{(1)} \ll u^{(0)} . \]

Proceeding as for eq. (11.55) we find

\[ (u^{(0)})'' + u^{(0)} - \frac{m}{L^2} - 3m(u^{(0)})^2 + (u^{(1)})'' + u^{(1)} - 3m(u^{(1)})^2 - 6mu^{(0)}u^{(1)} = 0 . \] \hspace{1cm} (11.107)

Since \( u^{(0)} \) satisfies (11.100), eq. (11.107) becomes

\[ (u^{(1)})'' + u^{(1)} - 3m(u^{(0)})^2 - 3m(u^{(1)})^2 - 6mu^{(0)}u^{(1)} = 0 . \] \hspace{1cm} (11.108)

The terms \( 3m(u^{(1)})^2 \) and \( 6mu^{(0)}u^{(1)} \) are of higher order with respect to \( 3m(u^{(0)})^2 \), therefore the leading terms in equation (11.55) are

\[ (u^{(1)})'' + u^{(1)} = 3m(u^{(0)})^2 , \] \hspace{1cm} (11.109)
CHAPTER 11. EXPERIMENTAL TESTS OF GENERAL RELATIVITY

i.e.

\[(u^{(1)})'' + u^{(1)} = 3 \frac{m^3}{L^4} (1 + e \cos \phi)^2. \tag{11.110}\]

Let us expand the right-hand side

\[
(u^{(1)})'' + u^{(1)} = 3 \frac{m^3}{L^4} \left[1 + e^2 \cos^2 \phi + 2e \cos \phi \right]
\]

\[
= 3 \frac{m^3}{L^4} \left[1 + \frac{1}{2} e^2 (1 + \cos 2\phi) + 2e \cos \phi \right]
\]

\[
= 3 \frac{m^3}{L^4} \left[ \cos + \frac{1}{2} e^2 \cos 2\phi + 2e \cos \phi \right].
\]

This is the equation of an harmonic oscillator with three forcing terms. They are all very small, because, as shown in eq. (11.105), \(\frac{m^2}{L^2} \ll 1\). However, the term

\[2e \cos(\phi),\]

is in resonance with the free oscillations of the harmonic oscillator, therefore, even if its amplitude is comparable to that of the other terms, it determines a secular perturbation of the planet motion which, after a long time, becomes relevant. For this reason, we will neglect the constant term and the term \(\frac{1}{2} e^2 \cos 2\phi\) and look for the solution of the resulting equation

\[(u^{(1)})'' + u^{(1)} = 6e \frac{m^3}{L^4} \cos(\phi). \tag{11.111}\]

As can be checked by direct substitution, the solution of this equation is

\[u^{(1)} = \frac{3em^3}{L^4} \phi \sin \phi ,\]

therefore, the complete solution is

\[u = \frac{m}{L^2} \left[1 + e \left( \cos \phi + 3 \frac{m^2}{L^2} \phi \sin \phi \right)\right].\]

At first order in \(m^2/L^2\),

\[\cos \left( \frac{3m^2}{L^2} \phi \right) \simeq 1\]

\[\sin \left( \frac{3m^2}{L^2} \phi \right) \simeq \frac{3m^2}{L^2} \phi\]

therefore we can write

\[u \sim \frac{m}{L^2} \left[1 + e \cos \phi \left(1 - 3 \frac{m^2}{L^2}\right)\right].\tag{11.112}\]

A comparison with the corresponding newtonian equation shows that the term \(\frac{3m^2}{L^2} \phi\) determines a secular precession of the periastron. When the argument of the sinusoidal function
in eq. (11.112) goes from zero to $2\pi$, i.e. when the planet reaches again the radial distance $r = r_{\text{periastron}}$, $\phi$ changes by

$$\Delta \phi = \frac{2\pi}{1 - \frac{3m^2}{L^2}} \simeq 2\pi \left(1 + \frac{3m^2}{L^2}\right).$$

Consequently, in one period the periastron is shifted by

$$\Delta \phi_P = \frac{6\pi m^2}{L^2}, \quad (11.113)$$

as shown in the following figure.

Thus, in general relativity the orbit of a planet around a central object is not an ellipse; it is an open orbit, and the periastron shifts by $\Delta \phi_P$ at each revolution.

For example, for Mercury equation (11.113) gives a precession of 42.98 arcsec/century. The observed value, after all effects which can be explained with newtonian theory (precession of the equinoxes, perturbations of other planets on Mercury’s orbit, etc) is 42.98 ± 0.04 arcsec/century.
Chapter 12

The Geodesic deviation

The Principle of equivalence establishes that we can always choose a locally inertial frame
where the affine connections vanish and the metric becomes that of a flat spacetime. Con-
versely, if the spacetime is flat we can always define a coordinate system which “simulates”,
locally, the existence of any arbitrary gravitational field. In this frame we could measure
the “simulated” gravitational force by studying the motion of a single particle, but these
measurements would never allow us to know whether that force is simulated or real: this
can be understood only by comparing the motion of close particles, i.e. by comparing the
behaviour of close geodesics.

12.1 The equation of geodesic deviation

Consider two particles moving along the trajectories $x^\mu(\tau)$ and $x^\mu(\tau) + \delta x^\mu(\tau)$, where $\delta x^\mu$
is the vector of separation between the two close geodesics, and $\tau$ is an affine parameter.
This is equivalent to say: consider a two-parameter family of geodesics $x^\mu(\tau, p)$, where the
parameter $p$ labels different geodesics

\[ t^\alpha = \frac{\partial x^\alpha}{\partial \tau} \quad (12.1) \]
CHAPTER 12. THE GEODESIC DEVIATION

162

the tangent vector to the geodesic line, and be

$$\delta x^\alpha = \frac{\partial x^\alpha}{\partial p}. \quad (12.2)$$

Note that

$$\frac{\partial t^\alpha}{\partial p} = \frac{\partial \delta x^\alpha}{\partial \tau}. \quad (12.3)$$

We now compute the covariant derivative of the vector $\vec{t}$ along the curve $\tau = \text{const}$ whose tangent vector is $\delta x^\mu$, i.e. $\nabla_{\delta x} \vec{t}$. The components of this vector are

$$\left( \nabla_{\delta x} \vec{t} \right)^\alpha = \frac{\partial x^\mu}{\partial p} \left[ \frac{\partial t^\alpha}{\partial x^\mu} + \Gamma^\alpha_{\mu\nu} t^\nu \right] = \frac{\partial t^\alpha}{\partial p} + \Gamma^\alpha_{\mu\nu} t^\nu \delta x^\mu. \quad (12.4)$$

Similarly, the covariant derivative of the vector $\delta \vec{x}$ along the curve $p = \text{const}$, i.e. along the geodesic, has components

$$\left( \nabla_{\vec{t}} \delta \vec{x} \right)^\alpha = t^\mu \delta x^\alpha;\mu = \frac{\partial \delta x^\alpha}{\partial \tau} + \Gamma^\alpha_{\mu\nu} \delta x^\nu t^\mu. \quad (12.5)$$

From eq. (12.3) and from the symmetry of $\Gamma^\alpha_{\mu\nu}$ in the lower indices it follows that

$$\nabla_{\vec{t}} \delta \vec{x} = \nabla_{\delta x} \vec{t}. \quad (12.6)$$

The quantities $\left( \nabla_{\vec{t}} \delta \vec{x} \right)^\alpha$ or $\left( \nabla_{\delta x} \vec{t} \right)^\alpha$ involve only the affine connections, and therefore they do not give significant information on the gravitational field. We then compute the second covariant derivative of the vector $\delta \vec{x}$ along the curve $p = \text{const}$, i.e. $\nabla_{\vec{t}} \left( \nabla_{\vec{t}} \delta \vec{x} \right)$. We define the following operator:

$$D \frac{d}{d\tau} \delta x^\alpha \equiv \left( \nabla_{\vec{t}} \delta \vec{x} \right)^\alpha = t^\mu \delta x^\alpha;\mu. \quad (12.7)$$

With this definition,

$$D^2 \delta x^\alpha \frac{d}{d\tau^2} = \left( \nabla_{\vec{t}} \left( \nabla_{\vec{t}} \delta \vec{x} \right) \right)^\alpha. \quad (12.8)$$

This quantity, called geodesic deviation, is a vector describing the relative acceleration of two nearby geodesics.

In order to compute the geodesic deviation, let us consider the commutator

$$\left[ \nabla_{\vec{t}}, \nabla_{\delta x} \right] \vec{t} = \nabla_{\vec{t}} \left( \nabla_{\delta x} \vec{t} \right) - \nabla_{\delta x} \left( \nabla_{\vec{t}} \vec{t} \right). \quad (12.9)$$

whose components are

$$\left[ \nabla_{\vec{t}} \left( \nabla_{\delta x} \vec{t} \right) \right]^\alpha = t^\mu \delta x^\nu t^\alpha;\nu;\mu - \delta x^\mu \left( t^\nu t^\alpha;\nu;\mu \right) \quad (12.10)$$

From eq. (12.6) we find that

$$t^\mu \delta x^\nu;\mu = \delta x^\nu t^\mu;\mu.$$
and eq. (12.11) becomes
\[ \left[ \nabla_t \left( \nabla_{\delta x} \, \dot{t} \right) \right]^{\alpha} = (t^{\alpha ; \nu \mu} - t^{\alpha , \mu ; \nu}) \, t^{\mu} \delta x^{\nu}. \] (12.11)

We now remind that, according to eq. (6.43), the commutator of covariant derivatives is
\[ (t^{\alpha ; \nu \mu} - t^{\alpha , \mu ; \nu}) = R^{\alpha}_{\beta \mu \nu} t^{\beta}, \] (12.12)
therefore eq. (12.11) becomes
\[ \left[ \nabla_t \left( \nabla_{\delta x} \, \dot{t} \right) \right]^{\alpha} = R^{\alpha}_{\beta \mu \nu} t^{\beta} t^{\mu} \delta x^{\nu}. \] (12.13)

Moreover, since \( t^{\mu} \) is the geodesic tangent vector, when it is parallel-transported along the geodesic it gives (see Section 5.9)
\[ \nabla_t \, \bar{t} = 0; \] (12.14)
as a consequence \( \nabla_{\delta x} \left( \nabla_t \, \bar{t} \right) = 0 \) and the commutator (12.9) can be rewritten as
\[ \left( \left[ \nabla_t, \nabla_{\delta x} \right] \, \bar{t} \right)^{\alpha} = \left( \nabla_t \left( \nabla_t \, \delta x \right) \right)^{\alpha} = R^{\alpha}_{\beta \mu \nu} t^{\beta} t^{\mu} \delta x^{\nu}. \] (12.15)

By direct substitution of this expression in eq. (12.8) we finally find
\[ \frac{D^2 \delta x^{\alpha}}{d\tau^2} = R^{\alpha}_{\beta \mu \nu} t^{\beta} t^{\mu} \delta x^{\nu}. \] (12.16)

This is the equation of geodesic deviation, which shows that the relative acceleration of nearby particles moving along geodesics depends on the curvature tensor. Since the Riemann tensor is zero if and only if the gravitational field is either zero or constant and uniform, the equation of the geodesic deviation really contains the information on the gravitational field in a given spacetime.
Chapter 13
Gravitational Waves

One of the most interesting predictions of the theory of General Relativity is the existence of gravitational waves. The idea that a perturbation of the gravitational field should propagate as a wave is, in some sense, intuitive. For example electromagnetic waves were introduced when the Coulomb theory of electrostatics was replaced by the theory of electrodynamics, and it was shown that they transport through space the information about the evolution of charged systems. In a similar way when a mass-energy distribution changes in time, the information about this change should propagate in the form of waves. However, gravitational waves have a distinctive feature: due to the twofold nature of \( g_{\mu\nu} \), which is the metric tensor and the gravitational potential, gravitational waves are metric waves. Thus when they propagate the geometry, and consequently the proper distance between spacetime points, change in time.

Gravitational waves can be studied by following two different approaches, one based on perturbative methods, the second on the solution of the non linear Einstein equations.

*The perturbative approach*

Be \( g^0_{\mu\nu} \) a known exact solution of Einstein’s equations; it can be, for instance, the metric of flat spacetime \( \eta_{\mu\nu} \), or the metric generated by a Schwarzschild black hole. Let us consider a small perturbation of \( g^0_{\mu\nu} \) caused by some source described by a stress-energy tensor \( T_{\mu\nu}^{\text{pert}} \). We shall write the metric tensor of the perturbed spacetime, \( g_{\mu\nu} \), as follows

\[
g_{\mu\nu} = g^0_{\mu\nu} + h_{\mu\nu}, \tag{13.1}
\]

where \( h_{\mu\nu} \) is the small perturbation

\[
|h_{\mu\nu}| << |g^0_{\mu\nu}|.
\]

It is clear that this assumption is ambiguous, because we should specify in which reference frame this is true; however we shall assume that this frame does exists.

The inverse metric can be written as

\[
g^{\mu\nu} = g^{0\mu\nu} - h^{\mu\nu} + O(h^2), \tag{13.2}
\]

where the indices of \( h^{\mu\nu} \) have been raised with the unperturbed metric

\[
h^{\mu\nu} \equiv g^{0\mu\alpha} g^{0\nu\beta} h_{\alpha\beta}. \tag{13.3}
\]
Indeed, with this definition,

\[
(g^0{}_{\mu \nu} - h^0{}_{\mu \nu})(g^0_{\alpha \alpha} + h_{\alpha \alpha}) = \delta^\mu_\alpha + O(h^2).
\]  

(13.4)

In order to find the equations that describe \( h_{\mu \nu} \), we shall write Einstein’s equations for the metric (13.1) in the form

\[
R_{\mu \nu} = \frac{8 \pi G}{c^4} \left( T_{\mu \nu} - \frac{1}{2} g_{\mu \nu} T^\lambda_\lambda \right),
\]

(13.5)

where \( T_{\mu \nu} \) is the sum of two terms, one associate to the source that generates the background geometry \( g^0_{\mu \nu} \), say \( T^0_{\mu \nu} \), and one associate to the source of the perturbation \( T^\text{pert}_{\mu \nu} \). We remind that the Ricci tensor \( R_{\mu \nu} \) is

\[
R_{\mu \nu} = \frac{\partial}{\partial x^\alpha} \Gamma^\alpha_{\mu \nu} - \frac{\partial}{\partial x^\nu} \Gamma^\alpha_{\mu \alpha} + \Gamma^\alpha_{\sigma \alpha} \Gamma^\sigma_{\mu \nu} - \Gamma^\alpha_{\sigma \nu} \Gamma^\sigma_{\mu \alpha},
\]

(13.6)

and that the affine connections \( \Gamma^\gamma_{\beta \mu} \) are

\[
\Gamma^\gamma_{\beta \mu} = \frac{1}{2} g^{\alpha \gamma} \left[ \frac{\partial}{\partial x^\mu} g_{\alpha \beta} + \frac{\partial}{\partial x^\beta} g_{\alpha \mu} - \frac{\partial}{\partial x^\alpha} g_{\beta \mu} \right].
\]

(13.7)

The \( \Gamma^\gamma_{\beta \mu} \) computed for the perturbed metric (13.1) are

\[
\Gamma^\gamma_{\beta \mu} (g_{\mu \nu}) = \frac{1}{2} g^{0\alpha \gamma} \left[ \left( \frac{\partial}{\partial x^\mu} g^0_{\alpha \beta} + \frac{\partial}{\partial x^\beta} g^0_{\alpha \mu} - \frac{\partial}{\partial x^\alpha} g^0_{\beta \mu} \right) + \left( \frac{\partial}{\partial x^\mu} h_{\alpha \beta} + \frac{\partial}{\partial x^\beta} h_{\alpha \mu} - \frac{\partial}{\partial x^\alpha} h_{\beta \mu} \right) \right]
\]

\[
= \frac{1}{2} g^{0\alpha \gamma} \left[ \frac{\partial}{\partial x^\mu} g^0_{\alpha \beta} + \frac{\partial}{\partial x^\beta} g^0_{\alpha \mu} - \frac{\partial}{\partial x^\alpha} g^0_{\beta \mu} \right] + \frac{1}{2} g^{0\alpha \gamma} \left[ \frac{\partial}{\partial x^\mu} h_{\alpha \beta} + \frac{\partial}{\partial x^\beta} h_{\alpha \mu} - \frac{\partial}{\partial x^\alpha} h_{\beta \mu} \right]
\]

\[
- \frac{1}{2} h^{\alpha \gamma} \left[ \frac{\partial}{\partial x^\mu} g^0_{\alpha \beta} + \frac{\partial}{\partial x^\beta} g^0_{\alpha \mu} - \frac{\partial}{\partial x^\alpha} g^0_{\beta \mu} \right] + O(h^2)
\]

\[
= \Gamma^\gamma_{\beta \mu} (g^0) + \Gamma^\gamma_{\beta \mu} (h) + O(h^2),
\]

(13.8)

where \( \Gamma^\gamma_{\beta \mu} (h) \) are the terms that are first order in \( h_{\mu \nu} \)

\[
\Gamma^\gamma_{\beta \mu} (h) = \frac{1}{2} g^{0\alpha \gamma} \left[ \frac{\partial}{\partial x^\mu} h_{\alpha \beta} + \frac{\partial}{\partial x^\beta} h_{\alpha \mu} - \frac{\partial}{\partial x^\alpha} h_{\beta \mu} \right] - \frac{1}{2} h^{\alpha \gamma} \left[ \frac{\partial}{\partial x^\mu} g^0_{\alpha \beta} + \frac{\partial}{\partial x^\beta} g^0_{\alpha \mu} - \frac{\partial}{\partial x^\alpha} g^0_{\beta \mu} \right].
\]

(13.9)

When we substitute these expressions of the \( \Gamma^\gamma_{\beta \mu} (g_{\mu \nu}) \) in the Ricci tensor we get

\[
R_{\mu \nu} (g_{\mu \nu}) = R_{\mu \nu} (g^0)
\]

\[
+ \frac{\partial}{\partial x^\alpha} \Gamma^\alpha_{\mu \nu} (h) - \frac{\partial}{\partial x^\nu} \Gamma^\alpha_{\mu \alpha} (h)
\]

\[
+ \Gamma^\alpha_{\sigma \alpha} (g^0) \Gamma^\sigma_{\mu \nu} (h) + \Gamma^\alpha_{\sigma \nu} (h) \Gamma^\sigma_{\mu \alpha} (g^0)
\]

\[
- \Gamma^\alpha_{\sigma \nu} (g^0) \Gamma^\sigma_{\mu \alpha} (h) - \Gamma^\alpha_{\sigma \alpha} (h) \Gamma^\sigma_{\mu \nu} (g^0) + O(h^2)
\]

\[
= \frac{8 \pi G}{c^4} \left( T_{\mu \nu} - \frac{1}{2} g_{\mu \nu} T^\lambda_\lambda \right).
\]

(13.10)
Since \( g^0_{\mu\nu} \) is by assumption an exact solution of Einstein’s equations in vacuum \( R_{\mu\nu} (g^0) = \frac{8\pi G}{c^4} \left( T^0_{\mu\nu} - \frac{1}{2} g^0_{\mu\nu} T^0_{\lambda\lambda} \right) \); thus, if we retain only first order terms, the equations for the perturbations \( h_{\mu\nu} \) reduce to

\[
\frac{\partial}{\partial x^\alpha} \Gamma^\alpha_{\mu\nu} (h) - \frac{\partial}{\partial x^\nu} \Gamma^\alpha_{\mu\alpha} (h) \tag{13.11}
\]

\[
+ \Gamma^\alpha_{\sigma\alpha} (g^0) \Gamma^\sigma_{\mu\nu} (h) + \Gamma^\alpha_{\sigma\alpha} (h) \Gamma^\sigma_{\mu\nu} (g^0)
\]

\[
- \Gamma^\alpha_{\sigma\nu} (g^0) \Gamma^\sigma_{\mu\alpha} (h) - \Gamma^\alpha_{\sigma\nu} (h) \Gamma^\sigma_{\mu\alpha} (g^0) = \frac{8\pi G}{c^4} \left( T^\text{pert}_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T^\text{pert}_{\lambda\lambda} \right).
\]

that are linear in \( h_{\mu\nu} \); their solution will describe the propagation of gravitational waves in the considered background.\(^1\) This approximation works sufficiently well in a variety of physical situations because gravitational waves are very weak. This point will be better understood in the next chapter, when we will discuss the generation of gravitational waves.

The ”exact” approach

The second approach to the study of gravitational waves seeks for exact solutions of Einstein’s equations which describe both the source and the emitted wave, but no solution of this kind has been found so far. Of course the non-linearity of the equations makes the problem very difficult; however, it may be noted that also in electrodynamics an exact solution of Maxwell’s equations appropriate to describe the electromagnetic field produced by a current which decreases in an electric oscillator due to the emission of electromagnetic waves has never been found, although Maxwell’s equations are linear.

Exact solutions of Einstein’s equations describing gravitational waves can be found only if one imposes some particular symmetry as for example plane, spherical, or cylindrical symmetry. The interaction of plane waves can also be described in terms of exact solutions, and due to the non-linearity of the equations of gravity it is very different from the interaction of electromagnetic waves.

In the following we shall use the perturbative approach to show that a weak perturbation of the flat spacetime satisfies the wave equation.

### 13.1 A perturbation of the flat spacetime propagates as a wave

Let us consider the flat spacetime described by the metric tensor \( \eta_{\mu\nu} \) and a small perturbation \( h_{\mu\nu} \), such that the resulting metric can be written as

\[
g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad |h_{\mu\nu}| << 1. \tag{13.12}
\]

The affine connections (13.8) computed for the metric (13.12) give

\[
\Gamma^\lambda_{\mu\nu} = \frac{1}{2} \eta^{\lambda\rho} \left[ \frac{\partial}{\partial x^\rho} h_{\mu\nu} + \frac{\partial}{\partial x^\nu} h_{\rho\mu} - \frac{\partial}{\partial x^\mu} h_{\rho\nu} \right] + O(h^2). \tag{13.13}
\]

\(^1\)Notice that the right-hand side of eq.(13.11) is a particular case of the Palatini identity.
CHAPTER 13. GRAVITATIONAL WAVES

167

Since the metric $g^0_{\mu\nu} \equiv \eta_{\mu\nu}$ is constant, $\Gamma^\lambda_{\mu\nu}(g^0) = 0$ and the right-hand side of eq. (13.11) simply reduces to

$$\frac{\partial \Gamma^\alpha_{\mu\nu}}{\partial x^\alpha} - \frac{\partial \Gamma^\alpha_{\mu\alpha}}{\partial x^\nu} + O(h^2),$$

(13.14)

The operator $\Box_F$ is the D’Alambertian in flat spacetime

$$\Box_F \equiv \eta^{\alpha\beta} \frac{\partial}{\partial x^\alpha} \frac{\partial}{\partial x^\beta} = -\frac{\partial^2}{c^2 \partial t^2} + \nabla^2.$$ (13.15)

Einstein’s equations (13.5) for $h_{\mu\nu}$ finally become

$$\left\{ \Box_F h_{\mu\nu} - \left[ \frac{\partial^2}{\partial x^\lambda \partial x^\mu} h^\lambda_{\nu} + \frac{\partial^2}{\partial x^\lambda \partial x^\nu} h^\lambda_{\mu} - \frac{\partial^2}{\partial x^\mu \partial x^\nu} h^\lambda_{\lambda} \right] \right\} = -\frac{16\pi G}{c^4} \left( T^{\mu\nu}_{\text{pert}} - \frac{1}{2} \eta_{\mu\nu} T^\lambda_{\text{pert} \lambda} \right).$$ (13.16)

As already discussed in chapter 8, the solution of eqs. (13.16) is not uniquely determined. If we make a coordinate transformation, the transformed metric tensor is still a solution: it describes the same physical situation seen from a different frame. But since we are working in the weak field limit, we are entitled to make only those transformations which preserve the condition $|h'_{\mu\nu}| \ll 1$ (note that in this Section we denote the transformed tensor as $h'_{\mu\nu}$ rather than as $h'_{\mu'\nu'}$, since this simplifies the discussion of infinitesimal coordinate transformations).

If we make an infinitesimal coordinate transformation

$$x'^\mu = x^\mu + \epsilon^\mu(x),$$ (13.17)

(the prime refers to the coordinate $x'^\mu$, not to the index $\mu$) where $\epsilon^\mu$ is an arbitrary vector such that $\frac{\partial \epsilon^\mu}{\partial x^\nu}$ is of the same order of $h_{\mu\nu}$, then

$$\frac{\partial x'^\alpha}{\partial x^\nu} = \delta^\alpha_\mu + \frac{\partial \epsilon^\alpha}{\partial x^\nu}.$$ (13.18)

Since

$$g_{\mu\nu} = \left( \eta_{\alpha\beta} + h'_{\alpha\beta} \right) \left( \delta^\alpha_\mu + \frac{\partial \epsilon^\alpha}{\partial x^\mu} \right) \left( \delta^\beta_\nu + \frac{\partial \epsilon^\beta}{\partial x^\nu} \right)$$

$$= \eta_{\mu\nu} + h'_{\mu\nu} + \epsilon_{\mu,\nu} + \epsilon_{\nu,\mu} + O(h^2),$$ (13.19)

and $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, then (up to $O(h^2)$)

$$h'_{\mu\nu} = h_{\mu\nu} - \frac{\partial \epsilon_{\mu}}{\partial x^\nu} - \frac{\partial \epsilon_{\nu}}{\partial x^\mu}.$$ (13.20)

In order to simplify eq. (13.16) it appears convenient to choose a coordinate system in which the harmonic gauge condition is satisfied, i.e.

$$g^{\mu\nu} \Gamma^\lambda_{\mu\nu} = 0.$$ (13.21)
Let us see why. This condition is equivalent to say that, up to terms that are first order in $h_{\mu\nu}$, the following equation is satisfied

$$\frac{\partial}{\partial x^\mu}h_{\mu\nu} = \frac{1}{2} \frac{\partial}{\partial x^\nu}h^{\mu}_{\mu}. \quad (13.22)$$

Using this condition the term in square brackets in eq. (13.16) vanishes, and Einstein’s equations reduce to a simple wave equation supplemented by the condition (13.22)

$$\begin{cases}
\Box_F h_{\mu\nu} = -\frac{16\pi G}{c^4} \left( T_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} T^{\lambda}_{\lambda} \right) \\
\frac{\partial}{\partial x^\nu}h^{\mu}_{\mu} = \frac{1}{2} \frac{\partial}{\partial x^\nu}h^{\mu}_{\mu},
\end{cases} \quad (13.23)$$

(to hereafter, we omit the superscript ‘pert’ to indicate the stress-energy tensor associated to the source of the perturbation). If we introduce the tensor

$$\bar{h}_{\mu\nu} \equiv h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h^{\lambda}_{\lambda}, \quad (13.24)$$

eqs. (13.23) become

$$\begin{cases}
\Box_F \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu} \\
\frac{\partial}{\partial x^\nu}h^{\mu}_{\mu} = 0,
\end{cases} \quad (13.25)$$

and outside the source where $T_{\mu\nu} = 0$

$$\begin{cases}
\Box_F \bar{h}_{\mu\nu} = 0 \\
\frac{\partial}{\partial x^\nu}h^{\mu}_{\mu} = 0.
\end{cases} \quad (13.26)$$

Thus, we have shown that a perturbation of a flat spacetime propagates as a wave travelling at the speed of light, and that Einstein’s theory of gravity predicts the existence of gravitational waves.

As in electrodynamics, the solution of eqs. (13.25) can be written in terms of retarded potentials

$$\bar{h}_{\mu\nu}(t, x) = \frac{4G}{c^4} \int \frac{T_{\mu\nu}(t - \frac{|x - x'|}{c}, x')}{|x - x'|} \, d^3 x', \quad (13.27)$$

and the integral extends over the past light-cone of the event $(t, x)$ . This equation represents the gravitational waves generated by the source $T_{\mu\nu}$.

We may now ask how eqs. (13.26) and (13.25) should be modified if, instead of considering the perturbation of a flat spacetime, we would consider the perturbation of a curved
background. For example, suppose \( g_{\mu\nu}^0 \) is the Schwarzschild solution for a non rotating black hole. In this case, it is possible to show that, by a suitable choice of the gauge, the Einstein equations written for certain combinations of the components of the metric tensor, can be reduced to a form similar to eqs. (13.25). However, since the background spacetime is now curved, the propagation of the waves will be modified with respect to the flat case. The curvature will act as a potential barrier by which waves are scattered and the final equation will have the form

\[
\Box_F \Phi - V(x) \Phi = -\frac{16\pi G}{c^4} T
\]

where \( \Phi \) is the appropriate combination of metric functions, \( T \) is a combination of the stress-energy tensor components, \( \Box_F \) is the d’Alambertian of the flat spacetime and \( V \) is the potential barrier generated by the spacetime curvature. In other words, the perturbations of a spherically symmetric, stationary gravitational field would be described by a Schroedinger-like equation! A complete account on the theory of perturbations of black holes can be found in the book *The Mathematical Theory of Black Holes* by S. Chandrasekhar, Oxford: Claredon Press, (1984).

### 13.2 How to choose the harmonic gauge

We shall now show that if the harmonic-gauge condition is not satisfied in a reference frame, we can always find a new frame where it is, by making an infinitesimal coordinate transformation

\[
x'\lambda = x^\lambda + \epsilon^\lambda,
\]

provided

\[
\Box_F \epsilon_\rho = \frac{\partial h_\rho^\beta}{\partial x^\beta} - \frac{1}{2} \frac{\partial h_\beta^\rho}{\partial x^\rho}. (13.30)
\]

Indeed, when we change the coordinate system \( \Gamma^\lambda = g^{\mu\nu} \Gamma_{\mu\nu}^\lambda \) transforms according to equation (8.63), i.e.

\[
\Gamma^\lambda = \frac{\partial x'\lambda}{\partial x^\rho} \Gamma^\rho + g^{\rho\sigma} \frac{\partial^2 x'\lambda}{\partial x^\rho \partial x^\sigma},
\]

where, from eq. (13.29)

\[
\frac{\partial x'\lambda}{\partial x^\rho} = \delta^\lambda_\rho + \frac{\partial \epsilon^\lambda}{\partial x^\rho}.
\]

If \( g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \) (see footnote after eq. (13.21))

\[
\Gamma^\rho = \eta^{\rho\sigma} \left\{ h^\mu_{\kappa;\mu} - \frac{1}{2} h^\nu_{\kappa;\nu} \right\};
\]

moreover

\[
g^{\rho\sigma} \frac{\partial^2 x'\lambda}{\partial x^\rho \partial x^\sigma} = g^{\rho\sigma} \left[ \frac{\partial}{\partial x^\rho} \left( \frac{\partial x^\lambda}{\partial x^\sigma} + \frac{\partial \epsilon^\lambda}{\partial x^\sigma} \right) \right] =
\]

\[
g^{\rho\sigma} \left[ \frac{\partial}{\partial x^\rho} \left( \delta^\lambda_\sigma + \frac{\partial \epsilon^\lambda}{\partial x^\sigma} \right) \right] \simeq \eta^{\rho\sigma} \left[ \frac{\partial^2 \epsilon^\lambda}{\partial x^\rho \partial x^\sigma} \right] = \Box_F \epsilon^\lambda,
\]
therefore in the new gauge the condition $\Gamma^\lambda = 0$ becomes
\[
\Gamma^\lambda = \left[ \delta^\lambda_\rho + \frac{\partial \lambda}{\partial x^\rho} \right] \eta^{\rho\mu} \left[ \frac{\partial h^\mu_\kappa}{\partial x^\nu} - \frac{1}{2} \frac{\partial h^\nu_\kappa}{\partial x^\mu} \right] + \Box F \epsilon^\lambda = 0. \tag{13.34}
\]
If we neglect second order terms in $h$ eq. (13.34) becomes
\[
\Gamma^\lambda = \eta^{\lambda\kappa} \left[ \frac{\partial h^\mu_\kappa}{\partial x^\mu} - \frac{1}{2} \frac{\partial h^\nu_\kappa}{\partial x^\nu} \right] + \Box F \epsilon^\lambda = 0.
\]
Contracting with $\eta_{\lambda\alpha}$ and remembering that $\eta_{\lambda\alpha} \eta^{\lambda\kappa} = \delta^\kappa_\alpha$ we finally find
\[
\Box F \epsilon_\alpha = - \left( \frac{\partial h^\mu_\alpha}{\partial x^\mu} - \frac{1}{2} \frac{\partial h^\nu_\alpha}{\partial x^\nu} \right).
\]
This equation can in principle be solved to find the components of $\epsilon_\alpha$, which identify the coordinate system in which the harmonic gauge condition is satisfied.

### 13.3 Plane gravitational waves

The simplest solution of the wave equation in vacuum (13.26) is a monocromatic plane wave
\[
\bar{h}_{\mu\nu} = \Re \left\{ A_{\mu\nu} e^{ik_\alpha x^\alpha} \right\}, \tag{13.35}
\]
where $A_{\mu\nu}$ is the polarization tensor, i.e. the wave amplitude and $\vec{k}$ is the wave vector. By direct substitution of (13.35) into the first equation we find
\[
\Box F \bar{h}_{\mu\nu} = \eta^{\alpha\beta} \frac{\partial}{\partial x^\alpha} \frac{\partial}{\partial x^\beta} (e^{ik_\gamma x^\gamma}) = \eta^{\alpha\beta} \frac{\partial}{\partial x^\alpha} \left[ i k_\gamma \frac{\partial x^\gamma}{\partial x^\beta} e^{ik_\gamma x^\gamma} \right] = \eta^{\alpha\beta} \frac{\partial}{\partial x^\alpha} \left[ i k_\alpha k_\beta e^{ik_\gamma x^\gamma} \right] = -\eta^{\alpha\beta} k_\alpha k_\beta e^{ik_\gamma x^\gamma} = 0, \quad \rightarrow \quad \eta^{\alpha\beta} k_\alpha k_\beta = 0,
\]
thus, (13.35) is a solution of (13.26) if $\vec{k}$ is a null vector. In addition the harmonic gauge condition requires that
\[
\frac{\partial}{\partial x^\mu} \bar{h}^\mu_\nu = 0, \tag{13.37}
\]
which can be written as
\[
\eta^{\mu\alpha} \frac{\partial}{\partial x^\mu} \bar{h}_{\alpha\nu} = 0. \tag{13.38}
\]
Using eq. (13.35) it gives
\[
\eta^{\mu\alpha} \frac{\partial}{\partial x^\mu} A_{\alpha\nu} e^{ik_\gamma x^\gamma} = 0 \quad \rightarrow \quad \eta^{\mu\alpha} A_{\alpha\nu} k_\mu = 0 \quad \rightarrow \quad k_\mu A^\mu_\nu = 0. \tag{13.39}
\]
This further condition expresses the orthogonality of the wave vector and of the polarization tensor.
CHAPTER 13. GRAVITATIONAL WAVES

Since $h_{\mu\nu}$ is constant on those surfaces where $k_\alpha x^\alpha = \text{const}$, (13.40)

these are the equations of the wavefront. It is conventional to refer to $k^0$ as $\frac{\omega}{c}$, where $\omega$ is the frequency of the waves. Consequently

$$\vec{k} = (\frac{\omega}{c}, k).$$ (13.41)

Since $k$ is a null vector

$$-(k_0)^2 + (k_x)^2 + (k_y)^2 + (k_z)^2 = 0,$$ i.e. (13.42)

$$\omega = ck_0 = c\sqrt{(k_x)^2 + (k_y)^2 + (k_z)^2},$$ (13.43)

where $(k_x, k_y, k_z)$ are the components of the unit 3-vector $k$.

13.4 The $TT$-gauge

We now want to see how many of the ten components of $h_{\mu\nu}$ have a real physical meaning, i.e. what are the degrees of freedom of a gravitational plane wave. Let us consider a wave propagating in flat spacetime along the $x^1 = x$-direction. Since $h_{\mu\nu}$ is independent of $y$ and $z$, eqs. (13.26) become (as before we raise and lower indices with $\eta_{\mu\nu}$)

$$\left(-\frac{\partial^2}{c^2 \partial t^2} + \frac{\partial^2}{\partial x^2}\right) \bar{h}^\mu_{\nu} = 0,$$ (13.44)

i.e. $\bar{h}^\mu_{\nu}$ is an arbitrary function of $t \pm \frac{x}{c}$, and

$$\frac{\partial}{\partial x^\mu} \bar{h}^\mu_{\nu} = 0.$$ (13.45)

Let us consider, for example, a progressive wave $\bar{h}^\mu_{\nu} = \bar{h}^\mu_{\nu} [\chi(t, x)]$, where $\chi(t, x) = t - \frac{x}{c}$. Being

$$\begin{align*}
\frac{\partial}{\partial t} \bar{h}^\mu_{\nu} &= \frac{\partial \bar{h}^\mu_{\nu}}{\partial x} \frac{\partial x}{\partial t} = \frac{\partial \bar{h}^\mu_{\nu}}{\partial x}, \\
\frac{\partial}{\partial x} \bar{h}^\mu_{\nu} &= \frac{\partial \bar{h}^\mu_{\nu}}{\partial x} \frac{\partial x}{\partial x} = -\frac{1}{c} \frac{\partial \bar{h}^\mu_{\nu}}{\partial x},
\end{align*}$$ (13.46)

eq. (13.45) gives

$$\frac{\partial}{\partial x^\mu} \bar{h}^\mu_{\nu} = \frac{1}{c} \frac{\partial \bar{h}^t_{\nu}}{\partial t} + \frac{1}{c} \frac{\partial \bar{h}^x_{\nu}}{\partial x} = \frac{1}{c} \frac{\partial}{\partial x} \left[ \bar{h}^t_{\nu} - \bar{h}^x_{\nu} \right] = 0.$$ (13.47)

This equation can be integrated, and the constants of integration can be set equal to zero because we are interested only in the time-dependent part of the solution. The result is

$$\bar{h}^t_t = \bar{h}^x_t, \quad \bar{h}^t_y = \bar{h}^x_y,$$ (13.48)

$$\bar{h}^t_x = \bar{h}^x_x, \quad \bar{h}^t_z = \bar{h}^x_z.$$
We now observe that the harmonic gauge condition does not determine the gauge uniquely. Indeed, if we make an infinitesimal coordinate transformation

$$x'^\mu = x^\mu + \epsilon^\mu,$$

from eq.  (13.31) we find that, if in the old frame $\Gamma^\rho = 0$, in the new frame $\Gamma^\lambda = 0$, provided

$$\eta^{\rho\sigma} \frac{\partial^2 x^\lambda}{\partial x^\rho \partial x^\sigma} = 0,$$

namely, if $\epsilon^\mu$ satisfies the wave equation

$$\Box_F \epsilon^\mu = 0.$$  

(13.51)

If we have a solution of the wave equation,

$$\Box_F \tilde{h}_{\mu\nu} = 0$$

(13.52)

and we perform a gauge transformation, the perturbations in the new gauge

$$h'_{\mu\nu} = h_{\mu\nu} - \partial_\mu \epsilon_\nu - \partial_\nu \epsilon_\mu$$

(13.53)

give

$$\tilde{h}'_{\mu\nu} = \tilde{h}_{\mu\nu} - \partial_\mu \epsilon_\nu - \partial_\nu \epsilon_\mu - \eta_{\mu\nu} \partial^\alpha \epsilon_\alpha$$

(13.54)

and, due to (13.51), the new tensor is solution of the wave equation,

$$\Box_F \tilde{h}'_{\mu\nu} = 0.$$  

(13.55)

It can be shown that the converse is also true: it is always possible to find a vector $\epsilon^\mu$ satisfying (13.51) to set to zero four components of $\tilde{h}_{\mu\nu}$ solution of (13.52).

Thus, we can use the four functions $\epsilon^\mu$ to set to zero the following four quantities

$$\tilde{h}'_x = \tilde{h}'_y = \tilde{h}'_z = \tilde{h}'_t = 0.$$  

(13.56)

From eq. (13.48) it then follows that

$$\tilde{h}'_x = \tilde{h}_x, \quad \tilde{h}'_y = \tilde{h}_y, \quad \tilde{h}'_z = \tilde{h}_z = \tilde{h}'_t = 0.$$  

(13.57)

The remaining non-vanishing components are $\tilde{h}'_y$ and $\tilde{h}'_z$. These components cannot be set equal to zero, because we have exhausted our gauge freedom.

From eqs. (13.56) and (13.57) it follows that

$$\tilde{h}'_{\mu\nu} = \tilde{h}'_t + \tilde{h}'_x + \tilde{h}'_y + \tilde{h}'_z = 0,$$

(13.58)

and since

$$\tilde{h}'_{\mu\nu} = h_{\mu\nu} - 2h_{\mu\nu} = -h_{\mu\nu},$$

(13.59)

it follows that

$$h_{\mu\nu} = 0, \quad \tilde{h}'_{\mu\nu} = h_{\mu\nu},$$

(13.60)
CHAPTER 13. GRAVITATIONAL WAVES

i.e. in this gauge \( h_{\mu\nu} \) and \( \tilde{h}_{\mu\nu} \) coincide and are traceless. Thus, a plane gravitational wave propagating along the \( x \)-axis is characterized by two functions \( h_{xy} \) and \( h_{yy} = -h_{zz} \), while the remaining components can be set to zero by choosing the gauge as we have shown:

\[
h_{\mu\nu} = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & h_{yy} & h_{yz} \\
0 & 0 & h_{yz} & -h_{yy}
\end{pmatrix}.
\] (13.61)

In conclusion, a gravitational wave has only two physical degrees of freedom which correspond to the two possible polarization states. The gauge in which this is clearly manifested is called the \( TT \)-gauge, where ‘\( TT \)’ indicates that the components of the metric tensor \( h_{\mu\nu} \) are different from zero only on the plane orthogonal to the direction of propagation (transverse), and that \( h_{\mu\nu} \) is traceless.

13.5 How does a gravitational wave affect the motion of a single particle

Consider a particle at rest in flat spacetime before the passage of the wave. We set an inertial frame attached to this particle, and take the \( x \)-axis coincident with the direction of propagation of an incoming \( TT \)-gravitational wave. The particle will follow a geodesic of the curved spacetime generated by the wave

\[
\frac{d^2x^\alpha}{d\tau^2} + \Gamma^\alpha_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = \frac{dU^\alpha}{d\tau} + \Gamma^\alpha_{\mu\nu} U^\mu U^\nu = 0.
\] (13.62)

At \( t = 0 \) the particle is at rest \((U^\alpha = (1, 0, 0, 0))\) and the acceleration impressed by the wave will be

\[
\left( \frac{dU^\alpha}{d\tau} \right)_{(t=0)} = -\Gamma^\alpha_{00} = -\frac{1}{2} \eta^{\alpha\beta} [h_{\beta0,0} + h_{0\beta,0} - h_{00,\beta}],
\] (13.63)

but since we are in the \( TT \)-gauge it follows that

\[
\left( \frac{dU^\alpha}{d\tau} \right)_{(t=0)} = 0.
\] (13.64)

Thus, \( U^\alpha \) remains constant also at later times, which means that the particle is not accelerated neither at \( t = 0 \) nor later! It remains at a constant coordinate position, regardless of the wave. We conclude that the study of the motion of a single particle is not sufficient to detect a gravitational wave.

13.6 Geodesic deviation induced by a gravitational wave

We shall now study the relative motion of particles induced by a gravitational wave. Consider two neighbouring particles \( A \) and \( B \), with coordinates \( x_A^\mu, x_B^\mu \). We shall assume that the two
particles are initially at rest, and that a plane-fronted gravitational wave reaches them at some time \( t = 0 \), propagating along the \( x \)-axis. We shall also assume that we are in the \( TT \)-gauge, so that the only non-vanishing components of the wave are those on the \((y,z)\)-plane. In this frame, the metric is

\[
d s^2 = g_{\mu\nu} \, dx^\mu \, dx^\nu = (\eta_{\mu\nu} + h^{TT}_{\mu\nu}) \, dx^\mu \, dx^\nu. \tag{13.65}
\]

Since \( g_{00} = \eta_{00} = -1 \), we can assume that both particles have proper time \( \tau = ct \). Since the two particles are initially at rest, they will remain at a constant coordinate position even later, when the wave arrives, and their coordinate separation

\[
\delta x^\mu = x_B^\mu - x_A^\mu \tag{13.66}
\]
remains constant. However, since the metric changes, the proper distance between them will change. For example if the particles are on the \( y \)-axis,

\[
\Delta l = \int ds = \int_{y_A}^{y_B} |g_{yy}|^{\frac{1}{2}} dy = \int_{y_A}^{y_B} \left[ 1 + h^{TT}_{yy}(t - x/c) \right]^{\frac{1}{2}} dy \neq \text{constant}. \tag{13.67}
\]

We now want to study the effect of the wave by using the equation of geodesic deviation. To this purpose, it is convenient to change coordinate system and use a locally inertial frame \( \{x^{\prime\alpha}\} \) centered on the geodesic of one of the two particles, say the particle \( A \); in the neighborhood of \( A \) the metric is

\[
d s^2 = \eta_{\alpha\beta} dx^{\alpha'} dx^{\beta'} + O(|\delta x|^2). \tag{13.68}
\]
i.e. it differs from Minkowski’s metric by terms of order \( |\delta x|^2 \). It may be reminded that, as discussed in Chapter 1, it is always possible to define such a frame.

In this frame the particle \( A \) has space coordinates \( x^{\prime\alpha}_A = 0 \) \((i = 1, 2, 3)\), and

\[
t_A = \tau/c, \quad \frac{dx^{\prime\mu}}{d\tau} \big|_A = (1, 0, 0, 0), \quad g_{\mu'\nu'} |_A = \eta_{\mu'\nu'}, \quad g_{\mu'\nu',\alpha'} |_A = 0 \quad (\text{i.e. } \Gamma_{\mu'\nu'}^{\alpha'} |_A = 0), \tag{13.69}
\]
where the subscript \( |_A \) means that the quantity is computed along the geodesic of the particle \( A \). Moreover, the space components of the vector \( \delta x^{\mu'} \) which separates \( A \) and \( B \) are the coordinates of the particle \( B \):

\[
x^{\prime\mu'}_B = \delta x^{\mu'}. \tag{13.70}
\]
To simplify the notation, in the following we will rename the coordinates of this locally inertial frame attached to \( A \) as \( \{x^\mu\} \), and we will drop all the primes.

The separation vector \( \delta x^\mu \) satisfies the equation of geodesic deviation (see Chapter 12):

\[
\frac{D^2 \delta x^\mu}{d\tau^2} = R^\mu_{\alpha\beta\gamma} \frac{dx^\alpha}{d\tau} \frac{dx^\beta}{d\tau} \delta x^\gamma. \tag{13.71}
\]
If we evaluate this equation along the geodesic of the particle \( A \), using eqs. (13.69) (removing the primes) we find

\[
\frac{d^2 \delta x^i}{dt^2} = R^i_{00j} \delta x^j. \tag{13.72}
\]
If the gravitational wave is due to a perturbation of the flat metric, as discussed in this chapter, the metric can be written as 

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

and the Riemann tensor

$$R_{\alpha\kappa\lambda\mu} = \frac{1}{2} \left( \frac{\partial^2 g_{\alpha\mu}}{\partial x^\kappa \partial x^\lambda} + \frac{\partial^2 g_{\kappa\lambda}}{\partial x^\alpha \partial x^\mu} - \frac{\partial^2 g_{\alpha\lambda}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 g_{\kappa\mu}}{\partial x^\alpha \partial x^\lambda} \right) + g_{\nu\sigma} \left( \Gamma^\nu_{\kappa\lambda} \Gamma^\sigma_{\alpha\mu} - \Gamma^\nu_{\kappa\mu} \Gamma^\sigma_{\alpha\lambda} \right),$$

(13.73)

after neglecting terms which are second order in \( h_{\mu\nu} \), becomes

$$R_{\alpha\kappa\lambda\mu} = \frac{1}{2} \left( \frac{\partial^2 h_{\alpha\mu}}{\partial x^\kappa \partial x^\lambda} + \frac{\partial^2 h_{\kappa\lambda}}{\partial x^\alpha \partial x^\mu} - \frac{\partial^2 h_{\alpha\lambda}}{\partial x^\kappa \partial x^\mu} - \frac{\partial^2 h_{\kappa\mu}}{\partial x^\alpha \partial x^\lambda} \right) + O(h^2);$$

(13.74)

consequently

$$R_{i00m} = \frac{1}{2} \left( \frac{\partial^2 h_{im}}{\partial x^0 \partial x^0} + \frac{\partial^2 h_{00}}{\partial x^i \partial x^m} - \frac{\partial^2 h_{i0}}{\partial x^0 \partial x^m} - \frac{\partial^2 h_{0m}}{\partial x^i \partial x^0} \right) = \frac{1}{2} h_{TT}^{im,00},$$

(13.75)

because in the \( TT \)-gauge \( h_{i0} = h_{00} = 0 \). \( i \) and \( m \) can assume only the values 2 and 3, i.e. they refer to the \( y \) and \( z \) components. It follows that

$$R^\lambda_{00m} = \eta^{\lambda i} R_{i00m} = \frac{1}{2} \eta^{\lambda i} \frac{\partial^2 h_{TT}^{im}}{c^2 \partial t^2},$$

(13.76)

and the equation of geodesic deviation (13.72) becomes

$$\frac{d^2}{dt^2} \delta x_\lambda = \frac{1}{2} \eta^{\lambda i} \frac{\partial^2 h_{TT}^{im}}{c^2 \partial t^2} \delta x_m.$$  

(13.77)

For \( t \leq 0 \) the two particles are at rest relative to each other, and consequently

$$\delta x^j = \delta x^j_0, \quad \text{with} \quad \delta x^j_0 = \text{const}, \quad t \leq 0.$$  

(13.78)

Since \( h_{\mu\nu} \) is a small perturbation, when the wave arrives the relative position of the particles will change only by infinitesimal quantities, and therefore we put

$$\delta x^j(t) = \delta x^j_0 + \delta x^j_1(t), \quad t > 0,$$

(13.79)

where \( \delta x^j_1(t) \) has to be considered as a small perturbation with respect to the initial position \( \delta x^j_0 \). Substituting (13.79) in (13.77), remembering that \( \delta x^j_0 \) is a constant and retaining only terms of order \( O(h) \), eq. (13.77) becomes

$$\frac{d^2}{dt^2} \delta x^j = \frac{1}{2} \eta^{ji} \frac{\partial^2 h_{TT}^{ik}}{\partial t^2} \delta x^k.$$  

(13.80)

This equation can be integrated and the solution is

$$\delta x^j = \delta x^j_0 + \frac{1}{2} \eta^{ji} h_{TT}^{ik} \delta x^k.$$  

(13.81)
which clearly shows the tranverse nature of the gravitational wave; indeed, using the fact
that if the wave propagates along $x$ only the components $h_{22} = -h_{33}, h_{23} = h_{32}$ are different
from zero, from eqs. (13.81) we find

$$
\delta x^1 = \delta x^1_0 + \frac{1}{2} \eta^{11} h^{TT}_{1k} \delta x^k_0 = \delta x^1_0
$$

$$
\delta x^2 = \delta x^2_0 + \frac{1}{2} \eta^{22} h^{TT}_{2k} \delta x^k_0 = \delta x^2_0 + \frac{1}{2} (h^{TT}_{22} \delta x^2_0 + h^{TT}_{23} \delta x^3_0)
$$

$$
\delta x^3 = \delta x^3_0 + \frac{1}{2} \eta^{33} h^{TT}_{3k} \delta x^k_0 = \delta x^3_0 + \frac{1}{2} (h^{TT}_{32} \delta x^2_0 + h^{TT}_{33} \delta x^3_0).
$$

Thus, the particles will be accelerated only in the plane orthogonal to the direction of
propagation.

Let us now study the effect of the polarization of the wave. Consider a plane wave whose
nonvanishing components are (we omit in the following the superscript $TT$)

$$
h_{yy} = -h_{zz} = 2 \Re \{A_+ e^{i \omega (t - \frac{x}{c})}\},
$$

$$
h_{yz} = h_{zy} = 2 \Re \{A_\times e^{i \omega (t - \frac{x}{c})}\}.
$$

Consider two particles located, as indicated in figure (13.1) at $(0, y_0, 0)$ and $(0, 0, z_0)$. Let us
consider the polarization ‘+’ first, i.e. let us assume

$$
A_\times = 0
$$

Assuming $A_+$ real eqs. (13.83) give

$$
h_{yy} = -h_{zz} = 2A_+ \cos \omega (t - \frac{x}{c}),
$$

$$
h_{yz} = h_{zy} = 0.
$$

If at $t = 0$, $\omega (t - \frac{x}{c}) = \frac{\pi}{2}$, eqs. (13.82) written for the two particles for $t > 0$ give

1) $z = 0$, $y = y_0 + \frac{1}{2} h_{yy} y_0 = y_0 [1 + A_+ \cos \omega (t - \frac{x}{c})]$,

2) $y = 0$, $z = z_0 + \frac{1}{2} h_{zz} z_0 = z_0 [1 - A_+ \cos \omega (t - \frac{x}{c})]$.

After a quarter of a period ($\cos \omega (t - \frac{x}{c}) = -1$)

1) $z = 0$, $y = y_0 [1 - A_+]$,

2) $y = 0$, $z = z_0 [1 + A_+]$.

After half a period ($\cos \omega (t - \frac{x}{c}) = 0$)

1) $z = 0$, $y = y_0$,

2) $y = 0$, $z = z_0$.

After three quarters of a period ($\cos \omega (t - \frac{x}{c}) = 1$)

1) $z = 0$, $y = y_0 [1 + A_+]$,

2) $y = 0$, $z = z_0 [1 - A_+]$. 

Figure 13.1:
Figure 13.2:
Similarly, if we consider a small ring of particles centered at the origin, the effect produced by a gravitational wave with polarization ‘+’ is shown in figure (13.2).

Let us now see what happens if \( A_\times \neq 0 \) and \( A_+ = 0 \):

\[
h_{yy} = h_{zz} = 0, \quad h_{yz} = h_{zy} = 2A_\times \cos \omega (t - \frac{x}{c}),
\]

Comparing with eqs. (13.82) we see that a generic particle initially at \( P = (y_0, z_0) \), when \( t > 0 \) will move according to the equations

\[
y = y_0 + \frac{1}{2} h_{yz} \ z_0 = y_0 + z_0 A_\times \cos \omega (t - \frac{x}{c}), \quad (13.91)
\]

\[
z = z_0 + \frac{1}{2} h_{zy} \ y_0 = z_0 + y_0 A_\times \cos \omega (t - \frac{x}{c}).
\]

Let us consider four particles disposed as indicated in figure (13.3)

1) \( y = r, \quad z = r \),

2) \( y = -r, \quad z = r \),

3) \( y = -r, \quad z = -r \),

4) \( y = r, \quad z = -r \).

As before, we shall assume that the initial time \( t = 0 \) corresponds to \( \omega (t - \frac{x}{c}) = \frac{\pi}{2} \). After a quarter of a period \( \cos \omega (t - \frac{x}{c}) = -1 \), the particles will have the following positions

1) \( y = r[1 - A_\times], \quad z = r[1 - A_\times] \),

2) \( y = r[-1 + A_\times], \quad z = r[1 + A_\times] \),

3) \( y = r[-1 + A_\times], \quad z = r[-1 + A_\times] \),

4) \( y = r[1 + A_\times], \quad z = r[-1 - A_\times] \).

After half a period \( \cos \omega (t - \frac{x}{c}) = 0 \), and the particles go back to the initial positions. After three quarters of a period, when \( \cos \omega (t - \frac{x}{c}) = 1 \)

1) \( y = r[1 + A_\times], \quad z = r[1 + A_\times] \),

2) \( y = r[-1 - A_\times], \quad z = r[1 - A_\times] \),

3) \( y = r[-1 - A_\times], \quad z = r[-1 - A_\times] \),

4) \( y = r[1 - A_\times], \quad z = r[-1 + A_\times] \).

The motion of the particles is indicated in figure (13.3).

It follows that a small ring of particles centered at the origin, will again become an ellipse, but rotated at \( 45^\circ \) (see figure (13.4)) with respect to the case previously analysed. In conclusion, we can define \( A_+ \) and \( A_\times \) as the polarization amplitudes of the wave. The wave will be linearly polarized when only one of the two amplitudes is different from zero.
Figure 13.3:
Figure 13.4:
Chapter 14

The Quadrupole Formalism

In this chapter we will introduce the quadrupole formalism which allows to estimate the gravitational energy and the waveforms emitted by an evolving physical system described by the stress-energy tensor $T_{\mu\nu}$. We shall solve eq. (13.25) under the following assumption: we shall assume that the region where the source is confined, namely

$$|x^i| < \epsilon, \quad T_{\mu\nu} \neq 0,$$

is much smaller than the wavelength of the emitted radiation, $\lambda_{GW} = \frac{2\pi c}{\omega}$. This implies that

$$\frac{2\pi c}{\omega} \gg \epsilon \quad \rightarrow \quad \epsilon \omega \ll c \quad \rightarrow \quad v_{\text{typical}} \ll c,$$

i.e. the velocities typical of the physical processes we are considering are much smaller than the speed of light; for this reason this is called the slow-motion approximation.

Let us consider the first equation in (13.25)

$$\Box_F \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} T_{\mu\nu},$$

where

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h$$

and

$$\Box_F = \left[ -\frac{1}{c^2} \frac{\partial^2}{\partial t^2} + \nabla^2 \right].$$

By Fourier-expanding both $\bar{h}_{\mu\nu}$ and $T_{\mu\nu}$

$$T_{\mu\nu}(t, x^i) = \int_{-\infty}^{+\infty} T_{\mu\nu}(\omega, x^i) e^{-i\omega t} d\omega,$$

$$\bar{h}_{\mu\nu}(t, x^i) = \int_{-\infty}^{+\infty} \bar{h}_{\mu\nu}(\omega, x^i) e^{-i\omega t} d\omega, \quad i = 1, 3$$

eq. (14.2) becomes

$$\left[ \nabla^2 + \frac{\omega^2}{c^2} \right] \bar{h}_{\mu\nu}(\omega, x^i) = -K T_{\mu\nu}(\omega, x^i)$$

where

$$K = \frac{16\pi G}{c^4}.$$
We shall solve eq. (14.4) outside and inside the source, matching the two solutions across the source boundary.

**The exterior solution**

Outside the source $T^{\mu\nu} = 0$ and eq. (14.4) becomes

$$\left[ \nabla^2 + \frac{\omega^2}{c^2} \right] \bar{h}_{\mu\nu}(\omega, x^i) = 0. \quad (14.6)$$

In polar coordinates, the Laplacian operator $\nabla^2$ is

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left[ r^2 \frac{\partial}{\partial r} \right] + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta} \left[ \sin \theta \frac{\partial}{\partial \theta} \right] + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2}.$$ 

We shall consider the simplest solution of this equation, i.e. one which does not depend on $\phi$ and $\theta$:

$$\bar{h}_{\mu\nu}(\omega, r) = \frac{A_{\mu\nu}(\omega)}{r} e^{i \frac{\omega x}{c}} + \frac{Z_{\mu\nu}(\omega)}{r} e^{-i \frac{\omega x}{c}}. \quad (14.7)$$

This solution represents a spherical wave, with an *ingoing* part ($\sim e^{-i \frac{\omega x}{c}}$), and an *outgoing* ($\sim e^{i \frac{\omega x}{c}}$) part; indeed, substituting in the second eq. (14.3) $\bar{h}_{\mu\nu}(\omega, x^i)$ by $\sim e^{\pm i \frac{\omega x}{c}}$ the result of the integration over $\omega$ gives a function of $(t \mp \frac{x}{c})$ respectively.

Since we are interested only in the wave emitted *from* the source, we shall set $Z_{\mu\nu} = 0$, and consider the solution

$$\bar{h}_{\mu\nu}(\omega, r) = \frac{A_{\mu\nu}(\omega)}{r} e^{i \frac{\omega x}{c}}. \quad (14.7)$$

This is the solution outside the source and on its boundary, where $T^{\mu\nu}$ vanishes as well. $A_{\mu\nu}$ is the wave amplitude to be found by solving the equations inside the source.

**The interior solution**

The wave equation

$$\left[ \nabla^2 + \frac{\omega^2}{c^2} \right] \bar{h}_{\mu\nu}(\omega, x^i) = -K T_{\mu\nu}(\omega, x^i) \quad (14.8)$$

can be solved for each assigned value of the indices $\mu, \nu$. To solve eq. (14.8) let us integrate over the source volume

$$\int_V \left[ \nabla^2 + \frac{\omega^2}{c^2} \right] \bar{h}_{\mu\nu}(\omega, x^i) d^3 x = -K \int_V T_{\mu\nu}(\omega, x^i) d^3 x.$$ 

The first term can be expanded as follows

$$\int_V \nabla \cdot \bar{h}_{\mu\nu}(\omega, x^i) d^3 x = \int_V \text{div} (\nabla \bar{h}_{\mu\nu}) d^3 x = \int_S (\nabla \bar{h}_{\mu\nu}) \times dS_k \quad (14.9)$$

where $\nabla \bar{h}_{\mu\nu}$ is the gradient of $\bar{h}_{\mu\nu}$, $S$ is the surface surrounding the source volume, and we have applied Gauss theorem to $\nabla \bar{h}_{\mu\nu}$. Using eq. (14.7) the surface integral can be approximated as follows
\[ \int_S (\nabla \tilde{h}_{\mu\nu})^k dS_k \simeq 4\pi \epsilon^2 \left( \frac{d}{dr} \frac{A_{\mu\nu}}{r} e^{i \omega r} \right)_{r=\epsilon} \]

\[ = 4\pi \epsilon^2 \left[ - \frac{A_{\mu\nu}}{r^2} e^{i \omega r} + \frac{A_{\mu\nu}}{r} \left( \frac{i \omega}{c} \right) e^{i \omega r} \right]_{r=\epsilon} ; \]

if we keep the leading term and discard terms of order \( \epsilon \), we find \(^1\)

\[ \int_V \nabla^2 \tilde{h}_{\mu\nu}(\omega, x^i) d^3x \simeq -4\pi A_{\mu\nu}(\omega), \]

and eq. (14.8) becomes

\[ -4\pi A_{\mu\nu} + \int_V \frac{\omega^2}{c^2} \tilde{h}_{\mu\nu}(\omega, x^i) d^3x = -K \int_V T_{\mu\nu}(\omega, x^i) d^3x. \tag{14.10} \]

The second term

\[ \int_V \frac{\omega^2}{c^2} \tilde{h}_{\mu\nu}(\omega, x^i) d^3x \]

satisfies the following inequality

\[ \int_V \frac{\omega^2}{c^2} \tilde{h}_{\mu\nu}(\omega, x^i) d^3x \lesssim |\tilde{h}_{\mu\nu}|_{max} \frac{\omega^2}{c^2} \frac{4}{3}\pi \epsilon^3, \tag{14.11} \]

where \( |\tilde{h}_{\mu\nu}|_{max} \) is the maximum reached by \( \tilde{h}_{\mu\nu} \) in the volume \( V \), and since the right-hand side of eq. (14.11) is of order \( \epsilon^3 \) it can be neglected. Consequently eq. (14.10) becomes

\[ -4\pi A_{\mu\nu}(\omega) = -K \int_V T_{\mu\nu}(\omega, x^i) d^3x \]

i.e.

\[ A_{\mu\nu}(\omega) = \frac{4G}{c^4} \int_V T_{\mu\nu}(\omega, x^i) d^3x. \tag{14.12} \]

Thus, the solution of the wave equation inside the source gives the wave amplitude \( A_{\mu\nu}(\omega) \) as an integral of the stress-energy tensor of the source over the source volume. Knowing \( A_{\mu\nu}(\omega) \) we finally find

\[ \tilde{h}_{\mu\nu}(\omega, r) = \frac{4G}{c^4} \frac{e^{i \omega r}}{r} \int_V T_{\mu\nu}(\omega, x^i) d^3x, \tag{14.13} \]

or, by the inverse Fourier transform

\[ \tilde{h}_{\mu\nu}(t, r) = \frac{4G}{c^4} \frac{1}{r} \int_V T_{\mu\nu}(t - \frac{r}{c}, x^i) d^3x. \tag{14.14} \]

This is the gravitational signal emitted by the source.

The integral in (14.14) can be further simplified, but in the meantime note that:

\(^1\)It should be noted that \( e^{i \omega r} \sim 1 \) since we have assumed that \( \lambda_{GW} >> \epsilon \).
1) The solution (14.14) for $\bar{h}_{\mu\nu}$ automatically satisfies the second eq. (13.25), i.e. the harmonic gauge condition

$$\frac{\partial}{\partial x^\mu}\bar{h}^\mu_{\nu} = 0.$$ 

To prove this, we first notice that the solution (14.14) is equivalent to the expression (13.27)

$$\bar{h}_{\mu\nu}(t, \mathbf{x}) = \frac{4G}{c^5} \int_V \frac{T_{\mu\nu}(t - \frac{|\mathbf{x}-\mathbf{x}'|}{c}, \mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|} \, d^3x', \quad (14.15)$$

indeed, since

$$|\mathbf{x}'| < \epsilon, \quad \text{and} \quad r \gg \epsilon, \quad (14.16)$$

then

$$r \equiv |\mathbf{x}| \simeq |\mathbf{x}-\mathbf{x}'|. \quad (14.17)$$

By defining the following function

$$g(\mathbf{x}-\mathbf{x}') \equiv \frac{4G}{c^5} \frac{1}{|\mathbf{x}-\mathbf{x}'|} \delta\left[t' - \left(t - \frac{|\mathbf{x}-\mathbf{x}'|}{c}\right)\right], \quad (14.18)$$

where $\mathbf{x} = (ct, \mathbf{x})$ and $\mathbf{x}' = (ct', \mathbf{x}')$, eq. (14.15) can be written as a four-dimensional integral as follows

$$\bar{h}_{\mu\nu}(\mathbf{x}) = \int_\Omega T_{\mu\nu}(\mathbf{x}') \, g(\mathbf{x}-\mathbf{x}') \, d^4x', \quad (14.19)$$

where $\Omega \equiv V \times I$, and $I$ is the time interval to be taken such that $g(\mathbf{x}-\mathbf{x'})$ vanishes at the extrema of $I$; this happens if $I$ is so large that, for all $\mathbf{x}' \in V$, the expression $t - \frac{|\mathbf{x}-\mathbf{x}'|}{c}$ is inside $I$; indeed, from the definition (14.18) $g$ is different from zero only for $t' = t - \frac{|\mathbf{x}-\mathbf{x}'|}{c}$.

Since $g$ is a function of the difference $(\mathbf{x}-\mathbf{x}')$, then

$$\frac{\partial}{\partial x^\mu} [g(\mathbf{x}-\mathbf{x}')] = -\frac{\partial}{\partial x^\mu} [g(\mathbf{x}-\mathbf{x}')]. \quad (14.20)$$

Consequently,

$$\frac{\partial}{\partial x^\mu} \bar{h}^{\mu\nu}(\mathbf{x}) = \int_\Omega T^{\mu\nu}(\mathbf{x}') \frac{\partial}{\partial x^\mu} g(\mathbf{x}-\mathbf{x}') \, d^4x' = -\int_\Omega T^{\mu\nu}(\mathbf{x}') \frac{\partial}{\partial x^\mu} g(\mathbf{x}-\mathbf{x}') \, d^4x'. \quad (14.21)$$

Integrating by parts, and using the fact that $T^{\mu\nu} = 0$ on the boundary of $V$ and $g = 0$ on the boundary of $I$, we have

$$\frac{\partial}{\partial x^\mu} \bar{h}^{\mu\nu}(\mathbf{x}) = \int_\Omega g(\mathbf{x}-\mathbf{x}') \frac{\partial}{\partial x^\mu} T^{\mu\nu}(\mathbf{x}') \, d^4x' = 0 \quad (14.22)$$

because the stress-energy tensor satisfies the conservation law $T^{\mu\nu,\nu} = 0$. Q.E.D.

2) In order to extract the physical components of the wave we still have to project $\bar{h}_{\mu\nu}$ on the TT-gauge.

3) Eq. (14.14) has been derived on two very strong assumptions: weak field ($g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$) and slow motion ($v_{\text{typical}} << c$). For this reason that expression has to be considered as an estimate of the emitted radiation by the system, unless the two conditions are really satisfied.
14.1 The Tensor Virial Theorem

In order to simplify the integral in eq. (14.14) we shall use the conservation law that \( T_{\mu \nu} \) satisfies (see chapter 7)

\[
\frac{\partial T_{\mu \nu}}{\partial x^\nu} = 0, \quad \rightarrow \quad \frac{1}{c} \frac{\partial T_{\mu 0}}{\partial t} = - \frac{\partial T_{\mu k}}{\partial x^k}, \quad \mu = 0..3, \quad k = 1..3. \tag{14.23}
\]

Let us integrate this equation over the source volume, assuming the index \( \mu \) is fixed

\[
\frac{1}{c} \frac{\partial}{\partial t} \int_V T_{\mu 0} d^3x = - \int_V \frac{\partial T_{\mu k}}{\partial x^k} d^3x.
\]

By Gauss’ theorem, the integral over the volume is equal to the flux of \( T_{\mu k} \) across the surface \( S \) enclosing that volume, thus the right-hand-side becomes

\[
\int_V \frac{\partial T_{\mu k}}{\partial x^k} d^3x = \int_S T_{\mu k} dS_k.
\]

By definition, on \( S \) \( T_{\mu \nu} = 0 \) and consequently the surface integral vanishes; thus

\[
\frac{1}{c} \frac{\partial}{\partial t} \int_V T_{\mu 0} d^3x = 0, \quad \rightarrow \quad \int_V T_{\mu 0} d^3x = \text{const.} \tag{14.24}
\]

From eq. (14.14) it follows that

\[ \bar{h}_{\mu 0} = \text{const}, \quad \mu = 0..3, \]

and since we are interested in the time-dependent part of the field we shall put

\[ \bar{h}_{\mu 0} = 0, \quad \mu = 0..3. \tag{14.25} \]

(Indeed, in the TT-gauge \( \bar{h}_{\mu 0} = 0 \).) We shall now prove the **Tensor-Virial Theorem** which establishes that

\[
\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \int_V T^{00} x^k x^n d^3x = 2 \int_V T^{kn} d^3x, \quad k, n = 1..3. \tag{14.26}
\]

Let us consider the space-components of the conservation law (14.23)

\[
\frac{\partial T^{n0}}{\partial x^0} = - \frac{\partial T^{ni}}{\partial x^i}, \quad i, n = 1..3;
\]

multiply both members by \( x^k \) and integrate over the source volume

\[
\frac{1}{c} \frac{\partial}{\partial t} \int_V T^{n0} x^k d^3x = - \int_V \frac{\partial T^{ni}}{\partial x^i} x^k d^3x
\]

\[ = - \left[ \int_V \frac{\partial (T^{ni} x^k)}{\partial x^i} d^3x - \int_V T^{ni} \frac{\partial x^k}{\partial x^i} d^3x \right] \]

\[ = - \int_S (T^{ni} x^k) dS_i + \int_V T^{nk} d^3x, \]
(remember that $\frac{\partial x^k}{\partial x^i} = \delta^k_i$). As before $\int_S \left(T^{ni} x^k\right) dS_i = 0$, therefore

$$\frac{1}{c} \frac{\partial}{\partial t} \int_V T^{n0} x^k d^3x = \int_V T^{nk} d^3x.$$  

Since $T^{nk}$ is symmetric we can rewrite this equation in the following form

$$\frac{1}{2c} \frac{\partial}{\partial t} \int_V \left(T^{n0} x^k + T^{k0} x^n\right) d^3x = \int_V T^{nk} d^3x. \quad (14.27)$$

Let us now consider the 0 component of the conservation law

$$\frac{1}{c} \frac{\partial T^{00}}{\partial t} + \frac{\partial T^{0i}}{\partial x^i} = 0, \quad i = 1..3$$

multiply by $x^k x^n$ and integrate over $V$

$$\frac{1}{c} \frac{\partial}{\partial t} \int_V T^{00} x^k x^n d^3x = - \int_V \frac{\partial T^{0i}}{\partial x^i} x^k x^n d^3x$$

$$= - \left[ \int_V \frac{\partial \left(T^{0i} x^k x^n\right)}{\partial x^i} d^3x - \int_V \left(T^{0i} \frac{\partial x^k}{\partial x^i} x^n + T^{0i} x^k \frac{\partial x^n}{\partial x^i}\right) d^3x \right]$$

$$= - \int_S \left(T^{0i} x^k x^n\right) dS_i + \int_V \left(T^{0k} x^n + T^{0n} x^k\right) d^3x$$

the first integral vanishes and this equation becomes

$$\frac{1}{c} \frac{\partial}{\partial t} \int_V T^{00} x^k x^n d^3x = \int_V \left(T^{0k} x^n + T^{0n} x^k\right) d^3x.$$  

If we now differentiate with respect to $x^0$ we find

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \int_V T^{00} x^k x^n d^3x = \frac{1}{c} \frac{\partial}{\partial t} \int_V \left(T^{0k} x^n + T^{0n} x^k\right) d^3x,$$

and using eq. (14.27) we finally find

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \int_V T^{00} x^k x^n d^3x = 2 \int_V T^{kn} d^3x, \quad k, n = 1, 3. \quad (14.28)$$

The left-hand-side of this equation is the second time derivative of the quadrupole moment tensor of the system

$$q^{kn}(t) = \frac{1}{c^2} \int_V T^{00}(t, x^i) x^k x^n d^3x, \quad k, n = 1, 3, \quad (14.29)$$

which is a function of time only. Thus, in conclusion

$$\int_V T^{kn}(t, x^i) d^3x = \frac{1}{2} \frac{d^2}{dt^2} q^{kn}(t).$$
By using eqs. (14.14) and (14.25) we finally find

\[
\begin{cases}
\bar{h}^{00} = 0, & \mu = 0 \ldots 3 \\
\bar{h}^{ik}(t, r) = \frac{2G}{c^4 r} \cdot \left[ \frac{d^2}{dt^2} q^{ik}(t - \frac{r}{c}) \right].
\end{cases}
\]

(14.30)

This is the gravitational wave emitted by a gravitating system evolving in time. It can be composed of masses or of any form of energy, because mass and energy are both sources of the gravitational field.

**NOTE THAT**

1) \( \frac{G}{c^4} \sim 8 \cdot 10^{-50} \) s\(^2\)/g cm : this is the reason why gravitational waves are extremely weak!!

3) In order to make the physical degrees of freedom explicitly manifest we still have to transform to the TT-gauge

4) These equations have been derived on very strong assumptions: one is that \( T^{\mu\nu,\nu} = 0 \), i.e. that the motion of the bodies is dominated by non-gravitational forces. However, and remarkably, the result (14.30) depends only on the sources motion and not on the forces acting on them.

5) Gravitational radiation has a quadrupolar nature. A system of accelerated charged particles has a time-varying dipole moment

\[ \vec{d}_{EM} = \sum_i q_i \vec{r}_i \]

and it will emit dipole radiation, the flux of which depends on the second time derivative of \( \vec{d}_{EM} \). For an isolated system of masses we can define a gravitational dipole moment

\[ \vec{d}_G = \sum_i m_i \vec{r}_i, \]

which satisfies the conservation law of the total momentum of an isolated system

\[ \frac{d}{dt} \vec{d}_G = \vec{0}. \]

For this reason, gravitational waves do not have a dipole contribution. It should be stressed that for a spherical or axisymmetric, stationary distribution of matter (or energy) the quadrupole moment is a constant, even if the body is rotating. Thus, a spherical or axisymmetric star does not emit gravitational waves; similarly a star which collapses in a perfectly spherically symmetric way has a vanishing \( \frac{d^2 q^{ik}}{dt^2} \) and does not emit gravitational waves. To produce these waves we need a certain degree of asymmetry, as it occurs for instance in the non-radial pulsations of stars, in a non-spherical gravitational collapse, in the coalescence of massive bodies etc.
14.2 How to transform to the TT-gauge

The solution (14.30) describes a spherical wave far from the emitting source. Locally, it looks like a plane wave propagating along the direction of the unit vector orthogonal to the wavefront

\[ n^\alpha = (0, n^i), \quad i = 1, \ldots, 3 \]  

(14.31)

where

\[ n^i = \frac{x^i}{r}. \]  

(14.32)

In order to express this waveform in the TT-gauge we shall make an infinitesimal coordinate transformation \( x^\mu' = x^\mu + \epsilon^\mu \) and choose the vector \( \epsilon^\mu \) which satisfies the wave equation \( \Box_F \epsilon^\mu = 0 \), so that the harmonic gauge condition is preserved, as explained in chapter 14.

The conditions to impose on the perturbed metric are

\[ \tilde{h}_{\alpha\beta} n^\beta = 0, \quad \text{transverse wave condition} \]

\[ \tilde{h}_{\alpha\beta} \delta^{\alpha\beta} = 0, \quad \text{vanishing trace}. \]

It should be mentioned that the transverse-wave condition implies that \( \tilde{h}^{\mu0} = 0, \quad \mu = 0, 3 \) as required in eq. (14.25). Indeed, given the wave-vector \( k^\mu = (k^0, k^0 n^i) \) we know by eq. (13.39) that \( k^\mu \tilde{h}_{\mu\nu} = 0 \), i.e.

\[ k^0 \tilde{h}'_{0\nu} + k^0 n^i \tilde{h}'_{i\nu} = 0. \]

The second term vanishes because of the transverse wave condition, therefore

\[ \tilde{h}'_{0\nu} = 0. \]

We remind here that, as shown in eq. (13.60), in the TT-gauge \( \tilde{h}_{\mu\nu} \) and \( h_{\mu\nu} \) coincide.

To hereafter, we shall work in the 3-dimensional euclidean space with metric \( \delta_{ij} \). Consequently, there will be no difference between covariant and contravariant indices.

We shall now describe a procedure to project the wave in the TT-gauge, which is equivalent to perform the coordinate transformation mentioned above. As a first step, we define the operator which projects a vector onto the plane orthogonal to the direction of \( n \)

\[ P_{jk} \equiv \delta_{jk} - n_j n_k. \]  

(14.33)

Indeed, it is easy to verify that for any vector \( V^j \), \( P_{jk} V^k \) is orthogonal to \( n^j \), i.e. \( (P_{jk} V^k) n^j = 0 \), and that

\[ P^j_k P^k_l V^l = P^j_l V^l. \]  

(14.34)

Note that \( P_{jk} = P_{kj} \), i.e. \( P_{jk} \) is symmetric. The projector is transverse, i.e.

\[ n^j P_{jk} = 0. \]  

(14.35)

Then, we define the transverse–traceless projector:

\[ P_{jkmn} \equiv P_{jm} P_{kn} - \frac{1}{2} P_{jk} P_{mn}. \]  

(14.36)

which “extracts” the transverse-traceless part of a \( \begin{pmatrix} 0 \\ 2 \end{pmatrix} \) tensor. In fact, using the definition (14.36), it is easy to see that it satisfies the following properties
Chapter 14. The Quadrupole Formalism

14.3 Gravitational wave emitted by a harmonic oscillator

Let us consider a harmonic oscillator composed of two equal masses $m$ oscillating at a frequency $\nu = \frac{\omega}{2\pi}$ with amplitude $A$. Be $l_0$ the proper length of the string when the system
is at rest. Assuming that the oscillator moves on the x-axis, the position of the two masses will be
\[
\begin{align*}
  x_1 &= -\frac{1}{2}l_0 - A \cos \omega t \\
  x_2 &= +\frac{1}{2}l_0 + A \cos \omega t
\end{align*}
\]

The 00-component of the stress-energy tensor of the system is
\[
T^{00} = 2 \sum_{n=1}^{2} c p^0 \delta(x - x_n) \delta(y) \delta(z);
\]
and since \( v << c \), \( \gamma \sim 1 \) \( \rightarrow \ p^0 = mc \), it reduces to
\[
T^{00} = mc^2 \sum_{n=1}^{2} \delta(x - x_n) \delta(y) \delta(z);
\]
the \( xx \)-component of the quadrupole moment
\[
q^{xx} = m \left[ \int_V \delta(x - x_1) x^2 dx \delta(y) dy \delta(z) dz + \int_V \delta(x - x_2) x^2 dx \delta(y) dy \delta(z) dz \right]
\]
(14.46)
\[
= m \left[ x_1^2 + x_2^2 \right] = m \left[ \frac{1}{2}l_0^2 + 2A^2 \cos^2 \omega t + 2Al_0 \cos \omega t \right]
\]
\[
= m \left[ \cos^2 + A^2 \cos 2\omega t + 2Al_0 \cos \omega t \right],
\]
where we have used the trigonometric expression \( \cos 2\alpha = 2 \cos^2 \alpha - 1 \).

The \( zz \)-component of the quadrupole moment is
\[
q^{zz} = m \left[ \int_V \delta(x - x_1) dx \delta(y) dy \delta(z) z^2 dz + \int_V \delta(x - x_2) dx \delta(y) dy \delta(z) z^2 dz \right] = 0
\]
because \( \int_V z^2 \delta(z) \, dz = 0 \). Since the motion is confined to the x-axis, all remaining components of \( q_{ij} \) vanish. Let us now compute the reduced quadrupole moment \( Q_{ij} = q_{ij} - \frac{1}{3} \delta_{ij} q^{kk} \); since \( q^{kk} = \eta^{lek} q_{lk} = \eta^{xx} q_{xx} = q_{xx} \) we find
\[
Q_{xx} = q_{xx} - \frac{1}{3} q_{xx} = \frac{2}{3} q_{xx} \\
Q_{yy} = Q_{zz} = -\frac{1}{3} q_{xx} \\
Q_{xy} = 0.
\]

We shall compute, as an example, the wave emerging in the z-direction; in this case \( n = \frac{x}{r} \to (0,0,1) \) and \( P_{jk} = \delta_{jk} - n_j n_k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \).

By applying to \( Q_{ij} \) the transverse-traceless projector \( P_{jkmn} \) constructed from \( P_{jk} \), we find
\[
Q_{TT}^{xx} = \left( P_{xm} P_{xn} - \frac{1}{2} P_{xx} P_{mn} \right) Q_{mn} \\
= \left( P_{xx} P_{xx} - \frac{1}{2} P_{xx}^2 \right) Q_{xx} - \frac{1}{2} P_{xx} P_{yy} Q_{yy} = \frac{1}{2} (Q_{xx} - Q_{yy}),
\]
\[
Q_{TT}^{xy} = \left( P_{xm} P_{yn} - \frac{1}{2} P_{xy} P_{mn} \right) Q_{mn} = P_{xx} P_{yy} Q_{xy} = 0,
\]
\[
Q_{TT}^{zz} = \left( P_{zm} P_{zn} - \frac{1}{2} P_{zz} P_{mn} \right) Q_{mn} = 0.
\]

Using these expressions it is easy to show that eqs. (14.41) become
\[
\begin{cases}
\hat{h}_{TT}^{\mu_0} = 0 \\
\hat{h}_{TT}^{zi} = 0, \\
\hat{h}_{TT}^{xx} = \frac{2G}{c^4 z} \frac{d^2}{dt^2} q_{xx} = 0 \\
\hat{h}_{TT}^{yy} = \frac{G}{c^4 z} \frac{d^2}{dt^2} (Q_{xx} - Q_{yy}),
\end{cases}
\]
and using eqs. (14.47) and (14.46)
\[
\hat{h}_{TT}^{xx} = -\hat{h}_{TT}^{yy} = \frac{2G}{c^4 z} \frac{d^2}{dt^2} q_{xx}(t - \frac{z}{c}) \left( \frac{d^2}{dt^2} q_{xx} \left( t - \frac{z}{c} \right) \right),
\]
\[
= -\frac{2Gm}{c^4 z} \omega^2 \left[ 2A^2 \cos 2\omega (t - \frac{z}{c}) + Al_0 \cos \omega \left( t - \frac{z}{c} \right) \right].
\]

Thus, radiation emitted by the harmonic oscillator along the z-axis is linearly polarized. If, for instance, we consider two masses \( m = 10^3 \) kg, with \( l_0 = 1 \) m, \( A = 10^{-4} \) m, and \( \omega = 10^4 \) rad/s, the term \( [2A^2 \cos 2\omega t] \) is negligible, and the dominant term is at the same frequency of the oscillations:
\[
\hat{h}_{TT}^{xx} \sim -\frac{2Gm}{c^4 z} \omega^2 Al_0 \cos \omega \left( t - \frac{z}{c} \right) \sim 1.6 \cdot 10^{-35} \frac{z}{c}.
\]
which is, as expected, very very small.

It should be noticed that due to the symmetry of the system, the wave emitted along \( y \) will be the same. To find the wave emitted along \( x \), we choose \( \mathbf{n} = (1, 0, 0) \) and use the same procedure: no radiation will be found.

### 14.4 How to compute the energy carried by a gravitational wave

In order to evaluate how much energy is radiated in gravitational waves by an evolving system, we need to define a tensor that properly describes the energy content of the gravitational field. Our effort will not be completely successful, since we will be able to define a quantity which behaves like a tensor only under linear coordinate transformations. However, this pseudo-tensor will be useful for the purpose we have in mind.

#### 14.4.1 The stress-energy pseudotensor of the gravitational field

In Chapter 7 we have shown that the stress-energy tensor of matter satisfies a divergenceless equation

\[
T_{\mu\nu} = 0. \tag{14.51}
\]

If we choose a locally inertial frame (LIF), the covariant derivative reduces to the ordinary derivative and eq. (14.51) becomes

\[
\frac{\partial T_{\mu\nu}}{\partial x^\nu} = 0. \tag{14.52}
\]

We shall now try to find a quantity, \( \eta^{\mu\nu\gamma} \), such that

\[
T_{\mu\nu} = \frac{\partial}{\partial x^\gamma} \eta^{\mu\nu\gamma}; \tag{14.53}
\]

In this way, if we impose that \( \eta^{\mu\nu\gamma} \) is antisymmetric in the indices \( \nu \) and \( \gamma \), the conservation law (14.52) will automatically be satisfied.

The problem now is: can we find the explicit expression of \( \eta^{\mu\nu\gamma} \)?

From Einstein’s equations we know that

\[
T_{\mu\nu} = \frac{c^4}{8\pi G} \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right); \tag{14.54}
\]

since we are in a locally inertial frame, the Riemann tensor, whose generic expression is

\[
R_{\gamma\alpha\beta\delta} = \frac{1}{2} \left[ \frac{\partial^2 g_{\gamma\beta}}{\partial x^\alpha \partial x^\delta} + \frac{\partial^2 g_{\alpha\delta}}{\partial x^\gamma \partial x^\beta} - \frac{\partial^2 g_{\alpha\beta}}{\partial x^\gamma \partial x^\delta} - \frac{\partial^2 g_{\gamma\delta}}{\partial x^\alpha \partial x^\beta} \right] \tag{14.55}
\]

reduces to the term in square brackets since all \( \Gamma^\gamma_{\alpha\beta} \)'s vanish; it follows that in this frame the Ricci tensor becomes

\[
R^\mu_\nu = g^{\mu\alpha} g^{\nu\beta} R_{\alpha\beta} = g^{\mu\alpha} g^{\nu\beta} g^{\gamma\delta} R_{\gamma\alpha\beta\delta} \tag{14.56}
\]

\[
= \frac{1}{2} g^{\mu\alpha} g^{\nu\beta} g^{\gamma\delta} \left( \frac{\partial^2 g_{\gamma\beta}}{\partial x^\alpha \partial x^\delta} + \frac{\partial^2 g_{\alpha\delta}}{\partial x^\gamma \partial x^\beta} - \frac{\partial^2 g_{\alpha\beta}}{\partial x^\gamma \partial x^\delta} - \frac{\partial^2 g_{\gamma\delta}}{\partial x^\alpha \partial x^\beta} \right).
\]
By using this equation, after some cumbersome calculations eq. (14.54) becomes

\[
T_{\mu\nu} = \frac{\partial}{\partial x^\alpha} \left\{ \frac{c^4}{16\pi G} \frac{1}{(-g)} \frac{\partial}{\partial x^\beta} \left[ (-g) \left( g^{\mu\nu} g^{\alpha\beta} - g^{\mu\alpha} g^{\nu\beta} \right) \right] \right\}.
\]  
(14.57)

The term in parentheses is antisymmetric in the indices \( \nu \) and \( \alpha \) and it is the quantity we were looking for:

\[
\eta_{\mu\nu\alpha} = \frac{c^4}{16\pi G} \frac{1}{(-g)} \frac{\partial}{\partial x^\beta} \left[ (-g) \left( g^{\mu\nu} g^{\alpha\beta} - g^{\mu\alpha} g^{\nu\beta} \right) \right] .
\]  
(14.58)

If we now introduce the quantity

\[
\zeta_{\mu\nu\alpha} = (-g) \eta_{\mu\nu\alpha} = \frac{c^4}{16\pi G} \frac{\partial}{\partial x^\beta} \left[ (-g) \left( g^{\mu\nu} g^{\alpha\beta} - g^{\mu\alpha} g^{\nu\beta} \right) \right],
\]  
(14.59)

since we are in a locally inertial frame \( \frac{\partial}{\partial x^\gamma} (-g) = 0 \), therefore we can write eq. (14.57) as

\[
\frac{\partial \zeta_{\mu\nu\alpha}}{\partial x^\alpha} = (-g) T_{\mu\nu}.
\]  
(14.60)

This equation has been derived in a LIF, where all first derivatives of the metric tensor vanish, but in any other frame this will not be true and the difference \( \frac{\partial \zeta_{\mu\nu\alpha}}{\partial x^\alpha} - (-g) T_{\mu\nu} \) will not be zero, but a quantity which we shall call \( (-g)t_{\mu\nu} \) i.e.

\[
(-g)t_{\mu\nu} = \frac{\partial \zeta_{\mu\nu\alpha}}{\partial x^\alpha} - (-g) T_{\mu\nu}.
\]  
(14.61)

\( t_{\mu\nu} \) is symmetric because both \( T_{\mu\nu} \) and \( \frac{\partial \zeta_{\mu\nu\alpha}}{\partial x^\alpha} \) are symmetric in \( \nu \) and \( \alpha \). The explicit expression of \( t_{\mu\nu} \) can be found by substituting in eq. (14.61) the definition of \( \zeta_{\mu\nu\alpha} \) given in eq. (14.59), and \( T_{\mu\nu} \) computed in terms of the Ricci tensor from eq. (14.54) in an arbitrary frame (i.e. starting from the full expression of the Riemann tensor given in eq. 14.55): after some careful manipulation of the equations it is possible to show that

\[
\begin{align*}
T_{\mu\nu} & = \frac{c^4}{16\pi G} \left\{ 2\Gamma^\gamma_{\alpha\beta} \Gamma^\sigma_{\delta\sigma} - \Gamma^\gamma_{\sigma\delta} \Gamma^\sigma_{\alpha\beta} + \Gamma^\delta_{\alpha\delta} \Gamma^\gamma_{\beta\sigma} \right\} \left( g^{\mu\alpha} g^{\nu\beta} - g^{\mu\beta} g^{\nu\alpha} \right) \\
& + 3 \Gamma^\alpha_{\gamma\delta} \left( \Gamma^\mu_{\gamma\alpha} \Gamma^\nu_{\delta\beta} - \Gamma^\mu_{\delta\alpha} \Gamma^\nu_{\beta\gamma} \right) \\
& + 3 \Gamma^\alpha_{\gamma\beta} \left( \Gamma^\mu_{\gamma\alpha} \Gamma^\nu_{\beta\delta} - \Gamma^\mu_{\beta\alpha} \Gamma^\nu_{\gamma\delta} \right) \\
& + \Gamma^\alpha_{\gamma\delta} \left( \Gamma^\mu_{\gamma\delta} \Gamma^\nu_{\alpha\beta} - \Gamma^\mu_{\alpha\delta} \Gamma^\nu_{\gamma\beta} \right) .
\end{align*}
\]

This is the stress-energy pseudotensor of the gravitational field we were looking for. Indeed we can rewrite eq. (14.61), valid in any reference frame, in the following form

\[
(-g) \left( T_{\mu\nu} + t_{\mu\nu} \right) = \frac{\partial \zeta_{\mu\nu\alpha}}{\partial x^\alpha},
\]  
(14.62)

and since \( \zeta_{\mu\nu\alpha} \) is antisymmetric in \( \mu \) and \( \alpha \)

\[
\frac{\partial}{\partial x^\mu} \frac{\partial \zeta_{\mu\nu\alpha}}{\partial x^\alpha} = 0,
\]
and consequently
\[ \frac{\partial}{\partial x^\mu} \left[ (-g)(t^{\mu\nu} + T^{\mu\nu}) \right] = 0. \] (14.63)

This equation expresses a conservation law, because, as explained in chapter 7, it has the form of a vanishing ordinary divergence of the quantity \([(-g)(t^{\mu\nu} + T^{\mu\nu})]\). Since \(t^{\mu\nu}\) when added to the stress-energy tensor of matter (or fields) satisfies a conservation law, and since it vanishes only in a locally inertial frame where gravity is suppressed, we interpret \(t^{\mu\nu}\) as the entity that contains the information on the energy and momentum carried by the gravitational field. Thus eq. (14.63) expresses the conservation law of the total energy and momentum. Unfortunately, \(t^{\mu\nu}\) is not a tensor; indeed it is a combination of the \(\Gamma\)'s that are not tensors.

14.4.2 The energy flux carried by a gravitational wave

Let us consider an emitting source and the associated 3-dimensional coordinate frame \(O(x, y, z)\). Be an observer located at \(P = (x1, y1, z1)\) as shown in figure 14.1. Be \(r = \sqrt{x1^2 + y1^2 + z1^2}\) its distance from the origin. The observer wants to detect the wave coming along the direction identified by the versor \(\mathbf{n} = \frac{\mathbf{r}}{r}\). As a pedagogical tool, let us consider a second frame \(O' (x', y', z')\), with origin coincident with \(O\), and having the \(x'\)-axis aligned with \(\mathbf{n}\). Assuming that the wave traveling along \(x'\) direction is linearly polarized and has only one polarization, the corresponding metric tensor will be
\[
g_{\mu'\nu'} = \begin{pmatrix} (ct) & (x') & (y') & (z') \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & [1 + h_{TT}'(t, x')] & 0 \\ 0 & 0 & 0 & [1 - h_{TT}'(t, x')] \end{pmatrix}. \]

The observer wants to measure the energy which flows per unit time across the unit surface orthogonal to \(x'\), i.e. \(t^{0x'}\), therefore he needs to compute the Christoffel symbols i.e. the derivatives of \(h_{\mu'\nu'}^{TT}\). According to eq. (14.41) the metric perturbation has the form \(h_{\mu'\nu'}^{TT}(t, x') = \frac{\text{const}}{x'} \cdot f(t - \frac{r}{c})\), and since the only derivatives which matter are those with respect to time and \(x'\)
\[
\frac{\partial h_{TT}^{TT}}{\partial t} = \tilde{h}_{TT}^{TT} = \frac{\text{const}}{x'} f, \\
\frac{\partial h_{TT}^{TT}}{\partial x'} = h_{TT}' = \frac{-\text{const}}{x^2} f + \frac{\text{const}}{x'} f' \sim -\frac{1}{c} \frac{\text{const}}{x'} f = -\frac{1}{c} \tilde{h}_{TT},
\]
where we have retained only the dominant \(1/x'\) term. Thus, the non-vanishing Christoffel symbols are:
\[
\Gamma^0_{y'y'} = -\Gamma^0_{z'z'} = \frac{1}{2} \tilde{h}_{TT}^{TT} \quad \Gamma^0_{0y'} = -\Gamma^0_{0z'} = \frac{1}{2} \tilde{h}_{TT}^{TT} \\
\Gamma^x_{y'y'} = -\Gamma^x_{z'z'} = \frac{1}{2c} \tilde{h}_{TT}^{TT} \quad \Gamma^y_{y'z'} = -\Gamma^y_{z'x'} = -\frac{1}{2c} \tilde{h}_{TT}^{TT}. \] (14.64)
By substituting the Christoffel symbols in $t^{\mu\nu}$ we find

$$c t^{0x'} = \frac{dE_{GW}}{dt dS} = \frac{c^3}{16\pi G} \left[ \left( \frac{dh_{TT}^{TT}(t, x')}{dt} \right)^2 \right].$$

If both polarizations are present

$$g_{\mu'\nu'} = \begin{pmatrix} (ct) & (x') & (y') & (z') \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & [1 + h_{TT}^{TT}(t, x')] & h_{TT}^{TT}(t, x') \\ 0 & 0 & h_{TT}^{TT}(t, x') & [1 - h_{TT}^{TT}(t, x')] \end{pmatrix},$$

and

$$c t^{0x'} = \frac{dE_{GW}}{dt dS} = \frac{c^3}{16\pi G} \left[ \left( \frac{dh_{TT}^{TT}(t, x')}{dt} \right)^2 + \left( \frac{dh_{TT}^{TT}(t, x')}{dt} \right)^2 \right]$$

$$= \frac{c^3}{32\pi G} \left[ \sum_{jk} \left( \frac{dh_{TT}^{TT}(t, x')}{dt} \right)^2 \right]. \quad (14.65)$$

This is the energy per unit time which flows across a unit surface orthogonal to the direction $x'$. However, the direction $x'$ is arbitrary; if the observer is located in a different position...
and computes the energy flux he receives, he will find formally the same eq. (14.65) but with $h_{jk}^{TT}$ referred to the TT-gauge associated with the new direction. Therefore, if we consider a generic direction $\mathbf{r} = r\mathbf{n}$

$$t^{0r} = \frac{c^2}{32\pi G} \left[ \sum_{jk} \left( \frac{dh_{jk}^{TT}(t, r)}{dt} \right)^2 \right].$$

(14.66)

In General Relativity the energy of the gravitational field cannot be defined locally, therefore to find the GW-flux we need to average over several wavelengths, i.e.

$$\frac{dE_{GW}}{dtdS} = \left\langle ct^{0r} \right\rangle = \frac{c^3}{32\pi G} \left( \sum_{jk} \frac{dh_{jk}^{TT}(t, r)}{dt} \right)^2.$$  \hspace{1cm} (14.67)

Since

$$\begin{cases} h_{\mu0}^{TT} = 0, \\ h_{ik}^{TT}(t, r) = 2G \frac{\ddot{Q}_{TT}^{jk} (t - \frac{r}{c})}{c^4} \end{cases}$$

by direct substitution we find

$$\frac{dE_{GW}}{dtdS} = \frac{G}{8\pi c^5 r^2} \left\langle \sum_{jk} \left( \frac{\ddot{Q}_{TT}^{jk} (t - \frac{r}{c})}{c^4} \right)^2 \right\rangle.$$  \hspace{1cm} (14.68)

If we now want to make explicit the dependence of the flux on the propagation direction $\mathbf{n}$, we can express $Q_{jk}^{TT}$ in terms of the projector $P_{jkmn}$ as in eq. (14.45)

$$\frac{dE_{GW}}{dtdS} = \frac{G}{8\pi c^5 r^2} \left\langle \sum_{jk} \left( P_{jkmn} \frac{\ddot{Q}_{mn} \dot{Q}_{rs} (t - \frac{r}{c})}{c^4} \right)^2 \right\rangle.$$  \hspace{1cm} (14.68)

From this formula we can now compute the gravitational luminosity $L_{GW} = \frac{dE_{GW}}{dt}$, i.e. the gravitational energy emitted by the source per unit time

$$L_{GW} = \int \frac{dE_{GW}}{dtdS} dS = \int \frac{dE_{GW}}{dtdS} r^2 d\Omega$$

$$= \frac{G}{2c^5} \frac{1}{4\pi} \int d\Omega \left\langle \sum_{jk} \left( P_{jkmn} \ddot{Q}_{mn} \left( t - \frac{r}{c} \right) \right)^2 \right\rangle,$$

where $d\Omega = (d\cos \theta)d\phi$ is the solid angle element. This integral can be computed by using the properties of $P_{jkmn}$:

$$\sum_{jk} \left( P_{jkmn} \ddot{Q}_{mn} \right)^2 = \sum_{jk} P_{jkmn} \ddot{Q}_{mn} P_{jkr} \ddot{Q}_{rs} =$$

$$= \left[ \sum_{jk} P_{mnjk} P_{jkr} \right] \ddot{Q}_{mn} \ddot{Q}_{rs} = P_{mnrs} \ddot{Q}_{mn} \ddot{Q}_{rs}$$

$$= \left( \delta_{mr} - n_m n_r \right) \left( \delta_{ns} - n_n n_s \right) - \frac{1}{2} \left( \delta_{mn} - n_m n_n \right) \left( \delta_{rs} - n_r n_s \right) \ddot{Q}_{mn} \ddot{Q}_{rs}.$$

(14.70)

If we expand this expression, and remember that
• \( \delta_{mn} \ddot{Q}_{mn} = \delta_{rs} \ddot{Q}_{rs} = 0 \)
because the trace of \( Q_{ij} \) vanishes by definition, and

• \( n_m n_r \delta_{ns} \ddot{Q}_{mn} \ddot{Q}_{rs} = n_n n_s \delta_{mr} \ddot{Q}_{mn} \ddot{Q}_{rs} \)
because \( Q_{ij} \) is symmetric,

at the end we find

\[
\sum_{jk} \left( P_{krmn} Q_{mn} \right)^2 = \ddot{Q}_{rm} \ddot{Q}_{rn} - 2n_m \ddot{Q}_{ms} \ddot{Q}_{nr} n_r + \frac{1}{2} n_m n_n n_r n_s \ddot{Q}_{mn} \ddot{Q}_{rs}.
\]  
(14.71)

This expression has to be substituted in eq. (14.70), and the integrals to be performed over the solid angle are:

\[
\frac{1}{4\pi} \int d\Omega n_i n_j, \quad \text{and} \quad \frac{1}{4\pi} \int d\Omega n_i n_j n_r n_s.
\]  
(14.72)

Let us compute the first.
In polar coordinates, the versor \( n \) can be written as

\[
n_i = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta).
\]  
(14.73)

Thus, for parity reasons

\[
\frac{1}{4\pi} \int d\Omega n_i n_j = 0 \quad \text{when} \quad i \neq j.
\]  
(14.74)

Furthermore, since there is no preferred direction in the integration (isotropy), it must be

\[
\int d\Omega n_i^2 = \int d\Omega n_j^2 = \int d\Omega n_3^2 \rightarrow \frac{1}{4\pi} \int d\Omega n_i n_j = \text{const} \cdot \delta_{ij}.
\]  
(14.75)

For instance,

\[
\frac{1}{4\pi} \int d\Omega (n_3)^2 = \frac{1}{4\pi} \int d\cos \theta d\phi \cos^2 \theta = \frac{1}{4\pi} \int_0^{2\pi} d\phi \int_{-1}^1 d\cos \theta \cos^2 \theta = \frac{1}{3},
\]  
(14.76)

and consequently

\[
\frac{1}{4\pi} \int d\Omega n_i n_j = \frac{1}{3} \delta_{ij}.
\]  
(14.77)

The second integral in (14.72) can be computed in a similar way and gives

\[
\frac{1}{4\pi} \int d\Omega n_i n_j n_r n_s = \frac{1}{15} (\delta_{ij} \delta_{rs} + \delta_{ir} \delta_{js} + \delta_{is} \delta_{jr}).
\]  
(14.78)

Consequently

\[
\frac{1}{4\pi} \int d\Omega \left( \ddot{Q}_{rm} \ddot{Q}_{rn} - 2n_m \ddot{Q}_{ms} \ddot{Q}_{nr} n_r + \frac{1}{2} n_m n_n n_r n_s \ddot{Q}_{mn} \ddot{Q}_{rs} \right) = \frac{2}{5} \ddot{Q}_{rm} \ddot{Q}_{rn}
\]  
(14.79)

and, finally, the emitted power is

\[
L_{GW} = \frac{G}{5e^5} \left\langle \sum_{k,n=1}^3 \ddot{Q}_{kn} \left( t - \frac{r}{c} \right) \ddot{Q}_{kn} \left( t - \frac{r}{c} \right) \right\rangle.
\]  
(14.80)
14.5 Gravitational radiation from a rotating star

In this section we shall show that a rotating star emits gravitational waves only if its shape deviates from axial symmetry.

Consider an ellipsoid of uniform density $\rho$. Its quadrupole moment is

$$q_{ij} = \int_V \rho \, x_i x_j \, dx^3, \quad i = 1, 3$$

and it is related to the inertia tensor

$$I_{ij} = \int_V \rho \left( r^2 \delta_{ij} - x_i x_j \right) \, dx^3$$

by the equation

$$q_{ij} = -I_{ij} + \delta_{ij} \text{Tr} \, q,$$

where $\text{Tr} \, q \equiv q_{m \ m}$. Consequently, the reduced quadrupole moment can be written as

$$Q_{ij} = q_{ij} - \frac{1}{3} \delta_{ij} \text{Tr} \, q = - \left( I_{ij} - \frac{1}{3} \delta_{ij} \text{Tr} \, I \right).$$

Let us first consider a non rotating ellipsoid, with semiaxes $a, b, c$, volume $V = \frac{4}{3} \pi abc$, and equation:

$$\left( \frac{x_1}{a} \right)^2 + \left( \frac{x_2}{b} \right)^2 + \left( \frac{x_3}{c} \right)^2 = 1.$$ 

The inertia tensor is

$$I_{ij} = \frac{M}{5} \begin{pmatrix} b^2 + c^2 & 0 & 0 \\ 0 & c^2 + a^2 & 0 \\ 0 & 0 & a^2 + b^2 \end{pmatrix} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix},$$

where $I_1, I_2, I_3$ are the principal moments of inertia.

Let us now consider an ellipsoid which rotates around one of its principal axes, for instance $I_3$, with angular velocity $(0, 0, \Omega)$. What is its inertia tensor in this case?

Be $\{x_i\}$ the coordinates of the inertial frame, and $\{x'_i\}$ the coordinates of a co-rotating frame. Then,

$$x_i = R_{ij} x'_j,$$

where $R_{ij}$ is the rotation matrix

$$R_{ij} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{with} \quad \varphi = \Omega t.$$ 

For instance, a point at rest in the co-rotating frame, with coordinates $x'_1 = (1, 0, 0)$, has, in the inertial frame, coordinates $x_i = (\cos \Omega t, \sin \Omega t, 0)$, i.e. it rotates in the $x-y$ plane with angular velocity $\Omega$. 

Since in the co-rotating frame \( \{ x'_i \} \)

\[
I'_{ij} = \begin{pmatrix}
I_1 & 0 & 0 \\
0 & I_2 & 0 \\
0 & 0 & I_3
\end{pmatrix},
\]

in the inertial frame \( \{ x_i \} \) it will be

\[
I_{ij} = R_{ik} R_{jl} I'_{kl} = (R I' R^T)_{ij}
\]

\[
= \begin{pmatrix}
I_1 \cos^2 \varphi + I_2 \sin^2 \varphi & - \sin \varphi \cos \varphi (I_2 - I_1) & 0 \\
- \sin \varphi \cos \varphi (I_2 - I_1) & I_1 \sin^2 \varphi + I_2 \cos^2 \varphi & 0 \\
0 & 0 & I_3
\end{pmatrix}.
\]

It is easy to check that \( \text{Tr} I = I_1 + I_2 + I_3 = \text{constant} \).
The quadrupole moment therefore is

\[
Q_{ij} = - \left( I_{ij} - \frac{1}{3} \delta_{ij} \text{Tr} I \right) = -I_{ij} + \text{constant}
\]

Using \( \cos 2\varphi = 2 \cos^2 \varphi - 1 \), etc., the quadrupole moment can be written as

\[
Q_{ij} = \frac{I_2 - I_1}{2} \begin{pmatrix}
\cos 2\varphi & \sin 2\varphi & 0 \\
\sin 2\varphi & - \cos 2\varphi & 0 \\
0 & 0 & 0
\end{pmatrix} + \text{constant}
\]

Since

\[
I_1 = \frac{M}{5} (b^2 + c^2), \quad \text{and} \quad I_2 = \frac{M}{5} (c^2 + a^2),
\]
if \( a, b \) are equal, the quadrupole moment is constant and **no gravitational wave is emitted.**

This is a generic result: an axisymmetric object rigidly rotating around its symmetry axis does not radiate gravitational waves.

In realistic cases, \( a \neq b \), and \( I_1 \neq I_2 \); however the difference is expected to be extremely small. It is convenient to express the quadrupole moment of the star in terms of a dimensionless parameter \( \epsilon \), the **oblateness**, which expresses the deviation from axisymmetry

\[
\epsilon \equiv \frac{a - b}{(a + b)/2}.
\]

It is easy to show that

\[
\frac{I_2 - I_1}{I_3} = \epsilon + O(\epsilon^3).
\]

Indeed,

\[
a - b = \frac{1}{2} \epsilon (a + b),
\]

thus

\[
\frac{I_2 - I_1}{I_3} = \frac{a^2 - b^2}{a^2 + b^2} = \frac{(a + b)(a - b)}{a^2 + b^2} = \frac{1}{2} \epsilon (a + b)^2.
\]

On the other hand, from (14.81) we have

\[
(a - b)^2 = O(\epsilon^2) = a^2 + b^2 - 2ab,
\]

therefore

\[
2ab = a^2 + b^2 + O(\epsilon^2)
\]

and

\[
\frac{I_2 - I_1}{I_3} = \frac{1}{2} \epsilon \left( \frac{a^2 + b^2 + 2ab}{a^2 + b^2} \right) = \epsilon + O(\epsilon^3).
\]

Consequently

\[
Q_{ij} = \frac{\epsilon I_3}{2} \begin{pmatrix}
\cos 2\varphi & \sin 2\varphi & 0 \\
\sin 2\varphi & -\cos 2\varphi & 0 \\
0 & 0 & 0
\end{pmatrix} + \text{constant.}
\]

Since \( \varphi = \Omega t \), eq. (14.86) shows that GW are emitted at twice the rotation frequency. From eq. (14.41) and (14.45), the waveform is

\[
h_{jk}^{TT}(t, r) = \frac{2G}{r c^4} P_{jklm} \left( \frac{d^2}{dt^2} Q_{lm}(t - \frac{r}{c}) \right),
\]

i.e., using eq. (14.86),

\[
h_{ij}^{TT} = h_0 \begin{pmatrix}
-\cos 2\varphi_{ret} & -\sin 2\varphi_{ret} & 0 \\
-\sin 2\varphi_{ret} & \cos 2\varphi_{ret} & 0 \\
0 & 0 & 0
\end{pmatrix}, \quad \varphi_{ret} = \Omega \left( t - \frac{r}{c} \right),
\]
CHAPTER 14. THE QUADRUPOLE FORMALISM

where

\[ h_0 = \frac{4G \Omega^2}{c^4 r} I_3 \epsilon = \frac{16\pi^2 G}{c^4 r T^2} I_3 \epsilon , \] (14.88)

where \( T \) is the rotation period; the term in square brackets in eq. (14.87) depends on the direction of the observer relative to the star axes. Eq. (14.87) shows that a triaxial star rotating around a principal axis emits gravitational waves at twice the rotation frequency

\[ \nu_{GW} = 2\nu_{\text{rot}} . \] (14.89)

Fastly rotating neutron stars have rotation period of the order of a few ms; a typical value of a neutron star moment of inertia is \( \sim 10^{38} Kg m^2 \). For a galactic source the distance from Earth is of a few kpc, thus, if we assume an oblateness as small as \( \sim 10^{-6} \) we find

\[ \frac{16\pi^2 G}{c^4 r T^2} I_3 \epsilon = \frac{16\pi^2 G}{c^4} \cdot (1\text{ms})^{-2} \cdot (1\text{Kpc})^{-1} \cdot (10^{38} Kg m^2) \cdot (10^{-6}) = 4.21 \cdot 10^{-24} . \]

This calculation indicates that the wave amplitude can be normalized as follows

\[ h_0 = 4.21 \cdot 10^{-24} \left[ \frac{\text{ms}}{T} \right]^2 \left[ \frac{\text{Kpc}}{r} \right] \left[ \frac{I_3}{10^{38} Kg m^2} \right] \left[ \frac{\epsilon}{10^{-6}} \right] . \] (14.90)

The rotation period and the star distance can be measured; the moment of inertia can be estimated, if we choose an equation of state among those proposed in the literature to model matter in the neutron star interior; conversely, \( \epsilon \) is unknown. However, we shall now show how astronomical observations allow to set an upper limit on this parameter. It is known that the rotation period of observed pulsars increases with time, i.e. the star rotational energy decreases. Pulsars slow down mainly because, having a time varying magnetic dipole moment, they radiate electromagnetic waves. A further braking mechanism is provided by gravitational wave emission. We shall now assume that the pulsar radiates its rotational energy entirely in gravitational waves and, using this very strong assumption and the expression of the gravitational luminosity (14.80) in terms of the source quadrupole moment (14.86), we shall show how to estimate the pulsar oblateness. This estimate will be an upper bound for \( \epsilon \) because we know that only a small fraction of the pulsar energy is dissipated in gravitational waves.

From eq (14.86) we find

\[ \ddot{Q}_{kn} = 4\Omega^3 \epsilon^2 f^2 \begin{pmatrix} \sin 2\phi & -\cos 2\phi & 0 & 0 \\ -\cos 2\phi & \sin 2\phi & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} ; \] (14.91)

by replacing this expression in (14.80) we find

\[ L_{GW} = \frac{32G}{5c^5} \Omega^6 \epsilon^2 f^2 . \] (14.92)
The rotational energy, in the Newtonian approximation, is
\[ E_{\text{rot}} = \frac{1}{2} I \dot{\Omega}^2, \quad (14.93) \]
and its time derivative
\[ \dot{E}_{\text{rot}} = I \Omega \dot{\Omega}. \quad (14.94) \]
Since \( L_{GW} \leq -\dot{E}_{\text{rot}} \) (with equality if the spin-down is entirely due to gravitational emission), then
\[ \frac{32G}{5c^5} (2\pi \nu)^6 \epsilon^2 I^2 \leq I (2\pi)^2 \nu |\dot{\nu}| \quad (14.95) \]
(note that \(|\dot{\nu}| = -\ddot{\nu}|\)), therefore the spin-down limit on \( \epsilon \) gives
\[ \epsilon \leq \epsilon_{sd} = \left( \frac{5c^5 |\dot{\nu}|}{512\pi^4 G \nu^5 I} \right)^{1/2}. \quad (14.96) \]
For instance, in the case of the Crab pulsar, for which \( \nu = 30 \text{ Hz} \) and \( r = 2 \text{ kpc} \), if we assume that the momentum of inertia is \( I = 10^{38} \text{ kg m}^2 \), eq. (14.96) gives
\[ \epsilon_{sd} = 7.5 \cdot 10^{-4}. \quad (14.97) \]
This calculation has been done for a number of known pulsars (A. Giazotto, S. Bonazzola and E. Gourgoulhon, *On gravitational waves emitted by an ensemble of rotating neutron stars*, Phys. Rev. D55, 2014, 1997) and the results are shown in Table 14.1.

**Table 14.1: Upper limits for the oblateness of an ensemble of known pulsars, obtained from spin-down measurements.**

<table>
<thead>
<tr>
<th>name</th>
<th>( \nu_{GW} ) (Hz)</th>
<th>( \epsilon_{sd} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vela</td>
<td>22</td>
<td>1.8 \cdot 10^{-3}</td>
</tr>
<tr>
<td>Crab</td>
<td>60</td>
<td>7.5 \cdot 10^{-4}</td>
</tr>
<tr>
<td>Geminga</td>
<td>8.4</td>
<td>2.3 \cdot 10^{-3}</td>
</tr>
<tr>
<td>PSR B 1509-68</td>
<td>13.2</td>
<td>1.4 \cdot 10^{-2}</td>
</tr>
<tr>
<td>PSR B 1706-44</td>
<td>20</td>
<td>1.9 \cdot 10^{-3}</td>
</tr>
<tr>
<td>PSR B 1957+20</td>
<td>1242</td>
<td>1.6 \cdot 10^{-9}</td>
</tr>
<tr>
<td>PSR J 0437-4715</td>
<td>348</td>
<td>2.9 \cdot 10^{-8}</td>
</tr>
</tbody>
</table>

As we said, these numbers are only upper bounds. Recent studies which take into account the maximum strain that the crust of a neutron star can support without breaking set a further constraint on \( \epsilon \)
\[ \epsilon \lesssim 5 \cdot 10^{-6} \quad (14.98) \]

The data collected during the past few years by the first generation of interferometric gravitational antennas VIRGO and LIGO are being analyzed; although waves have not been
detected yet, available data allow us to set more stringent constraints on the oblateness of some known pulsars. For instance, the present detectors sensitivity would allow to detect the gravitational signal emitted by the Crab pulsar if its amplitude would exceed $h_0 = 2.0 \cdot 10^{-25}$; since no signal has been detected, it means that

$$h_0 < 2.0 \cdot 10^{-25},$$

and using eq. (14.90), this equation implies that the Crab oblateness satisfies the following constraint

$$\epsilon \leq 1.1 \cdot 10^{-4}.$$  

This limit is more restrictive than the spin-down limit (14.97), even though it is larger than the theoretical value arising from the maximal strain sustainable by the crust, (14.98). However, this result is very important, since data analysis from LIGO/VIRGO tells us something which we did not know from astrophysical observation (Abbot et al., Astrophysical Journal, 713, 671, 2010, “Searches for gravitational waves from known pulsars with S5 LIGO data”).

Let us now consider the case in which the star rotates about an axis which forms an angle with one of the principal axes, say, $I_3$. The angle between the two axes is called “wobble angle”. In this case, the angular velocity precedes around $I_3$ (see figure 14.2). For simplicity, let us assume that $I_3$ is a symmetry axis of the ellipsoid, i.e.

$$a = b \quad \rightarrow \quad I_1 = I_2,$$

and that the wobble angle $\theta$ is small, i.e. $\theta \ll 1$. Be $\{x'\}$ the coordinates of the co-rotating frame $O'$ and $\{x\}$ those of the inertial frame $O$. As usual $x_i = R_{ij}x'_j$, where $R_{ij}$ is the rotation matrix.

The transformation from $O'$ to $O$ is the composition of two rotations:

- A rotation of $O'$ around the $x'$ axis by a small angle $\theta$ (constant); the new frame $O''$ has the $z''$ axis coincident with the rotation axis. The corresponding rotation matrix is

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \theta \\ 0 & -\theta & 1 \end{pmatrix} + O(\theta^2).$$

- A time dependent rotation around the $z''$ axis, by an angle $\varphi = \Omega t$; the corresponding rotation matrix is

$$R_z = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$  

After this rotation, the symmetry axis of the ellipsoid precedes around the $z$ axis, with angular velocity $\Omega$. 
Figure 14.2: From \( O' \) to \( O'' \)

Figure 14.3: From \( O'' \) to \( O \)
The rotation matrix from $O'$ to $O$ therefore is
\[
R = R_z R_x = \begin{pmatrix}
  \cos \varphi & -\sin \varphi & 0 \\
  \sin \varphi & \cos \varphi & 0 \\
  0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
  1 & 0 & 0 \\
  0 & 1 & \theta \\
  0 & -\theta & 1
\end{pmatrix} + O(\theta^2)
\]
\[
= \begin{pmatrix}
  \cos \varphi & -\sin \varphi & -\theta \sin \varphi \\
  \sin \varphi & \cos \varphi & \theta \cos \varphi \\
  0 & -\theta & 1
\end{pmatrix} + O(\theta^2).
\]  
(14.103)

Since in the co-rotating frame $O'$
\[
I'_{ij} = \begin{pmatrix}
  I_1 & 0 & 0 \\
  0 & I_1 & 0 \\
  0 & 0 & I_3
\end{pmatrix},
\]
in the inertial frame $O$ it will be
\[
I_{ij} = R_{ik} R_{jl} I'_{kl} = (R I' R^T)_{ij}
\]
\[
= \begin{pmatrix}
  I_1 & 0 & 0 \\
  0 & I_1 & (I_1 - I_3) \theta \sin \varphi \\
  (I_1 - I_3) \theta \sin \varphi & -(I_1 - I_3) \theta \cos \varphi & I_3
\end{pmatrix}
\]  
+ O(\theta^2).
\]  
(14.104)

The quadrupole moment can then be written as
\[
Q_{ij} = -I_{ij} + \text{const.} = (I_1 - I_3) \theta \begin{pmatrix}
  0 & 0 & -\sin \varphi \\
  0 & 0 & \cos \varphi \\
-\sin \varphi & \cos \varphi & 0
\end{pmatrix}
\]  
+ const + O(\theta^2),
\]  
(14.105)

and the wave amplitude therefore is
\[
h^{TT}_{jk}(t, r) = \frac{2G}{r c^4} \mathcal{P}_{jklm} \left[ \frac{d^2}{dt^2} Q^{lm}(t - \frac{r}{c}) \right],
\]
i.e., using eq. (14.106),
\[
h^{TT}_{ij} = h_0 \mathcal{P} \begin{pmatrix}
  0 & 0 & \sin \varphi_{ret} \\
  0 & 0 & -\cos \varphi_{ret} \\
 \sin \varphi_{ret} & -\cos \varphi_{ret} & 0
\end{pmatrix}, \quad \varphi_{ret} = \Omega \left( t - \frac{r}{c} \right),
\]  
(14.107)

where
\[
h_0 = \frac{2G}{c^4 \mathcal{P}} (I_1 - I_3) \theta = \frac{8\pi^2 G}{c^4 r T^2} (I_1 - I_3) \theta.
\]  
(14.108)

From eq. (14.107) we see that when the triaxial star rotates around an axis which does not coincide with a principal axis gravitational waves are emitted at the rotation frequency
\[
\nu_{GW} = \nu_{rot}.
\]

As the oblateness, the wobble angle is an unknown parameter.
14.6 Gravitational wave emitted by a binary system in circular orbit

We shall now estimate the gravitational signal emitted by a binary system composed of two stars moving on a circular orbit around their common center of mass. For simplicity we shall assume that the two stars of mass \( m_1 \) and \( m_2 \) are point masses. Be \( l_0 \) the orbital separation, \( M \) the total mass
\[
M \equiv m_1 + m_2,
\]
and \( \mu \) the reduced mass
\[
\mu \equiv \frac{m_1 m_2}{M}.
\]

Let us consider a coordinate frame with origin coincident with the center of mass of the system as indicated in figure (14.4) and be
\[
l_0 = r_1 + r_2, \quad r_1 = \frac{m_2 l_0}{M}, \quad r_2 = \frac{m_1 l_0}{M}.
\]

The orbital frequency can be found from Kepler’s law
\[
G \frac{m_1 m_2}{l_0^2} = m_1 \omega^2 \frac{m_2 l_0}{M}, \quad G \frac{m_1 m_2}{l_0^2} = m_2 \omega^2 \frac{m_1 l_0}{M},
\]
from which we find
\[
\omega_K = \sqrt{\frac{GM}{l_0^3}}
\]
is the Keplerian frequency. Be \((x_1, x_2)\) and \((y_1, y_2)\) the coordinates of the masses \( m_1 \) and \( m_2 \)

\[
x_1 = \frac{m_2 l_0}{M} \cos \omega_K t \quad x_2 = \frac{m_1 l_0}{M} \cos \omega_K t
\]
\[
y_1 = \frac{m_2 l_0}{M} \sin \omega_K t \quad y_2 = \frac{m_1 l_0}{M} \sin \omega_K t.
\]

Figure 14.4: Two point masses in circular orbit around the common center of mass
The 00-component of the stress-energy tensor of the system is
\[
T^{00} = c^2 \sum_{n=1}^{2} m_n \delta(x - x_n) \delta(y - y_n) \delta(z),
\]
and the non vanishing components of the quadrupole moment are
\[
q_{xx} = m_1 \int_V x^2 \delta(x - x_1) \, dx \delta(y - y_1) \, dy \delta(z) \, dz + m_2 \int_V x^2 \delta(x - x_2) \, dx \delta(y - y_2) \, dy \delta(z) \, dz = m_1 x_1^2 + m_2 x_2^2
\]
\[
= \mu l_0^2 \cos^2 \omega_K t = \frac{\mu}{2} l_0^2 \cos 2\omega_K t + \text{cost},
\]
\[
q_{yy} = m_1 \int_V y^2 \delta(x - x_1) \, dx \delta(y - y_1) \, dy \delta(z) \, dz + m_2 \int_V y^2 \delta(x - x_2) \, dx \delta(y - y_2) \, dy \delta(z) \, dz = m_1 y_1^2 + m_2 y_2^2
\]
\[
= \mu l_0^2 \sin^2 \omega_K t = -\frac{\mu}{2} l_0^2 \cos 2\omega_K t + \text{cost1},
\]
and
\[
q_{xy} = m_1 \int_V x y \delta(x - x_1) \, dx \delta(y - y_1) \, dy \delta(z) \, dz + m_2 \int_V x y \delta(x - x_2) \, dx \delta(y - y_2) \, dy \delta(z) \, dz
\]
\[
= m_1 x_1 y_1 + m_2 x_2 y_2 = \mu l_0^2 \cos \omega t \sin \omega_K t = \frac{\mu}{2} l_0^2 \sin 2\omega_K t.
\]
(we have used \(\cos 2\alpha = 2 \cos^2 \alpha - 1\), \(\sin^2 \alpha = \frac{1}{2} - \frac{1}{2} \cos 2\alpha\) and \(m_1 m_2 = \mu M\)).

In summary
\[
q_{xx} = \frac{\mu}{2} l_0^2 \cos 2\omega_K t + \text{cost}
\]
\[
q_{yy} = -\frac{\mu}{2} l_0^2 \cos 2\omega_K t + \text{cost1}
\]
\[
q_{xy} = \frac{\mu}{2} l_0^2 \sin 2\omega_K t,
\]
and
\[
q_k^k = \eta^{kk} q_{kk} = q_{xx} + q_{yy} = \text{constant}.
\]

The time-varying part of \(Q_{ij} = q_{ij} - \frac{1}{3} \delta_{ij} q_k^k\) therefore is:
\[
Q_{xx} = -Q_{yy} = \frac{\mu}{2} l_0^2 \cos 2\omega_K t
\]
\[
Q_{xy} = \frac{\mu}{2} l_0^2 \sin 2\omega_K t,
\]
and defining a matrix \(A_{ij}\)
\[
A_{ij}(t) = \begin{pmatrix}
\cos 2\omega_K t & \sin 2\omega_K t & 0 \\
\sin 2\omega_K t & -\cos 2\omega_K t & 0 \\
0 & 0 & 0
\end{pmatrix}
\]
we can write
\[ Q_{ij} = \frac{\mu}{2} \frac{l_0^2}{r^2} A_{ij} + \text{const.} \] (14.116)

Since the wave emitted along a generic direction \( \mathbf{n} \) in the \( \mathbf{TT} \)-gauge is
\[ h_{ij}^{TT}(t, r) = \frac{2G}{rc^4} \frac{d^2}{dt^2} \left[ Q_{ij}^{TT}(t - \frac{r}{c}) \right] \quad \text{where} \quad Q_{ij}^{TT}(t - \frac{r}{c}) = \mathcal{P}_{ijkl} Q_{kl}(t - \frac{r}{c}) \]
using eq. (14.112) we find
\[ h_{ij}^{TT} = -\frac{2G}{rc^4} \frac{\mu}{2} \frac{l_0^2}{r^2} (2\omega_K)^2 \left[ \mathcal{P}_{ijkl} A_{kl} \right] = -\frac{4\mu M G^2}{r l_0 c^4} \left[ \mathcal{P}_{ijkl} A_{kl} \right]. \] (14.117)

By defining a wave amplitude
\[ h_0 = \frac{4\mu M G^2}{r l_0 c^4} \] (14.118)
we can finally write the emitted wave as
\[ h_{ij}^{TT}(t, r) = -h_0 A_{ij}^{TT}(t - \frac{r}{c}), \] (14.119)
where
\[ A_{ij}^{TT}(t - \frac{r}{c}) = \left[ \mathcal{P}_{ijkl} A_{kl} \right] \] (14.120)
depends on the orientation of the line of sight with respect to the orbital plane.

From these equations we see that the radiation is emitted at twice the orbital frequency.

For example, if \( \mathbf{n} = \mathbf{z} \), \( P_{ij} = \text{diag}(1, 1, 0) \)
\[ A_{ij}^{TT}(t) = \begin{pmatrix} \cos 2\omega_K t & \sin 2\omega_K t & 0 \\ \sin 2\omega_K t & -\cos 2\omega_K t & 0 \\ 0 & 0 & 0 \end{pmatrix} \] (14.121)
and
\[ h_{xx}^{TT} = -h_{yy}^{TT} = -h_0 \cos 2\omega_K \left( t - \frac{z}{c} \right) \] (14.122)
\[ h_{xy}^{TT} = -h_0 \sin 2\omega_K \left( t - \frac{z}{c} \right). \]

In this case the wave has both polarizations, and since \( h_{xx}^{TT} = h_0 \Re \{ e^{i\omega(t-\frac{z}{c})} \} \) and \( h_{xy}^{TT} = h_0 \Im \{ e^{i\omega(t-\frac{z}{c})} \} \), the wave is circularly polarized.

If \( \mathbf{n} = \mathbf{x} \), \( P_{ij} = \text{diag}(0, 1, 1) \)
\[ A_{ij}^{TT} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\frac{1}{2} \cos 2\omega_K t & 0 \\ 0 & 0 & \frac{1}{2} \cos 2\omega_K t \end{pmatrix} \] (14.123)
and
\[ h_{yy}^{TT} = -h_{zz}^{TT} = \pm \frac{1}{2} h_0 \cos 2\omega_K \left( t - \frac{x}{c} \right), \] (14.124)
i.e. the wave is a linearly polarized wave.

If \( \mathbf{n} = \mathbf{y} \), \( P_{ij} = \text{diag}(1, 0, 1) \) and

\[
A_{ij}^{TT} = \begin{pmatrix}
\frac{1}{2} \cos 2\omega_K t & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & -\frac{1}{2} \cos 2\omega_K t
\end{pmatrix}
\]  

and again the wave is linearly polarized

\[
h_{TT}^{xx} = -h_{TT}^{zz} = -\frac{1}{2} h_0 \cos 2\omega_K (t - \frac{y}{c}).
\]

Eqs. (14.118) can be used to estimate the amplitude of the gravitational signal emitted by the binary system PSR 1913+16 discovered in 1975, (R.A. Hulse and J.H. Taylor, *Discovery Of A Pulsar In A Binary System*, Astrophys. J. 195, L51, 1975) which consists of two neutron stars orbiting at a very short distance from each other. The data we know from observations are:

\[
m_1 \sim m_2 \sim 1.4 M_\odot, \quad l_0 = 0.19 \cdot 10^{12} \text{ cm}
\]

\[
T = 7h 45m 7s, \quad \nu_K = \frac{\omega_K}{2\pi} \sim 3.58 \cdot 10^{-5} \text{ Hz}
\]

where \( T \) is the orbital period. Note that the two stars have nearly equal masses: they are comparable to that of the Sun, and their orbital separation is about twice the radius of the Sun! The orbit is eccentric with eccentricity \( \epsilon \simeq 0.617 \), however we shall assume it is circular and apply eqs. (14.118). For this system the emission frequency is

\[
\nu_{GW} = 2\nu_K \sim 7.16 \cdot 10^{-5} \text{ Hz},
\]
therefore the wavelength of the emitted radiation is

\[ \lambda_{GW} = \frac{c}{\nu_{GW}} \sim 10^{14} \text{cm} \quad \text{i.e.} \quad \lambda_{GW} >> l_0. \]  

(14.129)

Thus, the slow-motion approximation, on which the quadrupole formalism is based, is certainly satisfied in this case even though the two neutron stars are orbiting at such small distance from each other. The distance of the system from Earth is \( r = 5 \text{ kpc} \) and since

\[ 1 \text{ pc} = 3.08 \cdot 10^{18} \text{ cm}, \quad \rightarrow \quad r = 1.5 \cdot 10^{22} \text{ cm}. \]

The wave amplitude is

\[ h_0 = \frac{4 \mu M G^2}{r l_0 c^4} \sim 5 \cdot 10^{-23}. \]

A new binary pulsar has recently been discovered (M. Burgay et al., *An increased estimate of the merger rate of double neutron stars from observations of a highly relativistic system* Nature 426, 531, 2003) which has an even shorter orbital period and it is closer than PSR 1913+16: it is the double pulsar PSR J0737-3039, whose orbital parameters are

\[ m_1 = 1.337 M_\odot, \quad m_2 \sim 1.250 M_\odot \]
\[ T = 2.4 h, \quad e = 0.08 \]
\[ r = 500 \text{ pc} \quad l_0 \sim 1.2 R_\odot. \]

In this case the orbit is nearly circular,

\[ \mu = \frac{m_1 m_2}{m_1 + m_2} = 0.646 M_\odot \quad \rightarrow \quad h_0 = \frac{4\mu MG^2}{r l_0 c^4} \sim 1.1 \cdot 10^{-21}, \]

and waves are emitted at the frequency

\[ \nu_{GW} = 2 \nu_K = 2.3 \cdot 10^{-4} \text{ Hz}. \]

In this section we have considered only circular orbits; the calculations can be generalized to the case of eccentric or open orbits by replacing the equation of motion of the two masses (14.113) by those appropriate to the chosen orbit. By this procedure it is possible to show that when the orbits are ellipses, gravitational waves are emitted at frequencies multiple of the orbital frequency \( \nu_K \), and that the number of equally spaced spectral lines increases with the eccentricity.

### 14.7 Evolution of a binary system due to the emission of gravitational waves

In this section we shall show that the orbital period \( T \) of a binary system decreases in time due to gravitational wave emission. \( \dot{T} \equiv dT/dt \) has indeed been measured for PSR 1913+16, and the observations are in very good agreement with the predictions of General Relativity,
providing a first indirect proof of the existence of gravitational waves. The orbital evolution driven by gravitational wave emission quietly proceeds bringing the stars closer. As their distance decreases, the process becomes faster and the two stars spiral toward their common center of mass until they coalesce.

We shall now describe how the binary system evolves, explicitly computing $\dot{T}$ and the emitted signal up to the point when the quadrupole approximation is violated and the theory developed in this chapter can no longer be applied.

We shall start by computing the gravitational luminosity defined in eq. (14.114); using the reduced quadrupole moment of a binary system given by eqs. (14.115) and (14.116) we find

$$\sum_{k,n=1}^{3} \ddot{Q}_{kn} \dot{Q}_{kn} = 32 \mu^2 \frac{t_0^4}{l_0^4} \omega_K^3 = 32 \mu^2 G^3 \frac{M^3}{l_0^5},$$

and by direct substitution in eq. (14.80)

$$L_{GW} \equiv \frac{dE_{GW}}{dt} = \frac{32}{5} \frac{G^4}{c^5} \frac{\mu^2 M^3}{l_0^5}. \quad (14.130)$$

This expression has to be considered as an average over several wavelengths (or equivalently, over a sufficiently large number of periods), as stated in eq. (14.80); therefore, in order $L_{GW}$ to be defined, we must be in a regime where the orbital parameters do not change significantly over the time interval taken to perform the average. This assumption is called adiabatic approximation, and it is certainly applicable to systems like PSR 1913+16 or PSR J0737-3039 that are very far from coalescence. In the adiabatic regime, the system has the time to adjust the orbit to compensate the energy lost in gravitational waves with a change in the orbital energy, in such a way that

$$\frac{dE_{orb}}{dt} + L_{GW} = 0. \quad (14.131)$$

Let us see what are the consequences of this equation. The orbital energy is

$$E_{orb} = E_K + U$$

where the kinetic and the gravitational energy are, respectively,

$$E_K = \frac{1}{2} m_1 \omega_K^2 r_1^2 + \frac{1}{2} m_2 \omega_K^2 r_2^2$$

$$= \frac{1}{2} \omega_K^2 \left[ \frac{m_1 m_2 l_0^2}{M^2} + \frac{m_2 m_1 l_0^2}{M^2} \right]$$

$$= \frac{1}{2} \omega_K^2 \mu_0^2 = \frac{G \mu M}{2 l_0}$$

and

$$U = -\frac{G m_1 m_2}{l_0} = -\frac{G \mu M}{l_0}.$$

Therefore

$$E_{orb} = -\frac{1}{2} \frac{G \mu M}{l_0} \quad (14.132)$$
CHAPTER 14. THE QUADRUPOLE FORMALISM

and its time derivative is
\[
\frac{dE_{\text{orb}}}{dt} = \frac{1}{2} G \mu M \left( \frac{1}{l_0} \frac{dl_0}{dt} \right) = -E_{\text{orb}} \left( \frac{1}{l_0} \frac{dl_0}{dt} \right). \tag{14.133}
\]

The term \( \frac{dl_0}{dt} \) can be expressed in terms of the time derivative of \( \omega_K \) as follows
\[
\frac{\omega_K^2}{l_0} \rightarrow 2 \ln \omega_K = -3 \ln l_0 \rightarrow \frac{1}{\omega_K} \frac{d\omega_K}{dt} = -\frac{3}{2} \frac{1}{l_0} \frac{dl_0}{dt},
\]
and eq. (14.133) becomes
\[
\frac{dE_{\text{orb}}}{dt} = \frac{2}{3} E_{\text{orb}} \frac{d\omega_K}{dt}. \tag{14.134}
\]

Since \( \omega_K = 2\pi P^{-1} \)
\[
\frac{1}{\omega_K} \frac{d\omega_K}{dt} = \frac{1}{P} \frac{dP}{dt},
\]
and eq. (14.134) gives
\[
\frac{dE_{\text{orb}}}{dt} = -\frac{2}{3} \frac{E_{\text{orb}}}{P} \frac{dP}{dt} \rightarrow \frac{dT}{dt} = -\frac{3}{2} \frac{P}{E_{\text{orb}}} \frac{dE_{\text{orb}}}{dt}. \tag{14.135}
\]

Since by eq. (14.131) \( \frac{dE_{\text{orb}}}{dt} = -L_{\text{GW}} \) we finally find how the orbital period changes due to the emission of gravitational waves
\[
\frac{dP}{dt} = \frac{3}{2} \frac{P}{E_{\text{orb}}} L_{\text{GW}}. \tag{14.136}
\]

For example if we consider PSR 1913+16, assuming the orbit is circular we find
\[
P = 27907 \text{ s}, \quad E_{\text{orb}} \sim -1.4 \cdot 10^{48} \text{ erg}, \quad L_{\text{GW}} \sim 0.7 \cdot 10^{31} \text{ erg/s} \tag{14.137}
\]
and
\[
\frac{dP}{dt} \sim -2.2 \cdot 10^{-13}.
\]

As mentioned in section 14.6, the orbit of the real system has a quite strong eccentricity \( \epsilon \approx 0.617 \). By doing the calculation using the equations of motion appropriate for an eccentric orbit we would find
\[
\frac{dP}{dt} = -2.4 \cdot 10^{-12}.
\]

PSR 1913+16 has now been monitored for decades and the rate of variation of the period, measured with very high accuracy, is
\[
\frac{dP}{dt} = -(2.4184 \pm 0.0009) \cdot 10^{-12}.
\]

(J. M. Wisberg, J.H. Taylor *Relativistic Binary Pulsar B1913+16: Thirty Years of Observations and Analysis*, in Binary Radio Pulsars, ASP Conference series, 2004, eds. F.A.A.Rasio, I.H.Stairs). Thus, the prediction of General Relativity are confirmed by observations. This result provided the first indirect evidence of the existence of gravitational waves and for this
discovery Hulse and Taylor have been awarded of the Nobel prize in 1993.

For the recently discovered double pulsar PSR J0737-3039

\[ P = 8640 \text{ s}, \quad E_{\text{orb}} \sim -2.55 \times 10^{48} \text{ erg}, \quad L_{\text{GW}} \sim 2.24 \times 10^{32} \text{ erg/s} \]  

(14.138)

and

\[ \frac{dP}{dt} \sim -1.2 \times 10^{-12}, \]

which is also in agreement with observations.

Knowing the energy lost by the system, we can also evaluate how the orbital separation \( l_0 \) changes in time. From eq. (14.133) and remembering that \( \frac{dE_{\text{orb}}}{dt} = -L_{\text{GW}} \) we find

\[ \frac{1}{l_0} \frac{dl_0}{dt} = \frac{L_{\text{GW}}}{E_{\text{orb}}} = -\left[ \frac{64}{5} \frac{G^3}{c^5} \mu M^2 \right] \cdot \frac{1}{l_0^3} \]  

(14.139)

Assuming that at some initial time \( t = 0 \) the orbital separation is \( l_0(t = 0) = l_{0i} \) by integrating eq. (14.139) we easily find

\[ l_0^4(t) = (l_{0i})^4 - \frac{256}{5} \frac{G^3}{c^5} \mu M^2 t. \]  

(14.140)

If we define

\[ t_{\text{coal}} = \frac{5}{256} \frac{c^5}{G^3} (l_{0i})^4, \]  

(14.141)

eq. (14.140) can be written as

\[ l_0(t) = l_{0i} \left[ 1 - \frac{t}{t_{\text{coal}}} \right]^{1/4}. \]  

(14.142)

From this equation we see that when \( t = t_{\text{coal}} \) the orbital separation becomes zero, and this is possible because we have assumed that the bodies composing the binary system are pointlike. Of course, stars and black holes have finite sizes, therefore they start merging and coalesce before \( t = t_{\text{coal}} \) is reached. In addition, when the two stars are close enough both the slow motion approximation and the weak field assumption on which the quadrupole formalism relies fails to hold and strong field effects have to be considered; however, the value of \( t_{\text{coal}} \) gives an indication of the time the system needs to merge starting from a given initial distance \( l_{0i} \).

### 14.7.1 The emitted waveform

Since in the adiabatic regime the orbit evolves through a sequence of stationary circular orbits, using eq. (14.142) we can compute how \( \omega_K \) changes in time

\[ \omega_K(t) = \sqrt{\frac{GM}{l_0^3}} = \omega_K^{in} \left[ 1 - \frac{t}{t_{\text{coal}}} \right]^{-3/8}, \quad \omega_K^{in} = \sqrt{\frac{GM}{l_{0i}^3}}. \]  

(14.143)
Consequently, the frequency of the emitted wave at some time $t$ will be twice the orbital frequency at the same time, i.e.

$$\nu_{GW}(t) = \frac{\omega_K}{\pi} = \nu_{GW}^{\text{in}} \left[ 1 - \frac{t}{t_{\text{coal}}} \right]^{-3/8}, \quad \nu_{GW}^{\text{in}} = \frac{1}{\pi} \sqrt{\frac{GM}{(l_0^{\text{in}})^3}}. \quad (14.144)$$

Similarly, the instantaneous amplitude of the emitted signal can be found from eq. (14.118)

$$h_0(t) = \frac{4\mu MG^2}{r l_0(t) c^4} = \frac{4\mu MG^2}{r c^4} \cdot \frac{\omega_K^{2/3}(t)}{G^{1/3} M^{1/3}} \quad (14.145)$$

$$= \frac{4\pi^{2/3} G^{5/3} M^{5/3}}{c^4} \nu_{GW}^{2/3}(t) \quad (14.146)$$

where

$$\mathcal{M}^{5/3} = \mu M^{2/3} \rightarrow \mathcal{M} = \mu^{3/5} M^{2/5} \quad (14.147)$$

Eqs. (14.144) and (14.147) show that the amplitude and the frequency of the gravitational signal emitted by a coalescing system increases with time. For this reason this peculiar waveform is called \textit{chirp}, like the chirp of a singing bird, and $\mathcal{M}$ is said chirp mass. According to eq. (14.119) the emitted waveform is

$$h_{ij}^{\text{TT}}(t, r) = h_0 \left[ \mathcal{P}_{ijkl} A_{kl}(t - \frac{r}{c}) \right] \quad (14.148)$$

where $A_{kl}$ is given in eq. (14.115). Since $\omega_K$ changes in time as in eq. (14.143), the phase appearing in $A_{kl}$ has to be substituted by an integrated phase

$$\Phi(t) = \int^t 2\omega_K(t)dt = \int^t 2\pi \nu_{GW}(t) \ dt + \Phi_{\text{in}}, \quad \text{where} \quad \Phi_{\text{in}} = \Phi(t = 0)$$

Since

$$\nu_{\text{in}} t_{\text{coal}}^{3/8} = \left( \frac{5}{3} \right)^{3/8} \frac{1}{8\pi} \left( \frac{c^3}{G M} \right)^{5/8}$$

then

$$\nu_{GW}(t) = \frac{1}{8\pi} \left( \frac{c^3}{G M} \right)^{5/8} \left[ \frac{5}{t_{\text{coal}} - t} \right]^{3/8}$$

and the integrated phase becomes

$$\Phi(t) = -2 \left[ \frac{c^3 (t_{\text{coal}} - t)}{5G M} \right]^{5/8} + \Phi_{\text{in}}$$

which shows that if we know the phase we can measure the chirp mass. In conclusion, the signal emitted during the inspiralling will be

$$h_{ij}^{\text{TT}} = -\frac{4\pi^{2/3} G^{5/3} M^{5/3}}{c^4} \nu_{GW}^{2/3}(t) \mathcal{P}_{ijkl} A_{kl}(t - \frac{r}{c})$$

where

$$A_{ij}(t) = \begin{pmatrix} \cos \Phi(t) & \sin \Phi(t) & 0 \\ \sin \Phi(t) & -\cos \Phi(t) & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
Chapter 15

Variational Principle and Einstein’s equations

15.1 An useful formula

There exists an useful equation relating $g_{\mu\nu}$, $g^{\mu\nu}$ and $g = \det(g_{\mu\nu})$:

$$\frac{\partial g}{\partial x^\alpha} = gg^{\mu\nu}\frac{\partial g_{\mu\nu}}{\partial x^\alpha}. \quad (15.1)$$

The proof is the following. The determinant $g$ is a polynomial in $g_{\mu\nu}$

$$g = g(g_{\mu\nu}). \quad (15.2)$$

Let us first compute the derivatives

$$\frac{\partial g}{\partial g_{\mu\nu}}. \quad (15.3)$$

We remind that $g$ can be expressed by fixing one row, i.e. $\mu$, and computing the expansion

$$g = \sum_\nu g_{\mu\nu}M_{\mu\nu}(-1)^{\mu+\nu} \quad \text{(no sum over $\mu$)} \quad (15.4)$$

where $M_{\mu\nu}$ is the minor $\mu, \nu$, i.e. the determinant of the matrix obtained by cutting the row $\mu$ and the column $\nu$ from the matrix $g_{\mu\nu}$.

From (15.4) we find that

$$\frac{\partial g}{\partial g_{\mu\nu}} = (-1)^{\mu+\nu}M_{\mu\nu} \quad (15.5)$$

where, again, we are not summing on the indices $\mu, \nu$.

Furthermore, the components of $g^{\mu\nu}$, the matrix inverse of $(g_{\mu\nu})$, are given by

$$g^{\mu\nu} = \frac{1}{g}M_{\mu\nu}(-1)^{\mu+\nu}. \quad (15.6)$$

Therefore,

$$\frac{\partial g}{\partial g_{\mu\nu}} = gg^{\mu\nu}, \quad (15.7)$$
then
\[
\frac{\partial g}{\partial x^\alpha} = \frac{\partial g}{\partial g_{\mu\nu}} \frac{\partial g_{\mu\nu}}{\partial x^\alpha} = g^{\mu\nu} \frac{\partial g_{\mu\nu}}{\partial x^\alpha}
\] (15.8)
and (15.1) is proved.

The same rule applies to variations:
\[
\delta g = \frac{\partial g}{\partial g_{\mu\nu}} \delta g_{\mu\nu}
\] (15.9)
thus
\[
\delta g = g^{\mu\nu} \delta g_{\mu\nu}.
\] (15.10)
Furthermore, since
\[
\delta (g_{\mu\nu} g^{\nu\sigma}) = \delta (\gamma_{ij}) = 0
\]
\[
= \delta g_{\mu\nu} g^{\nu\sigma} + g_{\mu\nu} \delta g^{\nu\sigma}
\] (15.11)
multiplying with \(g^{\mu\rho}\), we find
\[
\delta g^{\rho\sigma} = -g^{\mu\rho} g^{\nu\sigma} \delta g_{\mu\nu}.
\] (15.12)
Therefore, equation (15.10) gives
\[
\delta g = -g_{\mu\nu} \delta g^{\mu\nu}.
\] (15.13)

### 15.2 Gauss’ Theorem in curved space

First of all we give a preliminary definition.
- Given a manifold \(\mathcal{M}\) described by coordinates \(\{x^\mu\}\), and a metric \(g_{\mu\nu}\) on \(\mathcal{M}\).
- Given a submanifold \(\mathcal{N} \subset \mathcal{M}\) described by coordinates \(\{y^j\}\), such that on \(\mathcal{N}\) \(x^\mu = x^\mu(y^j)\).

We define the metric induced on \(\mathcal{N}\) from \(\mathcal{M}\) as
\[
\gamma_{ij} \equiv \frac{\partial x^\mu}{\partial y^i} \frac{\partial x^\nu}{\partial y^j} g_{\mu\nu}.
\] (15.14)

We can now generalize Gauss’ theorem to curved space:
- Be \(\Omega\) an \(n\)-dimensional volume described by coordinates \(\{x^\mu\}\)_{\mu=0,...,n-1}, and \(g_{\mu\nu}\) the metric on \(\Omega\).
- Be \(\partial \Omega\) the boundary of \(\Omega\), described by coordinates \(\{y^j\}\)_{\mu=0,...,n-2} with normal vector \(n_\mu\) (having \(|n_\mu n^\mu| = 1\) for a timelike or spacelike surface); be \(\gamma_{ij}\) the metric induced on \(\partial \Omega\) from \(g_{\mu\nu}\).
- Given a vector field \(V^\mu\) defined in \(\Omega\), then
\[
\int_\Omega d^4x \sqrt{-g} V^\mu_{\mu} = \int_{\partial \Omega} d^3y \sqrt{-\gamma} V^\mu n_\mu.
\] (15.15)
If we define the surface integration element as
\[
dS_\mu \equiv \sqrt{-\gamma} n_\mu d^3y,
\] (15.16)
the Gauss’ theorem can also be written as
\[
\int_\Omega d^4x \sqrt{-g} V^\mu_{\mu} = \int_{\partial \Omega} V^\mu dS_\mu.
\] (15.17)
In particular, if one considers an infinite volume, and if \(V^\mu\) vanishes asymptotically, then the integral of its covariant divergence is zero.
15.3 Diffeomorphisms on a manifold and Lie derivative

Given an \( n \)-dimensional differentiable manifold \( \mathcal{M} \), let us consider a mapping \( \Psi \) of the manifold \( \mathcal{M} \) into itself

\[
\Psi : \mathcal{M} \rightarrow \mathcal{M},
\]

in which we map any point \( P \in \mathcal{M} \) to another point \( Q \) of the same manifold \( \mathcal{M} \):

\[
P \rightarrow \Psi(P) = Q.
\]

If such a map is invertible and regular, as we will always assume in the following, the mapping is named **diffeomorphism**.

In a local coordinate frame that, by definition, is a mapping of an open subset of \( \mathcal{M} \) to \( \mathbb{R}^n \) (cfr. Chapter 2 section 2.5)

\[
P \rightarrow \{x^\mu(P)\}_{\mu=1,...,n}
\]

a diffeomorphism \( \Psi \) can be expressed as a set of \( n \) real functions on \( \mathbb{R}^n \), \( \{\psi^\alpha\} \), that acting on the coordinates of the point \( P \), \( \{x^\alpha\} \), produce the coordinates of the point \( Q \), \( \{x'^\alpha\} \) in **the same coordinate frame** i.e.

\[
\psi : x^\alpha \rightarrow x'^\alpha = \psi^\alpha(x).
\]

The functions \( \psi^\alpha(x^\mu) \) are smooth and invertible.

Notice that the transformation (15.21) has the same form of a general coordinate transformation, but in the present context its meaning is very different: in a coordinate transformation we assign to the same point \( P \) of the manifold two different \( n \)-ples of \( \mathbb{R}^n \); thus, in this case the mapping \( \psi \) denotes, as shown in Fig. 15.1, the set of functions which allows to transform from the coordinates \( \{x^\alpha\} \) to the coordinates \( \{x'^\alpha\} \), i.e. it is a mapping of \( \mathbb{R}^n \) to \( \mathbb{R}^n \).

\[\text{Figure 15.1: A coordinate transformation: } \zeta \text{ and } \eta \text{ are the maps from } \mathcal{M} \text{ to } \mathbb{R}^n \text{ that associate different sets of coordinates (i.e. different } n\text{-ples of } \mathbb{R}^n \text{) to the same point } P.\]

Conversely, given a diffeomorphism \( \Psi \) which maps \( P \) to \( Q \), where \( P \) and \( Q \) are two different points of the same manifold \( \mathcal{M} \), we use the same coordinate frame to label both
\( P \) and \( Q \) as shown in Fig. 15.2. To hereafter we will use the following convention: when dealing with a coordinate transformation we shall write, as usual

\[
\psi : x^{\alpha} \rightarrow x'^{\alpha} = \psi^{\alpha}(x),
\]

(15.22)

while, a diffeomorphism \( \Psi : M \rightarrow M \) expressed in a coordinate frame will be indicated as

\[
\psi : x^{\alpha} \rightarrow x'^{\alpha} = \psi^{\alpha}(x).
\]

(15.23)

i.e. \( x'^{\alpha} \) indicate the coordinates of the point \( Q \).

\[\text{Figure 15.2: A diffeomorphism } \Psi \text{ on the manifold, and its representation in a coordinate frame, } \psi. \text{ The map of } M \text{ to the coordinate frame is denoted by } \zeta.\]

### 15.3.1 The action of diffeomorphisms on functions and tensors

Be \( f \) a real function defined on a manifold \( M \). A diffeomorphism \( \Psi \) changes \( f \) into a different function, which is called the pull-back of \( f \), and is usually denoted by the symbol \( \Psi^* f \):

\[
f \rightarrow \Psi^* f
\]

such that

\[
(\Psi^* f)(P) = f(Q) \quad \rightarrow \quad (\Psi^* f)(P) = f(\Psi(P))
\]

(15.25)

i.e.

\[
\Psi^* f \equiv f \circ \Psi.
\]

(15.26)

In a coordinate frame, where \( P \) has coordinates \( x^\mu \) and \( Q \) has coordinates \( x'^\mu \), eq. (15.25) takes the form

\[
(\psi^* f)(x) = f(x') \quad \rightarrow \quad (\psi^* f)(x) = f(\psi(x)).
\]

(15.27)

The function \( \Psi^* f(P) \), which is equal to \( f(Q) \), will be different from \( f(P) \); the difference is

\[
\delta f(P) = (\Psi^* f)(P) - f(P) = f(Q) - f(P) = f(\Psi(P)) - f(P)
\]

(15.28)

or, in a coordinate frame,

\[
\delta f(x) = (\psi^* f)(x) - f(x) = f(\psi(x)) - f(x).
\]

(15.29)
In a similar way, the diffeomorphism $\Psi$ also induces a change on tensors: due to the action of $\Psi$, a tensor field $T$ changes to a new tensor field called the pull-back of $T$, denoted by $\Psi^* T$, such that
\[
(\Psi^* T)(P) = T(Q).
\] (15.30)

The difference between the pull-back and the original tensor, evaluated in a given point $P$ of the manifold, is
\[
\delta T(P) = (\Psi^* T)(P) - T(P) = T(\Psi(P)) - T(P). \tag{15.31}
\]

An important point to stress

As discussed in Chapter 3, a tensor field depends on the point of the manifold, but also on vectors and one-forms defined in the tangent space (and its dual) at that point of the manifold. In eq. (15.30) we have explicitly written the dependence on the point of the manifold, but we have left implicit the dependence on vectors and one-forms. This dependence can be made explicit as follows:

$T$ acts on vectors and one-forms defined in the tangent space (and in its dual) in $Q$, while $\Psi^* T$ acts on vectors and one-forms which are defined in the tangent space (and in its dual) in $P$. The complete definition of the pull-back $\Psi^* T$ therefore is
\[
[(\Psi^* T)(\Psi^* \tilde{V}, \ldots, \Psi^* \tilde{q}, \ldots)](P) = [T(\tilde{V}, \ldots, \tilde{q}, \ldots)](Q). \tag{15.32}
\]

In this definition there are the pull-backs of vectors and one-forms, $\Psi^* \tilde{V}$ and $\Psi^* \tilde{q}$, which we still have to define.

To this purpose, we remind that vectors are in one-to-one correspondence with directional derivatives (see Chapter 3); indeed a vector $\tilde{V} \to \{\frac{df}{d\lambda}\}$ tangent to a given curve with parameter $\lambda$, associates to any function $f$ the directional derivative
\[
\tilde{V}(f) \equiv \frac{df}{d\lambda} = \frac{\partial f}{\partial x^\mu} \frac{dx^\mu}{d\lambda} = V^\mu \frac{\partial f}{\partial x^\mu}.
\]

Thus, we define the pull-back of a vector $\Psi^* \tilde{V}$ as follows: for any function $f$
\[
[(\Psi^* \tilde{V})(\Psi^* f)](P) = \tilde{V}(f)(Q). \tag{15.33}
\]
Let us now consider one-forms. By definition a one-form $\tilde{q}$ associates to any vector $\tilde{V}$ a real number $\tilde{q}(\tilde{V})$.

The pull-back of a one-form, $\Psi^* \tilde{q}$ is defined as follows: for any vector $\tilde{V}$,

$$[(\Psi^* \tilde{q})(\Psi^* \tilde{V})](P) = [\tilde{q}(\tilde{V})](Q).$$  \hfill (15.34)

In particular, the pull-backs of the coordinate basis vectors and of the coordinate basis one-forms are defined as follows: the pull-back maps the coordinate basis (of vectors or one-forms) in $Q$ in the coordinate basis in $P$:

$$\Psi^* \frac{\partial}{\partial x_\alpha} = \frac{\partial}{\partial x^\alpha} \frac{\partial x^\mu}{\partial x_\alpha} \frac{\partial}{\partial x^\mu},$$ \hfill (15.35)

$$\Psi^* \tilde{dx}^\alpha = \tilde{dx}^\alpha = \frac{\partial x^\alpha}{\partial x'^\mu} \tilde{dx}'^\mu,$$ \hfill (15.36)

where

$$\frac{\partial x^\alpha}{\partial x'^\mu} = \frac{\partial x^\mu}{\partial x'^\alpha},$$ \hfill (15.37)

and $\psi^\alpha(x)$ are the functions defined in eq. (15.21), whereas $\partial x'^\alpha/\partial x^\mu$ are the derivatives of the inverse functions.

### 15.3.2 The action of diffeomorphisms on the metric tensor

The pull-back of the metric tensor is, by definition,

$$(\Psi^* g)(P) = g(Q).$$ \hfill (15.38)

Let us write eq. (15.38) in components. The tensor $\Psi^* g$, evaluated in $P$, must be expanded in terms of the coordinate basis of one-forms in $P$, i.e. $\tilde{dx}^\alpha$, while the tensor $g$ must be expanded in terms of the coordinate basis of one-forms in $Q$, which is $\tilde{dx}'^\mu$, therefore

$$(\psi^* g)_{\alpha\beta}(x) \tilde{dx}^\alpha \tilde{dx}^\beta = g_{\mu\nu}(x') \tilde{dx}'^\mu \tilde{dx}'^\nu$$ \hfill (15.39)

and using eq. (15.36) we find

$$(\psi^* g)_{\alpha\beta}(x) = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x'^\nu}{\partial x^\beta} g_{\mu\nu}(\psi(x)).$$ \hfill (15.40)

The change induced by $\Psi$ on the metric tensor therefore is

$$\delta g_{\alpha\beta}(x) = (\psi^* g)_{\alpha\beta}(x) - g_{\alpha\beta}(x) = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x'^\nu}{\partial x^\beta} g_{\mu\nu}(\psi(x)) - g_{\alpha\beta}(x).$$ \hfill (15.41)

### 15.3.3 Lie derivative

Let us consider an infinitesimal diffeomorphism $\Psi$, which in a coordinate frame is

$$\psi^\mu(x) = x^\mu + \epsilon \xi^\mu(x),$$ \hfill (15.42)
where \( \xi^\mu \) is a vector field and \( \epsilon \) is a small parameter. To hereafter we shall neglect \( O(\epsilon^2) \) terms. From eq. (15.42) it follows that

\[
x'^\mu = x^\mu + \epsilon \xi^\mu.
\]  

According to eq. (15.29) the change induced by \( \Psi \) on a function \( f \) is

\[
\delta f(x) = f(x + \epsilon \xi) - f(x) = \epsilon \xi^\mu \frac{\partial f}{\partial x^\mu}.
\]  

We define the Lie derivative of a scalar function \( f \) with respect to the infinitesimal diffeomorphism associated to the vector field \( \xi \), as

\[
\mathcal{L}_\xi(f) \equiv \lim_{\epsilon \to 0} \frac{\Psi^* f(P) - f(P)}{\epsilon} = \lim_{\epsilon \to 0} \frac{f(\Psi(P)) - f(P)}{\epsilon},
\]  

or, in coordinates,

\[
\mathcal{L}_\xi(f) \equiv \lim_{\epsilon \to 0} \frac{f(x + \epsilon \xi) - f(x)}{\epsilon} = \xi^\mu \frac{\partial f}{\partial x^\mu}.
\]  

Therefore, eq. (15.44) can also be written as

\[
\delta f = \epsilon \mathcal{L}_\xi f.
\]  

In a similar way we can define the Lie derivative of a tensor field \( T \) with respect to the vector field \( \xi \) as

\[
\mathcal{L}_\xi T \equiv \lim_{\epsilon \to 0} \frac{\Psi^* T(P) - T(P)}{\epsilon} = \lim_{\epsilon \to 0} \frac{T(\Psi(P)) - T(P)}{\epsilon}.
\]  

Therefore, from eq. (15.31) we see that the change induced on a tensor by an infinitesimal diffeomorphism is given by the Lie derivative:

\[
\delta T = \epsilon \mathcal{L}_\xi T.
\]  

In the case of the metric tensor, since from eq.(15.42)

\[
\frac{\partial x'^\mu}{\partial x'^\nu} = \delta^\mu_\nu + \epsilon \frac{\partial \xi^\mu}{\partial x'^\nu},
\]  

equation (15.41) gives

\[
\delta g_{\alpha\beta}(x) = \left( \delta^\mu_\alpha + \epsilon \frac{\partial \xi^\mu}{\partial x'^\alpha} \right) \left( \delta^\nu_\beta + \epsilon \frac{\partial \xi^\nu}{\partial x'^\beta} \right) g_{\mu\nu}(x + \epsilon \xi) - g_{\alpha\beta}(x)
\]

\[
= \epsilon \left[ \frac{\partial \xi^\mu}{\partial x'^\alpha} g_{\mu\beta} + \frac{\partial \xi^\nu}{\partial x'^\beta} g_{\alpha\nu} + \xi^p \frac{\partial g_{\alpha\beta}}{\partial x'^p} \right],
\]  

therefore

\[
\mathcal{L}_\xi g_{\mu\nu} = \frac{\partial \xi^\mu}{\partial x'^\alpha} g_{\mu\beta} + \frac{\partial \xi^\nu}{\partial x'^\beta} g_{\alpha\nu} + \xi^p \frac{\partial g_{\alpha\beta}}{\partial x'^p}.
\]
This expression has already been derived in Chapter 7 when we studied how the metric tensor changes under infinitesimal translations along a given curve of tangent vector $\vec{\xi}$

$$x^\mu \rightarrow x^\mu + \xi^\mu(x)d\lambda$$  \hspace{1cm} (15.53)

Indeed, an infinitesimal translation is an infinitesimal diffeomorphism; thus the derivation of eqs. (15.51) and (15.52) repeats, in a more rigorous way, the derivation presented in Chapter ???. Furthermore, as shown in Chapter ??, the expression (15.52) can be rewritten in terms of covariant derivatives of $\vec{\xi}$

$$\mathcal{L}_\xi g_{\mu\nu} = \xi_{\mu\nu} + \xi_{\nu\mu}.$$  \hspace{1cm} (15.54)

### 15.3.4 General Covariance and the role of diffeomorphisms

Since the equation

$$\psi : x^\alpha \rightarrow x'^\alpha = \psi^\alpha(x)$$  \hspace{1cm} (15.55)

allows a double interpretation as a general coordinate transformation and as a diffeomorphism on the spacetime manifold, it follows that to any coordinate transformation we can associate a diffeomorphism, and viceversa.

The principle of general covariance states that since the laws of physics are expressed by tensorial equations, they retain the same form in any coordinate frame, i.e. they are invariant under a general coordinate transformation; since a coordinate transformation corresponds, in the sense explained above, to a diffeomorphism, we can restate the principle of general covariance as follows: **all physical laws are invariant for diffeomorphisms**, i.e. we restate general covariance as symmetry principle defined on the manifold.

### 15.4 Variational approach to General Relativity

In the variational approach to field theory, the dynamics of fields is described by the action functional.

**15.4.1 Action principle in special relativity**

Let us consider a collection of tensor (and, eventually, spinor) fields in special relativity

$$\left\{ \Phi^{(A)}(x) \right\}_{A=1,\ldots}.$$  \hspace{1cm} (15.56)

where $x$ denotes the point of coordinates $\{x^\mu\}$. We shall use symbols in boldface to denote a generic tensorial or spinorial object. For instance, for a vector field we shall write

$$\mathbf{V} = (V^\mu),$$  \hspace{1cm} (15.57)

and summation over the tensor indices will be left implicit:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{V}} \delta \mathbf{V} = \frac{\partial \mathcal{L}}{\partial V^\mu} \delta V^\mu.$$  \hspace{1cm} (15.58)
It should be noted that here we use a convention different from that used for vector fields in Chapter 3.

Typically, the fields involved are scalars, rank-one tensors (vectors and one-forms) and spinors. The action is a functional of these fields and of their first derivatives, written as an integral of a Lagrangian density over the 4-dimensional volume:

\[ I = \int d^4x \mathcal{L} \left( \Phi^{(1)}(x), \ldots, \Phi^{(A)}(x), \ldots, \partial_\mu \Phi^{(1)}(x), \ldots, \partial_\mu \Phi^{(A)}(x), \ldots \right). \]  

To hereafter \( \partial_\mu \equiv \frac{\partial}{\partial x^\mu} \). All fields are assumed to vanish on the boundary of the integration volume or asymptotically, if the volume is infinite.

Let us consider the variation of the action with respect to a given field \( \Phi^{(A)} \)

\[ \delta I = \int d^4x \sum_A \frac{\delta \mathcal{L}}{\delta \Phi^{(A)}} \delta \Phi^{(A)} = \int d^4x \sum_A \left( \frac{\partial \mathcal{L}}{\partial \Phi^{(A)}} \delta \Phi^{(A)} + \frac{\partial \mathcal{L}}{\partial (\partial_\alpha \Phi^{(A)})} \delta (\partial_\alpha \Phi^{(A)}) \right) = \int d^4x \sum_A \left( \frac{\partial \mathcal{L}}{\partial \Phi^{(A)}} - \partial_\alpha \frac{\partial \mathcal{L}}{\partial (\partial_\alpha \Phi^{(A)})} \right) \delta \Phi^{(A)} . \]  

Here we have used the general property that the operations of variation and differentiation commute, and then we have integrated by parts. The stationarity of \( \delta I \) with respect to the considered field gives the equation of motion for that field

\[ \delta I = 0, \quad \forall \delta \Phi^{(A)}, \]  

and since the integral (15.60) has to vanish for every \( \delta \Phi^{(A)}(x) \), it follows that

\[ \frac{\partial \mathcal{L}}{\partial \Phi^{(A)}} - \partial_\alpha \frac{\partial \mathcal{L}}{\partial (\partial_\alpha \Phi^{(A)})} = 0 \]  

which are the Euler-Lagrange equations for the field \( \Phi^{(A)} \).

### 15.4.2 Action principle in general relativity

In general relativity, besides the fields \( \{ \Phi^{(A)} \} \), which are the matter and gauge fields, there is the metric field

\[ g(x) = (g_{\mu\nu}(x)) \]  

which describes the gravitational field whose action is the Einstein-Hilbert action

\[ I^{\text{E-H}} = \frac{c^3}{16\pi G} \int d^4x \sqrt{-g} \, R . \]
Due to the strong equivalence principle, in a locally inertial frame the dynamics of all fields \( \{ \Phi^{(A)} \} \) except gravity is described by the action (15.59). Therefore, according to the principle of general covariance in a general frame the action, which is a scalar, retains the same form provided \( \eta_{\mu \nu} \to g_{\mu \nu} \), the partial derivatives \( \partial_\alpha \) are replaced by covariant derivatives \( \nabla_\alpha \), and the integration volume element \( d^4x \) is replaced by the covariant volume element \( \sqrt{-g} d^4x \).

With these replacements, we shall now show that the results of the previous section (in particular, the derivation of Euler-Lagrange equations) remain valid.

The total action is

\[
I = I^{E-H} + I^{FIELDS} \tag{15.64}
\]

with

\[
I^{FIELDS} = \int d^4x \sqrt{-g} L^{FIELDS} \left( \Phi^{(1)}(x), \ldots, \Phi^{(A)}(x), \ldots, \nabla_\alpha \Phi^{(1)}(x), \ldots, \nabla_\alpha \Phi^{(A)}(x), \ldots, g \right). \tag{15.65}
\]

Notice that now the Lagrangian density \( L^{FIELDS} \) depends explicitly on \( g \) because we have replaced \( \eta_{\mu \nu} \) by \( g_{\mu \nu} \) and \( \partial_\alpha \) by \( \nabla_\alpha \).

As in special relativity, the equations for a field \( \Phi^{(A)} \) are found by varying the action with respect to that field, and since the Einstein-Hilbert action does not depend on \( \Phi \)

\[
\delta I = \delta I^{FIELDS} = \int d^4x \sqrt{-g} \sum_A \frac{\delta L^{FIELDS}}{\delta \Phi^{(A)}} \delta \Phi^{(A)}
\]

\[
= \int d^4x \sqrt{-g} \sum_A \left( \frac{\partial L}{\partial \Phi^{(A)}} \delta \Phi^{(A)} + \frac{\partial L}{\partial \left( \partial_\alpha \Phi^{(A)} \right)} \nabla_\alpha \delta \Phi^{(A)} \right)
\]

\[
= \int d^4x \sqrt{-g} \sum_A \left( \frac{\partial L}{\partial \Phi^{(A)}} \delta \Phi^{(A)} + \frac{\partial L}{\partial \left( \partial_\alpha \Phi^{(A)} \right)} \nabla_\alpha \delta \Phi^{(A)} \right)
\]

\[
= \int d^4x \sqrt{-g} \sum_A \left( \frac{\partial L}{\partial \Phi^{(A)}} - \nabla_\alpha \frac{\partial L}{\partial \left( \partial_\alpha \Phi^{(A)} \right)} \right) \delta \Phi^{(A)} = 0, \quad \forall \delta \Phi^{(A)} \tag{15.66}
\]

where we have used the property \( \delta \nabla_\alpha = \nabla_\alpha \delta \). To obtain the last row of eq. (15.66) we have integrated by parts using the generalization of Gauss’ theorem in curved space (see Section 15.2) which assures that the integral of a covariant divergence is zero, provided all fields vanish at the boundary, or asymptotically if the volume is infinite, i.e.

\[
\int d^4x \sqrt{-g} V^{\mu;\mu} = 0. \tag{15.67}
\]

Thus, the equations of motion for the field \( \Phi^{(A)} \) are the Euler-Lagrange equations generalized in curved space:

\[
\frac{\partial L}{\partial \Phi^{(A)}} - \nabla_\alpha \frac{\partial L}{\partial \left( \partial_\alpha \Phi^{(A)} \right)} = 0. \tag{15.68}
\]
The field equations for the gravitational field are obtained by varying the action (15.64) with respect to \( g \). In section 15.4.3 we shall show that the variation of the Einstein-Hilbert action with respect to \( g \) gives the Einstein tensor; here we just write the result:

\[
\delta I^{E-H} = \frac{c^3}{16\pi G} \int d^4x \delta (\sqrt{-g} R) = \frac{c^3}{16\pi G} \int d^4x \sqrt{-g} \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) \delta g^{\mu\nu}.
\] (15.69)

The variation of \( I^{\text{FIELDS}} \) with respect to \( g \) it is easy to find if we use the property (15.13), from which

\[
\delta (\sqrt{-g})^{1/2} = -\frac{1}{2} (\sqrt{-g})^{1/2} g_{\mu\nu} \delta g^{\mu\nu},
\] (15.70)

thus

\[
\delta I^{\text{FIELDS}} = \int d^4x \sqrt{-g} \left[ \frac{\delta L^{\text{FIELDS}}}{\delta g^{\mu\nu}} - \frac{1}{2} L^{\text{FIELDS}} g_{\mu\nu} \right] \delta g^{\mu\nu}.
\] (15.71)

Combining eqs. (15.69) and (15.71), and defining the stress-energy tensor as

\[
T_{\mu\nu} \equiv -\frac{c^3}{2} \left[ \frac{\delta L^{\text{FIELDS}}}{\delta g^{\mu\nu}} - \frac{1}{2} L^{\text{FIELDS}} g_{\mu\nu} \right]
\] (15.72)

the variation of the total action (15.64) can be written as

\[
\delta I = \frac{c^3}{16\pi G} \int d^4x \sqrt{-g} \left[ G_{\mu\nu} - \frac{8\pi G}{c^4} T_{\mu\nu} \right] \delta g^{\mu\nu} = 0.
\] (15.73)

Thus, with this definition of \( T_{\mu\nu} \) the action principle gives Einstein’s equations

\[ G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \] (15.74)

The advantage of deriving field equations using a variational approach is that it makes explicit the connection between symmetries and conservation laws: any symmetry of a theory corresponds to a conservation law.

In Section 15.3 we have shown that the principle of general covariance implies that equations expressing the laws of physics are invariant for diffeomorphisms. If they can be derived from an Action Principle, this is equivalent to impose that the action is invariant for diffeomorphisms.

As an example, let us consider the action \( I^{\text{FIELDS}} \). It is invariant for diffeomorphisms, because the equations of motion (15.68) for the fields \( \Phi^{(A)} \), arising from \( \delta I^{\text{FIELDS}} = 0 \), are diffeomorphism invariant. We will now show that the diffeomorphism invariance of \( I^{\text{FIELDS}} \) implies the divergenceless equation \( T^{\mu\nu}_{\;\;;\nu} = 0 \), which generalizes the conservation law \( T^{\mu\nu}_{\;\;\nu} = 0 \) of flat spacetime.

Let us consider an infinitesimal diffeomorphism (see eq. 15.42)

\[ x^\mu \rightarrow x'^\mu = x^\mu + \epsilon \xi^\mu. \] (15.75)

In Section 15.3 we have shown how tensor fields change under infinitesimal diffeomorphisms, and in particular that the change of the metric tensor, written in components, is (eq. 15.54):

\[ \delta g_{\mu\nu} = \epsilon \left[ \xi_{\mu;\nu} + \xi_{\nu;\mu} \right]. \] (15.76)
The change in $g^{\mu\nu}$ can be derived by using the property (15.12):

$$\delta g^{\mu\nu} = -\epsilon \left[ \xi^{\mu\nu} + \xi^{\nu\mu} \right].$$

(15.77)

The diffeomorphism (15.75) will produce a variation of the fields $\Phi^{(A)}$'s, and a variation of the metric tensor given by (15.77); consequently $I^{\text{FIELDS}}$, defined in eq. (15.65), will vary as follows

$$\delta I^{\text{FIELDS}} = \int d^4x \sqrt{-g} \left[ \frac{\delta L^{\text{FIELDS}}}{\delta g^{\mu\nu}} - \frac{1}{2} L^{\text{FIELDS}} g^{\mu\nu} \right] \delta g^{\mu\nu}
+ \int d^4x \sqrt{-g} \sum_A \frac{\delta L^{\text{FIELDS}}}{\delta \Phi^{(A)}} \delta \Phi^{(A)}.$$

(15.78)

Since the action $I^{\text{FIELDS}}$ is invariant under diffeomorphisms, $\delta I^{\text{FIELDS}}$ must vanish. The term in square brackets is the stress-energy tensor $T_{\mu\nu}$ defined in eq. (15.72). Therefore, using eq. (15.77), we find

$$\delta I^{\text{FIELDS}} = \frac{\epsilon c^3}{2} \int d^4x \sqrt{-g} T_{\mu\nu} \xi^{\mu\nu} + \frac{\epsilon}{c} \int d^4x \sqrt{-g} \sum_A \frac{\delta L^{\text{FIELDS}}}{\delta \Phi^{(A)}} \delta \Phi^{(A)} = 0$$

(15.79)

where we have used: $T_{\mu\nu} (\xi^{\mu\nu} + \xi^{\nu\mu}) = 2T_{\mu\nu} \xi^{\mu\nu}$, which follows from the symmetry of $T_{\mu\nu}$.

Since the fields $\Phi^{(A)}$'s satisfy their equations of motion (15.68) the last integral in eq. (15.79) vanishes. Thus

$$\delta I^{\text{FIELDS}} = \frac{\epsilon}{c} \int d^4x \sqrt{-g} T_{\mu\nu} \xi^{\mu\nu} = \frac{\epsilon}{c} \int d^4x \sqrt{-g} T^{\mu\nu} \xi^\mu = 0 \ \forall \xi^\mu$$

(15.80)

and consequently

$$T^{\mu\nu} \xi^\mu = 0.$$  

(15.81)

We stress that eq. (15.81) is not satisfied for all field configurations, but only for the field configurations which are solutions of the field equations, i.e. of the Euler-Lagrange equations (15.68).

The diffeomorphism invariance of the Einstein-Hilbert action $I^{E-H}$, instead, implies the Bianchi identities. Indeed, the variation of $I^{E-H}$ is given by (15.69)

$$\delta I^{E-H} = \frac{\epsilon c^3}{16\pi G} \int d^4x \sqrt{-g} G_{\mu\nu} \delta g^{\mu\nu}$$

(15.82)

and if the variation is due to an infinitesimal diffeomorphism, from (15.77)

$$\delta g^{\mu\nu} = \epsilon \left[ \xi^{\mu\nu} + \xi^{\nu\mu} \right]$$

(15.83)

thus, integrating by parts,

$$\delta I^{E-H} = \frac{\epsilon c^3}{16\pi G} \int d^4x \sqrt{-g} G_{\mu\nu} 2\xi^{\mu\nu} = -\frac{\epsilon c^3}{8\pi G} \int d^4x \sqrt{-g} G^{\mu\nu} \xi^\mu = 0 \ \forall \xi^\mu$$

(15.84)

and consequently we find the Bianchi identities

$$G^{\mu\nu} \xi^\mu = 0.$$  

(15.85)
15.4.3 Variation of the Einstein-Hilbert action

In this section we shall prove that the variation of the Einstein-Hilbert action with respect to the metric tensor gives the Einstein tensor (equation (15.69)):

$$\int \delta (\sqrt{-g} R) = \int \sqrt{-g} \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \right) \delta g^{\mu\nu}. \quad (15.86)$$

As a first step we derive the Palatini identity:

$$\delta R_{\mu\nu} = (\delta \Gamma^\lambda_{\mu\nu})_{;\lambda} - (\delta \Gamma^\lambda_{\mu\lambda})_{;\nu}. \quad (15.87)$$

By varying the Ricci tensor:

$$R_{\mu\nu} = \Gamma^\lambda_{\mu\nu,\lambda} - \Gamma^\lambda_{\mu\lambda,\nu} + \Gamma^\alpha_{\mu\nu} \Gamma^\lambda_{\alpha\lambda} - \Gamma^\lambda_{\alpha\nu} \Gamma^\alpha_{\mu\lambda}, \quad (15.88)$$

we find

$$\delta R_{\mu\nu} = \left( \delta \Gamma^\lambda_{\mu\nu,\lambda} - \delta \Gamma^\lambda_{\mu\lambda,\nu} + \delta \Gamma^\alpha_{\mu\nu} \Gamma^\lambda_{\alpha\lambda} - \delta \Gamma^\lambda_{\alpha\nu} \Gamma^\alpha_{\mu\lambda} \right). \quad (15.89)$$

To evaluate this expression, we need to compute $\delta \Gamma^\lambda_{\mu\nu}$. To this purpose, we will use the property (15.12)

$$\delta g^\rho_\lambda \delta g^\sigma_\delta = -g^\rho^\lambda g^{\sigma\delta} \delta g^\rho_\sigma \delta g^\sigma_\rho. \quad (15.90)$$

If we define

$$\Gamma^\mu_{\nu\delta} = \frac{1}{2} (g_{\mu\delta,\nu} + g_{\nu\delta,\mu} - g_{\mu\nu}, \delta) \quad (15.91)$$

we can write the variation of Christoffel’s symbols as follows

$$\delta \Gamma^\lambda_{\mu\nu} = \delta \left[ g^\lambda_\mu \Gamma^\mu_{\nu\delta} \right] = \delta g^\lambda_\mu \Gamma^\mu_{\nu\delta} + g^\lambda_\mu \delta \Gamma^\mu_{\nu\delta}$$

$$= -g^\rho^\lambda g^{\sigma\delta} \delta g_{\rho\sigma} \Gamma^\mu_{\nu\delta} + g^\lambda_\rho \delta \Gamma^\mu_{\nu\rho}$$

$$= -g^\rho^\lambda \delta g_{\rho\sigma} \Gamma^\mu_{\nu\sigma} + \frac{1}{2} \delta g_{\mu\rho,\nu} + \delta g_{\nu\rho,\mu} - \delta g_{\mu\nu,\rho}$$

$$= \frac{1}{2} g^\lambda_\rho \left[ \delta g_{\mu\rho,\nu} + \delta g_{\nu\rho,\mu} - \delta g_{\mu\nu,\rho} - 2 \Gamma^\alpha_{\mu\nu} \delta g^\alpha_{\rho\sigma} \right]. \quad (15.92)$$

which can be recast in the form

$$\delta \Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^\lambda_\rho \left[ \left( \delta g_{\mu\rho,\nu} - \Gamma^\alpha_{\mu\rho} \delta g_{\alpha\rho} - \Gamma^\alpha_{\nu\rho} \delta g_{\alpha\mu} \right) + \left( \delta g_{\nu\rho,\mu} - \Gamma^\alpha_{\nu\rho} \delta g_{\alpha\rho} - \Gamma^\alpha_{\mu\rho} \delta g_{\alpha\mu} \right) \right.$$

$$- \left( \delta g_{\mu\nu,\rho} - \Gamma^\alpha_{\mu\rho} \delta g_{\alpha\nu} - \Gamma^\alpha_{\nu\rho} \delta g_{\alpha\mu} \right) \left. \right]$$

$$= \frac{1}{2} g^\lambda_\rho \left[ \delta g_{\mu\rho,\nu} + \delta g_{\nu\rho,\mu} - \delta g_{\mu\nu,\rho} \right]. \quad (15.93)$$

Eq. (15.93) shows that since $\delta g_{\mu\nu}$ is a tensor, $\delta \Gamma^\lambda_{\mu\nu}$ is also a tensor. Therefore, the expression (15.87),

$$(\delta \Gamma^\lambda_{\mu\nu})_{;\lambda} - (\delta \Gamma^\lambda_{\mu\lambda})_{;\nu} \quad (15.94)$$
can be evaluated with the usual rules of covariant differentiation of tensors. Thus we find

\[
\begin{align*}
\left( \delta \Gamma^\lambda_{\mu\nu} \right)_{,\lambda} - \left( \delta \Gamma^\lambda_{\mu\lambda} \right)_{,\nu} &= \delta \Gamma^\lambda_{\mu\nu,\lambda} - \delta \Gamma^\lambda_{\mu\lambda,\nu} + \delta \Gamma^\lambda_{\mu\nu} \Gamma^\alpha_{\nu\lambda} - \delta \Gamma^\lambda_{\mu\nu,\alpha} \Gamma^\alpha_{\nu\lambda} - \Gamma^\lambda_{\alpha\lambda} \delta \Gamma^\alpha_{\mu\lambda}.
\end{align*}
\] (15.95)

A comparison of this equation with eq. (15.89) shows that

\[
\delta R^\mu_{\nu\nu} = \left( \delta \Gamma^\lambda_{\mu\nu} \right)_{,\lambda} - \left( \delta \Gamma^\lambda_{\mu\lambda} \right)_{,\nu}.
\] QED.

Let us now go back to the variation of the Einstein-Hilbert action

\[
\delta \left( \sqrt{-g} R \right) = \delta \left( \sqrt{-g} g^\mu_{\nu} R^\mu_{\nu} \right).
\] (15.96)

Using eq. (15.70) we find

\[
\delta \left( \sqrt{-g} R \right) = \sqrt{-g} \left( g^\mu_{\nu} \delta R^\mu_{\nu} + \delta g^\mu_{\nu} R^\mu_{\nu} - \frac{1}{2} g_{\mu\nu} \delta g^\mu_{\nu} R \right).
\] (15.97)

Using the fact that \( g_{\mu\nu;\alpha} = 0 \), the first term in (15.97) becomes

\[
\begin{align*}
g^\mu_{\nu} \delta R^\mu_{\nu} &= g^\mu_{\nu} \left( \delta \Gamma^\lambda_{\mu\nu,\lambda} - \left( \delta \Gamma^\lambda_{\mu\lambda} \right)_{,\nu} \right) = \left( g^\mu_{\nu} \delta \Gamma^\lambda_{\mu\nu,\lambda} - (g^\mu_{\nu} \delta \Gamma^\lambda_{\mu\lambda})_{,\nu} \right) \\
&= \left( g^\mu_{\nu} \delta \Gamma^\lambda_{\mu\nu} - g^\mu_{\nu} \delta \Gamma^\lambda_{\mu\lambda} \right)_{,\alpha}.
\end{align*}
\] (15.98)

which is the divergence of a vector; therefore by Gauss’ theorem such term vanishes when integrated over the 4-volume \( \sqrt{-g} d^4x \). The remaining two terms in (15.97) give the Einstein tensor, therefore we can write

\[
\delta \left( \sqrt{-g} R \right) = \sqrt{-g} \left( R^\mu_{\nu\nu} - \frac{1}{2} g_{\mu\nu} R \right) \delta g^\mu_{\nu} + \text{surface terms}
\] (15.99)

which shows that eq. (15.86) is true.

**15.4.4 An example: the electromagnetic field**

The Lagrangian density of the electromagnetic field is

\[
\mathcal{L} = -\frac{1}{4c} F_{\mu\alpha} F_{\nu\beta} g^{\mu\nu} g^{\alpha\beta}
\] (15.100)

with

\[
F_{\mu\nu} = \partial_{\mu} A_{\nu} - \partial_{\nu} A_{\mu} = A_{\nu,\mu} - A_{\mu,\nu}
\] (15.101)

(the last equality arises from the symmetry property of the Christoffel symbols \( \Gamma^\alpha_{\mu\nu} = \Gamma^\alpha_{\nu\mu} \)).

The field equations for the electromagnetic field \( A^\mu \) are

\[
\frac{c \delta \mathcal{L}}{\delta A^\nu} = -\frac{1}{2} F^{\alpha\beta} \frac{\delta F_{\alpha\beta}}{\delta A^\nu} = -\frac{1}{2} F^{\alpha\beta} \frac{\delta}{\delta A^\nu} (A_{\beta;\alpha} - A_{\alpha;\beta})
\]

\[
= F^{\alpha\beta} \frac{\delta}{\delta A^\nu} A_{\alpha;\beta} = -F^{\alpha\beta;\beta} + \text{surface terms}.
\] (15.102)
as usual, we eliminate the surface terms on the assumption that the fields vanish at the boundary of a given volume, or at infinity. The field equations then are

$$F^{\alpha\beta}_{;\alpha} = 0,$$

and the stress-energy tensor is

$$T_{\mu\nu} = -2c \left[ \frac{\delta L}{\delta g^{\mu\nu}} - \frac{1}{2} \frac{L}{g_{\mu\nu}} \right] = -2 \left[ -\frac{1}{2} F_{\mu\alpha} F_{\nu}^{\alpha} + \frac{1}{8} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta} \right] = F_{\mu\alpha} F_{\nu}^{\alpha} - \frac{1}{4} g_{\mu\nu} F_{\alpha\beta} F^{\alpha\beta}. \quad (15.104)$$

### 15.4.5 A comment on the definition of the stress-energy tensor

In special relativity, typically one defines the canonical stress-energy tensor as

$$T_{\mu\nu}^{\text{can}} = c \left[ \partial_{\beta} \frac{\partial L}{\partial \Phi^{\alpha}} \partial^{\beta} \Phi - \eta^{\alpha\beta} L \right]. \quad (15.105)$$

In principle, we can generalize this definition in curved space by appealing to the principle of general covariance

$$T_{\mu\nu}^{\text{can}} = c \left[ \frac{\partial L}{\partial (\nabla \Phi)} \nabla^{\beta} \Phi - g^{\alpha\beta} L \right]. \quad (15.106)$$

However, the tensor (15.106) is not a good stress-energy tensor. Indeed, in general it is not symmetric, and it does not satisfy the divergenceless condition

$$T_{\mu\nu}^{\text{can}} \partial_{\beta} = 0. \quad (15.107)$$

The correct definition for the stress-energy tensor in general relativity is given by eq. (15.72)

$$T_{\mu\nu} = -2c \left[ \frac{\delta L}{\delta g^{\mu\nu}} - \frac{1}{2} g_{\mu\nu} L \right]. \quad (15.108)$$

Anyway, it can be shown that the difference between (15.108) and the canonical stress-energy tensor (15.106) is a total derivative, which disappears when integrated in the overall spacetime; furthermore, it can be shown that in the Minkowskian limit, where $g_{\mu\nu} \rightarrow \eta_{\mu\nu}$ and $\nabla_{\mu} \rightarrow \partial_{\mu}$ the two tensors (15.106), (15.108) coincide.